

Verdad:

Deflacionismo, Referencia y Paradojas

Autor:

Picollo, Lavinia María

Tutor:

Barrio, Eduardo Alejandro

2015

Tesis presentada con el fin de cumplimentar con los requisitos finales para la obtención del título Doctor de la Universidad de Buenos Aires en Filosofía.

Posgrado

[Universidad de Buenos Aires](http://www.ub.edu.ar)

Tesis Doctoral

VERDAD
De acionismo, Referencia y Paradojas

Autora:
Lavinia Mar a Picollo

Director y Consejero de
Estudios:
Dr Eduardo Alejandro Barrio



"In vino, possibly, veritas", but in a sober symposium verum".

John L. Austin

Agradecimientos

Quisiera agradecer primero y principalmente a mi director y consejero de estudios Eduardo Barrio, sin el cual esta tesis no habr a sido de ninguna manera posible. Gracias a Eduardo comenc a trabajar sobre la nocion de verdad y hered de el las ideas de acionistas que ac de endo. Le agradezco no solo por su gran labor de supervision de todos mis escritos y su continuo apoyo sino tambien por su constante incentivo a la autosuperacion, por sus innumerables esfuerzos para ubicarme a m y a todos los inte-grantes del Buenos Aires Logic Group al nivel de los mejores investigadores del mundo en el area.

Para citar algunos ejemplos, desde el a~no 2010 Eduardo coordina el seminario Work in progress, en el cual aprend a dar charlas as como a escucharlas y realizar buenas intervenciones. Tambien ha establecido v nculos en la forma de proyectos de investigacion colaborativos con la Universidad de Oxford y el MCMP (Munich Center for Mathematical Philosophy), actualmente los mejores centros donde trabajar los temas que ocupan esta tesis. Esto me permitio realizar estancias cortas en ambos lugares durante el tiempo de mi doctorado, en las que aprend a adoptar nuevos puntos de vista, nuevas metodolog as, y las cuales modelaron y enriquecieron mis ideas de modo sustancial. Como tercer ejemplo, aunque podr a citar muchos otros, Eduardo se ha encargado de organizar numerosos talleres y conferencias e invitar una gran cantidad de losofos de todas las latitudes a exponer su trabajo frente al grupo. En otras palabras, me abrio las puertas a la comunidad logico- loso ca internacional.

Estoy especialmente en deuda con Volker Halbach, quien me ha recibido en la Universidad de Oxford e integrado al grupo de trabajo que dirige durante mi estad a en los meses de abril, mayo y junio de 2012. Fue una grata sorpresa enterarme a mi llegada de que ambos estabamos interesados en el concepto de referencia aplicado al lenguaje de la aritmetica y sus extensiones, y comenzando a trabajar en ello. Esto se tradujo en una serie de discusiones loso cas al respecto que han tenido una vasta in uencia en el trabajo sobre referencia que presento en esta tesis. Sin obligacion alguna, desde entonces Volker ha supervisado extrao cialmente todo mi trabajo, dandome consejos y devoluciones invaluable. Asimismo, me puso en contacto con muchos especialistas en los temas de esta tesis, con los cuales he tenido conversaciones fruct feras que impulsaron mi investigacion en la direccion actual.

Estoy sumamente agradecida con Hannes Leitgeb por recibirme en el MCMP en tres oportu nidades diferentes, integrandome a su grupo de doctorandos, permitiendome participar de su seminario de trabajo en progreso, supervisandome durante mi estad a,

manteniendo fructíferas discusiones los casos, dándome valiosos consejos y apoyando mi aplicación a diversas conferencias y escuelas de verano. Gracias a Hannes logre superar el estancamiento en el cual se encontraban mis investigaciones en torno a la referencia a mediados de 2013.

Un agradecimiento muy especial es para Thomas Schindler, quien no solo ha colaborado conmigo en la escritura del capítulo 5 y en la prueba de la proposición [3.37](#) sino que también hemos mantenido numerosas discusiones que desembocaron en diversas ideas que acabo desarrollando, en particular en torno al deflacionismo y al minimalista acerca de la verdad. También quisiera agradecerle el enorme apoyo que me brindó estos últimos meses para poder completar la escritura de la tesis.

A lo largo de estos años el Buenos Aires Logic Group ha sido una ayuda invaluable para la exploración de distintos temas, producción y discusión de ideas, tanto en el seminario Work in Progress como en el grupo de lectura y en las tantas ocasiones en las cuales he tenido la oportunidad de discutir ideas los casos con sus integrantes. Les agradezco mucho a Natalia Buacar, Bruno Da Re, Juan Nascimbene, Federico Paislos, Ariel Roa, Lucas Rosenblatt, Mariela Rubin, Damian Szmuc, Diego Tajer, Paula Teijeiro y Joaquín Toranzo. Quisiera agradecer también a los miembros del seminario Work in progress que no pertenecen a este grupo, esto es, Ramiro Caso, Miranda del Corral, Justina Díaz Legaspe y Nicolás Loguercio, por sus excelentes comentarios en las oportunidades que tuve de presentar mi trabajo ante ellos.

Agradezco también a los miembros de los grupos de lógica de la Universidad de Oxford, del MCMP y del TiLPS (Tilburg Center for Logic, Ethics, and Philosophy of Science), por haber escuchado mis charlas sobre extractos en progreso de la tesis y por sus excelentes comentarios y objeciones, sin los cuales mi trabajo no sería lo que es. En particular quisiera agradecer a Catrin Campbell-Moore, Johannes Stern y Martin Fischer por discutir puntos específicos sobre la referencia conmigo y a Jonas Raab su lectura detallada de algunas partes de la tesis y sus útiles comentarios.

Estoy en deuda también con J.C. Beall, Filip Buekens, Cezary Cieślinski, Roy Cook, Je Ketland, Øystein Linnebo, Agustín Rayo y Albert Visser por sus excelentes devoluciones sobre diferentes partes de la tesis, así como sus buenos consejos para el futuro de la investigación.

He presentado mi trabajo en diversos eventos académicos, donde recibí muy buenos comentarios que resultaron de suma utilidad para esta investigación: charlas en las Universidades de Gante y Groningen; el 29th International Symposium Logica 2015 en Hejnice, República Checa; el II Congreso Internacional de la Sociedad Filosófica del

Uruguay en Montevideo; el Logic Colloquium 2013 en Evora, Portugal; el First MCMP-Munich Buenos Aires Workshop en Munich, Alemania; el Buenos Aires-Oxford Research Project: Truth, Open-Endedness and Inexpressibility, el II Congreso Latinoamericano de Filosofía Analítica, el Logic Workshop in Buenos Aires - A Tribute to Horacio Arlo-Costa y el Symposium on Yablo's Paradox, los cuatro en Buenos Aires; la Third New College Logic Meeting en Oxford, Reino Unido; y, por último, el Segundo Coloquio de Jóvenes Investigadores en Filosofía Analítica en Buenos Aires. Mi profundo agradecimiento a todos sus participantes.

Quisiera agradecer también a la Universidad de Buenos Aires, por haberme permitido realizar mis estudios doctorales y haber financiado parte de mi viaje a Oxford; al Conicet por haber financiado todos mis estudios doctorales y la parte restante de aquel viaje: al DAAD, el servicio alemán de intercambio académico, por haber cubierto por completo una de mis estancias de tres meses en el MCMP; y, por último pero no menos importante, a Sadaf, por ser el lugar donde todas las reuniones del Buenos Aires Logic Group, las charlas de los invitados y muchos eventos académicos han tenido y siguen teniendo lugar.

Finalmente, quisiera agradecer a mi familia y, principalmente, a mi madre, cuyo apoyo incondicional y sin precedentes es mi sosten, la condición de posibilidad de toda mi carrera.

Indice general

Agradecimientos	II
	V
Indice general	
Introduccion	VIII
1. Preliminares	1
1.1. Lenguajes y teor as de primer orden	1
1.2. La aritmetica de primer orden	3
1.2.1. Lenguaje y teor a	3
1.2.2. De nibilidad y representabilidad	3
1.2.3. Codigos de G•odel y aritmetica como teor a de la sintaxis	5
1.2.4. El lenguaje de la verdad	8
1.2.5. La jerarqu a aritmetica	8
1.2.6. Lemas de diagonalizacion	9
1.2.7. Resultados de incompletitud de G•odel	13
1.3. La teor a de conjuntos Zermelo-Fraenkel	14
1.4. Notacion para ordinales	16
2. Tarski o la verdad tipeada	18
2.1. Criterios para una de nicion de verdad	19
2.1.1. Correccion formal	19
2.1.2. Adecuacion material	20
2.2. De niciones tarskianas de verdad	25
2.2.1. El metodo recursivo	26
2.2.2. Una de nicion recursiva de verdad a secas	29
2.2.3. Nociones recursivas de verdad en una estructura	34
2.3. Limitaciones y jerarqu as	39
2.3.1. Inde nibilidad de la verdad	40
2.3.2. Jerarqu as semanticas	42
2.3.3. Jerarqu as axiomaticas	44
3. Teor as de la verdad no tipeadas	50
3.1. Las teor as de Punto Fijo de Kripke	51
3.1.1. Logicas paracompletas	53
3.1.2. Modelos de punto jo	57

3.1.3. <u>Contrapartes axiomáticas</u>	63
3.1.3.1. <u>Kripke-Feferman</u>	63
3.1.3.2. <u>Verdad Positiva</u>	66
3.2. <u>La dependencia de Leitgeb</u>	69
3.2.1. <u>Un resultado limitativo de McGee</u>	69
3.2.2. <u>Dependencia, autodependencia y d-fundación</u>	71
3.2.3. <u>Dependencia y desentrecomillación</u>	75
3.3. <u>Friedman-Sheard</u>	78
3.4. <u>Verdad transparente paraconsistente</u>	81
3.4.1. <u>Verdad transparente en la lógica paraconsistente básica</u>	81
3.4.1.1. <u>La lógica de las paradojas</u>	81
3.4.1.2. <u>Verdad transparente en LP</u>	83
3.4.2. <u>La teoría de la verdad de Priest</u>	84
3.4.2.1. <u>La lógica de Priest</u>	86
3.4.2.2. <u>Verdad transparente en PL</u>	88
3.4.3. <u>Abordajes subestructurales y verdad transparente</u>	90
3.4.3.1. <u>Una lógica intransitiva</u>	91
3.4.3.2. <u>Verdad transparente en ST</u>	93
4. <u>La verdad de acciónada</u>	95
4.1. <u>Breve prehistoria del de acciónismo</u>	96
4.2. <u>El rol de la verdad</u>	100
4.2.1. <u>La verdad como un mecanismo desentrecomillador</u>	100
4.2.2. <u>Las funciones lógico-expresivas de la verdad</u>	102
4.2.2.1. <u>Primera función</u>	102
4.2.2.2. <u>Segunda función</u>	103
4.2.3. <u>Los usos expresivos de la verdad</u>	105
4.2.3.1. <u>Usos retóricos y estilísticos</u>	106
4.2.3.2. <u>Usos epistémicos</u>	106
4.2.3.3. <u>Usos lógicos</u>	107
4.2.3.4. <u>Usos insurgentes de la verdad</u>	111
4.3. <u>El significado de la verdad</u>	117
4.3.1. <u>Intensión, extensión y reglas de uso</u>	118
4.3.2. <u>El carácter analítico de los principios de transparencia</u>	121
4.3.3. <u>Principios de transparencia vs de nociones explícitas de verdad</u>	125
4.3.4. <u>Principios de transparencia vs principios generales</u>	129
4.3.5. <u>El carácter extensional de los principios de transparencia</u>	133
5. <u>La quasi lógica de la verdad</u>	135
5.1. <u>Condiciones generales sobre una quasi lógica de la verdad</u>	136
5.2. <u>Expresando conjunciones infinitas mediante generalizaciones</u>	139
5.2.1. <u>Primer intento de elucidación: mismas consecuencias</u>	143
5.2.2. <u>Segundo intento de elucidación: mismo comportamiento inferencial</u>	147
5.2.2.1. <u>La propiedad de eliminación de la verdad</u>	147
5.2.2.2. <u>La propiedad de introducción de la verdad</u>	153
5.2.3. <u>El costo útil de las generalizaciones</u>	156
5.3. <u>Como debe lucir una teoría de acciónista de la verdad</u>	163

5.3.1.	El proyecto minimalista de Horwich	165
5.3.2.	De acionismo y para completitud	170
5.3.3.	De acionismo y composicionalidad	172
5.3.4.	De acionismo y para consistencia	173
6.	Condiciones sobre una noción de referencia formal	177
6.1.	La paradoja de Yablo	178
6.1.1.	Formalización y paradojidad de la secuencia de Yablo	181
6.1.2.	Autorreferencia en la secuencia de Yablo	183
6.2.	Condiciones de Leitgeb sobre la referencia	186
6.3.	Referencia y contenido en la literatura	189
6.3.1.	Referencia según Ryle	192
6.3.2.	Contenido lógico según Putnam	193
6.3.3.	Contenido lógico según Goodman	197
6.3.4.	Contenido lógico según Urbaniak	201
6.4.	La condición de equivalencia	205
7.	Referencia en el lenguaje de la aritmética de primer orden y sus extensiones	210
7.1.	Referencia a números	212
7.2.	Referencia a oraciones	216
7.3.	Referencia aletica	225
7.3.1.	Referencia aletica semantica	226
7.3.1.1.	Nociones basicas y su ejemplificación	227
7.3.1.2.	Indenibilidad y axiomatización	235
7.3.2.	Referencia aletica sintactica	239
7.3.2.1.	Nociones basicas, de indenibilidad y ejemplos	240
7.3.2.2.	Incompletitud y jerarquías	246
7.3.2.3.	Referencia estable	251
8.	Verdad minimalista	257
8.1.	Conexiones entre referencia aletica y dependencia	258
8.2.	Referencia y desentrecomillación	260
8.2.1.	Verdad bien fundada	262
8.2.2.	Verdad y referencia nita	270
8.2.3.	Verdad Equivalente	273
8.2.4.	Verdad y autorreferencia	275
	Conclusion	279
	Bibliografía	282
		296
	Indice alfabético	

Introducción

Alfred Tarski y la corriente de de acionista sacaron la verdad del ámbito de la metafísica para ubicarla en el ámbito de la lógica. Mientras que Tarski se encargó de despojarla de otros conceptos metafísicos comúnmente asociados (correspondencia, realidad, mundo, etc.) ofreciendo un abordaje puramente extensional, el de acionismo, impulsado por el trabajo de Tarski, la puso al nivel de una constante lógica.

De acuerdo con Tarski, el predicado de verdad satisface el esquema

`A' es verdadera si y solo si A. (Esquema T)

donde A es cualquier oración del lenguaje de cuyas expresiones se quiere predicar verdad y `A' un nombre para ella. En otras palabras, una oración es verdadera siempre y cuando lo que dice es el caso. Y, a la inversa, cualquier predicado que satisfaga el [Esquema T](#) y se aplique exclusivamente a oraciones de un lenguaje será un predicado de verdad para este lenguaje. En consecuencia, no parece haber necesidad de recurrir a conceptos metafísicos para caracterizar el predicado veritativo.

Los de acionistas, por su parte, fueron los primeros en advertir explícitamente que en enunciados como

La equivalencia entre masa y energía es verdadera. (1)

y

Todos los teoremas de la aritmética son verdaderos. (2)

el predicado de verdad no agrega contenido conceptual alguno, no ocurre para adscribir una propiedad a la equivalencia entre masa y energía o a los teoremas de la aritmética porque, como nota Tarski, una oración y su predicación de verdad son siempre equi-valentes. Ocurre, en cambio, para permitir expresar precisamente aquello de lo cual se predica verdad, en estos casos la equivalencia y los teoremas, sin la necesidad de articularlo explícitamente, i.e., sin la necesidad de enunciar la letra explícita de la equivalencia entre masa y energía o de todos y cada uno de los teoremas de la aritmética.

Si bien (1) puede tener un sentido diferente a "La energía en reposo de un cuerpo es el producto de su masa por la velocidad de la luz al cuadrado," su contenido proposicional es ciertamente el mismo y, por ende, son intercambiables en contextos extensionales salva veritate. Del mismo modo, (2) no expresa más que el contenido de todos y cada uno de los teoremas de la aritmética a la vez, de su 'conjunción in nita'.

Expresar uno o varios enunciados sin articularlos puede resultar necesario no solo por razones retóricas (énfasis) o epistémicas (podría no recordarse, por ejemplo, que establece la equivalencia entre masa y energía) sino incluso puramente lógicas: es imposible, al menos para los seres humanos, enunciar todos los teoremas de la aritmética porque estos son infinitos; el predicado de verdad, en cambio, permite hacerlo mediante una única expresión, (2).

El cuantificador universal posibilita expresar infinitas oraciones a la vez siempre y cuando estas difieren en un término. Por ejemplo, "Todos los seres humanos son mortales" puede verse como la expresión de todos los enunciados de la forma "m es mortal", donde m es el nombre de un ser humano. Pero si las oraciones no comparten su forma, como los teoremas de la aritmética, el cuantificador por sí solo no puede expresarlas todas a la vez. Sin embargo, con la ayuda del predicado veritativo esto sí es posible, porque predicar verdad de un enunciado no opera un cambio de contenido en el enunciado pero sí un cambio sintáctico; el enunciado, que originalmente estaba en posición oracional, pasa a estar en posición nominal, a ser un objeto en la oración, sobre el cual los cuantificadores sí son capaces de cuantificar. (2) puede verse como la expresión de todos los enunciados de la forma "m es verdadero", donde m es el nombre de un teorema de la aritmética. A su vez, estos enunciados expresan lo mismo que cada teorema, correspondientemente. Luego, los efectos de incorporar un predicado de verdad al lenguaje son comparables, al menos en este caso, a los de los cuantificadores, a los de una constante lógica.

Una de las tesis centrales del deíacionismo es que la única razón de ser del predicado de verdad es posibilitar la expresión de ciertos enunciados sin articularlos para, entre otras cosas, expresar 'conjunciones in nitas', posición a la cual adhiero y en favor de la cual argumento a lo largo de la tesis. Este rol lógico-expresivo de la verdad, que, sostengo, es su rol más importante aunque otros hayan sostenido posiciones diferentes, hace del predicado una cuestión lógica; exhorta a la búsqueda de lógicas o teorías formales de la verdad que extiendan los sistemas actuales de lógica para poder formular en ellas las teorías con las cuales se trabaja, v.g., la aritmética, y beneficiarse de estas bondades expresivas del predicado de verdad, esto es, la expresión de 'nuevas conjunciones in nitas'. El objetivo central de esta tesis es averiguar que principios formales

para el gobierno del predicado veritativo debe adoptarse para garantizar que el predicado pueda llevar a cabo este y sus otros roles expresivos, unica razon de su existencia.

La posicion mas popular al respecto sostiene que es necesario y suficiente adoptar formalizaciones del [Esquema T](#) o de sus versiones en forma de reglas, esto es, aquellas que permiten inferir 'A es verdadera' y viceversa, para todas las oraciones del lenguaje en el cual se trabaja. No obstante, Tarski ha demostrado que si la logica subyacente es clasica, esto da lugar a paradojas semanticas (que involucran el predicado de verdad) y, por tanto, a trivialidad. La mas famosa de estas antinomias es la paradoja del mentiroso, cuya formulacion original es atribuida a Eubulides de Mileto, del siglo IV antes de Cristo. Surge al considerar la oracion:

(L) es falsa. (L)

L es conocida como la 'oracion del mentiroso'. En virtud del [Esquema T](#), si se supone que es verdadera, se debe concluir que es falsa y viceversa, lo cual lleva a una contradiccion.

Por esta razon muchos teoricos de la verdad que buscan, como yo, principios que garanticen su rol expresivo, mayormente pero no necesariamente de acionistas, han optado por abandonar la logica clasica. En esta tesis de endo el camino contrario. Una elucidacion de que significa exactamente que expresiones como (2) expresen la 'conjuncion in nita' de los enunciados de los cuales predicen verdad dejar en claro que el [Esquema T](#) y principios semejantes no son ni necesarios ni suficientes para garantizar el rol logico-expresivo de la verdad. Aun mas, en contextos clasicos basta con adoptar la direccion de izquierda a derecha de los bicondicionales del [Esquema T](#), esto es, el esquema

Si 'A' es verdadera, entonces A. (T Out)

para toda oracion A del lenguaje, que, a diferencia del [Esquema T](#), no da lugar a contradicciones ceteris paribus. Consecuentemente, garantizar el rol expresivo de la verdad no es razon suficiente para adoptar logicas no clasicas.

Sin embargo, muchas instancias del [Esquema T](#) son correctas e inofensivas y puede resultar deseable, por razones de completitud o cuestiones practicas, contar tambien con ellas en una teoria clasica formal de la verdad. El problema yace en identificar cuales son estas instancias correctas, en dar criterios razonables de seleccion de instancias del [Esquema T](#). En la literatura se han propuesto muchos criterios de seleccion, ninguno de los cuales ha resultado del todo satisfactorio, al menos desde un punto de vista de acionista. Estos son o bien demasiado restrictivos, es decir, dejan de lado muchas instancias intuitivamente correctas del [Esquema T](#), o bien resultan arbitrarios, o bien demasiado complejos como para dar lugar a una teoria de la verdad cuyo proposito ultimo es hacer las veces de una logica, esto es, auxiliar al razonamiento.

Una estrategia prometedora e inexplorada consiste recurrir a la noción de referencia. La idea de que las expresiones que llevan a paradojas semánticas comparten ciertos patrones referenciales 'peligrosos' que son responsables por las contradicciones, que llamo 'visión referencialista de las paradojas semánticas', es considerada conocimiento común. Aun más, la postura ortodoxa sostiene que, más específicamente, estos patrones comunes se reducen a la autorreferencia, directa o indirecta, como en el caso de la oración L del mentiroso. En cualquier caso, si la visión referencialista es acertada, es posible un criterio de restricción de instancias del Esquema T basado en la noción de referencia que resulte consistente, amplio y con poder explicativo. Asimismo, la referencia es prima facie una noción de baja complejidad aritmética, por estar basada, aunque sea parcialmente, en la sintaxis de las expresiones referentes, lo cual apunta a la posibilidad de un criterio relativamente simple de restricción.

No obstante, el concepto de referencia que se necesita para dar cuenta de los patrones referenciales detrás de las paradojas semánticas no es fácil de caracterizar. Tampoco lo es la referencia para lenguajes formales en general. Lo evidencian los diversos intentos en la literatura pero, principalmente, el debate que se generó alrededor de la paradoja de Yablo. Estas dificultades han impedido hasta el momento verificar o refutar la visión referencialista y, a fortiori, su versión ortodoxa que, por ende, han adquirido cierto estatus de leyenda.

Stephen Yablo introdujo una antinomia que declaró 'no autorreferencial' para refutar la tesis referencialista ortodoxa, i.e., que todas las paradojas semánticas son autorreferenciales. La no autorreferencialidad de la paradoja de Yablo fue a su vez puesta en duda, suscitando un debate inextricable que culminó recién cuando Hannes Leitgeb advirtió que diversas nociones incompatibles y poco claras de autorreferencia estaban en juego en la discusión. Esto se tradujo en el 'desafío de Leitgeb' a encontrar una noción adecuada y precisa de referencia aplicable a lenguajes formales de primer orden y, en particular, al lenguaje de la verdad (usualmente el lenguaje de la aritmética de Peano de primer orden aumentado con un símbolo de predicado monádico para la verdad), que permita clasificar las expresiones de estos lenguajes según sus patrones de referencia. Como señala Leitgeb, de no encontrar una noción tal la autorreferencia debe ser eliminada del discurso científico como un pseudo concepto.

En esta tesis supero el desafío de Leitgeb, aunque dejo de lado su condición de equivalencia, un requisito de extensionalidad de acuerdo con el cual dos oraciones lógicamente equivalentes refieren a lo mismo, que, argumento, es inadecuado. En consecuencia, proporciono una serie de nociones de referencia hiperintensionales [i.e., que fallan la condición de equivalencia] y adecuadas desde un punto de vista material para el lenguaje de la aritmética de primer orden y para su extensión con un predicado de verdad.

De acuerdo con los nuevos conceptos, la paradoja de Yablo no es autorreferencial pero s mal fundada, i.e., no bien fundada, (el grafo que la relacion de referencia genera lo es), por lo cual queda refutada la postura referencialista ortodoxa. Pero la vision referencialista general resulta verdadera: todas las expresiones que dan lugar a paradojas son mal fundadas, lo cual queda claro porque la teoria de la verdad que proporciono dada por instancias del [Esquema T](#) generadas por las oraciones bien fundadas es no solo consistente sino tambien correcta. Como la nocion de referencia a partir de la cual de no buena fundacion, as como los otros criterios de seleccion, es de caracter sintactico y, por ende, de baja complejidad aritmetica, la nocion de buena fundacion, a diferencia de los conceptos usuales de fundacion como los de Kripke o Leitgeb, es de nivel en el lenguaje de la aritmetica y, en consecuencia, las teorías resultantes correspondientes son lo suficientemente sencillas como para ser empleadas como logicas, satisfaciendo uno de los requisitos mas importantes del de acionismo.

Adicionalmente, muestro que el criterio puede refinarse, sin aumentar la complejidad: las antinomias comparten patrones de referencia mal fundados con ciertas caracteristicas especificas, no cualquier mal fundacion es capaz de dar lugar a paradojas. Luego, introduzco teorías de la verdad mas abarcativas que aquella dada por instancias bien fundadas, para las cuales tambien pruebo correccion. En todos los casos, incluso cuando se emplea la buena fundacion como criterio de restriccion, las teorías tienen un gran poder de prueba, comparable con el de las mejores teorías a disposicion en la literatura, lo cual indica que el criterio de seleccion es lo suficientemente amplio. Ademias, en virtud de la tesis referencialista, los criterios empleados son claramente explicativos, no arbitrarios.

La tesis consta de 8 capitulos. El capitulo [1](#) establece los preliminares tecnicos y la notacion para toda la investigacion posterior.

El capitulo [2](#) esta dedicado a la presentacion del trabajo de Tarski en torno a la verdad. En [2.1](#) introduzco las condiciones que Tarski ha impuesto sobre toda definicion de verdad: la correccion formal y la adecuacion material. Mientras que la correccion formal (apartado [2.1.1](#)) establece que la definicion debe tomar la forma de una definicion explicita, la condicion de adecuacion material (apartado [2.1.2](#)) presupone una distincion entre el lenguaje para el cual se quiere definir verdad, el lenguaje objeto, y el lenguaje desde el cual se define verdad, el metalenguaje, y consiste en lo que Tarski denomina 'Convencion T': a grandes rasgos, una definicion de verdad dada en un metalenguaje debe implicar todas las instancias del [Esquema T](#) dadas por oraciones del lenguaje objeto, y el predicado de verdad solo puede aplicarse a estas ultimas. Por esta razon, se dice que las nociones tarskianas de verdad estan tipeadas o estratificadas.

La seccion [2.2](#) introduce el metodo de Tarski para definir verdad en [2.2.1](#) y luego expone en detalle los dos tipos de definicion de verdad que da para lenguajes formales: una absoluta (apartado [2.2.2](#)), en la cual las expresiones del lenguaje objeto tienen su significado usual, y otra relativa a una estructura (apartado [2.2.3](#)), donde las expresiones del lenguaje objeto son interpretadas de acuerdo con esta estructura dada.

Finalmente, la seccion [2.3](#) se ocupa de las limitaciones del abordaje tarskiano y esboza ciertas soluciones posibles. En [2.3.1](#) presento una version formal general del teorema de la indecibilidad de la verdad de Tarski (teorema [2.9](#)), que, grosso modo, prohíbe la identificacion de lenguaje y metalenguaje en toda definicion de verdad que se lleve a cabo en logica clasica, so pena de trivialidad, e introduzco el problema que presentan aquellos lenguajes y teorías que carecen de interpretacion 'natural' o estandar en el marco de Tarski: aquellas que se encuentran 'al nivel de la jerarquía' generada por la estratificación. En [2.3.2](#) considero entonces la posibilidad de extender la nocion de interpretacion para incluir estas interpretaciones 'naturales', mientras que en [2.3.3](#) exploro la via axiomática, sugerida pero desestimada por el mismo Tarski. El metodo axiomático introducido en este apartado se extendera a practicamente toda la tesis.

El capitulo [3](#) introduce una serie de teorías de la verdad que pretenden trascender la estratificación tarskiana, la distincion entre lenguaje y metalenguaje, si caer en trivialidad, ya sea abandonando la logica clasica o adoptando unicamente instancias correctas e inofensivas de versiones formales del [Esquema T](#), camino que tomo en esta tesis. Estas teorías son bien consideradas en la literatura y, al mismo tiempo, funcionales al hilo argumental de la tesis, en particular a los capitulos [5](#), [6](#) y [8](#). En [3.1](#) presento las teorías semanticas de Punto Fijo de Saul Kripke, que abandonan la logica clasica en favor de logicas para completas, seguidas por dos axiomatizaciones posibles de una de las teorías en logica clasica. [3.2](#) esta dedicada a la exposicion de la nocion de dependencia de Leitgeb y la teoría semántica de la verdad que en base a ella Leitgeb construye. Esta es una teoría que, en lugar de abandonar la logica clasica, busca dar un criterio de seleccion explicativo para la restriccion del [Esquema T](#), parte del objetivo de esta investigacion. En [3.3](#) presento una de las teorías axiomáticas clásicas más reconocidas en la literatura, el sistema Friedman-Sheard. Finalmente, en [3.4](#) exploro tres abordajes que abandonan la logica clasica y adoptan logicas paraconsistentes en su lugar para poder dar teorías de verdad libre de tipos.

El capitulo [4](#) es una reconstrucción racional de la postura de acionista con respecto a la verdad. Doy primero una breve historia de sus orígenes, para luego examinar lo que, considero, es la tesis central del de acionismo, i.e., que el predicado de verdad existe en el lenguaje unicamente para llevar a cabo funciones expresivas comparables a

las de las conectivas logicas usuales; principalmente, la expresion de 'conjunciones in-nitas' como (2). Argumento que estas funciones logico-expresivas son posibles gracias a lo que llamo 'funcion basica de la verdad' o 'naturaleza desentrecomilladora', y doy una serie de ejemplos. En la seccion 4.3 sostengo que la funcion basica de la verdad es la naturaleza de este predicado, esto es, la verdad es simplemente un mecanismo de desentrecomillacion, y de endo que esta naturaleza o signi cado es expresable mediante el [Esquema T](#) o principios analogos frente a una serie de objeciones en los apartados 4.3.1-4.3.5.

El cap tulo 5 esta basado en un trabajo conjunto con Thomas Schindler, titulado "Disquotation and Infinite Conjunctions" (cf. [Picollo & Schindler \(2015\)](#)). El objetivo es dar un conjunto de condiciones necesarias m nimas que toda teor a de la verdad de acio-nista |y, en general, toda teor a cuyo predicado sea capaz de llevar a cabo las funciones logicas que el de acionismo le adjudica| debe satisfacer, a la luz de la reconstruccion de esta posicion que ocupa el cap tulo anterior.

Tras dar ciertas condiciones generales, en 5.2 hago dos intentos por elucidar en que sentido expresiones como (2) son capaces de expresar la conjuncion in nita de todas las oraciones de las cuales predicen verdad, en este caso la conjuncion in nita de todos los teoremas de la aritmetica, de los cuales el segundo es exitoso (apartado 5.2.2), para luego investigar que principios formales para regular el comportamiento del predicado veritativo en el marco de una teor a de la verdad |v.g., el [Esquema T](#)| son necesarios para garantizar esta capacidad expresiva. Pero la cuestion central, de la cual me ocupo en 5.2.3, es si acaso estos principios son necesarios y su cientes para garantizar el cumplimiento de las funciones logico-expresivas del predicado veritativo y, en caso negativo, cuales lo son.

Contrario a la opinion generalizada, muestro que, si bien el [Esquema T](#) u otros principios semejantes expresan la naturaleza de la verdad, en sus versiones formales no son ni necesarios ni su cientes para garantizar el rol expresivo de la verdad. Mientras que ciertos principios logicos clasicos resultan indispensables, principios que algunas teor as presentadas en el cap tulo 3 dejan de lado, junto con ellos la formalizacion de [T](#) [Out](#) es no solo necesaria sino tambien su ciente. Luego, examino posibles objeciones y las respondo positivamente. Siendo esta 'mitad' del [Esquema T](#) consistente en logica clasica bajo condiciones normales, sostengo que las paradojas semanticas ya no son una razon para abandonar el terreno de la logica clasica.

Finalmente, en la seccion 5.3 listo las condiciones que toda teor a de acionista de la verdad debe satisfacer basandome en los resultados de las dos secciones anteriores y despues evaluo las teor as de la verdad introducidas en el cap tulo 3 a la luz de es-tos requisitos. Demuestro que muchas de estas teor as satisfacen versiones formales del

[Esquema T](#) o principios de verdad similares pero, no obstante, no logran garantizar las funciones logico-expresivas de su predicado de verdad, mientras que otras, sin satisfacer ningun principio de aquella ndole, s lo logran.

El capitulo [6](#) expone las dificultades que toda nocion de referencia para un lenguaje formal de primer orden debe enfrentar y, de este modo, permite identificar las condiciones que una definicion materialmente adecuada debe satisfacer necesariamente. En [6.1](#) hago una breve presentacion de la paradoja de Yablo, su formalizacion en el lenguaje de la verdad, su caracter paradójico y el debate en torno a su presunta no autorreferencialidad, junto con las objeciones de Leitgeb a las nociones de autorreferencia, confusas e incoherentes, que estaban en juego en la discusion. En [6.2](#) expongo las condiciones que de acuerdo con Leitgeb toda nocion de referencia formal debe satisfacer, en particular su 'condicion de equivalencia', junto con su mencionado desafio a encontrar nociones precisas que capturen las intuiciones basicas alrededor de la referencia. [6.3](#) es un repaso sucinto de los intentos previos de Gilbert Ryle, Hilary Putnam, Nelson Goodman y Rafal Urbaniak de definir referencia u otros conceptos cercanos, señalando sus aciertos y desaciertos a la luz de las condiciones de Leitgeb. Por ultimo, en [6.4](#) argumento en contra de la condicion de equivalencia de Leitgeb y, en general, a favor de la hiperintensionalidad de la referencia, de modo tal que queda claramente que condiciones debe satisfacer una nocion adecuada de referencia para lenguajes formales.

El capitulo [7](#) ofrece nociones de referencia para el lenguaje de la aritmetica de primer orden y el lenguaje de la verdad que satisfacen las condiciones enumeradas en el capitulo anterior. En [7.1](#) doy una definicion de referencia de oraciones del lenguaje de la aritmetica a numeros naturales, sus objetos primarios, mientras que en [7.2](#) doy una definicion de referencia de oraciones a oraciones del mismo lenguaje, a traves de los numeros naturales que las codifican y, a partir de ella, doy definiciones de autorreferencia y buena fundacion. En ambos casos las nociones son semanticas, definidas en funcion de la nocion de verdad en el modelo estandar de la aritmetica, y, en consecuencia, altamente complejas. Por otro lado, son hiperintensionales y materialmente adecuadas hasta donde las peculiaridades del lenguaje lo permiten, como ah ejemplo. La ultima nocion permite dar cuenta de la autorreferencialidad de las oraciones de Gödel y otros fenomenos, as como de la infundacion de las cadenas de oraciones, cada una de las cuales refiere a la que le sigue.

La seccion [7.3](#) se ocupa de definir referencia de oraciones a oraciones del lenguaje de la verdad, la extension del lenguaje de la aritmetica con un simbolo de predicado para expresar verdad, pero este es un tipo especial de referencia, una referencia que solo tiene en cuenta aquellas expresiones que caen bajo el alcance del predicado de verdad en una oracion, de la referencia a traves del predicado veritativo o, como la llamo, 'referencia

aletica'. Ofrezco dos nociones de referencia aleatica, una semantica (apartado [7.3.1](#)) y otra sintactica (apartado [7.3.2](#)), ya no en funcion de la verdad en un modelo dado sino del predicado de prueba de la teoria en la cual se trabaja, y muestro que ambas dan veredictos intuitivamente correctos sobre ciertas expresiones paradójicas y otras patologías. La paradoja de Yablo y sus formulaciones alternativas resultan no autorreferentes y, de este modo, queda refutada la version ortodoxa de la tesis referencialista acerca de las paradojas.

En el apartado [7.3.1](#) demuestro asimismo la indecibilidad de la referencia aleatica semantica, resultado que puede extenderse facilmente a las nociones de referencia semanticas introducidas en las dos secciones previas, y ofrezco una axiomatizacion posible por medio de un predicado primitivo. En [7.3.2](#), en cambio, muestro que la referencia aleatica sintactica es decible en el lenguaje de la aritmetica, así como las nociones derivadas de autorreferencia y buena fundacion aleaticas sintacticas, a diferencia de los conceptos de fundacion de Kripke o Leitgeb. No obstante, señalo que las nociones sintacticas generan cierta estratificación que es preciso neutralizar mediante jerarquías y la restriccion a oraciones con ciertas características particulares.

En el ultimo capítulo de la tesis, empleo la noción de referencia aleatica sintactica dada en el capítulo anterior para la formulacion de criterios de seleccion de instancias de una version formal del [Esquema T](#), dando lugar a una serie de teorías clásicas de la verdad. Tras ofrecer algunos resultados que servirán luego para establecer la consistencia de estos sistemas, en la seccion [8.2](#) exploro cuatro criterios posibles y las teorías de la verdad a las cuales dan lugar, correspondientemente.

A grandes rasgos, en [8.2.1](#) restrinjo el [Esquema T](#) a oraciones aleaticamente bien fundadas y sus equivalentes, y pruebo que el sistema resultante es correcto y deductivamente tan poderoso como las mejores teorías de la verdad disponibles en la literatura, lo cual muestra que es bien abarcativo. Esto constituye una verificación de la vision referencialista de las paradojas, de acuerdo con la cual estas comparten ciertos patrones referenciales que otras oraciones inofensivas no tienen: todas las oraciones paradójicas son infundadas, y no muchas expresiones infundadas son no paradójicas, porque pocas instancias inofensivas del [Esquema T](#) han quedado fuera.

En [8.2.2](#) y [8.2.3](#) muestro que el criterio de seleccion dado por la infundacion puede reñarse al menos en dos direcciones: o bien admitiendo, además de las instancias bien fundadas, aquellas que no son autorreferenciales y refieren directamente solo a un numero finito de otras oraciones, o bien aquellas que no son autorreferenciales ni refieren a expresiones autorreferenciales y refieren directamente solo a un numero finito de otras oraciones, o sus equivalentes. En ambos casos muestro que las teorías son correctas y, por incluir las instancias bien fundadas, deductivamente poderosas.

Dado que las tres teorías introducidas hasta el momento son consistentemente extensibles con una versión formal de [T_Out](#), que garantiza la capacidad de sus predicados de cumplir las funciones lógico-expresivas de la verdad, y sus criterios de restricción son explicativos, abarcativos y tan simples como de niveles en el lenguaje de la aritmética de primer orden, sostengo que estas extensiones son buenas candidatas a teoría de acciónista clásica de la verdad. En cambio, la teoría dada en [8.2.4](#), cuyas instancias del [Esquema T](#) están restringidas únicamente a oraciones no autorreferentes, si bien es consistente, no es correcta y, por tanto, no es una buena teoría de acciónista.

Finalmente, en el último apartado de la tesis recapitulo los resultados obtenidos y extraigo algunas conclusiones.

Existen al menos cinco modos de leer esta tesis. Uno de ellos es hacerlo simplemente de corrido, pero es posible seguir otros caminos, según los conocimientos e intereses que se tenga. Tengase en cuenta en cualquier caso que los capítulos [1-3](#) son introductorios. Mientras que el primero puede resultar necesario para establecer algunas cuestiones de notación, a menos que se esté familiarizada con la notación de Halbach, los capítulos [2](#) y [3](#) son meramente expositivos; su lectura puede omitirse si ya se conoce su contenido. En última instancia, a lo largo del texto habrá referencias a definiciones, teorías y resultados que aparecen en estos capítulos, a los cuales podrá volverse de ser necesario.

Aquellas personas que solo estén interesadas en lo escrito sobre verdad y no sobre referencia pueden saltar los apartados [6.1.2](#), [6.2-6.4](#) y [7.1-7.3.1](#). Aquellas, en cambio, que se interesen exclusivamente en las nociones de referencia pueden leer únicamente los capítulos [1](#), [6](#) y [7](#). Si se está interesado solo en lo que concierne a teorías formales de la verdad puede comenzarse por los capítulos [1-3](#) y pasar directamente al apartado [7.3.2](#) y al capítulo [8](#). Finalmente, si el interés está en lo escrito acerca del de acciónismo, puede leerse únicamente los capítulos [1-5](#).

1

Preliminares

En este capítulo doy los preliminares técnicos para toda la tesis y establezco cuestiones de notación.

Para evitar formulaciones engorrosas muchas veces omito las comillas que marcan la mención de una expresión para diferenciarla de su uso cuando se trata de una expresión perteneciente a un lenguaje formal.

Con frecuencia, abrevio expresiones castellanas como 'si y solo si' o 'siempre y cuando' mediante 'sii' o el símbolo '↔', y 'si . . . entonces . . .' mediante '⇒', por cuestiones de simpleza y legibilidad.

1.1. Lenguajes y teorías de primer orden

En esta investigación trabajo casi exclusivamente con lenguajes de primer orden. Un lenguaje de primer orden L es un conjunto recursivo de fórmulas construidas a partir de un vocabulario que contiene \neg , \wedge , \vee y $=$ como sus únicos símbolos lógicos, un stock infinito enumerable de variables de individuo $x_1, x_2, \dots, x_n, \dots$, y los parentesis izquierdo y derecho como símbolos auxiliares. Otros operadores lógicos como \rightarrow , \leftrightarrow , \exists y \forall son definidos en función de \neg , \wedge y \vee como es usual. Además, L cuenta con un stock enumerable (finito o infinito) de constantes de individuo, una cantidad enumerable de símbolos de función de cualquier aridad y una cantidad finita de símbolos de relación de cualquier aridad. Si bien esta última restricción no es necesaria, es extremadamente conveniente en términos de complejidad notacional.

Con frecuencia utilizo las letras x, y, z, \dots en lugar de x_1, x_2, x_3, \dots , por comodidad. Las nociones de término, variable libre, variable ligada, fórmula, oración, etc. son

las usuales. Como metavariables empleo las letras $v; u$ o $v_1; v_2; v_3; \dots$ para variables de individuo, $a; b; c; \dots$ o $a_1; a_2; a_3; \dots$ para constantes de individuo, $t; s$ o $t_1; t_2; t_3; \dots$ para términos en general, $f; g; h; \dots$ para símbolos de función, $P; Q; R; \dots$ para símbolos de relación, $'; \dots; \dots; \dots; \dots$ para fórmulas y $\{ \dots; \dots; \dots; \dots; \dots \}$ para conjuntos de oraciones del lenguaje. Muchas veces escribo $\langle v_0; \dots; v_n \rangle$ en lugar de $'$ para indicar que $v_0; \dots; v_n$ son las únicas variables libres en $'$ (y lo mismo para otras metaletas de fórmula y variables de individuo).

Sea $\langle t=s \rangle$ la fórmula que resulta de reemplazar en $'$ todas las ocurrencias libres de s por t , esto es, o bien s es un término cerrado, o bien contiene variables y se reemplazan únicamente las ocurrencias de s cuyas variables estén libres en $'$.

Un modelo M para L consta de un dominio $|M|$ sobre el cual varían los cuantificadores del lenguaje y asigna un objeto M^M construido a partir de $|M|$ a cada símbolo lógico de L como su interpretación, como es usual. $M \models '$ indica que $'$ es verdadera de acuerdo con la interpretación de L dada por M y la lógica I . Cuando I es la lógica clásica omito el subíndice.

Una teoría axiomática (o sistema formal) th formulada en L es un subconjunto recursivo de oraciones cerrado bajo los principios y reglas lógicas de un cálculo nitario o lógica I . $th \models_1 '$ indica que $' \in th$ y que I es la lógica subyacente. Si I es el cálculo usual de la lógica clásica, omito el subíndice. Cuando no de lugar a confusión, omito también el calificativo 'axiomática'.

Definición 1.1 (Interpretaciones relativas). Sean th y th^0 dos teorías clásicas formuladas en lenguajes de primer orden L y L^0 , respectivamente. th es relativamente interpretable en th^0 si existe una función de traducción $\cdot : L \rightarrow L^0$ tal que, para cualesquiera fórmulas $'$ de L (en adelante, $' \in L$)

- $\langle \cdot \rangle = \langle \cdot \rangle$,
- $\langle \neg \cdot \rangle = \langle \cdot \rangle \neg \langle \cdot \rangle$, y
- $\langle \exists v \cdot \rangle = \exists v \langle \cdot \rangle$, para alguna fórmula $\cdot \in L^0$; y

$$th \models \langle \cdot \rangle \iff th^0 \models \langle \cdot \rangle$$

1.2. La aritmética de primer orden

1.2.1. Lenguaje y teoría

L_{pa} es el lenguaje de primer orden en el cual se formulan usualmente los axiomas de la aritmética de Peano, pa . Cuenta únicamente con una constante de individuo 0 para el número 0, un símbolo monádico de función S para la función sucesor y dos símbolos diádicos de función $+$ y para la suma y la multiplicación, respectivamente. El término que consiste de n ocurrencias de S seguidas inmediatamente por 0 es el numeral de n , que noto \bar{n} .

pa contiene los axiomas usuales para 0; S ; $+$ y \cdot , más las infinitas instancias del axioma de inducción dadas por fórmulas de L_{pa} . La aritmética de Robinson q es idéntica a pa sin el esquema de inducción.

Sea \mathbb{N} el conjunto de los números naturales. \mathbb{N} es el modelo estándar de L_{pa} , donde $\mathbb{N} = \{n \in \mathbb{N}\}$ y todos los símbolos del vocabulario no lógico de L_{pa} reciben su interpretación pretendida.

Definición 1.2 (!-consistencia). Una teoría th formulada en un lenguaje L que extiende L_{pa} (en adelante, $L \supseteq L_{pa}$) es !-inconsistente si existe una fórmula $\exists v \varphi(v)$ tal que $th \vdash \varphi(\bar{n})$ para cada $n \in \mathbb{N}$ y, a la vez, $th \not\vdash \exists v \varphi(v)$.

Luego, si th es !-inconsistente, ninguna extensión de \mathbb{N} a L es un modelo de th .

1.2.2. Definibilidad y representabilidad

Sea $L \supseteq L_{pa}$ y th una teoría formulada en L (en adelante, $th \subseteq L$).

Definición 1.3 (Definibilidad). Una fórmula $\exists v_1 \dots \exists v_n \varphi(v_1, \dots, v_n)$ de L define una relación R de aridad n sobre \mathbb{N} (en adelante, $R \subseteq \mathbb{N}^n$) si

$$\mathbb{N} \models \varphi(k_1, \dots, k_n) \iff \langle k_1, \dots, k_n \rangle \in R$$

Definición 1.4 (Representabilidad). Una fórmula φ tal para R , R es definible en L .

Siendo los conjuntos o (la extensión de las) propiedades relaciones de aridad 1 y las funciones un tipo especial de relación, esta noción se aplica también a conjuntos y a funciones.

Definición 1.4 (Representabilidad débil). Una fórmula $\varphi(v_1; \dots; v_n) \in L$ representa débilmente una relación $R \subseteq \mathbb{N}^n$ en \mathcal{L} si

$$\mathcal{L} \models \varphi(k_1; \dots; k_n) \iff (k_1; \dots; k_n) \in R$$

De existir una fórmula φ tal, R es débilmente representable en \mathcal{L} .

Definición 1.5 (Representabilidad de relaciones). Una fórmula $\varphi(v_1; \dots; v_n) \in L$ representa una relación $R \subseteq \mathbb{N}^n$ en \mathcal{L} si

1. $\mathcal{L} \models \varphi(k_1; \dots; k_n) \iff (k_1; \dots; k_n) \in R$; y
2. $\mathcal{L} \models \varphi(k_1; \dots; k_n) \iff (k_1; \dots; k_n) \in R$.

De existir una fórmula φ tal, R es representable en L .

Para la representabilidad de funciones suele reforzarse el segundo requisito.

Definición 1.6 (Representabilidad de funciones). Una fórmula $\varphi(v_1; \dots; v_n; v) \in L$ representa una función n -ádica f sobre \mathbb{N} ($f: \mathbb{N}^n \rightarrow \mathbb{N}$) en \mathcal{L} si

1. $\mathcal{L} \models \varphi(k_1; \dots; k_n; k) \iff f(k_1; \dots; k_n) = k$; y
2. $\mathcal{L} \models \forall v_1 \dots \forall v_n \exists v (\varphi(v_1; \dots; v_n; v) \wedge \forall v_{n+1} (\varphi(v_1; \dots; v_n; v_{n+1}) \implies v_{n+1} = v))$.

De existir una fórmula φ tal, f es representable en L .

El punto 2 de esta definición implica el punto dos de la definición 1.5, si f hace las veces de la relación R .

Todos los subconjuntos recursivamente enumerables o semirecursivos de \mathbb{N}^n , con $n \in \mathbb{N}$, son definibles en L_{pa} y débilmente representables en q y, por tanto, en toda teoría que la contenga; v.g. pa . Adicionalmente, un conjunto o relación es representable en q si y solo si es recursivo. Por último, toda función primitiva recursiva (p.r.) es representable en q .¹ Por ejemplo, las relaciones 'es menor que' y 'es menor o igual que' entre números naturales son recursivas y, por tanto, definibles en L_{pa} y representables en q mediante fórmulas $u < v$ y $u \leq v$, respectivamente.

¹Vease [George Boolos, John P. Burgess, & Richard C. Jeffrey \(2007, caps 6 y 7\)](#) para una definición de las nociones de conjunto y función recursivos, y relaciones o conjuntos semirecursivos o recursivamente enumerables; vease [Peter Smith \(2007, cap. 9\)](#) para una exposición detallada de las relaciones y funciones que q es capaz de representar, débilmente y simpliciter.

1.2.3. Códigos de Gödel y aritmética como teoría de la sintaxis

Es posible codificar la sintaxis de cualquier lenguaje de primer orden L mediante números naturales llamados 'códigos de Gödel', esto es, asignando a cada secuencia finita de símbolos del vocabulario de L un y solo un número natural. Luego, a través de sus códigos, L_{pa} contiene un término que denota cada una de las expresiones de L y puede, por tanto, hablar de ellas (cf. Boolos et al. (2007, cap. 15)). Si n es el código de una secuencia finita de símbolos de L [que noto $\#()$] p q es el numeral de n , i.e., p q es un modo alternativo a n de notar el término que se obtiene al anteponer n veces el símbolo S al símbolo 0 . Asumo que el número 0 no codifica ningún símbolo o secuencia de símbolos de L , por cuestiones técnicas.

También es posible codificar secuencias finitas de números naturales o, si se quiere, códigos de expresiones de un lenguaje, mediante números naturales. Si la secuencia $\langle k_1; \dots; k_n \rangle \in \mathbb{N}^n$, $\#(\langle k_1; \dots; k_n \rangle)$ es su código. Por diversas razones que quedarán claras luego trabajo únicamente con códigos efectivos y monotónicos. Un código de Gödel es efectivo si y solo si existe un algoritmo tal que, dado un número natural, indica tras nito de pasos que objeto sintáctico codifica, si es que codifica uno, y, viceversa, dado un objeto sintáctico cualquiera devuelve su código en un número nito de pasos. Una codificación es monotona si, cada vez que una secuencia de símbolos ocurre propiamente en otra secuencia más extensa, el código de la primera es menor que el de la segunda. Diferentes codificaciones pueden dar lugar a diferentes resultados, por lo cual es preciso ser cuidadosa. Bajo codificaciones no monotónicas principios usualmente inofensivos se pueden tornar inconsistentes.²

A través de sus códigos, L_{pa} y q pueden definir y representar una serie propiedades, relaciones y funciones de los objetos codificados. Sea L un lenguaje de primer orden y $th L$ una teoría. En la siguiente tabla, las propiedades de secuencias o secuencias de secuencias de símbolos de L a la derecha encuentran a su izquierda (variantes notacionales de) las fórmulas de L_{pa} que definen los respectivos conjuntos de códigos de estas secuencias naturalmente.³ Para evitar este discurso engorroso, en adelante cuando sea conveniente y claro identifico directamente las secuencias de símbolos del lenguaje codificado L con sus respectivos códigos. Asimismo, con frecuencia identifico también los conjuntos $\langle ; ; ; \dots \rangle$ de oraciones de un lenguaje con los conjuntos de sus códigos correspondientes. Como los conjuntos a la derecha de la tabla son recursivos, las fórmulas a la izquierda representan las propiedades a su derecha en q .

²Vease, v.g., Richard Jr Heck (2007) y Volker Halbach & Albert Visser (2014a,b).

³Si bien no es demasiado preciso, por 'natural' me refero a una de nición que formalice lo más lealmente posible la de nición informal que se tiene en mente. De niciones no naturales de ciertos conjuntos pueden arrojar resultados indeseables; vease Halbach & Visser (2014a,b).

-
- CIT_{ermL}(v) := v es un termino cerrado de L
 - Var(v) := v es una variable de L de L
 - At_L(v) := v es una formula atomica
 - Form_L(v) := v es una formula de L
 - Sent_L(v) := v es una oracion de L de L
 - Seq_L(v) := v es una se secuencia de formulas
 - SubForm_L(u; v) := u es una subformula de v en L
 - U_L(v₁; v₂; v₃) := v₂ es una secuencia de variables de individuo y v₁ es una oracion de L que consta de la formula v₃ precedida por una secuencia de cuantificadores universales sobre las variables en v₂ en ese orden
 - Prf_{th}(u; v) := u es una prueba de v en th

'V ar' no lleva sub ndice dado que todos los lenguajes de primer orden comparten las mismas variables de individuo. Cuando es lo su cientemente claro de que lenguaje se trata, omito el sub ndice 'L'. 8' es un modo abreviado de escribir 8v(CIT_{ermL}(v) !').

Cuanti cando existencialmente sobre u en Prf_{th}(u; v) se obtiene una formula de L_{pa}, 9uPrf_{th}(u; v),⁴ que representa debilmente el conjunto semirecursivo (pero no recursivo) de teoremas de th en q, y que usualmente se nota Bew_{th}(v). Si Prf_{th}(u; v) de ne la relacion 'ser una prueba de' de modo natural, Bew_{th}(v) satisface las condiciones de teoremicidad de L•ob (cf. [Martin H. L•ob \(1955\)](#)) para cualesquiera oraciones ' y del lenguaje de la teor a:

1. th ` ') th ` Bew_{th}(p'q)
2. th ` Bew_{th}(p' ! q) ! (Bew_{th}(p'q) ! Bew_{th}(p q))
3. th ` Bew_{th}(p'q) ! Bew_{th}(pBew_{th}(p'q)q)

Ademas, las funciones que toman como argumentos las entradas de la segunda columna y devuelven aquellas de la tercera en la siguiente tabla son todas p.r. y, por tanto, representables en q.⁵ Sea R un s mbolo de relacion n-adico, t₁; : : : t_n; t; s terminos, v una variable y ; 2 L.

⁴Que, como 9 no es un s mbolo del lenguaje, es en realidad un modo abreviado de escribir :8v₁;Prf_{th}(v₁; v).

⁵Si estas funciones se aplican a argumentos que no responden a la forma indicada en la segunda columna, devuelven 0.

	$h\#(t_1; \dots; \#(t_n))i$	$\#(Rt_1 \dots t_n)$	$R_i(v_1; \dots; v_n)$
	$\#(')$	$\#(:')$	$:_v$
	$h\#('); \#()i$	$\#(' !)$	$u!_v$
	$h\#('); \#(v)i$	$\#(8v')$	$8_v u$
	$\#('(v_1; \dots; v_n))$	n	$free(v)$
Sustitucion	$h\#('); \#(t); \#(s)i$	$\#('[t=s])$	$v_1(v_2=v_3)$
	$h\#('(\sim v)); \#(hki)i$	$\#('(k))$	$\sim s_{(u; v)}$
Diagonalizacion	$\#('(v))$	$\#(8v(v = p'q ! '))$	
Numeral	n	$\#(n)$	$num(v)$ o \underline{v}
	$\#(hk_1; \dots; k_n)i$	n	$lg(v)$
	$h\#(hk_1; \dots; k_n)i; m_i$	k_m	$(u)_v$

Asimismo, si q th y L contiene exclusivamente s mbolos de funcion para funciones p.r. sobre ! de nidas en th de modo natural,⁶ la funcion valor, que toma el codigo de un termino y devuelve su denotacion, tambien es representable en q .

Sera a menudo conveniente y algunas veces necesario contar con s mbolos de funcion para las funciones p.r. detalladas para la formulacion de axiomas y principios a lo largo de la tesis. Sea L^+_{pa} el resultado de expandir L_{pa} con un numero nito de s mbolos de funcion para funciones p.r. (a especi car), incluyendo aquellos que se encuentran en la columna de la derecha de la tabla anterior. La funcion diagonalizacion no tendra un s mbolo de funcion propio sino que es representada en q por la formula de L_{pa} que noto $Diag(u; v)$, siguiendo la de nicion [1.6](#).

Por su parte, la funcion valor no puede tener un s mbolo de funcion que la presente, so pena de caer en contradicciones; solo puede ser representada en q por una formula $'(u; v)$. No obstante, por comodidad escribo $u = v$ en su lugar, aunque ` ` no sea un s mbolo de funcion del lenguaje.

Adopto dos s mbolos de funcion alternativos para la funcion numeral por cuestiones de legibilidad. Esta funcion suele utilizarse para cuanti car dentro de las comillas $p:q$ de Godel. Por ejemplo, si se quiere expresar que $'(x)$ es demostrable en pa para todos los numeros, $8xBew_{pa}(p'(x)q)$ no sirve, porque las comillas de Godel indican un numeral, un termino cerrado. Recurriendo a las funciones numeral y sustitucion es posible reintroducir la variable libre en $p'(x)q$ y expresar lo que se quer a mediante la formula $8xBew_{pa}(p'q(x\equiv pxq))$, que suelen abreviarse como $8xBew_{pa}(p'(x)q)$, siguiendo la notacion dot de Feferman. Asimismo, expresiones como $8tBew_{pa}(p'(t.)q)$ son abreviaturas de $8tBew_{pa}(p'q(t=pvq))$, donde v es la unica variable libre en $'$ y el punto debajo de t indica que t esta libre en la formula que sigue al cuanti cador.

⁶Uno de los requisitos de naturalidad es que los de niens sean lo mas simple posibles en terminos de complejidad cuanti cacional, i.e., que sean o (vease el apartado [1.2.5](#)).

Sean q^+ y pa^+ la aritmetica de Robinson y la de Peano formuladas en L_{pa}^+ con axiomas que den en los nuevos simbolos de funcion en terminos de (alguna de) las formulas que representaban naturalmente estas funciones en q . N^+ es el modelo estandar de L_{pa}^+ , que extiende N a todo el lenguaje. Nuevamente, $jN^+j = \omega$ y todos los nuevos simbolos de funcion L_{pa}^+ reciben tambien su interpretacion pretendida.

1.2.4. El lenguaje de la verdad

Sea L_T el resultado de extender L_{pa}^+ con un simbolo de predicado monadico T para la verdad. L_T es el lenguaje de buena parte de las teorías formales de la verdad que juegan un rol en esta investigacion.

Sea pat la teoría que resulta de formular pa^+ en L_T , incluyendo todas las instancias del esquema de axioma de induccion generadas por oraciones de L_T . Si \mathcal{I} , $\mathcal{I} \models N^+$; \mathcal{I} es el modelo que resulta de extender N^+ a todo L_T , asignando a T como su extension.

1.2.5. La jerarquía aritmetica

Las formulas de L_{pa}^+ se pueden clasificar de acuerdo con su complejidad cuantificacional, dando lugar a una jerarquía de conjuntos Σ_n y Π_n en L_{pa}^+ con $n \geq 1$ conocida como 'jerarquía aritmetica'. Si ϕ es logicamente equivalente a una formula que solo contiene cuantificadores acotados, $\phi \in \Sigma_1$ pertenece tanto a Σ_0 como a Π_0 . Si ϕ es logicamente equivalente a una formula de la forma $\exists v_1 \dots \exists v_n \psi$, donde $\psi \in \Sigma_n$, entonces $\phi \in \Sigma_{n+1}$. Si ϕ es logicamente equivalente a una formula de la forma $\forall v_1 \dots \forall v_n \psi$, donde $\psi \in \Sigma_n$, entonces $\phi \in \Pi_{n+1}$. Finalmente, si $\phi \in \Sigma_n \setminus \Pi_n$, $\phi \in \Sigma_n$. Toda formula de L_{pa}^+ es o bien Σ_n o bien Π_n para algun $n \geq 1$. Como es posible agregar cuantificadores superfluos a cualquier enunciado de L_{pa}^+ , si $\phi \in \Sigma_n$ o Π_n , tambien es Σ_m y Π_m para cada $m > n$.

Una relacion sobre ω es de nible en L_{pa}^+ mediante una formula ϕ si y solo si es recursiva; mientras que es de nible mediante una formula ψ si y solo si es semirecursiva. En cambio, si es de nible exclusivamente mediante formulas Σ_n o Π_{n+1} donde $n > 0$, la relacion no es siquiera semirecursiva. En consecuencia, $q(q^+)$ prueba todas las oraciones ψ verdaderas en $N(N^+)$. Este resultado se conoce como 'completitud ψ ' de $q(q^+)$.

Por ejemplo, si la formula $Prf_{th}(u; v)$ del apartado 1.2.3 define la relacion de prueba en th de modo natural, $Prf_{th}(u; v) \in \Sigma_0$, porque la relacion 'ser una prueba de'
⁷I.e., todos los cuantificadores $\exists v$ en ψ estan seguidos por $(v < t)$ para algun termino t y una formula ψ .

es recursiva. $Bew_{th}(v)$, en cambio, es Σ_1 , porque se obtiene anteponiendo un cuantificador existencial a $P_{rf_{th}}(u; v)$. Además, si th es consistente, $Bew_{th}(v) \neq \Sigma_1$, porque el conjunto de teoremas de th no es recursivo sino solo semirecursivo.

El lenguaje de la aritmética de segundo orden, esto es, el resultado de extender L_{pa}^+ con variables V que pueden ocupar la posición de predicados y sus cuantificadores $\forall V$ correspondientes (cf. [Stewart Shapiro \(1991\)](#)), permite continuar la jerarquía aritmética, ahora llamada 'jerarquía aritmética', en la cual hay lugar también para las nuevas expresiones. Luego, en vez de un único índice, ahora las letras v estarán acompañadas por un subíndice y un superíndice: Σ_n^1 y Σ_n^2 . El superíndice se mantiene fijo, es siempre 1; indica que se trata de una fórmula de primer orden, precedida por la cantidad de bloques de cuantificadores de segundo orden que indica el subíndice, que v varía, comenzando por el tipo de cuantificador que indica la letra que se utiliza: esta por los cuantificadores universales mientras que por los existenciales, o mejor dicho, por los universales precedidos por una negación.

Al igual que antes, si una fórmula ϕ del lenguaje es lógicamente equivalente a una fórmula que no contiene cuantificadores de segundo orden o donde estos están acotados,⁸ ϕ pertenece tanto a Σ_0^1 como a Σ_0^2 . Si ϕ es lógicamente equivalente a una fórmula de la forma $\forall v_1 : \dots : v_n, \psi$, donde $\psi \in \Sigma_n^1$, entonces $\phi \in \Sigma_{n+1}^1$. Si ϕ es lógicamente equivalente a una fórmula de la forma $\exists v_1 : \dots : v_n, \psi$, donde $\psi \in \Sigma_n^1$, entonces $\phi \in \Sigma_{n+1}^1$. Finalmente, si $\phi \in \Sigma_n^1 \setminus \Sigma_n^2$, $\phi \in \Sigma_n^2$. Toda fórmula es o bien Σ_n^1 o bien Σ_n^2 para algún $n \in \mathbb{N}$.

1.2.6. Lemas de diagonalización

Uno de los logros más grandes de [Kurt Gödel \(1931\)](#) fue mostrar que la autorreferencia es posible en teorías formuladas en lenguajes formales como L_{pa} y sus extensiones de primer orden. Esto es lo que establecen los lemas diagonales, piezas claves en la prueba de sus resultados de incompletitud. Sea L un lenguaje de primer orden y th una teoría formulada en L . Asumo una codificación de L en \mathbb{N} . Sea $\sim v$ una abreviatura de $\forall v_1; \dots; v_n$.

Teorema 1.7 (Lema diagonal débil). Para toda fórmula $\phi(v; \sim) \in L$ existe una fórmula $\psi(\sim v) \in L$ tal que

$$th \vdash (\sim v) \leftrightarrow \phi(\sim v; \sim v)$$

Demostración. Considerese la fórmula

$$\exists v(\text{Diag}(u; v) \wedge \neg \phi(v; \sim)) \tag{1.1}$$

⁸i.e., todos los cuantificadores $\forall V$ en ϕ están seguidos por $(\forall X!)$ para alguna variable de predicado X y una fórmula ψ .

$$\exists u(u = \exists v(\text{Diag}(u; v) \wedge \varphi(v; \sim v))) \wedge \exists v(\text{Diag}(u; v) \wedge \varphi(v; \sim v)) \quad (1.2)$$

Sea m el numeral del número de Gödel de (1.2). Por las leyes de la identidad, esta fórmula es equivalente en \mathcal{L} a

$$\exists v(\text{Diag}(\exists v(\text{Diag}(u; v) \wedge \varphi(v; \sim v)); v) \wedge \varphi(v; \sim v)) \quad (1.3)$$

que, en virtud de que $\text{Diag}(u; v)$ representa una función en \mathcal{L} y, por ende, debe satisfacer la condición 2 de la definición 1.6, es a su vez es equivalente en \mathcal{L} a

$$\exists m(\sim v)$$

puesto que (1.3) afirma precisamente que la diagonalización de (1.1) [alias (1.2) o la denotación de m] es un \exists . Luego,

$$\mathcal{L} \models \exists u(u = \exists v(\text{Diag}(u; v) \wedge \varphi(v; \sim v))) \wedge \exists v(\text{Diag}(u; v) \wedge \varphi(v; \sim v)) \iff \exists m(\sim v)$$

□

Grosso modo, el teorema 1.7 establece que para cada predicado φ de \mathcal{L} hay una fórmula del lenguaje que es equivalente en \mathcal{L} a $\exists x(\varphi(x))$, que dice de φ que es un \exists .

Dado cualquier predicado $\varphi(v)$, el lema diagonal débil permite probar únicamente una equivalencia entre una fórmula $\exists x(\varphi(x))$ y $\exists v(\varphi(v))$. Si el lenguaje contiene ciertos símbolos de función adicionales, es también posible probar una identidad. Si se cuenta con un símbolo de función para denotar la función sustitución (a partir de la cual se puede definir la función diagonalización), es posible reforzar el lema diagonal débil y obtener, para cada predicado φ , una fórmula de la forma $\exists t(\varphi(t))$ (y no meramente equivalente a) $\exists v(\varphi(v))$ que contiene un término que la denota a ella misma. Este resultado se conoce como 'lema diagonal fuerte'. La posibilidad de obtener identidades en lugar de equivalencias será crucial en los capítulos posteriores de esta tesis, particularmente cuando se hable acerca de la referencia en extensiones de \mathcal{L}_{pa} . Sea \mathcal{L}^+_{pa} . Nuevamente, asumo una codificación de \mathcal{L} mediante números naturales.

Teorema 1.8 (Lema diagonal fuerte). Para toda fórmula $\varphi(v; \sim)$ de \mathcal{L} existe un término t de \mathcal{L}^+_{pa} tal que

$$\mathcal{L}^+ \models t = \exists v(\varphi(v; \sim v))$$

Demostración. Sea $\varphi(v; \sim v)$ el resultado de reemplazar v en φ por $v(v=pvq)$. Como $v_1(v_2=v_3)$ representa la función sustitución en \mathcal{L}^+ , \mathcal{L}^+ prueba que el término $\exists v(\varphi(v; \sim v))$ es equivalente en \mathcal{L}^+ a $\exists v(\varphi(v; \sim v))$.

en adelante, t , es idéntico a $p'(p'(v(v=pvq); \sim v)q(p'(v(v=pvq); \sim v)q=pvq);$

$$\sim v)q. \text{ Luego, } q^+ \cdot t = p'(t; \sim v)q$$

□

Si la autorreferencia pareciera una característica dudosa de los lenguajes naturales, Gödel le devolvió toda su legitimidad al mostrar que es posible en el discurso matemático puro. No solo la autorreferencia directa es posible en el lenguaje de la aritmética sino también ciertas formas de referencia indirecta, v.g., ciclos de oraciones, cada una de las cuales refiere a la que le sigue en un ciclo. La existencia de ciclos es probable en q^+ mediante lemas diagonales alternativos para cualesquiera fórmulas y cualquier longitud. Sea L como en el teorema anterior y $\sim u$ una abreviatura de $u_1; \dots; u_m$.

Proposición 1.9 (Ciclos de longitud 2). Para cualesquiera fórmulas $(v; \sim v)$; $(u; \sim u) \in L$ existen términos t y s de L_{pa^+} tales que

$$q^+ \cdot t = p'(s; \sim v)q$$

$$q^+ \cdot s = p(t; \sim u)q$$

Demostración

. Sea $d(v) := v(v; puq)$. Entonces,

-

$$q^+ \cdot d \left(\underbrace{p \left(\underbrace{p'(d(u); \sim v)q(u=puq)}_s; \sim u \right) q}_t \right) =$$

$$p \left(\underbrace{p'(d(u); \sim v)q(\text{num}(p \left(\underbrace{p'(d(u); \sim v)q(u=puq)}_t; \sim u)q)=puq)}_t; \sim u \right) q$$

Luego,

$$q^+ \cdot s = p(t; \sim u)q$$

y

$$q^+ \cdot t = p'(d(u); \sim v)q(\text{num}(p \left(\underbrace{p'(d(u); \sim v)q(u=puq)}_t; \sim u)q)=puq)$$

$$= p'(d \left(\underbrace{p \left(\underbrace{p'(d(u); \sim v)q(u=puq)}_t; \sim u)q \right)}_s; \sim v)q$$

$$= p'(s; \sim v)q$$

□

La proposición 1.9 establece la existencia de ciclos de longitud 2 para cualesquiera predicados. En general es posible probar la existencia de ciclos de longitud n para

cualquier $n \geq 2$ y formulas $\varphi_1; \dots; \varphi_n$ de L con al menos una variable libre. Si para un ciclo de longitud n para las formulas $\varphi_1; \dots; \varphi_n$ se diagonaliza sobre (v) [v.g., sobre $(p(d(u); \sim v)q(u=puq); \sim u)$ en el caso de ciclos de longitud 2], para un ciclo de longitud $n + 1$

para las formulas $\varphi_1; \dots; \varphi_{n+1}$ se aplica d_v a $\varphi_{n+1}(p q(v=pvq); \sim v)$ donde v esta libre en φ_{n+1} . Por ejemplo, para formulas $\varphi_1; \dots; \varphi_n$ con una unica variable libre v , comienza

'desarrollandose'

$$\frac{d(p'_n(p'_n \varphi_1(\dots p'_1 (d_v)q(v=pvq) \dots)q(v=pvq))q)}{\{z_{t_n}\}}$$

Proposicion 1.10 (Ciclos de longitud n). Para cualesquiera formulas $\varphi_1(v); \dots; \varphi_n(v)$ de L existen terminos $t_1; \dots; t_n$ de L^+_{pa} tales que

$$\begin{aligned} q^+ \cdot t_1 &= p'_1(t_2)q \\ q^+ \cdot t_2 &= p'_2(t_3)q \\ &\vdots \\ q^+ \cdot t_{n-1} &= p'_{n-1}(t_n)q \\ q^+ \cdot t_n &= p'_n(t_1)q \end{aligned}$$

Este resultado puede generalizarse a formulas con cualquier numero de variables mayor o igual a 1, al igual que la proposicion que sigue.

Proposicion 1.11 (!-cadenas). Si $\varphi(v) \in L$, existe una secuencia de terminos diferentes t_i de L^+_{pa} con $i \geq 2$ tales que

$$q^+ \cdot t_i = p'(t_{i+1})q$$

Demostracion. Sea $g(v) := v (\text{num}(v (v=pvq))=pvq)$. Aplicada al codigo de un termino que denota (el codigo de) una formula $\varphi \in L$ con una variable libre v , g devuelve (el codigo de) $(p'(p'q)q)$. Luego, razonando en q^+ ,

$$\frac{g(\text{num}(p'(g(v))q))}{\{z_{t_0}\}} = \frac{p'(g(\text{num}(p'(g(\text{num}(p'(g(v))q)))q))q)}{\{z_{t_1}\}}$$

y, por tanto,

$$q^+ \cdot t_0 = p'(t_1)q$$

Ademas,

$$\begin{aligned}
 t_1 &= g(\text{num}(p'(g(\text{num}(p'(g(v))q))))q)) \\
 &= p'(g(\text{num}(p'(g(\text{num}(p'(g(\text{num}(p'(g(\text{num}(p'(g(v))q))))g))))q))))q))q)) \\
 &= \left(\begin{array}{c} t_2 \\ p \quad q \end{array} \right)_{t_2}
 \end{aligned}$$

y as siguiendo. Cada termino en la secuencia es el resultado de aplicar $g(v)$ a alguna formula. Como la funcion representada por $g(v)$ es monotonamente creciente, para cualesquiera $i; j \geq 2$!, si $i < j$, el valor de t_i resulta de aplicar $g(v)$ a una formula mas corta que la que es el valor de t_j y bajo codificaciones monotonas las formulas mas cortas que estan formadas a partir de los mismos simbolos que otras tienen codigos mas pequenos, si $i \neq j$, $t_i \neq t_j$. \square

No solo los ciclos de cualquier longitud son posibles en q^+ sino tambien las cadenas infinitamente descendientes. Naturalmente, tambien son posibles las cadenas infinitas hacia los dos lados, que pueden obtenerse a partir de las cadenas infinitamente descendientes, simplemente aplicando $'(v)$ al primer termino t_0 , obteniendo as un termino $t_1 = p'(t_0)q$, y luego aplicando $'(v)$ a t_1 , obteniendo un termino t_2 , etc.

1.2.7. Resultados de incompletitud de G•odel

G•odel empleo un lema de diagonalizacion analogo a los presentados en el apartado [1.2.6](#) para probar sus famosos teoremas de incompletitud. La prueba del primer resultado es particularmente iluminadora para los propositos de esta tesis. Sea th q una teoria clasica formulada en un lenguaje de primer orden L L_{pa} .

Teorema 1.12 (G•odel. Primer resultado de incompletitud). Si th es consistente, existe una oracion $th \in L$ tal que th no prueba ni th ni $\neg th$.

Demostracion. Aplicando el lema diagonal debil a $\neg \text{Bew}_{th}(x)$, se obtiene una oracion th tal que

$$th \leftrightarrow th \leftrightarrow \neg \text{Bew}_{th}(p \ thq) \tag{1.4}$$

Si $th \leftrightarrow th$, como Bew_{th} debilmente representa la teoremicidad de th en th , $th \leftrightarrow \text{Bew}_{th}(p \ thq)$, esto es, th sera inconsistente en virtud de [\(1.4\)](#). Si, en cambio, $th \leftrightarrow \neg th$, se tendra, por un lado, que $th \leftrightarrow \text{Bew}_{th}(p \ thq)$ por [\(1.4\)](#) y, por otro, por la representacion debil, que $th \leftrightarrow \neg \text{Bew}_{th}(p \ thq)$, y th sera inconsistente tambien. \square

⁹*Este resultado supone que Bew_{th} satisfaga las condiciones de teoremicidad de $L \cdot ob$.*

La oración th , conocida como 'oración de Gödel' de th , es frecuentemente descripta como una oración que dice de sí misma que no es demostrable en th , esto es, como una oración autorreferente. En virtud de lo que sigue, dire que th es una oración débil de Gödel de th .

Si se cuenta con suficientes símbolos de función como para probar el lema diagonal fuerte, es posible, no solo una oración equivalente a otra que dice de la primera que no es demostrable, sino una oración idéntica a 'otra' que dice de ella que no es demostrable. Sea ahora $\text{th} \vdash q$ formulada en una extensión L de L_{pa}^+ . Aplicando el lema de diagonalización fuerte al predicado $\text{Bew}_{\text{th}}(x)$, se obtiene un término g tal que

$$\text{th} \vdash g = \text{p:Bew}_{\text{th}}(g)q$$

para el cual es también posible probar que ni $\text{Bew}_{\text{th}}(g)$ ni $\neg \text{Bew}_{\text{th}}(g)$ son teoremas de th . $\neg \text{Bew}_{\text{th}}(g)$ también es una oración de Gödel de th , una oración fuerte de Gödel, y de ella suele decirse con más seguridad que es autorreferente.

Teorema 1.13 (Gödel. Segundo resultado de incompletitud). Si $\text{th} \vdash q$ es una teoría consistente, no es capaz de probar su propia consistencia, esto es, $\text{th} \vdash \neg \text{Bew}_{\text{th}}(t)$ para ningún término t del lenguaje de la teoría.

Corolario 1.14 (Löb). Si $\text{th} \vdash q$ es una teoría formulada en L_{pa} , para cualquier oración $\varphi \in L$, si $\text{th} \vdash \text{Bew}_{\text{th}}(\varphi) \rightarrow \varphi$, entonces $\text{th} \vdash \varphi$.

1.3. La teoría de conjuntos Zermelo-Fraenkel

L_{zfc} es el lenguaje de la teoría de conjuntos de Zermelo-Fraenkel de primer orden con el axioma de elección (zfc).¹⁰ Su vocabulario no lógico está dado por un símbolo de relación diádico primitivo \in para la pertenencia entre conjuntos y un término \emptyset para el conjunto vacío, de nido en L_{zfc} en función de \in .

pa y, a fortiori, q son relativamente interpretables en L_{zfc} . Esto significa que L_{zfc} prueba la existencia de objetos isomórficos a los números naturales. Si bien existen diversas traducciones posibles que dan cuenta de la interpretabilidad relativa, la más usual es la que 'identifica' cada número natural con un cardinal de Von Neumann (estos son $\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset, \{\emptyset\}\}\}, \dots$) (cf. Tarski, Andrzej Mostowski & Raphael M. Robinson (1953) y Enderton (1977, cap. 4)). \aleph es el conjunto de cardinales de Von Neumann que suele identificarse directamente con el conjunto de números naturales 'real', que noto \mathbb{N} .

¹⁰Herbert B. Enderton (1977), Yiannis Moschovakis (2006) y Azriel Levy (1979) contienen todo lo que cito sin prueba acerca de L_{zfc} .

En consecuencia, L_{zfc} también es capaz de definir y zfc de representar relaciones recursivas y funciones p.r. de alguna manera. Además, L_{zfc} puede codificar cualquier lenguaje de primer orden L dado, por ejemplo, a través de los cardinales de von Neumann y definir las nociones sintácticas [tanto conjuntos como funciones] de L , que zfc luego representará mediante fórmulas que notará igual que para el caso de q y pa . Sea L_{zfc}^+ el resultado de extender L_{zfc} con los símbolos de función de pa^+ y zfc^+ el resultado de incorporar a zfc axiomas para definir estos símbolos.

Sea L_{zfc}^{+u} el resultado de expandir L_{zfc}^+ con un nuevo símbolo monádico de predicado C y zfc_u^+ una teoría idéntica a zfc^+ excepto que los cuantificadores que ocurren en los axiomas y las definiciones están relativizados a C . Por ejemplo, relativizar los cuantificadores del axioma de extensionalidad

$$\exists x; \forall y(\exists z(z \in x \leftrightarrow z \in y) \rightarrow x = y)$$

al predicado C da por resultado

$$\exists x; \forall y(Cx \wedge Cy \rightarrow (\exists z(z \in x \leftrightarrow z \in y) \rightarrow x = y))$$

A diferencia de zfc^+ , esta teoría permite la incorporación de urelementos al universo de discurso, i.e., objetos que no son conjuntos. Por ejemplo, es posible extender el lenguaje y la teoría para hablar de números naturales y conjuntos de números naturales propiamente dichos. Sea L_{zfc}^{+n} el resultado de extender L_{zfc}^{+u} con dos nuevas constantes de individuo 0 y N y tres símbolos de función, S , $+$ y \cdot , y sea zfc_n^+ el resultado de formular zfc_u^+ sobre este lenguaje e incorporar los siguientes axiomas:

- $\exists x(x \in N \rightarrow Cx) \wedge CN$
- $0 \in N \wedge \exists x \in N (S(x) \neq 0 \wedge (\exists y \in N S(y) \neq x \rightarrow x = 0))$
- $\exists x \in N \exists y \in N S(x) = y$
- $\exists z(\exists x(x \in z \rightarrow Cx) \wedge Cz \wedge 0 \in z \wedge \exists x \in z(S(x) \neq 0 \wedge (\exists y \in z S(y) \neq x \rightarrow x = 0)) \wedge \exists x \in z \exists y \in z S(x) = y \rightarrow N z)$

junto con definiciones recursivas de $+$ y \cdot en términos de 0 y S , como es usual.

El primer axioma indica que los elementos de N están extendidos de satisfacer los axiomas de zfc^+ , mientras que N mismo no. El segundo establece que 0 es un miembro de N , que 0 no es una imagen de S y que es el único con esta propiedad en N . De acuerdo con el tercer axioma de zfc_n^+ , todos los miembros de N tienen una imagen en

N a través de S. Por último, el cuarto axioma afirma que N es el conjunto más pequeño

que satisface los tres anteriores. De este modo, es posible hablar del conjunto de números naturales \mathbb{N} (no \aleph , el de los cardinales de von Neumann) en el marco de la teoría de conjuntos de Zermelo-Fraenkel.

1.4. Notación para ordinales

Los números ordinales, con los cuales asumo cierta familiaridad, son entidades conjuntistas sobre las cuales se pueden realizar ciertas operaciones algebraicas (suma, exponenciación, etc) (cf. [Wolfram Pohlers \(2009, cap. 3\)](#)). Los cardinales de von Neumann son los ordinales finitos y \aleph_0 es el primer ordinal infinito. Sea On la clase de todos los ordinales cuya existencia es probada en zfc^+ . Como metavariables sobre ordinales emplear las letras griegas $\alpha, \beta, \gamma, \dots$, a veces con subíndices.

Ciertos segmentos iniciales contables X de On son codificables mediante números naturales, esto es, es posible asignar recursivamente a cada ordinal en X un número natural. Estos sistemas de códigos se conocen como notaciones para ordinales. El numeral de cada uno de estos números sirve luego como un nombre para el ordinal en L_{ω_1} : el lenguaje puede hablar de este modo también acerca de ordinales.

Teorema 1.15 (Cantor. Forma normal). Para todo $\alpha \neq 0$, existen únicos ordinales β_1, \dots, β_n tales que $\alpha = \aleph_{\beta_1} + \dots + \aleph_{\beta_n}$.

Codificando el ordinal 0 con un número natural cualquiera, es posible codificar nuevos ordinales $\alpha = \aleph_{\beta_1} + \dots + \aleph_{\beta_n}$ empleando la forma normal de Cantor, mediante el código de la secuencia de los códigos de β_1, \dots, β_n . Para eso es necesario que $\beta_1, \dots, \beta_n < \alpha$. Esto no lleva demasiado lejos: solo para pocos ordinales $\alpha > 0$ existen $\beta_1, \dots, \beta_n < \alpha$ tales que $\alpha = \aleph_{\beta_1} + \dots + \aleph_{\beta_n}$, a saber, solo aquellos que se encuentran por debajo de ω , el límite de los ordinales

$$1; \aleph_1; \aleph_1^{\aleph_1}; \aleph_1^{\aleph_1^{\aleph_1}}; \dots$$

Es posible obtener códigos para segmentos iniciales más largos de ordinales mediante la función de Veblen.

Definición 1.16 (Función de Veblen). La función de Veblen en $On^2 \rightarrow On$ es tal que, si α, β son ordinales,

1. $\text{veb}(\aleph_0; \alpha) = \aleph_{\alpha}$;
2. $\text{veb}(\alpha + 1; \beta) = \text{el } \beta\text{-ésimo ordinal tal que } \text{veb}(\alpha; \gamma) = \beta$;

3. si α es un ordinal límite, $\varphi(\alpha) = \text{el } \alpha\text{-ésimo ordinal tal que } \varphi(\beta) = \alpha \text{ para todo } \beta < \alpha$.

Esta es una definición de la función de Veblen por inducción transnita. Dado un ordinal α , $\varphi(\alpha)$ es el α -ésimo punto fijo de la función monádica $\varphi(\cdot)$. Por ejemplo, $\varphi(0) = 1$, $\varphi(1) = \omega$ y $\varphi(\omega)$ es el primer ordinal tal que $\varphi(\alpha) = \alpha$, esto es, el primer punto fijo de $\varphi(\cdot)$, φ_0 .

Algunos ordinales α son constructibles mediante la función de Veblen, esto es, existen $\beta < \alpha$ tales que $\varphi(\beta) = \alpha$. Luego, si se codifican los ordinales menores a α con la notación anterior, dada por la forma normal de Cantor, la función de Veblen ofrece códigos para nuevos ordinales $\alpha = \varphi(\beta)$ si $\beta < \alpha$. Lamentablemente, no todos los segmentos contables de ordinales pueden obtenerse de este modo tampoco.

Definición 1.17 (Ordinal de Feferman-Schütte). El ordinal θ de Feferman-Schütte es el menor ordinal > 0 tal que, para cualesquiera $\beta < \theta$, $\varphi(\beta) < \theta$.

Todo ordinal menor a θ puede codificarse mediante un número natural único de modo recursivo y, por ende, para cada $\alpha < \theta$ existe en L_{pa} un término que denota el código de α . Asumo una única notación ordinal de aquí en más, en la cual el ordinal 0 es codificado por el número 0. Sean $\text{Ord}(v)$ y $\forall v_1 \forall v_2$ las fórmulas de L_{pa} que denotan, respectivamente, el conjunto de códigos de ordinales menores a θ y la relación de menor entre estos ordinales bajo esta codificación. También es posible desarrollar sistemas de notación más abarcativos, para segmentos de ordinales más extensos que θ , pero no es necesario para los propósitos de esta investigación.

Sea δ para cualquier variable de ordinal un modo abreviado de escribir $\delta \forall v (\text{Ord}(v) \rightarrow \delta)$ en L_{pa} y sus extensiones. pa^+ solo es capaz de probar una versión acotada del principio de inducción transnita o, equivalentemente, buen orden,

$$\delta \rightarrow (\delta \rightarrow \delta) \rightarrow \delta$$

para cada $\delta \in L_{\text{pa}}(L_{\text{pa}}^+)$, para ordinales menores a θ . Esto es,

$$\text{pa}^+ \vdash \delta \rightarrow (\delta \rightarrow \delta) \rightarrow \delta \tag{IT}$$

si $\alpha < \theta$. El menor ordinal a partir del cual una teoría no es capaz de probar instancias del principio de inducción transnita se conoce como su 'ordinal prueba-teórico'. El ordinal prueba-teórico de pa^+ es θ , si bien hay codificaciones al menos hasta θ . Como consecuencia, en pa^+ no siempre es posible probar ciertas propiedades básicas acerca de los ordinales más allá de θ . V.g., los predicados $\text{Ord}(v)$ y $\forall v_1 \forall v_2$ de L_{pa} solo representan los conjuntos que denotan para ordinales menores a θ .

2

Tarski o la verdad tipeada

"If I had to name the theory, I would call it the 'Look No Semantics! Conception of Truth'."

{ Hilary Putnam, 'A Comparison of Something with Something Else'

La noción de verdad, junto con otros conceptos llamados 'semánticos', había caído en considerable descrédito hacia comienzos del siglo XX. Matemáticos y filósofos analíticos [el círculo de Viena en primer lugar, por sus consideraciones empiristas y anti-metafísicas] preferían evitar el uso de la palabra 'verdad', debido a las dificultades que enfrentaban los múltiples intentos de definir la y a las conocidas antinomias a las cuales da lugar.

No obstante, la noción de verdad era empleada informalmente como hoy, no solo en el discurso cotidiano, sino también en el filosófico y en el discurso de la lógica matemática, área de especialización de Tarski, para hablar de modelos de teorías, consistencia, satisfacibilidad y categoricidad, entre otros, ya desde el siglo XIX. Su uso parecía ser esencial y constituía por tanto una amenaza para la consistencia y precisión de cualquier disciplina que la involucrara.

Tarski fue el primero en buscar dar nociones formales de verdad que superasen estas dificultades, que dieran lugar a conceptos claros y libres de inconsistencias. Impuso primero dos condiciones sobre toda definición de verdad, una de corrección formal y otra de adecuación material, que expongo en la sección [2.1](#) de este capítulo. La sección [2.2](#) da una introducción general al método tarskiano y presenta los dos tipos de definición de verdad que Tarski introdujo: verdad absoluta y verdad relativa. En [2.2.2](#) ilustro el método tarskiano para definir verdad absoluta, esto es, para lenguajes formales interpretados, que introdujo Tarski en ([1933](#)), aplicándolo al lenguaje de la aritmética de primer orden.

En [2.2.3](#) presento la definición tarskiana de verdad para oraciones de un lenguaje formal no interpretado relativa a un modelo, estructura o interpretación cualquiera del lenguaje, que Tarski introdujo en ([1952](#)), y en ([1956](#)) junto a Robert L. Vaught.

Las definiciones tarskianas enfrentan conocidas limitaciones, principalmente la falta de universalidad. En [2.3](#), el último apartado del capítulo, las señalo e indico dos maneras posibles de superar algunas de las dificultades que estas limitaciones conllevan mediante jerarquías, sin apartarme del espíritu tarskiano.

2.1. Criterios para una definición de verdad

El desafío que Tarski se impuso fue legitimar la noción de verdad y otros conceptos semánticos ante las comunidades matemática y filosófica; mostrar que es posible superar no solo las dificultades conceptuales sino lógicas que la verdad trae consigo.

For a long time the semantical concepts have had an evil reputation among specialists in the study of language. They have resisted all attempts to define their meaning exactly, and the properties of these concepts, apparently so clear in their content, have led to paradoxes and antinomies. (Tarski, [1933](#), p. 252)

Este desafío toma la forma de dos condiciones que, desde el punto de vista de Tarski, toda definición de verdad debe satisfacer. Las presento a continuación, en los dos apartados que siguen, respectivamente. La primera concierne a la forma de la definición. La segunda refiere a la corrección de la definición, es un requisito de adecuación material, que responde a ciertos rasgos asignados usualmente al concepto de verdad y, de acuerdo con Tarski, hace de la definición una definición semántica, que respeta la intuición correspondentista.

2.1.1. Corrección formal

De acuerdo con [Tarski \(1933, 1944\)](#), una definición formalmente correcta de verdad debe tomar la siguiente forma:

Para toda oración x , x es verdadera si y solo si $P(x)$: (EXP)

donde P es un predicado posiblemente complejo expresado en terminos de nociones mas sencillas y bien entendidas, para las cuales se cuenta con formulaciones precisas (esto

es, ni la expresión 'es verdadera', ni otras expresiones sinónimas o igual o más oscuras ocurren en P), o bien ser equivalente a una definición de la forma de (EXP), y esta equi-valencia debe ser demostrable empleando principios que no contengan la expresión 'es verdadera', ni otras expresiones sinónimas o igual o más oscuras. En palabras más sencillas, lo que Tarski pide es una definición explícita del predicado de verdad en términos no semánticos.

For, if this postulate is satisfied, the definition of truth [. . .] will explain the meaning of the term being defined in terms whose meaning appears to be completely clear and unequivocal. And, moreover, we have then a kind of guarantee that the use of semantic concepts will not involve us in any contradictions. (Tarski, 1944, p. 351)

Una definición explícita del predicado veritativo permitiría eliminarlo, esto es, reemplazarlo por una fórmula dada en términos claros, que no engendren contradicciones. De este modo, se obtendría una prueba automática de la consistencia del concepto de verdad. A diferencia de los abordajes clásicos de la verdad, como la corriente correspondentista, Tarski no busca que (EXP) sea la intensión del predicado de verdad, el significado de este concepto, sino meramente la extensión.

2.1.2. Adecuación material

Tarski basó su investigación en torno a la verdad sobre sus intuiciones correspondentistas o, como él las llama, semánticas.

Amongst the manifold efforts which the construction of a correct definition of truth for the sentences of colloquial language has called forth, perhaps the most natural is the search for a semantical definition. By this I mean a definition which we can express in the following words:

a true sentence is one which says that the state of affairs is so and so, and the state of affairs indeed is so and so.

From the point of view of formal correctness, clarity, and freedom from ambiguity of the expressions occurring in it, the above formulation obviously leaves much to be desired. Nevertheless its intuitive meaning and general intention seem to be quite clear and intelligible. To make this intention more definite, and to give it a correct form, is precisely the task of a semantical definition. (Tarski, 1933, p. 155, sus *italicas*)

Desde este punto de vista, lo semántico es aquello que, de alguna manera, relaciona significantes con sus significados, el lenguaje con el mundo.¹

La condición de adecuación material que Tarski impone sobre toda definición de verdad está para garantizar que el predicado adquiera la extensión adecuada, que respete la intención correspondentista; que los objetos que resulten verdaderos de acuerdo con la definición sean precisamente aquellos que intuitivamente lo son. Comienza por considerar los siguientes dos principios:

(1) Todos los enunciados de la forma

'A' es verdadera si y solo si A: (Esquema T)

deben poder derivarse de una definición de verdad, donde A es una oración y 'A' es un nombre transparente para A, i.e., un término que indica, por su estructura, que denota A, v.g., dado por comillas o por una descripción estructural (la oración cuya primera palabra es . . .) de A.

(2) Solo las oraciones pueden ser verdaderas.

La condición (2) indica que, entre los diversos candidatos posibles a portadores de verdad, i.e., oraciones declarativas, proposiciones, creencias, enunciados, pensamientos, juicios, aserciones, etc., Tarski opta por las oraciones. Estas no deben ser entendidas como tal o cual soporte físico, ya sea gráfico o sonoro, porque, como señala [Tarski \(1933\)](#), pp. 166-167),

[...] we are not interested here in "formal" languages and sciences in one special sense of the word "formal", namely sciences to the signs and expressions of which no material sense is attached. For such sciences the problem here discussed [el de definir verdad] has no relevance, it is not even meaningful. We shall always ascribe quite concrete and, for us, intelligible meanings to the signs which occur in the languages we shall consider.

Tampoco son oraciones de cualquier tipo sino, como [Willard Van Orman Quine \(1992\)](#) las denomina, oraciones eternas, esto es, libres de indexicos y otras palabras sensibles al contexto, ya que una misma oración que contenga indexicos puede ser capaz de expresar en algunos casos algo verdadero y en otros algo falso (cf. [Donald Davidson \(1969\)](#)), como la oración "Mi nombre es Lavinia", que dicha por mí es verdadera, mientras que dicha por ustedes es seguramente falsa.

¹Tarski señala también otras nociones como típicamente semánticas: las de designación, satisfacción y de intención.

Las razones que da Tarski para preferir oraciones por sobre las proposiciones, la opción rival más saliente, son escasas y pueden resumirse en la siguiente cita:

[...] as regards the term "proposition", its meaning is notoriously a subject of lengthy disputations by various philosophers and logicians, and it seems never to have been made quite clear and unambiguous. For several reasons it appears most convenient to apply the term "true" to sentences, and we shall follow this course. (Tarski, 1944, p. 342)

En efecto, se cuenta con definiciones precisas de oración de prácticamente todos los lenguajes relevantes a esta investigación, mientras que no puede decirse lo mismo del concepto de proposición. Además, al elegir oraciones como las portadoras de verdad, las herramientas formales necesarias son mínimas, a saber, términos que denoten las oraciones en cuestión. Para estos últimos hay a disposición un mecanismo generador, i.e., las comillas, aunque puede haber términos de otras clases.

Por estas razones opto yo también por tomar oraciones como portadoras de verdad en toda esta investigación. No obstante, si se prefieren predicaciones de verdad como "A es verdadera", donde A es una oración, como diciendo de la proposición expresada por A que es verdadera en lugar de A misma, no encuentro razones para oponerme ni hay nada en esta tesis que lo impida.

Con respecto a la condición (1), el [Esquema T](#) implica, por ejemplo, que la oración "Virgilio fue un poeta" es verdadera si y solo si Virgilio fue un poeta, lo cual es bien sencillo y razonable. Tarski ve en el [Esquema T](#) una suerte de extensionalización de la teoría de la verdad por correspondencia; una realización de esta teoría sin el aura metafísica que suele acompañarla.

Pero Tarski advierte rápidamente que este esquema trae graves problemas. Considerese la siguiente oración 'del mentiroso':

L es falsa. (L)

L dice de sí misma que es falsa. Dado que 'L' es un nombre para L, de acuerdo con (1) una buena definición de verdad debería implicar la siguiente instancia del [Esquema T](#):

L es verdadera si y solo si L es falsa

esto es, una contradicción. El [Esquema T](#) es precisamente lo que habilita a derivar de la suposición de que L es verdadera que es falsa y viceversa. En esto consiste la paradoja del mentiroso. (cf. [Jose Martínez \(2014\)](#)).

Las causas por las cuales un principio tan intuitivo como el [Esquema T](#) engendra paradojas son, como indica Tarski, dos. Por un lado, el lenguaje en y para el cual se pretende dar una definición de verdad [el castellano] coinciden. Los lenguajes naturales son semanticamente cerrados, contienen su propio predicado de verdad. Se ha querido dar una definición del predicado de verdad de y para un mismo lenguaje y, en consecuencia, algunas de las oraciones por las cuales se puede reemplazar A en el [Esquema T](#) contienen ellas mismas la expresión 'es verdadera' (o expresiones sinónimas), así como nombres para estas mismas oraciones.

Por otro lado, la lógica subyacente que se supone ligada al lenguaje en el cual se trabaja es la clásica. Cualquier lenguaje que satisfaga estas dos condiciones [i.e., que sea semanticamente cerrado y opere con lógica clásica] será trivial. Este resultado se conoce informalmente como 'teorema de la indecidibilidad de Tarski'. El teorema de indecidibilidad de Tarski propiamente dicho (teorema [2.9](#)) tiene una formulación precisa y dependiente de ciertas características del lenguaje bajo consideración. En el apartado [2.3.1](#) de esta sección enuncio y pruebo el teorema para una amplia familia de lenguajes formales.

Dado que Tarski no considera seriamente la posibilidad de adoptar lógicas alternativas a la clásica, sostiene que todo lenguaje que contenga su propio predicado de verdad es o bien inconsistente o bien confuso, que no es posible que un lenguaje formal no trivial contenga un predicado aplicable a sus propias oraciones y satisfaga el [Esquema T](#). Luego concluye que "[a]ccordingly, we decide not to use any language which is semantically closed in the sense given" (Tarski, 1944, p. 349).

La paradoja del mentiroso ha sido presentada en castellano y una estar a tentada a inferir entonces que o bien el castellano es trivial o bien no puede contener un predicado que satisfaga el [Esquema T](#) y, a fortiori, tampoco un predicado de verdad. Sin embargo, Tarski se guarda de extraer estas conclusiones. Desde su punto de vista, que comparto, el lenguaje coloquial no tiene una estructura especificada, no se sabe con exactitud cuáles son sus oraciones declarativas pues está sujeto a cambio continuo y, por ende, no queda claro que sea trivial. Tampoco es claro cuál es la lógica que lo subyace, si es que tal cosa existe. La 'demostración' del teorema de Tarski en castellano es puramente ilustrativa y servirá de modelo para las pruebas formales.

We may at best only risk the guess that a language whose structure has been exactly specified and which resembles our everyday language as closely as possible would be inconsistent. (Tarski, 1944, p. 349)

Por esta razón, y lo seguiré en esto, Tarski abandona el lenguaje natural y opta por trabajar en lenguajes formales.

The problem of the definition of truth obtains a precise meaning and can be solved in a rigorous way only for those languages whose structure has been exactly specified. For other languages [thus, for all natural, "spoken" languages] the meaning of the problem is more or less vague, and its solution can have only an approximate character. Roughly speaking, the approximation consists in replacing a natural language (or a portion of it in which we are interested) by one whose structure is exactly specified, and which diverges from the given language "as little as possible". (Tarski, 1944, p. 347, sus *italicas*)

A la luz de las hipótesis de su teorema, Tarski deja de lado los lenguajes semánticamente cerrados y establece una distinción entre el lenguaje de cuyas oraciones se quiere predicar verdad [el lenguaje objeto] y el lenguaje desde el cual se hacen estas as-cripciones de verdad [el metalenguaje]. Así evita la paradoja del mentiroso, que no es una fórmula bien formada en los lenguajes tarskianos. Otros caminos a seguir para no caer en contradicciones consisten en dar únicamente teorías incompletas o parciales de verdad, abandonar la lógica clásica o restringir la capacidad expresiva del lenguaje de modo tal que o bien no contenga nombres para sus propias expresiones o bien no sea capaz de formular oraciones como L . Lo último es quizás una movida espuria, en tanto se está a recortando la expresividad de un lenguaje por un lado para incrementarla (agregando un predicado veritativo) por otro. Las primeras dos opciones, en cambio, son muy comunes en la literatura post tarskiana, como se verá más adelante.

Luego, para definir verdad para las oraciones de un lenguaje objeto L , Tarski propone operar desde un metalenguaje L^0 que contenga

- una copia o traducción de L , esto es, que todo lo que pueda decirse en L pueda también decirse en L^0 ,
- nombres para las oraciones de L y capacidad de hablar de la sintaxis de este lenguaje,
- un predicado monádico para expresar verdad en L ,

y sea esencialmente más rico que L . El metalenguaje L^0 deberá tener, por tanto, al menos los recursos de un lenguaje de predicados de primer orden y, adicionalmente, ser de alguna manera más poderoso que el lenguaje objeto. Tarski (1944, p. 351) mismo admite que el calificativo "esencialmente más rico" no es fácil de definir en términos generales, y todos los intentos formales hasta el momento han fracasado. Nótese que,

evidentemente, Tarski no esta pensando en lenguajes aislados sino en pares de lenguajes

y calculos formulados en ellos; identifica de cierto modo un lenguaje con la teoría para cuya expresión fue diseñado.

La distinción entre lenguaje y metalenguaje da lugar a nociones tipeadas de verdad, en analogía con la separación en niveles de la teoría de tipos de [Bertrand Russell \(1903\)](#). Esta distinción requiere a su vez la reformulación de las condiciones (1) y (2) de adecuación material de toda definición de verdad. El resultado es lo que [Tarski \(1933\)](#) llama 'Convención T'.

Definición 2.1 (Convención T). Una definición de T en L^0 es una definición materialmente adecuada de verdad para L solo si implica (traducciones adecuadas a L^0 de):

1. todas las instancias del esquema

'A' es verdadera si y solo si (A) (Esquema T Tipeado)

donde A es una oración del L, 'A' es un nombre transparente para A y (A) es una traducción de A a L^0 (si $L \subseteq L^0$, $(A) = A$), y

2. solo las oraciones de L son verdaderas.

A diferencia del [Esquema T](#), el [Esquema T Tipeado](#) permite que la ocurrencia entrecomillada de cada oración A a la izquierda de los bicondicionales pertenezca a un lenguaje diferente que su ocurrencia sin entrecomillar, a la derecha. En otras palabras, habilita una distinción tajante entre lenguaje objeto y metalenguaje. En el apartado que sigue presento los dos tipos de definición tarskiana de verdad y muestro como, en ambos casos, se satisface la Convención T.

2.2. Definiciones tarskianas de verdad

A fines de los años 1920' Tarski inició la búsqueda de una definición formal de verdad. En [\(1933\)](#) ofreció tres definiciones de verdad simpliciter (a secas, absoluta) con estas características para oraciones de tres lenguajes formales interpretados diferentes y mostró que las nociones resultantes satisfacen los criterios propuestos en la sección anterior. Como señala [Solomon Feferman \(2008a\)](#), este trabajo estuvo dirigido a la comunidad filosófica, mientras que en [\(1952\)](#), y en [\(1956\)](#) junto a Robert L. Vaught, Tarski presentó una definición precisa de verdad para oraciones de un lenguaje formal en un modelo, de la verdad relativa a una estructura o interpretación, orientada más bien hacia la comunidad lógico-matemática. Tras una breve introducción en [2.2.1](#) al método general que Tarski

emplea en sus definiciones, los apartados [2.2.2](#) y [2.2.3](#) que siguen están dedicados a la presentación de cada uno de estos tipos de definición, respectivamente.

El primer grupo de definiciones inusualmente radicalmente en (prácticamente) toda obra de filosofía analítica posterior acerca de la verdad. Tuvo una gran influencia en el desarrollo del formalismo, así como en la literatura actual sobre teorías de la verdad. El segundo grupo de definiciones dio lugar a lo que hoy se conoce como 'teoría de modelos', una de las ramas centrales de la lógica matemática. Ambas nociones tuvieron un éxito rotundo en la comunidad lógico-filosófica y también lógico-matemática, al menos a largo plazo.

Fue idea de Tarski que sus definiciones estén dadas en términos claros y precisos, que estuvieran más allá de toda duda. Por ende, considero que los definiendo no deben contener términos semánticos que no fueran a su vez reducibles a otros conceptos no semánticos, razón por la cual empleo exclusivamente términos lógicos y matemáticos, dándole consistencia y claridad, y por ende, legitimidad a la noción de verdad. Si bien eligió el lenguaje de la teoría general de clases o teoría simple de tipos como metalenguaje para las dos clases de teorías de la verdad que introdujo, hoy es más usual emplear el lenguaje y la teoría de conjuntos. Las definiciones tarskianas de verdad que presento en esta sección están formuladas en una extensión de L_{zfc}^+ y dadas en una metateoría que extiende zfc^+ (véase el apartado [1.3](#)); son definiciones conjuntistas de verdad. En matemática es usual definir un concepto en términos conjuntistas para establecer su legitimidad.

2.2.1. El método recursivo

Una definición de verdad dada por todas las instancias del [Esquema T Tipeado](#) resultará materialmente adecuada por estipulación. Sin embargo, en ese caso la definición no será formalmente correcta. En palabras de Tarski,

We can only say that every equivalence of the form [[Esquema T Tipeado](#)] obtained by replacing [A] by a particular sentence, and ['A'] by a name of this sentence, may be considered a partial definition of truth, which explains wherein the truth of this one individual sentence consists. The general definition has to be, in a certain sense, a logical conjunction of all these partial definitions. (Tarski, 1944, p. 344)

Si el lenguaje objeto contara con un número finito de oraciones $A_1; \dots; A_n$, las instancias del [Esquema T Tipeado](#) serían finitas, y la conjunción de todas ellas daría

lugar a una definición formalmente correcta de verdad, a saber

Para toda oración x , x es verdadera si y solo si $x = \neg A_1$ y A_1 , o . . . , o $x = \neg A_n$ y A_n ,
(2.1)

[si y el lenguaje objeto fuera parte del metalenguaje; de lo contrario habría que traducir el segundo conyunto en cada disyunción del lado derecho del bicondicional]. Lamentablemente, los lenguajes formales con los cuales se suele trabajar cuentan con un número infinito de oraciones y no admiten conjunciones infinitas, por lo cual no es posible poner en conjunción todas las instancias del [Esquema T Tipeado](#) y dar una definición analoga a (2.1).²

La Convención T parece sugerir otro modo aparentemente rápido y sencillo, alternativo a la conjunción infinita de las instancias del [Esquema T Tipeado](#), de dar una definición de verdad que satisfaga el requisito de corrección formal dado por (EXP): i.e., generalizar sobre el [Esquema T Tipeado](#). Supongase que $L \subseteq L^0$. Luego, se podrá definir verdad para L en L^0 del siguiente modo:

Para toda oración A de L , $\neg A$ es verdadera si y solo si A . (2.2)

Sin embargo, como señala Tarski, existen dificultades sintácticas que ponen en duda la viabilidad de (2.2). Para comenzar, será preciso extender los lenguajes primer orden para contar con cuantificadores sobre posición oracional en el lenguaje, i.e. los cuantificadores de segundo orden sobre variables 0-ádicas (cf. Boolos et al. (2007, cap. 22)). En segundo lugar, el predicado veritativo se aplicará exclusivamente a nombres transparentes de oraciones. Este es superficialmente un mal menor, pues si se sabe que una oración $\neg A$ es verdadera, y que $\neg A = s$, donde s no es un nombre transparente de A , por la ley de Leibniz se sabe también que s es verdadera.

Males mayores aparecen cuando se cuestiona la naturaleza de las comillas (o las descripciones estructurales). Usualmente los términos que se obtienen entrecomillando otras expresiones son considerados sintácticamente simples, es decir, carecen de complejidad interna, son tan complejos como un nombre de cualquier otra índole. No obstante, dado que es preciso cuantificar dentro de las comillas, hay que entenderlas como un símbolo de función $\neg x$, que toma fórmulas de L y las devuelve transformadas en objetos. De lo contrario, no será posible sustituir $\neg A$ por oraciones en (2.2) y, por ende,

$\neg A$ es verdadera si y solo si Virgilio fue un poeta.

²*Aun si admitieran conjunciones infinitas una definición de verdad en esta línea no sería posible, porque en lenguajes infinitarios normales, i.e., donde la definición de fórmula bien formada es recursiva, no existe la conjunción infinita de todas las fórmulas del lenguaje, ya que debería contenerse a ella misma.*

ser a una consecuencia de la definición, lo cual es absurdo. Además, podrá derivarse también esta otra instancia:

'A' es verdadera si y solo si Virgilio no fue un poeta.

esto es, la teoría a ser inconsistente.

Por un lado, las comillas deben comportarse como un operador, que se aplica a fórmulas como lo hace la negación '¬'. Pero en vez de dar lugar a nuevas fórmulas, como los operadores, las comillas dan lugar a objetos, y se comportan en esto como símbolos de función. Las comillas son un dispositivo lingüístico híbrido. Lamentablemente, no existe tal cosa en ninguno de los lenguajes artificiales construidos hasta ahora. Si se piensa la cuantificación sobre oraciones como cuantificación de segundo orden sobre variables 0-ádicas, las comillas tomarán objetos del dominio sobre el cual varían las variables de segundo orden y arrojarán elementos del dominio de primer orden.

Quizás sea posible agregar un dispositivo de este tipo a algún lenguaje formal que sea capaz de cuantificar sobre fórmulas. Pero, como señala Tarski, esto no es tan sencillo. Primero, el operador-función resultante no será extensional sino intensional, i.e., dadas dos oraciones A y B, podrá (y, en muchos casos, deberá) ocurrir que A y B sean (incluso lógicamente) equivalentes y, sin embargo, 'A' ≠ 'B'. Por ejemplo, dos tautologías diferentes.

Segundo, ya no tendrán sentido algunas expresiones como 'x es una variable', porque ya no serán oraciones sino funciones oracionales o proposicionales. Una opción será contar con dos tipos de comillas para evitar este problema: las funcionales y las que arrojan términos cerrados, como es usual.

Pero, tercero y más importantemente, habrá un alto riesgo de caer en contradicciones aun antes de introducir un predicado de verdad. Si el lenguaje tuviera capacidad autorreferencial, podrá formularse la siguiente oración:

Para toda oración A, si 'A'=(2.3), entonces no A. (2.3)

Instanciando el cuantificador universal de esta oración en (2.3), se tendrá que (2.3) implica su propia negación. Pero la negación de (2.3) implica a su vez que hay una oración 'A' tal que 'A'=(2.3) y A, esto es, (2.3). Luego, (2.3) y su negación serán equivalentes, lo cual es un absurdo.

En vista de estas dificultades, Tarski exploró un método alternativo para dar definiciones de verdad formalmente correctas y materialmente adecuadas para lenguajes con un número infinito de oraciones: el método recursivo. El resultado fue lo que el

llamo 'definiciones estructurales' o de definiciones recursivas de verdad. La intuición que está por detrás es la siguiente: la verdad de una oración en un lenguaje formal depende de la verdad de sus partes componentes, es composicional. Luego, se asigna a su lenguaje un conjunto de instancias del [Esquema T Tipeado](#) para oraciones atómicas del lenguaje objeto y luego se agregan cláusulas que permitan inferir la verdad o falsedad de las oraciones más complejas, que se forman a partir de estos componentes atómicos. Por ejemplo, cláusulas que digan que una conjunción es verdadera si y solo si ambos conjuntos lo son.

No obstante, así como se describió este método falla, porque no toda oración compleja está compuesta por oraciones atómicas. En particular, las oraciones universalmente cuantificadas [como 'para todo x , x es poeta'] están sintácticamente compuestas de fórmulas con variables libres [como ' x es un poeta'], de las cuales no tiene sentido predicar verdad o falsedad. Por esta razón Tarski da un rodeo y antes de definir verdad para las oraciones de un lenguaje de primer orden da una definición recursiva de la noción de satisfacción para sus fórmulas por ciertos objetos. Por ejemplo, ' x es un poeta' es satisfecha por Virgilio, y también por Homero, pero no por Augusto; y no es en sí ni verdadera ni falsa. Oraciones universalmente cuantificadas como 'para todo x , x es poeta' serán verdaderas en la medida en que todos los objetos de los cuales se está hablando satisfagan la subfórmula ' x es un poeta'.

De este modo, Tarski define verdad para las oraciones de un lenguaje objeto desde un metalenguaje a partir de la noción de satisfacción, que introduce recursivamente, como se verá en los siguientes dos apartados.

2.2.2. Una definición recursiva de verdad a secas

En 1933 Tarski ilustra su método recursivo para definir satisfacción y luego verdad cumpliendo con los requisitos de corrección formal [dado por ([EXP](#))] y de adecuación material [dado por la Convención T] con tres lenguajes objeto absolutamente interpretados, tres fragmentos del lenguaje de la teoría general de clases o teoría simple de tipos (cf. [Mario Gómez-Torrente \(2004\)](#)). Este es un lenguaje de segundo orden donde las variables de predicado están indexadas u ordenadas. Mientras que las variables de individuo varían sobre individuos, las de segundo orden que llevan el índice 1 varían sobre clases de individuos, las que tienen índice 2 sobre clases de clases de individuos, etc. El lenguaje de la teoría general de clases contiene todas las variables de índice o tipo nito. El fragmento para el cual Tarski da su más famosa definición de verdad es el que contiene únicamente variables de primer y segundo orden, el lenguaje de la teoría básica de clases. Como metateoría emplea la teoría general de clases o simple de tipos (cf. [Thierry Coquand \(2015\)](#)).

Sin embargo, ya en la traducción al alemán de su monografía de 1933 publicada en 1935 Tarski agregó una séptima sección titulada 'Postscript' en la cual señala la conveniencia de recurrir a una metateoría de clases más poderosa, que contenga también órdenes transnitos. Todos estos órdenes pueden darse en teoría de conjuntos (utilizando aritmética transnita) que, como sostienen [Gödel \(1933\)](#) y [Lýstein Linnebo & Agustín Rayo \(2012\)](#), puede verse como una generalización de la teoría simple de tipos. Trabajo por tanto directamente desde el lenguaje de la teoría de conjuntos e ilustro las definiciones tarskianas de 1933 mediante una definición análoga donde el lenguaje objeto es L_{pa} . La posibilidad de codificar las expresiones de cualquier lenguaje de primer orden mediante conjuntos, la capacidad del lenguaje de la teoría de conjuntos para hablar de estos códigos y de verificar ciertas propiedades recursivas de estos últimos, y la de la teoría de conjuntos para representar buena parte de estas propiedades (vease el apartado [1.3](#)) permite llevar a cabo la definición.

El metalenguaje extiende L_{zfc}^+ con un predicado monádico $T_{L_{pa}}$ para expresar verdad en L_{pa} (vease el apartado [1.2.1](#)) y un predicado diádico $Sat_{L_{pa}}$ para expresar satisfacción. La metateoría en la cual se define verdad para oraciones de L_{pa} es una extensión de zfc_n^+ formulada en esta expansión de L_{zfc}^+ (vease el apartado [1.3](#)).

Intuitivamente, los objetos que satisfacen fórmulas de L_{pa} con una variable libre son números, los que satisfacen fórmulas con dos variables libres son pares ordenados de números, etc. Para evitar estas consideraciones relativas a cada fórmula Tarski prefiere trabajar con asignaciones de números, funciones o secuencias que asignan números no solo a las variables que están libres en una fórmula dada, sino a todas las variables a la vez y, en general, a todos los términos del lenguaje, de modo tal que una misma asignación podrá satisfacer fórmulas con distinto número de variables libres.³ Trabajo sobre una codificación de expresiones de L_{pa} con conjuntos de la metateoría (vease el apartado [1.3](#)).

Definición 2.2 (Asignaciones para L_{pa}). Una asignación g para L_{pa} es una función cuyo dominio es el conjunto de (los códigos de) términos de L_{pa} y su codominio N , tal que

1. $g(p0q) = 0$;
2. $g(pSvq) = Sg(pvq)$;
3. $g(pv + uq) = g(pvq) + g(puq)$; y
4. $g(pv \cup uq) = g(pvq) \cup g(puq)$.

³Dado que los lenguajes que [Tarski \(1933\)](#) considera carecen de términos cerrados, sus asignaciones, que él denomina 'secuencias', solo asignan objetos a las

variables de individuo. Aca, como se vera a continuacion, es preciso extenderlas a todos los terminos del lenguaje.

Una asignacion para L_{pa} asigna a cada termino cerrado el numero que este denota, a cada variable un numero y a cada termino complejo con variables libres el numero que denota a el termino si las variables libres fueran reemplazadas por los numerales de los numeros que la asignacion les asigna.

Dado que el concepto de funcion y el conjunto de terminos de L_{pa} son definibles en zfc^+ , el metalenguaje contiene una formula $As_{L_{pa}}(v)$ que define el conjunto de funciones que son asignaciones para L_{pa} . Luego, se define la noción de satisfaccion de formulas de L_{pa} por una asignacion en la metateoria. Sea g' una abreviatura de $\exists v(As_{L_{pa}}(v) \wedge g = v)$.

Definición 2.3 (Satisfaccion de formulas de L_{pa} por una asignacion).

$$\begin{aligned}
 \text{Sat}_{L_{pa}}(g; x) &\stackrel{\text{def}}{=} \text{Form}_{L_{pa}}(x) \wedge \\
 &(\exists y; z(x = (y = z) \wedge g(y) = g(z)) \rightarrow \\
 &\exists y(x = \neg y \wedge \neg \text{Sat}_{L_{pa}}(g; y)) \rightarrow \\
 &\exists y; z(x = (y \rightarrow z) \wedge (\text{Sat}_{L_{pa}}(g; y) \rightarrow \text{Sat}_{L_{pa}}(g; z))) \rightarrow \\
 &\exists y; z(x = \exists yz \wedge \exists g^0(\exists w(\forall v(w = y \rightarrow g^0(w) = g(w)) \\
 &\wedge \text{Sat}_{L_{pa}}(g^0; z))))
 \end{aligned}$$

Una formula x de L_{pa} es satisfecha por una asignacion g si x es una formula atomica de la forma $t = s$ y t denota lo mismo que s de acuerdo con g ,⁴ o x es la negacion de una formula que no es satisfecha por g , o el condicional entre dos formulas tales que g no satisface el antecedente o s el consecuente, o una formula de la forma $\exists v'$ y todas las asignaciones que difieren de g a lo sumo en lo que asignan a v satisfacen x .

Por ejemplo, si $g(px_1q) = 0$ y $g(pvq) = 1$ para toda variable v distinta de x_1 , $\text{Sat}_{L_{pa}}(g; \exists x_2 \exists x_2 0 = x_1)$, porque toda asignacion g^0 que difiere de g a lo sumo en lo que asigna a x_2 es tal que $\text{Sat}_{L_{pa}}(g^0; \exists x_2 0 = x_1)$, ya que, para todo $n \in \mathbb{N}$, $n = g^0(x_1) = 0$, porque g^0 no puede diferir de g en cuanto a x_1 sino solo con respecto a x_2 . $\exists x_2 \exists x_2 0 = x_1$ es satisfecha exclusivamente por asignaciones que asignen 0 a x_1 . Por otro lado, $\exists x_2 \exists x_2 0 = 0$ es satisfecha por todas las asignaciones y $\exists x_2 \exists x_2 0 = S_0$ por ninguna. La asignacion particular que esta bajo consideracion solo sera relevante en la medida en que la formula en cuestion tenga variables libres. Si una formula no contiene variables libres, esto es, es una oracion, o bien todas las asignaciones la satisfacen o bien ninguna lo hace. Este concepto binario es lo que da lugar a la noción de verdad.

⁴Notese que las unicas formulas atomicas de L_{pa} son enunciados de identidad.

Definición 2.4 (Verdad de oraciones de L_{pa}).

$$T_{L_{pa}}(x) \text{ def Sent}_{L_{pa}}(x) \wedge \exists g \text{Sat}_{L_{pa}}(g; x)$$

Luego, una oración de L_{pa} es verdadera siempre y cuando toda $\langle \alpha, \beta \rangle$, lo que es lo mismo, alguna asignación para este lenguaje la satisface, y falsa en caso contrario.

$\exists x \exists y (x \neq y \wedge \text{True}(x) \wedge \text{False}(y))$ es verdadera, mientras que $\exists x \exists y (x \neq y \wedge \text{False}(x) \wedge \text{True}(y))$ es falsa.

La definición de verdad dada es formalmente correcta porque sigue el patrón de (EXP): su definición no contiene $T_{L_{pa}}$, aunque sí contiene el predicado $\text{Sat}_{L_{pa}}$, que está definido en términos de sí mismo. Sin embargo, siendo la definición 2.3 recursiva y categorica, es posible obtener una definición explícita de $\text{Sat}_{L_{pa}}$ en la metateoría.⁵

Además, aparte de sí mismo, $\text{Sat}_{L_{pa}}$ no está definido en términos semánticos de ninguna clase sino provenientes exclusivamente de la teoría de conjuntos. Luego, las definiciones 2.3 y 2.4 apuntan a una reducción de la noción de verdad a conceptos no semánticos sino matemáticos, conjuntísticos, lo cual implica la consistencia del concepto de verdad, provisto que la teoría de conjuntos sea consistente.

Naturalmente, esto no basta. Es preciso también establecer la adecuación material de las definiciones dadas para que sean totalmente aceptables. Como indica Tarski, la definición 2.4 satisface de hecho la Convención T y es, por tanto, materialmente adecuada. De acuerdo con la definición 2.1, deberían obtenerse a partir de 2.4

1. todas las instancias de

$$T_{L_{pa}}(p'q) \text{ def } (') \tag{2.4}$$

para cada oración $'$ de L_{pa} , donde $(')$ es una traducción (razonable) de $'$ al lenguaje de la metateoría α ; y

2. $\exists x (T_{L_{pa}}(x) \wedge \neg \text{Sent}_{L_{pa}}(x))$.

La condición 2 es fácilmente verificable leyendo el bicondicional de la definición 2.4 de izquierda a derecha y eliminando el segundo conyunto del definiens. La condición

⁵Se puede probar en la metateoría a que existe una relación R tal que

$$\begin{aligned} \exists g; x (R(g; x) \text{ def } \text{Form}_{L_{pa}}(x) \wedge (\exists y; z (x = (y \neq z) \wedge g(y) = g(z)) \wedge \\ \exists y (x = \neg y \wedge R(g; y)) \wedge \exists y; z (x = (y \neq z) \wedge (R(g; y) \wedge \neg R(g; z))) \wedge \\ \exists y; z (x = \exists y (y \neq z) \wedge \exists w (w \neq y \wedge \exists w (w \neq y \wedge g^0(w) = g(w)) \wedge \neg R(g^0; z)))))) \end{aligned}$$

y que para toda relación R^0 tal que

$$\begin{aligned} \exists g; x (R^0(g; x) \text{ def } \text{Form}_{L_{pa}}(x) \wedge (\exists y; z (x = (y \neq z) \wedge g(y) = g(z)) \wedge \\ \exists y (x = \neg y \wedge R^0(g; y)) \wedge \exists y; z (x = (y \neq z) \wedge (R^0(g; y) \wedge \neg R^0(g; z))) \wedge \\ \exists y; z (x = \exists y (y \neq z) \wedge \exists w (w \neq y \wedge \exists w (w \neq y \wedge g^0(w) = g(w)) \wedge \neg R^0(g^0; z)))))) \end{aligned}$$

$R = R^0$ (cf. Tarski (1933, p. 193, nota 1)).

1, en cambio, es levemente mas compleja y requiere, en primer lugar, de una traduccion 'razonable' de las oraciones de L_{pa} en el metalenguaje. Intuitivamente, lo razonable en este caso consiste en traducir, primero, los terminos de L_{pa} de acuerdo con el siguiente esquema:

$$\begin{aligned} (0) &= 0 \\ (v) &= v \\ (St) &= S(t) \\ (t + s) &= (t) + (s) \\ (t \cdot s) &= (t) \cdot (s) \end{aligned}$$

y luego las formulas:

$$\begin{aligned} (t = s) &= ((t) = (s)) \\ (\neg) &= (\neg) \\ (\forall) &= (\forall) \\ (\exists) &= (\exists) \end{aligned}$$

Luego, da la interpretacion estandar de L_{pa} y es, por ende, una buena traduccion de las formulas de este lenguaje a L_{zfc}^+ . Basandose en esta traduccion es posible probar que todas las instancias de (2.4) se siguen de la definicion 2.4 y, por tanto, esta es materialmente adecuada.⁶ Por ejemplo, se sabe que $T_{L_{pa}}(p \cdot x = 0)$ porque $\text{Sat}_{L_{pa}}(g; p \cdot x = 0)$, pues $\text{Sat}_{L_{pa}}(g; p \cdot x = 0) \iff \text{Sat}_{L_{pa}}(g; p \cdot x = 0)$, pues $\text{Sat}_{L_{pa}}(g; p \cdot x = 0) \iff \text{Sat}_{L_{pa}}(g; p \cdot x = 0)$ o, como x es la unica variable libre, $\text{Sat}_{L_{pa}}(g; p \cdot x = 0) \iff \text{Sat}_{L_{pa}}(g; p \cdot 0 = 0)$, ya que $\text{Sat}_{L_{pa}}(g; p \cdot x = 0) \iff \text{Sat}_{L_{pa}}(g; p \cdot 0 = 0)$, pues $\text{Sat}_{L_{pa}}(g; p \cdot x = 0) \iff \text{Sat}_{L_{pa}}(g; p \cdot 0 = 0)$.

Ademas, la definicion 2.4 permite derivar principios razonables que gobiernan el predicado de verdad para L_{pa} , como los siguientes principios de consistencia y completitud internas:

Consistencia $\forall x (\text{Sent}_{L_{pa}}(x) \rightarrow \neg \text{Neg}_{L_{pa}}(x))$

Completitud $\forall x (\text{Sent}_{L_{pa}}(x) \rightarrow \text{Ver}_{L_{pa}}(x))$

⁶Esto es una consecuencia inmediata de

$(\text{Sat}_{L_{pa}}(g; p'(v_1 :: v_n)q) \iff \text{Sat}_{L_{pa}}(g; p'(v_1 :: v_n)q)) \wedge (\text{Ver}_{L_{pa}}(g; p'(v_1 :: v_n)q) \iff \text{Sat}_{L_{pa}}(g; p'(v_1 :: v_n)q))$

que puede probarse por induccion sobre la complejidad de \cdot .

De acuerdo con el primero, no es posible que una oración y su negación sean ambas verdaderas, mientras que el segundo afirma que toda oración es tal que o bien ella es verdadera o su negación lo es. Como indica Tarski, juntos estos dos principios implican que el conjunto de verdades aritméticas es consistente y completo y, por tanto, diferente del conjunto de teoremas de pa , como muestra el primer resultado de incompletitud de Gödel (véase el apartado [1.2.7](#)), de acuerdo con el cual hay oraciones verdaderas de L_{pa} que no son demostrables en pa .

Tarski ha mostrado que es posible dar una definición de verdad consistente, formalmente correcta y materialmente adecuada [por supuesto, bajo su propia comprensión de estos términos] de verdad para un lenguaje formal, sin recurrir a conceptos semánticos sino exclusivamente a nociones matemáticas claras y precisas. Además, sugiere que su método puede generalizarse a cualquier lenguaje de orden finito o, en términos modernos, a cualquier lenguaje cuyo dominio de discurso tenga el tamaño de un conjunto (esto quedará claro más adelante en el capítulo); aunque no presenta ninguna definición general de verdad para lenguajes con estas características. Nótese que como llevar a cabo esta generalización no es obvio, en tanto las definiciones dadas en esta sección dependen fuertemente de la sintaxis de L_{pa} .

Si bien, al haber sido dadas para tres lenguajes en particular, las definiciones tarskianas de [1933](#) son, en algún sentido, nociones relativas de verdad, en otro sentido de verdad para el lenguaje objeto [L_{pa} en este caso] en términos absolutos, no dependientes de tal o cual interpretación del vocabulario no lógico de este lenguaje sino entendiendo sus símbolos bajo su interpretación usual, como hablando, en el caso de L_{pa} , sobre los números naturales. Por eso, como señala [Feferman \(2008b\)](#), esta noción de verdad está orientada a la comunidad lógica.

En la siguiente subsección presento, en cambio, una definición de verdad para un lenguaje relativa a una estructura, modelo o interpretación de ese lenguaje, destinada más bien a la comunidad lógico-matemática. Esta definición tiene un carácter general en tanto es aplicable a cualquier lenguaje de primer orden; es, en alguna medida, una versión parametrizada de la definición [2.4](#), donde L_{pa} se convierte en parámetro; pero, además, la interpretación canónica del lenguaje se convierte en una variable sobre la cual se cuantifica.

2.2.3. Nociones recursivas de verdad en una estructura

La noción de verdad en una estructura o modelo aparece de manera informal al menos desde el siglo XIX en la literatura lógico-matemática, por ejemplo, al hablar

de modelos para las geometrías no euclidianas, o de la caracterización de Dedekind de los números

naturales y de los reales hasta el isomorfismo en lenguajes de segundo orden. Luego vino el formalismo de Hilbert y el algebra de estructuras. La noción fue central también en la formulación del teorema de Löwenheim-Skolem y el teorema de completitud de Gödel.

[Tarski & Vaught \(1956\)](#) definieron formalmente por primera vez este concepto para, entre otras cosas, dar pruebas más sencillas y extender el resultado de Löwenheim-Skolem. La definición fue esencial para el desarrollo de la teoría de modelos, que resultó de suma utilidad luego en diversas áreas de conocimiento como la filosofía, la lógica, la matemática y otras ciencias deductivas (cf. [Chen Chung Chang & H. Jerome Keisler \(1990\)](#)).

A diferencia del concepto de verdad absoluto presentado en el apartado anterior, relativo a un lenguaje, la noción de verdad en una estructura es, obviamente, no solo relativa a un lenguaje sino también a una estructura. En lugar de trabajar con un lenguaje objeto absolutamente interpretado, se consideran lenguajes de primer orden cuyo vocabulario no lógico carece de significado junto con una familia de modelos posibles para este vocabulario, también llamados 'estructuras' o 'interpretaciones'. Si se entiende la interpretación usual [original, pretendida, canónica] del vocabulario no lógico de un lenguaje, sobre la cual se define la verdad en el apartado anterior, como una entre muchas, si se piensa la definición 2.4 como una definición de verdad para L_{pa} bajo la estructura de los números naturales, interpretando los símbolos no lógicos de este lenguaje como hablando de los números naturales, la definición de verdad en un modelo puede ser vista como una generalización de la definición 2.4 que la extiende a otras interpretaciones de su vocabulario no lógico. Asimismo, la definición tarskiana de verdad en un modelo es en realidad un esquema de definición, donde el lenguaje de primer orden L para el cual está dada es meramente un parámetro: la noción es directamente aplicable a cualquier lenguaje de primer orden.

A diferencia del caso anterior, los modelos o interpretaciones de un lenguaje objeto L no cuentan con la ontología propia de la interpretación canónica del lenguaje, como la de los números naturales, sino que sus dominios están dados exclusivamente por conjuntos. El metalenguaje es simplemente el resultado de incorporar un predicado T_L para la verdad de oraciones de L , ahora diádico para relativizar la verdad a una estructura, y un predicado triádico Sat_L , ahora con un lugar extra para modelos, simplemente a L_{zfc}^+ . En lugar de agregar como urelementos los objetos de los cuales hablar a L y una definición de estos a zfc_u^+ , como se hizo con N en la definición absoluta de verdad para L_{pa} , se trabaja directamente sobre zfc^+ y, en lugar de un conjunto de urelementos, se utiliza un conjunto de conjuntos que sea isomorfo al primero y definible en zfc^+ . Por ejemplo, en lugar de introducir el conjunto N para definir la estructura de los números

naturales, se recurre al \aleph_1 , esto es, el conjunto de los cardinales de von Neumann, el ordinal de los números naturales, o a otros conjuntos isomorfos. Esto siempre será posible en la medida en que el dominio de la estructura no conjuntista tenga una cardinalidad de nible en zfc^+ , siempre y cuando su cardinalidad sea lo suficientemente pequeña. Vuelvo sobre esto más adelante.

Trabajar únicamente con estructuras conjuntistas, por un lado, implica que ningún lenguaje recibirá su interpretación pretendida, excepto quizás ciertos fragmentos de L^+_{zfc} . Pero por otro, simplifica enormemente las definiciones y pruebas sobre la noción de estructura, asignación, satisfacción y verdad sin conllevar otras consecuencias significativas. A continuación ofrezco una definición actual de estructura, modelo o interpretación de un lenguaje L dado en zfc^+ , que difiere poco desde el punto de vista técnico de la original presentada en [Tarski & Vaught \(1956\)](#). Sea L un lenguaje cualquiera de primer orden codificado en \aleph_1 .

Definición 2.5 (Modelos de L). Una interpretación, modelo o estructura M de L es un par ordenado $\langle \text{dom} M; x^M \rangle$, donde $\text{dom} M$, el dominio de M , es un conjunto no vacío y x^M es una función de interpretación, que asigna miembros de $\text{dom} M$ a las constantes de individuo de L , funciones n -adicas de nidas de $\text{dom} M^n$ en $\text{dom} M$ a los símbolos de función n -adicos de L y subconjuntos de $\text{dom} M^n$ a las letras de predicado n -adicos de L .

Por ejemplo, N es una interpretación estándar conjuntista del vocabulario de L_{pa} : $\langle \aleph_1; p^N \rangle$, $p^N = \{ \langle x, y \rangle \mid x \neq y \}$, $f^N = \{ \langle x, y \rangle \mid x \neq y \}$, etc. Tras identificar cada número de von Neumann con un número natural $|x|$ con 0 , f^N con 1 , etc., la definición de verdad de oraciones de L_{pa} en N coincide prácticamente con la definición 2.4. Sea $\text{Mod}_L(M)$ la fórmula de L^+_{zfc} que define la clase de modelos de L , y sea δM una abreviatura de $\delta v(\text{Mod}_L(v))$ para todo L .

Al igual que antes, para dar una definición de verdad en un modelo es preciso primeramente definir satisfacción en un modelo, para poder lidiar con fórmulas, lo cual es necesario para dar las condiciones de verdad de oraciones universalmente cuantificadas. Nuevamente, la noción de satisfacción está dada recursivamente y en función de asignaciones. La noción de asignación para términos de L es relativa a la estructura en la cual se define satisfacción.

Definición 2.6 (Asignaciones de un modelo). Una asignación g de una interpretación M de L es una función cuyo dominio es el conjunto de términos de L y su codominio $\text{dom} M$, tal que

1. si a es una constante de individuo de L , $g(a) = a^M$, y

⁷A menos que se adopte una vision estructuralista de las matematicas. Vease, v.g., [Shapiro \(1997\)](#).

2. si f es un simbolo n -adico de funcion y t_1, \dots, t_n son terminos de L ,

$$g(pf t_1; \dots; t_nq) = pfq^M(g(pt_1q); \dots; g(pt_nq))$$

Sea $As_L(g; M)$ la formula de L^+_{zfc} que define la clase de asignaciones de cada interpretacion M para L . Las asignaciones deben respetar lo asignado por la estructura a las constantes de individuo y simbolos de funcion, mientras que tienen total libertad sobre las variables de individuo.

El predicado de satisfaccion de formulas de L , $Sat_L(v_1; v_2; v_3)$, tendra ahora tres argumentos: la asignacion v_1 del modelo v_3 de L satisface la formula $v_2 \in L$. Sean $R_1^{n_1}; \dots; R_m^{n_m}$ los unicos simbolos de predicado de L de aridad $n_k \geq 1$ con $1 \leq k \leq m$, correspondientemente.

Definicion 2.7 (Satisfaccion de formulas de L por una asignacion de un modelo).

$$Sat_L(g; x; M) \stackrel{\text{def}}{=} Form_L(x) \wedge As_L(g; M) \wedge$$

$$(\exists t_1 \dots t_n (x = (R_1^{n_1} t_1 \dots t_n) \wedge \bigwedge_{i=1}^n t_i \in M) \wedge$$

$$\bigwedge_{i=1}^m (R_i^{n_i} x \in M))$$

\dots

$$\exists t_1 \dots t_m (x = (R_1^{n_1} t_1 \dots t_m) \wedge \bigwedge_{i=1}^m t_i \in M)$$

$$\exists y (x = \lambda y. y \wedge Sat_L(g; y; M))$$

$$\exists y, z (x = \lambda yz. (Sat_L(g; y; M) \wedge Sat_L(g; z; M)))$$

$$\exists z (x = \lambda yz. (As_L(g^0; M) \wedge$$

$$\exists w (\forall v (v \in M \wedge w = y \wedge g^0(w) = g(w)) \wedge Sat_L(g^0; z; M)))$$

Una formula ϕ de L es satisfecha por una asignacion g de una interpretacion M si y solo si ϕ es una formula atomica de la forma $R_i^{n_i} t_1 \dots t_{n_i}$ con $1 \leq i \leq m$ y la n_i -tupla dada por los objetos de $|M|$ que g asigna a $t_1 \dots t_{n_i}$ pertenece a la relacion que x^M asigna a $R_i^{n_i}$, o ϕ es la negacion de ψ y no es satisfecha por g , o ϕ es el condicional de dos formulas ψ y χ , si la primera es satisfecha por g , la segunda tambien, o ϕ es la cuantificacion universal $\forall v \psi$ y todas las asignaciones de M que difieren de g a lo sumo en lo que asignan a v satisfacen ψ .

Notese que la forma de la definicion 2.7 depende esencialmente del vocabulario no logico de L : tanto los predicados y las funciones involucradas ($Sent_L(v), \dots$, etc.) como el

numero de lineas dedicadas a las formulas atomicas, que depende a su vez del numero de letras de predicado en el vocabulario no logico de L. Esto no podra ser de otro modo, en tanto estos predicados y funciones dependen del codigo de expresiones de

L con el cual se trabaje, que es algo externo a la metateoria L. L no ocurre en la definicion

[2.7](#) como variable, como lo hace M, sino como parametro. Estrictamente hablando, hay una definición del estilo de [2.7](#) para cada lenguaje de primer orden L; la definición [2.7](#) es esquemática.

A partir de la noción de satisfacción de fórmulas de L, al igual que antes, se puede definir verdad, solo que ahora, siendo L un lenguaje formal no interpretado, ambas nociones dependen de una u otra interpretación. Por esta razón se agregó un predicado diádico a L_{zfc}^+ para expresar verdad en un modelo. Esta noción de verdad para oraciones de L bajo una estructura, al estar definida en términos de Sat_L , hereda su esquematicidad. Como señala [Eduardo A. Barrio \(2014a, p. 28\)](#), "Tarski no ofreció una definición general, sino un método sistemático para obtener definiciones de verdad para una colección considerable de lenguajes formales."

Definición 2.8 (Verdad de oraciones de L en un modelo).

$$T_L(x; M) \text{ def } Sent_L(x) \wedge \exists g(As_L(g; M) \wedge Sat_L(g; x; M))$$

Si bien Tarski & Vaught no hablan de corrección formal y adecuación material para la definición de verdad relativa a un modelo, se puede probar que, dado un modelo, la definición [2.8](#) es formalmente correcta y materialmente adecuada. Es posible dar una definición explícita de Sat_L en L_{zfc}^+ utilizando la misma técnica esbozada en el apartado anterior y, por tanto, de acuerdo con [EXP](#), la definición [2.8](#) es formalmente correcta. Además, al igual que para la definición [2.4](#), dado un modelo M de L se puede probar todas las instancias de

$$T_L(p'q; M) \text{ def } (') \tag{2.5}$$

donde τ es una función que traduce oraciones de L en L_{zfc} tal que, si a es una constante de individuo, v una variable, t_1, \dots, t_n términos, y $'$ y τ fórmulas de L,

$$\begin{aligned} (a) &= paq^M \\ (v) &= v \\ (ft_1 \dots t_n) &= pfq^M((t_1); \dots; (t_n)) \end{aligned}$$

y luego las fórmulas:

$$\begin{aligned} (R_i^{n_i} t_1 \dots t_{n_i}) &= h(t_1; \dots; t_{n_i}) \text{ si } 1 \leq i \leq m \text{ (}') \\ &= (') \\ (') &= (') \wedge (') \\ (8v') &= \exists v \text{ (}') \end{aligned}$$

Además, dado un modelo M de L , la definición 2.8 implica que

$$\exists x(T_L(x; M) \leftrightarrow \text{Sent}_L(x))$$

Por tanto, para cada modelo M de L , $T_L(v; M)$ satisface la Convención T: la definición 2.8 es materialmente adecuada.

Notese que a partir de la definición 2.8, cuando sobre M es posible definir la noción de verdad en todo modelo o, lo que es lo mismo, de verdad lógica en sentido contemporáneo. También se puede obtener definiciones análogas de falsedad lógica, satisfacibilidad, consecuencia lógica, y otros conceptos lógicos, aunque se han dado argumentos en contra de la adecuación material de estas definiciones.⁸

2.3. Limitaciones y jerarquías

Tarski ha dado, por un lado, definiciones de verdad absoluta para ciertos lenguajes formales interpretados, como la definición 2.4, y, por otro, definiciones de verdad para cualquier lenguaje de primer orden relativas a un modelo de este lenguaje, que introduce en la definición 2.8, en ambos casos satisfaciendo las condiciones impuestas en el apartado 2.1. Como señalamos más de una vez, esta última definición puede entenderse como una generalización de las definiciones del estilo de 2.4 en dos sentidos. Por un lado, 2.4 está anclada a la interpretación canónica del lenguaje objeto, mientras que la definición 2.8 permite variar el significado de sus símbolos no lógicos. Por otro, la definición 2.4 concierne exclusivamente a L_{pa} , mientras que la otra es un esquema aplicable a cualquier lenguaje de primer orden.

Sin embargo, hay un sentido en el cual la definición de verdad en un modelo no es una generalización de la de verdad absoluta para L_{pa} . Si bien la primera ofrece muchas más interpretaciones que la segunda, la interpretación particular de la cual la segunda hace uso implícito no se encuentra entre las que ofrece la definición de verdad en un modelo, pues su dominio no está formado por conjuntos. Solo las interpretaciones conjuntistas son consideradas por la definición 2.8. Una podría conformarse con interpretaciones conjuntistas isomorfas a las originales; pero, lamentablemente, no toda interpretación canónica tiene una interpretación conjuntista isomorfa a ella.

En esta sección pruebo primero un resultado general que implica este hecho, el teorema de la indecibilidad de la verdad de Tarski en términos formales. Luego, en 2.3.2 presento un modo de solucionar la dificultad de dar nociones de verdad absolutas

⁸Vease, por ejemplo, [John Etchemendy \(1990\)](#), [Vann McGee \(1992b\)](#), [Manuel García-Carpintero](#)

(1993), [Gomez-Torrente \(1998/1999\)](#) para una discusion.

para cualquier lenguaje formal de primer orden relajando ciertas características de la noción de interpretación. Finalmente, en [2.3.3](#) exploro una segunda vía que circunviene el problema, evitando hablar de interpretaciones simpliciter. Ambos caminos desembocan en jerarquías.

2.3.1. Indenitabilidad de la verdad

Si se quisiera dar una definición de verdad absoluta para L_{zfc}^+ no se podría recurrir a una instancia de la definición [2.8](#), pues la estructura que se precisa debe ser la interpretación usual, estándar, de este lenguaje. I.e., debería existir un modelo M en donde $|M|$ contenga todos los conjuntos del universo de zfc^+ , lo cual es imposible, pues zfc^+ debería implicar la existencia de un conjunto de todos los conjuntos que, por la paradoja de Russell, se sabe que no existe.⁹

Naturalmente, esto no es ninguna novedad. De haber sido posible dar una definición de verdad absoluta para L_{zfc}^+ en términos de la definición [2.8](#), que está formulada en L_{zfc}^+ , se habría dado con un lenguaje semanticamente cerrado y, por tanto, inconsistente. En otras palabras, zfc^+ sería una teoría trivial. La imposibilidad de dar esta definición es un simple corolario del teorema de Tarski. Ahora se está en condiciones de dar una versión formal de este resultado.

Teorema 2.9 (Tarski. Indenitabilidad de la verdad). Si \mathcal{L} es una teoría capaz de interpretar relativamente q y existe una fórmula $T(v) \in \mathcal{L}$ tal que, para toda oración $\phi \in \mathcal{L}$, \mathcal{L} prueba

$$T(p \rightarrow q) \rightarrow \neg T(p) \quad (\text{Esquema-T})$$

donde $p \rightarrow q$ denota en \mathcal{L} el código de Gödel de $\phi \rightarrow \psi$, \mathcal{L} es inconsistente.¹⁰

Demostración. Si \mathcal{L} es capaz de interpretar relativamente q , es capaz, por un lado, de codificar la sintaxis de L y, por otro, de probar una traducción del lema diagonal débil. Aplicando este lema al predicado $\neg T(v)$, existe una oración tal que

$$\mathcal{L} \vdash \neg T(p \rightarrow q)$$

que contradice directamente su instancia correspondiente del [Esquema-T](#). □

⁹Si existiera, el axioma de separación de zfc^+ permitiría inferir la existencia del conjunto de todos los conjuntos que no son miembros de sí mismos, esto es, el conjunto de Russell, que lleva directamente a inconsistencias.

¹⁰Notese que el guion en [Esquema-T](#) diferencia al [Esquema T](#) de este principio formulado en un lenguaje de primer orden, una de sus posibles formalizaciones.

En otras palabras, si una teoría contiene su propio predicado de verdad y es tanto o más poderosa que la aritmética de Robinson [y ciertamente zfc^+ lo es] es posible formular una oración del mentiroso y reproducir en la teoría el razonamiento informal que lleva de este tipo de oraciones a una contradicción. Se obtiene así la paradoja del mentiroso en una de sus versiones formales. Si la definición 2.5 proveyera una estructura M en la cual el vocabulario de L^+_{zfc} sea interpretado de modo estándar, el predicado de L^+_{zfc}

$$\text{Sent}_{L^+_{zfc}}(v) \wedge \exists g(\text{As}_{L^+_{zfc}}(g; M) \wedge \neg \text{Sat}_{L^+_{zfc}}(g; v; M))$$

[que abrevio $T(v)$] ser a tal que

$$zfc^+ \vdash \neg T(p'q) \text{ } (\cdot)$$

para toda $'$ de L^+_{zfc} , donde $(\cdot) = \cdot$. Por lo tanto, zfc^+ probar a todas las instancias del [Esquema-T](#) dadas por oraciones de su propio lenguaje y, en virtud del teorema de indecibilidad de la verdad de Tarski, ser a inconsistente.

>Implica esto que se debe aceptar la tesis del pesimismo semántico, esto es, la idea de que hay lenguajes para los cuales no es posible dar una definición de verdad? En [1944](#) Tarski sostiene que

It should be noticed that these terms "object-language" and "metalanguage" have only a relative sense. If, for instance, we become interested in the notion of truth applying to sentences, not of our original object-language, but of its meta-language, the latter becomes automatically the object-language of our discussion; and in order to define truth for this language, we have to go to a new meta-language|so to speak, to a meta-language of a higher level. In this way we arrive at a whole hierarchy of languages. (Tarski, 1944, p. 350)

Luego, es posible dar una definición de verdad para un lenguaje cualquiera en la medida en que se encuentre un lenguaje de nivel superior o, como también suele decir Tarski, un lenguaje esencialmente más rico, que haga las veces de metalenguaje. Dado que Tarski ofrece sus definiciones de verdad absoluta en el marco de la teoría general de clases, traduce esta riqueza esencial para este caso particular en la idea de órdenes superiores. Un lenguaje de orden superior a otro [v.g., un lenguaje de segundo orden con respecto a uno de primer orden] contiene variables y cuantificadores que varían sobre una colección que contiene entre sus miembros a la clase de objetos de los cuales 'habla' el lenguaje de orden inferior y, por esta razón, es capaz de dar una definición de verdad para este último. En [1933](#) Tarski concluye que no es posible dar una definición de verdad absoluta para el lenguaje de la teoría general de clases pues, al contener todos los órdenes, no existe un lenguaje de orden mayor desde el cual dar la definición.

La situación para L_{zfc}^+ y zfc^+ parece ser analoga. Si bien las definiciones de verdad en zfc^+ no recurren a variables de distinto orden sino a conjuntos que contienen otros conjuntos, para poder dar una definición de verdad absoluta para un lenguaje L en zfc^+ se precisa ahora que esta teoría pruebe la existencia de un conjunto isomorfo al conjunto dado por la ontología de la interpretación pretendida de L . Así interpreto preliminarmente la idea de riqueza esencial para el caso particular de las definiciones dadas desde la teoría de conjuntos. Dado que la ontología de zfc^+ contiene todos los conjuntos, entre los cuales no se encuentra ningún conjunto isomorfo al conjunto de todos los conjuntos pues este no existe, no es posible dar una definición de verdad absoluta para L_{zfc}^+ .

No obstante, existen otros caminos a explorar. Por un lado, el método axiomático, en el que [Tarski \(1933, 1944\)](#) mismo incursiona. Por otro, interpretaciones alternativas razonables de la idea de riqueza esencial, que permitan trabajar con entidades diferentes a los conjuntos. Exploro primero esta última vía.

2.3.2. Jerarquías semánticas

[Raul Orayen \(2003\)](#) propone relajar el requisito de la teoría de modelos de acuerdo con el cual los conjuntos son las únicas entidades que pueden dar lugar a dominios e interpretaciones de lenguajes de primer orden, para poder contar con una interpretación canónica de L_{zfc} . Con este objetivo introduce la noción de conjunto de segundo tipo o, preferentemente, clase. En este apartado exploro la propuesta de Orayen.¹¹

La metateoría, una teoría de clases, ya no será zfc^+ sino más poderosa. Sean L_{tc0} otro nombre para L_{zfc}^+ , tc_0 otro nombre para zfc^+ , L_{tc1} el resultado de incorporar una constante de individuo c_0 a L_{tc0} y tc_1 la teoría de primer orden formulada en L_{tc1} que cuenta con los axiomas de zfc^+ tal como aparecen en esta teoría más estos mismos axiomas relativizados al predicado $\forall x \in c_0$.

En un sentido bastante claro, tc_1 es esencialmente más rica que tc_0 ya que, si bien no prueba la existencia del conjunto de todas las entidades de las cuales L_{tc0} naturalmente habla [i.e., todos los conjuntos] prueba la existencia de una colección c_0 llamada 'clase', que las contiene a todas. Además, en virtud del axioma no relativizado de separación, tc_1 es capaz de probar la existencia de una colección que contiene todos y exactamente aquellos pares ordenados $\langle x, y \rangle$ de miembros de c_0 tales que $x \neq y$, esto es,

¹¹Una vía alternativa pero semejante consiste en permitir modelos constituidos no ya por entidades de tipo conjuntista sino por pluralidades, colecciones que, a diferencia de aquellas, no dan lugar a nuevas entidades diferentes de los miembros que las componen. Veanse [Boolos \(1984, 1985\)](#), [Rayo & Uzquiano \(1999\)](#) y [Uzquiano \(2003\)](#). Para explorarla será necesario presentar el lenguaje y la teoría de las pluralidades, lo cual me llevaría demasiado lejos del punto central de esta tesis.

tc_1 es capaz de probar la existencia de la interpretación estándar de L_{zfc}^+ .¹² Asimismo, es posible definir en tc_1 la noción de asignación sobre esta interpretación y, por tanto, las nociones de satisfacción y verdad absoluta para oraciones de L_{zfc}^+ exactamente como en las definiciones 2.6, 2.7 y 2.8. La definición de verdad resultante es no solo formalmente correcta sino también materialmente adecuada.

La pregunta que surge naturalmente ahora es si acaso se puede definir verdad absoluta para oraciones de L_{tc_1} . Dada la nueva interpretación propuesta de la idea de riqueza esencial, lo que se precisa es un metateor a capaz de probar la existencia de una colección de todas las clases que pertenecen a la ontología pretendida de L_{tc_1} . Puesto que la existencia de una clase con estas características lleva a un absurdo en la teoría de clases tc_1 , el metateor a deberá postular un nuevo tipo de entidades, las hiperclases. Sea tc_2 un sistema formulado en el lenguaje L_{tc_2} , que resulta de expandir L_{tc_1} con una nueva constante de individuo c_1 para la hiperclase de todas las clases, y cuenta con los axiomas de tc_1 relativizados a $\forall z c_1$ más una copia de los axiomas de zfc^+ sin relativizar. Al igual que antes, tc_2 es capaz de definir la interpretación estándar de L_{tc_1} y las nociones de asignación, satisfacción y verdad adecuada y correctamente para este lenguaje.

La iteración de este proceso da lugar a una jerarquía de lenguajes $L_{tc_0}; L_{tc_1}; L_{tc_2}; \dots$, donde cada miembro de la jerarquía es el lenguaje objeto del que le sigue y cada lenguaje sucesor es el metalenguaje de su predecesor, y una jerarquía de teorías $tc_0, tc_1, tc_2; \dots$, donde cada una funciona como metateor a desde la cual definir verdad para oraciones del lenguaje de la teoría que la precede. De este modo, es posible definir verdad absoluta para las oraciones de cualquier lenguaje en la jerarquía. El proceso puede iterarse transitivamente, hasta que la notación para números ordinales a disposición lo permita (cf. Halbach (1995)).

Por supuesto, esta construcción genera las inquietudes que suelen plantearse al-rededor de las jerarquías, la más importante de las cuales es quizás como definir verdad para el lenguaje en el cual se construye la jerarquía misma, si es que es (y parece ser) un lenguaje formal [i.e., la unión de los lenguajes de cada uno de los niveles]. De no ser posible, esto implicaría que la tesis de la generalidad absoluta, esto es, la posibilidad de hablar acerca de todo como en "Todo es idéntico a sí mismo", es falsa. No me adelanto acá en estas cuestiones, que me llevarán demasiado lejos del objetivo de esta sección (cf. Rayo & Gabriel Uzquiano (2006a)).

¹²La siguiente instancia del axioma de separación o comprensión restringida da lugar a la interpretación canónica de \exists en L_{zfc}^+ :

$$\exists x \forall y (y \in x \leftrightarrow \exists z (c^2 \wedge \delta z; w(y = hz; \forall i \exists z w)))$$

Otra duda que despierta la jerarquía introducida es cuál es la diferencia, si es que existe, entre un conjunto, una clase, una hiperclase, etc. Todas estas entidades 'colectivas' obedecen axiomas análogos. El cambio de nombre parece responder exclusivamente a la necesidad de dar una definición de verdad para los lenguajes de cada una de estas teorías, es puramente ad hoc (cf. Boolos (1984)). No me queda claro que esto sea realmente un problema.

2.3.3. Jerarquías axiomáticas

El enfoque axiomático [el camino que sigue Tarski para ofrecer una noción de verdad para la teoría general de clases o, en este caso, para zfc^+] evita los problemas que he señalado para el enfoque semántico, aunque abandona el proyecto de dar una definición formalmente correcta, i.e., explícita, del predicado veritativo. Hay dos formas de proceder. Una primera opción que se encuentra ya en Tarski (1933, x5) consiste en tomar como axiomas todas las instancias del [Esquema T Tipeado](#) para oraciones de L_{zfc}^+ , aunque aquí presento un esquema de teoría para cualquier lenguaje L que no contenga el predicado T . Una segunda opción es tomar como axiomas cláusulas semejantes a las de la definición recursiva de verdad [2.8](#) para L_{zfc} , dejando de lado los modelos y empleando Sat como un símbolo de predicado diádico primitivo. A continuación exploro ambas vías en este orden.

Sea th una teoría de primer orden cualquiera formulada en L capaz de codificar su propia sintaxis que axiomatice (parte de) el conjunto de oraciones verdaderas de L bajo su interpretación pretendida.

Definición 2.10 ($tb(th)$). $tb(th)$ $L[T]$ extiende th mediante todas las instancias del [Esquema-T](#) dadas por oraciones $\ulcorner L$.

$tb(th)$ (por Tarski-biconditionals sobre th) es una teoría desentrecorolladora de la verdad, porque sus axiomas están dados exclusivamente por instancias del [Esquema-T](#) que, como se verá en el capítulo 4, es un principio de desentrecorollación. Las instancias del [Esquema-T](#) en $tb(th)$ garantizan automáticamente la adecuación material de esta teoría como una definición de verdad, aunque no la corrección formal. Ahora, para dar un concepto de verdad aplicable a oraciones del lenguaje de $tb(th)$ es preciso iterar el proceso: incorporar un nuevo predicado monádico al lenguaje para expresar verdad en el lenguaje de $tb(th)$ y nuevos axiomas, semejantes al [Esquema-T](#) pero ahora para oraciones \ulcorner del lenguaje de $tb(th)$. Así, se obtiene una jerarquía de lenguajes y otra de teorías axiomáticas de la verdad, que puede extenderse, al igual que las jerarquías anteriores, a niveles trans finitos.

También es posible dar una versión más completa de esta axiomatización mediante bicondicionales uniformes.¹³

Definición 2.11 (utb(th)). utb(th) L [FT] g que extiende th mediante todas las instancias de

$$\forall t \left(\left(\bigwedge_{p \in \mathcal{P}} (T(p; \dots; t)) \rightarrow \bigwedge_{q \in \mathcal{Q}} (q; \dots; t) \right) \rightarrow \left(\bigwedge_{p \in \mathcal{P}} (p; \dots; t) \rightarrow \bigwedge_{q \in \mathcal{Q}} (q; \dots; t) \right) \right) \quad (\text{Esquema-T Uniforme})$$

donde \mathcal{P} y \mathcal{Q} tienen n variables libres.

utb(th) (por Uniform Tarski-biconditionals sobre th) también es considerada una teoría desentrecorolladora. De acuerdo con [Tarski \(1933\)](#), las teorías cuyos axiomas están dados meramente por instancias formales del [Esquema T Tipeado](#), ya sean uniformes o no, son altamente incompletas, porque no son capaces de probar ciertos principios generales como los de consistencia y completitud introducidos en el apartado [2.2.2](#) (cf. [Halbach \(2011\)](#), teorema 7.6). Si bien es claro que tb(th) prueba todas las instancias de

$$\vdash T(p \rightarrow q) \rightarrow \vdash T(p) \rightarrow \vdash T(q)$$

con \mathcal{P} y \mathcal{Q} , no es capaz de probar el principio general de consistencia dado por

$$\exists x (\text{Sent}_L(x) \rightarrow \vdash T(x) \rightarrow \vdash T(\neg x)) \quad (2.6)$$

y lo mismo con respecto a otros principios generales que intuitivamente gobiernan el predicado de verdad. Esto evidencia que el requisito de adecuación material para una teoría de la verdad es solo una condición mínima para Tarski, no es suficiente como podría en principio esperarse.

Tarski considera la posibilidad de agregar este y otros principios generales como axiomas a tb(th) (y, correspondientemente, a cada teoría de la jerarquía), pero rápidamente la descarta por considerarla ad hoc. Desde su punto de vista, el único criterio razonable para una axiomatización del predicado de verdad es la categoricidad, i.e., que el sistema caracterice unívocamente la extensión de T: si se agrega a la teoría una copia de los axiomas para T pero esta vez para T^0 , debería poder probarse dentro del sistema que T y T^0 son coextensionales. Lamentablemente, Tarski advierte, esto no es posible. El teorema de Beth implica que si un predicado es categoricamente axiomatizable sobre una teoría, es posible también dar una definición explícita de este en aquella (cf. [Chang & Keisler \(1990\)](#), p. 90 y ss), lo cual contradice al teorema [2.9](#).

¹³Para los detalles sobre la notación empleada en la formulación del [Esquema-T Uniforme](#), véase el [apartado 1.2.3](#).

Por estas razones Tarski pre ere de nir el predicado de verdad a axiomatizarlo. Sin embargo, sus razones en contra del abordaje axiomático no pueden ser concluyentes. De serlo, se aplicar an a cualquier teor a incompleta en el sentido de G•odel, como a la aritmetica de Peano o a la teor a de conjuntos de Zermelo-Fraenkel. Ninguna de estas dos ofrece caracterizaciones categoricas de las propiedades de los numeros naturales o los conjuntos de los cuales hablan, respectivamente, como indican los teoremas de incompletitud de G•odel (vease el apartado [1.2.7](#)). Como sostiene Halbach,

The impossibility of a categorical axiomatization should not keep one from studying axiomatic theories of truth, any more than the impossibility of a complete axiomatization of arithmetic or of set theory should keep one from studying axiomatic systems for integers or sets. ([Halbach, 2011](#), p. 23)

Una segunda opcion axiomática que las de niciones tarskianas mismas sugieren consiste en tomar directamente las clausulas recursivas o composicionales de las de niciones de satisfaccion y transformarlas en axiomas, dando lugar a una teor a de la verdad composicional. Como una axiomatizacion de estas caracter sticas depende esencialmente de la sintaxis del lenguaje objeto, ilustro este camino mediante ejemplos: primero con

L^+_{zfc} y luego con L^+_{pa} .

De nicion 2.12 ($ct(zfc^+)$). Sea $ct(zfc^+)$ la teor a formulada en la expansion de L^+_{zfc} con un predicado diadico Sat y uno monadico T que extiende zfc^+ con los axiomas

$$ct(zfc^+)_1 \ \exists t; s \exists g (F \text{orm}_{L^+_{zfc}} (t=.s) \ \& \ (\text{Sat}(g; t=.s) \ \$ \ g(t) = g(s)))$$

$$ct(zfc^+)_2 \ \exists t; s \exists g (F \text{orm}_{L^+_{zfc}} (t2.s) \ \& \ (\text{Sat}(g; t2.s) \ \$ \ g(t) \ 2 \ g(s)))$$

$$ct(zfc^+)_3 \ \exists x (F \text{orm}_{L^+_{zfc}} (x) \ \& \ \exists g (\text{Sat}(g; x) \ \$ \ : \text{Sat}(g; :.x)))$$

$$ct(zfc^+)_4 \ \exists x; y (F \text{orm}_{L^+_{zfc}} (x!.y) \ \& \ \exists g (\text{Sat}(g; x!.y) \ \$ \ (\text{Sat}(g; x) \ \& \ \text{Sat}(g; y))))$$

$$ct(zfc^+)_5 \ \exists x; y (F \text{orm}_{L^+_{zfc}} (\exists .xy) \ \& \ \exists g (\text{Sat}(g; \exists .xy) \ \$ \ \exists g^0 (\exists z (V \text{ar}(z) \ \wedge \ z \ 6= x \ \& \ g^0(z) = g(y)) \ \& \ \text{Sat}(g^0; y))))$$

$$ct(zfc^+)_6 \ \exists x (T \ x \ \$ \ \text{Sent}_{L^+_{zfc}} (x) \ \wedge \ \exists g \text{Sat}(g; x))$$

donde expresiones de la forma $\exists g'$ abrevian $\exists v (As_{L^+_{zfc}} (v) \ \& \ ')$ y $As_{L^+_{zfc}} (v)$ de ne el conjunto al cual pertenecen todas y solo aquellas funciones que asignan conjuntos a los terminos de L^+_{zfc} tal que `?' tiene por imagen siempre al conjunto vac o y $f \ t_1 : : : t_n$, donde f es un s mbolo de funcion del lenguaje y $t_1; : : : t_n$ son terminos, recibe la imagen de la n-tupla dada por lo que es asignado a $t_1; : : : t_n$ a traves de la funcion de nida por f.

Procediendo por inducción sobre la complejidad de \cdot se puede probar que la teoría $ct(zfc^+)$ (por Compositional Truth sobre zfc^+) implica todas las instancias del [Esquema-T](#) para L_{zfc}^+ . Asimismo, es capaz de probar muchos de los principios generales que Tarski requiere, v.g., el principio de consistencia dado por [\(2.6\)](#) a partir de $ct(zfc^+)_3$ y $ct(zfc^+)_6$. Al igual que en el caso de $tb(zfc^+)$, es posible iterar esta axiomatización agregando un nuevo predicado de satisfacción y un nuevo predicado de verdad, ahora para fórmulas y oraciones del lenguaje de $ct(zfc^+)$, y luego incorporando a esta teoría axiomas análogos, ahora relativizados a fórmulas y oraciones del lenguaje de $ct(zfc^+)$. Llevar esto a cabo reiteradas veces da lugar nuevamente a una jerarquía de lenguajes, teorías y axiomatizaciones del predicado de verdad, que puede extenderse a niveles trans finitos (cf. [Halbach \(2011, cap. 9\)](#)). Naturalmente, estas axiomatizaciones tampoco serán completas o categóricas, en virtud del primer teorema de incompletitud de Gödel.

El requisito de corrección formal que Tarski impone a las definiciones de verdad tiene por objeto dar una garantía automática de consistencia al concepto de verdad. Pero, como señala [Etchemendy \(1988\)](#), esto no es necesario, en tanto dar una definición explícita no es el único modo de garantizar consistencia. Si el predicado de verdad no es introducido explícitamente, hay que dar una prueba de su consistencia aparte. Esto es efectivamente lo que hace [Tarski \(1933, x5, teorema III\)](#) mediante un argumento por compacidad cuando introduce un predicado de verdad como un primitivo en la meta-teoría, dando lugar a una teoría del estilo de $tb(zfc^+)$. Esta prueba de consistencia tarskiana puede modificarse levemente para probar la consistencia de $tb(th)$. La consistencia de $ct(zfc^+)$ también puede establecerse, pero su demostración es un tanto más compleja (cf. [Halbach \(2011, teorema 8.42\)](#)).

Como también es posible codificar las expresiones de cualquier lenguaje de primer orden con números, L_{pa} es capaz de hablar de ellos y de definir ciertas propiedades recursivas suyas, y pa de representar una cantidad considerable de estas propiedades (vease el apartado [1.2.3](#)), también es posible axiomatizar las cláusulas recursivas tarskianas sobre pa^+ . Dado que, además, L_{pa}^+ contiene nombres para cada uno de los individuos del dominio de su modelo pretendido, se puede prescindir del predicado de satisfacción y las asignaciones, y trabajar únicamente con el de verdad y el auxilio de la función sustitución (vease el apartado [1.2.3](#)). Por ejemplo, en lugar de $\exists g \text{Sat}(g; p'q)$ se puede escribir $\exists t T p'q(t; pvq)$ o, directamente, $\exists t T p'(t).q$, si \cdot contiene una única variable libre v . Recuerdese que L_T extiende L_{pa}^+ con T y pat es el resultado de formular pa^+ en L_T (vease el apartado [1.2.4](#)).

Definición 2.13 ($ct(pa^+)$). $ct(pa^+)_{L_T}$ es la teoría que extiende pat con los siguientes axiomas:¹⁴

$$ct(pa^+)1 \quad \forall t; s(T(t=s) \leftrightarrow t = s)$$

$$ct(pa^+)2 \quad \forall x(Sent_{L_{pa^+}}(x) \rightarrow (T(x) \leftrightarrow T(x)))$$

$$ct(pa^+)3 \quad \forall x; y(Sent_{L_{pa^+}}(x \rightarrow y) \rightarrow (T(x \rightarrow y) \leftrightarrow (T(x) \rightarrow T(y))))$$

$$ct(pa^+)4 \quad \forall x; v(Sent_{L_{pa^+}}(\forall v. x) \rightarrow (T(\forall v. x) \leftrightarrow \forall t T(x(t=v))))$$

Este sistema también puede iterarse hasta cualquier ordinal para el cual se dis-ponga de notación en L_{pa^+} (vease el apartado 1.4). Las teorías de la jerarquía resultante se conocen como 'teorías de la verdad ramificada' (Ramified Truth). Sea $L_{<\theta}$ otro nombre para L_{pa^+} y, para cada $0 < \theta$, sea $L_{<}$ el resultado de expandir L_{pa^+} mediante los nuevos predicados monádicos $T_{<}$ para cada $<$. $pat_{<}$ es el resultado de formular pa^+ en $L_{<}$, con todas instancias del axioma esquema de inducción dadas por las fórmulas de este lenguaje. Como antes, si θ , $Sent_{L_{<}}(v)$ representa en pa^+ el conjunto de oraciones de $L_{<}$. Sin embargo, para la formulación de las teorías que siguen va a ser preciso a veces cuantificar sobre $<$, lo cual es posible dado que existe una codificación recursiva de los ordinales hasta θ . Por eso trabajo en su lugar con un predicado diádico $Sent_{L_{<}}(v) := Sent_L(<; v)$.

Definición 2.14 ($rt_{<}$). Sea $rt_{<\theta}$ otro nombre para pa^+ , $rt_{<1}$ otro para $ct(pa^+)$, reemplazando las ocurrencias de T por T_{θ} y, para todo otro θ , sea $rt_{<}$ la teoría formulada en $L_{<}$ que extiende $pat_{<}$ con los siguientes axiomas para cada $< < :$

$$rt_{<}1 \quad \forall t; s(T_{<}(t=s) \leftrightarrow t = s)$$

$$rt_{<}2 \quad \forall x(Sent_{L_{<}}(x) \rightarrow (T_{<}(x) \leftrightarrow T_{<}(x)))$$

$$rt_{<}3 \quad \forall x; y(Sent_{L_{<}}(x \rightarrow y) \rightarrow (T_{<}(x \rightarrow y) \leftrightarrow (T_{<}(x) \rightarrow T_{<}(y))))$$

$$rt_{<}4 \quad \forall x; v(Sent_{L_{<}}(\forall v. x) \rightarrow (T_{<}(\forall v. x) \leftrightarrow \forall t T_{<}(x(t=v))))$$

$$rt_{<}5 \quad \forall t(Sent_{L_{<}}(t) \rightarrow (T_{<}(T_{<}(t)) \leftrightarrow T_{<}(t)))$$

$$rt_{<}6 \quad \forall t \quad \forall \theta \quad (Sent_{L_{<}}(t) \rightarrow (T_{<}(T_{<}(t)) \leftrightarrow T_{<}(t)))$$

Si $\theta > 1$, los axiomas $rt_{<}1$ - $rt_{<}4$ son simplemente generalizaciones de los axiomas de $ct(pa^+)$ para expresiones de $L_{<}$ y sus predicados de verdad. $rt_{<}5$ completa el axioma $rt_{<}1$ al lidiar con las nuevas fórmulas atómicas del lenguaje. El axioma $rt_{<}6$,

¹⁴*En algunas partes de la literatura esta teoría aparece bajo el nombre de t(pa).*

en cambio, no tiene análogo en $ct(pa^+)$ sino que establece la acumulatividad de los predicados de verdad: si $\alpha < \beta$, T_α y T_β coinciden en todas las oraciones de $L_{\alpha+1}$ y, si α es un ordinal límite, T_α es la unión de todos los predicados de nivel inferior. Notese que en este axioma no se cuantifica sobre los subíndices de los predicados de verdad, ya que esto no es posible, sino sobre los símbolos cuyos subíndices son menores a un ordinal dado α , que conforman un conjunto recursivo bajo ciertas codificaciones naturales de los predicados de la jerarquía con números naturales.

Si $\alpha < \beta$ y $\beta \in L_{\alpha+1}$, T_α prueba todas las instancias del esquema

$$T_\alpha(p'q) \leftrightarrow p'$$

satisfaciendo el requisito de adecuación material de Tarski para oraciones de $L_{\alpha+1}$. Las cláusulas composicionales implican también los principios de consistencia y completitud y otros principios generales deseables a para cada uno de los predicados de verdad del lenguaje de la teoría T_α .

Hay entonces infinitas teorías a disposición, una por cada ordinal α . Las más naturales parecen ser aquellas dadas por T_α , que contiene todos los niveles finitos; T_α , porque pa^+ es capaz de probar instancias del principio de inducción transfinita (IT) para (los códigos de los) ordinales menores a α (α , véase el apartado 1.4); y T_α , porque una vez que se avanza hasta α más instancias de (IT) son demostrables, con lo cual se puede iterar la teoría más allá, que a su vez dar lugar a nuevas instancias de (IT), etc. Este proceso alcanza un punto jón en α : T_α no prueba instancias de (IT) transfinita para segmentos de ordinales más grandes que α .

En consecuencia, la distinción tarskiana entre lenguaje objeto y metalenguaje, al trabajar únicamente con teorías de verdad tipeada, desemboca inevitablemente en jerarquías o en lenguajes formales para los cuales no es posible dar una definición de verdad a la Tarski, esto es, en el pesimismo semántico.

3

Teorías de la verdad no tipeadas

"So one often hears that the surest way of keeping a language paradox-free is to impose an absolute ban on all self-reference. This may be using a cannon against a y," it is said, "but at least it stops the y."

{ Stephen Yablo, 'Paradox without Self-Reference'

Si bien las nociones tarskianas resultaron de extrema utilidad y constituyeron una clara mejora de la situación en la cual se encontraban previamente las investigaciones en torno a la verdad, la distinción entre lenguaje y metalenguaje, la imposibilidad de utilizar un predicado de verdad para adscribir verdad a oraciones que contienen ya este predicado y, en consecuencia, la enorme brecha entre el predicado veritativo del lenguaje natural y las nociones tarskianas, dieron lugar a una extensa serie de críticas e intentos por dar de nociones o teorías de verdad que superen las dificultades del abordaje tipeado de Tarski, aunque sin apartarse de su metodología formal.

Naturalmente, si se quiere abandonar la distinción entre lenguaje y metalenguaje es preciso tomar ciertas medidas para evitar caer en las hipótesis del teorema de indecidibilidad de la verdad de Tarski y, por ende, en trivialidad. Grosso modo, las teorías de verdad no tipeada disponibles en la literatura han tomado alternativamente dos caminos: o bien debilitar la lógica, i.e., abandonar la lógica clásica, o bien restringir significativamente las instancias del [Esquema-T](#), si bien no únicamente a oraciones que no contienen el predicado veritativo.

En este capítulo exploro algunas de las teorías no tipeadas más salientes de la literatura, aquellas que son relevantes para esta investigación (vease el apartado [5.3](#)), todas ellas inspiradas en el trabajo de Tarski. En [3.1](#) presento las teorías semánticas paracompletas de Punto Fijo de Kripke, junto con dos axiomatizaciones clásicas. La

sección 3.2 considera la teoría semántica clásica de Leitgeb basada en su noción de dependencia. En 3.3 presento la teoría axiomática clásica Friedman-Sheard. Finalmente, en 3.4, introduzco tres teorías de la verdad cuya lógica subyacente es paraconsistente: una teoría formulada sobre la lógica paraconsistente básica, la teoría de Graham Priest y la teoría de David Ripley.

3.1. Las teorías de Punto Fijo de Kripke

Aunque vago y problemático, el lenguaje natural, al contener un único predicado de verdad aplicable a oraciones que ya lo contienen, presenta ciertos usos deseables del predicado veritativo que no es posible replicar con predicados tarskianos. En otras palabras, las nociones formales tipeadas de verdad dejan de lado ciertas características del lenguaje cotidiano, algunas de las cuales son necesarias para ciertos propósitos. Por ejemplo, no es posible utilizar un predicado tarskiano para afirmar que todas las oraciones de la forma "Si A entonces A" son verdaderas, pero esto es perfectamente posible e inofensivo en el lenguaje natural. Aun más, ningún predicado de verdad de la jerarquía de Tarski puede expresar la idea de que todos los teoremas de cada metateoría de esta misma jerarquía son verdaderos, lo cual, si la jerarquía es correcta, es ciertamente deseable.

Por otro lado, [Kripke \(1975\)](#) señala que la asignación sintáctica de niveles de verdad a una oración, el hecho de que deba usarse un predicado con el índice correcto según el caso, sabiéndolo de antemano, imposibilita las adscripciones ciegas de verdad. Se carece por tanto de predicados de verdad para decir que ciertos enunciados, de los cuales no se conoce su formulación explícita, son verdaderos. El ejemplo famoso de Kripke concierne a Jones y a Nixon. Supongase que Jones afirma el siguiente y único enunciado sobre Watergate:

La mayor parte de las afirmaciones de Nixon sobre Watergate son falsas.

mientras que Nixon dice la misma cantidad de enunciados verdaderos que falsos sobre Watergate más el que sigue:

Todo lo que Jones dice sobre Watergate es verdadero.

Si bien no hay nada intrínsecamente problemático con cada una de estas oraciones, puestas en conjunto llevan a una contradicción en el marco de la lógica clásica. Ambas son perfectamente gramaticales y es solo a causa de las particularmente desfavorables circunstancias empíricas que juntas llevan a paradojas. Para evitar este problema Jones o Nixon deberían haber averiguado primero cuáles son las aserciones hechas por

el otro acerca de Watergate, asignado niveles a cada uno de los predicados de verdad que ocurrieran en ellas y luego elegir en nivel apropiado de su propia asercion, lo cual es absurdo. De acuerdo con [Kripke \(1975, p. 696\)](#),

This means that in some sense a statement should be allowed to seek its own level, high enough to say what it intends to say. It should not have an intrinsic level fixed in advance, as in the Tarski hierarchy.

Kripke no fue el primero en criticar el enfoque tarskiano, pero sí, simultáneamente con [Robert L. Martin & Peter W. Woodru \(1975\)](#), en ofrecer una teoría formal en la cual las dificultades señaladas fueran superadas. En ambos casos, se ofrecen modelos para un lenguaje formal semanticamente cerrado, que contiene su propio predicado veritativo, esto es, un unico predicado aplicable tambien a oraciones que lo contienen. Me centro exclusivamente en las teorías de Kripke que, a pesar de haber sido influenciadas por la de Martin & Woodru (cf. [Kripke \(1975, p. 699\)](#)), la incluyen y son mejor conocidas.

La estrategia kripkeana para evadir las hipotesis del teorema de Tarski (teorema [2.9](#)) consiste, por un lado, en abandonar la logica clasica por una paracompleta, donde falla el principio de Tercero Excluido, permitiendo que ciertas oraciones de los lenguajes formales para los que se desea definir verdad, v.g., las paradójicas, no sean verdaderas ni falsas sino que carezcan de valor de verdad. Por otro lado, Kripke modifica levemente la Convencion T (definicion [2.1](#)): en lugar de satisfacer las instancias del [Esquema-T](#) como condicion de adecuacion material, valida las reglas

$$\frac{}{T p'q} \quad (T\text{-Intro}) \qquad \frac{T p'q}{'} \quad (T\text{-Elim})$$

tomadas en conjunto, para la introduccion y eliminacion del predicado veritativo. La concepcion kripkeana de la verdad puede resumirse en la siguiente afirmacion: "We may say that we are entitled to assert (or deny) of a sentence that it is true precisely under the circumstances when we can assert (or deny) the sentence itself." ([Kripke, 1975, p. 701](#))

Las reglas [T-Intro](#) y [T-Elim](#), tambien llamadas 'reglas de captura' y 'liberacion', son equivalentes al [Esquema-T](#) en logica clasica pero, como se vera mas adelante en esta seccion, no lo implican en ciertas logicas paracompletas adoptadas por Kripke, como k_3 y b_3 (definiciones [3.3](#) y [3.5](#)), donde falla el metateorema de la deduccion.¹ Tanto el [Esquema-T](#) como las reglas [T-Intro](#) y [T-Elim](#) son denominados 'principios de verdad'

¹Esto es, no siempre es el caso de que, si k_3 , entonces k_3 '!', y lo mismo para b_3 .

transparente' o, directamente, 'de transparencia', porque permiten 'ver' la oración de la cual se predica verdad 'a través' del predicado. También suele referirse a ellos como 'principios de equivalencia': en ambos casos, la oración de la cual se predica verdad y su predicación son equivalentes, aunque en un caso esta equivalencia se expresa en el lenguaje objeto mientras que en el otro se expresa en el metalenguaje. A diferencia del enfoque tarskiano, el kripkeano permite instancias de estos principios generadas por oraciones en las cuales T mismo ocurre y sus teorías son, en consecuencia, teorías libres de tipos.

En lo que resta de esta sección presento en detalle dos de los tres caminos que Kripke recorrió para ilustrar esta idea, según las necesidades posteriores de la presente investigación. En ambos casos, el objetivo de Kripke es extender el modelo estándar del lenguaje de la aritmética a un modelo del lenguaje ampliado mediante un predicado monádico T para expresar verdad que satisfaga las reglas [T-Intro](#) y [T-Elim](#). El resultado es una familia de modelos para este lenguaje ampliado, razón por la cual se llama a las teorías kripkeanas 'semánticas', en contraste con abordajes axiomáticos.²

En [3.1.1](#) presento las dos lógicas paracompletas sobre las cuales Kripke llevo a cabo cada una de las teorías semánticas que introduzco y cuyas propiedades exploro en [3.1.2](#). Posteriormente, en [3.1.3](#), presento también dos teorías axiomáticas inspiradas en las teorías semánticas de Kripke: kf, por Kripke-Feferman, una axiomatización más o menos directa de una de las teorías semánticas de Kripke en lógica clásica, en [3.1.3.1](#), y putb, por Positive Uniform Tarski Biconditionals, una subteoría de kf dada exclusivamente por instancias del [Esquema-T Uniforme](#), en [3.1.3.2](#).

3.1.1. Lógicas paracompletas

El objetivo es extender N^+ , el modelo estándar de L_{pa}^+ (que hace las veces de modelo base, provee los 'hechos no semánticos' del mundo y, a la vez, provee la teoría de la sintaxis necesaria para tener un predicado de verdad en el lenguaje, i.e., nombres para las expresiones y ciertos predicados sintácticos como los introducidos en el apartado [1.2.3](#)), a un modelo de todo L_T en el cual una oración $\phi \in L_T$ pertenezca a la extensión de T si y solo si ϕ es verdadera en el modelo.³ Como, por el teorema de indecibilidad de Tarski, esto no es posible para modelos y esquemas de valuación clásicos, Kripke elige modelos no clásicos que llama 'parciales' o 'paracompletos' y esquemas de valuación no clásicos en su lugar.

²Este calificativo no debe confundirse con el adjetivo homónimo que Tarski utiliza para remarcar que sus teorías son de corte correspondentista, i.e., entienden a la verdad como un tipo de relación entre el lenguaje y el mundo, de alguna manera. Véase la introducción al capítulo [2](#).

³A *decir verdad*, [Kripke \(1975\)](#) no es tan específico al introducir el lenguaje para el cual de ne verdad, pero este detalle no afecta en nada el espíritu de la presentación.

Definicion 3.1 (Modelos no clasicos). M es un modelo no clasico de un lenguaje de primer orden L sii se comporta como un modelo clasico con respecto a las constantes de individuo y s mbolos de funcion de L y el predicado de identidad, mientras que, para cada predicado n-adico R del lenguaje distinto de =, $R^M = hR^{M+}; R^M i$, donde $R^{M+}; R^M \quad jMj^n$ y $R^M \setminus R^{M+} = ?$.

Mientras que en un modelo clasico M cada predicado R de L recibe unicamente un subconjunto R^M de n-tuplas del dominio como su interpretacion, que usualmente se interpreta como la extension de R, mientras que su complemento sobre jMj^n es la antiextension, en los modelos no clasicos recibe un par ordenado de subconjuntos de jMj^n , $hR^{M+}; R^M i$. Desde el punto de vista de las logicas paracompletas, R^{M+} es la extension de R y R^M su antiextension, que puede no ser el complemento de R^{M+} sobre conjunto de n-tuplas de elementos del dominio.⁴ Por esta razon la logica paracompleta ve los modelos no clasicos como modelos parciales. Si bien no es posible que R^{M+} y R^M tengan miembros en comun, i.e., que un objeto o tupla de objetos del dominio del modelo pertenezca tanto a la extension como a la antiextension de un predicado, su union podra no agotar el dominio; podra haber objetos o tuplas de objetos que no pertenezcan ni a la extension ni a la antiextension de un predicado.

Dado un modelo parcial M, no existe un unico modo de asignar valores de verdad a las oraciones de L. Kripke considera tres esquemas de valuacion diferentes: Kleene fuerte, Kleene debil y el esquema supervaluacionista, cada uno de los cuales va a dar lugar a diferentes modelos de verdad transparente para L_T . Por esta razon se puede decir que Kripke no propone una unica teor a semantica, una unica familia de modelos, sino un metodo general para obtener teor as semanticas de la verdad. En funcion de esta investigacion, solo presento las familias basadas en los esquemas de Kleene fuerte y Kleene debil.⁵

Definicion 3.2 (Kleene fuerte). El esquema de valuacion Kleene fuerte sk para un lenguaje de primer orden L es el conjunto de funciones valuacion $v_{sk}^M : L \rightarrow \{1, 0, \frac{1}{2}\}$ para cada modelo no clasico M de L tales que⁶

$$v_{sk}^M(t = s) = \begin{cases} 1 & \text{si } t^M = s^M \\ 0 & \text{en caso contrario} \end{cases}$$

⁴En otros abordajes paracompletos tambien la relacion de identidad puede contar con una extension y una antiextension no exhaustivas. En esta investigacion bloqueo esta posibilidad para simplificar la presentacion.

⁵Para una presentacion mas detallada de los tres esquemas de valuacion y de las teor as de Kripke, vease Paula Teijeiro & Damian Szmuc (2014).

⁶Utilizo una clausula sustitucional para los cuantificadores porque en los modelos con los cuales trabajo, i.e., extensiones de N^* , todos los objetos del dominio tienen un nombre en L_T . Estrictamente hablando deber a haber introducido asignaciones en este caso, pero lo evito para facilitar la lectura.

$$v_{sk}^M(Rt_1; \dots; t_n) = \begin{cases} 1 & \text{si } t_1^M, \dots, t_n^M \in R^{M+} \\ > 0 & \text{si } t_1^M, \dots, t_n^M \in R^M \\ < 1 & \text{en caso contrario} \end{cases}$$

- $v_{sk}^M(') = 1$ $v_{sk}^M(')$
- $v_{sk}^M('!) = \max(1 - v_{sk}^M('); v_{sk}^M('))$
- $v_{sk}^M(8v') = \min v_{sk}^M(' [t=x]) : t \text{ es un termino cerrado}$

La logica basada en modelos no clasicos y el esquema de valuacion sk que interpreta en valor 1 como verdadero, 0 como falso y $\frac{1}{2}$ como ni verdadero ni falso se conoce como k3.

Definicion 3.3 (k3). Si L es un lenguaje de primer orden, L es un conjunto de oraciones y ' ∈ L una oracion, ' es consecuencia logica de en k3 (k3 ') sii no existe un modelo no clasico M de L tal que $v_{sk}^M(') = 1$ (o, alternativamente, $M \models_{k3} ')$ y $v_{sk}^M(') \neq 1$ (o $M \not\models_{k3} ')$.

Luego, si una oracion no es ni verdadera ni falsa en un modelo, tampoco lo sera su negacion; y si un condicional tiene como antecedente y como consecuente oraciones cuyo valor de verdad no es clasico, el valor del condicional tampoco es clasico. Una interpretacion usualmente considerada natural del valor de verdad $\frac{1}{2}$ en k3 es la indeterminacion o falta de informacion. Kripke lo entiende como un tipo especial de vac o de verdad.

Una de las consecuencias mas salientes de adoptar k3 es que diversas reglas y principios validos en logica clasica resultan invalidos. Por ejemplo, como se quer a, algunas instancias del principio de Tercero Excluido

$$' \vee ' : \quad \text{(Tercero Excluido)}$$

reciben valor $\frac{1}{2}$ en algunos modelos parciales sobre sk y, por ende, no son verdaderas. Pero la falla de [Tercero Excluido](#) trae tambien consigo ciertas consecuencias que pueden resultar indeseables. Por ejemplo, la regla de introduccion del condicional

$$; '= () = '! : \quad (II)$$

o metateorema de la deduccion, ya no vale en k3. V.g., ' k3 ', porque todo modelo en el cual ' es verdadera es tambien un modelo en el cual ' es verdadera, pero, a la vez, $2k3 ' !'$, i.e., el condicional de k3 no satisface identidad

'!

(Identidad)

porque si en una valuacion ' recibe valor $\frac{1}{2}$, de acuerdo con la definicion 3.2, el valor del condicional '!' tambien es $\frac{1}{2}$, i.e., no es verdadero de acuerdo con esa valuacion.

Definicion 3.4 (Kleene debil). El esquema de valuacion Kleene debil w_k para un lenguaje de primer orden L es el conjunto de funciones valuacion $v_{w_k}^M : L \rightarrow \{0, \frac{1}{2}, 1\}$ para cada modelo no clasico M de L tales que⁷

$$\begin{aligned}
 v_{w_k}^M(t = s) &= 1 \text{ si } t^M = s^M \\
 &< 0 \text{ en caso contrario} \\
 &: 1 \text{ si } \langle t_1^M, \dots, t_n^M \rangle \in R^{M+} \\
 v_{w_k}^M(\langle t_1, \dots, t_n \rangle) &= \begin{cases} > 0 & \text{si } \langle t_1^M, \dots, t_n^M \rangle \in R^M \\ > \\ < 1 \\ > \frac{1}{2} & \text{en caso contrario} \\ \vdots \end{cases} \\
 \begin{cases} v_{w_k}^M(\neg \phi) = 1 - v_{w_k}^M(\phi) \\ v_{w_k}^M(\phi \wedge \psi) = \min\{v_{w_k}^M(\phi), v_{w_k}^M(\psi)\} \\ v_{w_k}^M(\phi \vee \psi) = \max\{v_{w_k}^M(\phi), v_{w_k}^M(\psi)\} \\ v_{w_k}^M(\phi \rightarrow \psi) = \min\{1, \max\{1 - v_{w_k}^M(\phi), v_{w_k}^M(\psi)\}\} \\ v_{w_k}^M(\forall x \phi) = \inf\{v_{w_k}^M(\phi[t=x]) : t \text{ es un termino cerrado}\} \\ v_{w_k}^M(\exists x \phi) = \sup\{v_{w_k}^M(\phi[t=x]) : t \text{ es un termino cerrado}\} \end{cases} \text{ si } v_{w_k}^M(\phi) \in \{0, \frac{1}{2}, 1\} \\
 &: \text{en caso contrario}
 \end{aligned}$$

La logica cuya nocion de consecuencia esta definida en terminos de modelos no clasicos y el esquema de valuacion w_k e interpreta los valores 1; 0 y $\frac{1}{2}$ como k_3 se conoce como b_3 .

Definicion 3.5 (b_3). Si L es un lenguaje de primer orden, L es un conjunto de oraciones y ' \in L una oracion, ' es consecuencia logica de Γ en b_3 ($\Gamma \vdash_{b_3} \phi$) sii no existe un modelo no clasico M de L tal que $v_{w_k}^M(\psi) = 1$ ($\psi \in \Gamma$) y $v_{w_k}^M(\phi) \neq 1$ ($\phi \in L$).

Si bien la negacion se comporta igual que en k_3 , el condicional y el cuantificador universal son un tanto diferentes: basta que una formula contenga una subformula sin valor de verdad para que la formula original tambien carezca de valor de verdad. La interpretacion mas natural de $\frac{1}{2}$ en b_3 es quizas 'es un sinsentido', un modo alternativo a la falta de informacion de tener un vacuo de verdad.

⁷Al igual que en k_3 , me tomo la licencia de emplear una clausula sustitucional para los cuantificadores para no complicar la presentacion.

En b3 el [Tercero Excluido](#), [!!](#) e [Identidad](#) tampoco valen. Aun mas, falla tambien la regla de introduccion del cuantificador existencial:

$$\frac{'[t=v]}{9v'} \tag{I9}$$

De la definicion [3.4](#) y la definibilidad del cuantificador existencial en terminos del universal y la negacion ($9v' := \neg 8v'$) se sigue que una formula de la forma $9v'$ adquiere valor 1 en un modelo M de acuerdo con wk si y solo si existe un termino t tal que $v_{wk}^M('[t=v]) = 1$ y para ningun termino t sucede que $v_{wk}^M('[t=v]) = \frac{1}{2}$. Luego, si en un modelo parcial M $v_{wk}^M('[t=v]) = 1$ y $v_{wk}^M('[s=v]) = \frac{1}{2}$, $v_{wk}^M(9v') = 1$; la regla [I9](#) no es valida.

3.1.2. Modelos de punto jo

Los modelos no clasicos abren en principio la posibilidad lidiar con las oraciones problematicas de L_T no asignandoles ningun valor de verdad clasico, pero no indican cuales son estas oraciones ni si es posible un modelo de verdad transparente, i.e., donde valgan [T-Intro](#) y [T-Elim](#). Es preciso seleccionar al menos un modelo en el cual este sea el caso.

Como L_T contiene exclusivamente un predicado diferente de $=, T$, este es el unico predicado que admite una interpretacion parcial en un modelo parcial de L_T . Luego, solo formulas que contengan T podran obtener un valor de verdad no clasico de acuerdo con los esquemas sk y wk : la no clasicidad esta restringida a formulas que contienen el predicado veritativo. En efecto, en los modelos que Kripke considera, el vocabulario de L_T que pertenece a L_{pa}^+ es interpretado de acuerdo con N^+ . Si $!$, sea hN^+ ; i el modelo parcial que extiende N^+ a L_T asignando a T como su extension y el conjunto de formulas tales que su negacion pertenece a $!$ como su antiextension.

Para ambos esquemas de valuacion sk y wk Kripke muestra que algunos $!$ son tales que $\#(') = 2$ si y solo si $'$ recibe valor 1 en hN^+ ; i , i.e., que existen modelos hN^+ ; i en los cuales valen las reglas [T-Intro](#) y [T-Elim](#). Para probar la existencia de tales conjuntos introduce operadores ev llamados 'salto', para cada esquema de valuacion ev , v.g., sk y wk .

Definicion 3.6 (Operadores salto). Si ev es un esquema de valuacion para L_T , ev es una funcion que mapea subconjuntos de $!$ en subconjuntos de $!$ de modo tal que, si $!$,

$$ev() := \{ f \in \#(') \mid \exists g \in \#(') : v_{ev}^{hN^+; i}(') = 1 \} \tag{3.1}$$

Los conjuntos que Kripke quiere obtener son puntos bajos del operador salto.

Definicion 3.7 (Modelos de punto jo). \mathcal{M}^+ ; i es un modelo de punto jo de L_T sobre el esquema de valuacion \mathcal{V} sii $v_{\mathcal{V}}(p) = 1$.

Proposicion 3.8 (Kripke). \mathcal{M}^+ ; i es un modelo de punto jo de L_T sobre \mathcal{V} sii, para toda oracion $\phi \in L_T$,

$$v_{\mathcal{V}}^{\mathcal{M}^+; i}(T \phi) = 1, v_{\mathcal{V}}^{\mathcal{M}^+; i}(\perp) = 0:$$

Por esta razon el enfoque kripkeano se conoce como 'concepcion de la verdad como un punto jo'.

Corolario 3.9. Las reglas [T-Intro](#) y [T-Elim](#) preservan verdad en todos los modelos de punto jo de L_T sobre algun esquema de valuacion.

Esto hace de los modelos de punto jo modelos de verdad transparente, y de la teoria kripkeana a la cual pertenecen una teoria de la verdad transparente, porque su predicado de verdad lo es.

La existencia de puntos jos de s_k y w_k queda garantizada por el hecho de que ambas son funciones monotonamente crecientes: si $\alpha < \beta$, $s_k(\alpha) \subseteq s_k(\beta)$ y $w_k(\alpha) \subseteq w_k(\beta)$.⁸ Luego, si, para algun α , $s_k(\alpha) \subseteq w_k(\alpha)$, la secuencia dada por todas las iteraciones de $s_k \circ w_k$ sobre α , i.e.,

$$\alpha; s_k(\alpha); s_k^2(\alpha); \dots; s_k(\alpha); \dots; i;$$

$$(\alpha; w_k(\alpha); w_k^2(\alpha); \dots; w_k(\alpha); \dots; i)$$

donde, para cada ordinal α , $s_k(\alpha) \subseteq w_k(\alpha)$ es el conjunto que resulta de iterar veces el operador $s_k \circ w_k$ sobre α , ser a monotonamente creciente. Como solo hay una cantidad enumerable de oraciones de L_T , para algun α , $s_k(\alpha) = s_k^{+1}(\alpha)$ ($w_k(\alpha) = w_k^{+1}(\alpha)$), esto es, la secuencia alcanza a un punto jo.

Teorema 3.10 (Kripke). Existen conjuntos α tales que $s_k(\alpha) = \alpha$; y otros conjuntos β tales que $w_k(\beta) = \beta$.⁹

Naturalmente, la existencia de puntos jos para los operadores salto permanece sujeta a la existencia de conjuntos α tales que $s_k(\alpha) \subseteq w_k(\alpha)$. El conjunto mas pequeno que satisface estas condiciones es \perp . Los puntos jos de las secuencias

$$\alpha; s_k(\alpha); s_k^2(\alpha); \dots; s_k(\alpha); \dots; i \tag{3.2}$$

⁸Vease [Kripke \(1975, p. 703\)](#) para una prueba.

⁹Cf. [Kripke \(1975\)](#).

y

$$h^?; w_k(?)^2; \dots; w_k(?); \dots; i \tag{3.3}$$

se conocen como 'puntos fijos mínimos' de s_k y w_k .

Estas secuencias pueden entenderse como dando lugar a secuencias de modelos parciales $hN^+; s_k(?)^i (hN^+; w_k(?)^i)$ de L_T . Al inicio de cada una de estas secuencias tanto la extensión como la antiextensión del predicado de verdad están vacías. En el segundo lugar en las secuencias se produce la primera aplicación de s_k (w_k), e ingresan a la extensión de T todas las oraciones que reciben valor 1 en $hN^+; ?^i$ de acuerdo con k_3 (b_3) (y a la antiextensión las que reciben valor 0), entre ellas, todas las oraciones de L_{pa}^+ que son verdaderas en N^+ , v.g., $0 = 0$. En el paso siguiente, se suman aquellas que son verdaderas en $hN^+; s_k(?)^i (hN^+; w_k(?)^i)$ de acuerdo con k_3 (b_3), por ejemplo, $T p0 = 0q$. En el próximo paso ingresa, v.g., $T pT p0 = 0qq$, etc.

Como señala [Kripke \(1975, p. 701\)](#),

There is no reason to suppose that all statements involving 'true' will become decided in this way, but most will. Indeed, our suggestion is that the "grounded" sentences can be characterized as those which eventually get a truth value in this process.

Kripke adjudica el concepto de groundedness o fundación a [Hans G. Herzberger \(1970\)](#).¹⁰ Intuitivamente, las oraciones fundadas son aquellas cuya verdad o falsedad depende exclusivamente de la verdad o falsedad de oraciones que no contienen predicados semánticos, directa o indirectamente. Efectivamente, en los puntos fijos mínimos solo aquellas oraciones de L_{pa}^+ que, o predicán verdad de oraciones de L_{pa}^+ , o de estas últimas, o de estas últimas, etc., ingresan a la extensión o a la antiextensión del predicado veritativo, adquieren un valor de verdad clásico. Los modelos resultantes son, por tanto, modelos de una verdad superveniente sobre los 'hechos' no semánticos.

Por su parte, la oración del mentiroso $:T I$, que se obtiene aplicando el lema de diagonalización fuerte (teorema [1.8](#)) a $:T x$, esto es,

$$N^+ I = p:T Iq \tag{3.4}$$

jamás entrar ni en la extensión ni en la antiextensión de T sino que recibirá siempre el valor $\frac{1}{2}$ en cada modelo $hN^+; s_k(?)^i (hN^+; w_k(?)^i)$; y lo mismo vale para la oración del honesto, que se obtiene también aplicando el lema de diagonalización fuerte ahora a $T x$, i.e.,

$$N^+ h = pT hq:$$

¹⁰Quien lo llama
`groundlesness'.

Las oraciones del mentiroso y del honesto son consideradas infundadas (ungrounded). Su valor de verdad depende, no del de las oraciones de L_{pa} , sino de sí mismas.

Definición 3.11 (ev-fundación). ' L_T está ev-fundada si ' pertenece al punto mínimo de ev .

El punto mínimo parece ser la materialización perfecta de la metáfora kripkeana del nivel de una oración y de cómo debe permitirse que cada oración encuentre su propio nivel. Por 'nivel de una oración' puede entenderse el primer ordinal en la secuencia (3.2) o (3.3) en el cual ingresa a la extensión o a la antiextensión del predicado veritativo. Y este nivel dependerá enteramente del modelo que se está expandiendo (en este caso $hN^+; ?i$, y no de la estructura sintáctica o gramatical de la oración. Como sostiene [Michael Kremer \(1988, p. 237\)](#),

Associated with this explication of the notion of the "level" of a sentence is the idea that Kripke's construction of the minimal fixed point has provided us with an alternative hierarchy of languages to that provided by the "orthodox approach". The Tarskian hierarchy of object-language, meta-language, meta-meta-language, . . . is to be replaced by the ordinal sequence $[hN^+; ?i; hN^+; sk(?i); \dots; hN^+; sk(?i)]$.

En las clásicas palabras del mismo [Kripke \(1975, p. 80\)](#), un tanto más pesimistas, "The ghost of the Tarski hierarchy is still with us."

No obstante, los puntos fijos generados por τ para cada operador, los respectivos puntos mínimos, no son los únicos puntos fijos de sk ni de w_k , sino que existen otros, sobre los cuales no voy a entrar en detalle (cf. [Kripke \(1975\)](#)). Naturalmente, en estos puntos fijos 'más grandes' no solo oraciones fundadas entrarán en la extensión o antiextensión del predicado veritativo. Pero ninguna de las que ingresen será paradójica, como el mentiroso, porque las expresiones paradójicas son precisamente aquellas que bloquean la posibilidad de alcanzar puntos fijos, aquellas que no pueden ser ni verdaderas ni falsas. Si se comenzara una secuencia de iteraciones de un operador salto con un conjunto que tuviera a \perp como miembro, v.g., $f:T \text{ lg}$, jamás se alcanzará a un punto fijo, porque \perp pasará de la extensión a la antiextensión de T o viceversa en cada paso; la secuencia no será monotona, ya que $f:T \text{ lg}^* sk(f:T \text{ lg}) (w_k(f:T \text{ lg}))$. No hay un modelo de punto fijo en el cual el mentiroso adquiera un valor de verdad clásico. Aquellas oraciones que no ingresan ni a la extensión ni a la antiextensión del predicado de verdad en ningún modelo de punto fijo dado por un esquema de valuación ev suelen identificarse como ev -paradójicas.

Si bien [Kripke \(1975, p. 712, nota 29\)](#) admite que el punto jo m nimo \is singled out as natural in many respects" y ciertos ejemplos motivacionales al comienzo de su artículo favorecen el modelo de punto jo m nimo, todos los modelos de punto jo satisfacen la condición de adecuación material kripkeana, i.e., el principio de transparencia dado por las reglas [T-Intro](#) y [T-Elim](#). Aun más, de acuerdo con [Kremer \(1988, p. 241\)](#),

the fixed point conception of truth is at the heart of Kripke's paper. Other notions|groundedness, the ordinal construction of the minimal fixed point, the intrinsic fixed points|are best seen as products of his investigation into the existence and properties of fixed points, which he thinks of as languages containing their own truth predicates.

Las teorías kripkeanas de la verdad no están dadas por el modelo de punto jo m nimo exclusivamente sino por la clase de todos los modelos de punto jo para cada operador salto.

Como anticipo, el [Esquema-T](#) no es verdadero en ninguno de los modelos de punto jo. Esto se debe a la falla de [Identidad](#). Si bien ninguna de sus instancias resulta falsa, algunas adquieren valor $\frac{1}{2}$ en todos estos modelos. V.g., como la oración del mentiroso $:\text{T I}$ en [\(3.4\)](#) jamás pertenece ni a la extensión ni a la antiextensión del predicado de verdad, tanto ella como su predicación de verdad, T p:T I , adquieren siempre valor $\frac{1}{2}$. Por las condiciones de evaluación de condicionales en las definiciones [3.2](#) y [3.4](#), lo mismo sucede con las fórmulas $\text{T p:T I} \rightarrow \text{T I}$ y $:\text{T I} \rightarrow \text{T p:T I}$, tanto en sk como wk . Luego, por la interdefinibilidad de las conectivas proposicionales, el bicondicional

$$\text{T p:T I} \leftrightarrow \text{T I};$$

instancia del [Esquema-T](#) dada por $:\text{T I}$, vale $\frac{1}{2}$ en todos los modelos de punto jo de Kripke. Sin embargo, en cada modelo de punto jo hN^+ ; i las instancias del [Esquema-T](#) dadas por oraciones que reciben algún valor clásico de verdad s son verdaderas, esto es,

$$v_{\text{sk}}^{\text{hN}^+; i}(\text{T p}'\text{q} \leftrightarrow) = 1, \text{ ' } 2 \quad \text{o: ' } 2 \quad :$$

Y otro tanto para wk .

Las teorías paracompletas de la verdad que Hartry Field desarrolla en [\(2002, 2003, 2005, 2007, 2008\)](#) [foco de atención de buena parte de las publicaciones actuales en teorías no clásicas de la verdad] tienen como uno de sus objetivos principales superar esta dificultad, enriquecer el lenguaje L_T con un condicional que valide identidad con el cual formular el [Esquema-T](#); que permita definir un nuevo bicondicional que resulte

verdadero tambien cuando las formulas que se encuentran a derecha e izquierda reciban

ambas valor de verdad $\frac{1}{2}$ y, de este modo, cumplir con la Convención T de Tarski. El resultado es una serie de modelos altamente complejos pero que logran acomodar el [Esquema-T](#), las reglas [T-Intro](#) y [T-Elim](#) y un tercer principio de transparencia conocido como 'principio de intersustitutividad' o 'principio de transparencia', que Field y J.C. Beall preeren por sobre los otros:¹¹

$$\frac{\dots' \dots}{\dots T p'q \dots} \quad (\text{Intersustitutividad})$$

donde la doble línea indica que la regla puede aplicarse tomando como premisa la fórmula que se encuentra por arriba y la que está por debajo como conclusión o viceversa, y $\dots' \dots$ indica que ' ocurre como subfórmula. Informalmente, si dos oraciones B y C son idénticas excepto en que (en algún contexto no opaco) B contiene la subfórmula A mientras que C contiene 'A es verdadera' en su lugar, entonces B implica C y viceversa (cf. [Field \(2008, p. 12\)](#)). En virtud de [T-Intro](#) y [T-Elim](#), [Intersustitutividad](#) resulta válido también en los modelos de punto jo de Kripke.

Proposición 3.12. Las reglas que constituyen el principio de [Intersustitutividad](#) preservan verdad en todos los modelos de punto jo sobre algún esquema de valuación.

Demostración. Pruebo que, si M es un modelo de punto jo sobre ev (sk o wk), $v_{ev}^M(') = v_{ev}^M(T p'q)$ para toda oración $' \in L_T$. Como todas las conectivas de L_T aplicables a oraciones son composicionales, esto es suficiente para establecer la corrección de [Intersustitutividad](#) en M.

La proposición [3.8](#) garantiza que $v_{ev}^M(') = v_{ev}^M(T p'q)$ si $v_{ev}^M(') = 1$. Sea $M = \langle hN^+, i \rangle$; i. Si $v_{ev}^M(') = 0$, como $=_{ev}(\)$, la definición [3.1](#) garantiza que $' \in f \in L_T : \dots' \dots g$, esto es, que ' pertenece a la antiextensión de T en hN^+ ; i y, por ende, $v_{ev}^M(T p'q) = 0$. Finalmente, si $v_{ev}^M(') = \frac{1}{2}$, entonces, nuevamente por la definición [3.1](#), ' no pertenece ni a la extensión ni a la antiextensión de T en el modelo y, en consecuencia, $v_{ev}^M(T p'q) = \frac{1}{2}$ también. □

Si bien las lógicas k_3 y b_3 son completamente axiomatizables (cf. [Kremer \(1988\)](#)), los conjuntos de oraciones que resultan verdaderas en cada modelo de punto jo sobre cada esquema no son axiomatizables. Por un lado, presuponen el modelo estándar de L_{pa}^+, N^+ , esto es, las oraciones verdaderas en N^+ son verdaderas en todos los modelos

¹¹Para una exposición detallada, véase también [Lucas Rosenblatt & Federico Pailos \(2014\)](#).

de punto jo, lo cual introduce complejidad ¹₁ (vease el apartado [1.2.5](#)).¹² Si una teoría consistente axiomatizase estos conjuntos de modo acabado, debería ser capaz de probar todas las verdades de la aritmética, lo cual no es posible por los teoremas de incompletitud de Gödel (teoremas [1.12](#) y [1.13](#)). Pero, además, el predicado veritativo agrega asimismo complejidad: las extensiones de T en los modelos de punto jo de Kripke son ¹₁. Sin embargo, ciertas axiomatizaciones parciales son posibles, como se verá en el apartado que resta de esta sección.

Las teorías kripkeanas de la verdad no solo inspiraron a nuevas teorías semánticas paracompletas y versiones axiomáticas sino que inspiraron en prácticamente todas las investigaciones formales sobre la verdad en el ámbito de la lógica clásica. Casos obvios son la teoría de la dependencia de Leitgeb, que introduzco en la sección [3.2](#), la teoría de la revisión (cf. [Herzberger \(1982a,b\)](#), [Anil Gupta \(1982\)](#), [Nuel D. Belnap \(1982\)](#) y [Gupta & Belnap \(1993\)](#)), para la cual lamentablemente no hay lugar en esta tesis,¹³ y quizás también las teorías de la verdad dialécticas que ocupan la sección [3.4](#).

3.1.3. Contrapartes axiomáticas

3.1.3.1. Kripke-Feferman

La teoría kf que presento en este apartado se debe a [William Reinhardt \(1986\)](#) y a [Feferman \(1991\)](#).¹⁴ Si bien suele decirse que es una axiomatización de la clase de modelos de punto jo sobre sk en k₃, esto no es estrictamente cierto. Lo naturalmente esperable de una axiomatización de una familia de modelos parciales bajo k₃ es que se lleve a cabo sobre un cálculo correcto y completo con respecto a k₃. Sin embargo, kf está formulada sobre un cálculo clásico. Como consecuencia, no es correcta con respecto a la familia de modelos de punto jo en k₃, v.g., kf prueba que $\vdash \neg \text{Tr} \text{Tr}$, pero esta oración no recibe valor 1 en ninguno de estos modelos,¹⁵ aunque la identidad entre $\text{Tr} p$ y p también vale en kf, por ser una extensión de q^+ , donde

$$q^+ \text{ ` } \vdash p \text{ :Tr } lq \tag{3.5}$$

por el lema diagonal fuerte aplicado a $\text{Tr} \text{Tr}$.

¹² Esto es, el conjunto de oraciones de L^+_{pa} verdaderas en N^+ solo puede ser de nido mediante una fórmula del lenguaje de la aritmética de segundo orden que sea ¹₁, i.e., conste de una fórmula de primer orden precedida por un cuantificador universal o existencial de segundo orden. Esto queda claro en parte a través del teorema de la indecidibilidad de la verdad de Tarski (teorema [2.9](#)), que muestra que ninguna fórmula de la jerarquía aritmética, esto es, de L^+_{pa} , es capaz de definir el conjunto de verdades de este mismo lenguaje.

¹³ Para una exposición detallada vease [Lavinia Picollo & Natalia Buacar \(2014\)](#).

¹⁴ Feferman la denomina 'Ref(pa)'.
¹⁵ Para una axiomatización de la familia de modelos de punto jo sobre sk en k₃, pkf (por Paracomplete Kripke-Feferman), vease [Halbach & Leon Horsten \(2006\)](#).

kf es una axiomatizacion, no de la familia de modelos de punto jo sobre sk en k3 sino de su clausura. La clausura de un modelo parcial M es el modelo clasico M^0 que asigna a las letras de predicado del lenguaje la misma extension que M. La diferencia entre M y M^0 es que M asigna a cada predicado una antiextension que puede no ser el complemento de la extension sobre el dominio del modelo, mientras que en M^0 este siempre es el caso; M^0 pone en la antiextension de cada predicado todo lo que M pone a mas aquello que no estaba ni en la extension ni en la antiextension. Supongase, por ejemplo, que en un modelo parcial $hN^+; i$ de punto jo ' no pertenece ni a la extension ni a la antiextension de T. Luego, $v_{sk}^{hN^+; i}(T p'q) \neq 1$ y $v_{sk}^{hN^+; i}(\neg T p'q) \neq 1$. Sin embargo, en la clausura de $hN^+; i$, ' pertenecera a la antiextension de T, con lo cual $\neg T p'q$ sera verdadera. Notese que la clausura de un modelo de punto jo no es ya ella misma un modelo de punto jo, porque eso es imposible por el teorema de indecibilidad de la verdad de Tarski.

Definicion 3.13 (kf). Sea kf_{L_T} la teoria que extiende pat con los siguientes axiomas:

$$kf1 \quad \forall t, s (T(t=s) \rightarrow t=s)$$

$$kf2 \quad \forall x (\text{Sent}_{L_T}(x) \rightarrow (T \neg x \rightarrow T x))$$

$$kf3 \quad \forall x, y (\text{Sent}_{L_T}(x \rightarrow y) \rightarrow (T(x \rightarrow y) \rightarrow (T x \rightarrow T y)))$$

$$kf4 \quad \forall x, y (\text{Sent}_{L_T}(x \wedge y) \rightarrow (T(x \wedge y) \rightarrow (T x \wedge T y)))$$

$$kf5 \quad \forall x, v (\text{Sent}_{L_T}(\exists v x) \rightarrow (T \exists v x \rightarrow \exists t T x(t=v)))$$

$$kf6 \quad \forall x, v (\text{Sent}_{L_T}(\forall v x) \rightarrow (T \forall v x \rightarrow \forall t T x(t=v)))$$

$$kf7 \quad \forall t (\text{Sent}_{L_T}(t) \rightarrow (T T t \rightarrow T t))$$

$$kf8 \quad \forall t (\text{Sent}_{L_T}(t) \rightarrow (T \neg T t \rightarrow T \neg t))$$

$$\text{cons} \quad \forall x (\text{Sent}_{L_T}(x) \rightarrow (T x \wedge T \neg x))$$

El primer axioma es comun a las axiomatizaciones tarskianas. Si bien kf2 valida una version interna de la ley de Doble Negacion, al no haber un axioma que establezca la conmutatividad del predicado veritativo con la negacion, como en el caso de, v.g., $ct(pa^+)$ o $rt_<$, es preciso desglosar los axiomas para el condicional y el cuantificador universal en un principio positivo y uno negativo, como sucede en los axiomas kf3 kf6. Los axiomas kf7 y kf8 establecen la iterabilidad del predicado de verdad, mientras que cons bloquea la posibilidad de que una oracion y su negacion sean ambas verdaderas.

Teorema 3.14 ((Feferman, Reinhardt)). $hN^+; i$ es la clausura de un modelo de punto jo bajo k3 sii $hN^+; i$ kf. [16](#)

¹⁶Vease [Feferman \(1991\)](#) para una prueba.

Este teorema no solo establece que k_f es correcta con respecto a la clase de modelos de punto jo clausurados sino tambien que k_f es completa con respecto a esta clase, si se tienen en cuenta unicamente modelos de la forma hN^+ ; i, i.e., extensiones de N^+ . La propiedad de k_f que el teorema 3.14 establece suele llamarse 'N⁺-categoricidad' de k_f con respecto a la clase de clausuras de los modelos de punto jo bajo k_3 : los modelos de k_f estandares con respecto a L_{pa}^+ son exactamente los modelos pertenecientes a esta clase.

Corolario 3.15. k_f es !-consistente.¹⁷

La propiedad de !-consistencia resulta deseable para teor as de la verdad formuladas sobre la aritmetica, donde esta opera como teor a de la sintaxis. Grosso modo, mientras que el predicado veritativo no requiere mas objetos sintacticos que los que la aritmetica provee, una teor a !-inconsistente prueba que estos objetos deben existir; en otras palabras, contiene teoremas falsos (cf. [Leitgeb \(2007\)](#), [Barrio & Picollo \(2013\)](#)).

Hay tambien una relacion mas o menos directa entre k_f y la familia de modelos de punto jo sobre sk en k_3 . Si bien la logica externa de k_f , esto es, el conjunto de teoremas de esta teor a, esta cerrado bajo logica clasica, su logica interna, el conjunto de oraciones de las cuales k_f prueba que son verdaderas, obedece los principios de k_3 : la logica interna de k_f es correcta con respecto a la familia de modelos parciales de punto jo sobre sk en k_3 .

Proposicion 3.16. Si $k_f \vdash T p \rightarrow q$ y M es un modelo de punto jo sobre sk , $M \models k_3 \vdash$.

Si bien varios autores son partidarios de k_f (cf. [McGee \(1991\)](#), [Scott Soames \(1999\)](#)), esta teor a tiene al menos una caracter stica que suele considerarse indeseable: sus logicas externa e interna no coinciden. Como la logica externa, al ser clasica, es mas amplia que la interna; existen oraciones que k_f prueba pero que son declaradas no verdaderas por la misma teor a. k_f prueba todas las instancias de

$$T p \rightarrow q \vdash ! \quad (T\text{-Out})$$

donde \vdash es una oracion de L_T , i.e., la direccion de izquierda a derecha de los bicondicionales que conforman el [Esquema-T](#) (cf. [Cantini \(1996\)](#), p. 54, Teorema 8.8(i)). Luego, k_f prueba $T p \rightarrow T q \vdash ! \rightarrow T p \rightarrow T q$, lo que es lo mismo (vease (3.5)), $T p \rightarrow ! \rightarrow T p$. En consecuencia, $k_f \vdash ! \rightarrow T p$, es decir, k_f prueba la oracion del mentiroso y, a la vez, que esta no es verdadera.

¹⁷Vease la definicion 1.2.

Como se~nalan [Halbach & Horsten \(2006\)](#), basta con abandonar *cons* para evitar este problema.¹⁸ De hecho, McGee es un defensor de *kf* sin este axioma. Sin embargo, en el capitulo 5 argumento que, al contrario, *T-Out* es un principio altamente deseable, al menos desde una perspectiva de acionista. Por otro lado, *cons* permite probar el siguiente principio general intuitivo:

$$\exists x; y(\text{Sent}_{LT}(x!y) \rightarrow (T(x!y) \rightarrow (T x \rightarrow T y))) \quad (\text{imp})$$

una version interna de la regla Modus Ponens o de eliminacion del condicional

$$\frac{}{\quad} \quad \quad \quad (E!)$$

que vale tanto en *k3* como en la logica clasica. *(imp)* no es demostrable en *kf* - *cons*.

El siguiente resultado establece una relacion entre *kf* y *rt_o* (vease la definicion [2.14](#)) en terminos de su poder de prueba.

Teorema 3.17 (Halbach). *kf* es relativamente interpretable en *rt_o* y viceversa.¹⁹

3.1.3.2. Verdad Positiva

Algo que las oraciones que dan lugar a paradojas semanticas tienen en comun es que en ellas el predicado de verdad se encuentra de alguna manera negado. En efecto, si se restringe el [Esquema-T Uniforme](#) a formulas en las cuales esto no sucede, se obtiene una teoria no solo consistente sino *!*-consistente, *putb* (cf. [Cantini \(1989\)](#), [Halbach \(2009\)](#)).

La teoria desentrecorolladora resultante es mucho mas abarcativa que *utb(pa⁺)* porque incluye instancias del [Esquema-T Uniforme](#) dadas no solo por todas y cada una de las oraciones de *L⁺_{pa}* sino tambien por muchas oraciones en las cuales *T* ocurre, como *T p0 = 0q* y $\exists x(\text{Bew}_{pa^+}(x) \rightarrow Tx)$. *putb* es una teoria de la verdad en la cual el predicado veritativo no esta tipeado: puede y de hecho se aplica a oraciones que ya lo contienen. A diferencia de las teorias semanticas de punto jo de Kripke, *putb* no abandona el ambito de la logica clasica. Para sortear las hipotesis del teorema de Tarski (teorema [2.9](#)), relaja el requisito de que todas las instancias del [Esquema-T](#) se sigan de los axiomas de la teoria. Si bien, desde cierto punto de vista, el predicado de verdad resultante es parcial o incompleto, *putb* constituye una mejora con respecto a las versiones axiomatizadas tipeadas, si lo que se desea es un unico predicado de verdad autoaplicable y no una jerarquía

¹⁸En cuyo caso habra que agregar el axioma $\exists x(T x \rightarrow \text{Sent}_{LT}(x))$ para preservar la *N⁺-categoricidad de la teoria*.

¹⁹Vease la definicion [1.1](#) y [Halbach \(2011, x15.3\)](#).

con infinitos predicados y sus subíndices. Alternativamente, putb puede verse como una teoría de la verdad completa para el sublenguaje de L_T que comprende exclusivamente fórmulas donde el predicado de verdad ocurre positivamente. En palabras de [Leitgeb \(2005, p. 155-156\)](#),

knowing that there is a correct and adequate definition of truth for L_{pa}^+ , and knowing that there is no such definition for L_T , we might still be looking for a formally correct definition of truth that is applicable to a class of sentences with L_{pa}^+ (L_T), such that the definition is materially adequate with respect to the members of .

Es preciso tener cuidado al identificar los casos en los cuales el predicado de verdad ocurre negado en una fórmula, porque L_T cuenta con conectivas que pueden utilizarse para definir negaciones, como \neg .

Definición 3.18 (T-positividad). Una fórmula $\varphi \in L_T$ es T-positiva si la suma de símbolos de negación bajo cuyo alcance cada aparición de T ocurre sumada a la cantidad de veces que esa aparición ocurre en el antecedente de un condicional es par.²⁰

Definición 3.19 (putb). $\text{putb } L_T$ es la teoría que extiende pat mediante las instancias del [Esquema-T Uniforme](#)

$$\exists t_1 \dots t_n (T \varphi(t_1, \dots, t_n) \supset \varphi(t_1, \dots, t_n))$$

donde φ es una fórmula T-positiva de L_T con exactamente n variables libres.

Notese que, como todas las fórmulas de L_{pa}^+ son T-positivas, $\text{utb } \text{putb}$. putb también puede ser entendida como una axiomatización de la familia de las clausuras de los modelos de punto jo sobre \mathcal{K}_3 , en tanto es una subteoría de \mathcal{K}_f .

Proposición 3.20 (Cantini). Todas las instancias del [Esquema-T Uniforme](#) dadas por fórmulas T-positivas de L_T son teoremas de \mathcal{K}_f .²¹

En consecuencia, \mathcal{K}_f también puede verse como una teoría de la verdad libre de tipos que sortea las hipótesis del teorema de Tarski relajando la adecuación material pero preservando la lógica clásica.

Corolario 3.21. putb es \neg -consistente.

²⁰La definición de [Halbach \(2009, p. 788, de definición 2.1\)](#) concierne únicamente a los símbolos de negación, porque el lenguaje con el cual trabaja no contiene condicionales sino disyunciones y conjunciones como únicas conectivas diádicas. Su definición es equivalente a la mía en virtud de la definibilidad del condicional mediante la negación y la disyunción.

²¹Vease, v.g., [Cantini \(1989, lema 3.2\(ii\)\)](#).

Demostración. Por el corolario [3.15](#) y la proposición anterior. \square

Corolario 3.22. $\text{putb} + \text{T-Out}$ es ω -consistente.

Demostración. Por el corolario [3.15](#), la proposición anterior y el hecho de que kf prueba todas las instancias de T-Out . \square

No obstante, a diferencia de kf , putb no es N^+ -categorica con respecto a los modelos clausurados de punto ω sobre sk . putb tiene modelos de la forma hN^+ ; i donde contiene únicamente oraciones T -positivas, pero ninguno de estos modelos pertenece a la clase señalada.

Al ser una subteoría sencilla y no N^+ -categorica de kf , podría pensarse que putb es menos poderosa que kf en cuanto a su capacidad de prueba. Sin embargo, la sencillez de putb no guarda relación con su poder de prueba. putb es mucho más poderosa que $\text{utb}(\text{pa}^+)$ e igualmente poderosa que kf .

Teorema 3.23 (Halbach). kf es relativamente interpretable en putb .²²

Como putb es a la vez una subteoría de kf , también es relativamente interpretable en ella, lo cual muestra que tienen exactamente el mismo poder de prueba.

Corolario 3.24. putb es relativamente interpretable en $\text{rt}_{<\omega}$ y en kf y viceversa.

Demostración. Por el teorema [3.17](#) y el corolario [3.23](#). \square

Estos resultados no significan que putb sea una teoría composicional de la verdad. Si bien es capaz de definir los predicados de verdad de kf y $\text{rt}_{<\omega}$ y probar, para ellos, los principios generales que constituyen los axiomas de estas teorías, putb no es capaz de probar ningún principio general no trivial para su propio predicado de verdad.²³

§

Kripke no solo mostró cómo contar con un predicado de verdad transparente y, a la vez, trascender la separación tarskiana entre lenguaje objeto y metalenguaje abandonando la lógica clásica sino que también inspiró teorías no tipeadas de la verdad que, si bien no ofrecen predicados de verdad completamente transparentes, están dadas en el marco de la lógica clásica, como kf y putb . Asimismo, sus teorías semánticas fueron fuente de inspiración del concepto de dependencia de Leitgeb y la teoría de la verdad a la cual dio lugar, que presento en la siguiente sección.

²²Vease (Halbach, 2009, teorema 5.1). ²³Vease Halbach (2009, x6).

3.2. La dependencia de Leitgeb

Trabajando dentro de los límites de la lógica clásica, [Leitgeb \(2005, p. 156, sus itálicas\)](#) se pregunta "What kinds of sentences with truth predicate may be inserted plausibly and consistently into the T-scheme?" Como evidencia putb en el apartado anterior, es posible insertar al menos algunas oraciones que contienen el predicado veritativo en el [Esquema-T](#) sin caer en contradicciones, esto es, la restricción tarskiana que recomienda tomar únicamente aquellas instancias de este principio dadas por oraciones en las cuales el predicado veritativo no ocurre no es necesaria para evitar inconsistencias. Pero quizás es posible adoptar aun más instancias del [Esquema-T](#) que las T-positivas, o adoptar un conjunto diferente de instancias que est mejor motivado.

Con este objetivo, Leitgeb desarrolla la noción de dependencia. Grosso modo, una oración de L_{pa}^+ no depende de nada, mientras que las oraciones que contienen T dependen del conjunto de oraciones de las cuales predicen verdad. Por ejemplo, $T p0 = 0q$ depende del conjunto que contiene la oración $0 = 0$, $T pT p0 = 0qq$ del conjunto que contiene la oración $T p0 = 0q$ y $:T I$ en (3.5) depende del conjunto que la tiene a ella misma por elemento. La propuesta de Leitgeb consiste en tomar únicamente aquellas oraciones cuya verdad o falsedad dependa, directa o indirectamente [como $T p0 = 0q$ o $T pT p0 = 0qq$, respectivamente], de la verdad o falsedad de las oraciones de L_{pa}^+ .

En esta sección presento primero los resultados limitativos de McGee a la búsqueda de criterios plausibles de selección de instancias seguras del [Esquema-T](#), que llevan a Leitgeb al desarrollo de su noción de dependencia. En 3.2.2 introduzco las definiciones formales de dependencia, autodependencia, d-fundación y otras nociones asociadas, para las cuales pruebo diversos resultados que serán de utilidad en el apartado 3.2.3 que sigue. Ah presento en primer lugar la teoría semántica de Leitgeb, esto es, el modelo de L_T donde las instancias del [Esquema-T](#) están restringidas a oraciones d-fundadas, y, posteriormente, tres modelos adicionales de L_T cuyo criterio de restricción es más permisivo que la d-fundación. Estos resultados serán de suma utilidad en el capítulo 8, para establecer la consistencia de las teorías formales de la verdad que introduzco ah .

3.2.1. Un resultado limitativo de McGee

Una opción prima facie natural para restringir el [Esquema-T](#) es tomar simplemente todas aquellas instancias que son consistentes. Sea T el conjunto de instancias del [Esquema-T](#) generadas por miembros de L_T , i.e.,

$$T := \{ \langle T, p \rangle \mid p \in L_T \}$$

El conjunto ideal sería entonces aquel que diera lugar a un conjunto de instancias del [Esquema-T](#) maximal consistente. Lamentablemente, [McGee \(1992a\)](#) mostró que esta no es una vía razonable.

Teorema 3.25 (McGee). Para cada L_T consistente con pat existe un L_T tal que

1. para todo $\phi \in L_T$, $\text{pat}[\phi]$;
2. $\text{pat}[L_T]$ es consistente;
3. si L_T , $\text{pat}[L_T]$ es inconsistente;
4. $\text{pat}[L_T]$ es completa con respecto a la negación.

Luego, toda oración indecidible en pat es decidible positivamente en alguna teoría desentrecorolladora maximal $\text{pat}[L_T]$ y negativamente en otra $\text{pat}[L_T^0]$, i.e., no existe un único conjunto maximal consistente de instancias del [Esquema-T](#) sino una cantidad no numerable de ellos.

Aun más, esta cantidad no se reduce siquiera restringiendo la atención a conjuntos tales que $\text{pat}[L_T]$ decida positivamente aquellas oraciones ϕ de L_{pa}^+ que son verdaderas en N^+ , pues quedan todavía en libertad las oraciones de L_T que contienen T . Por la proposición [1.9](#) aplicada a los predicados Tx y $\neg Tx$, existen dos términos I_1 y I_2 de L_{pa}^+ tales que

$$q \in I_1 \iff p \in I_2 \wedge I_2 = p \in I_1 \quad (3.6)$$

esto es, dos oraciones $T I_2$ y $\neg T I_1$ tales que la primera dice de la segunda que es verdadera, mientras que la segunda dice de la primera que no lo es. Juntas estas oraciones conforman lo que se conoce como un 'ciclo de mentirosos de longitud 2'. A partir de las instancias del [Esquema-T](#) para cada una de ellas es posible derivar fácilmente una contradicción, usualmente denominada 'paradoja de la postal'. Luego, ningún dado por el teorema [3.25](#) contiene ambas oraciones. Sin embargo, las instancias del [Esquema-T](#) dadas por $T I_2$ y $\neg T I_1$ son consistentes tomadas por separado con cualquier otro conjunto consistente de instancias. Luego, todo conjunto que satisfaga las condiciones 1-4 del teorema [3.25](#) contiene, por ser maximal, o bien $T I_2$ o bien $\neg T I_1$. Si se adoptara la consistencia maximal como criterio de selección de instancias del [Esquema-T](#) se debería tomar decisiones arbitrarias, donde ninguna de las opciones resulta razonable.

Esta es la razón por la cual Leitgeb agrega el calificativo 'plausibly' a su pregunta por las oraciones que pueden consistentemente insertarse en el [Esquema-T](#). Lo que busca

es un criterio de selección amplio pero más restrictivo que la consistencia y que, adicionalmente, tenga cierta razón de ser, cierto poder explicativo. Su propuesta está dada en función de su noción de dependencia, que presento a continuación.

3.2.2. Dependencia, autodependencia y d-fundación

La noción de dependencia de Leitgeb tiene claramente visos de la noción kripkeana de fundación (véase el apartado [3.1.2](#)). Conceptualmente aquellas oraciones cuya verdad o falsedad depende de la verdad o falsedad de las oraciones de L_{pa}^+ son precisamente las oraciones fundadas. Como el mismo Kripke sostiene,

In general, if a sentence [. . .] asserts that (all, some, most, etc.) of the sentences of a certain class are true, its truth value can be ascertained if the truth values of the sentences in the class C are ascertained. If some of these sentences themselves involve the notion of truth, their truth value in turn must be ascertained by looking at other sentences, and so on. If ultimately this process terminates in sentences not mentioning the concept of truth, so that the truth value of the original statement can be ascertained, we call the original sentence grounded; otherwise, ungrounded. (Kripke, 1975, p. 693-694, sus *italicas*)

Efectivamente, en (1982) Yablo argumenta que la noción intuitiva de fundación tiene dos caras. Por un lado, un aspecto ascendente, hereditario, que es capturado por el modo en que Kripke construye sus modelos de punto fijo mínimos, i.e., por las iteraciones de los operadores salto sobre el conjunto vacío. Las oraciones de L_{pa}^+ transmiten su fundación a las oraciones que predicán verdad sobre ellas, que a su vez transmiten su fundación a las que predicán verdad sobre estas últimas, etc. Pero, por otro lado, la fundación tiene una cara descendente, dependiente, descrita en la cita de Kripke del párrafo anterior. De acuerdo con Yablo, Kripke dejó esta cara de la fundación sin explorar y, en consecuencia, ofrece el mismo un abordaje de la dependencia.

Mientras que la noción de dependencia de Yablo representa simplemente otro ángulo desde el cual mirar la fundación kripkeana, como señala Leitgeb, el concepto que él propone tiene marcadas diferencias con estos últimos, dadas principalmente por cuestiones conceptuales y por trabajar en lógica clásica en lugar de lógicas paracompletas, además de ciertas divergencias técnicas (cf. [Leitgeb \(2005, x5\)](#)). Consecuentemente, el conjunto de oraciones fundadas de nido en términos de la noción de dependencia de Leitgeb va a divergir de los conjuntos de oraciones fundadas de Kripke y Yablo. ²⁴

²⁴Vease espec camente [Leitgeb \(2005, x5, teorema 21\)](#).

Como las oraciones de L_T pueden predicar verdad sobre más de una expresión,

la dependencia es definida como una relación entre oraciones, por un lado, y conjuntos de oraciones, por otro. Intuitivamente, una oración de L_T depende de un conjunto de oraciones en la medida en que su valor de verdad supervenga sobre la presencia o ausencia de los miembros de este conjunto en la extensión del predicado veritativo, es decir, en la medida en que ningún otro factor excepto la presencia o ausencia de los miembros del conjunto del cual depende en la extensión de T pueda afectar su valor de verdad. Por ejemplo, la verdad o falsedad de $T \text{ p}0 = 0q$ depende de si $0 = 0$ es un miembro de la extensión de T o no y de nada más, mientras que la de $\exists x(\text{Bew}_{\text{pa}^+}(x) \rightarrow T x)$ depende de si todos y cada uno de los teoremas de pa^+ pertenecen a la extensión del predicado veritativo.

Definición 3.26 (Dependencia). Dadas una oración $\phi \in L_T$ y un conjunto $\Gamma \subseteq L_T$, ϕ depende de Γ si, para todo L_T ,

$$M \models \phi \text{ si y sólo si } M \models \bigwedge \Gamma$$

Si ϕ depende de Γ , todo lo que está por fuera de Γ puede ser agregado o quitado de la extensión del predicado de verdad sin afectar el valor de verdad de ϕ en el modelo. Esto significa que cuando se dice que una oración depende de un conjunto, se dice que todo aquello que es necesario para determinar su valor de verdad está en este conjunto, nada que pueda afectar su valor de verdad quedo fuera. Pero puede pasar que muchas cosas irrelevantes hayan quedado dentro. De hecho, vale el siguiente resultado.

Proposición 3.27 (Leitgeb). Sean $\phi \in L_T$ y $\Gamma \subseteq L_T$. Si ϕ depende de Γ entonces también depende de todo conjunto Δ que contenga a Γ ; en particular, todas las oraciones dependen de L_T .

La proposición que sigue muestra que agregar a la extensión del predicado veritativo oraciones que pertenecen a un conjunto cuya intersección con otro del cual ϕ depende es vacía no puede afectar el valor de verdad de esta oración.

Proposición 3.28. Sean $\phi \in L_T$ y $\Gamma \subseteq L_T$. Si existe un conjunto del cual ϕ depende tal que $\Gamma \cap \Delta = \emptyset$,

$$M \models \phi \text{ si y sólo si } M \models \bigwedge \Gamma \text{ y } M \models \bigwedge \Delta$$

²⁵ Véase Leitgeb (2005, p. 161, de definición 1 y lema 2).

²⁶ Véase Leitgeb (2005, p. 161, lema 3).

Demostración. Como $\forall i \in I, \forall \alpha \in \mathcal{L}_T, \alpha \in \mathcal{L}_T$. Luego,

$\mathcal{M} \models \alpha$; $\mathcal{M} \models \alpha$; $\forall i \in I$ De noción [3.26](#)

$\mathcal{M} \models \alpha$; $\mathcal{M} \models \alpha$; $\forall i \in I$
 $\mathcal{M} \models \alpha$; $\mathcal{M} \models \alpha$; $\forall i \in I$ De noción [3.26](#)

□

Proposición 3.29. Sean Γ y $\Delta \subseteq \mathcal{L}_T$. Si, para todo $\mathcal{M} \models \Gamma, \mathcal{M} \models \Delta$ sii $\mathcal{M} \models \Gamma$;

entonces, para todo $\mathcal{M} \models \Gamma$, Δ depende de Γ sii Δ depende de Γ .

Demostración. Por la de noción [3.26](#),

Δ depende de Γ , para todo $\mathcal{M} \models \Gamma; \mathcal{M} \models \Delta$ sii $\mathcal{M} \models \Gamma; \mathcal{M} \models \Delta$

Δ depende de Γ , para todo $\mathcal{M} \models \Gamma; \mathcal{M} \models \Delta$ sii $\mathcal{M} \models \Gamma; \mathcal{M} \models \Delta$
 Δ depende de Γ

□

La noción de dependencia de Leitgeb es semántica, en tanto queda determinada por lo que es verdadero o falso en el modelo estándar de \mathcal{L}_{pa}^+ , y, por ende, también extensional, como muestra esta última proposición: dos oraciones equivalentes en el modelo estándar dependen necesariamente de los mismos conjuntos.

Algunas oraciones tienen un conjunto mínimo del cual dependen, que está incluido en todos los otros conjuntos de los que dependen, su dependencia esencial.

Definición 3.30 (Dependencia esencial). Una oración $\alpha \in \mathcal{L}_T$ depende esencialmente de $\Gamma \subseteq \mathcal{L}_T$ sii α depende de Γ , para todo $\Gamma' \subseteq \mathcal{L}_T$ del cual α depende, $\Gamma \subseteq \Gamma'$.

Por ejemplo, $\exists x(Tx)$ depende esencialmente de $\{Tx\}$ y $\exists x(\text{Sent}_{\mathcal{L}_T}(x) \rightarrow Tx)$ depende esencialmente de \mathcal{L}_T , porque basta con quitar una oración de \mathcal{L}_T de la extensión de T para que el valor de verdad de este enunciado varíe. Otras oraciones, en cambio, no dependen esencialmente de ningún conjunto, v.g., $\exists x > 0 \forall y (y > x \rightarrow Ty)$ (cf. [Leitgeb \(2005, p. 164, lista de ejemplos 1\)](#)). Para que esta oración sea verdadera, no puede existir un código a partir del cual ninguna oración pertenezca a la extensión de T , pero no importa cuán grande sea el menor código con estas características. Luego, esta oración depende de $\{f_1; 2; \dots; g; f_2; 3; \dots; g; \dots\}$, pero no hay un conjunto más chico del cual depende.

Proposición 3.31 (Leitgeb). Sean Γ ; Δ oraciones de \mathcal{L}_T y $\Gamma_i \subseteq \mathcal{L}_T$, con $i \in \mathbb{N}$.

1. Si $\phi \in L_{pa}^+$, ϕ depende esencialmente de ψ ?
2. Si $\phi := \bigwedge_{i < n} \psi_i$ y $\psi_i = p_i \rightarrow q_i$ para alguna $\langle p_i, q_i \rangle \in L_T$, ϕ depende esencialmente de $\langle p_i, q_i \rangle$.
3. Si ϕ depende de ψ , ψ también depende de ϕ .
4. Si ϕ depende de ψ y ψ depende de χ , ϕ depende de χ .
5. Si, para cada $n \in \mathbb{N}$, ϕ_n depende de ψ_n , $\bigwedge_{i \in \mathbb{N}} \phi_i$ depende de $\bigwedge_{i \in \mathbb{N}} \psi_i$.²⁷

Luego, la dependencia obedece cierto tipo de composicionalidad. A partir ella es posible definir nociones de autodependencia y d-fundación. La autodependencia es una noción cercana de alguna manera a la autorreferencia pero, como se verá en los capítulos 6 y 7, se diferencia de aquella por su carácter semántico. Como el mismo Leitgeb (2005) sugiere, su noción de dependencia es comparable a pero diferente del concepto de contenido (aboutness) o referencia. Me ocuparé de esta cuestión en detalle en el capítulo 6.

Definición 3.32 (Autodependencia). $\phi \in L_T$ es autodependiente si existen oraciones $\psi_1, \dots, \psi_n \in L_T$ tales que, para todo $\psi \in L_T$ del cual ϕ depende, $\psi_i \in L_T$ para $1 \leq i < n$, para todo del cual ψ_i depende, $\psi_{i+1} \in L_T$; y, para todo ψ del cual ψ_n depende, $\psi \in L_T$.

Esta definición incluye casos de autodependencia directa e indirecta. Por ejemplo, la oración $\neg \text{TI}$ del mentiroso depende esencialmente de f:TI , todo conjunto del cual esta oración depende la tiene como uno de sus miembros y, por ende, es autodependiente. Pero, además, si TI_1 y TI_2 son como en (3.6), como TI_2 depende esencialmente de f:TI_1 y TI_1 depende esencialmente de f:TI_2 , TI_1 y TI_2 son asimismo autodependientes.²⁸

Para definir fundación utilizando la noción de dependencia, lo que llamo 'd-fundación', hay que dar un rodeo. La idea general consiste en que lo fundado es aquello que depende, o bien del conjunto vacío, o bien del conjunto de oraciones que dependen del conjunto vacío, o bien del conjunto de oraciones que dependen de conjuntos de oraciones que a su vez dependen del conjunto vacío, etc.

Definición 3.33 (pf). Sea $\phi := \psi$, $\psi_{i+1} := \bigwedge_{s < i} \psi_s \rightarrow \psi$ para cada ordinal i , y, si α es un ordinal límite, $\psi_\alpha := \bigwedge_{i < \alpha} \psi_i$. Como esta jerarquía es monótona, i.e., si $\alpha < \beta$ entonces $\psi_\alpha \rightarrow \psi_\beta$, por consideraciones de cardinalidad existe un ordinal mínimo tal que $\psi_\alpha = \psi_{\alpha+1}$. Sea pf este punto fijo mínimo.

²⁷ Véase Leitgeb (2005, p. 162, lista de ejemplos 1 y p. 165, lema 5).

²⁸ Leitgeb (2005) de ne únicamente la autodependencia (o, como él la denomina, 'autorreferencia') directa y exclusivamente para oraciones con dependencia esencial. Extiendo a su noción del modo más natural por cuestiones de necesidad relativas a lo que viene luego.

Γ_1 contiene todo lo que depende del conjunto vacío, Γ_2 todo lo que depende del conjunto vacío o más todo lo que depende de conjuntos de oraciones que dependen del conjunto vacío, etc. Intuitivamente, pf contiene aquellas oraciones de L_T cuyo valor de verdad está determinado exclusivamente por el modelo estándar de L_{pa}^+ , las oraciones fundadas. La similitud con Kripke en este punto es clara.

Definición 3.34 (D-fundación). Γ_2 L_T está d-fundada si y sólo si pf .

3.2.3. Dependencia y desentrecomillación

La noción de dependencia, como quiere Leitgeb, puede utilizarse para discriminar instancias del [Esquema-T](#) de modo tal que la teoría resultante sea consistente e, incluso, ω -consistente. A continuación presento una serie de resultados que muestran modos diferentes de discriminar.

Teorema 3.35 (Leitgeb). Existe un único pf tal que hN^+ ; pf i T pf . [29](#)

El concepto de d-fundación puede funcionar entonces como criterio de selección de instancias consistentes del [Esquema-T](#), dando lugar a la teoría a semántica de Leitgeb, el modelo hN^+ ; pf i . Es innegable que el criterio empleado tiene una gran fuerza intuitiva. Este criterio puede relajarse todavía más, de dos modos diferentes, obteniendo modelos que validen todavía más instancias del [Esquema-T](#), generando teorías semánticas más abarcativas.

Proposición 3.36. Si L_T es un conjunto finito de oraciones no autodependientes, existe un L_T tal que hN^+ ; i T (pf []).

Demostración. Asumo que $\setminus \text{pf} = ?$ (de haber oraciones en la intersección, quedarán automáticamente cubiertas por el teorema [3.35](#)). Sea $f : \mathcal{P}(L_T) \rightarrow \mathbb{N}$ tal que, si $\Gamma \subseteq L_T$ es la intersección de todos los subconjuntos de L_T de los cuales Γ depende,

$$f(\Gamma) := \begin{cases} 80 & \text{si } \Gamma = \emptyset \\ 2 \setminus g & \text{si } \Gamma \neq \emptyset \end{cases}$$

$$f(\Gamma) := \max_{\Gamma' \subseteq \Gamma} f(\Gamma') + 1 \text{ en cualquier otro caso}$$

f existe porque, por un lado, debe haber oraciones en tales que algún conjunto del cual dependen no contenga otras oraciones de L_T . De lo contrario, habría autodependencia dado que es finito. Por otro lado, como no hay autodependencia, directa o indirecta, y es finito, siempre es posible ordenar los miembros de L_T de acuerdo con la relación de dependencia comenzando por aquellos que no dependen de los otros.

²⁹Vease [Leitgeb \(2005\)](#), p. 172, teorema 17) para una prueba.

Si $n \geq 2$!, sea $\mathcal{L}_n := \{f^2 : f(\cdot) = ng\}$. Como es nito, solo un numero nito de estos conjuntos es no vac o. Sea $m \geq 2$! el ultimo numero tal que $m \neq ?$. Claramente, $\mathcal{L}_m \cap \mathcal{L}_n = \emptyset$. Si f^2 no tiene dependencia esencial, sea \mathcal{L}_T tal que f^2 depende de \mathcal{L}_T y $\mathcal{L}_T \neq \emptyset$. Siempre se puede elegir un \mathcal{L}_T con estas caracter sticas porque es nito. Si f^2 s tiene dependencia esencial, sea $\mathcal{L}_T = \emptyset$.

Por el teorema [3.35](#), $\mathcal{L}_T \in \text{hN}^+$; $\text{pf } i \in T$ pf . Sean

$$\begin{aligned} \mathcal{L}_0 &:= \text{pf } [f^2 \mathcal{L}_0 : \text{hN}^+; \text{pf } i \in 'g; \\ \mathcal{L}_{n+1} &:= \mathcal{L}_n [f^2 \mathcal{L}_{n+1} : \text{hN}^+; \text{ni } 'g \end{aligned}$$

Se vera que $\mathcal{L}_T \in \text{hN}^+$; $\text{mi } T$ ($\text{pf } [$). Primero, como $\mathcal{L}_m = \text{pf } [$, donde y, por tanto, $\mathcal{L}_T \in \text{hN}^+$; $\text{mi } T$ pf . Segundo, si f^2 , existe un $n \geq m$ tal que $f^2 \in \mathcal{L}_n$. Sea $\mathcal{L}_T := \text{pf}$. Luego,

$$\begin{aligned} \mathcal{L}_T \in \text{hN}^+; \text{mi } ' , \text{hN}^+; \mathcal{L}_T \in \text{hN}^+; \text{ni } ' & \text{Proposicion } \a href="#">3.28 \\ , \text{hN}^+; \text{ni } T p'q & \\ , \text{hN}^+; \text{mi } T p'q & \text{Proposicion } \a href="#">3.28 \end{aligned}$$

porque cada $f^2 \in \mathcal{L}_n$ ($f(\cdot) = n$) depende de \mathcal{L}_T (en [3.28](#)), que es disjunto con $\mathcal{L}_{n+1} : \text{hN}^+; \text{ni } 'g_{n+1}$ (en [3.28](#)), ya que, si $\mathcal{L}_T \in \mathcal{L}_{n+1} \neq ?$, $f(\cdot) > n + 1$, lo cual es imposible, y $\mathcal{L}_T \neq ?$. \square

Proposicion 3.37. Si \mathcal{L}_T es un conjunto in nito de oraciones, cada una de las cuales depende esencialmente de un conjunto nito y no es autodependiente, existe un \mathcal{L}_T tal que $\mathcal{L}_T \in \text{hN}^+; \text{mi } T$ ($\text{pf } [$).³⁰

Demostracion. Sea $\mathcal{L}_T := \{f^2_1; f^2_2; \dots; g\}$. Nuevamente, asumo que $\mathcal{L}_T \in \text{hN}^+; \text{mi } T$. Para cada $n \geq 2$!, sean \mathcal{L}_n el conjunto del cual f^2_n depende esencialmente, $\mathcal{L}_n := \{f^2_1; \dots; f^2_n; g\}$ y

$\mathcal{L}_n := \{f^2_1; \dots; f^2_n; g\}$! : pf $\bigcup_{i=1}^n \text{pf } [\mathcal{L}_i (\mathcal{L}_i [f^2_i g) \text{ y } \text{hN}^+; \text{mi } T p'q \text{ } \mathcal{L}_n$ existen $f^2_1; \dots; f^2_n$ tales $f^2_1 \dots f^2_n \dots f^2_n \dots f^2_n$. Cada \mathcal{L}_n es nito, porque \mathcal{L}_n ($\mathcal{L}_n [f^2_i g$) lo es. Adem as, se vera que para cada $n \geq 2$! que \mathcal{L}_n donde \mathcal{L}_n para cada $1 \leq i \leq n$.

Primero, por la proposicion [3.36](#), para cada $n \geq 2$! hay un modelo $\mathcal{L}_T \in \text{hN}^+; \text{mi } T$ en el cual pf y T ($\text{pf } [\mathcal{L}_n$) es verdadera. Para cada $1 \leq i \leq n$, sea

$$\mathcal{L}_i := \mathcal{L}_n \setminus (\text{pf } [\mathcal{L}_i (\mathcal{L}_i [f^2_i g))$$

³⁰Este resultado fue probado en colaboracion con Thomas Schindler.

Luego, para cada $i \in \mathbb{N}$, $i \geq 1$. Y, como por la proposicion 3.29 cada $\alpha \in \Sigma^i$ depende de $\text{pf}[\Sigma^k]$ ($k \in \mathbb{N}^+$; $i \geq k$), con lo cual $\alpha \in \Sigma^i$.

De no recursivamente un arbol A como un conjunto de secuencias formadas a partir de miembros de Σ^i del siguiente modo:

1. $\alpha \in \Sigma^i$.
2. Si $\alpha = \alpha_1 \dots \alpha_n \in \Sigma^i$ y $\alpha_{n+1} \in \Sigma^{i-n}$, $\alpha_1 \dots \alpha_n \alpha_{n+1} \in \Sigma^i$.
3. Nada mas esta en A .

Notese que, para cada $n \in \mathbb{N}$, A contiene una secuencia $\alpha = \alpha_1 \dots \alpha_n$ de longitud n .

Ademas, para cada secuencia tal, $\alpha = \alpha_1 \dots \alpha_n \in \Sigma^i$, si $0 < i < n$, con lo cual todos los elementos de las secuencias estan conectados a traves de α . Y, como cada $\alpha \in \Sigma^i$ es nito pero tiene al menos un elemento, A es un arbol finito de ramiacion nita [esto es, cada nodo pertenece a un numero finito de secuencias]. Luego, por el lema de Konig, debe haber en A una secuencia infinita $\alpha = \alpha_1 \alpha_2 \dots$ con $\alpha_i \in \Sigma^i$ para todo $i > 0$.³¹

Sea $\alpha = \alpha_1 \alpha_2 \dots$. Se vera que $\alpha \in \Sigma^i$ ($\text{pf}[\Sigma^i]$). Como, para cada $n \in \mathbb{N}$, $\text{pf}[\Sigma^n] = \text{pf}[\Sigma^{n+1}]$ y, por la proposicion 3.28, $\alpha \in \Sigma^i$.

Dada una $\alpha \in \Sigma^i$, sea $m > i$ el menor numero natural tal que, para todo $k \leq m$, α_k no contiene miembros de Σ^i que no estuvieran ya en Σ^m . m existe porque Σ^i es siempre nito y las secuencias de Σ^j con $j \geq i$ son monotonas. Como $i < m$ y $\alpha \in \Sigma^i$, $\alpha_m \in \Sigma^i$; $\alpha_m \in \Sigma^i \implies \alpha_m \in \Sigma^i$. Dado el m que se eligio, la proposicion 3.28 implica que $\alpha \in \Sigma^i$; $\alpha \in \Sigma^i \implies \alpha \in \Sigma^i$. \square

Si ningun miembro de Σ^i es autodependiente y es, o bien nito, o bien tal que cada uno de sus elementos depende esencialmente de un conjunto nito de oraciones, es posible extender el conjunto pf de instancias aceptables del Esquema-T mediante los miembros de Σ^i . Ademas, el modelo resultante es un modelo de T-Out.

Corolario 3.38. Sean Σ^i y Σ^j como en la proposicion anterior. Para toda oracion α de L_T , $\alpha \in \Sigma^i \implies \alpha \in \Sigma^j$.

Demostracion. La prueba es trivial, en especial si $\alpha \in \Sigma^i$ ($\text{pf}[\Sigma^i]$). Si $\alpha \in \Sigma^i$ ($\text{pf}[\Sigma^i]$), como $\text{pf}[\Sigma^i] \subseteq \text{pf}[\Sigma^j]$. Por tanto, $\alpha \in \Sigma^i \implies \alpha \in \Sigma^j$; esto es, $\alpha \in \Sigma^i \implies \alpha \in \Sigma^j$. \square

³¹El lema de Konig establece que si A es un grafo conectado (para cualesquiera dos nodos hay un camino que los conecta) con infinitos vertices, cada uno de los cuales pertenece a un numero nito de secuencias, entonces A contiene una secuencia infinita.

§

Leitgeb ha logrado poner en términos precisos la idea de dependencia y, con ella, la de fundación, independientemente de su teoría semántica de la verdad, a diferencia de Kripke (vease el apartado [3.1.2](#)), quien extrae el conjunto de oraciones fundadas de sus modelos de punto fijo mínimos. Luego, muestra como es posible restringir el [Esquema-T](#) a instancias fundadas sin dar lugar a contradicciones ni consecuencias contraintuitivas, construyendo un modelo de L_T que extiende N^+ en el cual estas instancias resultan válidas. Basándose en estos resultados, muestro que también es posible construir modelos en los cuales se validen aún más instancias del [Esquema-T](#), delimitadas también en función de la noción de dependencia pero ya no solo por la noción de fundación.

Para saber si una oración depende de un conjunto de expresiones es preciso recurrir al valor de verdad de esta u otras expresiones en el modelo estándar de la aritmética. Por eso, al igual que las nociones de verdad de Kripke, la noción de Leitgeb dada por el modelo $hN^+; \text{pt}$ es altamente compleja. En efecto, [Philip Welch \(2014\)](#) ha probado que es Σ_1^1 , eso es, solo puede ser definida por una fórmula que comienza con un cuantificador universal de segundo orden seguido de una fórmula de primer orden, o con expresiones aún más complejas y, plausiblemente, lo mismo se aplica a los modelos que se obtienen mediante las proposiciones [3.36](#) y [3.37](#). No obstante, al igual que para los modelos de Kripke, existen axiomatizaciones parciales de $hN^+; \text{pt}$, en las cuales no entro en detalle (cf. [Schindler \(2014a\)](#)). En la sección que sigue presento una teoría axiomática clásica que es independiente de toda construcción semántica.

3.3. Friedman-Sheard

La teoría fs que presento en esta sección se debe originalmente a [Harvey Friedman & Michael Sheard \(1987\)](#), quienes ofrecieron una axiomatización significativamente diferente a la que aquí introduzco. Esta se debe a [Halbach \(1994, 2011\)](#), quien hizo famosa la teoría.

fs consiste en una generalización de la teoría de la verdad composicional $ct(pa^+)$ a oraciones de todo L_T , esto es, relativiza las cláusulas composicionales, ya no al predicado $Sent_{L^+pa}$ sino ahora a $Sent_{L_T}$. Como los axiomas resultantes no dicen nada acerca de la verdad de oraciones atómicas de la forma de $T p \rightarrow q$ y adoptar las instancias del [Esquema-T](#) para estas oraciones llevar a inmediatamente a contradicción a través de la oración del mentiroso $\neg T T$ en [\(3.5\)](#), fs cuenta con dos reglas especiales de introducción y eliminación del predicado, semejantes a [T-Intro](#) y [T-Elim](#) pero mucho más restrictivas.

Definición 3.39 (fs). Sea fs_{LT} la teoría que extiende pat con los siguientes axiomas:

mas:

$$fs1 \quad t; s(T(t=s) \quad \$t = s)$$

$$8 \quad .$$

$$fs2 \quad \exists x(Sent_{LT}(x) \rightarrow (T \rightarrow x \rightarrow T x))$$

$$fs3 \quad \exists x; y(Sent_{LT}(x \rightarrow y) \rightarrow (T(x \rightarrow y) \rightarrow (T x \rightarrow T y)))$$

$$fs4 \quad \exists x; v(Sent_{LT}(\exists v x) \rightarrow (T \exists v x \rightarrow \exists t T x(t=v)))$$

y esta cerrada por las siguientes metareglas:

$$fs \quad ' \rightarrow fs \quad ' \rightarrow T p'q \quad (nec)$$

$$fs \quad ' \rightarrow T p'q \rightarrow fs \quad ' \quad (conec)$$

La regla [nec](#) se diferencia de [T-Intro](#) en que, para poder deducir $T p'q$ a partir de $'$, esta oración no puede ser meramente una premisa, no basta con que est hipoteti-camente asertada en un razonamiento, sino que es necesario que haya sido demostrada, que sea un teorema, que su asercion sea categorica. Y, a la inversa, lo mismo sucede con [conec](#). Luego, de suponer que la oración $\rightarrow T I$ del mentiroso es verdadera, i.e., que $T I$, no puede deducirse que $\rightarrow T I$ ni viceversa; fs es consistente.

Teorema 3.40 (Friedman & Sheard). fs es consistente. ³²

A pesar del atractivo intuitivo que poseen los axiomas y las reglas de fs , la teoría es, lamentablemente, incorrecta, lo cual ha llevado a algunos a rechazarla (cf. [Leitgeb \(2007\)](#), [Barrio & Picollo \(2013\)](#)).

Proposición 3.41 (Friedman & Sheard). fs es !-inconsistente.

Demostración. Sea f un símbolo de función diádica que representa en q^+ la función recursiva que asigna a cada par hn ; $\#(')i$ la fórmula que resulta de iterar n veces el predicado $T a'$, i.e., en q^+

$$f(0; p'q) := p'q$$

$$f(Sx; p'q) := pT \text{ num}(f(x; p'q)) q$$

Aplicando el lema diagonal fuerte a $\exists x T f(x; y)$, se obtiene la oración de McGee, que dice que no toda iteración de T sobre ella misma es verdadera, i.e.,

$$q^+ \quad \neg m = p: \exists x T f(x; m) q \quad (3.7)$$

³²Vease [Halbach \(1994\)](#) para una prueba.

Este enunciado es una suerte de !-mentiroso que, bajo ciertos principios para la verdad como los que valida fs lleva a !-inconsistencias, a la paradoja de McGee (cf. [McGee \(1985\)](#)). Razonando en fs , supongase que $\exists x T f(x; m)$. Luego, por un lado, $\exists x T f(x + 1; m)$ y, por otro, $T f(0; m)$ o, lo que es lo mismo, $T \neg \exists x T f(x; m)$. De esto último, junto con los axiomas fs_2 y fs_4 , se sigue que $\exists x T f(x + 1; m)$, una contradicción. En consecuencia, el supuesto inicial es incorrecto; se debe concluir que $\exists x T f(x; m)$ es un teorema. Aplicando [nec](#) y el axioma fs_3 ! veces sobre este teorema, en virtud de (3.7) se tiene que $T f(0; m)$, $T f(1; m)$, $T f(2; m)$, etc, i.e., una !-inconsistencia. \square

A diferencia de kf , como establece el axioma fs_2 , a rmar la falsedad de una oración ' en fs ($T \neg p : q$) es equivalente a a rmar su no verdad ($\neg T p : q$). En consecuencia, las lógicas interna y externa de fs s coinciden y, por lo tanto, fs es incompatible con [T-Out](#).

Observación 3.42. $fs + \text{T-Out}$ es inconsistente.

Demostración. Porque la instancia de [T-Out](#) para la oración $\neg T \neg$ del mentiroso, $T \neg : T \neg$, implica $\neg T \neg$ que, junto con [nec](#), implica a su vez que $T \neg$.³³ \square

Sin embargo, fs tiene el mismo poder deductivo que $rt_{<}$, no es tan poderosa como kf o $putb$, que prueban los mismos enunciados de L^+_{pa} que $rt_{<}$.

Teorema 3.43 (Halbach). Si ' es una oración de L^+_{pa} ,³⁴

$$fs \vdash ' , rt_{<} \vdash ' :$$

Este resultado implica también que la incorrección de fs no afecta las fórmulas de L^+_{pa} sino que viene únicamente por el lado de la verdad, porque $rt_{<}$ es correcta con respecto a N^+ .

Como $ct(pa^+)$ fs , esta teoría prueba todas las instancias del [Esquema-T](#) para oraciones de L^+_{pa} . Además, por las reglas [nec](#) y [conec](#), es capaz también de demostrar todas las instancias dadas por oraciones que son verificables o refutables, v.g., los axiomas mismos y sus negaciones. En consecuencia, es claro que fs no es una teoría tipeada de la verdad, que logra trascender la distinción tarskiana entre lenguaje y metalenguaje permaneciendo en el marco de la lógica clásica, aunque también cae en !-inconsistencias.

³³ La incompatibilidad de [nec](#) y [T-Out](#) se conoce como 'paradoja de Montague' (cf. [Montague \(1963\)](#)).

³⁴ Véase [Halbach \(1994\)](#) para una prueba.

3.4. Verdad transparente paraconsistente

Si se desea un predicado de verdad transparente y, por ende, se está dispuesta a abandonar el ámbito de la lógica clásica, las lógicas paracompletas no son el único refugio. Las lógicas paraconsistentes, si bien no logran evitar que se deriven contradicciones a partir de las oraciones paradójicas, como la oración del mentiroso, sí bloquean el pasaje de una contradicción a la absoluta trivialidad. Aquellos que prefieren la paraconsistencia a la paracompletitud como lógica base para una teoría de la verdad sostienen que no hay nada malo en aceptar ciertas contradicciones, ya que algunas de ellas son verdaderas y, en consecuencia, aceptarlas no lleva a trivialidad. Se los conoce como 'dialectistas'.

El texto dialéctista fundamental se debe a [Priest \(1987\)](#), quien introduce la lógica paraconsistente lp , por 'Logic of Paradox', originalmente desarrollada de modo independiente por Florencio G. Asenjo y Newton da Costa (en sus tesis doctorales de 1954 y 1963, respectivamente; cf. [Asenjo \(1966\)](#) y [da Costa \(1974\)](#)), ciertas ampliaciones y una teoría de verdad transparente basada en una de estas ampliaciones. Posteriormente, [Priest \(2006b\)](#) y [Ripley \(2012a,b\)](#), entre otros,³⁵ presentaron sus propias variaciones sobre lp y teorías transparentes de verdad. En esta sección introduzco primeramente lp ,³⁶ seguida de una teoría semántica de la verdad transparente formulada sobre ella. Luego presento las variaciones sobre lp de Priest y Ripley y las teorías semánticas de la verdad más actuales defendidas por ellos en las secciones [3.4.2](#) y [3.4.3](#), correspondientemente.³⁷

3.4.1. Verdad transparente en la lógica paraconsistente básica

3.4.1.1. La lógica de las paradojas

La lógica lp , al igual que las lógicas paracompletas $k3$ y $b3$, trabaja con modelos no clásicos. No obstante, la lectura de estos modelos que desde una perspectiva paraconsistente varía de la paracompleta. Dado un predicado n -ádico R y un modelo M del lenguaje de primer orden L , $R^M = \{ \langle \dots \rangle \in M^n \mid R(\dots) \}$; R^M es al igual que antes, pero R^{M+} ya no es interpretada como la extensión de R en M sino que $\{ \langle \dots \rangle \in M^n \mid R^M(\dots) \}$ lo es (i.e., el conjunto de n -tuplas de objetos del dominio que no pertenecen a R^M). Del mismo modo, $\{ \langle \dots \rangle \in M^n \mid \neg R^M(\dots) \}$ es la antiextensión de R en lugar de R^M . Desde un punto de vista paraconsistente, la extensión y antiextensión pueden solaparse [v.g., aquellos objetos del dominio que no pertenezcan ni a R^{M+} ni a R^M estarán tanto en la extensión como

³⁵Vease también [Alan Weir \(2005\)](#), [Walter Carnielli, Marcelo Coniglio, & Joao Marcos \(2007\)](#) y [Beall \(2009\)](#).

³⁶ Mi presentación diverge técnicamente de, por ejemplo, la de [Priest \(2006b, cap. 5\)](#), pero ambas son extensionalmente equivalentes.

³⁷ Para otras teorías como la que introduce Beall en [2009](#), vease [Diego Tajer \(2014\)](#).

en la antiextension de R pero deben agotar todo el dominio del modelo, a diferencia de lo que sucede en las logicas paracompletas. Luego, R^{M+} es meramente el conjunto de objetos del dominio que estan en la extension pero no en la antiextension de R , mientras que R^M es el de aquellos objetos que estan en la antiextension pero no en la extension de este predicado. Bajo la lectura paraconsistente, los modelos no clasicos se denominan tambien 'modelos paraconsistentes', en lugar de 'modelos parciales'.

lp y $k3$ comparten tambien el mismo esquema de valuacion: Kleene fuerte (vease el apartado 3.1.1). La diferencia entre $k3$ y lp es que, mientras la primera tiene un unico valor designado, un unico modo de que una oracion sea verdadera en un modelo, la segunda tiene dos: 1 y $\frac{1}{2}$. Del mismo modo, mientras que en $k3$ el unico valor de falsedad es 0 , en lp tanto 0 como $\frac{1}{2}$ lo son. En consecuencia, en lp el valor $\frac{1}{2}$ no signi ca ya 'ni verdadero ni falso' sino, al contrario, 'verdadero y falso a la vez'. Esto se corresponde con la lectura paraconsistente de los modelos no clasicos. Si $v_{sk}^M(Rt_1; \dots; t_n) = \frac{1}{2}$, $Rt_1; \dots; t_n$ es verdadera y falsa a la vez, porque $ht^M_{t_1; \dots; t_n}$ pertenece tanto a la extension como a la antiextension de R .

Definicion 3.44 (lp). Si L es un lenguaje de primer orden, L es un conjunto de oraciones y $\vdash_{lp} A$ una oracion, ' A es consecuencia logica de Γ en lp ' ($\Gamma \vdash_{lp} A$) sii no existe un modelo no clasico M de L tal que $v_{sk}^M(\Gamma) \neq 0$ (i.e., $M \models_{lp} \Gamma$) y $v_{sk}^M(A) = 0$ ($M \not\models_{lp} A$).

Luego, si una oracion es tanto verdadera como falsa en un modelo, lo mismo sucedera con su negacion; y si en un condicional tanto el antecedente como el consecuente reciben valor $\frac{1}{2}$, otro tanto ocurre con el condicional. Como consecuencia, teniendo en cuenta la interdefinibilidad de la disyuncion en terminos de la negacion y el condicional, el Tercero Excluido es verdadero en todo modelo paraconsistente en lp . En general, todas las verdades logicas de la logica clasica resultan asimismo verdaderas en lp , aunque a veces tambien falsas.³⁸ Por otra parte, la regla $\vdash_{lp} A \rightarrow B$, lo que es lo mismo, el metateorema de la deduccion, es valida. Ademias, como un enunciado universal cuyas instancias reciben valor mayor o igual a $\frac{1}{2}$ tambien le es asignado un valor mayor o igual a $\frac{1}{2}$, en virtud de la definibilidad del cuantificador existencial a partir de la negacion y el cuantificador universal, la regla \exists tambien es valida en lp .

No obstante, algunas inferencias validas en logica clasica fallan en lp . Como se quer a, la regla Ex falso sequitur quodlibet

$$\frac{\Gamma \wedge \perp}{\Gamma \wedge A} \quad (\text{EFSQ})$$

³⁸ Para una exposicion detallada vease Asenjo (1966), Priest (1979, 2006a,b, 2008).

no preserva verdad en los modelos no clásicos bajo sk en lp . Porque si en un modelo M $v_{sk}^M(\perp) = \frac{1}{2}$ y $v_{sk}^M(\top) = 0$, $\perp \wedge \top$ (i.e., \perp) también recibe $\frac{1}{2}$ y, por tanto, $M \models_{lp} \perp \wedge \top$, pero $M \not\models_{lp} \perp$. Luego, en lp una contradicción no implica cualquier cosa; en caso de que una contradicción fuese un teorema de una teoría sobre lp , la teoría no sería necesariamente trivial.

Algunas otras fallas de reglas clásicas no son tan deseables para algunos defensores de la paraconsistencia. Por ejemplo, la regla [E1](#), el Modus Ponens, resulta inválida en lp . Si en un modelo M $v_{sk}^M(\perp) = \frac{1}{2}$ y $v_{sk}^M(\top) = 0$, por el esquema sk para el condicional (definición [3.2](#)), $v_{sk}^M(\perp \rightarrow \top) = \frac{1}{2}$ y, por ende, $M \models_{lp} \perp \rightarrow \top$, pero $M \not\models_{lp} \perp$. Otras reglas intuitivamente válidas como el silogismo disyuntivo también resultan inválidas en lp .

Al igual que $k3$ y $b3$, lp es correcta y completamente axiomatizable mediante diversos cálculos alternativos. Dado que en esta sección deseo considerar la posibilidad de abandonar no solo reglas y principios lógicos que conciernen a las constantes lógicas sino también reglas estructurales, propias exclusivamente de la noción de consecuencia lógica, opto en lo que sigue por cálculos de secuentes al estilo de Gerhard Gentzen, donde esto es posible (cf. [Greg Restall \(2000\)](#)). \cdot representa la echa de secuentes.

Con respecto a un cálculo de secuentes para lp , cabe destacar la validez de la regla estructural Cut:

$$\frac{\cdot \vdash ; \top \quad \cdot \vdash ; \perp}{\cdot \vdash ;} \quad (\text{Cut})$$

esto es, Cut establece la transitividad de la relación de consecuencia lógica de lp . Finalmente, al igual que en $k3$ y $b3$, en lp \models es un predicado lógico que solo admite su interpretación usual, clásica. ³⁹

3.4.1.2. Verdad transparente en LP

Para evitar ambigüedades opto por una presentación axiomática de las teorías paraconsistentes de la verdad que quiero considerar en esta investigación. Naturalmente, los modelos correspondientes a estas teorías continúan jugando un rol central, son un medio para establecer la no trivialidad o consistencia absoluta de los sistemas. ⁴⁰

La teoría paraconsistente más sencilla, lp_{tt} (por lp Transparent Truth) (cf. [Beall \(2009, x1.4\)](#)), es la que resulta de cerrar lp junto con cierta teoría de la sintaxis bajo las

³⁹ Este no es necesariamente el caso en toda lógica paraconsistente. La relación de identidad también admite interpretaciones dialécticas en algunos sistemas. Véase, por ejemplo, [Beall \(2013\)](#).

⁴⁰*Al estar formuladas sobre lógicas que permiten que ciertas contradicciones sean (i.e., reciban valor $\frac{1}{2}$), con frecuencia las teorías de la verdad paraconsistentes no son simples pero sí absolutamente consistentes.*

reglas que constituyen el principio de [Intersustitutividad](#), expresadas ahora en el calculo de secuentes:⁴¹

$$\begin{array}{l} ; ' \dots ' ; \\ ; T p'q \dots T p'q ; \end{array} \quad (IS)$$

De nicion 3.45 (lptt). lptt L_T es la teor a que extiende un calculo para lp con [IS](#) e implica todos los teoremas de pat.⁴²

Como el unico predicado involucrado en L_{pa}^+ es la identidad, todos los modelos de lptt deben interpretar esta porcion de L_T de modo clasico. La existencia de estos modelos queda garantizada por el teorema [3.10](#) de Kripke. Nuevamente, si L_T es un conjunto de oraciones, sea hN^+ ; i el modelo paraconsistente de L_T que extiende N^+ asignando a T el par ordenado h ; $f \in L_T : : f \in L_T$ gi como su interpretacion.

Corolario 3.46. Existen conjuntos de oraciones L_T tales que hN^+ ; i lp lptt.

Demostracion. Como todo modelo paraconsistente es un modelo de lp y cualquier hN^+ ; i extiende N^+ , resta probar unicamente que existe un tal que las reglas que conforman el principio de [IS](#) preservan verdad (i.e., 1 o $\frac{1}{2}$) en hN^+ ; i. Sea tal que $sk(\) =$ (cf. de nicion [3.1](#)). Por la proposicion [3.12](#), $v_{sk}^{hN^+; i}(T p'q) = v_{sk}^{hN^+; i}(')$ para toda oracion $'$ del lenguaje. \square

Luego, todos los modelos no clasicos de punto jo sobre sk son modelos de lptt. En ellos, las oraciones que, desde el punto de vista de la paracompletitud resultaban no ser ni verdaderas ni falsas, ahora son verdaderas y falsas al mismo tiempo, como la oracion $:T I$ del mentiroso; mientras que las expresiones que adquieren valores clasicos de verdad los mantienen.

Corolario 3.47. lptt cuenta entre sus teoremas con todas las instancias del [Esquema-I](#).

Demostracion. Directamente de [IS](#), porque en lp la regla [II](#) es valida. \square

3.4.2. La teor a de la verdad de Priest

Si bien algunos paraconsistentes han mostrado cierta simpat a por lptt (cf. [Bradley Dowden \(1984\)](#), [Beall \(2009, cap. 1\)](#)) la mayor a ha argumentado en contra de ella, en

⁴¹ Mejor dicho, la intersustitutividad de una oracion y su predicacion de verdad puede probarse a partir de estas reglas.

⁴² Si bien el resultado de formular los axiomas de pat sobre lp no es una teor a que contiene todos los teoremas de pat [v.g., porque algunas reglas clasicas no son validas], [Beall \(2013\)](#) muestra como complementar $lp + pat$ con una serie de reglas para que todas las consecuencias clasicas de pat se sigan tambien. Asumo que lptt cuenta con estos recursos.

tanto carece de un condicional que satisfaga la regla de Modus Ponens. Suele considerarse que para que una conectiva ! pueda llamarse propiamente 'condicional' debe satisfacer tanto Identidad como EI, a diferencia del condicional de I_p .

Any conditional worth its salt [. . .] should satisfy the modus ponens principle [. . .]. This is, indeed, analytically part of what implication is. Yet this principle fails for material implication [!] as we saw. [Modus Ponens] is not, in general, true. Hence, material implication is not the conditional. (Priest, 2006b, p. 83)

Ademas, de acuerdo con Priest, el Esquema T establece el significado del predicado veritativo y, por ende, las equivalencias que se empleen en su formulacion no pueden ser meramente materiales sino intensionales; deben satisfacer Modus Ponens para permitir el uso del predicado veritativo (vease el apartado 5.2), i.e., la inferencia de la predicacion de verdad de una oracion a la oracion misma; y no deben ser contraponibles, esto es, si una oracion ' implica , no debe seguirse necesariamente que : implique :' porque, en palabras del mismo Priest (2006b, p. 79), \There seems to be no reason why, in general, if ' is a dialetheia, T p'q is too. If ' is a dialetheia, T p'q is certainly true, but it might be simply true, and not also false."

En consecuencia, al igual que Field en el caso de las logicas paraconsistentes, muchos paraconsistentes han buscado suplementar I_p con nuevos condicionales (cf. Priest (2006b), Beall (2009), Federico Pailos & Lucas Rosenblatt (2014)) que, en principio, di eran del condicional material y, centralmente, satisfagan Modus Ponens y la regla de introduccion del condicional. Lamentablemente, la tarea no es sencilla a causa de paradojas como la de Curry (cf. Haskell B. Curry (1942), Pailos (2014)). Si L es un lenguaje cualquiera de primer orden, sea L el resultado de expandir L con un nuevo operador binario .

Proposicion 3.48. Sea I una extension de I_p formulada en L_T donde las siguientes reglas resultan validas:

$$\frac{; ' \quad ;}{; ' \quad ;} \quad (I)$$

$$; ; ' \quad ; \quad (E)$$

o Modus Ponens para , y

$$\frac{; ' \quad (')}{; ' \quad ;} \quad (Contr)$$

o Contraccion para , y sea th I_{ptt} una teoria formulada sobre I . Luego, para toda ' $2 L_T$, th \vdash '.

Demostración. Sea ϕ una oración cualquiera de L_T . Como $\phi \vdash \text{Ip}\phi$, en virtud del lema fuerte de diagonalización (teorema 1.8) aplicado a $Tx\phi$ sobre la variable x , existe un término c de L_{pa}^+ tal que

$$\text{th } \vdash \text{Ip}c = \text{pT}c\phi \tag{3.8}$$

$Tc\phi$ es una oración que dice de sí misma que implica ϕ , una oración de Curry si ϕ es una falsedad. Luego, en th (la omito como premisa),

$$\frac{\frac{\frac{Tc\phi}{Tc\phi} \text{ (I)}}{Tc\phi} \text{ (IS)}}{\frac{\frac{Tc\phi}{Tc\phi} \text{ (I)}}{Tc\phi} \text{ (Contr)}} \quad \frac{\frac{\frac{Tc\phi}{Tc\phi} \text{ (IS)}}{Tc\phi} \text{ (L)}}{\frac{\frac{Tc\phi}{Tc\phi} \text{ (IS)}}{Tc\phi} \text{ (Contr)}} \quad \frac{\frac{\frac{Tc\phi}{Tc\phi} \text{ (IS)}}{Tc\phi} \text{ (Contr)}}{Tc\phi} \text{ (E)} \quad \frac{\frac{\frac{Tc\phi}{Tc\phi} \text{ (IS)}}{Tc\phi} \text{ (Contr)}}{Tc\phi} \text{ (Cut)}$$

□

En consecuencia, un condicional que valide las reglas de introducción y eliminación y contraiga sobre Ip no puede convivir con un predicado de verdad que satisfaga las reglas que conforman el IS sin tornar la teoría trivial, algo que ni siquiera una paraconsistente aceptaría. Esta derivación de trivialidad se conoce como 'paradoja de Curry'.

La propuesta de Priest (2006b) consiste entonces en extender Ip con un nuevo condicional que satisface las reglas de introducción y eliminación usuales pero no contrae. A continuación introduzco la lógica resultante y en el apartado subsiguiente una teoría de la verdad transparente formulada sobre ella.

3.4.2.1. La lógica de Priest

Como se desea que sea un condicional que satisfaga las reglas I y E, no podrá ser el condicional material sino que, al contrario, deberá tener un carácter intensional. En lugar de cláusulas composicionales, Priest utiliza mundos posibles para dar el significado de Ip . Luego, los modelos de los lenguajes L cuentan ahora con un conjunto de mundos posibles y una relación de accesibilidad entre ellos.

Definición 3.49 (Modelos de Priest). Si M es un modelo no clásico de un lenguaje de primer orden L , M^P es un modelo de Priest si extiende M con un conjunto W_{MP} de mundos posibles y una relación binaria $R_{MP} \subseteq W_{MP} \times W_{MP}$ entre mundos posibles, tal que, para todo $w \in W_{MP}$, existe un w^0 tal que hw^0 ; $w_i \in R_{MP}$ (R_{MP} es sobreyectiva).

El esquema de valuacion de lp, sk , debera extenderse ahora a formulas que con-tengan el nuevo condicional .

Definicion 3.50 (Esquema de Priest). Sea L un lenguaje de primer orden. El esquema de valuacion de Priest ep para L es el conjunto de funciones valuacion $v_{ep}^M : L \rightarrow \{1, 0, \frac{1}{2}\}$ para cada modelo de Priest M de L tales que $v_{ep}^M(\phi) = \min_{w \in W_M} v_{ep}^w(\phi)$, donde, para cada $w \in W_M$,

$$v_{ep}^w(t = s) = \begin{cases} 1 & \text{si } t^M = s^M \\ < 0 & \text{en caso contrario} \end{cases}$$

$$v_{ep}^w(\phi) = \begin{cases} 1 & \text{si } \exists t_1^M, \dots, t_n^M \in R^{M+} \end{cases}$$

$$v_{ep}^w(\phi) = \begin{cases} > \\ > \\ < \frac{1}{2} \\ > \frac{1}{2} & \text{en caso contrario} \\ \vdots \end{cases}$$

- $v_{ep}^w(\phi) = 1 - v_{ep}^w(\neg\phi)$
- $v_{ep}^w(\phi \rightarrow \psi) = \max(1 - v_{ep}^w(\phi), v_{ep}^w(\psi))$
- $v_{ep}^w(\phi) = \frac{1}{2}$ sii, para todo w^0 tal que $w \in R_M$, w^0 , si $v_{ep}^{w^0}(\phi) = \frac{1}{2}$, $v_{ep}^{w^0}(\psi) = \frac{1}{2}$

$$v_{ep}^w(\exists x \phi) = \min_{t \text{ es un termino cerrado}} v_{ep}^w(\phi[t=x])$$

Notese que para las formulas de L que no contienen el esquema ep se comporta

exactamente igual que sk . Notese tambien que la anteultima clausula para las funciones v_{ep}^w hace que para cada mundo posible w en un modelo no haya una unica funcion v_{ep}^w sino varias. Por ejemplo, si en todo w^0 al cual w accede se cumple que, si $v_{ep}^{w^0}(\phi) = \frac{1}{2}$, $v_{ep}^{w^0}(\psi) = \frac{1}{2}$, entonces $v_{ep}^w(\phi \rightarrow \psi)$ puede valer tanto 1 como $\frac{1}{2}$.

La logica basada en modelos de Priest y el esquema de valuacion ep que interpreta los valores de verdad 1; 0 y $\frac{1}{2}$ igual que lp es lo que denomino 'logica de Priest', pl .

Definicion 3.51 (pl). Si L es un lenguaje de primer orden, L un conjunto de oraciones y $\phi \in L$ una oracion, ϕ es consecuencia logica de Γ en pl ($\Gamma \vdash_{pl} \phi$) sii no existe un modelo de Priest M de L tal que $v_{ep}^M(\Gamma) = 0$ (o $M \not\models_{pl} \Gamma$) y $v_{ep}^M(\phi) = 0$ (o $M \not\models_{pl} \phi$).

Como quer a Priest, las reglas L y E son validas en pl . Para el caso de L , notese que, si $\Gamma \vdash_{pl} \phi$, entonces no hay un modelo M de Priest y un mundo $w \in W_M$ tales que $v_{ep}^w(\Gamma) = \frac{1}{2}$ y $v_{ep}^w(\phi) = 0$. Luego, para cualquier $w \in W_M$, $v_{ep}^w(\Gamma) = \frac{1}{2}$, i.e., $M \models_{pl} \Gamma$ para todo modelo M de Priest. Asimismo, si $M \not\models_{pl} \Gamma \wedge \phi$,

para todo $w \in W_M^{ep}$, $v_M^{ep}(w) = 1$ y $v_M^{ep}(w) = 1$, con lo cual, para cualquier $w \in W_M^{ep}$ tal que $v_M^{ep}(w) = 1$, $v_M^{ep}(w) = 1$. Como R_M es sobreyectiva, $M \models \phi$, esto es, vale E.

Adicionalmente, el condicional no es contraponible y Contr resulta inválida.⁴³ Para ver lo primero, notese que un condicional $\phi \rightarrow \psi$ tal que $v_M^{ep}(\phi) = 1$ y $v_M^{ep}(\psi) = 0$ es verdadero en M , mientras que su contraposición $\neg \psi \rightarrow \neg \phi$ es tal que $v_M^{ep}(\neg \psi) = 1$ y $v_M^{ep}(\neg \phi) = 0$. Luego, por la sobreyectividad de R_M , existen mundos $w; w \in W_M^{ep}$ tales que $v_M^{ep}(w) = 1$ y $v_M^{ep}(w) = 0$ y, por ende, $v_M^{ep}(\phi \rightarrow \psi) = 1$ y $v_M^{ep}(\neg \psi \rightarrow \neg \phi) = 0$. En consecuencia, el condicional contrapuesto es falso en M .

Del mismo modo, Contr falla. Supongase que ϕ y ψ no contienen \rightarrow , y considerese un modelo M de Priest tal que $W_M = \{w_1, w_2\}$ y $R_M = \{(w_1, w_1), (w_2, w_2), (w_1, w_2)\}$.⁴⁴ Aun así, supongase que $v_M^{ep}(w_1) = 0$, $v_M^{ep}(w_2) = 1$ y $v_M^{ep}(w_2) = 0$. Luego, todo mundo w de M es tal que, si $v_M^{ep}(w) = 1$, entonces $v_M^{ep}(w) = 1$ (w_1 falsifica el antecedente de este condicional mientras que w_2 verifica trivialmente el consecuente, porque w_2 no accede a ningún mundo) y, por ende, $M \models \phi \rightarrow \psi$. Ahora bien, como $v_M^{ep}(w_1) = 0$ y, por tanto, $M \not\models \psi$.

La falla de Contr va a permitir dar en PL una teoría de la verdad que contenga todas las instancias del Esquema T formuladas con un condicional que satisface las reglas L y E sin llevar a trivialidad, como se verá en el apartado que sigue.

3.4.2.2. Verdad transparente en PL

Una de las razones principales por las cuales Priest quiere un condicional intensional que satisfaga las reglas de introducción y eliminación usuales es para formular con el Esquema T en el lenguaje de una lógica paraconsistente y, de este modo, poder realizar ciertas inferencias que involucran el predicado veritativo de modo esencial (véase el apartado 5.2 y Priest (2006b, x4.2)). En efecto, Priest no considera el Esquema-T sino el siguiente principio formulado en L_T para cada oración ϕ de este lenguaje:

$$T \vdash \phi \rightarrow \phi \quad (\text{Esquema-T de Priest})$$

donde ϕ abrevia $\phi \wedge \phi$ (que a su vez abrevia una fórmula ψ en la cual ψ no ocurre sino que esta dada en términos de \neg y \rightarrow).

Definición 3.52 (pltt). pltt L_T es la teoría que extiende un cálculo para PL con todas las instancias del Esquema-T de Priest e implica todos los teoremas de pl .

⁴³ Esto es posible porque Priest (2002) probó la no trivialidad de pl .
⁴⁴ Notese que R_M es sobreyectiva, como requiere la definición 3.49.

pltt, por Priest Logic Transparent Truth puede ser extraída de [Priest \(2006b\)](#), pero jamás ha sido explícitamente formulada. En consecuencia, tampoco se conoce una prueba de consistencia. No obstante, es sencillo ver como esta teoría evita la paradoja de Curry presentada al comienzo de esta sección. Como [Contr](#) no vale en pl, es claro donde la derivación de cualquier oración \perp es bloqueada, pero esto no es suficiente para ver como la lógica de Priest acomoda el fenómeno de Curry.

Sea \perp una falsedad, v.g., $0 \neq 0$. De acuerdo con [\(3.8\)](#)

$$\text{pltt} \vdash_{\text{pl}} \text{c} = \text{pT c} \quad \perp \text{q}$$

Claramente, $\text{T c } \perp$ no puede ser verdadera en un modelo M de Priest tal que $M \models_{\text{pl}} \text{pltt}$, porque, en ese caso, para todo $w \in W_M$, $v_{\text{ep}}^w(\text{T c } \perp) = \frac{1}{2}$ y, por el [Esquema-T de Priest](#), $v_{\text{ep}}^w(\text{T c}) = \frac{1}{2}$ también, con lo cual en todo mundo w que acceda a algún otro $v_{\text{ep}}^w(\perp) = \frac{1}{2}$, lo cual es imposible. Luego, $\text{T c } \perp$ debe ser falsa en todo modelo. En efecto, esto es perfectamente posible en los modelos de Priest para pltt (recuérdese que para que una oración sea falsa en un modelo basta que lo sea en un mundo posible).

Supongase que M es un modelo tal, $w_1; w_2 \in W_M$, $w_1 R_M w_2$ y w_2 no accede a ningún mundo posible en M . Aun más, supongase que $v_{\text{ep}}^{w_2}(\text{T c}) = \frac{1}{2}$ y $v_{\text{ep}}^{w_2}(\perp) = 0$. Esto no es un problema porque, si bien implica que $v_{\text{ep}}^{w_2}(\text{T c } \perp) = \frac{1}{2}$, lo cual queda también implicado por el hecho de que w_2 no ve ningún otro mundo, los mundos posibles no obedecen necesariamente Modus Ponens; los modelos lo hacen. Luego, $v_{\text{ep}}^{w_1}(\text{T c } \perp) = 0$ y, por tanto, $M \not\models_{\text{pl}} \text{T c } \perp$.

Finalmente, cabe destacar que pltt valida las reglas [T-Intro](#) y [T-Elim](#) pero no el principio [IS](#). Si bien esta es una consecuencia deseada por Priest, como se vio en la segunda cita al comienzo de este apartado [3.4.2](#), también es la razón principal por la cual, v.g., [Beall \(2009\)](#) critica la teoría de Priest y opta por desarrollar un condicional intensional alternativo a que valide este principio.

Proposición 3.53. Mientras que [T-Intro](#) y [T-Elim](#) son válidas en pltt, falla [IS](#), si la teoría es no trivial.

Demostración. La validez de las reglas [T-Intro](#) y [T-Elim](#) se sigue inmediatamente del [Esquema-T de Priest](#) y [E](#). La falla de [IS](#) se debe a la posibilidad de que en el [Esquema-T de Priest](#) un lado de la equivalencia obtenga valor 1 en un modelo mientras que el otro valor $\frac{1}{2}$. V.g., si M es un modelo de Priest de pltt, la oración \perp del mentiroso obtiene valor $\frac{1}{2}$ en M , pero su predicación de verdad puede valer 1. Luego, la negación de la primera, $\neg \perp$ obtiene valor $\frac{1}{2}$ en M y es, por ende, verdadera ahí, mientras que la negación de la segunda, $\neg \text{p} \perp$ obtiene valor 0. En consecuencia, la inferencia de

$\vdash T \mid a : T \mid p : T \mid q$ no preserva verdad en M. Pero esta inferencia se sigue del [IS](#):

$$\frac{\frac{\frac{\vdash T \mid . \vdash T \mid}{\vdash T \mid p : T \mid q . \vdash T \mid}}{\vdash T \mid p : T \mid q . \vdash T \mid}}{\vdash T \mid . \vdash T \mid p : T \mid q} \quad (IS)$$

□

3.4.3. Abordajes subestructurales y verdad transparente

Un tercer modo de debilitar la logica para evitar las hipotesis del teorema de Tarski (teorema [2.9](#)) y as permitir a incorporacion de un predicado de verdad transparente a una teor a de la sintaxis consiste en abandonar reglas o principios clasicos, no ya concernientes a tales o cuales operadores logicos, como el [Tercero Excluido](#) o el [EFSQ](#), sino unicamente a la nocion misma de consecuencia logica, i.e., reglas estructurales. Las logicas resultantes se conocen como 'logicas subestructurales'.

Las reglas estructurales basicas de la logica clasica son Re exividad

$$\frac{\vdash \Gamma}{\vdash \Gamma} \quad (Re\ ex)$$

Monoton a

$$\frac{\vdash \Gamma, \Delta}{\vdash \Gamma, \Delta}$$

Contraccion

$$\frac{\vdash \Gamma, \Delta, \Delta \quad \vdash \Gamma, \Delta}{\vdash \Gamma, \Delta} \quad (Mon)$$

$$\frac{\vdash \Gamma, \Delta \quad \vdash \Gamma, \Delta}{\vdash \Gamma, \Delta} \quad (Contr)$$

y Transitividad o [Cut](#) (vease el apartado [3.4.1.1](#)).

Las logicas subestructurales que se emplean con mas frecuencia en la formulacion de teor as de verdad transparente son aquellas que abandonan o bien [Cut](#) o bien [Contr](#). Naturalmente, abandonar la regla de Contraccion presupone que los secuentes no esten dados en terminos de conjuntos sino otro tipo de entidades colectivas a las cuales un objeto pueda pertenecer mas de una vez, v.g., secuencias. En vista al apartado [5.3](#), considero exclusivamente la logica subestructural intransitiva st, por Strict Tolerant, que rechaza Transitividad pero no Contraccion u otras reglas estructurales como Monoton a o Re exividad;⁴⁵ y la teor a de verdad transparente sttt formulada en st.

⁴⁵Para mas detalles acerca de las logicas subestructurales, veanse [Restall \(2000\)](#) y [Francesco Paoli \(2002\)](#). Para un ejemplo de teor a de la verdad transparente formulada sobre una logica subestructural en la cual no vale [Contr](#), vease [Elia Zardini \(2011\)](#).

3.4.3.1. Una lógica intransitiva

La lógica intransitiva que introduzco en este apartado, originalmente desarrollada por [Robert van Rooij \(2011\)](#), trabaja con modelos no clásicos (de noción 3.1) y los interpreta paraconsistentemente, esto es, si R es un predicado n -ádico de un lenguaje de primer orden L y M es un modelo no clásico de M , la extensión y la antiextensión que M asigna a R no son R^{M+} y R^M sino $\{M\}^n R^M$ y $\{M\}^n R^{M+}$, respectivamente (vease el apartado 3.4.1.1). Aun más, el esquema de valuación en función del cual esta lógica está definida es el mismo sk . La diferencia entre st y lp (de noción 3.44) yace en el modo en el cual la noción de consecuencia lógica está definida.

Definición 3.54 (st). Si L es un lenguaje de primer orden, Σ es un conjunto de oraciones y $\phi \in \Sigma$ una oración, ϕ es consecuencia lógica de Γ en st ($\Gamma \vdash_{st} \phi$) si y sólo si no existe un modelo no clásico M de L tal que $v_{sk}^M(\Gamma) = 1$ y $v_{sk}^M(\phi) = 0$ (i.e., $M \not\models_{st} \phi$).

En st tanto 1 como $\frac{1}{2}$ representan verdad, aunque el último valor también representa falsedad, pero la noción de consecuencia lógica no está definida como preservación de verdad. De hecho, st no preserva ningún valor de verdad, lo cual permite la falla de transitividad. Lo que debe ocurrir para que un argumento resulte válido en st es que, si las premisas son estricta o absolutamente verdaderas [i.e., valen 1], la conclusión debe ser tolerantemente verdadera [i.e., puede valer 1 o $\frac{1}{2}$, pero no 0]. En st hay una diferencia sustancial entre la aserción hipotética de un conjunto de premisas Γ , esto es, $\Gamma \vdash_{st} \phi$ y su aserción categórica, i.e., $\Gamma \models_{st} \phi$.

Al igual que en lp , todas las verdades lógicas de la lógica clásica, v.g., el [Tercero Excluido](#), resultan verdaderas en st . Nótese que, de acuerdo con la definición 3.54, ϕ si y sólo si ϕ recibe valor 1 o $\frac{1}{2}$ en todo modelo no clásico. Por ende, ϕ es una verdad lógica de st si y sólo si ϕ es una verdad lógica de lp . Pero, a diferencia de lp , las inferencias de la lógica clásica también resultan válidas en st , porque si existe un modelo no clásico en el cual las premisas de un argumento reciben valor 1 y la conclusión 0, también hay un modelo clásico en el cual esto ocurre (cf. [Ripley \(2012b\)](#), lema 1), razón por la cual sus defensores con frecuencia la consideran una lógica clásica.

No obstante, como señalan [Barrio et al. \(forthcoming\)](#), st no es lógica clásica. La diferencia radica en que algunas metainferencias clásicas fallan en esta lógica intransitiva. Mientras que en un cálculo de secuentes las inferencias toman la forma de secuentes, i.e.

las metainferencias permiten inferir secuentes de otros secuentes. Por ejemplo, la inferencia

$$\frac{}{; ? .}; \quad (\text{EFSQ})$$

ahora expresada en forma de secuentes, resulta valida en st porque una contradiccion no recibe valor 1 en ningun modelo no clasico; mientras que la metainferencia correspondiente

$$\frac{. ?;}{. !}; \quad (\text{EFSQ}^c)$$

donde ? no es asertada hipotetica sino categoricamente, es valida en logica clasica pero invalida en st. Porque aunque implique logicamente una contradiccion (en caso en que = ?), puede ocurrir que en algun modelo reciba valor 1 y ? $\frac{1}{2}$. Como ' es cualquier oracion, basta con elegir una falsedad para que la conclusion del [EFSQ^c](#) no se siga. Por esta razon y porque st se de ne en terminos de modelos paraconsistentes clasico esta logica como paraconsistente, aunque estar a dispuesta tambien a asignarle una categoria por fuera de la oposicion paracompleta-paraconsistente.

Asimismo, si bien la regla Modus Ponens

$$\frac{}{; ' ! .}; \quad (\text{E!})$$

ahora expresada en secuentes, es valida en st, porque solo hay que evaluarla en los casos en los cuales tanto ' como reciben valor 1, su version categorica, de metainferencia

$$\frac{. ' . !;}{. ;}; \quad (\text{E!}^c)$$

que es valida en logica clasica, resulta invalida en st. Porque, si = ?, es posible un modelo en el cual todos los miembros de reciben valor 1, ' y, por ende, tambien ' ! , reciben valor $\frac{1}{2}$, mientras que valor 0.

Sin embargo, no todas las metainferencias clasicas fallan en st. V.g., la regla de introduccion del condicional

$$\frac{}{; ' .;}; \quad (\text{!})$$

ahora en terminos de secuentes, es valida, porque, si no hay un modelo en el cual ' valga 1 y 0, de acuerdo con sk el condicional ' ! no vale 0 en ningun modelo tampoco. Luego, st no enfrenta los problemas de las logicas paracompletas como k3 y b3 o las paraconsistentes como lp con respecto a los condicionales: tanto [!!](#) como [E!](#), e [Identidad](#), valen.

Usualmente, cuando se ofrece un calculo de secuentes para la logica clasica, no se incluye en la presentacion la regla estructural [Cut](#) (cf. [Boolos et al. \(2007, cap. 14\)](#)) porque esta se deriva directamente de las otras reglas. Sin embargo, si se incorporan nuevas reglas o axiomas no logicos [Cut](#) puede dejar de valer (cf. Pablo Cobreros, Paul Egge, Ripley & Van Rooij (2013)). Por ende, si se desea permanecer en un sistema clasico, es preciso agregarla como una regla primitiva.

En consecuencia, si se quiere un calculo para st es posible trabajar directamente con un calculo clasico en el cual la regla [Cut](#) no haya sido explícitamente agregada (cf. [Ripley \(2012a\)](#)), aunque existen otros calculos de secuentes correctos y completos, dados, por ejemplo, por secuentes de tres lados (cf. [Ripley \(2012b, x4.2\)](#)). En el apartado que sigue presento una teoría de la verdad transparente formulada sobre un calculo clasico sin [Cut](#).

3.4.3.2. Verdad transparente en ST

La falla de transitividad de st permite incorporar a un calculo para esta logica todos los teoremas de pat junto con principios de verdad transparente sin llevar a trivialidad.

Definición 3.55 (sttt). sttt L_T es la teoría que extiende el calculo de secuentes de la logica clasica (sin [Cut](#)) con los secuentes \cdot para cada axioma \cdot de pat y esta cerrada bajo las reglas que conforman el [IS](#).

Como st valida todos los teoremas e inferencias de la logica clasica y q^+ sttt, sttt, por Strict Tolerant Transparent Truth, prueba la oración $\cdot T \mid$ del mentiroso en (3.5):

$$\frac{\frac{\cdot T \mid, \cdot T \mid (IS)^\top}{\cdot \cdot T \mid; \cdot T \mid}}{\cdot \cdot T \mid}$$

Ademas, en sttt el mentiroso implica cualquier conjunto de oraciones:

$$\frac{\frac{\cdot T \mid, \cdot T \mid}{(IS) \cdot T \mid, \cdot T \mid}}{\cdot T \mid}$$

No obstante, nada de esto implica que sttt sea una teoría trivial, eso es, que el secuyente vacío \cdot valga en la teoría, porque esta tiene un modelo no trivial. Si el secuyente vacío fuera demostrable en sttt, por la regla estructural [Mon](#), todas las inferencias estarían permitidas en esta teoría.

Proposición 3.56 (Ripley). Existe un modelo no clásico $\mathcal{hN}; i$ de L_T tal que $\mathcal{hN}; i$ satisface sttt .⁴⁶

Es precisamente la falla de [Cut](#) la que bloquea que se infera el seciente vacío de los secientes $\vdash :T I$ y $\vdash :T I \dots$. En efecto, en los modelos no clásicos $\mathcal{hN}; i$ de sttt , $v_{sk}^{\mathcal{hN}; i}(\vdash :T I) = \frac{1}{2}$, lo cual explica la validez de los secientes $\vdash :T I$ y $\vdash :T I \dots$ pero también la falla de transitividad.

Corolario 3.57. sttt cuenta entre sus teoremas con todas las instancias del [Esquema-T](#) y valida las reglas [T-Intro](#) y [T-Elim](#).

Demostración. Porque el [IS](#) es un principio de intersustitutividad, [Identidad](#) es una verdad lógica de st y esta lógica valida [Re ex](#), i.e., es re exiva. \square

§

En este capítulo he mostrado que es perfectamente posible abandonar la distinción tarskiana entre lenguaje y metalenguaje pero que, a la vez, esto tiene sus costos: debe dejarse de lado, o bien la lógica clásica, o bien algunas instancias del [Esquema-T](#) u otros principios de transparencia. Asimismo, las teorías resultantes pueden ser demasiado complejas, restrictivas, arbitrarias o incorrectas. Por estas razones, no es sencillo optar por una de ellas sin dejar primero ciertos objetivos. En los capítulos que siguen exploro las condiciones que debe cumplir un sistema formal de verdad para poder ser adoptado por defensionistas. Si bien el resultado no es un vacío, muchas teorías que aquí introduje no son viables desde un punto de vista de defensionista y deben ser descartadas, mientras que otras resultan más o menos exitosas. En el último capítulo de esta tesis ofrezco diversas teorías de la verdad que, argumentado, se adecuan mejor a las demandas del defensionismo que las presentadas en este capítulo.

⁴⁶Vease [Ripley \(2012a\)](#) para una prueba de un resultado semejante.

4

La verdad de acionada

\. . . the relevance of a theory of truth may lie in its import regarding the irrelevance of truth."

{ Paul Horwich, Truth

El de acionismo es una postura filosófica en torno a la verdad que aglutina diversas tesis heterogéneas. Esto dificulta enormemente su exposición. Cuando se quiere hablar del de acionismo en términos generales y no de tal o cual corriente las formulaciones resultan vagas y a veces contradictorias; y a medida en que se adquiere más claridad y precisión se pierde el carácter general de la descripción. No pretendo exponer todas y cada una de las posiciones que se identifican a sí mismas con el de acionismo ni discutir todas las tesis que son usualmente asociadas a esta postura sino hacer más bien lo primero: explorar lo que podría llamarse 'núcleo común' de hipótesis de acionistas, para luego darle precisión y coherencia, y estudiar sus implicancias en relación a teorías formales de la verdad en el capítulo siguiente.

Hay al menos una tesis que es característica de la corriente de acionista, i.e., que la expresión 'es verdadera' no existe en el lenguaje natural para adscribir una propiedad sustantiva a ciertos los portadores de verdad, v.g., a aquellos que guardan cierta relación de correspondencia con el mundo, sino que solo cumple roles expresivos de índole lógica, semejantes a los de las conectivas y operadores lógicos usuales. Fundamentalmente, permite expresar ciertos conjuntos infinitos de oraciones mediante una única expresión, su 'conjunción infinita', como, por ejemplo, el conjunto de todos los teoremas de la aritmética, mediante

Todos los teoremas de la aritmética son verdaderos. (4.1)

Dentro (y también fuera) del de la oraciónismo suele considerarse que la razón por la cual el predicado de verdad es capaz de cumplir estas funciones yace en que predicar verdad de una oración es simplemente otro modo de afirmar el contenido de la oración: una oración y su predicación de verdad son siempre equivalentes. Esta equivalencia suele expresarse por medio de principios de transparencia como el [Esquema T](#) que, en consecuencia, se considera que expresan si no todo al menos parte del significado del predicado veritativo.

En este capítulo doy primero una breve historia heurística de los orígenes del de la oraciónismo: la teoría de la redundancia. Luego, analizo en detalle la función expresiva que el de la oraciónismo asigna al predicado de verdad: que la posibilita y en que consiste. Por último, la sección [4.3](#) se ocupa de los principios que expresan las características del predicado veritativo que le permiten cumplir las funciones que el de la oraciónismo le asigna, esto es, de los principios que expresan el significado de la verdad y las principales objeciones que han encontrado en la literatura.

4.1. Breve prehistoria del de la oraciónismo

Hasta comienzos del siglo XX toda investigación filosófica acerca de la verdad estaba enfocada en la búsqueda de una definición intensional de esta palabra, de la esencia o la naturaleza del concepto que ella denota. La teoría correspondentista de la verdad era la más aceptada a pesar de la oscuridad de sus diversas formulaciones. Sin embargo, nada se había dicho aún de una supuesta función expresiva de corte lógico que el predicado de verdad cumple en el lenguaje. Este rol comenzó a tornarse visible solo después de que se advirtiera explícitamente que, prima facie, el contenido de una oración A puede transmitirse igualmente predicando verdad sobre A; que el predicado de verdad no agrega nada al contenido proposicional de una oración.¹ Así surgieron los primeros (pre)de la oraciónistas, hoy llamados 'redundantistas', entre ellos Gottlob Frege, Frank P. Ramsey y Alfred J. Ayer. Por ejemplo, de acuerdo con Frege,

It is also worthy of notice that the sentence 'I smell the scent of violets' has the same content as the sentence 'it is true that I smell the scent of violets'. So it seems, then, that nothing is added to the thought [proposición] by my ascribing to it the property of truth. [. . .] The meaning of the word 'true' seems to be altogether unique. May we not be dealing here with something which cannot, in the ordinary sense, be called a quality at all? ([Frege, 1918-19](#), p. 88)

¹El contenido proposicional de una oración puede entenderse informalmente como aquello que la

oracion comparte con todas las otras que son necesariamente equivalentes a ella.

La idea no es que el predicado de verdad no agrega absolutamente nada a una oración sobre la cual se predica. Podría, en algunos casos, agregar algo al sentido de la oración, enfatizarla, presentarla de otro modo, etc. Lo que se sostiene es que, de alguna manera, la predicación de verdad no agrega nada al contenido proposicional de la oración que se dice que es verdadera sino que es equivalente a ella. Llamo a esto 'tesis de la redundancia conceptual de la verdad'.

En consecuencia, siguiendo a Frege, el predicado de verdad no parece ser un predicado como todos los otros; al no decir nada del objeto al cual se aplica, no puede denotar un concepto ordinario. La tesis según la cual la verdad directamente no es un concepto o propiedad genuina fue heredada posteriormente por varios defensores del afirmacionismo (cf. [Peter F. Strawson \(1949\)](#) y Dorothy L. Grover, Joseph L. Camp & Belnap (1975)), mientras que otros adoptaron una versión debilitada, como se verá más adelante. Junto con la tesis de la redundancia conceptual de la verdad, esta idea ubicó al afirmacionismo incipiente en las antipodas del correspondentismo y, en general, de cualquier posición que entienda la verdad como una propiedad compleja que debe poder darse intensionalmente (y no meramente de manera extensional, como hizo Tarski, capítulo 2) en términos de propiedades más simples, esto es, en términos de

$$x \text{ es verdadero} \quad \text{def } P x \quad (\text{SUST})$$

donde P es una conjunción de propiedades más simples y comprensibles que la verdad (v.g., las corrientes coherentista y pragmatista).

Las corrientes que buscan definiciones de verdad en estos términos son con frecuencia llamadas 'sustantivistas' o 'afirmacionistas'. El afirmacionismo redundante considera estéril esta búsqueda, y el problema de la verdad un pseudo problema. Ante las preguntas '¿Qué propiedades se adscriben a una oración cuando se dice que es verdadera?' o '¿Qué se está diciendo de ella?', el afirmacionismo responde 'Ninguna' y 'Nada', respectivamente. En palabras de [Ramsey \(1927, p. 157\)](#), "[...] there is really no separate problem of truth but merely a linguistic muddle." Como señala [Gupta \(1993\)](#), el afirmacionismo es una postura meta-teórica sobre la verdad, que no está al mismo nivel de las posiciones sustantivistas sino en un metanivel [por ejemplo, el de la idea sostenida por los positivistas lógicos según la cual la verdad es un concepto metafísico que debe ser eliminado de todo discurso con pretensiones de rigurosidad].

Precisamente en un intento por mostrar que la predicación de verdad no agrega en ningún contexto contenido alguno a las oraciones sobre las cuales se predica, [Ramsey \(1927\)](#) fue el primero en poner en evidencia el rol expresivo que el predicado veritativo juega en el lenguaje y su importancia, al hallar contextos en los cuales, si bien aún conceptualmente redundante, resulta ineliminable [i.e., mostrar que el predicado de verdad

no es expresivamente redundante o, en palabras de [Hartry Field \(1999\)](#), expresivamente liviano].

Como nota Ramsey, cuando el predicado de verdad se aplica a nombres explícitos o transparentes de oraciones, como en

"La nieve es blanca" es verdadera. (4.2)

es no solo conceptual sino expresivamente redundante; puede ser de alguna manera eliminado (junto con las comillas), porque el contenido de la oración puede expresarse directamente por medio de la oración que era objeto de la predicación, v.g.,

La nieve es blanca.

Su presencia en estos casos se debe exclusivamente a razones retóricas o estilísticas, para enfatizar lo que se quiere decir o presentarlo de un modo alternativo.

No obstante, el predicado de verdad, como cualquier predicado, puede aplicarse a términos de cualquier tipo sintáctico, ya sean variables, constantes o términos complejos, y no solamente a nombres explícitos de oraciones. Llamo 'adscripciones singulares de verdad' a aquellas en las cuales la verdad se aplica a términos cerrados, y 'adscripciones generales de verdad' o 'generalizaciones' cuando se aplica a términos abiertos. Por ejemplo, tanto (4.2) como

La equivalencia entre masa y energía es verdadera. (4.3)

son adscripciones singulares de verdad, mientras que en (4.1), en

Algunas ecuaciones de Maxwell son verdaderas. (4.4)

y, si se entienden las descripciones de nidas a la Russell, en

La fórmula más famosa de la física es verdadera. (4.5)

las adscripciones son generales. Notese que, al tratarse de oraciones, las variables libres bajo el alcance del predicado de verdad deben estar ligadas por un cuantificador de uno u otro modo.

Las adscripciones de verdad pueden a su vez clasificarse en explícitas e implícitas, según el término al cual se adjunte el predicado veritativo. Si este es un nombre transparente de alguna oración, la adscripción de verdad es explícita, mientras que en el

resto de los casos es implícita. Por ejemplo, (4.2) es una adscripción explícita de verdad, mientras que en (4.1), (4.3) y (4.4) las adscripciones son implícitas.

Ramsey nota que en las adscripciones implícitas de verdad ya no hay garantía de que el predicado pueda simplemente ser eliminado y la predicación reemplazada por las oraciones que son su objeto. Por ejemplo, en

Lo que dice Oscar es verdadero. (4.6)

el predicado veritativo sería eliminable solo si se supiera que es lo que dice Oscar, en pasado, presente y futuro, lo cual es físicamente imposible. Sin embargo, Ramsey argumenta, en este caso la verdad es conceptualmente tan redundante como en (4.2). La forma lógica de (4.6) está dada por

Para toda oración x ; si Oscar dice x ; entonces x es verdadera: (4.7)

Si el lenguaje permitiera no solo cuantificar sobre posición nominal [como sobre x en 'Oscar dice x ' o ' x es verdadera'] sino también en posición oracional, sería posible eliminar el predicado veritativo de (4.7) y decir simplemente

Para toda oración A ; si Oscar dice ' A ', entonces A : (4.8)

La cuantificación proposicional, como suele llamarse a la cuantificación sobre posición oracional, podría entonces tornar completamente eliminable el predicado de verdad. Efectivamente, como se señaló en el apartado 2.2, enriqueciendo previamente el lenguaje con una función que, dada una oración A cualquiera, devuelva un nombre para A [cuidando de no caer en contradicciones], es posible introducir el predicado de verdad como un término de nido. De hecho, en tiempos de Ramsey la lógica de segundo orden, que permite cuantificadores de este tipo, era lo usual, dados los trabajos de Russell y Alfred N. Whitehead en *Principia Mathematica* (cf. [Russell & Whitehead \(1910\)](#)).

No obstante, ni el español ni el inglés, ni ningún otro lenguaje natural del que tenga yo conocimiento, son capaces de cuantificar en posición oracional.² Como señala Ramsey, estos lenguajes, a diferencia de los lenguajes de segundo orden, requieren de un verbo en el consecuente de (4.8), que es exactamente lo que el predicado de verdad provee. El predicado de verdad no es eliminable en (4.6) porque está cumpliendo un rol que, al menos en este caso, consiste meramente en suplantar la cuantificación proposicional y,

²[Arthur N. Prior \(1971\)](#) y [Grover et al. \(1975\)](#) exploran extensiones del inglés con prosentencias, sintagmas semejantes a pronombres pero que a diferencia de estos ocupan el lugar de oraciones, sobre las cuales se puede cuantificar. El resultado es similar a la extensión del lenguaje natural con cuantificadores proposicionales, y el predicado de verdad es eliminable en casos como (4.6).

por ende, no es de naturaleza conceptual pero sí, de alguna manera, lógica.³ Este rol es lo que el de acciónismo posteriormente va a identificar como la única razón, importante por cierto, por la cual es necesario contar con un predicado de verdad en el lenguaje. En palabras de Horwich (1998b, p. 2), "[...] the truth predicate exists solely for the sake of a certain logical need." En la sección que sigue exploro la función lógica que el de acciónismo adscribe a la verdad.

4.2. El rol de la verdad

El primero en mencionar el rol lógico de la verdad explícitamente fue Quine, para quien justamente la lógica de segundo orden no era admisible sino solo la de primer orden, y cuya opinión está vigente todavía hoy prácticamente en toda la comunidad lógico-lingüística. Desde su punto de vista, el predicado veritativo es indispensable por su función tanto en los lenguajes naturales como en los formales. Sin tomar parte en la disputa acerca de la legitimidad de la lógica de segundo orden en tanto lógica (cf. Quine (1970), Shapiro (1991)), el de acciónismo mantiene que en lenguajes donde la cuantificación en posición oracional no es posible el predicado de verdad cumple un rol lógico importante e irremplazable.⁴

En esta sección argumento que la verdad tiene una función básica que lo caracteriza, esto es, ser un mecanismo desentrecomillador, que encapsula pero excede su redundancia conceptual, y que le permite cumplir dos funciones lógico-expresivas diferentes, a través de las cuales pueden explicarse todos los usos conocidos del predicado veritativo, todos los cuales son, asimismo, de naturaleza lógico-expresiva, no conceptual. El primer apartado se ocupa de la función básica del predicado. 4.2.2 introduce sus dos funciones lógico-expresivas y muestra como se derivan de la función básica. Finalmente, 4.2.3 examina los tipos de usos que suelen hacerse del predicado de verdad e indica como sus funciones lógico-expresivas son capaces de dar cuenta de estos usos.

4.2.1. La verdad como un mecanismo desentrecomillador

El predicado de verdad tiene diversos usos expresivos, algunos de los cuales quizás no hayan sido descubiertos aun. De acuerdo con el de acciónismo, se puede dar cuenta

³*Sin entrar en la discusión sobre que es una constante lógica y que no lo es, dados casos positivos claros, como la conjunción y el cuantificador universal, y también negativos, como la suma u otras operaciones entre números naturales, opero meramente por analogía; baso mis juicios de logicidad de la verdad en la cercanía o lejanía conceptual del predicado de verdad con respecto a otras nociones lógicas y no lógicas.*

⁴*Veáanse Field (2008) y Horwich (1998b).* También se han identificado a sí mismos como de acciónistas Beall, Robert Brandom, Arthur Fine, Grover, Horsten, Stephen Leeds, Michael Williams, Quine, Richard Rorty, entre otros.

de aquellos de los que se tiene conocimiento, incluido el que tiene en el ejemplo de Ramsey en (4.6), por medio de lo que denomino 'función básica de la verdad'. Esta función consiste en lo siguiente: el predicado veritativo transforma una oración en un término, sustituye sintácticamente una oración [A] por otra ['A es verdadera] donde la oración original es ahora el objeto de la predicación, pasa de ocupar una posición oracional a una nominal, sin alterar sus propiedades semánticas, esto es, salva veritate.⁵ Si bien sintácticamente la oración pasa a ocupar el lugar de un término, una posición nominal, desde un punto de vista semántico mantiene su posición oracional: predicar verdad de una oración da por resultado un enunciado equivalente a la oración en cuestión.

Luego, la función básica de la verdad garantiza que adscribir verdad a una oración sea un modo sintácticamente alternativo de expresar el contenido de esta oración. Esta idea suele resumirse diciendo que el predicado de verdad es un dispositivo (des)entrecorillador [en terminología quineana] o (de)nominalizador [de acuerdo con Horwich, otro exponente fundamental del de acciónismo]. Permite el ascenso a un plano lingüístico, donde las oraciones que antes eran 'sujeto' son ahora objeto de discurso. Pero, en palabras de Quine,

This ascent to a linguistic plane of reference is only a momentary retreat from the world, for the utility of the truth predicate is precisely the cancellation of linguistic reference. The truth predicate is a reminder that, despite a technical ascent to talk of sentences, our eye is on the world. (Quine, 1970, p. 12)

Si bien el predicado de verdad da sintácticamente lugar a un discurso acerca de oraciones, el contenido que transmite no es lingüístico, no es semánticamente acerca de oraciones, sino que es el de las oraciones mismas. Desde el punto de vista de Horwich (1998b, p. 5), "[...] the truth predicate serves merely to restore the structure of a sentence: it acts simply as a de-nominalizer."

Para llevar a cabo su función básica, el predicado de verdad no debe ser solo conceptualmente redundante, lo cual le permite no alterar el contenido de una oración cuando se aplica a ella, sino también un predicado, en contraste con, v.g., un operador, para poder efectivamente ascender a un plano lingüístico. >Para que se necesita un dispositivo con estas características?

⁵Muchos de acciónistas, como Horwich por ejemplo, consideran que la verdad es un predicado que se aplica a proposiciones en lugar de oraciones. Sin embargo, como señale en el apartado 2.1.2, emplear oraciones simplemente técnicamente la presentación y deja abierta la posibilidad de entender su predicación de verdad como hablando de proposiciones a través de las oraciones que las expresan, como es frecuente, [v.g., en Horwich (1997, 1998b)] mismo.

4.2.2. Las funciones lógico-expresivas de la verdad

Su naturaleza (des)entrecomilladora permite al predicado veritativo cumplir dos funciones esenciales que lo convierten en un predicado de suma utilidad, tanto en lenguajes formales como naturales. En este apartado presento estas dos funciones 'lógico-expresivas' de la verdad y muestro como dependen directamente de su función básica. En el apartado siguiente exploro los usos expresivos que suelen adjudicarse a la verdad y señalo de que manera están basados en sus funciones lógico-expresivas.

4.2.2.1. Primera función

La función básica de la verdad garantiza por sí sola que todas las adscripciones explícitas, esto es, de la forma

$$'A' \text{ es verdadera.} \quad (4.9)$$

sean equivalentes a la oración A. Pero también, junto con las leyes de la identidad, garantiza que toda adscripción singular de verdad simpliciter, esto es, toda oración de la forma

$$a \text{ es verdadera.} \quad (4.10)$$

donde a podrá no ser el nombre transparente de una oración, sea equivalente a A, provisto que 'a' denote A [i.e., que $a = 'A'$]. Por ejemplo, (4.3) expresa el mismo contenido que

$$\text{La energía en reposo de un cuerpo es el producto de su masa por la velocidad de la luz al cuadrado.} \quad (4.11)$$

porque el término 'la equivalencia entre masa y energía' denota esta oración.

La primera función lógico-expresiva de la verdad consiste entonces en posibilitar la expresión del contenido de una oración aplicándose únicamente a un nombre no transparente de esa oración. Sin la ayuda del predicado veritativo no es posible transmitir el contenido de una oración contando exclusivamente con un nombre no transparente para ella. Por ejemplo, no es gramatical decir 'la equivalencia entre masa y energía' para transmitir el contenido de (4.11). Pero anteponer el predicado de verdad, que, como diría Ramsey, provee el verbo necesario para restablecer la gramaticalidad de la expresión, el nombre de la oración es suficiente para expresar su contenido, sin la necesidad de

articular la oración misma.⁶ El predicado veritativo está ejerciendo esta función, v.g., en (4.3).

4.2.2.2. Segunda función

La segunda función lógico-expresiva de la verdad consiste en posibilitar la expresión del contenido de oraciones cuantificando sobre ellas. El predicado de verdad, como ya notó Ramsey, permite de alguna manera cuantificar sobre posición oracional, emular la cuantificación proposicional. Esta es su función más importante, por tornarlo irremplazable en lenguajes coloquiales y en los lenguajes formales más usados.

Como señala Quine (1970), se puede generalizar sobre 'Tom' en 'Tom es idéntico a Tom' simplemente diciendo 'Todo es idéntico a sí mismo' o 'Algo es idéntico a sí mismo', porque los cuantificadores del lenguaje natural permiten cuantificar sobre posición nominal. Cualquier oración B que contenga una ocurrencia de un nombre A puede generalizarse sobre A recurriendo a los cuantificadores usuales del lenguaje. Ahora bien, si A es, no un nombre, sino una suboración en B, i.e., no ocurre en posición nominal sino oracional, los cuantificadores de los cuales se dispone en el lenguaje natural y en los lenguajes de primer orden no permiten generalizar B sobre A.

Por ejemplo, no es posible generalizar sobre 'La nieve es blanca' en

Si Oscar dice 'La nieve es blanca', entonces la nieve es blanca. (4.12)

porque, como señala Ramsey (1927), (4.8) el resultado no es gramatical. El predicado de verdad permite remediar esta situación a través de su función básica, esto es, ubicando la suboración 'La nieve es blanca' en posición nominal, reescribiendo (4.12) como

Si Oscar dice 'La nieve es blanca', entonces 'La nieve es blanca' es verdadera.

sobre lo cual los cuantificadores de los cuales se dispone sí permiten cuantificar, obteniendo las oraciones (4.6) o (4.7), y 'Algo de lo que dice Oscar es verdadero', según el cuantificador que se quiera utilizar.

En general, dada una oración B y una suboración A de B, el predicado veritativo permite 'cuantificar sobre A' en B, reemplazando primero A por 'A es verdadera' en B y luego generalizando sobre el nombre 'A', exactamente como en 'Tom es idéntico a Tom' sobre el nombre 'Tom'. En términos de Ramsey, el predicado de verdad se aplica redundantemente en adscripciones generales pero provee un verbo a la variable sobre

⁶Tras una larga búsqueda, Filip Buekens me sugirió el verbo 'articular' para reemplazar la larga frase

`exhibir una formulacion expl cita de una oracion'.

la cual se cuantifica que es gramaticalmente necesario. El ascenso al plano lingüístico tiene como único propósito ubicar las oraciones cuyo contenido se quiere transmitir en posición de objeto, para poderlas cuantificar sobre ellas. Esta capacidad se debe al aspecto entrecomillador de la función básica de la verdad. El aspecto inverso de esta función, la desentrecomillación, como también nota Quine, garantiza que el ascenso semántico sea inocuo, que no cambie el contenido del discurso a oraciones sino que el foco permanezca en el contenido de las oraciones sobre las cuales se generaliza.

En consecuencia, a través de su función básica el predicado de verdad permite, por ejemplo, expresar el contenido de una oración recurriendo a una descripción de nada a la cual la oración responda, esto es, mediante una adscripción general de verdad de la forma

Existe una única oración que es P y es verdadera. (4.13)

Si la oración A es la única que satisface P, (4.13) adscribe verdad a A y, por tanto, la función básica de la verdad, junto con las leyes de la identidad y el cuantificador existencial, garantiza que sea equivalente a A. Por ejemplo, como (4.11) es fórmula más famosa de la física, (4.5) expresa su contenido gracias a la función básica de la verdad. Sin la ayuda del predicado veritativo no es posible transmitir el contenido de una oración contando exclusivamente con una descripción de nada de ella; v.g., la oración "La fórmula más famosa de la física" no es equivalente a (4.11), ni siquiera es una oración declarativa.

Asimismo, es posible expresar el contenido de varias oraciones a la vez adscribiendo verdad a todas ellas mediante generalizaciones de la forma

Todas las P s son verdaderas. (4.14)

donde la cuantificación sobre oraciones está relativizada a un predicado P que las oraciones cuyo contenido se quiere expresar satisfacen. V.g., mediante (4.6) se expresa el contenido de todo lo que dice Oscar o, equivalentemente, el de la conjunción de todos sus dicta. Por esta razón generalizaciones de la forma de (4.14) suelen denominarse 'conjunciones', la conjunción de los enunciados que satisfacen P.

Análogamente, la cuantificación sobre oraciones posibilita expresar el contenido de al menos una oración cualquiera entre varias o, si se quiere, de su disyunción, mediante adscripciones generales de la forma

Algunas P s son verdaderas.

(4.15)

donde P es satisfecho exactamente por estas oraciones. Por ejemplo, (4.4) expresa la dis-yuncion de las leyes de Gauss, Gauss para el campo magnetico, Faraday-Lenz y Ampere generalizada.

Las dos funciones logico-expresivas del predicado de verdad ponen en evidencia su caracter logico-sintactico o logico-linguistico. En el apartado siguiente se vera como estas dos funciones dan poder expresivo al predicado de verdad y lo convierten en un dispositivo invaluable, si es que no hablan ya por s solas.

4.2.3. Los usos expresivos de la verdad

Como se vio en el apartado anterior, junto con cierta informacion externa (premisas) y en combinacion con ciertos operadores y leyes logicas, la funcion basica de la verdad garantiza que oraciones donde el predicado de verdad se antepone sintacticamente a cual-quier tipo de termino, y no unicamente a nombres transparentes de oraciones, tambien expresen el contenido de otras oraciones, una o varias a la vez, sus conjunciones o sus disyunciones. Permitiendo expresar el contenido de oraciones conociendo solo nombres no transparentes de ellas o descripciones a las cuales responden, el predicado veritativo sirve una variedad interesante de propositos expresivos en el lenguaje. Aquellos de los cuales se tiene conocimiento hoy pueden agruparse, grosso modo, en usos retoricos y estilisticos, usos epistemicos y usos logicos.

La tesis central del de acionismo actual en su version mas precisa consiste quizas en sostener que todo uso del predicado de verdad, no solo aquellos de los cuales se tiene conocimiento, puede ser explicado en terminos de sus dos funciones logico-expresivas y, por tanto, indirectamente a traves de su funcion basica. Luego, de acuerdo con el de acionismo el predicado veritativo tiene unicamente roles logico-expresivos en el lenguaje. Tampoco esta para adscribir una propiedad sustantiva a los portadores de verdad como, por ejemplo, cierta correspondencia con el mundo (no es una relacion entre el lenguaje y el mundo porque, de hecho, no es una relacion), sino para facilitar o posibilitar la expresabilidad del contenido de ciertas oraciones. Luego, un modo de refutar el de acionismo es hallar un contexto en el cual la verdad est cumpliendo un rol que no pueda ser explicado en terminos de las funciones presentadas en el apartado anterior.

En este apartado presento los diversos usos o propositos expresivos del predicado veritativo tiene en el lenguaje, algunos de los cuales son indispensables, y los ilustro mediante ejemplos. Luego, en 4.2.3.4, considero ciertas objeciones que se han hecho al de acionismo señalando usos del predicado veritativo de los cuales no se podr a dar cuenta a traves de su funcion basica. Como se vera,

algunos de estos `contraejemplos' apuntan a un cuarto uso del predicado de verdad.

4.2.3.1. Usos retóricos y estilísticos

Como ya señalamos [Ramsey \(1927\)](#), ciertas adscripciones de verdad pueden utilizarse con propósitos retóricos o estilísticos. Por un lado, las adscripciones singulares explícitas de verdad en [\(4.9\)](#), si bien no agregan nada al contenido de la oración que es objeto de la predicación, lo cual queda garantizado por la función básica de la verdad, pueden ser utilizadas para dar énfasis a lo que se quiere expresar.

Por otro, el predicado de verdad también puede utilizarse para evitar repeticiones o formulaciones tediosas, mediante adscripciones implícitas de verdad, ya sean singulares o generales. Por ejemplo, si se quieren explorar en un discurso ciertas consecuencias de la equivalencia entre masa y energía, puede utilizarse frases como "Si la energía en reposo de un cuerpo es el producto de su masa por la velocidad de la luz al cuadrado, entonces . . .", pero adicional o alternativamente adscripciones singulares como "Si la equivalencia entre masa y energía es verdadera, entonces . . .", o generales, como "Si la fórmula más famosa de la física es verdadera, entonces . . .", que son mucho más breves.

Asimismo, si se desea argumentar a favor de las ecuaciones de Maxwell, parece mucho más conveniente utilizar frases como "Por lo tanto, todas las ecuaciones de Maxwell son verdaderas" que enunciar las cuatro ecuaciones explícitamente después de "Por lo tanto"; y si se desea argumentar a favor de una de ellas sin saber específicamente cuál, puede recurrirse a expresiones como "Por lo tanto, algunas ecuaciones de Maxwell son verdaderas", en lugar de su extensa disyunción.

En todos estos casos, se recurre a las funciones lógico-expresivas de la verdad para garantizar la existencia de formulaciones alternativas equivalentes. También es claro que, en materia de transmisión del contenido proposicional (pero no del modo o el sentido), el predicado de verdad resulta totalmente dispensable en estos casos. Los usos retóricos o estilísticos de la verdad son denominados por [Grover et al. \(1975\)](#) 'lazy uses of truth'.

4.2.3.2. Usos epistémicos

Decir que el predicado de verdad tiene ciertos propósitos epistémicos podría resultar confuso. El predicado veritativo ciertamente no cumple roles semejantes a los predicados u operadores de conocimiento, creencia, etc. bajo ningún concepto sino que en su función 'epistémica' se limita a permitir expresar el contenido de una oración sin conocer o recordar su formulación explícita o quizás su contenido mismo. Solo en este sentido puede decirse que la verdad tiene un uso epistémico.

La verdad cumple este propósito mediante adscripciones implícitas de verdad. Por ejemplo, cuando se realizan adscripciones de la forma [\(4.10\)](#), donde a no es un

nombre transparente de una oración A, y no se sabe exactamente que oración denota a. Esto último implica que la adscripción de verdad es ciega. Naturalmente, no es una nota sintáctica inherente a la adscripción sino que depende del estado mental de quien la enuncie. La primera función lógico-expresiva de la verdad garantiza que (4.10) transmite el contenido de A aun cuando la persona que la utilice no conozca la relación de denotación relevante. V.g., se puede expresar el contenido de (4.11) mediante (4.3) sin recordar o conocer (4.11).

Del mismo modo, adscripciones implícitas generales de verdad también pueden ser adscripciones ciegas. Mediante oraciones de la forma (4.13), se puede expresar ciegamente el contenido de una oración solo con disponer de una descripción de nada a la cual esta oración responda. Por ejemplo, supongase que Lorna es una experta en música nacional, además de una persona honesta, y alguien le pregunta en qué fecha Charly García sacó su primer álbum solista, pero yo no llegué a escuchar su respuesta. Supongase también que después de este episodio Lorna se va y el resto de las personas presentes, que escucharon la fecha que dijo Lorna, discute todavía a cuando es que Charly García sacó su primer álbum. Yo creo en la respuesta de Lorna, cualquiera sea, y puedo, utilizando el predicado de verdad, expresar lo que ella dijo mediante "La respuesta de Lorna es verdadera".

Otras generalizaciones como (4.14) y (4.15) también pueden ser ciegas y, en general, cualquier adscripción implícita puede serlo. Basta que no se sepa de qué oraciones exactamente se está predicando verdad pero sí se conozca una descripción a la cual todas estas respondan. Por ejemplo, siguiendo la forma de (4.14), el ejemplo de Ramsey en (4.6) es un modo de expresar todo lo que Oscar dice sin que sea necesario saber que dirá mañana o que dijo exactamente en el pasado.

En las adscripciones ciegas el predicado de verdad ya no es dispensable para la transmisión del contenido. Simplemente no hay otro modo para la persona que no recuerda o conoce la articulación de una o varias oraciones de transmitir su contenido que mediante nombres no transparentes o descripciones a las cuales las oraciones respondan.

4.2.3.3. Usos lógicos

Finalmente, lo que llamo 'uso' o 'propósito lógico' de la verdad suele describirse como la capacidad que el predicado de verdad tiene de permitir expresar varias oraciones a la vez que incluso pueden ser infinitas o, si se quiere, conjunciones y disyunciones posiblemente infinitas. Nuevamente, el calificativo 'lógico' podría ser un abuso terminológico. Las razones quedarán más claras en el apartado 4.3.3 y en el capítulo siguiente.

Como se~nala [Quine \(1970\)](#), el cuantificador universal puede verse como un dispositivo que permite expresar el contenido de varias, posiblemente infinitas, oraciones

a la vez o, en algún sentido, su conjunción posiblemente infinita. Por ejemplo, "Todo es idéntico a sí mismo" puede entenderse como expresando la 'conjunción infinita' de cada una de sus instancias, i.e., de todas las oraciones de la forma "a es idéntico a a", donde a es un término cerrado, mientras que "Todos los humanos son mortales" puede pensarse como la 'conjunción' (no infinita) de todas las oraciones de la forma "m es mortal", donde m es el nombre de un humano. Pero para que el cuantificador universal permita expresar varias oraciones mediante una única expresión es necesario que todas las oraciones del conjunto diéran entre sí solo en un término, esto es, debe existir una función proposicional $A(x)$ tal que todas ellas se obtengan reemplazando x en A por un nombre. Cuando este no es el caso, el predicado veritativo, combinado con el cuantificador universal, puede hacer el trabajo mediante expresiones como (4.14), en la medida en que las oraciones que se quiera expresar, si bien no la forma, compartan alguna propiedad expresable en el lenguaje por medio de un predicado P. Las siguientes palabras de Quine se convirtieron en el locus classicus:

We may affirm the single sentence by just uttering it, unaided by quotation or by the truth predicate; but if we want to affirm some infinite lot of sentences that we can demarcate only by talking about the sentences, then the truth predicate has its use. (Quine, 1970, p. 12, mis itálicas)

Si P es satisfecho por un número infinito de oraciones, la segunda función lógico-expresiva de la verdad implica que (4.14), esto es,

Todas las P's son verdaderas,

expresa estas infinitas oraciones. Dado que, como señale en el apartado 4.2.2.2, las expresiones de esta forma pueden verse como la conjunción de las fórmulas que satisfacen el predicado P y estas pueden ser infinitas, suele decirse de este tipo de generalizaciones que expresan o son conjunciones infinitas cuando ese es el caso.

Alternativamente, sin la necesidad de que P se aplique a un número infinito de enunciados, estas generalizaciones pueden entenderse como la conjunción infinita de todos los condicionales de la forma "Si P ($\neg A$), entonces A". Y, en general, toda oración de la forma "Para todo x, B" donde "x es verdadera" ocurra como subfórmula de B y x está ligada por el cuantificador del comienzo podrá ser vista como la conjunción infinita de todas las instancias que se obtienen al reemplazar "x es verdadera" por una oración A del lenguaje, en virtud de la segunda función lógico-expresiva de la verdad, del mismo modo en que "Todos los humanos son mortales" puede verse como la conjunción infinita de todas las instancias "si a es humano entonces a es mortal" donde a no es

necesariamente ya un nombre de un humano sino cualquier constante de individuo del lenguaje. Pero, al menos en el caso de (4.14), si P se aplica únicamente a un número finito de enunciados, aunque se entienda que se está expresando la conjunción finita de las expresiones de la forma $\neg P(\ulcorner A \urcorner)$, entonces A , (4.14) es equivalente a la mera conjunción finita de las oraciones que son P s. Solo cuando P es satisfecho por infinitas expresiones puede decirse que (4.14) expresa una conjunción infinita, y lo mismo para el caso de los cuantificadores universales solos.

De modo análogo, el cuantificador existencial puede ser visto como un dispositivo que permite expresar disyunciones, pero solo es capaz de hacerlo en la medida en que los 'disyuntos' diéresis eran entre sí exclusivamente en un término. El predicado veritativo puede pensarse entonces como un complemento del cuantificador existencial que, al permitir cuantificar sobre oraciones, habilita la expresión de nuevas disyunciones posiblemente infinitas, v.g., a través de expresiones como (4.15), esto es,

Algunas P s son verdaderas.

en la medida en que los 'disyuntos' compartan alguna propiedad expresable en el lenguaje por medio de un predicado P .

Un ejemplo sencillo de una 'conjunción infinita' está dado por la oración

Todos los teoremas de la aritmética son verdaderos. (4.16)

Los teoremas de la aritmética son infinitos y no comparten su forma lógica y, por ende, no pueden expresarse todos a la vez sin la ayuda del predicado veritativo. A través de la segunda función lógico-expresiva de este predicado, es posible cuantificar sobre todos y cada uno de los teoremas de la aritmética predicando verdad sobre ellos y, por ende, expresar el contenido de todos ellos a la vez.

Formulaciones como (4.16) pueden resultar útiles, por ejemplo, a la hora de expresar acuerdo con pa o, alternativamente, su corrección. (4.16) puede ser vista también como una axiomatización finita de pa que, como es sabido, no es nitamente axiomatizable en L_{pa} . En efecto, como prueban [Stephen C. Kleene \(1952\)](#) y [William Craig & Vaught \(1958\)](#), el predicado de verdad (o la relación de satisfacción, a partir de la cual la verdad es definible) es un dispositivo idóneo para estos fines.⁷

Otros casos de axiomatizabilidad o expresabilidad finita provenientes de la filosofía resultan también interesantes y desmitificadores. Por ejemplo, el realismo científico, la

⁷Por ejemplo, es sencillo probar que pa es nitamente axiomatizable en L_T sobre $tb(pa)$ (vease la definición 2.10, apartado 2.3.3) mediante la fórmula $\exists x(Bew_{pa}(x) \rightarrow T x)$, porque Bew_{pa} debilmente representa el conjunto de teoremas de pa (vease el apartado 1.2.3).

idea de que hallar la verdad es el objetivo de la ciencia, puede expresarse en los siguientes términos:

Es deseable que algo sea un teorema de una teoría científica si y solo si es verdadero. (4.17)

Este enunciado no está poniendo como meta de todo proceso científico un concepto metafísico oscuro, sino que es meramente la 'conjunción infinita' de todas las expresiones de la forma 'Es deseable que 'A' sea un teorema de una teoría científica si y solo si A'. La verdad, a través de su segunda función lógico-expresiva, está permitiendo generalizar sobre A y, por ende, garantiza que (4.17) exprese el contenido deseado. Gracias al predicado veritativo, es posible expresar nitamente el objetivo de la ciencia.

Asimismo, cuando se sostiene que el conocimiento es creencia verdadera justificada, 'verdadera' no agrega conceptualmente nada a la definición de conocimiento sino que está cumpliendo rol expresivo, está permitiendo definir conocimiento mediante una única oración,

Se conoce un enunciado si y solo si se lo cree justificadamente y es verdadero. (4.18)

la 'conjunción infinita' de todas las instancias de 'Se conoce 'A' si y solo si se cree justificadamente 'A' y A'.

Una vez más, cuando se dice que la verdad es la norma de la aserción, no se adscribe una propiedad metafísica a lo que legítimamente puede afirmarse; el predicado de verdad está ahí para permitir expresar a la vez todas las instancias de 'Es legítimo afirmar 'A' solo si A'. Si las instancias de la norma de la aserción fueran finitas, la generalización no sería necesaria, porque se podría tomar simplemente su conjunción. Pero al ser infinitas, el predicado veritativo y su segunda función lógico-expresiva se tornan indispensables.

O, cuando se dice, siguiendo a Frege, que dar el significado de una oración es dar sus condiciones de verdad, esto no implica que toda teoría del significado dependa necesariamente de una teoría de la verdad, como quieren, v.g., [Davidson \(1967\)](#), [Michael Dummett \(1978, cap. 14\)](#) y [Priest \(2006b\)](#), sino que es simplemente un modo sucinto de expresar todas las instancias de 'Dar el significado de 'A' es identificar bajo que condiciones A', para cada oración A del lenguaje.

Como un último ejemplo, cuando se define consecuencia lógica como preservación de verdad en todo mundo posible, la palabra 'verdad' está para reunir en una única oración todas las instancias de 'B es consecuencia lógica de 'A' si y solo si en todo mundo posible en el cual A, B'. El rol conceptual de la verdad en la definición de

la noción de consecuencia lógica tiene tanto peso como el del cuantificador universal ("... en todo mundo posible ...").

Sin el predicado de verdad, expresar ciertos conjuntos infinitos de oraciones sería completamente imposible en lenguajes sin cuantificación sobre oraciones. No importa que se sepa o recuerde, enumerar todos los miembros de un conjunto infinito está fuera de las capacidades humanas. Luego, los usos lógicos del predicado de verdad son de algún modo más importantes que los usos epistémicos y retórico-estilísticos, y mucho más frecuentes también en el discurso científico y filosófico. Por esta razón suele identificarse el rol expresivo de la verdad directamente con la expresabilidad de conjunciones y disyunciones infinitas (cf. [Horwich \(1998b\)](#), [Leeds \(1978\)](#), [Quine \(1970\)](#)). De acuerdo con Horwich,

the raison d'être of the truth predicate is to provide a device enabling us to formulate propositions that can be the objects of belief, desire, etc., in cases where the propositions of primary concern are inaccessible|or (what comes to the same thing) to allow us to formulate 'infinite conjunctions' [...] ([Horwich, 1998b](#), p. 52, mis itálicas)

Si bien es común referirse a expresiones de la forma de (4.14) como 'conjunciones infinitas' y, análogamente, a expresiones de la forma de (4.15) como 'disyunciones infinitas', como hice en este apartado, en ambos casos la terminología puede resultar confusiva, en tanto los lenguajes más usuales, como los naturales o los de primer orden, o incluso los de orden superior, no pueden expresar conjunciones y disyunciones infinitas y, por tanto, no hay posibilidad de contrastar directamente expresiones como (4.16) con la conjunción infinita de todos los teoremas de la aritmética, v.g., si son o no equivalentes o bajo qué circunstancias podrían serlo. Luego, las frases 'conjunciones infinitas' y 'disyunciones infinitas' deben entenderse por ahora solo en un sentido metafórico; en el mismo sentido en que podría decirse que los cuantificadores universal y existencial expresan conjunciones y disyunciones infinitas, respectivamente. En el capítulo siguiente, en el apartado 5.2, doy un análisis detallado de la cuestión.

4.2.3.4. Usos insurgentes de la verdad

El deíctismo es una postura fácilmente refutable; basta con encontrar un retazo de discurso en el cual el predicado de verdad sea correctamente utilizado pero no está cumpliendo un rol expresivo explicable en términos de su naturaleza desentremilladora. Naturalmente me refero a un predicado de verdad

dise~nado para aplicarse a oraciones o, si se desea, a proposiciones, pero no a objetos, como en "Jane es una amiga verdadera".

Un ejemplo que suele citarse para refutar ciertas versiones del de acionismo es la predicación de verdad de oraciones que pertenecen a un lenguaje diferente de aquel en el cual se realiza la predicación. Por ejemplo, en el marco de esta investigación yo puedo decir sin incurrir en ningún tipo de error gramatical

Todo lo que Quine dice en *Philosophy of Logic* es verdadero. (4.19)

aun cuando la copia del libro de Quine en mi biblioteca, la única a la cual he tenido acceso, está escrita en inglés. Intuitivamente, (4.19) expresa el contenido de la conjunción de todas las oraciones de *Philosophy of Logic*, pero la función básica de la verdad no permite concluir esto último. Porque, de acuerdo con lo dicho en el apartado 4.2.1, la expresión 'es verdadero' predicada únicamente de una oración del lenguaje al cual el predicado pertenece, en este caso el castellano, genera una expresión semanticamente equivalente a esta oración.

Lo que este contraejemplo muestra es simplemente que la descripción de la función básica que di en el apartado 4.2.1 es demasiado acotada; lo que [Michael Resnik \(1997\)](#) denomina 'inmanente'. En el lenguaje natural, la (des)entrecomillación es 'trascendente', i.e., alcanza oraciones en principio de cualquier lenguaje. Dada una oración A cualquiera, ya sea del lenguaje propio o de uno foráneo, el predicado de verdad aplicado a A expresa el contenido de A, lo cual explica porque (4.19) expresa el contenido de todo lo que Quine dice en *Philosophy of Logic*, del mismo modo en que la función básica acotada explica porque (4.11) expresa el contenido de la equivalencia entre masa y energía de Einstein (formulada en castellano).

Contrario a lo que creen [Resnik \(1997\)](#) y [Shapiro \(2003\)](#), esta extensión de la descripción de la función básica de la verdad no altera en nada la naturaleza desentrecomilladora de la verdad sino que meramente amplía el rango de aplicación de la operación sintáctica. De acuerdo con Shapiro,

To say that a sentence is true is tantamount to removing the quotes and asserting the sentences. In order to do that, the given sentence must be in one's own language. ([Shapiro, 2003](#), p. 115, *mis italicas*),

Pero, justamente, la función básica de la verdad garantiza que predicar verdad de una oración sea equivalente (tantamount) a decir esta oración, no idéntico a enunciar la oración desentrecomillada misma. En esta diferencia radica toda la utilidad del predicado veritativo: expresar el contenido de una oración sin articularla. Nada de esto implica la segunda oración en la cita de Shapiro; nada fuerza que el enunciado original deba estar dado en el mismo lenguaje que su predicación de verdad.

Aun más, el 'contraejemplo' propuesto apunta quizás a un cuarto uso del predicado veritativo explicado en términos de su función básica: expresar el contenido de oraciones formuladas en lenguajes extranjeros mediante expresiones del lenguaje propio sin exhibir una traducción de estas oraciones.⁸ Como nota [Graham Priest](#) (2006b, p. 55),

[. . .] suppose the Pope utters A. We would like to assert [a translation of A to english] (the Pope will, of course, speak in Latin), but we cannot. Instead, we form a noun phrase which refers to A, i.e. 'A', and assert ['A is true].

Al aplicar el predicado de verdad de un lenguaje a una oración de un lenguaje diferente, no solo se produce un pasaje de la oración de posición oracional a posición nominal sino que tiene lugar un cambio de lenguaje. El predicado de verdad posibilita 'intercalar' o 'utilizar' oraciones en lenguajes foráneos en medio de un discurso en el propio lenguaje.

Podrá surgir al menos un problema al extender la desentramación a oraciones de cualquier lenguaje. Es posible que una misma secuencia de símbolos sea una oración de dos lenguajes diferentes a la vez, y que en cada uno de ellos signifique dos cosas diferentes. Este tipo de inconvenientes es un fenómeno común en el lenguaje natural, donde, por ejemplo, la asignación de nombres a objetos no es inyectiva [v.g., 'Rivadavia' es el nombre de una avenida en la ciudad de Buenos Aires y de un ex presidente argentino a la vez]. Como es usual, es el contexto el que provee información suficiente para evitar la confusión; pero incluso aquellos casos en los cuales el contexto no logra indicar un vocábulo de qué oraciones se trata no ponen en riesgo la descripción de la función básica, o al menos no en mayor medida que los casos de homonimia comprometen el uso de nombres propios en el lenguaje.

No todas las versiones del deísmo pueden responder al desafío mencionado simplemente reconociendo que la función básica de la verdad tiene un rango más amplio que el conjunto de oraciones del lenguaje al cual pertenece el predicado de verdad. [Field \(1994\)](#), por ejemplo, sostiene que una oración A y su predicación explícita de verdad son siempre cognitivamente equivalentes. Luego, "a person can meaningfully apply 'true' in the pure disquotational sense only to utterances that he has some understanding of" ([Field, 1994](#), p. 250); el rango de la función básica de la verdad solo puede extenderse a las oraciones que el individuo que utiliza el predicado comprende, no arbitrariamente a cualquier lenguaje, ni siquiera necesariamente a cualesquiera oraciones del lenguaje propio.⁹

⁸Incluso podrá resultar que no exista una traducción (cf. [Shapiro \(2003\)](#)).

⁹De hecho, Field sostiene que un individuo podrá entender ciertas oraciones de un lenguaje foráneo y a la vez no comprender otras de su propio lenguaje, en cuyo caso podrá predicar verdad de las primeras pero no de las segundas.

No obstante, Field luego considera la posibilidad incluir en el rango de aplicación del predicado veritativo oraciones que el individuo potencialmente comprender a, si contara con una buena traducción de ellas a su propio idiolecto. Esto ser a su vez importante para dar cuenta de los usos del predicado de verdad en el lenguaje natural aplicado a oraciones de lenguajes foráneos si hubiera traducciones de cada una de las oraciones de las cuales un individuo puede en principio predicar verdad a su idiolecto y, además, esas traducciones fueran cognitivamente equivalentes a las originales, de alguna manera. [Shapiro \(2003\)](#) pone ambas condiciones en duda.¹⁰

Una segunda objeción que suele hacerse al de acionismo es que su predicado de verdad desentrecomillador no permite decir que ciertas oraciones no son verdaderas en casos en los que no son ni verdaderas ni falsas. Haciendo ciertas concesiones mínimas a las lógicas paraconsistentes,¹¹ supongase que existen oraciones que no son ni verdaderas ni falsas. Por ejemplo, la oración del mentiroso dada por

L no es verdadera. (L)

podría no ser ni verdadera ni falsa, o una oración en la cual se predique una propiedad vaga (v.g., 'ser extensa') de un objeto que no es un caso claramente positivo ni claramente negativo (v.g., esta tesis). Supongase ahora que una persona dice

L es verdadera. (4.20)

De acuerdo con la noción de verdad de acionista, si L no es ni verdadera ni falsa entonces (4.20) tampoco, porque debe ser equivalente a la primera. Ahora bien, desde otro punto de vista, suele argumentarse, (4.20) es simplemente falsa, porque afirma lo contrario de algo que es cierto, esto es, de la oración

L no es verdadera. (4.21)

(porque L no es verdadera ni es falsa). La función básica de la verdad no solo no explica sino que contradice este segundo uso de la verdad.

[Gupta & Belnap \(1993\)](#) explican esta confusión postulando dos nociones diferentes de verdad aplicables a oraciones declarativas: una débil y una fuerte. De acuerdo con la noción débil de verdad, la verdad de acionista, toda predicación de verdad sobre una oración es equivalente a la oración que es objeto de la predicación y, por ende, tienen ambas el mismo estatus semántico. La noción fuerte, en cambio, al predicarse de

¹⁰Vease [Halbach \(2003\)](#) para una respuesta a Shapiro.

¹¹Un caso semejante puede hacerse con respecto a las lógicas paraconsistentes. Vease [Priest \(2006b\)](#), p. 79).

oraciones que no son verdaderas, ya sea porque son falsas o porque no son ni verdaderas ni falsas, dan lugar a oraciones que son simplemente falsas.

Sin embargo, no me es claro que (4.20) sea en efecto falsa bajo algún punto de vista, ni que haya tal cosa como una noción fuerte de verdad. Lo que esta verdaderamente operando detrás del impulso a tomarla por falsa es que, si \underline{L} no es ni verdadera ni falsa, no es correcto afirmar que es verdadera, porque no lo es, lo cual no significa que sea correcto afirmar que es falsa, porque el mismo razonamiento aplicado a \underline{L} le asigna a valores clásicos.

Ahora bien, si \underline{L} no es verdadera, (4.21) sí lo es, porque dice precisamente lo que se está afirmando. Pero (4.21) parece ser idéntica a \underline{L} , lo cual llevaría directamente a contradicciones. De acuerdo con Gupta & Belnap, estas contradicciones no tienen lugar porque las oraciones no son idénticas. En \underline{L} se está empleando una noción débil de verdad, mientras que en (4.21) opera una noción fuerte (podría diferenciarse, v.g., por medio de subíndices).

El de condicionismo no tiene esta estrategia a su disposición, pero puede igualmente dar cuenta de la diferencia entre estas dos oraciones, notando que, en lugar de dos nociones de verdad, lo que hay en juego son dos negaciones diferentes: una débil y una fuerte.¹² Ambas negaciones se comportan como es usual cuando niegan expresiones que adquieren valores de verdad clásicos, pero mientras que la primera, al anteponerse a una expresión que no recibe un valor clásico, devuelve una expresión que tampoco recibe un valor clásico, la negación fuerte devuelve el valor clásico 'verdad'. En \underline{L} , la negación es débil, lo cual le permite a la oración no ser ni verdadera ni falsa. Luego, por la función básica de la verdad, (4.20) tampoco es ni verdadera ni falsa. Pero (4.21) es verdadera, porque aplica la negación fuerte a (4.20). Si la negación ahí fuese la débil, la oración tampoco tendría un valor clásico, sería idéntica a \underline{L} . Estableciendo una distinción entre estos dos tipos de negaciones, el de condicionismo puede explicar perfectamente como se puede afirmar correctamente de oraciones que no son verdaderas o falsas que no son verdaderas.

Finalmente, una tercera objeción que puede hacerse al de condicionismo consiste en su prima facie incapacidad de dar cuenta del rol que el predicado veritativo ejerce en expresiones como

Una conjunción es verdadera si y solo si ambos conjuntos lo son. (4.22)

¹²*Priest (2006b, p. 79) da una explicación similar donde, en lugar de negaciones, hay condicionales in-volucrados, distinguiendo dos tipos de condicionales: el condicional material y el condicional intensional, que introduce en el apartado 3.4.2.*

Notese que una explicación en las líneas de las de (4.17) o (4.18) no está disponible para esta oración. Siendo el resultado de generalizar sucesivamente sobre 'A' y 'B' en

La conjunción de 'A' y 'B' es verdadera si y solo si 'A' es verdadera y 'B' es verdadera.

si se dijera que en (4.22) el predicado veritativo está permitiendo una conjunción infinita, deber a concluirse que está proveyendo los medios para expresar, mediante una única expresión, todas las oraciones de la forma

A y B si y solo si A y B

lo cual no parece ser el caso.

Efectivamente, el predicado de verdad no está jugando ningún rol lógico-expresivo en (4.22) ni en otros principios semejantes sino que estos, al contrario, colaboran con la expresión del significado del predicado veritativo, son analíticos, i.e., se siguen directamente de los principios que dan el significado de la verdad, como se verá en el apartado 4.3.4.

Asimismo, contrario a lo que sostienen, v.g., Davidson (1967), Dummett (1978, cap. 14) y Priest (2006b), el predicado veritativo no juega ningún rol expresivo en oraciones como

"La nieve es blanca" es verdadera si y solo si la nieve es blanca. (4.23)

sino que esta y otras expresiones semejantes son parte de los principios que dan el significado del predicado, a lo cual dedico la sección siguiente; es gracias a estos y otros principios, el predicado de verdad es capaz de cumplir sus roles lógico-expresivos. De acuerdo con estos autores, el predicado en (4.23) cumple un rol: permite dar el significado de la oración "La nieve es blanca", porque dar el significado de una oración es dar sus condiciones de verdad. Pero como el mismo Priest (2006b, p. 60) admite, "[4.23] cannot be considered as simultaneously specifying both the sense of ['La nieve es blanca'] and what it is for ['La nieve es blanca'] to be true." Mientras que Davidson, contrario al de acciónismo, sostiene que oraciones como (4.23) no constituyen el significado del predicado veritativo, Dummett y Priest entienden la verdad y el significado como variables interdependientes. El de acciónismo, al contrario, rechaza que la expresión "dar el significado de una oración es dar sus condiciones de verdad" implique la dependencia de toda teoría del significado de una teoría de la verdad [de hecho, algunos de acciónistas han desarrollado teorías del significado en donde este no es el caso (cf. Horwich (1998a), Field (1994))], como se vio en el apartado anterior.

§

Creo haber mostrado como las dos características del predicado de verdad, i.e., su redundancia conceptual y el hecho de ser un predicado, constituyen su naturaleza desentrecomilladora y le permiten cumplir su función básica que, a su vez, habilita las funciones lógico-expresivas del predicado de verdad. Estas funciones explican todos los usos que se hacen del predicado veritativo tanto en lenguajes formales como coloquiales y, por ende, todos sus roles son lógico-expresivos, no conceptuales. El más importante, aquel en el cual el predicado es irremplazable, es cuando juega un rol lógico y, más particularmente, cuando lo hace para expresar infinitas oraciones a la vez, conjunciones o disyunciones infinitas. Por estas razones puede decirse, prima facie, que el predicado veritativo es semejante a una constante lógica. Los matices a esta afirmación ocupan el apartado [4.3.3](#) y el capítulo que sigue.

4.3. El significado de la verdad

Si el predicado de verdad existe en el lenguaje exclusivamente para servir ciertos usos expresivos y, además, todos estos usos pueden explicarse en términos de lo que llamo 'función básica de la verdad', parece que no hay nada más que decir acerca de la naturaleza de la verdad que esta capacidad de (des)entrecomillar, de permitir la transformación de oraciones en términos sin alterar sus propiedades semánticas. De acuerdo con el deacacionismo, el significado del predicado de verdad es su naturaleza desentrecomilladora, dada por su redundancia conceptual y el hecho de ser un predicado. Como sostienen [Gupta \(1993\)](#) y [Horsten \(2011\)](#), el deacacionismo es una dupla de ideas, una de las cuales depende de la otra: primero, se posiciona con respecto al significado de la verdad; segundo, da cuenta del rol que este predicado juega en las prácticas discursivas en virtud de su significado.

En esta sección examino la viabilidad de ciertos principios diseñados para expresar la naturaleza desentrecomilladora de la verdad, los principios de equivalencia o transparencia, para expresar parte o todo el significado de este predicado, a la luz de diversas objeciones que se aparecen en la literatura. Estas objeciones conciernen únicamente a la (in)capacidad de los principios de transparencia de expresar la función básica de la verdad; parten de la premisa de que el significado del predicado veritativo es su función básica, i.e., no son objeciones sustancialistas (correspondentistas, coherentistas, verificacionistas, etc).

4.3.1. Intension, extension y reglas de uso

Como quedo evidenciado en la sección anterior, el predicado veritativo funciona como una pieza dentro de un mecanismo, con una forma específica (su naturaleza desentramadora) que, precisamente por su forma o significado, engrana con otras piezas en este mecanismo, i.e., los operadores lógicos (los cuantificadores, la relación de identidad, el condicional, la conjunción, etc.), para permitir la expresabilidad de ciertas cosas que de otro modo no es posible expresar. El predicado de verdad cumple una función comparable a la de los operadores lógicos, tiene roles lógico-expresivos. De acuerdo con el de acciónismo, no es un predicado introducido en el lenguaje para designar un concepto ordinario, una característica compartida por ciertas oraciones u otros portadores de verdad como quiere la teoría de la verdad por correspondencia, sino por razones lógicas, expresivas, semejantes a las razones por las cuales se han introducido cuantificadores u otras conectivas. El predicado de verdad no establece una relación entre el lenguaje y el mundo ni la explica; realiza una operación meramente sintáctica.

El hecho de que se trate de un predicado puede resultar engañoso. Por un lado, la semántica del lenguaje natural apunta a la búsqueda de una propiedad o concepto ordinario subyacente, una naturaleza compartida por las cosas verdaderas, como sucede con 'es una mesa'. A esto se deben los intentos sustantivistas (correspondentistas, coherentistas, pragmatistas) de dar definiciones explícitas de verdad en términos de (SUST). Por otro, la semántica formal (v.g., la teoría de modelos originada en Tarski, apartados 2.2.2 y 2.2.3) pretende asignarle una extensión como su interpretación, que puede entenderse como la extensión de esta propiedad o concepto. En ambos casos, el objetivo es errado.

Si bien algunas de acciónistas aún sostienen que la verdad no es un predicado o que el predicado de verdad no expresa un concepto genuino, la posición más popular al respecto hoy en día, lo que Horsten (2011, p. 143) denomina 'the misty core of de-actionism', consiste en defender que la verdad es un predicado y expresa un concepto, pero, a diferencia de los predicados ordinarios, este concepto es metafísicamente liviano (cf. Horsten (2011), Horwich (1998b)); un concepto de acciónado, lo cual suele explicarse diciendo que expresa una noción de índole lógica o cuasi-lógica (cf. Field (1992, p. 322), Field (1999, p. 534)). De acuerdo con Horwich (1998b, p. 5),

[. . . T]he traditional attempt to discern the essence of truth|to analyse that special quality which all truths supposedly have in common|is just a pseudo-problem based on syntactic overgeneralization. Unlike most other properties, being true is unsusceptible to conceptual or scientific analysis. No wonder

that its 'underlying nature' has so stubbornly resisted philosophical elaboration; for there is simply no such thing.

La situación es quizás comparable con la relación de identidad, que también es un predicado, o con el predicado de existencia, caído en desuso, y volverse más comprensible a través de esta comparación. Al igual que la identidad y el 'predicado' de existencia, el propósito del predicado de verdad en el lenguaje no es la expresabilidad de una propiedad que algunos objetos comparten mientras que otros no [en el caso de la identidad y la existencia las propiedades y extensiones correspondientes serían triviales], sino cumplir un rol expresivo [la identidad sirve para expresar cuestiones de cardinalidad, o de n-er funciones, mientras que la 'existencia' permite expresar 'disyunciones (posiblemente) infinitas']. Y al igual que la identidad y la existencia, la verdad es capaz de cumplir este rol a través de los principios que la gobiernan. Si bien se puso fin a las confusiones generadas por el predicado de existencia adoptando en su lugar un operador,¹³ esta estrategia no está disponible para el predicado de verdad. Si se lo reemplazara por un operador, ya no sería capaz de 'transformar' sintácticamente oraciones en términos sino que estas seguirían ocupando una posición oracional, esto es, ya no sería un dispositivo desentrecomillador, no podría cumplir su función básica, y, por ende, ninguna de sus funciones lógico-expresivas.

Luego, si la búsqueda de la naturaleza profunda de la verdad es estéril, la de su extensión es inútil. Contrario a lo que sostiene Barrio (1998, p. 5), i.e. que "[t]odo lo que un de acciónista aconseja hacer es dejar la extensión de la verdad, sin postular ninguna propiedad especial que supuestamente todas las oraciones verdaderas tendrían en común y que las falsas carecerían", lo relevante no es tampoco la extensión del predicado veritativo sino los principios que lo gobiernan y le permiten cumplir sus roles expresivos. Nuevamente, en términos de Horsten (2011, p. 146),

So this is what de actionism about truth comes to in the end. We should not aim at describing the nature of truth because there is no such thing. Rather, we should aim at describing the inferential behavior of truth.

Consiguientemente, la naturaleza desentrecomilladora del predicado veritativo, su función básica y significado, suele expresarse en forma de principios o reglas de verdad transparente o de transparencia. El más popular es el

'A' es verdadera si y solo si A: (Esquema T)

¹³Para un pantallazo general del problema vease Michael Nelson (2012).

introducido por Tarski (v.g., [Williams \(1988\)](#), [Horwich \(1998b\)](#) adhieren particularmente a este principio de transparencia), pero también suele recurrirse a las reglas de inferencia

De A se puede inferir 'A es verdadera'.	(Introducción T)
De 'A es verdadera' se puede inferir A.	(Eliminación T)

tomadas en conjunto (cf. [Horsten \(2011\)](#)), o al principio de Intersustitutividad (cf. [Field \(2008\)](#), p. 12), [Beall \(2009\)](#)):

Si dos oraciones B y C son idénticas excepto en que (en algún contexto no opaco) B contiene la subfórmula A mientras que C contiene 'A es verdadera' en su lugar, entonces B implica C y viceversa.

Los tres principios parecen poner igualmente en evidencia las dos notas que caracterizan a un dispositivo desentrecorillador: la equivalencia entre una oración y su predicación de verdad, la redundancia o transparencia de la verdad, por un lado, y el pasaje sintáctico de posición oracional a posición nominal, de oración a término, por otro. Contrario a lo que el correspondentismo puede creer y Tarski mismo sostuvo (véase el apartado [2.1](#)), los principios de equivalencia no establecen que la verdad es una relación entre el lenguaje y el mundo, porque no hay ninguna relación involucrada sino a lo sumo un bicondicional; establecen meramente la naturaleza desentrecorilladora de la verdad. La extensión de la verdad, dada a través de estos principios, dependerá en cada caso de lo que se afirma y niega, esto es, de cada ciencia o disciplina particular.

Al expresar de alguna manera la naturaleza desentrecorilladora de la verdad, de acuerdo con el de acciónismo (algunos de) estos principios dan, al menos, parte del significado del predicado de verdad. Algunas versiones del de acciónismo, hoy llamadas 'minimalismo' (cf. [Horwich \(1998b\)](#)) o 'desentrecorilladorismo' ('disquotationalism', como las llama [Halbach \(1999\)](#)), sostienen que, de hecho, estos principios agotan el significado de la verdad (v.g., [Williams \(1988\)](#), [Horwich \(1998b\)](#)), aunque este último prefiere proposiciones como los portadores de verdad). Si su naturaleza desentrecorilladora de hecho completamente la verdad y (alguno de) los principios de transparencia expresan esta naturaleza de modo acabado, entonces no parece haber necesidad ni lugar para otros principios para dar el significado del predicado veritativo. De acuerdo con [Williams \(1988\)](#), p. 424), por ejemplo, el [Esquema T](#) es 'just about everything there is to be said about truth'. En palabras de Leeds,

What we have sketched is not a theory of truth|a theory that shows the class of true sentences to be an important subset of the sentences in our language|but a theory of the concept of truth. [Leeds \(1978\)](#), p. 122)

Una objeción obvia que vale la pena explorar preliminarmente es la ya enunciada por el mismo Tarski (vease el apartado [2.1.2](#)): si bien en un plano informal los principios de transparencia pueden resultar inocuos, cuando se asume cierta lógica subyacente, como la clásica (pero también la intuicionista), llevan directamente a contradicciones, debido a la existencia de oraciones del mentiroso como L en el lenguaje. Si se supone que L es verdadera, la dirección de izquierda a derecha del [Esquema T](#), i.e.¹⁴

Si 'A' es verdadera, entonces A. (T Out)

[Eliminación T](#) o el principio de intersustitutividad permiten derivar que no lo es; y si se parte en cambio de esto último, la dirección de derecha a izquierda del [Esquema T](#), esto es,

Si A, entonces 'A' es verdadera. (T In)

[Eliminación T](#) o el principio de intersustitutividad permiten derivar que no lo es; y si se parte en cambio de esto último, la dirección de derecha a izquierda del [Esquema T](#), [Introducción T](#) o, nuevamente, el principio de intersustitutividad implican la verdad de L. Luego, en virtud del principio de tercero excluido, es posible derivar una contradicción, y por la regla Ex falso sequitur quodlibet, cualquier cosa.

En consecuencia, aquellos que consideran que la lógica clásica (intuicionista) es la lógica 'verdadera' incluso en el contexto de los lenguajes ordinarios deben, o bien tener a la verdad como un concepto inconsistente y, por ende, inutilizable, o bien rechazar que los principios de transparencia sean parte constitutiva del significado del predicado veritativo, al menos en su versión irrestricta o no tipeada. No obstante, esta no es una posición demasiado extendida, ni siquiera entre aquellos que favorecen la lógica clásica por sobre otras (cf. [Horwich \(1998b\)](#)). Como se señaló en el capítulo 2 (apartado [2.1.2](#)), no es claro que el lenguaje natural obedezca una lógica en particular. Buena parte de la discusión acerca del significado y la naturaleza de la verdad suele llevarse a cabo en un plano intuitivo, desde el lenguaje natural, y las paradojas son puestas momentáneamente entre parentesis. En esta sección me mantendré siempre en este plano, a menos que aclare lo contrario.

4.3.2. El carácter analítico de los principios de transparencia

Si las instancias del [Esquema T](#) expresan al menos parte del significado del predicado veritativo, deben ser verdaderas en virtud de este significado, esto es, deben ser analíticas

¹⁴Notese que, al igual que la diferencia entre el [Esquema T](#) y el [Esquema-T](#), el guion, ausente en [T Out](#), indica que se trata de la versión informal, en castellano, de [T-Out](#), y otro tanto va para [T In](#) y su contraparte formal [T-In](#), que aparecer en el próximo capítulo.

y, por tanto, necesarias en algún sentido. Análogamente, las reglas [Introducción T](#) y [Eliminación T](#) y el principio de intersustitutividad deben preservar verdad necesariamente. Sin embargo, una clásica crítica a la necesidad del [Esquema T](#), que se extiende naturalmente a la preservación necesaria de verdad de los otros principios de transparencia, es que, por ejemplo, la verdad de la instancia

"La nieve es blanca" es verdadera si y solo si la nieve es blanca. (4.24)

depende, entre otras cosas, de que 'nieve' signifique nieve. Si, en cambio, signifique pasto, lo cual es perfectamente posible en virtud de la arbitrariedad del signo, la instancia sería falsa y, en su lugar, se debería tener bicondicionales como

"La nieve es blanca" es verdadera si y solo si el pasto es blanco. (4.25)

Luego, (4.24) no puede ser siquiera necesariamente verdadera desde un punto de vista físico (cf. [Putnam \(1988\)](#)). Para que "La nieve es blanca" sea verdadera no basta con que la nieve sea blanca sino que las palabras que conforman esta oración deben significar lo que significan actualmente y no otra cosa, mostrando que el [Esquema T](#) no puede ser parte constitutiva del significado de la expresión 'es verdadera'. De la misma manera, las reglas [Introducción T](#) y [Eliminación T](#) y el principio de intersustitutividad no pueden expresar parte del significado del predicado de verdad en la medida en que no son capaces de preservar verdad necesariamente.

Notese que Tarski no debe enfrentar esta objeción porque las condiciones que impone a toda definición de verdad para un lenguaje objeto a través de la Convención T (definición 2.1) no requieren que estas definiciones impliquen principios de transparencia como el [Esquema T](#) sino una versión alternativa, el [Esquema T Tipado](#):

"A" es verdadera si y solo si (A): (Esquema T Tipado)

donde A pertenece al lenguaje objeto, para el cual se define verdad y (A) es la traducción de A al lenguaje en el cual se enuncian las instancias del [Esquema T Tipado](#), el metalenguaje. Luego, si 'nieve' signifique pasto, la instancia que habría que aceptar, según el [Esquema T Tipado](#), no es (4.24) sino (4.25). [Roy T. Cook \(2014b, p. 6, nota 8\)](#) argumenta erróneamente cuando sostiene que, plausiblemente, como las definiciones tarskianas de verdad son analíticas (porque toda definición lo es), la consecuencia lógica preserva analiticidad y el [Esquema T](#) se sigue de estas definiciones, este principio es también analítico; porque su última premisa es falsa: es el [Esquema T Tipado](#), no el [Esquema T](#), el que se sigue de las definiciones tarskianas.

El de la acción puede en principio extender el [Esquema T](#) admitiendo las instancias del [Esquema T Tipado](#), por ejemplo, dadas por las mismas oraciones que el [Esquema T](#) (y otro tanto con [Introducción T](#) y [Eliminación T](#) o el principio de intersustitutividad). En el mundo actual, el [Esquema T Tipado](#) coincide completamente con el [Esquema T](#), pero en otros mundos posibles, donde los significados de las oraciones a las cuales se les predica verdad varían y el lenguaje objeto difiere del metalenguaje, los principios divergen. Algunas de las acciones que han favorecido esta posición argumentan a favor de ella sosteniendo que el predicado de verdad se aplica a oraciones interpretadas, todo principio que regule el comportamiento de este predicado supone un significado de las oraciones del lenguaje en el cual se está trabajando que no contengan el predicado de verdad.¹⁵ Si bien la verdad parece depender ahora de ciertas relaciones entre las palabras que constituyen las oraciones a las cuales se aplica el predicado y el mundo (v.g., la referencia o denotación de estas palabras), como quiere la teoría de la correspondencia, la mención de estas propiedades no es necesaria a menos que no se comprenda el lenguaje objeto: la parte inaccionaria no proviene de la verdad sino de esta posible incomprensión.

Una consecuencia inmediata de esta postura 'semanticalista' es que ya no puede, como suele hacerse y como [Davidson \(1969\)](#) demanda, explicarse la noción de significado en términos del concepto de verdad.¹⁶ Pero, como argumenta en el apartado [4.2.3.4](#), esto no es evidencia concluyente en contra del de la acción en tanto existen teorías alternativas del significado que no están formuladas en términos del predicado de verdad.

En cualquier caso, esta medida no es necesaria. Si A es una oración declarativa eterna (vease el apartado [2.1.2](#)), no se necesita saber nada acerca del significado de A o de las palabras o símbolos que la componen para formular su instancia correspondiente del [Esquema T](#): basta que la oración que está entre comillas sea la misma que la que está desentrecomillada en cada bicondicional. Como señala [Leeds \(1978, p. 122\)](#), "[T]o explain the utility of disquotation, we need say nothing about the relations between language and the world", o prácticamente nada: tan solo que las comillas (o cualquier otro mecanismo para asignar nombres transparentes a oraciones), al ubicarse al comienzo y al final de una oración, dan lugar a nombres para las oraciones que encierran, esto es, un mecanismo de sintaxis acerca del lenguaje sobre el cual se trabaja, una relación de denotación acotada. Esta relación, siguiendo a [Kripke \(1980\)](#), es la referencia de todos los nombres transparentes de oraciones en todo mundo posible, aunque claramente puede

¹⁵ [Barrio \(1998\)](#) identifica al de la acción que sostiene esta tesis como de la acción semanticalista. Vease, v.g., [Field \(1994\)](#).

¹⁶ Cf. [Dummett \(1958-1959\)](#), [Davidson \(1990\)](#), [Horwich \(1998b\)](#). Otros filósofos también han sugerido que la verdad, desde un punto de vista de la acción, no puede utilizarse para explicar la noción de significado porque, en general, al ser conceptualmente redundante, no puede explicar nada (cf. [Davidson \(1990\)](#), [Field \(1986, 1994\)](#)).

haber sido de otro modo. En palabras de [Gupta \(1993, p. 81, nota 43\)](#), "It is unreasonable to expect a substantial answer to the question 'Why are all bachelors males?' but not to the question 'Why does 'snow' mean snow?'."

Las instancias del [Esquema T](#) y los otros principios de transparencia postulados por distintos de acciónistas están dadas por oraciones pertenecientes al mismo lenguaje desde el cual se enuncian los principios, al cual pertenece el predicado de verdad que predicado, y solo en estas condiciones puede decirse que son necesarias.¹⁷ Si 'nieve' signifique para pastos, también lo hará del lado derecho del bicondicional en (4.24) que, por ende, seguirá a siendo verdadero.¹⁸ Luego, el predicado de verdad es libre de aparecer en enunciaciones de significado pero, si lo hiciera su rol no sería de ninguna manera conceptual sino lógico-expresivo, como argumenté en la sección anterior.

Ahora bien, si las instancias de los principios de transparencia se limitan al lenguaje al cual pertenece el predicado de verdad, estos principios no logran capturar la versión de la función básica de la verdad extendida a oraciones de lenguajes foráneos que menciono en el apartado 4.2.3.4. En este caso, las instancias del [Esquema T Tipado](#) dadas por cada oración de cada lenguaje traducible al lenguaje coloquial en el cual se enuncia el principio, resultan necesarias. Y, si todas las oraciones de todos los lenguajes 'significativos', cuyas oraciones pueden en principio ser verdaderas o falsas, son traducibles al castellano (o a cualquiera sea el lenguaje natural desde el cual se enuncian los principios), el [Esquema T Tipado](#) resulta también suficiente, provistas las respuestas a las otras objeciones a la tesis minimalista que considero en esta sección. Pero si esto último no es el caso o, alternativamente, siguiendo a [Field \(1994, sec. 8\)](#) y sacrificando levemente la gramática, podrán adoptarse instancias del [Esquema T](#) mismo dadas, no ya solamente por oraciones del castellano (o del lenguaje natural en cuestión) sino de todo otro lenguaje (interpretado) existente. Por ejemplo, entre las nuevas instancias ahora se contará con la siguiente:

"Schnee ist wei" es verdadera si y solo si Schnee ist wei.

Quizás este sea el modo más útil de representar la función básica extendida de la verdad, en cuya descripción ninguna traducción ha sido mencionada.

¹⁷ *Notese que es posible que las equivalencias entre oraciones y adscripciones implícitas de verdad sobre ellas ya no tengan el carácter necesario de las instancias del [Esquema T](#), esto es, de las equivalencias entre oraciones y sus adscripciones explícitas. Por ejemplo, es una contingencia que (4.11) sea la fórmula más famosa de la física y, por ende, la equivalencia entre (4.11) y (4.5) tiene un componente contingente, ya no puede ser necesaria.*

¹⁸ [Barrio \(1998\)](#) llama 'de acciónismo sintacticista' al de acciónismo explicado en estos términos.

4.3.3. Principios de transparencia vs de nociones explícitas de verdad

Prima facie, si un principio agota el significado de un término puede decirse que constituye una de noción, ya sea explícita o implícita. Tradicionalmente, se impone como requisito sobre las de nociones, cualesquiera sean, la eliminatividad, i.e., que el de nien-dum resulte eliminable, que la de noción sea equivalente a una expresión de la forma de (EXP) (vease el apartado 2.1.1), a una de noción explícita.

Es evidente que los principios de transparencia enunciados no tienen ellos mismos la forma de una de noción explícita. Si bien tienen condiciones necesarias y suficientes para la aplicación del predicado de verdad para cada uno de los objetos en su rango de aplicación, esto es, cada una de las oraciones del lenguaje para el cual se quiere dar el significado de la expresión 'es verdadera', no lo hacen de manera general sino a través de una instancia de instancias. Por esta razón Tarski (1933) denomina estas instancias meramente 'de nociones parciales' (vease el apartado 2.2) del predicado veritativo.

Sin embargo, esto no es suficiente para descartar la existencia de una única expresión, una de noción normal del predicado de verdad, que sea equivalente o implique, de alguna manera, alguno de los principios de transparencia. Como señala Tarski (vease el apartado 2.2), esto puede hacerse fácilmente si se dispone de cuantificadores sobre oraciones, esto es, de cuantificación proposicional y cierto operador entrecomillador, simplemente cuantificando sobre A en el Esquema T, i.e., mediante

Para toda oración A, 'A' es verdadera si y solo si A.

No obstante, como indica Horwich (1998b, p. 25), "the use of substitutional [propositional] quantification does not square with the *raison d'être* of our notion of truth, which is to enable us to do without substitutional quantification." De hecho, como se vio en la sección anterior, el de acciónismo sostiene que al menos parte de las razones por las cuales el predicado veritativo existe en el lenguaje natural (y en ciertos lenguajes formales) es precisamente para emular la cuantificación sobre oraciones.

Tras explorar otras (im)posibles alternativas para definir verdad explícitamente, Horwich (1998b, cap. 5) rechaza la existencia de una de noción con estas características, aunque su razonamiento inductivo no puede ser concluyente. Autores como Halbach & Horsten (2005) y Horsten (2011) han argumentado que el de acciónismo se compromete con la indecibilidad de la verdad de manera explícita, ya que "[...] de actionists have tried to describe the purpose of truth that can only be achieved if truth is available. Or they say that truth would be dispensable if we could use infinite conjunctions or certain forms of quantification." (Halbach & Horsten, 2005, p. 204) Si el predicado de verdad es imprescindible, no puede ser explícitamente definible porque, de lo contrario, sería a

eliminable. Luego, concluyen, la verdad de la acción es una noción primitiva, indefinida (o, agregado, al menos indefinible de manera explícita).

Hay ciertamente un salto inferencial al pasar de la idea de que el predicado de verdad no es necesario si se dispusiera, por ejemplo, de cuantificación proposicional, a que el predicado es necesario ahora. No es lo mismo decir que la única razón de ser del predicado veritativo es emular la cuantificación proposicional (mas ciertos usos epistémicos, etc.) que decir que el predicado veritativo es el único mecanismo disponible para esta función en el lenguaje. Es epistémicamente posible que existan mecanismos alternativos aun no descubiertos (por mí, al menos) que cumplan esta función, al menos en el lenguaje natural.

Lo que se queda claro es que no existe una definición explícita de verdad (en lenguajes sin capacidad de cuantificación proposicional) en términos de nociones clásicamente consistentes, por el teorema de la indefinibilidad de Tarski (teorema 2.9). Luego, aun si una definición explícita del predicado de verdad que implique todas las instancias de alguno de los principios de transparencia fuera posible, estar a dada, directa o indirectamente, en términos de nociones clásicamente inconsistentes para las cuales una definición explícita no es posible, que a su vez habrá que elucidar.

En cualquier caso, como nota Horwich (1998b, p. 25), "there does not have to be any succinct, explanatorily adequate theory of truth", pace Tarski. So pena de caer en circularidades, no todo término puede introducirse al lenguaje mediante una definición explícita sino que necesariamente algunos deben ser introducidos como primitivos, agregando axiomas o reglas que los regulen, como la relación de pertenencia en la teoría de conjuntos, o las constantes lógicas en sus cálculos correspondientes (al menos que se las defina desde un metalenguaje que contenga a su vez su propio conjunto de conectivas lógicas, dado lugar al trilema de Fries/Münchhausen).

Siendo la pertenencia un concepto matemático ordinario, desde un punto de vista tradicional los axiomas de ZFC no constituyen una definición, no agotan el significado de '2', en tanto no son conservativos sobre la lógica subyacente y, por ende, no permiten la eliminatividad de las ocurrencias de este término. Sin embargo, el requisito de eliminatividad suele relajarse cuando se trata de constantes lógicas. La tesis según la cual es posible fijar el significado de las constantes lógicas mediante reglas de introducción y eliminación que los gobiernen (algunos sostienen que con las de introducción basta) sin la necesidad de dar definiciones explícitas, originalmente defendida por Frege (1879) y Gentzen (1934), se conoce como 'inferencialismo lógico'. Si el predicado veritativo es de naturaleza lógica, reglas como [Introducción T](#) y [Eliminación T](#), o quizás también axiomas como el [Esquema T](#) o el principio de intersustitutividad, podrán ser suficientes

para dar el significado de este predicado, a pesar de no permitir su eliminatividad y dar lugar a definiciones en el sentido ordinario.

No obstante, como señala [Prior \(1960/61, 1964\)](#), la tesis del inferencialismo lógico no puede valer irrestrictamente porque, a diferencia de las definiciones explícitas, no todo conjunto de reglas puede estipularse para dar el significado de una conectiva. Prior pone como ejemplo dos reglas que dan el significado de una conectiva binaria llamada 'tonk':

De A se puede inferir A tonk B.	(Introducción Tonk)
De A tonk B se puede inferir B.	(Eliminación Tonk)

Aplicando estas reglas sucesivamente, es posible inferir cualquier oración B a partir de cualquier oración A, por ejemplo, una contradicción a partir de una tautología. [Introducción Tonk](#) y [Eliminación Tonk](#) no logran asignar un significado coherente a 'tonk'.

Consecuentemente, [Belnap \(1962\)](#) propone debilitar la tesis del inferencialismo lógico, restringiéndola a conjuntos conservativos de reglas: la estipulación de reglas para una conectiva ya su significado en la medida en que no permita derivar nuevas oraciones que no contengan la conectiva correspondiente. Como [Introducción Tonk](#) y [Eliminación Tonk](#) generan problemas porque no son conservativas sobre los cálculos más usuales (i.e., aquellos en los que la noción de consecuencia lógica es transitiva, v.g., la lógica clásica pero también muchas lógicas no clásicas, como se vio en el capítulo 3), Belnap logra evitar el problema señalado por Prior. Luego, si los principios de transparencia fueran conservativos sobre el sistema de lógica preferido, podrían resultar suficientes para dar el significado del predicado veritativo.

Al mismo tiempo, [Horsten \(1995\)](#), [Shapiro \(1998\)](#) y [Jeremy Ketland \(1999\)](#) han argumentado que "in one form or another, conservativeness is essential to deflationism" ([Shapiro, 1998](#), p. 497) porque, si los principios de la verdad que se incorporan a una teoría o lógica permiten probar nuevas oraciones en el lenguaje de esta teoría o lógica, están agregando contenido semántico, contrario a lo que sostiene el deflacionismo.

Lamentablemente, los principios de transparencia no son conservativos sobre buena parte de los sistemas lógicos y teorías más utilizadas. Por ejemplo, ninguno de ellos es conservativo sobre la lógica clásica (cf. [Halbach \(2001\)](#)), ni siquiera en su versión tipeada, restringida a oraciones que no contienen ellas mismas el predicado de verdad. Por la reflexividad de la identidad, que vale en lógica clásica, se sigue que

Todo es idéntico a sí mismo.

(4.26)

Luego, las instancias de cualquiera de los principios de transparencia correspondientes a las oraciones (4.26) y su negación implican

\Todo es idéntico a sí mismo" es verdadera.

\No todo es idéntico a sí mismo" no es verdadera.

respectivamente. Por último, estas dos oraciones junto con el principio de indiscernibilidad de los idénticos, esto es, que si 'dos' objetos son idénticos entonces comparten todas sus propiedades, implican que las oraciones (4.26) y su negación son diferentes entre sí, i.e., que existen al menos dos objetos, lo cual no es una verdad lógica. La prueba de Halbach puede extenderse fácilmente a otras lógicas elegidas frecuentemente por de acciónistas.

Siguiendo a Belnap en su inferencialismo lógico debilitado, Cook (2005) muestra que reglas como [Introducción Tonk](#) y [Eliminación Tonk](#) pueden legítimamente fijar el significado de 'tonk' si se adopta una lógica subestructural (vease el apartado 3.4.3) en la cual la regla estructural de transitividad o [Cut](#) (vease el apartado 3.4.1.1)¹⁹ no vale irrestrictamente. Porque en una lógica sin transitividad [Introducción Tonk](#) y [Eliminación Tonk](#) pueden ser agregadas conservativamente. Ripley (2012a) muestra un resultado semejante para (versiones del cálculo de secuentes de) el principio de Intersustitutividad (IS, vease el apartado 3.4.1.2). Luego, pareciera que es preciso adoptar una lógica no transitiva u otra donde las reglas sean conservativas para que los principios de transparencia puedan fijar el significado del predicado de verdad.

Sin embargo, el requisito de conservatividad para la verdad ha sido puesto en duda tanto por de acciónistas (cf. Field (1999), Picollo (2010)) como por personas que no se identifican necesariamente con el de acciónismo (cf. Halbach (2001)). El contraargumento principal es que el predicado de verdad no es completamente lógico,²⁰ por no ser neutral al tópico; precisa una ontología específica, dada previamente, que es la verdadera responsable de las fallas de conservatividad. Al ser un predicado y no un operador, se aplica sintácticamente a términos, no a fórmulas, a diferencia, v.g., de la negación. Pero no se aplica a cualquier tipo de términos, como el predicado de identidad, sino solo a aquellos que denotan portadores de verdad, en este caso, oraciones. Los principios de transparencia requieren, como mínimo, de nombres en el lenguaje para cada una de las oraciones del fragmento del lenguaje para el cual se quiera definir verdad. En otras palabras, para formular principios de verdad es preciso contar con cierta teoría de la sintaxis en la base, que provea un discurso sobre las expresiones de las cuales se predica luego verdad. De acuerdo con, Horsten (2011, p. 65), por ejemplo,

¹⁹ i.e., *la propiedad de la noción de consecuencia lógica según la cual, si B es consecuencia lógica de A y C de B, implica que C es consecuencia lógica de A, para cualesquiera oraciones A, B y C del lenguaje.*

²⁰ *Contrario a lo que sostiene Field (1992, 1994, 1999).*

Definicionistas a veces dicen que la verdad es una noción lógica, o que la verdad expresa una propiedad lógica (Field (1992, p. 322), Field (1999, p. 534)). Al menos en parte, la verdad es una noción lógica. Hemos visto que, como los cuantificadores, el predicado de verdad es una herramienta para expresar generalidad. [...] Pero esta afirmación de la noción que la verdad es debería ser enmendada. En nuestro marco, el concepto de verdad se retrata como una noción lógico-lingüística porque, en parte, la verdad es una noción lingüística. Después de todo, los portadores de la propiedad de verdad son entidades lingüísticas. Más en particular, los portadores de verdad son oraciones significativas (i.e., objetos sintácticos interpretados).

Por otra parte, Cook (2014b) muestra, contra Priest (2007), que el hecho de haber verdades lógicas de la forma "A es verdadera" o "A no es verdadera" no lo son en virtud de su forma o, lo que es lo mismo, no satisfacen el principio de sustitutividad: para toda oración A, símbolo lógico ϕ y expresión (posiblemente compleja) ψ del mismo tipo lógico que ϕ ,²¹ si A es una verdad lógica, el resultado de sustituir ψ por A en ϕ también es una verdad lógica. Si ψ es una falsedad lógica, "A" y " ψ " son del mismo tipo lógico, con lo cual " ψ es verdadera" también debería ser una verdad lógica.

Luego, la verdad no es una noción propiamente lógica sino, como suele decirse, es un dispositivo lógico-lingüístico, y consecuentemente sus roles, podría decirse, no son plenamente lógicos sino casi lógicos o, más precisamente, lógico-expresivos. Al no ser completamente lógico sino precisar una sintaxis subyacente, la falta de conservatividad no puede evaluarse sobre el sistema lógico elegido sino sobre la teoría de la sintaxis que, como argumentan Halbach (2001) y Picollo (2010), bajo condiciones mínimas es ya capaz, sin la ayuda de ningún principio que gobierne el predicado veritativo, de probar que toda oración es diferente de su negación y muchas otras cosas más.²² De no dar lugar a trivialidad (por ejemplo, en sus versiones tipeadas), las versiones formales de los principios de transparencia son, hasta donde llega mi conocimiento, conservativas sobre teorías de la sintaxis para los lenguajes correspondientes.

4.3.4. Principios de transparencia vs principios generales

Desde cierto punto de vista, el mismo Tarski consideró que, en materia de contenido, el Esquema T era requisito suficiente para fijar el significado de la verdad, en tanto es el único principio (junto con el requisito de que solo oraciones pueden formar parte de la

²¹ V.g., si ϕ es una constante de individuo, ψ es un término; si ϕ es una letra monádica de predicado, ψ es una fórmula con una variable libre, etc.

²² Esta salvedad al requisito de conservatividad es semejante a la cualificación de la cual habla Field (1994, p. 250). Simplemente habría que reemplazar 'cognitivamente equivalente' por 'analíticamente equivalente'.

extensión de la verdad, un principio que establece meramente el rango de aplicación del predicado) para el predicado veritativo que forma parte de las condiciones de adecuación material enunciadas en la Convención T (definición 2.1). Pero, como señalan Martin Fischer, Halbach, Johannes Stern & Johannes Kriener (2015), esto entra en conflicto con el rechazo de Tarski a teorías axiomáticas dadas únicamente por instancias versiones formales del [Esquema T](#) (solo para oraciones del lenguaje objeto), como también (véase el apartado 2.3.3).

A theory of truth founded on them would be a highly incomplete system, which would lack the most important and most fruitful general theorems. (Tarski, 1933, p. 257)

esto es, ser incapaz de probar versiones formales de ciertos principios generales, como (4.22),

Una oración no es verdadera si y solo si su negación lo es. (4.27)

y

Una oración condicional es verdadera si y solo si la verdad de su antecedente implica la verdad de su consecuente. (4.28)

que se siguen directamente de las definiciones tarskianas.

Prima facie, es razonable requerir que estos principios generales se sigan de todo conjunto de expresiones que se precie de agotar el significado del predicado veritativo, en tanto resultan intuitivamente verdaderos y su verdad no parece depender de ninguna otra cosa que de la naturaleza del predicado veritativo y ciertas leyes lógicas. Como los principios de transparencia en sus versiones formales no implican estos enunciados generales, se debería concluir que no están agotando el significado de la verdad.

Sin embargo, hay dos cuestiones a tener en cuenta antes de llegar a esta conclusión. Por un lado, que el carácter intuitivo de principios generales como (4.27), (4.22) y (4.28) depende de la lógica que se adopte. Si esta es la lógica clásica, efectivamente, los principios resultan correctos. Pero si, por ejemplo, se prefiere una lógica en la cual el principio de tercero excluido no vale, (4.27) puede no resultar intuitivo ya, porque bajo ciertas condiciones mínimas este principio, junto con los principios de transparencia, implica todas las instancias del tercero excluido. Asimismo, si se opta por una lógica donde la regla de introducción del condicional o el Modus Ponens fallan, (4.28) debe ser rechazado. Luego, lo que parece razonable es requerir de un conjunto de oraciones que agote el significado del predicado veritativo es que implique, bajo cierta lógica, los principios generales de la verdad que resulten correctos de acuerdo con esta lógica misma.

V.g., en la vasta mayoría de los casos (4.22) resultar deseable, porque las reglas que regulan el comportamiento de la conjunción son excepcionalmente cuestionadas.

Dejando de lado los casos en los cuales los principios de transparencia resultan triviales, esto es, permiten probar todas las expresiones del lenguaje (como cuando se toman no tipeados y se adopta la lógica clásica), las versiones formales de estos principios no suelen ser suficientes para la derivación de los principios generales correspondientes. Por ejemplo, como nota Tarski (véase el apartado 2.3.3), $tb(pa^+)$ no es capaz de probar

$$\exists x(\text{Sent}_{L+pa}(x) \rightarrow (\exists T x \rightarrow \exists T :x))$$

ni

$$\exists x(\text{Sent}_{L+pa}(x) \rightarrow ((\exists T (x \wedge y) \rightarrow \exists T (x \wedge T y)))$$

aunque la lógica subyacente a la teoría es clásica. En general, si el sistema no es trivial, la lógica subyacente es nitaria y los principios generales no valen previamente a la incorporación de los principios de transparencia, las versiones formales de estos últimos no tienen su suficiente poder deductivo para implicar los principios generales.

No obstante, esta sección no se ocupa de las versiones formales de los principios de transparencia y sus propiedades en diferentes sistemas también formales, lo cual recién será objeto del capítulo siguiente, sino, como se señaló al comienzo, de su formulación en los lenguajes coloquiales, donde no parece haber tal cosa como una lógica subyacente particular, y las inferencias que se llevan a cabo son más bien informales. En este plano, como nota Horwich (2005), las instancias de los principios de transparencia junto con los principios lógicos que se adopten efectivamente implican, aunque tal vez no lógicamente, los principios generales correspondientes.

A diferencia de los lenguajes formales que se emplean usualmente, el lenguaje natural no parece ser 'compacto',²³ esto es, hay reglas nitarias cuyas versiones formales resultan inválidas en los lenguajes formales compactos, pero que preservan verdad en el lenguaje coloquial. Por ejemplo, la que Tarski denomina 'regla de inducción in nita': dado un predicado monádico P , si todos los objetos del dominio de discurso tienen un nombre en el lenguaje \mathcal{L} , y para cada término cerrado t del lenguaje que denota uno de estos objetos, se tiene que t es un P , se puede concluir legítimamente que todos los objetos son P s.²⁴ Como en el lenguaje coloquial las comillas, por ejemplo, proveen nombres para todas las oraciones del lenguaje, un caso especial de la regla de inducción in nita

²³Por supuesto, esta es meramente una manera de hablar, ya que el lenguaje natural no tiene una semántica precisamente de nita y la propiedad de compactidad correspondiente no puede ni siquiera enunciarse.

²⁴En el lenguaje de la aritmética, donde el dominio pretendido es el conjunto de números naturales, la regla se conoce como 'regla I': si ' ϕ ' es una fórmula de L_{pa}^+ con una variable libre v , de las in nitas instancias ' $\phi[n=v]$ dadas por nada número natural n puede inferirse $\exists v \phi$ '.

es el siguiente: dado un predicado monadico P , a partir de las instancias de ' A ' es una P ' dadas por cada oracion A del lenguaje, se puede inferir que todas las oraciones del lenguaje son P 's.

Junto con, v.g., el principio de Intersustitutividad, la regla de induccion in nita permite, bajo ciertas suposiciones logicas m nimas, derivar el principio general (4.22). Para cualesquiera oraciones del lenguaje A y B ,

$$A \text{ y } B \text{ si y solo si } A \text{ y } B:$$

y el [Esquema T](#) implican

$$\text{'A y B' es verdadera si y solo si 'A' es verdadera y 'B' es verdadera.}$$

Aplicando la regla de induccion in nita para dominios de oraciones en la ultima l nea sobre ' A ', para cada oracion B del lenguaje se tiene que

$$\text{Todas las oraciones son tales que su conjuncion con 'B' es verdadera si y solo si ellas son verdaderas y 'B' es verdadera.}$$

Una nueva aplicacion de la regla sobre ' B ' da por resultado (4.22), como se quer a.

Los principios generales que gobiernan el predicado veritativo estan de alguna ma-nera contenidos ya en los principios de transparencia. En palabras de [Tarski \(1933, p. 260\)](#), "from the intuitive standpoint the rule of in nite induction seems to be as reliable as the rules normally applied: it always leads from true sentences to true sentences" y, de acuerdo con [Horwich \(2005, p. 84, nota 14\)](#), la regla "enables the general facts about truth to be explained by their instances." Ciertamente, esta regla es empleada con bastante frecuencia, al menos en el discurso loso co que compete a esta investigacion. Por ejemplo, cuando se dice que los principios de transparencia dan condiciones necesarias y su cientes para la aplicacion del predicado de verdad a cada uno de los objetos en su rango de aplicacion, se esta pasando de una in nidad de instancias a un unico enunciado general. O, tal vez mas evidentemente, se dice que una generalizacion como

$$\text{Todas las } P \text{'s son verdaderas.} \tag{4.29}$$

expresa el contenido de todas las instancias de

$$\text{Si 'A' es } P \text{ entonces } A.$$

a la vez y no más que eso porque, por la función básica de la verdad, cada una de estas instancias es equivalente a

$$\text{Si } 'A' \text{ es } P \text{ entonces } 'A' \text{ es verdadera.} \quad (4.30)$$

y todas las instancias de (4.30) juntas implican (4.29) a través de la regla de inducción in nita.

Si bien en un plano informal los principios de transparencia pueden resultar suficientes para derivar los principios generales deseados para el gobierno del predicado de verdad y, por ende, agotar su significado, sus contrapartes formales podrán necesitar ciertos principios complementarios, si la completitud del sistema resultante es uno de los objetivos, como se verá en las secciones siguientes.

4.3.5. El carácter extensional de los principios de transparencia

De acuerdo con la función básica de la verdad, su naturaleza desentremilladora, la equivalencia entre una oración y su predicación explícita de verdad es de naturaleza meramente extensional. Lo que importa es la preservación de las propiedades semánticas y no, por ejemplo, del sentido; la transformación que el predicado de verdad opera es meramente salva veritate. En consecuencia, si el [Esquema T](#) captura la naturaleza desentremilladora de la verdad, los bicondicionales que lo constituyen deben ser materiales, lo cual queda quizás ya reflejando en las inferencias de [Introducción T](#), [Eliminación T](#) y el principio de Intersustitutividad.

Sin embargo, autores como [Gupta \(1993\)](#), [Gupta & Belnap \(1993\)](#) y [Priest \(2006b\)](#) sostienen que para establecer el significado de un predicado, como 'esta soltera' o 'es verdadera', se precisa un bicondicional más fuerte que el material. Por ejemplo, en

x esta soltera si y solo si x no esta casada

el bicondicional 's y solo si' debe tener cierto estatus de intensional o intensional, porque lo que se pretende con el es decir, no solo la extensión, sino la intensión del predicado en cuestión. Como el [Esquema T](#) captura la naturaleza de la verdad y, por tanto, su significado; como cada una de sus instancias ya las condiciones bajo las cuales el predicado de verdad se aplica a la oración correspondiente y son, por tanto, definiciones parciales del predicado veritativo, de acuerdo con Gupta & Belnap las equivalencias que lo constituyen no deben ser materiales sino de intensionales. Decir que una oración es verdadera es, por definición, lo mismo que decir la oración; existe una relación de sinonimia entre ambas.

[Gupta & Belnap \(1993\)](#), cap. 5) ofrecen una teoría de las deducciones mediante la cual explican que significa que una equivalencia sea deducible, y cuáles son sus diferencias con respecto a las equivalencias materiales necesarias o analíticas. No obstante, su teoría de las deducciones está diseñada para lidiar con predicados ordinarios, que existen en el lenguaje para expresar un concepto, entre los cuales incluyen a la verdad. De acuerdo con el acciónismo, la naturaleza del predicado veritativo es en cambio más afín a la de las constantes lógicas (para las cuales no es posible dar deducciones en el marco de la teoría de Gupta & Belnap). Como, v.g., la conjunción, el predicado veritativo admite una deducción mediante reglas de uso, como [Introducción T](#) y [Eliminación T](#) o, equivalentemente, el [Esquema T](#) o el principio de intersustitutividad. En ambos casos, las reglas que definen los términos en cuestión están expresadas en términos de equivalencias o inferencias materiales, lo cual es suficiente precisamente por la naturaleza inferencial de los términos que definen. El estatus especial que hace que determinadas reglas y no otras constituyan deducciones está dado por su analiticidad, por la cual argumenté en el apartado [4.3.2](#), y por la intención de quien las enuncia.

Otro tanto puede decirse de la formulación de Priest del [Esquema T](#) por medio de un (bi)condicional (la conjunción de dos condicionales) no material sino intensional. [Priest \(2006b\)](#), cap. 6) define su condicional en términos de mundos posibles, como mostre en el apartado [3.4.2](#), pero su intención no es darle un estatus deducible sino principalmente imprimirle ciertas propiedades lógicas: por un lado, restaurar la validez del Modus Ponens, que el condicional material pierde en la lógica sobre la cual Priest trabaja; por otro, obtener un condicional para el cual la regla de contraposición no valga irrestrictamente.

§

En este capítulo he argumentado, en primer lugar, que el predicado veritativo cumple únicamente roles lógico-expresivos en el lenguaje, todos los cuales pueden explicarse por medio de su función básica o naturaleza desentremilladora. En consecuencia, esta naturaleza desentremilladora constituye o encapsula el significado de la verdad. Además, en esta última sección he argumentado que este significado puede expresarse por medio de principios de transparencia, i.e., el [Esquema T](#), las inferencias de [Introducción T](#), [Eliminación T](#) o el principio de Intersustitutividad. Estas tesis, sostengo, conforman el núcleo del acciónismo. Si no me equivoco, difícilmente una persona que se identifica a sí misma con el acciónismo pueda diferir con respecto a ellas.

5

La quasi logica de la verdad

"Like the logical connectives, the concept of truth is a handmaiden of reasoning."

{ Leon Horsten, The Tarskian Turn

Como se~nale en el cap tulo anterior, el predicado veritativo cumple una serie de roles logico-expresivos de suma importancia, algunos de los cuales tienen una utilidad puramente logica.¹ Los usos logicos del predicado veritativo [i.e., la expresion de conjunciones y disyunciones posiblemente in nitas] inducen la formulacion de `logicas' o, mejor dicho, teor as formales de la verdad, dado que se requiere cierta sintaxis en la base, que extiendan la logica preferida con principios o reglas para el gobierno del predicado veritativo que garanticen, primeramente, estos usos logicos y, de ser posible, todos sus usos logico-expresivos. Los sistemas resultantes tendr an cierto estatus logico o quasi logico, porque la verdad cumplir a un rol inferencial en ellos y, por ende, jugar a un papel en el estudio de la validez de argumentos, verdades logicas, falsedades logicas, etc.

>Cuales son estos principios para la verdad capaces de garantizar sus roles logico-expresivos y, en particular, sus usos logicos? Dado que, como mostre en el apartado [4.2](#), todos los usos del predicado veritativo pueden explicarse a traves de su funcion basica y los principios de transparencia expresan esta funcion, tanto dentro como fuera del de acionismo suele considerarse que toda teor a formal de la verdad cuyo predicado pueda cumplir sus roles logico-expresivos es capaz de derivar versiones formales de (al menos alguno de) estos principios y, viceversa, toda teor a capaz de esto ultimo permite la expresion de conjunciones y disyunciones posiblemente in nitas.

¹*Este cap tulo es una ampliacion de un art culo escrito en colaboracion con Thomas Schindler (cf. [Picollo & Schindler \(2015\)](#)).*

No obstante, a causa del teorema de Tarski (teorema [2.9](#)), esto no es posible en el marco de la lógica clásica. Muchos teóricos de la verdad que apuntan a la utilidad lógica del predicado han, en efecto, adoptado lógicas no clásicas, aunque poco se ha dicho acerca del sentido en el cual ciertas adscripciones de verdad de un lenguaje formal serían capaces de expresar conjunciones o disyunciones posiblemente infinitas y como los principios de transparencia lo garantizarían. Falta una investigación detallada de las condiciones que deben darse para que el predicado de verdad de un lenguaje formal sea capaz de cumplir sus roles lógico-expresivos, bajo cuya luz evaluar si es realmente necesario o si puede contar con principios de transparencia en la teoría. Si acaso principios más débiles fuesen suficientes, esta ya no sería una razón para abandonar la lógica clásica. Al afirmar que la única razón de su existencia está dada por sus roles lógico-expresivos, de los cuales los únicos absolutamente irremplazables son los que tienen usos de índole lógica, una de acciónista no debe sentir demasiado apego por principios formales a menos que sean necesarios para garantizar el rol del predicado de verdad. Y si acaso otros principios en lugar de, o adicionalmente a, los de equivalencia fuesen necesarios, habría que cuidarse de excluirlos de las teorías formales de la verdad. Una investigación tal permitiría evaluar que teorías de la verdad son convenientes y cuáles no, y cuáles lo son más que otras, desde un punto de vista de acciónista.

En este capítulo llevo precisamente esta investigación a cabo. Primero doy condiciones generales que toda quasi lógica de la verdad debe cumplir. En [5.2](#) examino en qué sentido puede decirse que ciertas adscripciones generales de verdad expresan conjunciones posiblemente infinitas y qué principios para la verdad deben valer para garantizarlo; en otras palabras, doy condiciones necesarias y suficientes para posibilitar los usos lógicos del predicado veritativo en una teoría formal de la verdad. En [5.3](#) establezco criterios generales que una teoría de la verdad de acciónista debe satisfacer basándome en los resultados de los apartados previos y en otras consideraciones que aquí introduzco, y examino diversas teorías formales de la verdad introducidas en los capítulos [2](#) y [3](#) a la luz de estos criterios.

5.1. Condiciones generales sobre una quasi lógica de la verdad

Como se señaló, toda quasi lógica de la verdad debe estar formulada sobre una teoría a base que o bien sea de teoría de la sintaxis. De ella se espera, no solo que provea nombres para los portadores de verdad, las oraciones del lenguaje, sino posiblemente también predicados que expresen ciertas propiedades sintácticas que estos portadores puedan tener, como ser un enunciado condicional o pertenecer a cierta porción del lenguaje, y

funciones sintácticas como aquella que toma una fórmula y devuelve su negación. Estos predicados y funciones, junto con otros que se deseen incorporar según las necesidades particulares, como ser un teorema de tal o cual teoría, son los que permitan formular generalizaciones en las cuales se adscribe verdad a ciertas oraciones y no a otras. Por ejemplo, si se cuenta con un predicado P que expresa la propiedad de ser una oración condicional del lenguaje en la cual el antecedente y el consecuente coinciden, el enunciado

Todos los P s son verdaderos.

predica verdad exactamente de aquellas expresiones de la forma "Si A entonces A " y expresa, por tanto, la conjunción infinita de todas estas oraciones.

Con respecto a los principios para la verdad, como lo que se desea es un predicado quasi lógico, que permita expresar conjunciones y disyunciones posiblemente infinitas de oraciones de cualquier lenguaje, incluso de aquellas que contienen ellas mismas el predicado de verdad, el de acionismo [y presuntamente toda persona que quiera una teoría en la cual sea posible hacer uso lógico irrestricto del predicado de verdad] suele preferir teorías no tipeadas de la verdad, en las cuales el predicado de verdad no pertenezca en un metalenguaje (como, v.g., el predicado de verdad para el lenguaje de la aritmética

L_{pa} en la definición tarskiana 2.4) sino al mismo lenguaje objeto, el lenguaje de la quasi lógica resultante. Esto permite iterar el predicado veritativo, esto es, predicar verdad de oraciones que lo contienen, dando lugar a fórmulas bien formadas del lenguaje, pero no es suficiente. Adicionalmente, los principios que se adopten para el predicado de verdad deben admitir aplicaciones correctas del predicado a oraciones que lo contengan; deben o bien implicar la verdad de oraciones que contengan el predicado veritativo o bien carecer de restricciones a la aplicación del predicado a oraciones que ya lo contienen. Una teoría de la verdad verdaderamente libre de tipos debe cumplir estas condiciones (cf. Halbach (2011, cap. 10)).

Como se vio en el capítulo 3, sistemas de ese tipo son ciertamente posibles. Asimismo, al estar formulados sobre la aritmética, que opera como teoría de la sintaxis, cuentan con predicados que tienen una amplia variedad de propiedades sintácticas.² Siendo esto lo más común, al igual que en el capítulo 3, en este apartado considero exclusivamente teorías de la verdad formuladas en (posibles extensiones de) L_T que extienden la teoría base pa^+ con axiomas o reglas o el modelo base N^+ a modelos de todo L_T (veanse los apartados 1.2.1 y 1.2.4).

²Se ha probado que una teoría de la sintaxis decente es relativamente interpretable en pa y viceversa. Por ejemplo, Halbach y Graham Leigh están trabajando en teorías naturales de la sintaxis para las cuales prueban este resultado. Por esta razón suele decirse también que la verdad es una noción lógico-matemática (cf. Halbach & Horsten (2005), Horsten (2011)).

Como el resultado es meramente una lógica junto con principios para T y una teoría de la sintaxis necesaria para su formulación, debe entenderse exclusivamente como una base, precisamente una quasi lógica sobre la cual debe ser posible formular todo tipo de teorías, v.g., la teoría de conjuntos o cualquier otra cosa. Mientras que los abordajes axiomáticos (por ejemplo, $tb(pa^+)$, $utb(pa^+)$, $ct(pa^+)$ en el apartado 2.3.3, kf y $putb$ en 3.1.3.1 y 3.1.3.2, respectivamente, fs en 3.3 y $lptt$, $pltt$ y $sttt$ en 3.4) deberán ofrecer axiomas o reglas lógicamente exhibibles como para adaptarse fácilmente a cualquier lenguaje de primer orden $L_{L^+_{pa}}$, las teorías semánticas (v.g., las teorías de Punto Fijo de Kripke y la teoría d -fundada de Leitgeb; en los apartados 3.1.2 y 3.2, correspondientemente) deberán mostrar como extender modelos arbitrarios de L a L_T manteniendo las propiedades deseadas del predicado veritativo.

[Halbach & Horsten \(2005\)](#) argumentan en contra de los abordajes semánticos, señalando que su aplicabilidad a un lenguaje interpretado solo es posible en la medida en que la teoría de conjuntos, la metateoría en la cual se obtienen los modelos correspondientes, provea un modelo isomorfo a la interpretación pretendida de este lenguaje. Como no existe tal cosa para L_{zfc} , las teorías semánticas no son aplicables a todos los lenguajes de primer orden. Naturalmente, podría simplemente tomarse como metateoría una teoría más poderosa, como la teoría de clases tc_1 en el apartado 2.3.2, aunque esto daría lugar a jerarquías, con los inconvenientes que usualmente conllevan.

Otro argumento que proponen en contra de la adecuación de las teorías semánticas al proyecto de acionista se basa en su tesis, expuesta en el apartado 4.3.3, según la cual el de acionismo está comprometido con la ineliminabilidad y, por ende, con la indebididad explícita de la verdad [lo cual, como señalé entonces, no es claro]. Desde su punto de vista, las teorías de la verdad semánticas disponibles proveen de definiciones explícitas del predicado veritativo en la metateoría, esto es, la teoría de conjuntos. Si bien no es mi objetivo discutir esto en detalle acá, notese que, v.g., las teorías de Punto Fijo de Kripke (vease el apartado 3.1.2) no permiten la eliminabilidad del predicado de verdad porque, estrictamente hablando, no proveen una definición explícita de T de la forma de $T(x, \phi)$, donde ϕ no contiene T , sino que meramente definen la extensión que se le asigna a T en el metalenguaje. Pero el lenguaje a utilizar, que contiene su propio predicado de verdad, es el lenguaje objeto.

Un argumento que considero es concluyente, no necesariamente en contra de los abordajes semánticos sino a favor de los axiomáticos, es el siguiente. Desde el punto de vista de acionista, la verdad es semejante a una constante lógica, juega un rol expresivo e inferencial en el lenguaje. Luego, como toda lógica, la quasi lógica de la verdad debe poder ser recursivamente axiomatizable, para poder ser utilizada efectivamente al

expresar y razonar (cf. [Fischer et al. \(2015\)](#)). Los abordajes semánticos no anulan necesariamente esta posibilidad pero los conjuntos de oraciones que implican y las inferencias que validan suelen no ser recursivamente axiomatizables.

Esto puede deberse, como a veces sucede, al modelo base N^+ , que o sea de teoría de la sintaxis, cuya complejidad es 1_1 y, por ende, no es axiomatizable en un lenguaje de primer orden, como muestran los resultados de incompletitud de Gödel. No obstante, en buena parte de los casos sucede que los principios que conciernen a la verdad exclusivamente introducen aun más complejidad, como es el caso de los modelos de punto fijo de Kripke y el modelo d-fundado de Leitgeb (veanse los apartados [3.1](#) y [3.2](#)), que son ambos 1_1 .

Consecuentemente, estas teorías de la verdad resultan en buena parte inutilizables. Los abordajes axiomáticos no son en sí mismos; sus contrapartes axiomáticas lo son. Podrán servir como heurística, pero el foco debe estar en el cálculo y en su utilización, como argumentan [Halbach & Horsten \(2005\)](#).

Asimismo, es claro que la no trivialidad es un requisito fundamental. Adicionalmente, como ya mencione, los principios que se adopten para la verdad no pueden exceder la función básica de la verdad: aunque sea de manera informal, las consecuencias que una teoría de acionista de la verdad tenga con respecto al predicado veritativo deben desprenderse de su función básica o, más precisamente, de los principios de transparencia. Incluso podrá argumentarse que, si bien informalmente, todos los principios para la verdad que se sigan en una teoría de acionista deben ser explicables en términos de las instancias de las versiones formales de los principios de transparencia que la teoría misma valide, aunque podrá replicarse que la teoría es simplemente incompleta con respecto a las instancias correctas de los principios de equivalencia relevantes (si es que no prueba su negación o la lógica subyacente es paraconsistente).

En la sección que sigue dejo momentáneamente de lado estas cuestiones generales para volver a ellas en [5.3](#), y me ocupo de la pregunta central del capítulo, puntualmente si acaso en los sistemas introducidos en los capítulos [2](#) y [3](#) o en otros sistemas los usos lógicos del predicado veritativo son posibles y, más generalmente, que condiciones debe cualquier teoría de la verdad satisfacer para garantizar esta posibilidad.

5.2. Expresando conjunciones internas mediante generalizaciones

>Que principios deben valer en una teoría formal de la verdad para que ciertas adscripciones de verdad expresen conjunciones y disyunciones posiblemente internas? Como

anticipe, muchos filósofos apuntan a versiones formales de los principios de transparencia (cf. [Horwich \(1998b\)](#), [Field \(1994\)](#), [Beall \(2005\)](#), [Field \(2008\)](#), [Horsten \(2011\)](#)) basándose en un razonamiento como el siguiente: dado que la función básica del predicado de verdad es necesaria y suficiente para garantizar sus funciones lógico-expresivas, en las cuales se fundan todos los usos de este predicado, y los principios de transparencia expresan esta función básica, los principios de transparencia son necesarios y suficientes para posibilitar todos los usos del predicado veritativo. Por ejemplo, [Horwich \(1998b\)](#), p. 124) sostiene que

The generalising function of truth is perfectly well fulfilled as long as instances of the [Esquema T], understood as material biconditionals, are accepted.

Más aun, algunos teóricos de la verdad que no se identifican a sí mismos con el deacionismo mantienen tesis semejantes, como [Priest \(2006b\)](#), x 4.2) y [Cobrerros et al. \(2013\)](#). En ese caso, la consecuencia inevitable sería el abandono de la lógica clásica o la mutilación del poder expresivo de la verdad. Pero esto todavía está por verse.

La idea general más aceptada sobre cómo la noción de verdad permite que ciertas generalizaciones expresen conjunciones posiblemente infinitas puede resumirse en la siguiente cita de [Horwich \(1998b\)](#), p. 3):

Consider, for example,

(1) What Oscar said is true.

Here we have something of the form

(2) x is F ,

whose meaning is such that, given further information about the identity of x |given further premises of the form

(3) x =the proposition that p

|we are entitled to infer

(4) p .

And it is from precisely this inferential property that propositions involving truth derive their utility. For it makes them, in certain circumstances, the only appropriate object of our beliefs, suppositions, desires, etc. Suppose, for example, I have great confidence in Oscar's judgment about food; he has just asserted that eels are good but I didn't quite catch the remark. Which

belief might I reasonably acquire? Well, obviously not that eels are good. Rather what is needed is a proposition from which that one would follow, given identification of what Oscar said|a proposition equivalent to

If what Oscar said is that eels are good then eels are good, and if he said that milk is white then milk is white, ... and so on;

and the raison d'etre of the concept of truth is that it supplies us with such a proposition: namely (1).

Luego, el análisis informal de Horwich está en la misma línea que el que di en los apartados [4.2.2.2](#) y [4.2.3.3](#): una adscripción general de verdad o generalización como (1) es, en virtud de la naturaleza desentrecomilladora de la verdad, de alguna manera equivalente a la conjunción in nita

$$\begin{matrix} \wedge \\ \text{Si Oscar dijo } A \text{ entonces } A; \\ \text{A2L} \end{matrix}$$

y, por lo tanto, implica [dada una identificación de lo que Oscar dijo] aquello afirmado por Oscar, esto es, que las anguilas son ricas. Es esta propiedad inferencial lo que torna útiles a las generalizaciones como (1). Claramente, en el pasaje a un plano formal, i.e., a una teoría formal de la verdad \mathcal{L}_T formulada en (una extensión de) \mathcal{L}_T , esta explicación suscita una serie de preguntas importantes que es necesario distinguir:

(a) >En que sentido pueden una generalización de la forma

$$\exists v ('(v) \text{ T } v) \tag{5.1}$$

y la conjunción in nita correspondiente

$$\begin{matrix} \wedge \\ ('(p \text{ q}) \text{ ! }) \\ \text{2LT} \end{matrix} \tag{5.2}$$

ser equivalentes, relativamente a una teoría de la verdad? O, alternativamente, >en que sentido puede una generalización de la forma de (5.1) expresar una conjunción in nita?

Notese que ningún lenguaje de primer orden permite conjunciones in nitas, con lo cual la equivalencia, de existir, debe ser de alguna manera externa. Asimismo, hay al menos un sentido claro en el cual las conjunciones in nitas reales, como (5.2), pertenecientes a lenguajes in nitarios (cf. [John L. Bell \(2012\)](#)), difieren de expresiones de la forma de (5.1): como se señaló en el capítulo 2, mientras que una conjunción in nita jamás puede contenerse a sí misma como conyunto, porque las

formulas bien formadas de los lenguajes in nitarios estan de nidas por medio de definiciones recursivas; las oraciones de la forma (5.1) pueden ellas mismas expresar su propio contenido a traves del predicado de verdad, simplemente satisfaciendo '(v). Por ejemplo, "Todas las instancias del Principio del Tercero Excluido son verdaderas" expresa todas las instancias del Principio del Tercero excluido, incluso aquella dada por esta misma oracion. En un lenguaje in nitario, en cambio, no es posible expresar ciertos principios logicos como el Tercero Excluido mediante una unica oracion.

- (b) >Que condiciones necesarias y suficientes debe satisfacer una teoria de la verdad para conllevar esta equivalencia? En particular: >presupone esta equivalencia el Esquema-T, esto es,

$$T p \wedge q \rightarrow T p \wedge T q \quad (\text{Esquema-T})$$

u otros principios de transparencia formales como³

$$\forall t (T \neg t \rightarrow \neg T t) \quad (\text{Esquema-T Uniforme})$$

$$\frac{T p \wedge T q}{T p \wedge q} \quad \frac{T p \wedge T q}{T p \wedge q}$$

las reglas

$$\frac{}{T p \wedge q} \quad (\text{T-Intro}) \qquad \frac{T p \wedge q}{p} \quad (\text{T-Elim})$$

o el principio de Intersustitutividad dado por las siguientes reglas?

$$\frac{\dots \wedge \dots}{\dots \wedge \dots} \quad (\text{Intersustitutividad})$$

- (c) >De ser posible en algun sentido la equivalencia entre (5.1) y (5.2), implica las propiedades inferenciales de las cuales (5.1) deriva su utilidad, esto es, su capacidad de expresar todas las oraciones que satisfacen el predicado '(v)? >Presuponen estas propiedades inferenciales de las cuales las generalizaciones de la forma de (5.1) obtienen su utilidad la equivalencia en cuestion? En otras palabras, >son necesarios o suficientes los principios que garantizaran la equivalencia entre (5.1) y (5.2) para permitir que el predicado de verdad cumpla sus propositos logicos?

Dejando de lado las reconstrucciones informales de casos particulares dadas por [Quine \(1970\)](#), [Horwich \(1998b\)](#) y [Field \(2008\)](#), entre otros, [Halbach \(1999\)](#) ha sido el unico en intentar elucidar que significa en terminos formales que un predicado de verdad permita la expresion de conjunciones in nitas y como es que los principios de

³Vease el apartado [1.2.3](#) para aclarar detalles sobre la notacion.

transparencia garantizan esta posibilidad, aunque solo ha respondido estas preguntas para sistemas tipeados dados por instancias del [Esquema-T](#), i.e., para tb (vease el apartado [2.3.3](#)). En el siguiente apartado muestro que el abordaje de Halbach no es fácilmente extensible a otras teorías y, en [5.2.2](#), ofrezco una segunda elucidación, que resulta en cambio adecuada.

Siendo la verdad, desde un punto de vista de acionista, un predicado cuya existencia se debe exclusivamente a su capacidad de cumplir roles lógico-expresivos, de los cuales los más importantes son de índole lógica, lo que se espera de una 'lógica' de acionista es que provea un dispositivo que garantice esos roles, que permita expresar conjunciones (y disyunciones) posiblemente infinitas como aquellas enunciadas en el apartado [4.2.3.3](#), una reglamentación (en el sentido de Quine) de un cierto fragmento del lenguaje coloquial que no debe necesariamente preservar todas las características que el predicado de verdad posee en este lenguaje. Es completamente irrelevante si acaso la verdad es esencialmente desentrecorilladora o si los principios de transparencia dan el significado del predicado veritativo para la confección de una teoría o quasi lógica de la verdad. De hecho, la segunda elucidación pondrá de manifiesto que el [Esquema-T](#) no es estrictamente ni necesario ni suficiente, conclusión que extraigo en [5.2.3](#). En consecuencia, las paradojas semánticas no son razón suficiente para el abandono de la lógica clásica, al menos desde un punto de vista de acionista.

5.2.1. Primer intento de elucidación: mismas consecuencias

[Halbach \(1999\)](#) ofrece una respuesta a la primera de las tres interrogantes, la (a), esto es, propone un sentido en el cual un enunciado de la forma [\(5.1\)](#) es capaz de expresar la conjunción infinita en [\(5.2\)](#), para luego evaluar si acaso las teorías de acionistas desentrecorilladoras son capaces de garantizar esta equivalencia. Para evitar los problemas provenientes de paradojas como el mentiroso, considera exclusivamente predicados ' (v) ' y oraciones que no contienen el predicado veritativo T , y la teoría desentrecorilladora $tb(th)$ formulada en $L [\text{fT} g$, donde th es (o contiene) una teoría de la sintaxis formulada en L , un lenguaje de primer orden que no contiene T .

Desde su punto de vista, enunciados de la forma [\(5.1\)](#) expresan la conjunción infinita correspondiente en [\(5.2\)](#) en una teoría axiomática de la verdad th^0 formulada sobre una teoría base th siempre y cuando la teoría base junto con los infinitos conjuntos de [\(5.2\)](#) tenga las mismas consecuencias que la teoría de la verdad junto con [\(5.1\)](#). Formalmente, sea th^0 una teoría axiomática de la verdad formulada en $L [\text{fT} g$. De la Definición 5.1 (Equivalencia entre generalizaciones y conjunciones infinitas según Halbach). Una generalización de la forma $\forall x(\neg(x) \rightarrow Tx)$, donde ' $(v) \in L$ ', es equivalente a

la conjunción innita $\bigvee_{2L}(\ulcorner p \urcorner \ulcorner q \urcorner)!$ en th^0 (formulada sobre la teoría a base $th\ L$) de acuerdo con Halbach *sii*, para toda $2\ L$,

$$th + \ulcorner \ulcorner p \urcorner \ulcorner q \urcorner \urcorner! \quad : \quad 2\ Lg \quad , \quad th^0 + 8v(\ulcorner (v) \urcorner! \ulcorner Tv \urcorner) \quad `$$

Luego, Halbach propone una solución al problema del carácter externo de la equivalencia entre generalizaciones como (5.1) y sus correspondientes conjunciones innitas en (5.2) reemplazando esta última por el conjunto de sus innitos conyuntos.

A continuación, Halbach muestra que, de acuerdo con la definición 5.1, la teoría desentrecorolladora $tb(th)$ es suficiente para garantizar la equivalencia de generalizaciones dadas por (5.1) y las correspondientes conjunciones innitas dadas por (5.2), si en $\ulcorner (v) \urcorner$ y en T no ocurre.

Proposición 5.2 (Halbach). Si $\ulcorner (v) \urcorner \in 2\ L$, entonces, para toda oración $\in 2\ L$,⁴

$$th + \ulcorner \ulcorner p \urcorner \ulcorner q \urcorner \urcorner! \quad : \quad 2\ Lg \quad , \quad tb(th) + 8v(\ulcorner (v) \urcorner! \ulcorner Tv \urcorner) \quad `$$

Halbach propone una interpretación de este resultado en las siguientes líneas:

I take the result to be an exact formulation of the disquotationalist [minimalist] claim that innite conjunctions can be expressed in a language containing a truth predicate which is characterized by the Tarskian equivalences. The innite set $\ulcorner (p \ q) \urcorner!$ of axioms replaces the innite conjunction; therefore it can be avoided in order to introduce a formal system for a language comprising innite conjunctions. (Halbach, 1999, p. 14)

Si acaso todos los usos de enunciados de la forma (5.1) pueden ser explicados sobre la base de su equivalencia con (5.2) solo puede saberse si se tuviera una lista completa de estos usos; pero al menos puede verse claramente que los más importantes quedan explicados: si th y th^0 son teorías axiomáticas clásicas, la noción de consecuencia lógica es efectiva (i.e., recursivamente axiomatizable), y se cumple la definición 5.1, generalizaciones como (5.1) constituyen axiomatizaciones innitas del conjunto de conyuntos de (5.2).

No obstante, como respuesta razonable y satisfactoria a la pregunta (a), tiene un alcance limitado, en tanto se aplica exclusivamente a teorías axiomáticas de la verdad en el marco de la lógica clásica. Asimismo, permite responder solo parcialmente a la

⁴Vease Halbach (1999) para una demostración. Para evitar formulaciones engorrosas, supongo en esta sección que $\ulcorner 2\ L \urcorner$ significa que \ulcorner es una oración de L .

pregunta (b): da únicamente condiciones suficientes para que una teoría de la verdad clásica garantice la equivalencia entre cierto tipo de generalizaciones y conjunciones infinitas correspondientes, pero no necesarias y debe, por ende, ser completada.

Lamentablemente, la definición 5.1 tiene, adicionalmente, consecuencias contra-intuitivas. Notese que, de acuerdo con ella, cualquier teoría de la verdad que permita expresar conjunciones infinitas debe ser conservativa sobre su teoría base. En otras palabras, si $\forall L$ es una oración y th^0 implica \forall , entonces \forall debe ser ya una consecuencia de la teoría base th . Si bien algunos autores han argumentado a favor de la conservatividad de toda teoría de la verdad, estas ideas enfrentan serias réplicas (vease el apartado 4.3.3). Pero aun si toda teoría de la verdad debiera efectivamente ser conservativa sobre su teoría base, esto no es algo que deba seguirse de un criterio para la expresabilidad de conjunciones infinitas.

En primer lugar, un buen criterio debería ser aplicable también a teorías no decididas o indefinidas de la verdad, para las cuales ningún argumento de conservatividad ha sido esbozado. Aquellos que no se identifican con el decididismo no niegan necesariamente que la verdad cumple roles expresivos de índole lógica, sino que suelen sostener que el predicado veritativo no es meramente un dispositivo lógico-lingüístico. Luego, tiene sentido preguntarse también bajo qué condiciones una teoría sustantivista de la verdad permite la expresión de conjunciones infinitas.

Segundo, y más importantemente, si th^0 es una teoría de la verdad que permite expresar conjunciones infinitas en L y th^{00} , entonces intuitivamente th^{00} debería habilitar también la expresión de conjunciones infinitas en L . No obstante, la definición 5.1 y el requisito de conservatividad que conlleva no lo permiten. Por ejemplo, considerese la teoría (tipada) de la verdad $ct(pa^+)$. Como indique en el apartado 2.3.3, $tb(pa^+) \supset ct(pa^+)$. De acuerdo con la proposición 5.2, $tb(pa^+)$ posibilita la expresión de conjunciones infinitas, mientras que $ct(pa^+)$, al no ser conservativa sobre pa^+ , no es capaz de hacerlo de acuerdo con la definición 5.1. Estas dificultades pueden superarse reemplazando la última ocurrencia de 'th' en la versión extendida de la definición por th^0 .

Definición 5.3 (Equivalencia entre generalizaciones y conjunciones infinitas modificada). Si $\forall L$, una generalización de la forma $\forall v('v) \supset T v$ es equivalente a la conjunción infinita $\bigwedge_{i \in \mathbb{N}} ('p_i \supset q_i) \supset T$ en th de acuerdo con una versión modificada de Halbach si $th + \forall v('v) \supset T v$ y $th + \bigwedge_{i \in \mathbb{N}} ('p_i \supset q_i) \supset T$ tienen las mismas consecuencias en las cuales T no ocurre.

Pero nuevas dificultades emergen. Si bien esta definición no implica la conservatividad de la teoría de la verdad th sobre su teoría base, todavía es posible encontrar

teoras th tales que th satisface el criterio pero th^0 no. Por ejemplo, las teorías $pat + \text{T-Out} L_{pa}^+$ y $pat + \text{T-Out} L_{pa}^+ + :T p q$ (vease el apartado 1.2.4), donde T-Out es el principio

$$T p'q !' \quad (T-Out)$$

T-Out L_{pa}^+ (que abrevio T-Out) es el principio dado por las instancias de T-Out generadas por oraciones de L_{pa}^+ , una versión tipeada de T-Out, y es una oración de L_{pa}^+ tal que ni ella ni $:P rov_{pa^+} (p q)$ son teoremas de $pat + \text{T-Out} + fP rov_{pa^+} (p q) !$

: $2 L_{pa}^+ g$. Por un lado, $pat + \text{T-Out}$ satisface la definición 5.3.

Lema 5.4. Si $pat^+ th L_{pa}^+, pat + th + \text{T-Out}$ es una extensión conservativa de th .

Demostración. Toda teoría clásica th contiene ya un predicado que satisface T-Out, i.e., la fórmula $x \models x$. Luego, todo modelo de th puede ser expandido a un modelo de T-Out más las nuevas instancias del axioma esquema de inducción dadas por fórmulas que contienen T simplemente interpretando el predicado veritativo con el conjunto que se utiliza para interpretar $x \models x$, esto es, el conjunto vacío. \square

Proposición 5.5. Si $(v) 2 L_{pa}^+, entonces, para toda 2 L_{pa}^+, pat + \text{T-Out} + f(p q) ! : 2 L_{pa}^+ g, pat + \text{T-Out} + 8v('v) ! Tv)$

Demostración. Como $pat + \text{T-Out} tb(pa^+)$, la proposición 5.2 implica que

$$pat + \text{T-Out} + 8v('v) ! Tv) tb(pa^+) + 8v('v) ! Tv) (pa^+ + f(p q) ! : 2 L_{pa}^+ g) pat + \text{T-Out} + f(p q) ! : 2 L_{pa}^+ g$$

Además, por el lema 5.4,

$$pat + \text{T-Out} + f(p q) ! : 2 L_{pa}^+ g) pa^+ + f(p q) ! : 2 L_{pa}^+ g) pat + \text{T-Out} + 8v('v) ! Tv) (pa^+ + f(p q) ! : 2 L_{pa}^+ g)$$

porque solo un número finito de oraciones en $f(p q) ! : 2 L_{pa}^+ g$ pueden utilizarse

en una prueba de y , claramente, todas ellas se siguen de $8v('v) ! Tv)$ junto con las instancias relevantes de T-Out. \square

Pero, por otro lado, $\text{T-Out} + :T p q$ no satisface la definición 5.3, porque

$$\text{T-Out} + :T p q + 8x(Prov_{pa^+}(x) ! Tx) : Prov_{pa^+}(p q)$$

mientras que, por hipotesis,

$$\text{T-Out} \quad +:T p q + fP \text{rov}(p q) ! \quad : \quad 2 L^+_{\text{pag } 0} : \text{Prov}_{pa^+} (p q)$$

En consecuencia, hay dos caminos posibles: o bien rechazar la explicacion propuesta por Halbach en la definicion 5.1 y en su version mejorada dada por la definicion 5.3 del sentido en el cual una adscripcion general de la forma $\forall v('v) ! Tv$ puede ser equivalente a una conjuncion infinita por ser incorrectas; o bien concluir que las propiedades inferenciales que hacen utiles a estas generalizaciones en el razonamiento diario no presuponen la equivalencia total entre las generalizaciones y las correspondientes conjunciones infinitas. Para clarificar este dilema, propongo una segunda explicacion de la supuesta equivalencia entre generalizaciones y conjunciones infinitas, ahora en terminos de su comportamiento inferencial. El resultado no se aleja demasiado en espíritu del analisis de Halbach, pero resalta un aspecto importante de la relacion entre una generalizacion y la conjuncion infinita a la cual es presuntamente equivalente y permite extraer conclusiones importantes.

5.2.2. Segundo intento de elucidacion: mismo comportamiento inferencial

El deflacionismo sostiene que el predicado veritativo es un dispositivo quasi logico, comparable con conectivas y operadores. El uso e incluso el significado de una conectiva es comunmente caracterizado por reglas de introduccion y eliminacion (vease el apartado 4.3.3). En el caso de las conjunciones posiblemente infinitas suele adoptarse las siguientes dos reglas:

$$\begin{array}{l} f \quad i i ' \\ \quad \quad \quad i \\ \hline i : i \quad 2 \lg \\ \hline V \quad 2 \end{array} \quad (I) \quad V \qquad \begin{array}{l} V \quad i \\ \quad \quad \quad i \\ \hline i 2 i \quad i \\ \hline 2 \end{array} \quad \text{si } i \quad I \quad (E) \quad V$$

A continuacion examino cada regla por separado, comenzando por la regla de eliminacion.

5.2.2.1. La propiedad de eliminacion de la verdad

E permite inferir $'(p q) !$ a partir de $2L('p q) !$). Como la generalizacion $\forall v('v) ! Tv$ es supuestamente equivalente \forall a la conjuncion infinita, debera ser posible, por analogia, inferir $'(p q) !$ a partir de ella, para todo . Dado lo poco controvertido

del caso, propongo este requisito como una primera condición sobre la expresabilidad de conjunciones infinitas: T es capaz de expresar conjunciones infinitas en L_T (o cualquier lenguaje de primer orden que extienda L_T) mediante expresiones de la forma

$\forall v ('(v) \vdash T v)$ solo si, para todo predicado $'(v)$ y oración de L_T , la siguiente regla vale en th:

$$\frac{\forall v ('(v) \vdash T v)}{'(p \wedge q) \vdash T} \quad (\text{ET } 1)$$

A diferencia del apartado anterior, la atención no está restringida a oraciones que no contienen el predicado veritativo sino a cualesquiera enunciados del lenguaje de la teoría de la verdad th.

Asumiendo que la teoría de la verdad y su noción de consecuencia lógica subyacente son recursivamente axiomatizables, resulta evidente porque una teoría que satisfaga la condición [ET 1](#) es útil en el razonamiento real o cotidiano: bajo condiciones mínimas [esto es, la transitividad de la relación de consecuencia lógica (véase el apartado [3.4](#))] [ET 1](#) implica que, relativamente a la teoría de la verdad th, toda conjunción infinita $\forall_{L_T} ('(p \wedge q) \vdash T)$ o, si se quiere, el conjunto de sus conjuntos $\{ '(p \wedge q) \vdash T : \forall_{L_T} \}$, es reducible a o nitamente axiomatizable por la generalización correspondiente $\forall v ('(v) \vdash T v)$; si una oración se sigue del conjunto infinito $\{ '(p \wedge q) \vdash T : \forall_{L_T} \}$, entonces también se sigue de $\forall v ('(v) \vdash T v)$.

Hasta ahora he hablado de la generalización $\forall v ('(v) \vdash T v)$ como expresando la conjunción infinita $\forall_{L_T} ('(p \wedge q) \vdash T)$, pero, como se señaló en el apartado [4.2.3.3](#), es igualmente posible \forall las oraciones que son 's, i.e., de la conjunción de todas las oraciones que satisfacen '(x). Desde este punto de vista, los principios que gobiernen el predicado veritativo de una teoría en la cual sea capaz de expresar conjunciones infinitas deben permitir la derivación de todas las 's a partir del supuesto de que todas las 's son verdaderas, provista una identificación de las 's [esto es, la generalización $\forall v ('(v) \vdash T v)$] debe 'capturar' todas las 's. En efecto, si las reglas de introducción

$$; '=) = ' ! : \quad (II)$$

y eliminación de condicional

$$\frac{' !}{'} \quad (E!)$$

[o metateorema de la deducción y Modus Ponens, respectivamente] valen en th, la primera condición es equivalente a la segunda condición para la expresabilidad de conjunciones infinitas: T es capaz de expresar conjunciones infinitas en th mediante

L_T , la siguiente regla vale en th:

$$\frac{\forall v('v) ! T v; '(p q)}{\quad} \quad (ET 2)$$

Antes que [ET 1](#), es precisamente esta propiedad inferencial de la verdad lo que Horwich pretende enfatizar mediante el ejemplo de Oscar. Si se quiere afirmar una o varias oraciones a traves del predicado veritativo, la oracion que se utilice, que contiene este predicado, debe indefectiblemente implicar, relativamente a la teoria de la verdad en la que se est trabajando, por supuesto, todas las oraciones que originalmente se quer an aseverar; de otro modo la asercion yerra el objetivo. Si digo que todo lo que Lorna dijo es verdadero, deseo [provista una identificacion de lo que Lorna dijo] poder derivar todas las sentencias que Lorna hizo. En palabras de [Horwich \(1998b, p. 124\)](#), "[the generalising] function of truth requires merely that the generalisations permit us to derive the statements to be generalized [. . .]"

Afirmar $\forall v('v) ! T v$ implica un compromiso con todas las oraciones que satisfacen el predicado $'$. Supongase ahora que una de las $'s$ es refutable. Luego, el enunciado que sostiene que todas las $'s$ son verdaderas debe ser asimismo refutable. Por ejemplo, si una de las cosas que Lorna dijo es que es $\sqrt{2}+2=5$, la afirmacion de que todo lo que Lorna dijo es verdadera debe quedar refutada. Por ende, puede imponerse como una tercera condicion para la expresabilidad de conjunciones infinitas que la afirmacion de que todas las $'s$ son verdaderas sea refutable en el caso en que una de las $'s$ sea ella misma refutable, provisto que se la haya identificado como una $'$, i.e: T es capaz de expresar conjunciones infinitas en th mediante expresiones de la forma $\forall v('v) ! T v$ solo si, para todo predicado $'(v)$ y oracion

de L_T , la siguiente regla vale en th:

$$\frac{'(p q); \quad}{\forall v('v) ! T v} \quad (ET 3)$$

En un contexto clasico, este tercer requisito sobre la expresabilidad de conjunciones infinitas del predicado de verdad es equivalente al segundo y al primero. No obstante, existen logicas no clasicas donde ese no es el caso; razon por la cual demando el cumplimiento de las tres condiciones por separado. Juntas, estas tres condiciones expresan lo que denomino 'propiedad de eliminacion' del predicado veritativo.

Definicion 5.6 (Propiedad de eliminacion). El predicado de verdad de una teoria th satisface la propiedad de eliminacion sii las reglas [ET 1](#), [ET 2](#) y [ET 3](#) valen en th.

He indicado como traducir de alguna manera la regla de eliminacion de la conjuncion en nitaria \underline{E}_V al lenguaje de la verdad, lo cual da lugar a la propiedad de eliminacion del predicado veritativo. Antes de considerar las reglas de introduccion, examino brevemente las caracteristicas que una teoria de la verdad debe poseer para que su(s) predicado(s) de verdad tenga(n) la propiedad de eliminacion. La primera observacion es poco sorprendente.

Observacion 5.7. Si \mathcal{L} es una teoria clasica de la verdad en la cual $\underline{T-Out}$ (o, equivalentemente, $\underline{T-Elim}$) valen, el predicado de verdad de \mathcal{L} tiene la propiedad de eliminacion.

Demostracion. Considerense las siguientes reglas:

Modus Tollens $\frac{\Gamma, \Delta \vdash \neg \phi \quad \Gamma, \Delta \vdash \phi}{\Gamma, \Delta \vdash \perp}$	silogismo hipotetico $\frac{\Gamma, \Delta \vdash \phi \quad \Gamma, \Delta, \phi \vdash \psi}{\Gamma, \Delta \vdash \psi}$	(MT)	(SH)
---	--	------	------

introduccion de la conjuncion (en terminos de las conectivas reales de \mathcal{L}_T , i.e., \wedge y \neg)

$$\frac{\Gamma \vdash \phi \quad \Gamma \vdash \psi}{\Gamma \vdash \phi \wedge \psi} \quad (I^\wedge)$$

eliminacion del universal

$$\frac{\Gamma \vdash \forall x \phi \quad \Gamma \vdash t = v}{\Gamma \vdash \phi} \quad (E8)$$

introduccion del existencial (en terminos de \exists y \forall)

$$\frac{\Gamma \vdash \phi \quad \Gamma \vdash t = v}{\Gamma \vdash \exists x \phi} \quad (I9)$$

La siguiente derivacion constituye una prueba de \underline{ET}_1 :

- | | |
|---|------------------------|
| 1. $\forall v(\forall v \neg T v)$ | premise |
| 2. $\neg(p \wedge q) \wedge T p \wedge q$ | 1, $\underline{E8}$ |
| 3. $T p \wedge q$ | $\underline{T-Out}$ |
| 4. $\neg(p \wedge q) \wedge$ | 2, 3, \underline{SH} |

La siguiente derivacion prueba \underline{ET}_2 :

1. $\exists v('v) ! T v)$	premisa
2. $'(p q)$	premisa
3. $'(p q) ! T p q$	1, E8
4. $T p q$	2, 3, E!
5. $T p q !$	T-Out
6.	4, 5, E!

Por su parte, [ET 3](#) puede obtenerse del siguiente modo:

1. $'(p q)$	premisa
2. $:$	premisa
3. $T p q !$	T-Out
4. $:T p q$	2, 3, MT
5. $:('(p q) ! T p q)$	1, 4, I^
6. $:\exists v('v) ! T v)$	5, I9

□

Luego, en contextos clásicos, [T-Out](#) o, equivalentemente, [T-Elim](#), es suficiente para garantizar la propiedad de eliminación. En contextos no clásicos esto puede variar.

Corolario 5.8. Si th es una teoría de la verdad en la cual valen [T-Out](#), [E!](#), [MT](#), [SH](#), [I^](#), [E8](#) y [I9](#), el predicado de verdad de th tiene la propiedad de eliminación.

Esta es una consecuencia directa de la prueba de la observación anterior. Ahora bien, si en lugar de [T-Out](#) se cuenta con [T-Elim](#) en un contexto no clásico donde los principios no son necesariamente equivalentes, las reglas que se necesitan son otras. Considerese las siguientes metainferencias:

$$; \exists v('v) ! T v); '=) ; \exists v('v) ! T v)= ' ! : \quad (!!)$$

$$; '=) ; : = ': \quad (\text{contraposición})$$

Observación 5.9. Si th es una teoría de la verdad en la cual valen [T-Elim](#), [E!](#), [I^](#), [I!](#), [contraposición](#), [E8](#) y [I9](#), el predicado de verdad de th tiene la propiedad de eliminación.

Demostración. La prueba de [ET 2](#) es idéntica a la de la observación [5.7](#). La siguiente es una derivación de [ET 1](#):

1. $\exists v('v) ! T v)$	premisa
2. $'(p q) ! T p q$	1, E8
3. $'(p q)$	supuesto
4. $T p q$	2, 3, E!
5.	4, T-Elim
6. $'(p q) !$	2-5, I!

La siguiente derivación prueba [ET 3](#):

1. $'(p q)$	premisa
2. $:$	premisa
4. $:T p q$	2, T-Elim , contraposicion
5. $:(' (p q) ! T p q)$	1, 4, I^
6. $: \exists v('v) ! T v)$	5, I9

□

Luego, [T-Out](#) o [T-Elim](#) son condiciones suficientes para que el predicado de verdad de una teoría tenga la propiedad de eliminación provistas ciertas leyes lógicas. De hecho, en la sección [5.3](#) muestro que si estos requisitos no se cumplen, aun teorías que contienen [T-Out](#) o [T-Elim](#) e incluso principios de transparencia completos pueden no contener un predicado veritativo con la propiedad de eliminación.

Por otro lado, es sencillo ver que bajo condiciones mínimas [T-Elim](#) [y, si la regla de introducción del condicional vale, [T-Out](#) también] es una condición necesaria prácticamente en todo contexto.

Observación 5.10. Sea th_{LT} una teoría en la cual, en todos sus modelos, la relación de identidad se comporte clásicamente,⁵ las expresiones condicionales sean verdaderas si tanto el antecedente como el consecuente lo son o el primero obtiene un valor no designado, y los enunciados universales sean verdaderos cuando todas sus instancias lo sean. Si el predicado de verdad de th tiene la propiedad de eliminación, [T-Elim](#) vale en th . Y si [I!](#) resulta válida, [T-Out](#) vale también.

Las condiciones impuestas sobre th son satisfechas en lógica clásica pero también en muchas lógicas no clásicas, como $k3$, lp , pl y st , aunque no por $b3$, que viola el segundo requisito impuesto sobre los condicionales (vease el capítulo [3](#)).

⁵Esto es, $t = t$ obtiene el valor clásico designado (i.e., 1) en todos los modelos de th , y $'(t)$ puede inferirse a partir de $s = t$ y $'(s)$, para cualquier fórmula $'(x)$ y cualesquiera términos s ; t .

Demostración. Sea M un modelo de \mathcal{L}_T y sea $T \vdash p \wedge q$ verdadera en M . Dadas las condiciones impuestas sobre la identidad, condicionales y enunciados universales, $p \wedge q \equiv \forall x (x = p \wedge q \rightarrow Tx)$ son verdaderas en M . Luego, por [ET 2](#), $\forall x (x = p \wedge q \rightarrow Tx)$ es verdadera en M .⁶ \square

En consecuencia, bajo ciertas condiciones lógicas, [T-Elim](#), y, en algunos casos también [T-Out](#), es una condición necesaria y suficiente para garantizar que el predicado de verdad tenga la propiedad de eliminación.

5.2.2.2. La propiedad de introducción de la verdad

Resta ahora estudiar la posibilidad de definir, además de una propiedad de eliminación, una propiedad razonable de introducción del predicado veritativo que se corresponda con la regla de introducción de la conjunción intuitiva [L](#). En ese caso podrá decirse que una teoría de la verdad permite la expresión de conjunciones intuitivas siempre y cuando su predicado de verdad posea tanto la propiedad de eliminación como la de introducción.

Dada la regla de introducción de la conjunción intuitiva, se puede inferir

$$\frac{}{T \vdash (p \wedge q)}$$

a partir de

$$T \vdash p \quad T \vdash q$$

De ella podrá extraerse una cuarta condición para la expresabilidad de conjunciones intuitivas: T es capaz de expresar conjunciones posiblemente intuitivas en \mathcal{L}_T mediante expresiones de la forma $\forall v (v = (p \wedge q) \rightarrow Tv)$ solo si, para cualquier predicado $'(v) \in \mathcal{L}_T$, vale la siguiente regla:

$$\frac{T \vdash (p \wedge q)}{\forall v (\text{Sent}_{\mathcal{L}_T}(v) \wedge '(v) \rightarrow Tv)} \quad (IT 1)$$

Notese que es preciso incorporar el predicado $\text{Sent}_{\mathcal{L}_T}$ en la conclusión porque las premisas solo valen para oraciones de \mathcal{L}_T . Si acaso algo que no (codifica) es una oración satisface $'$, no debería concluirse a partir de $T \vdash (p \wedge q)$ que es verdadero. Esta inferencia podrá conformar la propiedad de introducción del predicado veritativo.

Definición 5.11 (Propiedad de introducción). El predicado de verdad de una teoría \mathcal{L}_T satisface la propiedad de introducción si la regla [IT 1](#) vale en \mathcal{L}_T .

⁶Notese que, como solo considero teorías formuladas en lenguajes de primer orden, la noción de consecuencia lógica subyacente es siempre efectiva y, por ende, los cálculos correspondientes pueden ser y son completos.

En consecuencia, podr a decirse, respondiendo a la pregunta (a) al comienzo de la seccion, que en una teoria de la verdad th expresiones de la forma $\forall v ('(v) \rightarrow T v)$ expresan conjunciones infinitas, que $\forall v ('(v) \rightarrow T v)$ es equivalente a $\bigwedge_{2L_T} ('(p \rightarrow q) \rightarrow T p \rightarrow T q)$, si y solo si '(n) implica en th que $\text{Sent}_{L_T}(n)$ para todo $n \in 2L_T$ y T tiene en th tanto la propiedad de eliminacion como la de introduccion. Las siguientes observaciones son practicamente triviales.

Observacion 5.12. Si th es una teoria de la verdad en cuyos modelos un condicional verdadero con antecedente verdadero tiene consecuente asimismo verdadero y valen la regla de induccion infinita

$$\frac{\forall (p \rightarrow q) : \bigwedge_{2L_T} g}{\forall v (\text{Sent}_{L_T}(v) \rightarrow '(v))} \quad (II)$$

y o bien [T-In](#), esto es, la direccion de derecha a izquierda del [Esquema-T](#)

$$\frac{\forall (p \rightarrow q)}{\forall v (\text{Sent}_{L_T}(v) \rightarrow '(v))} \quad (T\text{-In})$$

o bien [T-Intro](#), el predicado veritativo de th tiene la propiedad de introduccion.

Demostracion. Sea M un modelo de th donde todos los miembros de $\forall (p \rightarrow q) : \bigwedge_{2L_T} g$:

$\bigwedge_{2L_T} g$ son verdaderos. Si '(p q) es tambien verdadera, por la condicion impuesta a los condicionales, lo es y, por [T-In](#) y nuevamente la condicion sobre los enunciados condicionales, o por [T-Intro](#), T p q es verdadera en M. Una vez mas, por las condiciones sobre los condicionales, esto implica que '(p q) \rightarrow T p q. Luego, el conjunto $\forall (p \rightarrow q) : \bigwedge_{2L_T} g$ vale en M el cual, por la regla II, implica $\forall v (\text{Sent}_{L_T}(v) \rightarrow '(v) \rightarrow T v)$. □

En consecuencia, provisto que se cumplan algunas propiedades logicas, [T-In](#) o [T-Intro](#) son principios suficientes para garantizar que el predicado de verdad tenga la propiedad de introduccion. La siguiente observacion muestra que bajo ciertas condiciones, [T-In](#) o [T-Intro](#) tambien son necesarios para esta tarea.

Observacion 5.13. Sea th una teoria de la verdad tal que, en todos sus modelos, la relacion de identidad se comporta classicamente, los condicionales con antecedente verdadero tienen consecuente verdadero, y los enunciados universales son verdaderos solo si todas sus instancias lo son. Si el predicado de verdad de th tiene la propiedad de introduccion y $\text{Sent}_{L_T}(v)$ representa el conjunto de oraciones de L_T en la teoria, [T-Intro](#) vale en th. Y si la regla II resulta valida, [T-In](#) vale tambien.

Estas condiciones son satisfechas en logica clasica, pero tambien en k_3 , b_3 , lp , pl y st , entre otras.

Demostracion. Sea M un modelo de th en el cual ' es verdadera. Dadas las condiciones impuestas sobre la identidad y los condicionales, todos los miembros de $\forall p \supset q = p \supset q$!

: $\exists \text{L}_{\supset}$ g son verdaderos tambien. Luego, por la regla [IT 1](#), $\exists v(\text{Sent}_{\supset} (v) \wedge p \supset q = v \supset \text{Tv})$ es verdadera en M con lo cual, por las condiciones sobre enunciados universales, $\text{Sent}_{\supset} (p \supset q) \wedge p \supset q = p \supset q$! $\text{T} p \supset q$ tambien lo es. Como $\text{Sent}_{\supset} (p \supset q) \wedge p \supset q = p \supset q$ es verdadera, $\text{T} p \supset q$ es asimismo verdadera en M. \square

Luego, ahora se esta en condiciones de responder tambien a la pregunta (b) al comienzo de la seccion: bajo ciertas condiciones logicas (las hipotesis de las observaciones [5.9](#), [5.10](#), [5.12](#) y [5.13](#)) las reglas [T-Elim](#) y [T-Intro](#) son tanto necesarias como suficientes para garantizar la equivalencia entre generalizaciones de la forma $\exists v'(v) \supset \text{Tv}$ y las conjunciones in nitarias correspondientes $\forall \text{L}_{\supset} '(p \supset q)$! porque, de acuerdo con la respuesta dada a la pregunta (a), para ello el predicado de verdad debe poseer las propiedades de eliminacion e introduccion que, por las observaciones [5.9](#), [5.10](#), [5.12](#) y [5.13](#) requieren e implican estas reglas bajo aquellas condiciones. Asimismo, bajo condiciones levemente diferentes, los principios [T-Out](#) y [T-In](#) son tambien necesarios y suficientes, como muestran las observaciones [5.8](#), [5.10](#), [5.12](#) y [5.13](#).

Naturalmente, no tiene sentido esperar que una regla como [IT 1](#) ni una inferencia como [IT 1](#) valgan en una teoria de la verdad de primer orden, nitaria, a menos que se trabaje con una teoria semantica dada por una familia de modelos en los cuales esta regla preserve verdad. Por tanto, tampoco es razonable esperar de teorias de la verdad no semanticas que provean predicados de verdad con la propiedad de introduccion.

Existe una inferencia similar a [IT 1](#) pero mas debil, implicada por ella, que se puede enunciar sin recurrir a reglas in nitarias, i.e., una version restringida donde el conjunto de premisas es nito:

$$\frac{\forall i, 1 \leq i \leq n; \exists v'(v) \supset \text{Tv} \quad W}{\forall i, 1 \leq i \leq n; \exists v'(v) \supset \text{Tv}} \quad \text{(IT 2)}$$

Supongase que el Papa dijo exactamente $\forall i, 1 \leq i \leq n$. Luego, a partir de $\forall i, 1 \leq i \leq n$

podra esperarse que en una teoria de la verdad se pueda derivar que todo lo que el Papa dijo es verdadero. Por supuesto, esto funciona exclusivamente en la medida en que ' se aplique solo a un numero nito de oraciones. Esta regla podra constituir una propiedad de introduccion debil del predicado veritativo.

Definicion 5.14 (Propiedad de introduccion debil). El predicado de verdad de una teoria de th satisface la propiedad de introduccion debil sii la regla [IT 2](#) vale en th.

Tal vez sea más razonable, como argumento en la sección que sigue, no exigir la propiedad de introducción a un predicado veritativo o, al menos, exigir esta versión debilitada en su lugar. Las observaciones 5.12 y 5.13 pueden fácilmente adaptarse a esta nueva propiedad.

Observación 5.15. Si \mathcal{L} es una teoría de la verdad en cuyos modelos la relación de identidad se comporta clásicamente, un condicional verdadero con antecedente verdadero tiene consecuente asimismo verdadero y valen, o bien [T-In](#) o bien [T-Intro](#), el predicado veritativo de \mathcal{L} tiene la propiedad de introducción débil.

Observación 5.16. Sea \mathcal{L} una teoría de la verdad tal que, en todos sus modelos, la relación de identidad se comporta clásicamente, los condicionales con antecedente verdadero tienen consecuente verdadero, y los enunciados universales son verdaderos solo si todas sus instancias lo son. Si el predicado de verdad de \mathcal{L} tiene la propiedad de introducción débil, [T-Intro](#) vale en \mathcal{L} . Y si la regla [!!](#) resulta válida, [T-In](#) vale también.

Luego, aun debilitando la propiedad de introducción, principios como [T-Intro](#) y [T-In](#) siguen siendo necesarios y suficientes para garantizar esta propiedad al predicado de verdad, bajo ciertas condiciones lógicas.

5.2.3. El costado útil de las generalizaciones

En el apartado anterior he propuesto un sentido en el cual adscripciones generales de verdad de la forma $\forall v(\forall v \text{ ! } T v)$ pueden resultar equivalentes a las conjunciones infinitarias $\bigwedge_{v \in V} \text{! } (p \rightarrow q)$, i.e., si se comportan del mismo modo desde un punto de vista \forall predicado de verdad tiene las propiedades de eliminación e introducción.

Asimismo, he mostrado que si se cumple una serie de requisitos lógicos, los principios [T-Intro](#) y [T-Elim](#) resultan tanto necesarios como suficientes para garantizar estas dos propiedades al predicado veritativo, y otro tanto vale para [T-Out](#) y [T-In](#). En particular, ambos pares de principios son necesarios y suficientes para dar al predicado de verdad las propiedades de eliminación e introducción en el marco de la lógica clásica.

Dado que [T-Out](#) y [T-In](#) tomados en conjunto son inconsistentes en la lógica clásica más cierta sintaxis debido al teorema de la indecibilidad de la verdad de Tarski (teorema 2.9), no existe una teoría clásica de la verdad que satisfaga la equivalencia completa entre generalizaciones de la forma $\forall v(\forall v \text{ ! } T v)$ y las conjunciones infinitarias

$\bigwedge_{v \in V} \text{! } (p \rightarrow q)$. ¿Implica esto que el terreno de la lógica clásica debe ser abandonado si se desea un predicado de verdad capaz de cumplir sus roles lógico-expresivos? En este apartado argumento que la respuesta a esta pregunta es definitivamente "No", porque

la respuesta a la pregunta (c) a comienzos de la sección es también negativa. La idea central es que las características que tornan útiles a las generalizaciones de la forma $\forall v ('(v) \rightarrow T v)$ en el discurso científico y filosófico quedan absolutamente garantizadas por la propiedad de eliminación de la verdad. En otras palabras, [T-Out](#) o [T-Elim](#) son suficientes para asegurar los usos del predicado veritativo que lo vuelven indispensable. Los principios que garantizan la equivalencia entre las generalizaciones y las conjunciones infinitarias, la equivalencia misma entre estos enunciados, no son necesarios para que el predicado de verdad pueda cumplir sus propósitos lógicos, como muestro a continuación.

En efecto, una teoría cuyo predicado de verdad posee la propiedad de eliminación permite ya la axiomatizabilidad finita de cualesquiera conjuntos de oraciones de niveles en el lenguaje de la teoría. La proposición [5.2](#) de Halbach muestra que, relativamente a los bicondicionales tipados que constituyen tb , las conjunciones infinitarias y sus generalizaciones correspondientes son equivalentes con respecto a sus implicancias en las cuales T no ocurre. Pero note que este resultado no depende de las instancias de [T-Intro](#), [T-In](#) o de la propiedad de introducción de la verdad en modo alguno. Sea L un lenguaje que no contiene T , th una teoría axiomática formulada en L y th^0 una teoría de la verdad expresada en $L \cup \{T\}$ que contiene, como únicos axiomas, todas las instancias de [T-Out](#) dadas por oraciones de L . Supongase adicionalmente que ambas teorías son clásicas.

Proposición 5.17. Si $\phi(v) \in L$, entonces, para toda oración $\psi \in L$,

$$th \vdash \phi(v) \rightarrow \psi \quad : \quad \vdash Lg \quad , \quad th^0 \vdash \forall v ('(v) \rightarrow T v) \rightarrow \psi$$

Demostración. Si $\psi \in L$ es un teorema de $th^0 \vdash \forall v ('(v) \rightarrow T v)$, entonces es también un teorema de $th \vdash \forall v ('(v) \rightarrow T v)$, porque $th^0 \vdash \psi$. Luego, por la proposición [5.2](#), $th \vdash \phi(v) \rightarrow \psi : \vdash Lg$. A la inversa, si es un teorema de $th \vdash \phi(v) \rightarrow \psi$

$: \vdash Lg$, solo finitas de oraciones de $\phi(v) \rightarrow \psi : \vdash Lg$ pueden haber sido utilizadas en la prueba. Todas ellas se siguen de $\forall v ('(v) \rightarrow T v)$ junto con las instancias relevantes de [T-Out](#). \square

Este resultado no se limita únicamente a 'conjunciones infinitarias' de oraciones del lenguaje de la teoría base, que no contienen T . La propiedad de eliminación permite ya que oraciones de la forma $\forall v ('(v) \rightarrow T v)$ encapsulen todas las oraciones que caen bajo el predicado $'$, sean finitas o infinitarias, contengan o no contengan T . ¿Cuál sería entonces la ventaja de contar con una teoría de la verdad cuyo predicado, adicionalmente, tenga la propiedad de introducción? Suele argumentarse que la propiedad de introducción garantiza que generalizaciones como $\forall v ('(v) \rightarrow T v)$ impliquen nada más allá de las $'$ s sino todas y precisamente aquellas oraciones que satisfacen este predicado. En

otras palabras, permite justificar o explicar la generalización en base a las 's, porque, de acuerdo con [IT 1](#), estas implican $\forall v('v) ! T v$.

Por ejemplo, Horwich, en su abordaje del problema de las paradojas, señala lo siguiente:

[. . .] in so far as 'p' is not invariably equivalent to 'p is true', then a generalisation of the form 'Every instance of schema S is true' will not invariably entail every instance of S; nor will it always be justified or explained on the basis of those sentences. ([Horwich, 1998b](#), p. 42, nota 21)

Horwich menciona dos razones según las cuales el [Esquema T](#) completo es necesario: por un lado, garantiza que la generalización "Todos los P s son verdaderos" implique todos los P s; por otro lado, permite justificar o explicar la generalización en base a los P s. La primera razón corresponde a la propiedad de eliminación, para la cual se ha visto que, en contextos clásicos formales, [T-Out](#) es necesario y suficiente [i.e., la primera razón de Horwich justifica la adopción de únicamente una parte del [Esquema T](#), su dirección de izquierda a derecha]. La segunda razón se corresponde con la propiedad de introducción, y formalmente implica la necesidad de [T-In](#), la parte faltante del [Esquema T](#): de acuerdo con Horwich, al menos en un plano informal, [T-In](#) puede justificar o explicar generalizaciones de la forma $\forall v('v) ! T v$ sobre la base de las infinitas premisas del conjunto $\{ (p q) ! : 2 L_T g$, como se vio en el apartado [4.3.4](#).

No obstante, dos cosas. Primero, queda claro en virtud de la proposición [5.17](#) que en el plano formal, cuando se restringe la atención a oraciones en las cuales el predicado de verdad no ocurre, la propiedad de eliminación es suficiente, no solo para que expresiones de la forma $\forall v('v) ! T v$ encapsulen todas las 's sino también para garantizar que sus consecuencias no vayan más allá de las 's, i.e., la propiedad de eliminatividad es suficiente para asegurar que el predicado veritativo sea un dispositivo de axiomatizabilidad nita.

En segundo lugar, si esta restricción a oraciones del lenguaje sin T se abandona, el único modo de garantizar que $\forall v('v) ! T v$ no implique nada más allá del conjunto de enunciados que satisfacen ' o, equivalentemente, del conjunto de todas las instancias de $(p q) !$, la única manera de que estos conjuntos expliquen o justifiquen la adscripción general de verdad en cuestión, es mostrado que ellos mismos implican la generalización $\forall v('v) ! T v$, relativamente por supuesto a la teoría de la verdad en la cual se trabaja. Pero esto solo es posible en la medida en que, o bien se cuente con una regla infinitaria como [II](#), esto es, se trabaje con una teoría semántica de la verdad, o bien el conjunto de oraciones que caen bajo ' sea nito, en cuyo caso se estará aplicando la regla [IT 2](#) (o

bien en la teoría ya un número finito de las 's implique la generalización en cuestión, lo cual sería fortuito y, en una teoría de la verdad consistente, un fenómeno particular de algunas generalizaciones y sus instancias).

En un sentido informal, la propiedad de introducción de la verdad provee efectivamente una justificación de las generalizaciones a partir de sus instancias in nitas como quiere Horwich, pero formalmente no hace ningún trabajo en estos casos, a menos que se trabaje con una familia acotada de modelos, en los cuales valga una regla de inducción in nita. Sin embargo, queda claro que ninguna axiomatización de esta teoría semántica sería capaz de capturar esta justificación de las adscripciones generales de verdad a partir de sus instancias, lo cual pone seriamente en duda su viabilidad y su utilidad.

En consecuencia, en el plano formal la utilidad de la propiedad de introducción se reduce a la justificación o explicación de conjunciones in nitas.⁷ Sin embargo, notese que las generalizaciones empleadas en reemplazar conjunciones in nitas son en principio dispensables [es posible utilizar la conjunción in nita misma en su lugar]. Desde un punto de vista formal, el uso más importante de una adscripción general de verdad de la forma $\forall v ('v) ! T v$) es la axiomatizabilidad in nita o expresión de un conjunto in nito de oraciones, aunque en la práctica con frecuencia se utilizan también generalizaciones in nitas. En el marco de teorías científicas o filosóficas formales, esto es, al dar una quasi lógica de la verdad, estos usos no son razón suficiente para debilitar la lógica clásica.

Aun más, tiene poco sentido incorporar principios (como [T-Intro](#) o [T-In](#)) a una teoría (de la verdad) para asegurarse de que ciertas oraciones (en este caso las generalizaciones de la forma $\forall v ('v) ! T v$) no impliquen lo que no deben (algo más que las 's). Si ya lo hacen, agregar nuevos principios no va a bloquear la inferencia; y si no implicaban enunciados indeseados, entonces no hay razón para agregar los nuevos principios. El único objetivo razonable sería que la incorporación de los nuevos principios agregue poder inferencial a aquello que debe ser implicado. En este caso, [T-Intro](#) o [T-In](#) no estarían quitando poder deductivo a $\forall v ('v) ! T v$) sino agregandose a los miembros del conjunto $\{ (p \rightarrow q) ! : \exists L_T g$, para que estos puedan implicar la generalización. Pero no lo logran, porque en cualquier axiomatización consistente de la teoría fallaría la regla in nitaria [II](#), que solo puede valer en abordajes semánticos.

>Para que tomarse el trabajo entonces de incluir principios como [T-In](#) o [T-Intro](#) para obtener la propiedad de introducción de la verdad? La propiedad de eliminación es suficiente para garantizar los usos lógicos del predicado veritativo, esto es, posibilitar

⁷Si bien aun cuando ' es un predicado que se aplica exclusivamente a un número finito de oraciones, $\forall v ('v) ! T v$) puede entenderse como la conjunción in nita de todas las oraciones de la forma $(p \rightarrow q) !$, a diferencia de los casos en los cuales ' se aplica a in nitas expresiones, esta generalización es siempre equivalente a la conjunción in nita de las 's, como ya mencioné.

que generalizaciones de la forma $\exists v('v) ! T v)$ encapsulen las posiblemente infinitas oraciones que satisfacen ' y 'nada mas', porque, como se~nale, de garantizar algo, la propiedad de introduccion solo servir a para los casos menos interesantes, i.e., los casos finitos.

Asimismo, al implicar [T-Elim](#), la propiedad de eliminacion garantiza, en primer lugar, que oraciones de la forma $T p'q$, donde ' es una oracion, impliquen '. Segundo, enunciados de la forma $T t$, donde t es un termino cerrado de L_T que no es un numeral y denota una oracion ', i.e. t es un nombre no transparente de ', tambien implican ', as como, tercero, aquellas adscripciones $\exists v('v) ! T v)$ donde ' es satisfecha por un unico, varios o infinitos enunciados implican estos enunciados. Luego, la propiedad de eliminacion cubre todos los usos del predicado veritativo mencionados en el apartado [4.2.3](#) (con excepcion de la expresion de oraciones en otros lenguajes, porque solo se esta trabajando con codificaciones del mismo L_T), no solo los usos logicos.

Consecuentemente, es suficiente debilitar el [Esquema-T](#), tomar unicamente su direccion de izquierda a derecha, esto es, [T-Out](#), para reconstruir el rol logico-expresivo del predicado veritativo. Siendo [T-Out](#) no solo consistente sino conservativo sobre pat , es perfectamente posible permanecer dentro de los limites de la logica clasica y capturar, a la vez, todo lo expresivamente util del predicado de verdad. Esto es contrario, no solo a lo que cree buena parte del de acionismo, que ha abandonado al logica clasica con el objetivo preciso de garantizar los roles logico-expresivos de la verdad, sino tambien a lo que considera [Horwich \(1998b, p. 42, nota 21\)](#), de acuerdo con quien 'the need to restrict instantiation of the [[Esquema-T](#)] is somewhat in tension with the minimalist thesis about the function of our concept of truth|namely that it enables us to capture schematic generalisations.'

Sin embargo, quizas haya otros casos en los cuales la propiedad de introduccion o la equivalencia completa entre las adscripciones generales de verdad de la forma $\exists v('v) ! T v)$ y las correspondientes conjunciones infinitarias resulten utiles. Considere el siguiente escenario planteado por [Field \(2008, p. 210\)](#). Supongase que no recuerdo exactamente lo que Lorna dijo, pero quiero afirmar que eso que dijo implica cierta proposicion . Luego, podra decir

$$\exists x('x) ! T x) ! \tag{5.3}$$

donde ' x ' se aplica unicamente a lo que dijo Lorna. Relativamente a la suposicion de que las aserciones de Lorna fueron $p_1; \dots; p_n$, seria deseable entonces que su conjuncion implicara ϕ , o, si se quiere, que [\(5.3\)](#) implicara

$$p_1 \wedge \dots \wedge p_n \tag{5.4}$$

Field usa este ejemplo como un argumento en contra de las teorías de la verdad clásicas.

Señala que, para poder derivar (5.4) a partir de (5.3), \neg y \top p q_i deben ser intersustituibles, lo cual no puede ocurrir en ninguna teoría consistente clásica (para toda \mathcal{L}).

Una vez más, podrá descartar este argumento señalando que en este tipo de casos no se está haciendo un uso lógico del predicado veritativo sino meramente epistémico y que, por tanto, en contextos formales es enteramente dispensable: el predicado veritativo se aplica únicamente a un número finito de oraciones y, por ende, puede reemplazarse por la mera conjunción de aquellas.

No obstante, pueden proponerse otros casos análogos donde \mathcal{L} parece necesaria y donde se está haciendo uso lógico del predicado veritativo. El siguiente me ha sido sugerido por Halbach. En epistemología suelen recurrirse al predicado de verdad para formular definiciones de conocimiento de modo no esquemático, como se señala en el apartado 4.2.3.3. Se dice de un agente que conoce un enunciado siempre y cuando lo crea, está justificado en creerlo y, adicionalmente, el enunciado sea verdadero (y algunas condiciones de Gettier sean satisfechas). Formalmente, se dice

$$\forall x(K(a; x) \leftrightarrow C(a; x) \wedge Tx) \tag{5.5}$$

en lugar de las infinitas instancias del siguiente esquema:

$$K(a; p \rightarrow q) \leftrightarrow C(a; p \rightarrow q) \wedge \dots \tag{5.6}$$

donde $C(x; a)$ resume todas las condiciones para el conocimiento excepto la verdad. Supongase ahora que hay un agente a y una oración ϕ tales que $C(a; p \rightarrow q) \wedge \dots$. Sería claramente deseable poder concluir, a partir de (5.5), que a conoce ϕ , pero esto no es posible sin la instancia correspondiente de \mathcal{L} .

Sin embargo, esto no representa mayor problema. Una teoría de la verdad cuyo predicado veritativo cuenta con la propiedad de eliminación es capaz todavía de dar una formulación finita, no esquemática, de conocimiento, que abarque todas las instancias de (5.6), y derivar que un agente conoce una oración a partir de $C(a; p \rightarrow q) \wedge \dots$ si, en lugar de derivar conocimiento mediante (5.5), lo hace a través de

$$\forall x(\phi(a; x) \rightarrow Tx)$$

donde $\phi(a; x)$ es verdadera exactamente de todas las instancias de (5.6).

Concedo que esta formulación alternativa de conocimiento tiene sus inconvenientes: el predicado K ya no es eliminable y la definición no satisface el requisito de conservatividad, principalmente porque ya no es una definición explícita. Sin embargo, estos

parecen solo males menores si se los compara con el abandono de la lógica clásica (en casos en los cuales la única razón que se tiene para abandonarla es evitar las paradojas semánticas) y, aun más, la introducción de un predicado de verdad no clásico en la definición de conocimiento. La no clasicidad de la verdad es contagiosa; inevitablemente se transmite a la noción de conocimiento, lo cual no es evidentemente deseable.

Este no es un caso aislado, sino que pueden plantearse muchos otros ejemplos por el estilo. En general, debe tenerse en cuenta lo siguiente. Si se quiere expresar una 'conjunción in nita' de oraciones en una teoría de la verdad cuyo predicado posea la propiedad de eliminación, basta con recurrir a oraciones de la forma $\exists v(\forall v)!$

$\forall v$), donde v se aplica precisamente a las infinitas oraciones cuyo contenido se quiere expresar. El predicado de verdad de una teoría no transparente pero eliminativa de la verdad no es capaz, dada cualquier oración ϕ en la cual v ocurre como suboración, de expresar la conjunción in nita de todas las instancias de ϕ dadas por cada una de las oraciones del lenguaje generalizando sobre v , esto es, reemplazándola por $\forall v$ y cuantificando sobre v para obtener $\exists v(\forall v)\phi$, sino que debe primeramente encontrar un predicado que se aplique a todas las instancias de ϕ para luego expresar $\exists v(\forall v)!\phi$.

Esto permite también dar cuenta del caso anterior, donde quiero derivar a partir de lo que Lorna dijo. En lugar de (5.3), se puede utilizar

$$\exists x(\forall x)!\phi(x) \tag{5.7}$$

donde ϕ se aplica a la oración que se obtiene al concatenar la conjunción de ϕ_1, \dots, ϕ_n ,

i.e., lo que dijo Lorna, con la expresión $\forall x$.⁸ Luego, en cualquier teoría cuyo predicado de verdad tenga la propiedad de eliminación, (5.4) es derivable a partir de (5.7) y el supuesto de que se aplica exactamente a ϕ_1, \dots, ϕ_n .

La sugerencia de adoptar una formalización que impone modificaciones drásticas a la forma lógica original de expresiones como "Si lo que dijo Lorna es verdadero entonces A" o la definición coloquial de conocimiento puede resultar levemente ad hoc. Sin embargo, debe entenderse simplemente como una nueva regimentación. Del mismo modo en que la conjunción debe necesariamente cambiar de lugar al formalizar expresiones como "Lorna y Nixon votaron a Cristina", el predicado de verdad debe hacerlo en otras. En última instancia, parecería que los beneficios de esta regimentación superan ampliamente sus costos: entrometerse con la lógica clásica.

§

⁸Esto es, $\phi(x) := (\phi = (\phi_1 \wedge \dots \wedge \phi_n) \wedge \phi)$.

He comenzado esta sección con tres preguntas vitales para el de acionismo y cualquiera que desee una 'lógica' de la verdad que garantice los roles lógico-expresivos del predicado veritativo y, principalmente, sus usos lógicos. La primera pregunta fue, por ende, (a) en qué sentido puede decirse que generalizaciones de la forma $\forall v (v \rightarrow T v)$ son equivalentes a la conjunción in nitaria $\text{2LT} ((p \rightarrow q) \rightarrow T (p \rightarrow q))$, para lo cual he explorado dos caminos. Mientras que la primera propuesta no resultó adecuada, la segunda sí: la equivalencia entre generalizaciones y conjunciones in nitas está dada en función de su comportamiento inferencial. El de las conjunciones in nitas consiste de dos reglas: de eliminación y de introducción, que encuentran sus analogas en el plano de las generalizaciones en las propiedades de eliminación e introducción de la verdad, respectivamente. Estas, en lógica clásica y en otras lógicas también, se corresponden de alguna manera con los principios [T-Out](#) y [T-In](#) o, alternativamente, [T-Intro](#) y [T-Elim](#), lo cual responde la pregunta (b), esto es, qué principios para regular el comportamiento del predicado veritativo son necesarios y suficientes para asegurar la equivalencia de generalizaciones y conjunciones in nitarias.

Quedo claro [espero] que la propiedad de introducción y, con ella, [T-In](#) y [T-Intro](#), para cumplir la función deseada de justificación o explicación, requiere del auxilio de reglas de inducción in nita, que no están disponibles en las lógicas más comúnmente utilizadas, esto es, las lógicas nitarias. No obstante, muestro que esto no es realmente un problema, que los usos lógicos del predicado de verdad quedan ya garantizados por la propiedad de eliminación. En otras palabras, solo se necesita una parte de los principios de transparencia para garantizar estos usos, [T-Out](#) o [T-Elim](#), y, por ende, todo esto es posible en el marco de la lógica clásica, lo cual responde a la pregunta (c).

5.3. Como debe lucir una teoría de acionista de la verdad

Basándome en los resultados obtenidos e ideas defendidas a lo largo del capítulo, en esta sección enumero las una serie de requisitos que, sostengo, toda teoría de acionista de la verdad debe satisfacer y exploro diversos sistemas ya existentes a la luz de estos requisitos.

En virtud de la sección anterior estoy en condiciones de afirmar que una teoría de la verdad de acionista decente debe proveer un predicado de verdad que tenga la propiedad de eliminación, esto es, debe ser una teoría, no necesariamente transparente, pero sí eliminativista de la verdad. Pero, si bien bajo ciertas condiciones lógicas mínimas [T-Out](#) o [T-Elim](#) son suficientes para garantizar la propiedad de eliminación, esto no implica que la teoría cuyos únicos principios para el gobierno del predicado de verdad sean estos sea una teoría deseable o adecuada de la verdad.

Por ejemplo, podría resultar conveniente contar con principios generales correctos, versiones formales de algunos de los principios examinados en el apartado [4.3.4](#), o ciertas instancias inofensivas de [T-In](#) o [T-Intro](#), por diversas razones, como fines descriptivos o la posibilidad de no solo expresar sino también probar ciertas generalizaciones. En la medida en que estos principios se desprendan de alguna manera de los principios o subprincipios (partes) de transparencia que se adopten la posición de acionista no se verá afectada.

Como sucede con otros operadores lógicos, como el cuantificador universal o la negación, se espera de una lógica que los contenga que pruebe todos los principios y valide todas las reglas correctas con respecto a estos operadores, i.e., que la lógica sea completa, aunque este no es un requisito decisivo. Análogamente, podría quererse que la teoría de la verdad que se adopte sea completa con respecto al predicado veritativo, que sea capaz de derivar todos los principios correctos que atañen al significado de la verdad. Naturalmente, esto dependerá de que principios se consideren correctos y cuáles no. Pero, por otro lado, dado que la verdad no es un predicado puramente lógico sino lógico-lingüístico o, lo que es lo mismo, lógico-matemático (véase el apartado [4.3.3](#)),

[i]t may be too much to ask of any theory of truth to completely capture the meaning of the concept of truth. After all, axiomatic theories of truth would appear to be on a par with axiomatic theories of other concepts. ([Horsten, 2011](#), p. 62)

Por otro lado, también podría resultar deseable que la teoría de la verdad pruebe al menos algunos principios generales verdaderos que relacionen el predicado de verdad con otras conectivas, como, por ejemplo, si se trabaja en lógica clásica, la fórmula

$$\forall x(\exists y(\text{Sent}_{LT}(y) \wedge x = (y!y)) \rightarrow T x)$$

que establece que todas las oraciones de la forma '!' son verdaderas. Para ello es preciso contar con principios generales sobre la verdad en la teoría.

Finalmente, podría ser conveniente contar con al menos ciertas instancias de [T-In](#) o [T-Intro](#) [v.g., las instancias que no contienen T, las fundadas en el sentido de Kripke o Leitgeb (véanse los apartados [3.1](#) y [3.2](#)), las T-positivas (definición [3.18](#)), etc.] para poder derivar una generalización de la forma $\forall v(\forall v \rightarrow T v)$ de las 's en los casos en los cuales estas sean nitas, incluso en una teoría clásica de la verdad.

Por todo lo dicho, una buena teoría de la verdad de acionista debe

- (1) validar los principios necesarios para garantizar los usos lógicos del predicado de verdad, su única razón de ser; i.e., debe ser otorgar al predicado veritativo la propiedad de eliminación;
- (2) validar principios para el predicado veritativo que no excedan los principios de transparencia, porque de lo contrario la tesis según la cual estos principios capturan el significado de la verdad ser a falsa; y, por último,
- (3) poder ser efectivamente utilizada como una lógica, i.e., contar al menos con una axiomatización parcial que garantice los usos lógicos de la verdad.

Notese que estas condiciones excluyen la teoría trivial en tanto esta exceder a la función básica de la verdad, aunque estas nociones son informales. Con respecto a que constituye una axiomatización parcial, tampoco está demasiado claro, como evidencian [Fischer et al. \(2015\)](#), pero basta con que haya, y los hay, casos claros positivos, como kf con respecto a la clausura de los modelos de punto fijo de Kripke con Strong Kleene (véase el apartado [3.1](#)), y casos claros negativos, como $putb$ con respecto a esta misma familia de modelos, por ser demasiado incompleta.

A continuación evalúo, a la luz de estas condiciones, las teorías formales de la verdad introducidas en los capítulos [2](#) y [3](#). Primero presento y examino brevemente el proyecto minimalista de Horwich y algunas teorías que podrían en principio considerarse realizaciones del proyecto, como las teorías axiomáticas de Tarski, $putb$ y la teoría de la verdad d -fundada de Leitgeb. En [5.3.2](#) examino las teorías de la verdad de Kripke y su axiomatización más conocida, kf . [5.3.3](#) se ocupa de las teorías composicionales axiomáticas $ct(pa^+)$, $rt<_o$ y fs mientras que [5.3.4](#) estudia las teorías paraconsistentes $lptt$, $pltt$ y $sttt$. Por supuesto, es también interesante evaluar la adecuación de otras teorías formales de la verdad según las condiciones (1)-(3), pero me llevaría demasiado lejos.

5.3.1. El proyecto minimalista de Horwich

De acuerdo con Horwich, el predicado veritativo requiere del [Esquema-T](#) completo, no tipeado, para garantizar de modo absoluto sus usos lógicos. Siendo esto imposible, opta por la mutilación de este principio, en la menor medida posible. El siguiente es el locus classicus:

Given our purposes, it suffices for us to concede that certain instances of the equivalence [T]-schema are not to be included as axioms of the minimal theory, and to note that the principles governing our selection of excluded

instances are, in order of priority: (a) that the minimal theory not engender 'liar-type' contradictions; (b) that the set of excluded instances be as small as possible; and perhaps just as important as (b) (c) that there be a constructive specification of the excluded instances that is as simple as possible. (Horwich, 1998b, p. 42)

La búsqueda de teorías formales de la verdad con estas condiciones es lo que de-nomino 'proyecto minimalista'. Horwich busca entonces una teoría a desentrecomilladora (dada por instancias del [Esquema-T](#)) clásica de la verdad que sea (a) consistente, (b) abarcativa (i.e., lo menos tipeada posible) y (c) cuyo criterio de selección de instancias sea 'simple y constructivo'. Las razones por las cuales la teoría debe ser desentrecomilladora y no estar dada, por ejemplo, también por principios generales son, posiblemente, que Horwich advirtió solo posteriormente, en (2005), que estos principios no exceden el [Esquema T](#) porque pueden ser, si bien no derivados, explicados por este, como indique en el apartado [4.3.4](#). Desde un punto de vista de acionista, no es necesario que el sistema formal que se adopte sea desentrecomillador; la búsqueda de acionista no se enmarca necesariamente en el proyecto minimalista.

El punto más oscuro de esta cita es obviamente (c). >Que quiere decir Horwich con 'constructivo' y 'simple'? Por un lado, la simpleza o constructividad del criterio de selección puede entenderse como un requisito de axiomatizabilidad. Si el conjunto de instancias aceptables es semirecursivo, la teoría dada por estos axiomas es una teoría axiomatizable. Entendido así, (c) ser semejante a (3), aunque un poco más exigente. La axiomatizabilidad completa o completitud podrá resultar un requisito demasiado ambicioso, que no suele pedirse, v.g., de la aritmética, como señale al comienzo de esta sección. Podrá bastar, por ejemplo, con que el criterio de restricción de instancias del [Esquema-T](#) sea lo suficientemente simple como para ser definible en L_{pa}^+ , v.g., mediante un predicado (ν) (veanse los apartados [1.2.1](#) y [1.2.2](#)), para poder ofrecer una axiomatización parcial razonable del conjunto correspondiente de instancias en términos de

$$(\nu) ! (T p'q \$')$$

Por otro lado, podrá leerse las palabras 'constructivo' y 'simple' en la línea de [Michael Glanzberg](#) (2003, 2005). Desde su punto de vista, el criterio de restricción de instancias del [Esquema-T](#) que una teoría minimalista adopte no puede estar dado en términos de propiedades sustantivas, como la fundación (grounding) o la representacionalidad (expresar proposiciones).

A minimalist approach, however needed, is committed to there being nothing but overt surface properties that determine whether something is a

truth bearer or not. For there to be anything else would eo ipso provide an 'underlying nature' of truth. This would make truth 'substantial' in just the way the minimalist says it is not. ([Glanzberg, 2003](#), p. 31)

Quizás un modo de entender esta idea [caritativo desde el punto de vista de-acionista] es sosteniendo que la restricción de las oraciones portadoras de verdad mediante cierta propiedad sustantiva ser a una manera de exceder la función básica de la verdad y violar, en consecuencia, la condición (2) que he impuesto sobre las teorías de acionistas formales al comienzo de la sección, aunque no es claro que se entiende por 'propiedades sustantivas'. Dado que el de acionismo, al menos en el ámbito formal, ya permite ciertas restricciones a los objetos que pueden caer bajo el predicado veritativo, esto es, las oraciones, y estas restricciones son sintácticas o, equivalentemente, aritméticas, las restricciones sintáctico-aritméticas parecen ser seguras, y quizás aquellas que las excedan, como la fundación, puedan considerarse sustantivas, porque no son de niveles en la teoría de la sintaxis, i.e., pa^+ , sino solo en lenguajes de segundo orden.

Concediendo que Glanzberg está en lo cierto, que relativizar el [Esquema-T](#) u otros principios para la verdad a propiedades que exceden la sintaxis está vedado al de acionismo, la siguiente cita de Horwich parece socavar la posibilidad de que fuera esto lo que tenía en mente cuando escribí su condición (c).

The intuitive idea is that an instance of the equivalence [T]-schema will be acceptable, even if it governs a proposition concerning truth (e.g. "What John said is true"), as long as that proposition (or its negation) is grounded|i.e. is entailed either by the non-truth-theoretic facts, or by those facts together with whichever truth-theoretic facts are 'immediately' entailed by them (via the already legitimized instances of the equivalence schema), or . . . and so on. ([Horwich, 2005](#), p. 81)

Desafortunadamente, esta cita no solo vuelve dudosa la lectura de la condición (c) en términos de Glanzberg sino también la lectura previa, en términos de axiomatizabilidad o de nibilidad. Porque tan solo para las oraciones de L_{pa}^+ el conjunto de 'non-truth theoretic facts' o hechos no semánticos |i.e., el conjunto de oraciones de L_{pa}^+ verdaderas en N^+ | ya es lo suficientemente complejo como para ser expresable en un lenguaje de primer orden y, a fortiori, en L_{pa}^+ . Parece que Horwich simplemente no estaba al tanto de estas complicaciones, ni tampoco de la tensión entre la condición (c) bajo alguna interpretación plausible y la condición (b), como se verá a continuación.

En efecto, tomar la mayor cantidad posible de instancias del [Esquema-T](#) sin caer en inconsistencias lleva a considerar conjuntos maximales consistentes de estas instancias. Como muestra el teorema de McGee (teorema [3.25](#)), no solo hay una infinidad de

tales conjuntos, cada uno de los cuales toma decisiones sobre la verdad o falsedad de las oraciones indecidibles de L_T (vease el apartado 3.2), sino que todos ellos son altamente complejos: ninguno es definible en \mathcal{P}^+ , porque cada uno contiene todas las oraciones verdaderas en un modelo de $L_{\mathcal{P}^+}$, lo cual es imposible por el primer resultado de incompletitud de Gödel (teorema 1.12). Aun más, Cezary Cieslinski (2007) ha mostrado un resultado análogo al de McGee para conjuntos maximales conservativos de instancias del Esquema-T. El teorema de McGee no solo muestra que la condición (b) de Horwich sobre un criterio de selección de instancias del Esquema-T entra en una tensión irresoluble con la condición (c) y debe, por ende, ser debilitada, sino que también apunta hacia una cuarta condición o condición (d) para el proyecto minimalista, i.e., que el criterio que se adopte sea plausible, no arbitrario sino explicativo.

Como señala Halbach (2009, p. 787), "it has proved difficult to find a sensible restriction on instances of the T-schema that is more liberal than Tarski's and that avoids awkward consequences and that is not too artificial." Pero si bien no hay una teoría hasta el momento que satisfaga concluyentemente todos los requisitos del proyecto minimalista, con o sin la condición (d), diversas teorías desentrecomilladoras de la verdad disponibles podrían ser buenas candidatas al proyecto minimalista, algunas mejores que otras. Por ejemplo, $tb(th)$ o $utb(th)$ (vease el apartado 2.3.3) son consistentes y su criterio de restricción es simple, incluso recursivo (no contener T) y plausible, en tanto fue fundamentado lógicamente por Russell y Tarski. Pero, por otro lado, el criterio es demasiado restrictivo, y es posible trascenderlo en diversas direcciones.

Por su parte, el modelo de Leitgeb en donde las instancias del Esquema-T quedan restringidas a oraciones fundadas de acuerdo con su noción de dependencia, esto es, $hN^+; pf i$ (cf. teorema 3.35), es una dirección posible. Las instancias son obviamente consistentes y el criterio es sumamente abarcativo y plausible. Todavía más, los teoremas 3.36 y 3.37 muestran que es posible permitir aun más instancias del Esquema-T sin caer en inconsistencias (cf. definición 1.2). Sin embargo, no se cumple la condición (c): como mencionamos en el apartado 3.2, la noción de d-fundación de Leitgeb (definición 3.34) es demasiado compleja, plausiblemente sustantiva, ya que no puede ser definida por una fórmula de un lenguaje de primer orden.⁹

Finalmente, $putb$ es también consistente, incluye una buena cantidad de instancias del Esquema-T, lo cual queda evidenciado por su poder de prueba (cf. teorema 3.23) y el criterio de restricción, la T-positividad, es tan simple como recursivo (definición 3.18). No obstante, la T-positividad es un criterio sumamente arbitrario, poco explicativo y radical. Muchas oraciones que no son T-positivas parecen ser inocuas que

⁹Además, al no estar dada por axiomas sino por un modelo no queda claro en que sentido la teoría semántica de Leitgeb es una teoría desentrecomilladora.

instancias del [Esquema-T](#) desde todo punto de vista, como $\vdash T p0 = 0q$. Por otro lado, putb permite instancias que no parecen del todo deseables. Por ejemplo, volviendo al ciclo de mentirosos del capitulo 3,

$$q^+ \setminus I_1 = pT I_2q \wedge I_2 = p:T I_1q$$

como $T I_2$ es T -positiva, hay una instancia correspondiente del [Esquema-T](#) en putb para ella, pero no para $\vdash T I_1$. Luego, putb decide arbitrariamente entre ciertas oraciones que contienen T , del mismo modo que los conjuntos maximales consistentes de McGee. Otro ejemplo esta dado por $T I$, donde I denota la oracion del mentiroso:

$$q^+ \setminus I = p:T Iq$$

T ocurre solo positivamente en $T I$ y, por ende, hay una instancia del [Esquema-T](#) asociada a esta oracion en la teoria. Un tercer caso esta dado por el una version mas debil del mentiroso.¹⁰ Sea I^0 el termino que se obtiene al aplicar el lema diagonal fuerte (teorema 1.8) a $T \setminus v$. Luego,

$$q^+ \setminus I^0 = T \setminus p \setminus q \tag{5.8}$$

Nuevamente $T \setminus I^0$ es T -positiva, genera una instancia del [Esquema-T](#) en putb .¹¹

Desde el punto de vista de las condiciones (1)-(3), ninguna de estas teorías resulta adecuada. El siguiente resultado es evidente.

Proposicion 5.18. Los predicados de verdad de las teorías $\text{tb}(p^+)$, $\text{utb}(p^+)$ y putb no tienen la propiedad de eliminacion.

Demostracion. Siendo teorías clásicas, por la observacion 5.10, si sus predicados de verdad tuvieran la propiedad de eliminacion, todas las instancias de [T-Out](#) dadas por oraciones de L_T podrian derivarse como teoremas. Pero este no es el caso porque, por ejemplo, ninguna de estas teorías prueba $T I \vdash T I$, la instancia de [T-Out](#) para la oracion $\vdash T I$ del mentiroso, ya que esta contiene T y no es T -positiva. □

De hecho, es claro que toda teoría axiomática de la verdad que restrinja el [Esquema-T](#) a un subconjunto propio de oraciones no validar todas las instancias de [T-Out](#) y, por tanto, no sera una teoría eliminativa. De acuerdo con Horwich, si la teoría es lo su cientemente abarcativa, esto es, si la condicion (b) se cumple en buena medida, \such problematic cases are few and far between; so the utility of truth as a device of

¹⁰En realidad, $T \setminus I^0$ es una version formal del mentiroso tradicional, la oracion que dice de s misma que es falsa, mientras que $\vdash T I$ es lo que se conoce como 'mentiroso reforzado'.

¹¹Otra candidata a realizacion del proyecto minimalista de Horwich en la teoria de putb pero deduc-tivamente mas poderosa es la teoría a presentada por [Schindler \(2014b\)](#).

generalisation is not substantially impaired by their existence." (Horwich, 1998b, p. 42, nota 21) Sin embargo, también es claro que toda teoría desentrecorolladora clásica puede extenderse consistentemente mediante [T-Out](#) y, por ende, no hay necesidad alguna de resignar siquiera esta mínima porción de utilidad de la verdad.

En cambio, el modelo de Leitgeb sí cumple con la condición (1).

Corolario 5.19. El predicado de verdad de la teoría fhN^+ ; pf ig tiene la propiedad de eliminación.

Demostración. Este es un corolario del corolario [3.38](#) (= pf) y la observación [5.7](#). \square

Aun más, siendo esta porción de la teoría semántica de Leitgeb evidentemente axiomatizable (podría tomarse simplemente $\text{pat} + \text{T-Out}$), prima facie la condición (3) también se satisface. No obstante, esta axiomatización no refleja ninguna característica especial del modelo de Leitgeb, es altamente incompleta. Fischer et al. (2015) investigan diversos modos de evaluar en qué casos un sistema constituye una axiomatización de cierta teoría semántica, y es claro que bajo ninguno de los criterios que consideran seriamente $\text{pat} + \text{T-Out}$ puede axiomatizar fhN^+ ; pf ig . Schindler (2014a) ofrece una versión axiomática más razonable del modelo de Leitgeb que también incluye [T-Out](#), pero en ella las adscripciones de verdad están relativizadas a un predicado primitivo para expresar el concepto de d -fundación, presumiblemente sustantivo, ciertamente no de nivel L^+_{pa} y, por ende, se aplican las objeciones de Glanzberg en este caso, con lo cual la condición (2) falla a.¹²

En el capítulo [8](#) ofrezco teorías de la verdad desentrecorolladoras que cumplen considerablemente mejor las exigencias (a)-(d) del proyecto minimalista extendido. De hecho, argumento que pueden entenderse como realizaciones totales de este proyecto. Aun más, muestro que también se satisfacen las condiciones (1)-(3) que he impuesto sobre toda teoría de acionista formal de la verdad: las teorías que ofrezco son correctas, no exceden la función básica de la verdad y, más importantemente, son consistentes con [T-Out](#), por tanto, su predicado veritativo puede fácilmente adquirir la propiedad de eliminación y servir los roles lógico-expresivos deseados.

5.3.2. De acionismo y paracompletitud

Una segunda corriente de acionista que vio en la transparencia de la verdad la pieza necesaria para que el predicado pueda ser utilizado para expresar conjunciones infinitas

¹²Existía una posibilidad de eliminar esta relativización en la axiomatización de Schindler, pero estudiarla me llevaría demasiado lejos.

opt por debilitar la lógica lo justo y necesario para evitar inconsistencias, en lugar de los principios de equivalencia, como el proyecto minimalista. Uno de los caminos a tomar es el rechazo del principio de Tercero Excluido, la elección de lógicas paracompletas (vease el apartado 3.1.1). En este apartado investigo las teorías paracompletas semánticas de Punto Fijo de Kripke, las cuales son transparentes (corolario 3.9), y el sistema axiomático k_f , al cual una de ellas dio lugar, a la luz de las condiciones (1)-(3).

Un primer caso es la teoría de Punto Fijo de Kripke sobre el esquema de valuación Weak Kleene. Si bien válida todas las instancias de $T\text{-Elim}$, no todas las inferencias de $\vdash v'$ a partir de $\vdash [t=v]$ son válidas, a menos que todas las instancias de $\vdash (v)$ reciban un valor de verdad clásico, i.e., falla la regla de introducción del existencial 19. Luego, el paso 7 de la derivación de (ET 3) (observación 5.9) resulta inválido en la teoría. Se puede probar que no existe modo de reparar la derivación:

Proposición 5.20. El predicado de verdad de la teoría de Punto Fijo de Kripke sobre el esquema de valuación Weak Kleene no tiene la propiedad de eliminación.

Demostración. Sea I^0 como en (5.8), i.e.,

$$N^+ \quad I^0 = p_T : I^0 q$$

Es sencillo ver que $T : I^0$ recibe valor $\frac{1}{2}$ en todo modelo de punto fijo sobre w_k , por ser paradójica. A continuación exhibo dos fórmulas $\vdash (x)$ y $\vdash L_T$ para las cuales (ET 3) no preserva verdad.

Sea $\vdash (0 \neq 0)$ y $\vdash (x) := (x = p_0 \neq 0q)$. Si M es un modelo de punto fijo sobre w_k , $v_{w_k}^M(0 \neq 0) = 0$ y, por ende, $v_{w_k}^M(\vdash (x)) = 1$. Asimismo, $v_{w_k}^M((0 \neq 0) \neq 0) = 1$. Luego, las premisas de ET 3 son verdaderas en todo modelo de punto fijo. No obstante, este no

es el caso de la conclusión. Notese que una de las instancias de $\vdash (x) \neq Tx$ es el condicional $\vdash (I^0) \neq TI^0$, cuyo valor de verdad es $\frac{1}{2}$, porque $v_{w_k}^M(TI^0) = \frac{1}{2}$ en todo modelo de punto fijo (vease la derivación 3.4). Como en w_k un enunciado universal recibe valor $\frac{1}{2}$ si alguna de sus instancias tiene este valor, $v_{w_k}^M(\vdash (x) \neq Tx) = \frac{1}{2}$ y, por tanto, $\vdash (x) \neq Tx$ recibe valor $\frac{1}{2}$ en todo modelo de punto fijo sobre w_k . \square

En cambio, la teoría de punto fijo de Kripke sobre el esquema Strong Kleene sí tiene la propiedad de eliminación, porque, de acuerdo con la derivación 3.2, es suficiente que una de las instancias de un enunciado universal reciba valor 0 para que el enunciado mismo reciba este valor y, en consecuencia, su negación valga 1.

Proposición 5.21. El predicado de verdad de la teoría de Punto Fijo de Kripke sobre el esquema de valuación Strong Kleene tiene la propiedad de eliminación.

Demostración. Es sencillo ver que k_3 valida las reglas [EI](#), [I \$\wedge\$](#) , [II](#), [contraposición](#), [E8](#)

e [I9](#). Además, por el corolario [3.9](#), [T-Elim](#) vale en todos los modelos de punto jo. Luego, por la observación [5.9](#), en estos modelos el predicado de verdad tiene la propiedad de eliminación. \square

Luego, la teoría de punto jo sobre sk satisface la condición (1) pero, como se~nale en el apartado [3.1.2](#), es demasiado compleja para ser axiomatizada. No obstante, existen axiomatizaciones parciales, como k_f en lógica clásica y otra denominada 'pkf' en k_3 , de la cual no me ocupé en esta investigación (cf. [Horsten \(2011\)](#), [Halbach & Horsten \(2005\)](#)), porque no es necesario. En efecto, k_f es ya una teoría adecuada de la verdad.

Corolario 5.22. El predicado de verdad de k_f tiene la propiedad de eliminación.

Demostración. Directamente de la observación [5.7](#) y el hecho de que k_f prueba todas las instancias de [T-Out](#) (vease el apartado [3.1.3.1](#)). \square

Además, los principios composicionales pueden ser todos explicados en función de los principios de transparencia que la clausura de los modelos de Punto Fijo de Kripke sobre sk validan.

5.3.3. De acionismo y composicionalidad

En este apartado evalúo las teorías composicionales axiomáticas y clásicas $ct(pa^+)$, $rt_{< \theta}$ (vease el apartado [2.3.3](#)) y fs (vease el apartado [3.3](#)) a la luz de las condiciones de acionistas (1)-(3) que di al comienzo de la sección.

Proposición 5.23. Los predicados de verdad de las teorías $ct(pa^+)$ y $rt_{< \theta}$ no tienen la propiedad de eliminación.

Demostración. Nuevamente, al ser teorías clásicas, por la observación [5.10](#), si sus predicados de verdad tuvieran la propiedad de eliminación, deberían implicar las instancias de [T-Out](#), o las instancias análogas para los predicados de verdad de $rt_{< \theta}$, dadas por oraciones de todo L_T o todo $L_{< \theta}$. Pero mientras que $ct(pa^+)$ no prueba $T \mid \mid :T \mid$, la instancia de [T-Out](#) para la oración $:T \mid$ del mentiroso, $rt_{< \theta}$ no prueba $T \mid \mid :T \mid$ para ningún predicado T , con $< \theta$, donde $q^+ \mid = p:T \mid q$. \square

No obstante, ambas teorías son consistentemente extensibles con [T-Out](#), mientras que sus axiomas composicionales están motivados por las instancias fundadas del

[Esquema-T](#) y, por ende, los sistemas resultantes podrán ser buenos candidatos para el de acionismo. fs, en cambio, no solo no es eliminativista sino que es inconsistente con [T-Out](#).

Proposición 5.24. Si th fs es una teoría consistente, T no tienen la propiedad de eliminación en th.

Demostración. Razono en fs + [T-Out](#). Si $\vdash T \mid$ es la oración del mentiroso (reforzado), por [T-Out](#), se tiene que $T \mid \vdash \neg T \mid$ y, por ende, que $\vdash T \mid$. Luego, fs2 implica que $T \vdash \mid$ que, nuevamente por [T-Out](#), implica $\vdash \neg T \mid$ o, lo que es lo mismo, $\vdash T \mid$. En consecuencia, fs th + [T-Out](#) es inconsistente. Como th fs es consistente, th $\not\vdash$ [T-Out](#) y, por la observación [5.7](#), th no tiene la propiedad de eliminación. \square

Por lo tanto, fs es una teoría inviable para el de acionismo.

5.3.4. De acionismo y paraconsistencia

Finalmente, un segundo camino que han tomado aquellos que sostienen que la transparencia es necesaria y suficiente para garantizar los usos lógicos del predicado veritativo y están dispuestos a abandonar la lógica clásica para poder adoptar principios de equivalencia completos consiste en elegir lógicas paraconsistentes, donde la regla Ex Falso Sequitur Quodlibet resulta inválida (vease el apartado [3.4.1](#)). A continuación evalúo las teorías paraconsistentes de verdad transparente lptt, pltt y sttt a la luz de las condiciones

(1)-(3). Al ser axiomáticas y estar dadas únicamente por principios de transparencia, las condiciones (2) y (3) son satisfechas automáticamente. No obstante, muestro que la condición (1) falla en los tres casos.

Como mencioné en el capítulo [3](#), lp no satisface la regla [E!](#) y, por tanto, los pasos 4 y 6 en la derivación de la inferencia [\(ET 2\)](#) (observación [5.7](#)) no son válidos en pltt. Una vez más, puede probarse que no hay modo de reparar la prueba.

Proposición 5.25. lptt no tiene la propiedad de eliminación.

Demostración. Muestro que existen dos fórmulas $\varphi(x)$ and ψ tales que, en todo modelo de lptt, las premisas de [\(ET 2\)](#) resultan verdaderas mientras que su conclusión, falsa. Sean $M \models_{lp} lptt$, $\models 0 \neq 0$ y $\varphi(x) := T(x \& !^0)$, donde $!^0 = pT \vdash !^0 q$. Notese que $\varphi(x)$ es verdadera en M de toda oración de L_T , porque, al ser paradójica, $v_{lp}^M(T \vdash !^0) = \frac{1}{2}$ y, por las cláusulas del condicional de lp y las derivadas de la conjunción (cf. de nición [3.2](#)), para toda $2 \in L_T$, $v_{lp}^M(! T \vdash !^0)$; $v_{lp}^M(T \vdash !^0 !)$; $v_{lp}^M(\$ T \vdash !^0) = \frac{1}{2}$. Luego, por inter-sustitutividad y la denotación de $!^0$, para toda x tal que $Sent_{LT}(x)$, $v_{lp}^M(T(x \& !^0)) = \frac{1}{2}$.

Adicionalmente, $\exists x(T(x, I) \rightarrow \neg T(x))$ es tambien verdadera en M. Luego, ambas premisas de [ET 2](#) son verdaderas en M, pero, natural-mente, $0 \neq 0$ no lo es. \square

Si bien Priest ha introducido un condicional a I_p para remediar la falla de [EI](#), esto es, del Modus Ponens, obteniendo as la logica pl (cf. de nicion [3.51](#)), en ella el condicional no satisface [MT](#), Modus Tollens, y, en consecuencia, la regla [ET 3](#) resulta invalida en su teoria transparente de la verdad $pltt$, donde el condicional utilizado no es ya el condicional material sino el nuevo, (de lo contrario [ET 2](#) falla, porque el condicional material de pl es el de I_p).

Proposicion 5.26. El predicado de verdad de $pltt$ no tiene la propiedad de eliminacion.

Demostracion. Nuevamente, exhibo dos formulas, $\neg(x)$ y $\neg\neg(x)$ para las cuales las premisas de [ET 3](#) resultan verdaderas en un modelo de $pltt$, pero su conclusion falsa. Sea $\neg(x) := (x = I^0)$ y $\neg\neg(x) := T(I^0)$. Por [\(5.8\)](#), $T(I^0) = I$ es verdadera en todo modelo de la teoria y, por el [Esquema-T de Priest](#), esto es,

$$T p'q' \quad \text{(Esquema-T de Priest)}$$

el [Esquema T](#) formulado con el nuevo condicional \rightarrow y las clausulas \neg para la negacion

y el condicional del esquema de valuacion de Priest ep (cf. de nicion [3.50](#)), tanto $T : I^0$ como $\neg T : I^0$ obtienen el valor $\frac{1}{2}$. En consecuencia, ambas premisas de [ET 3](#) son verdaderas en todo modelo. Pero a la vez $\neg\exists x(x = I^0 \rightarrow Tx)$, la conclusion de esta inferencia, no es verdadera en todos ellos. Si M_{pltt} , el [Esquema-T de Priest](#) implica que, en todo mundo $w \in W_M$, $v_{pl}^w(T(I^0) \rightarrow T(I^0)) = \frac{1}{2}$, que, junto con la clausula semantica para \rightarrow , implica que, en todo w tal que $w_{R_M} w = I^0$, $v_{pl}^{w_0}(T(I^0)) = \frac{1}{2}$ sii $v_{pl}^{w_0}(T : I^0) = \frac{1}{2}$. Como esto ultimo es el caso, lo primero tambien y, por ende, $T(I^0)$ puede obtener valor 1 o $\frac{1}{2}$ en cada mundo, por la sobreyectividad de R_M . Luego, por las clausulas del condicional en el esquema de valuacion ep , como los enunciados de identidad adquieren valores clasicos en todo mundo de todo modelo, existe un modelo M de $pltt$ y un $w \in W_M$ tales que $v_{pl}^w(\exists x(x = I^0 \rightarrow Tx)) = 1$ y, por ende, $v_{pl}^w(\neg\exists x(x = I^0 \rightarrow Tx)) = 0$. En consecuencia, existe un modelo de $pltt$ en el cual $\neg\exists x(x = I^0 \rightarrow Tx)$, la conclusion de [ET 3](#), es falsa. \square

Finalmente, muestro que, bajo ciertas condiciones razonables, la teoria de verdad transparente $sttt$ tampoco satisface la condicion de adecuacion (1), esto es, su predicado de verdad no tiene la propiedad de eliminacion si se la considera de cierta manera (vease el apartado [3.4.3](#)).

Dado que la lógica subyacente de sttt, st, no es transitiva, a diferencia de otras lógicas como la clásica, existe una disparidad entre las reglas y las metareglas que valen en la teoría. Esto abre una brecha entre lo que puede derivarse a partir de ciertas oraciones asertadas hipotéticamente, como supuestos o premisas, y lo que puede derivarse de ellas cuando se las afirma categóricamente, como axiomas o teoremas (vease el apartado 3.4.3.1). Luego, si bien, v.g., ET 2 vale cuando sus premisas ocurren en contextos hipotéticos, falla cuando se las afirma categóricamente.

Claramente, es deseable poder derivar no solo cuando se supone que todas las 's son verdaderas (y que es una '), pero también, y quizás aun más todavía, cuando se sabe o se afirma categóricamente que todas las 's son verdaderas. Es consecuente demandar entonces que ET 2 valga en la forma de regla en el cálculo de secuentes, en el cual sttt está formulada, i.e.,

$$\frac{}{; \exists v('v) ! Tv); '(p \rightarrow q) . ;} \quad (5.9)$$

pero también en forma de metaregla, es decir,

$$\frac{. \exists v('v) ! Tv); '(p \rightarrow q);}{; . ; ;} \quad (5.10)$$

Y otro tanto con respecto a ET 1 y ET 3. Lamentablemente, mientras que (5.9) vale en sttt, (5.10) no.

Proposición 5.27. El predicado de verdad de sttt no tiene la propiedad de eliminación.

Demostración. Sean $'(x) := (x = p \rightarrow q \wedge \$. I^0)$ y $0 = 0 \rightarrow T(x) := (0 = 0)$. A continuación

muestro que los secuentes

$$. \exists x(x = 0 = 0 \rightarrow T(x, \$. I^0) ! Tx) \quad (5.11)$$

$$p \rightarrow q \wedge$$

y

$$. \frac{0 = 0 = 0 = 0 \wedge T(0 = 0, I^0)}{p \rightarrow q \rightarrow p \rightarrow q \rightarrow p \rightarrow q, \$.} \quad (5.12)$$

son válidos en sttt. Sea M un modelo de sttt. Es sencillo ver que $v_{sk}^M(T : I^0) = \frac{1}{2}$ y que, por las cláusulas del condicional y la conjunción de sk (cf. definición 3.2), $v_{sk}^M(0 = 0 \rightarrow T : I^0) = \frac{1}{2}$ también. Por IS, $v_{sk}^M(T(p \rightarrow 0q, \$. I^0)) = \frac{1}{2}$. Por lo tanto, como todo modelo de sttt es un modelo clásico de pat, por la cláusula de la conjunción, se tiene que $v_{sk}^M(p \rightarrow 0q = p \rightarrow 0q \wedge T(p \rightarrow 0q, \$. I^0)) = \frac{1}{2}$, esto es, vale el secuyente (5.12). A su vez, por la cláusula del condicional, esto implica que $v_{sk}^M(p \rightarrow 0q = p \rightarrow 0q \wedge T(p \rightarrow 0q, \$. I^0) ! T p \rightarrow 0q) = \frac{1}{2}$. Por las condiciones impuestas sobre la identidad y

los enunciados universales, $\forall x (x = 0 \rightarrow T(x) \rightarrow T(x)) = 1$ completa la prueba del seciente (5.11). Pero, a la vez, es claro que, en ningún modelo M de sttt, $\forall x (0 = 0 \rightarrow 1) = 1$, porque las identidades se comportan de modo clásico. En consecuencia, hay instancias de ET 2 para las cuales las premisas son tolerantemente verdaderas en todo modelo, mientras que su conclusión es falsa, i.e., (5.10) es inválida en sttt. □

§

A lo largo de este capítulo he identificado las características principales que toda teoría formal de la verdad debe satisfacer para cuadrar con las tesis principales del de acionismo. A grandes rasgos, estos sistemas deben ser (1) capaces de garantizar los roles lógico-expresivos del predicado veritativo, i.e., asegurar la propiedad de eliminación; (2) correctos con respecto a la naturaleza desentremilladora de la verdad, esto es, no afirmar más que lo que se desprende del significado de la verdad; y (3) ser parcialmente axiomatizables, de modo tal que la teoría axiomática resultante cumpla asimismo con (1).

He mostrado que para satisfacer la condición (1) no es necesario contar con principios de transparencia completos sino solo una parte de ellos, la parte eliminativa. Dado que estas partes son ceteris paribus consistentes en el marco de la lógica clásica, argumenté que no hay necesidad de adoptar lógicas no clásicas en su lugar para obtener una buena teoría de la verdad de acionista. Aun más, mostré que algunas teorías no clásicas con predicados de verdad transparentes no logran satisfacer (1), por la falla de ciertos principios lógicos. En otras palabras, los principios de transparencia no son necesarios ni suficientes para garantizar los roles lógico-expresivos del predicado veritativo, contrario a la opinión generalizada.

6

Condiciones sobre una noción de referencia formal

"We either suspect that much philosophical work lies ahead of us before the question is finally settled, or that otherwise the question is ill-posed, i.e., that the talk of self-referentiality is to be banished from scientific contexts."

{ Hannes Leitgeb, 'What is a Self-Referential Sentence?'

Como se vio en el capítulo anterior, si bien el [Esquema-T](#) no es estrictamente ni necesario ni suficiente para garantizar los roles lógico-expresivos de la verdad sino que, en contextos clásicos, [T-Out](#) lo es, ciertas instancias correctas y consistentes del [Esquema-T](#) pueden resultar deseables por otros motivos, v.g., descriptivos, prácticos (vease el apartado [5.3](#)). En otras palabras, el proyecto minimalista de Horwich (vease el apartado [5.3.1](#)) resulta atractivo desde un punto de vista filosófico. No obstante, hallar un buen criterio de selección de instancias del [Esquema-T](#) que cuadre con el proyecto no es sencillo, como mostre en el apartado [5.3.1](#): los criterios disponibles resultan o bien demasiado restrictivos, o bien demasiado complejos, o bien demasiado arbitrarios.

Una estrategia prometedora e inexplorada consiste recurrir a la noción de referencia. La idea de que las expresiones que llevan a paradojas semánticas comparten ciertos patrones referenciales 'peligrosos' que son responsables por las contradicciones, lo que llamo 'visión referencialista de las paradojas semánticas', es considerada conocimiento común. Aun más, la postura ortodoxa sostiene que, más específicamente, estos patrones comunes se reducen a la autorreferencia. En cualquier caso, si la visión referencialista es acertada, es prima facie posible un criterio de restricción de instancias del [Esquema-T](#) basado en la noción de referencia que resulte consistente, amplio y poco ad hoc. Asimismo, a diferencia, v.g., de la noción de dependencia (vease el apartado [3.2](#)), la

referencia tiene en principio un carácter más sintáctico por estar basada en la sintaxis de las expresiones, como muestro en este capítulo, lo cual apunta a la posibilidad de un criterio simple de restricción. En consecuencia, un criterio tal parecería satisfacer todas las condiciones minimalistas.

No obstante, el concepto de referencia que se necesita no es fácil de caracterizar. Lo evidencian los diversos intentos en la literatura pero, principalmente, el debate que se generó alrededor de la paradoja de Yablo. Yablo introdujo una antinomia diseñada para refutar la tesis referencialista ortodoxa, i.e., que todas las paradojas semánticas son autorreferenciales. Esto fue a su vez puesto en duda, suscitando un debate inextricable que culminó cuando Leitgeb notó que diversas nociones incompatibles y poco claras de autorreferencia estaban en juego al mismo tiempo en la discusión. Esto se tradujo en el desafío de encontrar una noción adecuada y precisa de referencia aplicable al estudio de las paradojas semánticas, desafío que, modulo ciertas modificaciones, creo superar en el capítulo siguiente.

En este capítulo me ocupé de señalar las dificultades que, desde mi punto de vista, toda noción de referencia adecuada y precisa debe superar; así como de identificar las condiciones que debe satisfacer. Para eso, introduzco primero la paradoja de Yablo en detalle, seguida por el debate en torno a su presunto carácter no autorreferencial. En [6.2](#) presento el desafío de Leitgeb, esto es, las condiciones que él impone sobre toda noción de referencia. [6.3](#) se ocupa de abordajes previos a la noción de referencia y otros conceptos cercanos, señalando sus aciertos y sus desaciertos, correspondientemente. Finalmente, en [6.4](#) argumento en contra de una de las condiciones de Leitgeb para un concepto adecuado de referencia y concluyo con las condiciones que un concepto tal debe satisfacer.

6.1. La paradoja de Yablo

[Yablo \(1985, 1993\)](#) desafía la idea instalada de que todas las antinomias semánticas involucran algún tipo de autorreferencia, la visión referencialista ortodoxa (cf. [Henri Poincaré \(1906\)](#), [Russell \(1908\)](#)) mediante una paradoja informal [hoy llamada 'paradoja de Yablo'] dada por la siguiente lista infinita de oraciones, cada una de las cuales

parece referir únicamente a las que se encuentran más adelante en la lista:

Para todo $x > 0$; Y_x no es verdadera. (Y₀)

Para todo $x > 1$; Y_x no es verdadera. (Y₁)

:::

Para todo $x > n$; Y_x no es verdadera. (Y_n)

:::

Prima facie, al igual que en el caso del mentiroso, no es posible asignar valores de verdad clásicos consistentemente a las oraciones de esta lista. Supongase que alguna de ellas, (Y_n) , es verdadera. Luego, ninguna cuyo subíndice sea mayor a n lo es, i.e., (Y_{n+1}) no es verdadera y tampoco lo son aquellas que se encuentran por debajo de (Y_{n+1}) . Pero esto último es precisamente lo que (Y_{n+1}) afirma, razón por la cual (Y_{n+1}) debería ser verdadera después de todo. Dado que suponer que una oración cualquiera es verdadera ha llevado a una contradicción, se debe concluir que ninguna lo es. Pero si ninguna oración debajo de (Y_0) es verdadera, (Y_0) debe ser a la vez verdadera, lo cual es imposible.

Prima facie, nuevamente, no hay rastros de autorreferencia en la secuencia de Yablo sino que, como sostiene [Roy A. Sorensen \(1998, p. 139\)](#), "The technique substitutes the cramped circularity of self-reference with the luxuriant linearity of an infinite series." En otras palabras, Yablo sustituye la autorreferencia con la infundación. Una paradoja con estas características sacudir a los fundamentos de teorías de la verdad como las de Tarski y, trasladando la idea de Yablo al ámbito de los conjuntos y las clases, teorías como la de tipos de [Russell \(1903\)](#) se ven afectadas del mismo modo.¹

Como se vio en el capítulo 2, Tarski propone sortear las paradojas semánticas mediante una distinción entre lenguaje y metalenguaje, donde 'verdad' es siempre 'verdad en el lenguaje objeto' y solo puede ser aplicada desde un metalenguaje esencialmente más rico. En lugar de un único predicado veritativo para un único lenguaje, se obtiene una jerarquía de predicados y lenguajes tales que el predicado de verdad de cada lenguaje se encuentra en los lenguajes que están por encima en la jerarquía. Esto bloquea la paradoja del mentiroso y otros ciclos pero no es suficiente para bloquear la paradoja de Yablo. Como afirma [Kripke \(1975, p. 697\)](#), "[...] the orthodox approach by no means obviously guarantees groundedness [...] Even if unrestricted truth definitions are in question, standard theorems easily allow us to construct a descending chain of first-order languages L_0, L_1, L_2, \dots , such that L_i contains a truth predicate for L_{i+1} ."

¹Vease [Yablo \(2004\)](#) para paradojas conjuntistas in nitarias semejantes a la paradoja de Yablo.

Siendo esta jerarquía acumulativa $|L_0 L_1 L_2 : : |$ siguiendo a [Visser \(1989\)](#) (cf. [Thomas Forster \(2004\)](#) también) es posible formular una lista semejante a la de Yablo en estos lenguajes: en cada L_n hay una oración F_n que afirma que ninguna oración F_m de L_m , donde $m > n$, es verdadera, y es posible obtener una contradicción a partir de esta secuencia mediante un argumento análogo al anterior.

Naturalmente, las teorías de la verdad de Tarski no son inconsistentes, porque Tarski siempre comenzó por un primer lenguaje objeto y a partir de este construyó lenguajes más y más ricos que sirvieran de metalenguajes desde los cuales dar definiciones de predicados de verdad para los anteriores. Sin embargo, la idea tarskiana de que es suficiente establecer una distinción entre lenguaje objeto y metalenguaje para evitar las paradojas es en alguna medida (ya se verá por qué el matiz) falsa. Como sostiene Yablo,

So the Tarskian way of avoiding paradox relies on more than a rigid object-/metalanguage distinction. It is also required that the sequence of languages eventually grounds out in a bottom-level object language. ([Yablo, 2004](#), p. 141)

Dada la importancia de las consecuencias de una paradoja semántica no circular, poco tiempo después de su aparición la paradoja de Yablo fue puesta bajo un minucioso escrutinio. El primer paso fue su formalización en una extensión del lenguaje de la aritmética de primer orden con un predicado de verdad, L_T , como es usual en el caso de antinomias semánticas y más aun en el caso de la secuencia de Yablo, donde hay vocabulario aritmético involucrado. Luego se pasó a evaluar tres cuestiones en torno a esta formalización: primero, si su existencia puede probarse en la aritmética o en alguna de sus extensiones, como puede probarse, mediante el lema diagonal, la de la oración del mentiroso, y, segundo y tercero, si, de existir, la secuencia es realmente paradójica y no autorreferencial. Este último punto fue quizás el más controversial y llevó el debate a un punto muerto, debido a la falta de nociones adecuadas y correctas de autorreferencia y a la dificultad que encontrar nociones tales conlleva (cf. [Leitgeb \(2002\)](#)).

Las subsecciones que siguen están dedicadas a la breve exposición de estas cuestiones. Primero presento una versión formal de la paradoja de Yablo en L_T , muestro como puede probarse la existencia de la lista en q^+ y analizo su paradójicidad. Luego, en [6.1.2](#), el último apartado de la sección, presento el debate alrededor de la autorreferencialidad.

6.1.1. Formalización y paradojicidad de la secuencia de Yablo

El modo más natural de formalizar la lista de Yablo en L_T parece ser el siguiente:

$$\begin{aligned}
 (0) &= \neg \exists x > 0: T(x) \\
 (1) &= \neg \exists x > 1: T(x) \\
 &\vdots \\
 (n) &= \neg \exists x > n: T(x) \\
 &\vdots
 \end{aligned} \tag{6.1}$$

donde (x) es un término abierto del lenguaje. Para probar la existencia de estos términos no se puede recurrir a una definición recursiva porque, al revés de lo usual, los que tienen números más pequeños requieren de los que tienen números más grandes para su definición. El único modo de obtener términos (n) para cada $n \in \mathbb{N}$ como los de (6.1) es mediante un procedimiento diagonalizador, v.g., el lema diagonal fuerte (teorema 1.8) aplicado a $\exists y > x: T(y = px)$ sobre la variable z .² De acuerdo con este lema, existe un término t tal que

$$\neg \exists t = \neg \exists y > x: T(y = px)$$

Aplicando la función $y(x = px)$ a ambos lados del igual,

$$\neg \exists t(x = px) = \neg \exists y > x: T(y = px)(x = px)$$

Sea $(v) := t(v = px)$. Luego,

$$\neg \exists (x) = \neg \exists y > x: T(y)$$

o, lo que es lo mismo, $\neg \exists$ tiene como teorema al Principio Uniforme de Punto Fijo de Yablo (Uniform Fixed-Point Yablo Principle), i.e.:

$$\exists x (x) = \neg \exists y > x: T(y) \tag{UFPYP}$$

Claramente, del UFPYP se siguen todas las oraciones de la lista en (6.1). Llamo 'oración de Yablo' a cada una de estas oraciones y 'predicado de Yablo' a la fórmula $\exists y > v: T(y)$, denotada por (v) .

Al igual que la oración del mentiroso en

$$\neg \exists l = \neg \exists T l \tag{6.2}$$

²Donde $v_1(v_2=v_3)$ representa la función sustitución y v la función numeral; véase el apartado 1.2.3.

o el ciclo de mentirosos en

$$q^+ \cdot l_1 = pT \mid_2 q \wedge l_2 = p:T \mid_1 q$$

el patrón de referencia de la lista de Yablo puede obtenerse en la aritmética vía el lema diagonal.

Basándose en la formalización de la lista de Yablo en (6.1), la siguiente es una formalización del argumento mediante el cual Yablo deriva una contradicción a partir de ella:

1. $T(n)$	supuesto
2. $\exists y > n : T(y)$	1
3. $\neg T(n+1)$	2
4. $\exists y > n+1 : T(y)$	2
5. $T(n+1)$	4
6. ?	3, 5
7. $\neg T(n)$	1-6
8. $\exists y : T(y)$	7
9. $\exists y > 0 : T(y)$	8
10. $T(0)$	9
11. $\neg T(0)$	8

A primera vista todo parece estar en orden, pero una mirada de cerca encuentra que no es claro que principios justificados con los pasos 2 y 8 en esta derivación, porque no es claro si la letra n es una variable en la derivación o el numeral de un número dado. Si n es el numeral de un número dado, no se está frente a una derivación sino a un esquema de derivación; se tiene una inferencia por cada $n \geq 1$. En ese caso, el paso 2 está dado por (6.1) junto con el [Esquema-T](#) para oraciones de Yablo, esto es, el principio desentrecomillador local de Yablo (Local Yablo Disquotation Principle):

$$T \mid p \exists y > n : T(y) \iff \exists y > n : T(y) \quad (\text{LYDP})$$

para cada $n \geq 1$. Pero el paso 8 queda completamente injustificado, a menos que se cuente con reglas inferenciales como la regla \exists , que permitan pasar de todas las oraciones de la forma $\neg T(n)$ con $n \geq 1$ a $\exists y : T(y)$, que no es el caso.³ En efecto, como se señalan [James Hardy \(1995\)](#) y [Ketland \(2005\)](#),⁴

Proposición 6.1 (Hardy, Ketland). $\text{pat} + \text{LYDP}$ es consistente pero \neg -inconsistente.

³La regla \exists permite derivar $\exists v'(v)$ a partir del conjunto de instancias de $'$, esto es, a partir de las instancias premisas de la forma $'[n=v]$ con $n \geq 1$. Para una discusión en torno a la posibilidad de obtener una contradicción a partir de la lista de Yablo empleando la regla \exists u otras reglas inferenciales, por supuesto, no en lenguajes de primer orden, véase [Priest \(1997\)](#), [Beall \(2001\)](#) y [Selmer Bringsjord & Bram van Heuveln \(2003\)](#).

⁴[Visser \(1989\)](#) ha probado ya un resultado equivalente para su formulación de la paradoja.

Si, en cambio, n es una variable en la derivación, el paso 2 no puede justificarse mediante (6.1) y el LYDP sino que es preciso recurrir a versiones uniformes de estos principios, es decir, al UFPYP y al Esquema-T Uniforme para el predicado de Yablo (Uniform Yablo Disquotation Principle), i.e. (vease el apartado 1.2.3):

$$\begin{array}{l} \exists x (T(x) \rightarrow \neg T(x)) \\ \text{.} \\ \exists x (T(x) \rightarrow \neg T(x)) \end{array} \quad \text{(UYDP)}$$

Proposición 6.2. $\text{pat} + \text{UYDP}$ es inconsistente.

En este caso, el paso 8 es legítimo, pero se ha recurrido para obtener una contradicción a algo más que las oraciones de Yablo. No son las oraciones de Yablo aquellas que originan la paradoja sino el predicado. Luego, esta paradoja será autorreferencial o no en la medida en que el predicado de Yablo lo sea, sin importar que las oraciones no tengan ningún patrón autorreferencial.

Muchos no han siquiera considerado esta versión de la paradoja por no ser lo suficientemente el original.⁵ Después de todo, Yablo pretendió derivar una contradicción de las oraciones de la lista y no de las peculiaridades del predicado de Yablo. El hecho de que la contradicción no fuera tal por la ilegitimidad de la regla !, que solo sea posible probar una !-inconsistencia y no una inconsistencia a secas, hace de su paradoja una !-paradoja, como la llama Ketland (2004).⁶ En cualquier caso, la lista resulta problemática en el contexto de la aritmética en tanto las !-inconsistencias son casi tan indeseables como las contradicciones, por bloquear la posibilidad de tener interpretaciones estándar, expansiones de \mathbb{N} , de las teorías en las que surgen (veanse el apartado 3.1.3.1, Leitgeb (2007), Barrio & Picollo (2013)).

6.1.2. Autorreferencia en la secuencia de Yablo

El primero en desarrollar el contraejemplo de Yablo a la autorreferencialidad de todas las paradojas semánticas fue Priest (1997). Desde su punto de vista, oraciones como el mentiroso en (6.2) son autorreferentes porque son puntos fijos de algún predicado y , en sus propias palabras,

[. . .] the paradox concerns a predicate, T , of the form $\exists x (T(x) \rightarrow \neg T(x))$; and the fact that $\exists x (T(x) \rightarrow \neg T(x))$ shows that we have a fixed point, y , here, of exactly the same self-referential kind as in the liar paradox. In a nutshell,

⁵Vease Picollo (2013), Picollo & Buacar (2014) para una argumentación explícita.

⁶Aunque la regla ! no es necesaria para obtener una simple paradoja. En (2012) muestro que, expresada en el lenguaje de la aritmética de segundo orden con un predicado de verdad la lista es insatisfacible.

is the predicate 'no number greater than x satisfies this predicate'. The circularity is now manifest. (Priest, 1997, p. 238)

Podría argumentarse que, en la medida en que no se recurra al UFPYP y se tenga simplemente una λ -paradoja, no hay circularidad involucrada. Sin embargo, Priest responde, la circularidad surge del hecho de que cada oración de Yablo, (n) , es un valor de la función, que está definida en términos de sí misma, que es un punto fijo. Las oraciones de Yablo son todas instancias de un predicado circular y, por ende, son circulares ellas mismas.⁷

Como nota Leitgeb (2002), parece haber dos nociones diferentes de autorreferencia en juego: una más bien intuitiva y una más bien técnica. La primera entiende la autorreferencia de oraciones como referencia de estas a sí mismas, y se basa en un concepto intuitivo de referencia o contenido de oraciones a oraciones, que a su vez se sostiene sobre la referencia o denotación de los términos que en ellas ocurren, esto es, en sus rasgos sintácticos, y el rango de cuantificación de sus cuantificadores.

This usage of 'refers to' and 'says of' seems to presuppose that the usual reference relation ref , which holds between (singular or general) terms and their referents, is extended or complemented by a reference relation holding between sentences and objects (but where the referents of the sentences are not the truth values of the sentences). (Leitgeb, 2002, p. 4)

De acuerdo con este concepto, una oración tiene dos medios de referir a otra: por mención y por cuantificación, y puede hacerlo directa o indirectamente: la referencia es transitiva.⁸ Una oración, v.g., de la forma $\lambda x (x > 0) : T(x)$ refiere intuitivamente a lo que x denota. Si $x = 0$, la oración no es autorreferencial, pero si $x = \lambda y (y > 0) : T(y)$, como en el caso del mentiroso, sí lo es. En general, una oración 'refiere a otra directamente por mención en la medida en que contenga un término que denote.

Además, intuitivamente, oraciones como $\lambda x (\lambda y (y > x) : T(y))$ refieren a los λ 's directamente por cuantificación. En este sentido, la n -ésima oración de Yablo refiere a las oraciones cuyo código es mayor a n . Como, por ejemplo, es claro que el código de la primera oración, $\lambda x > 0 : T(x)$, es mayor que 0, estas son autorreferenciales. Sin embargo, es posible reformular las oraciones de Yablo de modo tal que esto no ocurra:

$$(n) = \lambda y (\lambda z (y < z) : T(z))$$

⁷Para una discusión alrededor de este punto vease también Sorensen (1998) y Beall (2001). ⁸El esbozo de esta noción intuitiva puede ya encontrarse en Herzberger (1970).

Ahora, cada oración $\exists y(\exists z(y = (z) \wedge z > n) \rightarrow \neg y)$ refiere intuitivamente solo a las oraciones en la secuencia que se encuentran por debajo. Luego, aun considerando la referencia indirecta de cada una de estas oraciones, i.e. la clausura transitiva de la referencia directa, ninguna de ellas refiere a sí misma. Esta parece ser la noción que Yablo y Sorensen, entre otros, tienen en mente cuando sostienen que la lista de Yablo no es autorreferencial.

Por otro lado, Leitgeb nota la noción de autorreferencia de Priest, ligada a la presencia de puntos jos. Este concepto está inspirado en los trabajos de Gödel. Suele decirse que es la presencia de puntos jos, que pueden obtenerse via lemas diagonales, lo que codifica la autorreferencia en lenguajes aritméticos, razón por la cual la oración de Gödel sería autorreferencial. De acuerdo con este concepto 'técnico' de autorreferencia, una fórmula ϕ de un lenguaje de primer orden L es autorreferencial si es equivalente a otra fórmula ψ que contiene un término que denota ϕ , esto es,

$$M \models \phi \leftrightarrow \psi(\phi) \quad (6.3)$$

en todo modelo M que extiende N a todo L . Por ejemplo, la oración $\neg \text{TI}$ (es equivalente a sí misma) contiene un término, TI , que la denota; y el predicado de Yablo $\exists y > x: \neg y$ contiene un término, y que lo denota. En consecuencia, en las oraciones de Yablo hay 'involucrado' algún tipo de autorreferencia. Claramente ambas nociones de autorreferencia empleadas en el debate acerca de la paradoja de Yablo son incompatibles entre sí.

§

La paradoja de Yablo y el debate que le siguió no solo hicieron tambalear la visión referencialista ortodoxa que considera que la autorreferencia es común a todas las paradojas semánticas sino que pusieron en evidencia el carácter informal, la poca precisión con la cual esta postura y la tesis referencialista en general están formuladas. De este modo, impulsaron la búsqueda de nociones adecuadas y precisas de autorreferencia y referencia que pudieran utilizarse para evaluar los patrones referenciales detrás de expresiones paradójicas del lenguaje de la verdad, como la oración del mentiroso o las oraciones de Yablo, pero también de las oraciones mismas del lenguaje de la aritmética, principalmente para poder dar una explicación unificada de los patrones de referencia circulares detrás, v.g., de las oraciones del mentiroso y Gödel.

Lamentablemente, toda búsqueda de este tipo enfrenta serias dificultades, como mostramos [Leitgeb \(2002\)](#) y como evidencian los sucesivos intentos que han resultado prácticamente en fracasos. En las dos secciones que siguen presento las dificultades

por Leitgeb y varios intentos previos de dar nociones de referencia, que considero poco exitosos, con el objetivo de identificar los obstáculos principales que tendré que superar en el próximo capítulo.

6.2. Condiciones de Leitgeb sobre la referencia

Como señala [Leitgeb \(2002\)](#), las nociones intuitiva y técnica en juego en el debate sobre la autorreferencialidad de la lista de Yablo no son solo incompatibles entre sí sino de -cientos. La noción intuitiva es un concepto incompleto, en tanto no se ha especificado la referencia por cuantificación de cualquier oración sino solo de aquellas que constan de un cuantificador universal seguido por una expresión condicional. Para poder emplear este concepto en contextos científicos es necesaria su compleción.

De acuerdo con Leitgeb, esto no es tarea sencilla, porque toda noción de referencia debe a satisfacer lo que el denomina 'condición de equivalencia', esto es, la idea de que dos oraciones lógicamente equivalentes deben referir a lo mismo. Leitgeb da tres argumentos parcialmente solapados a favor de la condición de equivalencia.

En primer lugar, sostiene que dos oraciones lógicamente equivalentes son extensionalmente equivalentes, no solo en este mundo, sino en todo mundo posible y, por ende, indistinguibles desde un punto de vista lógico. Segundo, que la referencia es una noción semántica, en tanto esta se define en términos de la referencia de construcciones nominales a objetos. Y, tercero, que si la condición de equivalencia no se satisface y, es más, si dos oraciones aritmeticamente equivalentes no refieren a lo mismo, "[...] no philosopher may any longer argue in the following way: 'By Gödel's diagonalization lemma, we know that there is a sentence ϕ such that ϕ is equivalent to $\neg T(\phi)$ ' in arithmetic. Thus there is a self-referential sentence, that is, ϕ ." ([Leitgeb, 2002](#), p. 9) Y concluye: "If [the equivalence condition] is not true, the self-referentiality or circularity of a sentence does not only depend on what the sentence says, but also in which way its content is being expressed." ([Leitgeb, 2002](#), p. 7). Leitgeb considera que la referencia debe ser extensional.

No obstante, si la referencia intuitiva satisface la condición de equivalencia, nota Leitgeb, todas las oraciones refieren entre sí tanto por mención como por cuantificación.

Dadas dos fórmulas ϕ y ψ de L_T , existe una fórmula lógicamente equivalente a ϕ que contiene un término que denota ψ , porque

$$\neg \exists x (\phi \wedge x = \psi) \quad (6.4)$$

Además, toda oración de la forma $\exists v((v) \rightarrow (v))$ es lógicamente equivalente a

$$\exists v((v) \rightarrow ((v) \rightarrow (v))):$$

De valer la condición de equivalencia todas las oraciones serían autorreferenciales, se trivializaría el concepto.

Con respecto a la noción 'técnica' de autorreferencia, como se señalan [Leitgeb \(2002\)](#) y [Cook \(2006\)](#), es directamente trivial: toda fórmula ϕ es un punto fijo de algún predicado; v.g., $\phi \wedge x = x$, como se ve en (6.4). Aun más, es posible también construir predicados que no sean trivialmente (lógicamente) equivalentes a la fórmula original ϕ , como muestra [Cook \(2006\)](#): diagonalizando débilmente el predicado $\phi(x) \rightarrow (p \rightarrow q)$ sobre x se obtiene una fórmula tal que

$$\text{pat } \phi \rightarrow (\phi \rightarrow (p \rightarrow q))$$

lo cual implica que

$$\text{pat } \phi \rightarrow (\phi \rightarrow (p \rightarrow q))$$

Luego, toda oración sería autorreferente.

Podría replicarse que la equivalencia en (6.3) no es suficiente para que ϕ sea un punto fijo de sí mismo sino que es preciso una identidad. Una oración sería autorreferencial en este nuevo sentido si contiene un término t tal que

$$\phi^+(t) = p(t)q$$

para alguna fórmula ϕ del lenguaje. Pero si este fuera el caso, como sostiene Leitgeb, se debería concluir que en lenguajes sin símbolos de función adicionales, como L_{pa} , donde el lema diagonal fuerte no es un teorema sino solo el débil, no hay autorreferencia. Por ejemplo, oraciones del mentiroso 'débiles' dadas por equivalencias como en la prueba del teorema 2.9 de Tarski [i.e., $\phi : T p \rightarrow q$] no serían autorreferenciales, aunque dan lugar a paradojas tanto como sus contrapartes 'fuertes', dadas por identidades. Y, si bien las oraciones de Yablo resultarían autorreferenciales, se podría formular otras recurriendo a un predicado ϕ a el lema de diagonalización débil que no lo serían.

Como un segundo camino alternativo para bloquear la trivialidad podría argumentarse que no todas las equivalencias en (6.3) han de ser consideradas sino solo aquellas que se obtienen mediante un proceso de diagonalización, pero, hasta que no esté especificado que procesos darían lugar a puntos fijos 'legítimos' y cuáles no, no se puede avanzar con esta propuesta (cf. [Halbach & Visser \(2014a,b\)](#)).

Mientras que Cook decide, ante la trivialización de la autorreferencia, abandonar el terreno de la aritmética para la formulación de paradojas intuitivas y el estudio de los patrones referenciales detrás de las expresiones que dan lugar a paradojas semánticas,⁹

[Leitgeb \(2002](#), p. 9), por su parte, plantea una pregunta: "[. . .] what might a formally correct and materially adequate definition of self-referentiality look like?" y concluye:

[. . .] we either suspect that much philosophical work lies ahead of us before the question is finally settled, or that otherwise the question is ill-posed, i.e., that the talk of self-referentiality is to be banished from scientific contexts. ([Leitgeb, 2002](#), p. 11)

Hasta que no se cuente con una noción adecuada de autorreferencia [lo cual podría ser imposible] no solo no es posible responder a la pregunta por la autorreferencialidad de la lista de Yablo sino que se debe suspender el juicio con respecto al carácter autorreferencial de las oraciones de Gödel, la oración del mentiroso y otras expresiones que usualmente son clasificadas de ese modo, así como la visión referencialista, ortodoxa y no ortodoxa, sobre las paradojas semánticas.

En el capítulo siguiente respondo al desafío de Leitgeb y, consecuentemente, a la pregunta por la autorreferencialidad de la paradoja de Yablo y los patrones de referencia subyacentes a las paradojas semánticas, y a otras cuestiones los casos que tienen que ver con la noción de autorreferencia y que permanecen sin responder. Las nociones de autorreferencia que ofrezco son compleciones diferentes de la noción intuitiva pero, además, capturan la intuición detrás de la noción técnica de Priest. La estrategia yace, si se quiere, en morder la bala: las nociones que proveo están de nidad a partir de conceptos de referencia que, a diferencia de lo que quiere Leitgeb, son altamente intensionales o, mejor dicho, hiperintensionales (cf. [Björn Jespersen & Marie Duzi \(en prensa\)](#)), i.e., no satisfacen la condición de equivalencia, ni siquiera en su versión más débil: dos oraciones lógicamente equivalentes no referirán necesariamente a lo mismo. Sin embargo, como argumento en la sección 6.4, esta hiperintensionalidad es en realidad una característica deseable y que no conlleva las consecuencias indeseadas que señala Leitgeb.

⁹En ([Cook, 2004, 2014a](#)) introduce lenguajes intuitivos diseñados para este propósito y da definiciones de referencia, autorreferencia, fundamentación, etc. para sus expresiones, de acuerdo con las cuales las oraciones de Yablo formuladas en el nuevo lenguaje no resultan autorreferenciales en ningún sentido. Posteriormente, [Landon Rabern, Brian Rabern, & Matthew Macauley \(2013\)](#) y [Timo Beringer & Schindler \(2014\)](#) progresaron en esta dirección.

6.3. Referencia y contenido en la literatura

Los términos 'referencia' y 'contenido' ('aboutness') son muy frecuentes en diversas ramas de la filosofía, especialmente en lógica y filosofía del lenguaje, pero también en epistemología, metafísica, etc. La paradoja de Yablo y las paradojas en general no son la única razón para trabajar sobre los conceptos correspondientes y dar definiciones formalmente correctas y materialmente adecuadas de estos términos. Tampoco lo son las oraciones de Gödel ni otras cuestiones lógico-matemáticas. No hay un único sentido de 'referencia' o 'contenido' sobre el cual se haya trabajado en la literatura; se ha escrito sobre el contenido o la referencia de estados mentales, proposiciones, hechos, nombres, descripciones, oraciones, etc. Por obvias razones, solo me interesa aplicar los términos 'referencia' o 'contenido' a estas últimas, y voy a ignorar deliberadamente todo lo escrito en otras direcciones.

Pero incluso restringiéndose a oraciones hay nociones diferentes en la literatura, de las cuales no todas son relevantes o tendrán un lugar en esta sección, por llevarme demasiado lejos (cf. [Rudolf Carnap \(1937\)](#), [David Lewis \(1988b\)](#) y [Yablo \(2014\)](#)). [Ryle \(1933a\)](#), quizás el primero en tratar el tema explícitamente, distingue dos sentidos generales de 'contenido' aplicado al lenguaje: el contenido lingüístico y el contenido referencial. Mientras que el contenido lingüístico se ocupa de lo convenido por la oración o de lo que los términos que en ella ocurren designan, el contenido referencial atiende a su estructura sintáctica y lógica, se basa únicamente en los nombres propios, descripciones (definidas e indefinidas) y predicados que ocurren en estas oraciones. En palabras de [Goodman \(1961, p. 3\)](#), "Our sole problem (and it will prove troublesome enough) is to determine what a sentence is about, given what its terms designate" (mis énfasis). El contenido referencial es una relación entre las oraciones y los objetos a los cuales sus componentes refieren, sean estos los que fueren. Por ejemplo, el contenido lingüístico de "La estrella más brillante sale al atardecer" puede ser Venus o las estrellas, las cosas brillantes, el espacio exterior, o también el conjunto de mundos posibles sobre el cual el valor de verdad de esta oración superviene, etc., pero referencialmente su contenido es, al menos, cualquier cosa que sea una estrella y brille más que las otras, en la medida en que algo así exista.

El contenido referencial de una oración es el tipo de concepto que estoy buscando regimentar, una noción de corte lógico aplicable a lenguajes formales. Es la clase de noción a la que se apunta cuando se habla de autorreferencia relativamente a las paradojas, como la del mentiroso o la de Yablo, o cuando se dice que la autorreferencia es posible en la aritmética y puede ser empleada para probar la incompletitud de esta teoría. >¿Qué principios debe satisfacer una noción de este tipo? [Leitgeb \(2002\)](#), como se

vio en la sección anterior, menciona tres características deseables, no necesariamente en orden de importancia.

Primero, que respete ciertas intuiciones provenientes de la gramática del lenguaje natural, como que $P(a)$ era solo a lo denotado por a y que $\forall v(\neg(v) \rightarrow (v))$ era únicamente a los 's. Basándose en intuiciones de este tipo es que se dice que el mentiroso es autorreferencial mientras que las oraciones de Yablo no lo son. Un concepto de contenido referencial desligado completamente del lenguaje natural carece de sentido. Segundo, que las equivalencias e identidades que se obtienen v a los lemas diagonales débil y fuerte sean evidencia suficiente para la autorreferencia. Y, en tercer lugar, que satisfaga la condición de equivalencia.

Como mostró Leitgeb, estas tres condiciones no parecen ser compatibles, y en la literatura no hay nociones que las satisfagan simultáneamente. En su lugar, es posible distinguir dos sentidos de 'contenido referencial': uno, usualmente llamado 'referencia' (reference), ligado a la estructura sintáctica y lógica, hiperintensional, y otro, bajo el nombre de 'contenido' (aboutness) o 'contenido lógico' [para diferenciarlo de otros tipos de contenido], que depende de alguna manera de la estructura sintáctica pero es extensional, esto es, satisface la condición de equivalencia. Por ejemplo, mientras que, si a y b no denotan lo mismo, las oraciones $a = a$ y $b = b$ eran intuitivamente a distintos objetos, su contenido es el mismo, son ambas verdades lógicas.

Si bien son muy frecuentes en el lenguaje cotidiano de la comunidad logicoparlante, las nociones de contenido lógico y referencia son utilizadas únicamente en sentido intuitivo y son rara vez el foco de atención, quizás porque la investigación es percibida como un callejón sin salida, como se vio hacia el final del apartado anterior. Incluso cuando se emplean nociones de este tipo central y quasi técnicamente suele renunciarse a la precisión, como el caso de Herzberger:

Momentarily conceding sense to the notion of aboutness, each sentence has a certain domain. [. . .] The general notion of a domain is more readily indicated than explicated, but the analysis to follow depends on no problematic cases, and ultimately proves independent of any particular explication of 'domain'. (Herzberger, 1970, p. 147)

La desatención al concepto de contenido referencial quedó en evidencia en el debate sobre la paradoja de Yablo, pero tiene también consecuencias en otras áreas. Recientemente, Halbach & Visser (2014a,b) han resaltado la falta y la necesidad de nociones formalmente correctas y materialmente adecuadas de referencia y autorreferencia (entre

otras) en lenguajes formales, exhibiendo una serie de problemas los cuales que, sostienen, requieren formulaciones precisas de estos conceptos para ser siquiera propiamente enunciados.

La pregunta de [Leon Henkin \(1952\)](#) si la oración que dice de sí misma que es demostrable en una teoría \mathcal{T} es demostrable o independiente en \mathcal{T} es el primero de los problemas que Halbach y Visser identifican. Las oraciones de este tipo se conocen como 'oraciones de Henkin' de \mathcal{T} . Supongase que la fórmula $Bew_{\mathcal{T}}(v)$ del lenguaje de \mathcal{T} es un predicado de demostrabilidad para \mathcal{T} en la medida en que represente debilmente el conjunto de teoremas de esta teoría (veanse los apartados [1.2.2](#) y [1.2.3](#)). Por el teorema de Löb (corolario [1.14](#)), se sabe hoy que, si el predicado de prueba $Bew_{\mathcal{T}}$ que se utiliza satisface adicionalmente las condiciones de Löb (vease el apartado [1.2.3](#)), toda oración de Henkin para una teoría que contenga pa^+ debe ser demostrable en esa teoría, pero en 1952 Löb aun no había publicado sus resultados sino que, al contrario, estos estuvieron inspirados en el problema de Henkin (cf. [Löb \(1955\)](#)).

Como correctamente señalan Halbach y Visser, no es suficiente con tener que

$$\mathcal{T} \vdash \ulcorner Bew_{\mathcal{T}}(p) \urcorner \tag{6.5}$$

para que $\ulcorner \cdot \urcorner$ sea una oración de Henkin, porque todo teorema de \mathcal{T} , $0 = 0$ por ejemplo, satisface [\(6.5\)](#). Pero es ciertamente incorrecto decir que todo teorema de \mathcal{T} es, no solo autorreferencial, sino además una oración de Henkin. Una observación análoga podrá hacerse con respecto a las oraciones de Gödel. Aquí queda claro que tanto Henkin como Halbach y Visser no piensan en la noción de contenido sino la de referencia.

Al igual que en el caso de los puntos bajos en la lista de Yablo, es tentador reforzar el criterio y pedir una identidad, un punto bajo demostrable mediante el lema diagonal fuerte, esto es, llamar 'oración de Henkin' solo a expresiones de la forma $Bew_{\mathcal{T}}(h)$ tales que

$$\mathcal{T} \vdash \ulcorner h \urcorner = \text{p}Bew_{\mathcal{T}}(h) \tag{6.6}$$

De hecho, esto es lo que [Henkin \(1952\)](#) y [Georg Kreisel \(1953\)](#), el primero en ofrecer una solución al problema, tienen en mente: que una oración refiere a sí misma en la medida en que contenga un término que la denote; lo que Halbach y Visser llaman 'criterio de Kreisel-Henkin' para la autorreferencia. Sin embargo, nuevamente, esto choca con la intuición de que la autorreferencia está detrás de oraciones como el mentiroso débil en la prueba del teorema [2.9](#) de Tarski, que la autorreferencia es posible incluso en aquellos lenguajes aritméticos que no contienen símbolos de función, como L_{pa} , como el que Gödel utilizó en su prueba.

En esta sección recorro los primeros intentos de Ryle, Putnam, Goodman y Urbaniak de dar nociones más precisas y adecuadas de contenido referencial, señalando sus ventajas y desventajas. Algunas de ellas estarán más cerca de la noción intuitiva de referencia, otras de la de contenido. En cada caso muestro por qué esos intentos son insuficientes, por qué es necesario profundizar la investigación hasta dar con un criterio plenamente satisfactorio, como creo dar en el capítulo siguiente. Al analizar el recorrido va a quedar claro que condiciones debe cumplir la noción formal de referencia que busco, lo cual va a servir posteriormente como hilo conductor.

6.3.1. Referencia según Ryle

Respondiendo a la clásica pregunta por la estructura lógica de las oraciones que contienen términos propios y predicados ontológicos como 'existencia', 'ser un objeto', etc., [Ryle \(1933b\)](#) se ve en la necesidad de dar ciertas precisiones para una noción de contenido referencial. Su propuesta es la siguiente:

When a proposition asserts that a something named or described in that proposition has a certain quality, is of a certain kind, is in a certain state, and when further something does, in fact, possess that name or answer to the description, then the proposition is about the thing with that name or with the properties compiled in the description. ([Ryle, 1933b](#), p. 25)

De acuerdo con Ryle, "El Aconcagua tiene la cima nevada" refiere al Aconcagua y "El pico más elevado de América tiene la cima nevada" también, pero "La primera dama de Argentina es morocha" no refiere a nada. Todo esto parece correcto y se corresponde con la noción intuitiva de referencia por mención de [Leitgeb](#) que introduce en el apartado [6.1.2](#).

La propuesta de Ryle es, sin embargo, a todas luces limitada. En primer lugar porque se ocupa exclusivamente de descripciones de cosas y deja de lado oraciones con descripciones que se aplican a más de un objeto, como "Los picos de los Andes tienen la cima nevada", que intuitivamente refiere a los picos de los Andes (y a los Andes), aunque da una pista de cómo lidiar con ellas. En segundo lugar, porque, al igual que el esbozo de la noción intuitiva de referencia de [Leitgeb](#), no da idea alguna de la referencia de oraciones que no contienen ni nombres propios ni descripciones de ningún tipo, como "Todo es idéntico a sí mismo" o "Ninguna montaña que se encuentre en el hemisferio tiene la cima nevada". En otras palabras, deja de lado la referencia por cuantificación.

Por otro lado, es claro que una noción como la de Ryle no puede satisfacer la condición de equivalencia, so pena de ser trivial, por ejemplos como [\(6.4\)](#). Ryle ofrece

una noción que depende absolutamente de las construcciones nominales que ocurren en las oraciones; es hiperintensional y, por ende, más cerca de la referencia que del contenido.

6.3.2. Contenido lógico según Putnam

[Putnam \(1958\)](#), al contrario de Ryle, se centra en el contenido referencial de oraciones cuantificadas y busca una noción que satisfaga la condición de equivalencia. Además, también desea rescatar la primera condición de Leitgeb: desde su punto de vista, una oración como "Todas las montañas tienen la cumbre nevada" es acerca de las montañas y no de las cosas que tienen la cumbre nevada. Pero como esta expresión es lógicamente equivalente a "Las cosas que no tienen la cumbre nevada no son montañas", que es acerca de las cosas que no tienen la cumbre nevada y no sobre las que no son montañas, hay una tensión que resolver.

Putnam propone una noción de contenido lógico que relaciona oraciones con clases de objetos. Sin embargo, formalmente la de noción entre oraciones y fórmulas del lenguaje con una variable libre que, de alguna manera, identifica con la clase de objetos que estas fórmulas describen. Sea L un lenguaje de primer orden con un número finito de símbolos de predicado monádicos P_1, \dots, P_m y constantes de individuo a_1, \dots, a_n . Putnam sugiere primero sustituir la pregunta por si una oración es acerca de una clase con la pregunta por la cantidad de información que esta oración da sobre esta clase. Para eso recurre a las descripciones de estado y a la cantidad de información simpliciter que una oración da.

Definición 6.3 (Descripción de estado). Una oración $\phi \in L$ es una descripción de estado si es una conjunción que, para cada letra de predicado P y constante de individuo a , tiene como conjuntos $\{a \mid P a\}$ o $\{a \mid \neg P a\}$, pero no a ambos.

Notese que toda oración de L que no sea una contradicción puede escribirse como una disyunción de descripciones de estado, porque los cuantificadores en el lenguaje no son más que conjunciones lógicas.

Definición 6.4 (Cantidad de información). Sea P una función de probabilidad tal que, si a es una constante de individuo y P una letra de predicado, $P(Pa) = \frac{1}{2}$. La cantidad de información que da una oración $\phi \in L$ (ci(ϕ)) es igual a $1 - P(\phi)$.¹⁰

Luego, una contradicción da el máximo de información, seguida por las descripciones de estado, y las tautologías dan 0 información. Una conjunción de fórmulas atómicas

¹⁰

Esta es, en realidad, una simplificación de la propuesta de Putnam, con idénticas consecuencias.

o sus negaciones da siempre más información que cada uno de sus conjuntos, que a su vez dan más información que puestos en disyunción con alguna otra oración independiente.

Para establecer la cantidad de información que una oración da acerca de una clase es preciso que se sepa que cuenta como información acerca de una clase: todas las fórmulas que afirman o niegan que algo pertenece a esta clase y, una vez que se ha dicho que tales o cuales cosas están en la clase, todo lo que se afirma o niega sobre estas cosas. Por ejemplo, $(a_1) \wedge \neg(a_2) \wedge P_1(a_1) \wedge P_3(a_2)$ da información sobre la 'clase' (que puede ser una fórmula compleja) pero no toda la información que da es acerca de a_1 , i.e., $P_3(a_2)$ no es acerca de a_1 porque se dijo que $\neg(a_2)$.

Definición 6.5 (Información sobre una clase). La información que una oración ϕ da sobre una clase L ($I(L)$) es la conjunción de $P_1 a_1 \wedge \dots \wedge P_n a_n$ con

1. todas las fórmulas ϕ de la forma $(a) \vee \neg(a)$ que ϕ implica lógicamente, donde a es una letra de individuo, y
2. todas las fórmulas ϕ de la forma $P a$ o $\neg P a$ que ϕ implica lógicamente, donde a es una letra de individuo tal que ϕ implica lógicamente (a) y P una letra de predicado.

Agrego $P_1 a_1 \wedge \dots \wedge P_n a_n$, cuya cantidad de información es siempre 0, como conjunto a todo $I(L)$ para que $I_x(y)$ tenga una imagen cuando y no da información alguna sobre x . En cualquier otro caso voy a omitir este conjunto.

Definición 6.6 (Cantidad de información acerca de una clase). La cantidad de información que una oración ϕ da sobre una clase es la cantidad de información que da la fórmula que resulta de

1. escribir ϕ como una disyunción de descripciones de estado $\phi_1 \vee \dots \vee \phi_k$ contradictorias tomadas de a dos y, luego,
2. reemplazar cada ϕ_i por $I(\phi_i)$.

Es preciso dejar de lado las contradicciones porque no pueden expresarse como disyunciones de descripciones de estado contradictorias tomadas de a dos.

Definición 6.7 (Contenido lógico). Si ϕ no es una contradicción, $I(\phi)$ es acerca de la clase L si la cantidad de información que da ϕ y la cantidad de información que da $I(\phi)$ sobre L coinciden, i.e., $I(\phi) = I(I(\phi))$.

Grosso modo, lo que se esta pidiendo es que ' solo a rme cosas sobre los miembros de . Oraciones como $\exists x(P_1x \wedge P_2x)$ son acerca de P_1 , porque, si a_1 y a_2 son las unicas constantes de individuo del lenguaje, expresarlas como una disyuncion de descripciones de estado da por resultado la formula

$$\begin{aligned} & (P_1(a_1) \wedge P_1(a_2) \wedge P_2(a_1) \wedge P_2(a_2)) \vee (\neg P_1(a_1) \wedge \neg P_1(a_2) \wedge \neg P_2(a_1) \wedge \neg P_2(a_2)) \vee \\ & (P_1(a_1) \wedge \neg P_1(a_2) \wedge P_2(a_1) \wedge P_2(a_2)) \vee (\neg P_1(a_1) \wedge P_1(a_2) \wedge P_2(a_1) \wedge P_2(a_2)) \vee \\ & (\neg P_1(a_1) \wedge \neg P_1(a_2) \wedge P_2(a_1) \wedge P_2(a_2)) \vee (P_1(a_1) \wedge P_1(a_2) \wedge \neg P_2(a_1) \wedge \neg P_2(a_2)) \vee \\ & (\neg P_1(a_1) \wedge P_1(a_2) \wedge \neg P_2(a_1) \wedge \neg P_2(a_2)) \vee (P_1(a_1) \wedge \neg P_1(a_2) \wedge \neg P_2(a_1) \wedge \neg P_2(a_2)) \vee \\ & (P_1(a_1) \wedge P_1(a_2) \wedge \neg P_2(a_1) \wedge \neg P_2(a_2)) \end{aligned} \tag{6.7}$$

cuya cantidad de informacion es $\frac{7}{16}$. Reemplazar cada uno de estos disyuntos por su

informacion sobre P_1 resulta en

$$\begin{aligned} & (P_1(a_1) \wedge P_1(a_2) \wedge P_2(a_1) \wedge P_2(a_2)) \vee (\neg P_1(a_1) \wedge \neg P_1(a_2) \wedge P_2(a_1) \wedge P_2(a_2)) \vee \\ & (P_1(a_1) \wedge \neg P_1(a_2) \wedge P_2(a_1)) \vee (\neg P_1(a_1) \wedge P_1(a_2) \wedge P_2(a_1)) \vee \\ & (\neg P_1(a_1) \wedge \neg P_1(a_2) \wedge P_2(a_2)) \vee (P_1(a_1) \wedge \neg P_1(a_2) \wedge P_2(a_2)) \vee \\ & (\neg P_1(a_1) \wedge P_1(a_2) \wedge P_2(a_2)) \vee (P_1(a_1) \wedge \neg P_1(a_2) \wedge P_2(a_2)) \vee \\ & (\neg P_1(a_1) \wedge \neg P_1(a_2)) \end{aligned}$$

que se puede abreviar como

$$\begin{aligned} & (P_1(a_1) \wedge P_1(a_2) \wedge P_2(a_1) \wedge P_2(a_2)) \vee (\neg P_1(a_1) \wedge \neg P_1(a_2) \wedge P_2(a_1) \wedge P_2(a_2)) \vee \\ & (P_1(a_1) \wedge \neg P_1(a_2) \wedge P_2(a_1)) \vee (\neg P_1(a_1) \wedge P_1(a_2) \wedge P_2(a_1)) \end{aligned}$$

cuya cantidad de informacion tambien es $\frac{7}{16}$. Por otro lado, $\exists x(P_1(x) \wedge P_2(x))$ no es acerca de P_2 , porque reemplazar los disyuntos de (6.7) por su informacion sobre P_2 da

$$\begin{aligned} & (P_1(a_1) \wedge P_1(a_2) \wedge P_2(a_1) \wedge P_2(a_2)) \vee (\neg P_1(a_1) \wedge \neg P_1(a_2) \wedge P_2(a_1) \wedge P_2(a_2)) \vee \\ & (P_1(a_1) \wedge \neg P_1(a_2) \wedge P_2(a_1) \wedge P_2(a_2)) \vee (\neg P_1(a_1) \wedge P_1(a_2) \wedge P_2(a_1) \wedge P_2(a_2)) \vee \\ & (P_1(a_2) \wedge \neg P_2(a_1) \wedge P_2(a_2)) \vee (\neg P_1(a_1) \wedge P_2(a_1) \wedge \neg P_2(a_2)) \vee \\ & (\neg P_1(a_2) \wedge P_2(a_1) \wedge P_2(a_2)) \vee (\neg P_1(a_1) \wedge P_2(a_1) \wedge \neg P_2(a_2)) \vee \\ & (\neg P_2(a_1) \wedge \neg P_2(a_2)) \end{aligned}$$

cuya cantidad de informacion es 0.

Estos resultados pueden extenderse a cualesquiera oraciones de L de la forma $\exists v('v) \wedge ('v)$, donde ' y ' no son necesariamente letras de predicado. Ahora bien,

$\exists x(P_1(x) \leftrightarrow P_2(x))$ no solo es acerca de P_1 sino también de P_2 . Esta es una consecuencia inevitable de la condición de equivalencia.

Proposición 6.8 (Putnam). Si dos oraciones ' ϕ ' y ' ψ ' de L son lógicamente equivalentes, tienen el mismo contenido lógico.

Luego, si bien la noción de contenido lógico de Putnam depende en alguna medida de las notas particulares de L , es un concepto extensional. Otra consecuencia de la definición 6.7 es que las verdades lógicas, al dar 0 información, son acerca de todo aunque no digan nada, mientras que el contenido lógico no está siquiera definido para contradicciones. Y de esto se sigue que la mención de una clase en una oración no es condición suficiente para que esta oración sea acerca de esa clase, a diferencia del concepto de referencia de Ryle.

Esta definición de contenido lógico, también a diferencia de la de Ryle, se aplica a todas las oraciones de L , aunque L no puede ser cualquier lenguaje de primer orden; la noción de contenido de Putnam tiene un alcance limitado. En primer lugar, tal como está formulada, solo se aplica a lenguajes con un número finito de términos cerrados y letras de predicado, esto es, no sirve como criterio de contenido de oraciones de L_{pa} o sus extensiones. Segundo, el contenido lógico solo está definido para predicados/clases monádicas, aunque la definición 6.7 sugiere un modo natural de extenderse a predicados de cualquier aridad, trabajando con clases de pares ordenados en lugar de objetos sueltos.

Además, si bien se comporta perfectamente con respecto a oraciones universales seguidas por condicionales, la noción de Putnam no es del todo adecuada para lidiar con el contenido de otro tipo de oraciones y, por lo tanto, no puede dar lugar a un concepto satisfactorio de autorreferencia. Supongase que se trabaja en un sublenguaje de L_T dado por dos símbolos de predicado Tv y $At(v)$ (aparte de $=$, que es un símbolo lógico que no aporta información) para la verdad y el conjunto de oraciones atómicas del lenguaje y tres términos, I , $p_0 = 0q$ y $p_T p_0 = 0qq$, que denotan $\langle T \rangle$, $0 = 0$ y $T p_0 = 0q$, respectivamente. De acuerdo con [Putnam \(1958, p. 128\)](#),

If we adopt Quine's suggestion that any individual be identified with its unit class, then we also have the result that in the case of singular sentences, $\langle a \text{ is } B \text{ or } B(a) \rangle$, [the traditional subject term is the strict subject [contenido lógico] of the sentence.

Efectivamente, oraciones como $\langle T \rangle$ son acerca de $x = I$, porque se pueden reescribir equivalentemente como $\exists x(x = I \leftrightarrow Tx)$. Identificando la clase $x = I$ ($f: T \rangle$) con $\langle T \rangle$ (la denotación de I), la oración es acerca de $\langle T \rangle$, de sí misma. Pero también es

acerca de (la extensión de) T y, si se extiende la sugerencia de Quine para clases unitarias a todas las clases, TI es acerca de todas las oraciones que caen bajo la extensión de T , sean estas cuales fueren. Si $T \text{ p}0 = 0q$ está en la extensión de T , como es acerca de T , también debería decirse que es acerca de sí misma. En general, todas las oraciones de la forma $T \text{ t}$ que pertenezcan a la extensión de T son acerca de sí mismas.

Por otro lado, $\text{TI} \wedge \text{At}(\text{p}0 = 0q)$ no es acerca de TI de acuerdo con la definición 6.7, porque da más información que TI . En general, una conjunción no es necesariamente acerca de las clases sobre las cuales son los conjuntos. Esto puede resultar peligroso, porque, v.g., $\text{TI} \wedge \text{At}(\text{p}0 = 0q)$ en pat es equivalente a TI , pero el enfoque de Putnam no puede detectar ningún patrón de referencia sospechoso ahí. La autorreferencia de TI queda invisibilizada por la conjunción. Una solución posible consiste en modificar la definición de contenido lógico de modo tal que una oración sea acerca de una clase en la medida en que de una cantidad de información mayor a 0 sobre esa clase. Lamentablemente, este cambio no solucionará a los otros problemas que padece la noción de Putnam.

El contenido lógico de Putnam subgenera y sobregenera casos de contenido lógico y, en particular, para oraciones que son acerca de sí mismas o que tienen un patrón de referencia circular. No es una noción materialmente adecuada y, por tanto, no puede utilizarse para estudiar las paradojas y otros fenómenos semejantes.

6.3.3. Contenido lógico según Goodman

[Goodman \(1961\)](#) comparte con Putnam la intuición de que la condición de equivalencia debe satisfacerse, pero rechaza la primera condición de Leitgeb: que oraciones como "Todas las montañas tienen la cumbre nevada" sea acerca de las montañas y no de las cosas que tengan la cumbre nevada. En sus propios términos,

By our definition of absolute aboutness, the statement "Crows are black" is about black things as well as about crows; and this seems to me quite as it should be. Ryle and Putnam treat [that statement] as about crows but not about black things; and no doubt it is often thought of as 'telling us something about' the one but not the other. But whether we consider it on a given occasion as telling us about crows or black things may depend upon whether the statement occurs as an answer to a question like "What is the colour of crows?" or to a question like "What are some black things?" ([Goodman, 1961](#), pp. 7-8)

Notese que Goodman está suponiendo, aristotelicamente, que en todo enunciado uni-versal seguido por un condicional, como "Los cuervos son negros", la fórmula que ocupa el lugar del antecedente es satisfecha siempre por al menos un objeto, lo cual no es necesariamente el caso en la lógica actual.

Goodman propone dos nociones de contenido lógico. Una de ellas, el contenido lógico absoluto, satisface la condición de equivalencia desatendiendo un poco la forma de las oraciones, mientras que la otra, el contenido lógico inmediato, atiende más a la sintaxis pero la condición de equivalencia falla. Esta última parece ser más bien un concepto de referencia. Ambas relaciones se dan entre oraciones, por un lado, y objetos o clases por otro, esto es, tanto los objetos como las clases de las cuales habla una expresión pueden formar parte de su contenido lógico. Esta diversidad de contenido es razonable en tanto Goodman trabaja informalmente sobre lenguajes de segundo orden, donde las clases son un segundo tipo de objeto sobre los cuales se cuantifica. Dado que esta investigación se limita a lenguajes de primer orden, se podrá ignorar lo referente a clases en las definiciones que siguen, aunque no es la intención original del autor.

Sea L un lenguaje de segundo orden cualquiera (vease el apartado [1.2.5](#)). La lógica de segundo orden expande la de primer orden con principios para los nuevos cuantificadores. La regla de eliminación es análoga a la de los cuantificadores de primer orden, mientras que su introducción está dada por el esquema de axioma de comprensión (cf. [Shapiro \(1991\)](#)).

Definición 6.9 (Contenido lógico absoluto). Una oración ϕ de L es absolutamente acerca de un objeto a o un conjunto C si hay un término o fórmula ψ tal que ϕ implica lógicamente ψ de L , a ocurre en ψ , y denota a o C , correspondientemente, pero ϕ no implica ninguna generalización sobre ninguna parte de ψ . ¹¹

Al estar dada en términos de implicación lógica, la noción de contenido lógico absoluto está cerrada bajo la condición de equivalencia, es una noción plenamente extensional. Al igual que para la noción de Putnam, no es suficiente que una oración mencione un objeto o una clase para que sea acerca de ella, como evidencian las verdades lógicas y las contradicciones. Las verdades lógicas no son acerca de ningún objeto y, prácticamente, de ninguna clase, mientras que las contradicciones no son absolutamente acerca de nada. Por ejemplo $\exists x(Px \rightarrow Px)$ no es acerca de ningún objeto porque las verdades lógicas no implican verdades sobre objetos particulares que no impliquen a su

¹¹*Strictamente hablando, aunque usa el término, Goodman no trabaja con una noción de consecuencia lógica sino con un concepto de implicación analítica según el cual, por ejemplo, la oración del mentiroso es autocontradictoria. Esto aleja mucho su noción de los propósitos de esta investigación, como señala [Urbanjak \(2009\)](#), porque si el mentiroso es contradictoria, como se ve más adelante, no es absolutamente acerca de nada y, por ende, jamás puede resultar autorreferencial.*

Consecuentemente, modifícame un poco las definiciones de Goodman entendiendo el término 'implicación lógica' literalmente.

vez sobre todos los otros objetos, pero tampoco es acerca de la clase de nida por P , porque implica $\exists x \exists x (Xx \rightarrow \neg Xx)$.

Que una oración sea acerca de una clase no implica que también sea acerca de los objetos que le pertenecen. V.g., si $\exists x (x)$ es una contingencia, es acerca de la clase de nida por la fórmula ' \exists ', sea lo que fuere, pero no necesariamente acerca de un objeto particular. De acuerdo con Goodman, la noción de contenido guarda cierta analogía con la idea de elegir: del mismo modo en que no se puede elegir todo, porque eso no sería ya elegir, no se puede ser acerca de todos los objetos, o de todas las clases. Tampoco es $\exists x (\neg (x))$ acerca de los ' \exists ' sino acerca de las clases que las fórmulas ' \exists ' y de \neg .

El requisito de que ' \exists ' no implique ninguna generalización sobre ninguna parte de \exists en la definición 6.9 sirve para evitar la sobregeneración del criterio. Toda fórmula ' \exists ' L implica lógicamente $\exists a (a)$ para cualquier L y cualquier a , sea que denote un objeto o un conjunto. Como ' \exists ' también implica $\exists x (\neg (x))$, la generalización de ' \exists ' (a) sobre a , el criterio no permite inferir que ' \exists ' es acerca de lo que a denota del simple hecho de que ' \exists ' implica ' \exists ' (a) .

Pero Goodman no solo solicita que (a) no sea generalizable sobre a sino sobre ninguna de sus componentes, porque quiere evitar que una oración que es explícitamente acerca de una clase lo sea también de todas las clases que la contienen. Tómese como ejemplo una oración contingente cualquiera ' (a) '. ' (a) ' es acerca de la clase de los ' \exists '. $\exists x ((\neg (x)) \rightarrow (a))$ se sigue lógicamente de ' (a) ' y, a la vez, no es generalizable sobre a . No obstante, la definición de Goodman no permite inferir que entonces ' (a) ' es acerca de la clase de los ' \exists 's, porque, v.g., $\exists x \exists x ((\neg (x)) \rightarrow Xx) \rightarrow Xa$ es una generalización de $\exists x ((\neg (x)) \rightarrow (a))$ que también se sigue lógicamente de ' (a) '.

Sin embargo, tomando una a particular, se puede exhibir una fórmula en la cual ocurre cuya generalización sobre a no se sigue lógicamente de ' (a) '. Sea b una constante de individuo que denota un objeto que no pertenece a la clase de nida por ' \exists ' y sea

$(x) := (\neg (x)) \rightarrow x = b$. Luego, (a) se sigue lógicamente de ' (a) ', mientras que $\exists x Xx$ no.

Si ' (a) ' fuera acerca de la clase de los ' \exists 's en la medida en que sea posible dar una fórmula que contenga ' \exists ' y que se siga lógicamente de ' (a) ' pero que no sea generalizable sobre ' \exists ', debería concluirse que toda fórmula acerca de la clase de los ' \exists 's es asimismo acerca

de todas las clases que los contienen. Como señala Goodman, la definición 6.9 bloquea este tipo de casos porque, si bien $\exists x Xx$ no se sigue de ' (a) ', existe otra generalización sobre 'alguna parte' de (a) que sí se sigue, i.e., $\exists x (\exists x ((\neg (x)) \rightarrow Xx) \rightarrow Xa)$. Goodman entiende entonces por 'generalización sobre alguna parte de ' \exists ' no solo la cuantificación sobre alguna de sus componentes sino la relativización de estos cuantificadores a otras expresiones.

Lamentablemente, este último requisito es demasiado restrictivo. Si $'(a)$ no es

acerca de la clase de nada por $'(x) _ x = b$ porque $\exists x(\exists x('x) ! Xx) ! Xa$ es una generalización sobre alguna parte de $'(a) _ a = b$ y es consecuencia lógica de $'(a)$, mediante un razonamiento análogo puede concluirse que $'(a)$ no es tampoco acerca de la clase de los 's. Porque $\exists x(\exists x('x) ! Xx) ! Xa$ también es una generalización sobre alguna parte de $'(a)$, sobre la relativización de $'(a)$ a $\exists x('x) ! 'x)$, i.e., $\exists x('x) ! 'x) ! '(a)$, que es lógicamente equivalente a $'(a)$. Pero, al menos a la luz del contraejemplo dado, no veo necesidad de adoptar esta versión amplia de 'generalización sobre alguna parte de '. Si $(x) := '(x) _ x = b$, (a) es una instancia de $\exists x('a) _ Xa$ que, no solo se sigue lógicamente de $'(a)$ sino que también es una generalización sobre una parte de .

Aun dejando de lado esta ampliación inconveniente de la idea de generalización, la noción de contenido lógico absoluto de Goodman tiene ciertas consecuencias contrain-tuitivas. Al igual que el contenido lógico de Putnam y acarreado los problemas que esto genera, si una oración ' de L es absolutamente acerca de un objeto x, su conjunción con otra oración no es necesariamente absolutamente acerca de x. Por ejemplo, si $'(a)$ es acerca del objeto denotado por a, $'(a) \wedge \exists x[x=a]$ ya no es acerca de ese objeto, porque puede generalizarse sobre a.

Por otro lado, el contenido lógico absoluto de Goodman, aun con las restricciones sugeridas, parece sobregenerar casos de referencia a clases. $\exists x('x) ! (x)$ es acerca de las clases de los 's y de los s, pero por la condición de equivalencia también acerca de las clases de los : s y los : 's. Aun más, toda oración que es absolutamente acerca de algo, ya sea un objeto o una clase, es acerca de la clase universal (en un modelo, el dominio), porque implica lógicamente $\exists x x = x$ pero no $\exists x \exists x Xx$. De acuerdo con Goodman estas son consecuencias deseables de su definición, pero para esta investigación son perjudiciales. Oraciones 'mentirosas' como $\exists x('x) ! :Tx$ donde $'(x)$ se aplica exclusivamente a $\exists x('x) ! :Tx$, si bien son acerca de la clase de los 's, o sea, de $\exists x('x) ! :Tx$, no son acerca de $\exists x('x) ! :Tx$.

Aca no hay lugar para la sugerencia de Quine de identificar objetos con sus clases unitarias o, más generalmente, de entender que una oración es acerca de ciertos objetos cuando en realidad es acerca del conjunto que los tiene por miembros; porque en ese caso la noción de Goodman se trivializaría: como toda oración que es acerca de algo es acerca de la clase universal, toda oración sería absolutamente acerca de todos los objetos o acerca de nada. Tampoco tiene sentido realizar esta identificación con clases unitarias exclusivamente, porque se pueden formular otras oraciones del mentiroso de la forma $\exists x('x) ! :Tx$ en las cuales, v.g., $'(x)$ es satisfecha por $\exists x('x) ! :Tx$ y por cualquier verdad lógica.

Estos casos de sub y sobregeneración de contenido lógico absoluto hacen que el concepto no sea adecuado ni sirva para formular una noción precisa y correcta de autorreferencia. Goodman da una segunda noción de contenido, que se aproxima más a un concepto de referencia.

Definición 6.10 (Contenido lógico inmediato). Una oración ϕ es inmediatamente acerca de un objeto o o un conjunto C si hay un término o fórmula A que ocurre en ϕ y denota o o C , según corresponda, y ϕ no implica lógicamente ninguna generalización sobre ninguna parte de ϕ .

Una oración es inmediatamente acerca de un objeto o clase solo si la menciona; el contenido inmediato depende de la sintaxis de la oración. A diferencia del absoluto, no satisface la condición de equivalencia: $\exists x(\phi(x) \leftrightarrow \psi(x))$ puede ser inmediatamente acerca de las clases de los ϕ 's y los ψ 's, pero no acerca de la de los ϕ : s o los ψ :s, porque estas no están mencionadas en la oración. El contenido lógico inmediato es una noción hiperintensional. Como consecuencia, las oraciones que son acerca de algo ya no son necesariamente acerca de la clase universal; por ejemplo, $\exists x x = 6$ es solo acerca del conjunto vacío, lo cual puede significar una ventaja a la luz de esta investigación.

El contenido lógico inmediato tiene el mismo problema que menciono acerca del contenido absoluto, i.e., que es demasiado restrictivo debido al requisito de 'ninguna generalización sobre ninguna parte', y puede arreglarse también no permitiendo que una generalización relativice sus cuantificadores. En ese caso, a pesar de que Goodman considera que su noción de contenido inmediato es inútil por no satisfacer la condición de equivalencia, el concepto sería mucho más fructífero, en tanto permitiría la identificación de las clases con sus objetos como contenidos, aunque no evitaría las dificultades que tiene en común con la noción de Putnam.

6.3.4. Contenido lógico según Urbaniak

[Urbaniak \(2009\)](#) es el primero en participar activamente del desafío de Leitgeb de dar una noción formalmente correcta y materialmente adecuada de autorreferencia para poder responder a la pregunta por la autorreferencialidad de la paradoja de Yablo y otras cuestiones semejantes. Busca una noción de contenido referencial que complete lo que llame 'noción intuitiva' en el apartado [6.1.2](#). Para poder satisfacer la condición de equivalencia, se centra en una noción de informatividad diferente de la de Putnam. El resultado es un concepto bastante cercano al contenido lógico absoluto de Goodman.

Tras considerar ejemplos como [\(6.4\)](#), [Urbaniak \(2009, p. 243\)](#) sostiene, contrario a Ryle, que

This however indicates that the bare fact that a constant that names an object occurs in a sentence is not sufficient for that sentence to be about that object. This motivates certain Wittgensteinian intuitions about what sentences are about ("Therefore the propositions of logic say nothing."). In a sense, to be about x , ϕ would have to say something about x , provide us with some contingent information about it.

Claramente, estas intuiciones wittgensteinianas se corresponden más con una noción de contenido que con una de referencia. Una oración puede hacer referencia a un objeto diciendo únicamente algo trivial o, si se quiere, 'nada', acerca de él.

>En qué casos una oración provee información contingente acerca de algo? Sea L un lenguaje de primer orden con 0 y más de un término y con 0 o más de una letra de predicado n -ádica no lógicamente, para cada n .

Definición 6.11 (Informatividad). Si ϕ es una oración de L , t un término cerrado y P una letra de predicado no lógicamente n -ádica, $t(P)$ ocurre informativamente en ϕ si ocurre en ella y existe un término cerrado $s(Q)$, de la misma aridad que P y diferente de $t(P)$ tal que ϕ no es lógicamente equivalente a $\phi[s=t]$ ($\phi[Q=P]$).

Esta noción se acerca bastante a la idea de Goodman de no permitir que los términos y predicados en una oración que es acerca de lo que estos denotan sean generalizables módulo equivalencia lógica, pero deja de lado las restricciones adicionales de Goodman que llevan a una noción demasiado estrecha. Sea M una interpretación de L .

Definición 6.12 (Contenido lógico informativo). Una oración ϕ de L es informativa-mente acerca de un objeto o relativamente a M si y sólo si

1. hay un término t que ocurre informativamente en ϕ tal que $t^M = o$, o
2. hay un predicado no lógicamente n -ádico P que ocurre informativamente en ϕ tal que o es un componente de alguna n -tupla en P^M .

Este concepto se comporta adecuadamente con respecto a los casos intuitivos marcados por Leitgeb. Una oración de la forma $\phi(a)$ en la cual a ocurre informativamente es acerca de lo denotado por a . Por ejemplo, $\neg \exists x (x \neq x)$ en (6.2) es acerca del mentiroso. Y oraciones de la forma $\exists v_1 \dots \exists v_n (P_1(v_1; \dots; v_n) \wedge \neg P_2(v_1; \dots; v_2))$ son acerca de los objetos pertenecientes a la extensión de P_1 en M . Además, el contenido lógico informativo está definido para todas las oraciones de L y satisface la condición de equivalencia.

Proposición 6.13 (Urbaniak). Si ϕ y ψ son oraciones lógicamente equivalentes de L , ϕ es acerca de o relativamente a M si y sólo si ψ es acerca de o relativamente a M .

El contenido lógico informativo es de carácter extensional pero depende en alguna medida de la sintaxis de las oraciones, aunque no de su estructura lógica, porque solo atiende a los símbolos no lógicos que ocurren en ellas. A partir de la definición de contenido lógico informativo Urbaniak da un concepto de autorreferencia.

Definición 6.14 (Autorreferencia informativa). Una oración $\phi \in L$ es informativamente autorreferente relativamente a M si el par ordenado $\langle \phi, i \rangle$ pertenece a la clausura transitiva del contenido lógico informativo relativo a M .

Como Urbaniak mismo admite, el criterio no es lo suficientemente selectivo, so-bregenera casos de contenido. Por un lado, $\langle \phi, i \rangle$ es no solo acerca de ϕ (lo que i denota) sino de los objetos en la extensión de ϕ . Dado que ϕ ocurre informativamente en cualquier oración de la forma $\phi(t)$, al igual que en el caso de Goodman, todas ellas son acerca de los objetos en la extensión de ϕ y, por ende, si $\phi(t) \in L^M$, $\phi(t)$ es acerca de sí misma, sin importar lo que t denote: $\langle \phi, i \rangle = \langle \phi, i \rangle$ es autorreferente bajo las mejores interpretaciones de L .

Por otro lado, como para Goodman, $\exists v_1 \dots \exists v_n (P_1(v_1; \dots; v_n) \wedge P_2(v_1; \dots; v_2))$ no solo es acerca de los miembros de P_1^M sino también de los de P_2^M . Pero oraciones de la forma $\exists v_1 \dots \exists v_n (\phi)$, donde ϕ y ψ son funciones proposicionales contingentes que no son símbolos de predicado, no son necesariamente acerca de las clases que estas fórmulas de nenen en M . Por ejemplo, si se trabaja en L_{pa}^+ con N^+ ,

$$\exists x (\text{Bew}_{pa^+}(x) \wedge \text{Bew}_{pa^+}(\neg x))$$

no es acerca de (los códigos de) las oraciones de L_{pa}^+ que no son teoremas de pa^+ ni de los teoremas que son negaciones. Ni siquiera es acerca de los teoremas de pa^+ , porque $\text{Bew}_{pa^+}(x)$ y $\text{Bew}_{pa^+}(\neg x)$ son fórmulas complejas, sino de todos los objetos mencionados en estas fórmulas y todos aquellos que caen bajo la extensión de algún predicado primitivo no lógico que ocurra en ella. Como L_{pa}^+ no contiene predicados no lógicos, el contenido lógico informativo de esta oración se aleja bastante de lo esperado.

Además, nuevamente siguiendo a Goodman, las oraciones $\exists x Px$, $\exists x \neg Px$ y $\exists x Px \wedge \exists x \neg Px$, donde P es un símbolo de predicado no lógico, son todas acerca de los elementos de P^M .

A favor de Urbaniak, y eso pone su noción en una mejor posición con respecto a las de Putnam y Goodman, si una oración ϕ es acerca de x y su conjunción con otra oración ψ no da lugar a una verdad lógica o a una contradicción, $\langle \phi, i \rangle$ es también acerca de x . Esto evita problemas como el señalado en el antepenúltimo párrafo del apartado 6.3.2. Pero si $\langle \phi, i \rangle = \langle \psi, j \rangle$, al ser la conjunción una contradicción, no será acerca de nada. La

¹²O a través de una codificación, i.e., el par ordenado $h(i); \#(i)$.

diferencia con Goodman esta dada por el requisito adicional que impone Urbaniak sobre la condición de no informatividad: para que a ocurra informativamente en una oración de la forma $P a$ es suficiente que $P a$ no sea lógicamente equivalente a toda expresión de la forma $P b$, donde b es un término, no es necesario que no implique todas esas expresiones.

Urbaniak reformula la paradoja de Yablo con una serie de predicados primitivos $P_1; \dots; P_n; \dots$ para los antecedentes de los condicionales, donde P_n se aplica a todas y solo aquellas oraciones que se encuentran por debajo de la n -ésima en la lista, y otra serie de predicados primitivos U_n para los consecuentes, cada uno de los cuales se comporta como un predicado de no verdad solo para las oraciones de índice mayor a n . Las oraciones de Yablo ahora son de la forma

$$\neg \exists x (P_n(x) \wedge U_n(x))$$

para cada $n \geq 1$. Bajo esta reformulación (y no bajo la que se vio en [6.1.1](#)) es posible responder a la pregunta por la autorreferencia de (esta versión de) la lista de Yablo utilizando la noción de contenido informativo, negativamente. Con respecto a las oraciones de Henkin, el criterio carece de utilidad, al igual que para evaluar oraciones del mentiroso y otras expresiones paradójicas en lenguajes aritméticos.

Finalmente, Urbaniak responde al tercer argumento de Leitgeb para la condición de equivalencia, aquel en el cual Leitgeb sostiene que la referencia o el contenido no deben ser preservados únicamente bajo equivalencia lógica sino únicamente bajo equivalencia en la aritmética.

What should we make of Leitgeb's complaint that we won't be able to say that the diagonalization lemma allows us to construct self-referential sentences? Well, the answer is simply that on this notion of aboutness, the diagonal construction doesn't give us a self-referential sentence, and that's it. This, however, doesn't mean that we cannot use the phrase 'self-referential' in a weaker sense, which allows us to call a formula self-referential if it is equivalent (modulo syntactic encoding and Peano arithmetic) to a formula in which its code number occurs. ([Urbaniak, 2009](#), p. 284)

De acuerdo con el mismo Urbaniak, su noción de autorreferencia no es capaz de explicar porque el lema diagonal débil (teorema [1.7](#)) da lugar a oraciones autorreferenciales, uno de los desiderata fundamentales. Si bien es probablemente en lo cierto, no es por las razones que él cree. Ciertamente, una equivalencia no lógica sino aritmética no es suficiente, de acuerdo con la noción de autorreferencia informativa, para establecer autorreferencia. Pero hay más que eso detrás de los puntos débiles y sus equivalencias.

Si se atiende a la prueba del teorema [1.7](#) para una teoría \mathcal{T} se ve que las oraciones tales que $\mathcal{T} \vdash \exists v (\varphi(v))$ son de la forma

$$\exists v (\text{Diag}(\exists v (\text{Diag}(u; v) \wedge \varphi(v))) \wedge \varphi(v)):$$

Si Diag fuera un primitivo del lenguaje, estas últimas expresiones serían acerca de las componentes de algún par ordenado que pertenezca a Diag^N y, por ende, acerca de : ser informativamente autorreferencial, aunque también serían acerca de todo, porque Diag representa una función, toda expresión de L_{pa} es diagonalizable. Sin embargo, como Diag no es un primitivo, el contenido informativo de depende de la fórmula particular que se utilice para representar la función de diagonalización; no es claro que sea de hecho informativamente autorreferencial.

Urbaniak sugiere que aun es posible llamar autorreferentes a expresiones como $\exists v (\varphi(v))$ (donde $\varphi : \mathcal{T} \rightarrow \mathcal{T}$) en un sentido más débil de acuerdo con el cual una oración es autorreferencial si es aritmeticamente equivalente a otra en la cual 'su código ocurre'. Pero como se vio en [6.1.2](#), este sentido 'más débil' es trivial y es, por ende, justamente lo que había que precisar, regimentar, transformar en una noción formalmente correcta y materialmente adecuada que pueda con legitimidad formar parte del discurso científico. El sentido en el cual es autorreferente mientras que $0 = 0$ no lo es es el sentido de 'autorreferencia' que importa, sobre el cual debe trabajarse, porque es de carácter intuitivo todavía, permanece en las penumbras de la precientificidad. El concepto de contenido informativo de Urbaniak, tal como está formulado, no es útil para el estudio de las paradojas semánticas.

6.4. La condición de equivalencia

Ninguna de las nociones que he introducido hasta ahora resulta satisfactoria a la luz de las condiciones de Leitgeb. Por un lado, las de Putnam, Goodman y Urbaniak relajan demasiado la primera condición, no respetan las intuiciones básicas de la llamada 'noción intuitiva'. Su estrategia consiste en agregar requisitos adicionales a la mera mención de un objeto o cuantificación sobre una clase para extraer conclusiones con respecto al contenido de una oración. Los tres autores coinciden en que el requisito adicional debe tener que ver con la no trivialidad, la contingencia, la informatividad, tanto de las ocurrencias de los términos como de los predicados o fórmulas, y dejan de lado completamente la forma lógica de las oraciones. Como los símbolos de predicado son tratados del mismo modo que los nombres de individuo, independientemente de la posición

que ocupen en una oración, las consecuencias son contraintuitivas. Por ejemplo, oraciones de la forma

P a son acerca, no solo de lo denotado por a, sino de los P s lo cual, como se vio, no es deseable.

Por otro lado, si no se refuerza el requisito de la mención o cuantificación para establecer referencia o contenido, si, como en el caso de Ryle, se atiende plenamente las intuiciones de la primera condición de Leitgeb, la condición de equivalencia es insatisfacible, so pena de trivialidad.

No es claro entonces que una noción de contenido referencial que satisfaga las tres condiciones de Leitgeb sea lógicamente posible. Como ya sugerí, estas pretenden fusionar dos conceptos: la referencia y el contenido lógico, los cuales son diferentes y podrán no ser fusionables. Un intento de hacerlo consistiría en agregar un tercer requisito a la ocurrencia y la informatividad, una condición que atienda a la forma lógica de las oraciones. Por ejemplo, que las oraciones atómicas solo sean acerca de los objetos que mencionan informativamente, que oraciones como $\exists v('v) \neg (v))$ sean acerca de los 's y los no s en la medida en que ' (o ' y) ocurra informativamente, etc. Un abordaje de este tipo no ha sido aun explorado y tal vez valga la pena hacerlo. Sería preciso proceder con cuidado para no caer en inconsistencias, armonizando la condición de equivalencia con las cláusulas que indican el contenido de cada oración según su forma y, por otra parte, las cláusulas composicionales, sin las cuales la definición sería imposible. En cualquier caso, algunas intuiciones deberán dejarse de lado: algunas oraciones atómicas que mencionen ciertos objetos no serán acerca de ellos o serán acerca de todo y algunas oraciones que se obtienen mediante un mecanismo de diagonalización no resultarán autorreferenciales. Para verlo, tomese como ejemplo la oración que resulta de diagonalizar fuertemente $x = x$, i.e., hay un término de L^+_{pa} tal que $q^+ \neg t = pt = tq$. $t = t$ es una verdad lógica y, como tal, o bien no es acerca de nada o bien es acerca de todo, pero en ningún caso será exclusivamente acerca de t y, por ende, nunca va a ser propiamente autorreferencial.

De cualquier modo, mucho menos queda claro que la condición de equivalencia sea una condición necesaria de toda definición de contenido referencial cuyo fin sea colaborar en el debate alrededor de la paradoja de Yablo y sirva para estudiar los patrones referenciales de las paradojas y otras oraciones en el lenguaje de la aritmética. A continuación examino los argumentos que Leitgeb da a su favor.

En primer lugar, Leitgeb aduce que dos oraciones que son lógicamente equivalentes son indistinguibles desde un punto de vista lógico y, por ende, extensionalmente equivalentes en todo mundo posible. Como señala [Urbaniak \(2009\)](#), esto es cierto, pero también es cierto que cuando pasa de utilizarse las fórmulas a hablar acerca de ellas estas se tornan distinguibles. De otro modo no podría decirse, justamente, que hay oraciones

diferentes que son lógicamente equivalentes, que dos oraciones diferentes pero lógicamente equivalentes son indistinguibles desde un punto de vista lógico, etc. En otras palabras, ninguna noción hiperintensional har a sentido. El punto de vista que se toma cuando se habla del contenido referencial, al estar hablando de fórmulas como objetos, ya no es un punto de vista lógico sino, si se quiere, lógico-lingüístico, se atiende a la sintaxis.

Una respuesta en la misma línea sirve para refutar el segundo argumento de Leitgeb: que el contenido referencial es una noción semántica porque está definido a partir de la referencia o denotación de construcciones nominales a objetos, que es obviamente semántica. El contenido referencial de oraciones está definido, efectivamente, en términos de la denotación de construcciones nominales, pero no exclusivamente en estos términos. La sintaxis de la oración o del lenguaje al cual esta pertenece también juega un rol: la presencia o ausencia de una constante, la presencia o ausencia de una letra de predicado, es crucial para determinar la referencia o el contenido de una expresión. Luego, nuevamente, el contenido referencial no puede ser una noción plenamente semántica, tiene elementos de la sintaxis.

Leitgeb da un tercer argumento a favor de la condición de equivalencia o, mejor dicho, a favor de la condición más fuerte de equivalencia aritmética: que dos oraciones aritmeticamente equivalentes deben referir a lo mismo. Desde su punto de vista, este es el único modo de garantizar que las equivalencias que se pueden probar a partir del lema diagonal débil sean evidencia suficiente para la autorreferencia. Es aritmética pero no lógicamente equivalente a $\vdash p \leftrightarrow q$; para concluir que es autorreferente a partir de esta equivalencia debe tenerse que, como $\vdash p \leftrightarrow q$ refiere claramente a p y es aritmeticamente equivalente a p , también refiere a p .

Pero si la condición de equivalencia aritmética es el único modo de permitir que el lema diagonal débil sea un mecanismo de autorreferencia no hay esperanzas ya para una noción materialmente adecuada de contenido referencial. Porque si dos oraciones aritmeticamente equivalentes refieren a lo mismo, entonces, como se vio al comienzo en la sección 6.2, todos los teoremas de pa^+ refieren a lo mismo; el concepto se vuelve prácticamente trivial antes de considerar ningún otro principio para la referencia, como los que conforman las nociones intuitiva y técnica del apartado 6.1.2. Todos los teoremas de la aritmética son o bien oraciones de Henkin. Y, aun más, como son todos aritmeticamente equivalentes a una verdad lógica, la condición de equivalencia aritmética también trivializa las nociones de contenido lógico presentadas en los apartados anteriores.

Afortunadamente, hay mejores razones para concluir que una oración que ha sido obtenida por medio del lema diagonal débil es autorreferente, i.e., que esta dada por un cuantificador universal seguido por una expresión condicional cuyo antecedente es satisfecho exclusivamente por ella misma, esto es, estas oraciones autorrefieren porque

cuanti can sobre s mismas. La intuición según la cual expresiones de la forma $\exists v(\forall v)!$ (v)) son acerca de los 's' es suficiente para concluir que el lema diagonal débil garantiza autorreferencia. No parece haber razones a favor más que, tal vez, la elegancia, para adoptar un concepto de contenido referencial que satisfaga la condición de equivalencia.

[Leitgeb \(2002\)](#) concluye su argumentación con una advertencia: si la condición de equivalencia no vale, la referencia o contenido de una oración depende no solo de lo que dice sino de cómo lo dice. En efecto, si se quiere un concepto en el que la sintaxis y la estructura lógica de las oraciones jueguen un rol, un concepto más bien de referencia, el resultado es hiperintensional. Si lo que se busca es un concepto extensional, más bien de contenido lógico, la estructura lógica no puede tener demasiada injerencia y las intuiciones que Leitgeb quiere rescatar deben quedar a un lado. Pero estas intuiciones son esenciales al debate sobre la paradoja de Yablo, como Leitgeb mismo señaló, y también para la evaluación de la autorreferencialidad de oraciones de Henkin y Gödel, entre otras, mientras que la condición de equivalencia no parece serlo. Por esta razón, en lo que resta de esta tesis me centro en la primera vía, en un concepto de referencia hiperintensional.

De hecho, Leitgeb parece haber cambiado de opinión. En [\(2005, x5\)](#) considera y descarta inmediatamente la posibilidad de que su noción de dependencia (vease el apartado [3.2](#)), que satisface la condición de equivalencia, pueda hacer las veces de una noción de referencia o contenido lógico, por dos razones. Primero porque la dependencia de Leitgeb solo se ocupa, en última instancia, del contenido de oraciones 'a través' del predicado veritativo o, en palabras de Leitgeb, las relaciones de dependencia con frecuencia "[...] match our intuitions concerning aboutness with respect to truth-theoretic concerns." [Leitgeb \(2005, p. 175\)](#). Por ejemplo, mientras que la oración $\text{Bew}_{\text{pa}^*}(h)$ en [\(6.6\)](#) refiere intuitivamente a sí misma, depende esencialmente del conjunto vacío o, porque no contiene el predicado de verdad (cf. proposición [3.35](#)). Pero, en segundo lugar, y con más importancia, porque considera que ni siquiera cuando se trata del predicado veritativo la noción de dependencia respeta las intuiciones compartidas sobre la noción de contenido lógico o referencia:

Obviously, our pre-theoretic intuitions concerning aboutness are more fine-grained than any semantic notion of dependence such as ours can ever be. While our relation of dependency is closed under logical equivalence and even arithmetic equivalence [(cf. proposición [3.29](#))], such that logically equivalent sentences depend on the same sets of sentences, this is not the case according to the intuitions described above: $_$: may said to be about the semantic

properties of and not of ¹³ despite the fact that $_ :$ and $_ :$ are

logically equivalent; $2 + 2 = 4$ $_ T$ (c) is about the entity denoted by c, or so it seems, while the arithmetically equivalent $2 + 2 = 4$ is not, but both depend on the empty set. Contrary to our notion of dependency, the informal notion of aboutness is not just a matter of semantics, but also of syntax: what a sentence is about does not only depend on the proposition that the sentence expresses but also on how this proposition is expressed.

Sin ir tan lejos como Leitgeb, dir a que, en lo que al predicado de verdad concierne, la nocion de dependencia que ofrece puede o ciar de contenido logico, en la l nea de Putnam, Goodman y Urbaniak, pero ciertamente no de referencia, como el mismo indica.

§

A lo largo del cap tulo creo haber establecido que una nocion de contenido referencial que sirva para evaluar los patrones referenciales detras no solo de expresiones paradójicas del lenguaje de la verdad sino tambien de enunciados aritmeticos debe res-petar los casos intuitivos que conforman la nociones informales tanto 'intuitiva' como 'tecnica' de referencia y autorreferencia que se~nala Leitgeb |lo cual mostre que es po-sible por la estructura especial que poseen los enunciados que se obtienen a partir de un proceso de diagonalizacion| y, en consecuencia, no satisfacer la condicion de equi-valencia. Esto signi ca que lo que se precisa es una nocion de referencia mas que de contenido logico, un concepto hiperintensional.

Resta entonces encontrar un modo adecuado de completar la nocion intuitiva en detalle, que cubra absolutamente todos los casos, todas las expresiones posibles de los lenguajes de la aritmetica y de la verdad. Este es precisamente el objetivo del proximo cap tulo.

¹³ es la oracion que se obtiene diagonalizando debilmente el predicado $T x$, i.e.,

$\text{pat } _ \$ T p q$

es una oracion del honesto, que dice de s misma que es verdadera.

7

Referencia en el lenguaje de la aritmetica de primer orden y sus extensiones

"Wir haben also einen Satz vor uns, der seine eigene Unbeweisbarkeit behauptet."

{ Kurt Godel, "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I"

En este capítulo ofrezco una serie de nociones de referencia para extensiones del lenguaje de la aritmetica de primer orden que superan las dificultades y satisfacen las condiciones señaladas en el capítulo anterior: ¹ todas ellas respetan las características de la noción intuitiva de referencia introducida en el apartado [6.1.2](#) y son, por tanto, hiperintensionales, no satisfacen la condición de equivalencia de Leitgeb.

Siendo los objetos primarios del lenguaje de la aritmetica los números naturales, la sección [7.1](#) se ocupa de la referencia de oraciones de este lenguaje a números, y sus propiedades. En [7.2](#) de no referencia de oraciones de este lenguaje a otras oraciones del lenguaje mismo, a través de sus códigos de Godel. Muestro que el concepto de nido no solo es una compleción de la noción intuitiva sino que a la vez rescata la intuición detrás de la noción técnica (que solo tiene sentido cuando la referencia es a otras fórmulas). Como consecuencia, da el veredicto correcto con respecto a oraciones de Godel, Henkin y otros fenómenos de esta índole. Ambas nociones, de referencia a números y a oraciones,

¹Seguramente sea posible modificar levemente los conceptos que introduzco a continuación para elaborar nociones análogas de referencia de oraciones pertenecientes a otros lenguajes formales, como el lenguaje de la teoría de conjuntos, pero esto queda pendiente para un trabajo futuro.

son semánticas, de nidas en función de la noción de verdad en el modelo estándar de la aritmética.

Finalmente, en [7.3](#) me ocupo de la referencia de oraciones a oraciones del lenguaje de la aritmética con un predicado verdad únicamente a través de este predicado de verdad, que llamo 'referencia aletica'.² Contrario a lo que se espera, la referencia aletica no puede obtenerse simplemente generalizando la noción de referencia de la sección anterior a todo el lenguaje de la verdad, por razones que ahí indico. Ofrezco dos nociones de referencia aletica, una semántica y otra sintáctica, ya no en función de la verdad en un modelo sino del predicado de prueba de una teoría, y muestro que ambas capturan la noción técnica, esta vez restringida al predicado de verdad, y dan veredictos intuitivamente correctos sobre las expresiones paradójicas y otras patologías conocidas. La paradoja de Yablo y sus formulaciones alternativas resultan no autorreferentes y, por ende, queda refutada la versión ortodoxa de la tesis referencialista acerca de las paradojas.

Dado el objetivo general de lo que resta de la tesis, i.e., dar teorías de la verdad que cuadren con el de acionismo, y más específicamente, hacerlo restringiendo el [Esquema-T](#) a instancias generadas por expresiones que tengan patrones de referencia 'seguros', lo cual har de los sistemas resultantes buenos candidatos para el proyecto minimalista también (veanse los apartados [5.3](#) y [5.3.1](#)), me inclino por la referencia aletica sintáctica por sobre la semántica. En [7.3.1](#), tras de nimir referencia aletica semántica y otros conceptos cercanos, muestro que esta noción es indénible en la aritmética y, en consecuencia, los criterios de restricción de instancias del [Esquema-T](#) que en base a ella se formulen no serán meramente sintácticos, esto es, aritméticos; no podrán ser empleados directamente en la construcción de teorías desentrecorilladoras de acionistas (vease el apartado [5.3](#)).

En cambio, los conceptos de referencia aletica sintáctica que introduzco en [7.3.2](#) son denibles en el lenguaje de la aritmética de primer orden, como ahí indico. En el capítulo siguiente empleo esta noción sintáctica en la formulación de criterios de selección de instancias correctas y consistentes del [Esquema-T](#). El resultado es una serie de teorías desentrecorilladoras de la verdad sumamente abarcativas que, por ende, con rman la adecuación de la visión referencialista general sobre las paradojas semánticas: estas en efecto comparten ciertos patrones de referencia 'peligrosos'. Como se verá entonces, estas teorías satisfacen los criterios de adecuación tanto del de acionismo como del proyecto minimalista de Horwich.

²*El adjetivo 'aletica' esta simplemente para indicar que la noción es relativa a la verdad, al predicado veritativo. Nada tiene que ver con las modalidades aleticas de necesidad, posibilidad, etc.; no es en ningun sentido una noción modal.*

7.1. Referencia a números

Si bien L_{pa}^+ puede entenderse como un lenguaje que habla sobre las expresiones de un lenguaje formal vía códigos de Gödel, sus objetos primarios son los números. Por esta razón empiezo el capítulo analizando la referencia de oraciones de L_{pa}^+ a números.³

El objetivo es dar una definición de referencia que complete la noción intuitiva que esboza Leitgeb (2002), presentada en el apartado 6.1.2. De acuerdo con esta noción informal, las oraciones de un lenguaje de primer orden pueden referir a objetos por mención, conteniendo un término que los denote, o por cuantificación, cuantificando de alguna manera sobre ellos. Comienzo por analizar la referencia por mención.

Intuitivamente, $2 + 2 = 4$ refiere a los números 2 y 4 porque los menciona. Sin embargo, en esta oración no ocurren únicamente términos que denotan estos números sino también 0, 1 y 3, porque 4 es un numeral, el término complejo $SSSS0$ que contiene a los numerales de 0, 1, 2 y 3 como subtérminos propios (vease el apartado 1.2.1). Si que un término ocurra en una oración es suficiente para decir que esta refiere al número denotado por el, toda oración que contenga el numeral n referirá indefectiblemente a

todos los números menores o iguales a n . Para evitar decir que $2 + 2 = 4$ refiere a todos estos números, podrá exigirse que solo los términos que no ocurren como subtérminos propios de otros en una oración sean causa de referencia. Lamentablemente, esta no es una buena idea, porque, en $2 + 2 = 4$, 2 aparece siempre como subtérmino.

Para solucionar este inconveniente tomo la siguiente decisión. Dado que los numerales son los nombres canónicos de los números en L_{pa}^+ , a falta de constantes de individuo los considero como tales, i.e., las nociones de referencia a números que propongo no miran dentro de los numerales, no consideran los subtérminos de estos términos para establecer la referencia de una oración que los contenga, pero sí atienden a los subtérminos de otros

terminos, como 2 en $2 + 2$.

$= 0$, por ejemplo, no

Esta medida puede resultar un tanto arbitraria. $(2 + 1) = 5$ estar a refiriendo al número 3 aunque, en algún sentido, podrá creerse que dice que 3 por 5 es 0. Sin embargo, para no trivializar la noción es preciso tomar algunas decisiones. Estas decisiones dependen del peso de las intuiciones que se tenga y del uso que se le vaya a dar al concepto de referencia. No parecen ser conceptualmente esenciales a lo que sigue, aunque extensionalmente los resultados van a depender completamente de ellas.

³Intuitivamente, no tiene sentido decir de las fórmulas abiertas de un lenguaje que refieren, porque no 'dicen' nada acerca de ninguna cosa, porque simplemente no dicen nada, no 'expresan proposiciones'. Los conceptos de referencia y sus derivados que introduzco a lo largo del capítulo están definidos únicamente para oraciones. Sin embargo, algunas de estas nociones pueden extenderse sin demasiados cambios a fórmulas con variables libres.

Agradezco a Je Ketland por sugerirme esta idea.

Sea L una extensión de L_{pa}^+ con un número finito de símbolos de predicado que cuente con un modelo pretendido N^L que expanda N^+ a todo L .

Definición 7.1 (M-referencia a números). La oración ϕ de L m-referirá a $n \in \mathbb{N}$ si y sólo si n ocurre al menos una vez en ϕ no precedido inmediatamente por S .

Por ejemplo, $S^3 = 0$ m-referirá por mención o m-referirá a 3 y a 0, pero $S(3) = 0$ m-referirá únicamente a 4 y a 0. Los componentes atómicos de una oración determinan su m-referencia. Una negación m-referirá a lo que m-referirá la oración negada, un condicional a lo que m-referen su antecedente o consecuente, etc.

La m-referencia a números es una noción prácticamente sintáctica: habiendo asignado nombres a los números naturales en L_{pa} , basta estudiar los términos que ocurren en una oración para saber a qué números esta m-referirá. Esto la hace una relación recursiva, de muy baja complejidad.

Las oraciones también pueden referir a números vía cuantificadores. Intuitivamente oraciones de la forma $\exists x(x = 2 + 1!)$ m-referen de esta manera al número 3 y $\exists x(x \geq 8!)$ a todos los números mayores que 7. Por otra parte, aquellas expresiones universales cuyos cuantificadores no están relativizados, como $\forall x x = x$, parecen referir a todos los números, precisamente porque la cuantificación es irrestricta. Siguiendo la noción intuitiva de referencia indicada por Leitgeb, me rijo por la siguiente idea básica: el único modo de restringir la referencia por cuantificación es que el cuantificador esté seguido por una expresión condicional, en cuyo caso la referencia se reduce a los objetos que satisfacen el antecedente.

Si el lenguaje contuviera cuantificadores existenciales y conjunciones, otro modo de restringir la referencia sería mediante expresiones de la forma $\exists v(\wedge)$, que referirían a los v 's. En efecto, de acuerdo con la definición de referencia vía cuantificación que doy más adelante, las oraciones de L de las cuales expresiones como $\exists v(\wedge)$ son, mutatis mutandis, abreviaturas o variantes notacionales m-referen a los v 's.

Luego, una oración dada por un cuantificador universal seguido de cualquier fórmula que no sea un condicional referirá absolutamente a todos los objetos del universo de discurso. Esto evidencia también el carácter hiperintensional de la noción de referencia que pretendo dar: no referirán necesariamente a las mismas oraciones de la forma $\forall v(!)$ que aquellas de la forma $\forall v::(!)$.

> ¿A qué m-referen por cuantificación $\exists x(2 = 2! : x = 3)$, $\exists y(0 + 1 = 0 \wedge x)$; $y(x > y! : x + y = 15)$? Para reducir el número de decisiones arbitrarias al mínimo aplico el mismo criterio a $\exists x(2 = 2! : x = 3)$ que a oraciones en las cuales x está libre en el antecedente del condicional, esto es, $\exists x(2 = 2! : x = 3)$ referirá a todo número n que

satisfaga $\forall x (x = 2)$, esto es, a todo. Si el antecedente fuera falso, no referir a a nada. En general, oraciones de la forma $\forall x (A(x))$ refieren a n si y solo si $\forall x [A(x)]$ es verdadera.

Con respecto a $\forall x (0 + 1 = 0)$, el criterio tambien se desprende de lo dicho en el parrafo anterior. Como el cuantificador $\forall x$ no esta relativizado en modo alguno, esta oracion referira a todo, ignorando por completo que el cuantificador es superuo.

Finalmente, en cuanto a $\forall x; y (x > y \wedge x + y = 15)$, la extension del criterio para formulas de tipo $\forall v_1 \dots v_m (A)$ parece indicar que refieren a los pares ordenados de numeros tales que la primera componente es mayor que la segunda. Sin embargo, para que la referencia sea estrictamente una relacion entre oraciones y los objetos que menciona o sobre los cuales cuantifica, y evitar las dificultades de los enfoques de Putnam, Goodman y Urbaniak introducidos en el capitulo anterior, dire que $\forall x; y (x > y \wedge x + y = 15)$ y, en general, oraciones de la forma $\forall v_1 \dots v_m (A)$ refieren a n siempre y cuando $\exists v_1 \dots v_m [A]$ se de para algun $1 \leq i \leq m$, esto es, n sea una componente de alguna m-tupla $\langle k_1, \dots, k_m \rangle$ que satisfaga A. La siguiente definicion rescata las intuiciones y decisiones delineadas.

Definicion 7.2 (Q-referencia a numeros). Una oracion A refieren a n si y solo si se cumple una de las condiciones 1-3:

1. A es de la forma $\forall v_1 \dots v_m (A)$
 - a) A es una formula atomica o $\forall v_i (A)$, o
 - b) $\exists v_i (A)$ y, para algun i tal que $1 \leq i \leq m$,

$$N^L \exists v_1 \dots v_m [A]$$

2. A := B y B q-refieren a n.

3. A := B y B o q-refieren a n.

Al igual que la m-referencia, la q-referencia tambien es composicional, pero no depende ya de lo que sucede al nivel atomico, i.e., nada, sino de lo que ocurre con las subformulas universalmente cuantificadas.

Si bien parece que la definicion de referencia por cuantificacion o q-referencia es demasiado generosa con la referencia, por ejemplo, en oraciones universales seguidas por expresiones no condicionales, al contrario y a diferencia de otros abordajes como los expuestos en 6.3, abre la posibilidad de la referencia restringida mediante condicionales, de utilizar predicados en ciertas oraciones y que estas no refieren necesariamente a los objetos que los satisfacen. Todo depende de la expresion que se utilice.

La conjunción y el cuantificador existencial son símbolos de nidos en el lenguaje. Si, como es usual, se denota \wedge como $(\cdot \wedge \cdot)$ y $\exists v$ como $(\exists v \cdot)$, oraciones de la forma $\exists v_1 \dots \exists v_n (\cdot \wedge)$, que intuitivamente son casos de referencia restringida a los 's, son abreviaturas de $(\exists v_1 \dots \exists v_n : (\cdot \wedge))$, que q-refiere a todo, por la cláusula 1.a). Sin embargo, eliminando la doble negación, $(\exists v_1 \dots \exists v_n (\cdot \wedge))$ q-refiere a los 's, como se quería. Luego, la definición 7.2 es naturalmente extensible a un lenguaje con cuantificadores existenciales y conjunciones.

Naturalmente, una oración puede m y q-referir a números a la vez. Por ejemplo, $(\exists x (x = 2 \wedge 3 = 5) \wedge \forall x (x < 3 \rightarrow x = 6))$ m-refiere a 2; 3 y 5 y q-refiere a 0; 1 y 2. Luego, refiere simpliciter a 0; 1; 2; 3 y 5.

Definición 7.3 (Referencia a números). Una oración ϕ m-refiere a n si ϕ m-refiere a n y q-refiere a n .

Dada la composicionalidad de la m y la q-referencia, la referencia simpliciter es también de alguna manera composicional. En particular, si una expresión refiere a un número, negarla o ubicarla en el antecedente o consecuente de un condicional no cambia este hecho. Este es un segundo inconveniente de los conceptos de contenido referencial de Putnam, Goodman y Urbaniak que mi noción logra superar.

Para saber si oraciones de la forma $(\exists v (\cdot))$ q-refieren a un número n , hay que saber primero si (\cdot) es verdadera, esto es, si $N^L(\cdot)$; para saber que pares ordenados de oraciones, por un lado, y números, por el otro, están en la relación de q-referencia es preciso saber que oraciones son verdaderas en el modelo estándar de la aritmética. Como consecuencia, la q-referencia no es definible en L_{pa} . Al estar definida en términos de la q-referencia, la referencia también es una noción altamente compleja, no definible en L_{pa} . Luego, si se quiere contar con esta noción en el lenguaje objeto mismo, i.e., en L o una de sus extensiones, es preciso ofrecer una axiomatización de la noción de referencia, esto es, introducir en L un símbolo de predicado primitivo para esta relación e incorporar axiomas a una teoría que axiomatice parcialmente N^L que lo gobiernen correctamente. Lo mismo puede decirse de las nociones de referencia a oraciones y referencia aleática semántica que introduzco en los apartados subsiguientes. Para evitar repeticiones tediosas, pruebo un resultado de indefinibilidad y doy una axiomatización sobre pa solo para esta última (véase el apartado 7.3.1.2), lo cual podrá traducirse por analogía en resultados y teorías acerca de las otras dos nociones.

7.2. Referencia a oraciones

Gracias a la posibilidad de codificar las expresiones de un lenguaje natural mediante números naturales, el lenguaje de la aritmética y sus extensiones pueden entenderse también como hablando de sus propias u otras expresiones, y la aritmética como una teoría de la sintaxis para aquel lenguaje. En esto se basan los lemas de diagonalización, las oraciones de Gödel y Henkin y la posibilidad de formular teorías de la verdad en expansiones de L_{pa}^+ y sus posibles paradojas. Para dar cuenta de estos y otros fenómenos, en esta sección y en la que sigue doy nociones de referencia de expresiones a expresiones a través de sus códigos, a partir de las cuales se puede definir autorreferencia, buena fundación y otros patrones referenciales.

Mientras que L_T , las expresiones paradójicas y la referencia aleática aparecen en la sección siguiente, esta sección se ocupa únicamente de la referencia en L_{pa}^+ . Como anticipo, por su adecuación material, esta noción permitirá resolver una serie de cuestiones, como el problema de Henkin y las preguntas alrededor de la autorreferencialidad de las oraciones de Gödel y de toda expresión que es el resultado de un proceso de diagonalización.

En primer lugar ofrezco nociones de m-referencia y q-referencia, referencia directa y referencia simpliciter a oraciones. Dado que se trata de una relación entre oraciones, también de no autorreferencia y buena fundación a partir de los conceptos anteriores, y muestro que los procesos de diagonalización presentados en el apartado [1.2.6](#) dan lugar a expresiones autorreferenciales, mientras que las cadenas que se obtienen mediante, por ejemplo, a la manera de Yablo, resultan no bien fundadas (y, lamentablemente, en varios casos también autorreferenciales). Finalmente, evalúo a la luz de las nociones introducidas las oraciones de Henkin y Gödel, entre otras.

Definir referencia de expresiones a expresiones rescatando las intuiciones más básicas del lenguaje natural no es tarea sencilla. Tanto la referencia por mención como por cuantificación se vuelven más complicadas que la referencia a números. Analizo primero la m-referencia.

Sea L , como en el apartado anterior, una expansión de L_{pa}^+ con un número finito de símbolos de predicado que cuente con un modelo pretendido N^L que expanda N^+ a todo L . Tómese como ejemplo la oración de Gödel de pa^+ que se obtiene a través del lema fuerte de diagonalización (teorema [1.8](#)):

$$q^+ \text{ } \dot{=} \text{ } g = p: \text{Bew}_{pa^+}(g)q$$

Naturalmente sería deseable decir que $\text{Bew}_{pa^+}(g)$ refiere a sí misma, porque contiene un término, g , que la denota. Luego, no es posible tomar los numerales como única fuente de m -referencia. Si se siguiera la definición 7.1 de m -referencia a números, sería necesario que $p:\text{Bew}_{pa^+}(g)q$ ocurriera en $\text{Bew}_{pa^+}(g)$, esto es, que el numeral de código de esta oración ocurriera sintácticamente en la oración misma, lo cual no es posible bajo codificaciones monótonas de oraciones a través de números (véase el apartado 1.2.3). De acuerdo con el lema diagonal fuerte, g no es un numeral sino un término de la forma $t_1(t_2=t_3)$, esto es, una instancia del símbolo de función para la función sustitución.

La opción más natural parecería ser abandonar la restricción a numerales y considerar que cualquier término (como, prima facie, g) que no ocurra como subtérmino propio de otros en la fórmula, es fuente de m -referencia. Sin embargo, esta alternativa no es viable. Como L_{pa}^+ (y también L_{pa}) contiene términos de función, en sus predicados las variables libres pueden aparecer como subtérminos propios de otros términos, v.g., puede que en $\text{Bew}_{pa^+}(x)$, la única ocurrencia libre de x está en la subfórmula $f(x) = t$, donde f es un símbolo de función y t un término. En ese caso, g ocurre en $\text{Bew}_{pa^+}(g)$ como un subtérmino propio y, en consecuencia, $\text{Bew}_{pa^+}(g)$ no refiere a sí misma.

Aparentemente, la única opción que queda consiste en considerar que cualquier término que ocurra en una expresión, ya sea como subtérmino propio o no, ya sea un numeral u otro término complejo, puede ser fuente de m -referencia, lo cual está en línea con la cláusula de la referencia por mención de la noción intuitiva esbozada en el apartado 6.1.2.

Definición 7.4 (M -referencia a oraciones). Dadas dos oraciones ϕ ; $\psi \in L$, ϕ m -refiere a ψ sii ϕ contiene un término cerrado t tal que $q^+ \vdash t = \psi$.

Notese que, como q^+ prueba todas las identidades verdaderas en N^+ y, por tanto, en N^L , el concepto de m -referencia es extensionalmente equivalente a uno en el cual, en lugar de que q^+ pruebe $t = \psi$, se pide que $N^L \models t = \psi$: no hay más casos de denotación en el modelo estándar que los que q^+ puede probar.

Este es el llamado 'criterio de Kreisel-Henkin', como indique en el apartado 6.3. Las consecuencias de adoptarlo no son del todo deseables, porque se multiplican sospechosamente los casos de m -referencia (aunque ahora solo los términos que denotan oraciones pueden dar lugar a una relación de m -referencia). Por ejemplo, como $\text{Bew}_{pa^+}(x)$ es una fórmula compleja que puede contener sintácticamente otros términos aparte de x (puede ser de la forma $\phi(t; x)$, donde t es un término cerrado que denota una oración) todas las expresiones de la forma $\text{Bew}_{pa^+}(s)$ no solo refieren a lo que s denota sino también a lo que los otros términos en $\text{Bew}_{pa^+}(x)$ denotan (si t denota una oración, $\text{Bew}_{pa^+}(s)$ refiere a la denotación de t , para todo término cerrado s). Además, todos los

subterminos de un termino que ocurre en una oracion tambien ocurren sintacticamente

en la oracion, con lo cual cualquier formula $\ulcorner(p \ q) \ m\urcorner$ referira a $\ulcorner p \urcorner$ y a todas las oraciones de L_{pa^+} cuyo codigo es menor que el de $\ulcorner p \urcorner$.

Por otra parte, el criterio de Kreisel-Henkin obvia una gran cantidad de casos claros de m-referencia. Los ciclos de oraciones, que son tambien paradigmaticamente clasi cados como autorreferentes, no lo son segun este criterio. Como los ciclo de mentirosos que introduce en cap tulos anteriores (cf. (3.6)), existen en q^+ ciclos como el de Henkin de longitud 2 en pa^+ :

$$h_1 = pBew_{pa^+}(h_2)q \tag{7.1}$$

$$h_2 = pBew_{pa^+}(h_1)q$$

y el ciclo de Godel de longitud 3 en pa^+ :

$$g_1 = p:Bew_{pa^+}(g_2)q \tag{7.2}$$

$$g_2 = p:Bew_{pa^+}(g_3)q$$

$$g_3 = pBew_{pa^+}(g_1)q$$

entre otros, en virtud de los lemas diagonales alternativos (proposicion 1.10). El ciclo en (7.1) consta de dos oraciones, cada una de las cuales dice de la otra que es un teorema de pa^+ , mientras que en (7.2) hay tres oraciones, cada una de las cuales dice que la oracion que se encuentra por debajo no es un teorema excepto la ultima, que dice que la primera es un teorema.

Dado que las identidades en (7.1) y (7.2) son demostrables en q^+ , en ambos casos los fenomenos que ocurren para la oracion de Henkin y la oracion de Godel se reproducen. Por un lado, tanto $Bew_{pa^+}(h_2)$ como $Bew_{pa^+}(h_1)$ son demostrables en pa^+ , porque

$$pa^+ \vdash Bew_{pa^+}(h_2) \wedge Bew_{pa^+}(pBew_{pa^+}(h_2)q) \quad \text{3ra condicion de teoremicidad de } L \cdot ob^5$$

$$Bew_{pa^+}(h_1) \quad \text{Por (7.1)}$$

y

$$pa^+ \vdash Bew_{pa^+}(h_1) \wedge Bew_{pa^+}(pBew_{pa^+}(h_1)q) \quad \text{3ra condicion de teoremicidad de } L \cdot ob$$

$$Bew_{pa^+}(h_2) \quad \text{Por (7.1)}$$

con lo cual $pa^+ \vdash Bew_{pa^+}(h_2) \wedge Bew_{pa^+}(h_1)$. Por el teorema de L.ob (corolario 1.14), esto implica que $Bew_{pa^+}(h_2)$ y $Bew_{pa^+}(h_1)$ son teoremas de pa^+ .

Por otro lado, al igual que la oracion de G•odel, ni $\text{Bew}_{\text{pa}^+}(g_2)$, ni $\text{Bew}_{\text{pa}^+}(g_3)$ ni $\text{Bew}_{\text{pa}^+}(g_1)$ son teoremas de esta teor a. Para las dos primeras es obvio, por el segundo resultado de incompletitud de G•odel (teorema 1.13). Si la tercera fuera un teorema de pa^+ , como $\text{Bew}_{\text{pa}^+}(v)$ debilmente representa el conjunto de teoremas de esta teor a, $\text{Bew}_{\text{pa}^+}(g)$ tambien ser a un teorema de pa^+ . (7.2) permite dar una prueba alternativa al primer resultado de incompletitud de G•odel.

Si la autorreferencia es parte de la causa detras de la no demostrabilidad de las oraciones de G•odel y la demostrabilidad de las oraciones de Henkin, tambien debe serlo en el caso de los ciclos de Henkin y G•odel como (7.1) y (7.2). Las primeras pueden ser vistas como un caso limite de las ultimas. Pero de acuerdo con el criterio de Kreisel-Henkin, ninguna de estas oraciones refiere a s misma.

No solo los ciclos de cualquier longitud son posibles en q^+ sino tambien, en virtud de la proposicion 1.11, las cadenas infinitamente descendientes, en donde cada una de las oraciones refiere por mencion, y no por cuantificacion como en casos semejantes a la secuencia de Yablo, a la que le sigue. Por ejemplo, es posible obtener una secuencia de longitud ω de terminos $h_0; \dots; h_n; \dots$ tales que

$$\begin{aligned} h_0 &= \text{pBew}_{\text{pa}^+}(h_1)q & (7.3) \\ h_1 &= \text{pBew}_{\text{pa}^+}(h_2)q \\ h_2 &= \text{pBew}_{\text{pa}^+}(h_3)q \\ &\vdots \end{aligned}$$

esto es, una cadena de oraciones de Henkin.

Prima facie, en todos los ejemplos de ciclos y cadenas que he dado, las oraciones no se limitan a referir a aquellas que mencionan directamente, sino tambien a aquellas a las cuales las que mencionan refieren. La referencia en estos casos parece estar cerrada transitivamente, lo cual no esta contemplado por el criterio de Kreisel-Henkin, que solo establece m-referencia directa. Para evitar la multiplicacion innecesaria de referencias, en lugar de dar una nocion de m-referencia transitiva, trabajo primero sobre una nocion de q-referencia directa a oraciones, para luego definir referencia directa simpliciter como la disyuncion entre la m y la q-referencia y cerrarla bajo transitividad, obteniendo finalmente la nocion transitiva de referencia simpliciter.

Con respecto a la q-referencia a oraciones los problemas son aun mas serios. El criterio dado para la q-referencia a numeros es demasiado sencillo y deja muchos casos claros de q-referencia de lado. Si se recuerda que el objetivo principal de este capitulo es dar con una nocion de referencia que permita restringir el [Esquema-T](#) a instancias seguras, esto es inaceptable. Si una oracion refiere intuitivamente a s misma y, de ese modo,

lleva a contradicciones, pero la nocion con la cual se trabaja pasa por alto este caso de referencia, la teor a resultar inconsistente. Luego, es mejor un criterio que sobregenere casos de referencia a uno que se quede corto.

Considerese la oracion

$$\exists x(x < h \wedge \neg \text{Bew}_{pa^+}(\ulcorner x \urcorner)) \tag{7.4}$$

Intuitivamente, esta oracion refiere no solo a las oraciones que satisfacen el predicado $x < h$ sino tambien a sus negaciones. Si $g < h$, la oracion dice, intuitivamente, que la negacion de la oracion de Godel de pa^+ no es un teorema de pa^+ , esto es, refiere directamente a $\neg \text{Bew}_{pa^+}(g)$. Considerese este otro ejemplo:

$$\exists x(x = g \wedge \exists y(y = \ulcorner x \urcorner \wedge \neg (y))) \tag{7.5}$$

Esta oracion no solo parece referir a g sino tambien a su negacion, y es de esta ultima de la que dice que es $\ulcorner \cdot \urcorner$. La solucion mas sensata en ambos casos parece ser decir que una oracion de la forma $\exists v(\ulcorner \cdot \urcorner) \text{ q-re refiere}$, no solo a las oraciones que satisfacen su antecedente, sino tambien a las oraciones a las cuales $(\ulcorner \cdot \urcorner) [n=v]$ (m - o q -) refiere, donde n es un numero que satisface $\ulcorner \cdot \urcorner$.

Definicion 7.5 (Q-referencia a oraciones). Dadas dos oraciones $\ulcorner \cdot \urcorner$; $\ulcorner \cdot \urcorner \in L$, $\ulcorner \cdot \urcorner$ q-re refiere a $\ulcorner \cdot \urcorner$ sii se cumple una de las condiciones 1-3:

1. $\ulcorner \cdot \urcorner := \exists v_1 \dots v_m \ulcorner \cdot \urcorner$ y

- a) $\ulcorner \cdot \urcorner$ es una formula atomica o $\ulcorner \cdot \urcorner := \ulcorner \cdot \urcorner$, o
- b) $\ulcorner \cdot \urcorner := \ulcorner \cdot \urcorner$ y, para algun i tal que $1 \leq i \leq m$,

$$\exists v_1 \dots v_i \exists v_{i+1} \dots v_m [p \text{ q} = v_i]$$

o existe un $k \in \mathbb{N}$ tal que

$$\exists v_1 \dots v_i \exists v_{i+1} \dots v_m [k = v_i]$$

y $\exists v_1 \dots v_i \exists v_{i+1} \dots v_m (\ulcorner \cdot \urcorner) [k=v_i]$ q-re refiere a $\ulcorner \cdot \urcorner$ o contiene una ocurrencia de un termino cerrado t que no esta en L tal que $q^+ \ulcorner t \urcorner = p \text{ q}$.

2. $\ulcorner \cdot \urcorner := \ulcorner \cdot \urcorner$ y q-re refiere a $\ulcorner \cdot \urcorner$.

3. $\ulcorner \cdot \urcorner := \ulcorner \cdot \urcorner$ y o q-re refiere a $\ulcorner \cdot \urcorner$.

Esta es una definicion inductiva. Para cada oracion, en el dominio el predicado 'q-re ere' se aplica exclusivamente a formulas de menor complejidad. Es preciso indicar ! que la m-referencia de () [k=vi] debe ser 'nueva', esto es, el termino que denota no debe ya ocurrir en la formula original porque, de lo contrario, claros casos de m-referencia de oraciones universales seguidas por condicionales cuyo antecedente sea satisfecho por al menos un numero ser an declarados a la vez casos de q-referencia por la definicion. Por ejemplo, (7.4) m-re ere a la oracion de Henkin denotada por h pero no q-re ere a ella, porque su codigo no es menor a s mismo, su conectiva principal no es una negacion y, si $k < h$, $k < h$! : $Bew_{pa^+}(\cdot, k)$ tiene una ocurrencia de un termino h que denota la oracion de Henkin pero esta ya aparece a en la oracion original.

La clausula 1.b) permite acomodar las intuiciones sobre ejemplos como (7.4) y (7.5). Asimismo, la definicion 7.5 ayuda a esclarecer el criterio tecnico de autorreferencia de Priest (1997) expuesto en el apartado 6.1.2 y termina con el problema de las equivalencias entre oraciones y otras que las mencionan como fuente de autorreferencia.

De acuerdo con Leitgeb (2002) toda nocion de referencia para expresiones de L_{pa}^+ enfrenta serias dificultades porque, por un lado, toda oracion de L_{pa}^+ es equivalente en pa^+ a otra que la menciona pero, a la vez, las equivalencias no pueden ser abandonadas porque son el medio de establecer autorreferencia en lenguajes sin simbolos extra de funcion, como L_{pa} [v.g., en casos como los de las oraciones de Godel y el mentiroso debiles] (vease el apartado 6.2).

De acuerdo con las definiciones que introduce, la equivalencia entre ϕ y $(p \phi)$ no es una razon valida para concluir que ϕ es autorreferente, ya que ni la definicion 7.4 ni la definicion 7.5 (ni su union y posterior clausura transitiva) permiten tal inferencia. Pero, como anticipa y como quiere Priest, si ϕ ha sido obtenida mediante un proceso de diagonalizacion debil sobre '(v)', es autorreferencial y, mas especificamente, q-autorreferencial, porque, como muestra la prueba del teorema 1.7, ϕ es la oracion

$$\exists u(u = p \exists v(\text{Diag}(u; v) \neg '(v)) \neg q \exists v(\text{Diag}(u; v) \neg '(v))) \tag{7.6}$$

que, por la definicion 7.5, q-re ere a todo lo que

$$\exists v(\text{Diag}(p \exists v(\text{Diag}(u; v) \neg '(v)) \neg q; v) \neg '(v))$$

q-re ere, que a su vez, nuevamente por la definicion de q-referencia, q-re ere a todas las oraciones que satisfacen el predicado

$$\text{Diag}(p \exists v(\text{Diag}(u; v) \neg '(v)) \neg q; v)$$

i.e., a la diagonalizacion de $\delta v(\text{Diag}(u; v) ! '(v))$, esto es, a (7.6), a . En consecuencia,

es autorreferente; ni el lema de diagonalizacion fuerte ni la condicion de equivalencia son necesarios para establecer autorreferencia.

Contrario a lo que creyo Leitgeb, no es en virtud de las equivalencias que el lema de diagonalizacion debil logra establecer autorreferencia, sino mediante los cuantificadores involucrados en las oraciones obtenidas para cada predicado. Luego, si bien toda oracion de L es aritmeticamente equivalente a una oracion que la menciona, no toda oracion de L es autorreferente. Como quieren Halbach & Visser (2014a), $0 = 0$ no es autorreferencial, a pesar de ser demostrablemente equivalente a $\text{Bew}_{pa^+}(p0 = 0q)$, porque solo menciona a 0, que no denota una oracion, y carece de cuantificadores, mientras que $\text{Bew}_{pa^+}(p0 = 0q)$ s menciona a $p0 = 0q$ y, por eso, refiere a ella.

Como en el caso de la m-referencia, es tanto posible la q-referencia directa como la indirecta. Las cadenas y los ciclos pueden reescribirse facilmente con cuantificadores. Ademas, los cuantificadores posibilitan otros patrones referenciales semejantes a los de la secuencia de Yablo y otros. Siguiendo un procedimiento analogo al que se empleo en el apartado 6.1.1 para la construccion de la lista de Yablo pero en su version reformulada, dada una formula '(v) de L, se aplica el lema diagonal fuerte (teorema 1.8) a

$$\delta w(\exists u > v w = z(u=pvq) ! '(w))$$

sobre la variable z, obteniendo la identidad

$$q^+ \cdot t = p\delta w(\exists u > v w = t(u=pvq) ! '(w))q$$

Aplicando a derecha e izquierda de esta identidad la funcion $x(v=pvq)$ sobre x, se obtiene el principio general

$$q^+ \cdot t(v) = p\delta w(\exists u > v w = t(u=pvq) ! '(w))q$$

del cual se derivan todas las instancias

$$q^+ \cdot t(n) = p\delta w(\exists u > v w = t(u=pvq) ! '(w))q: \tag{7.7}$$

As , se obtiene una lista in nita de oraciones cada una de las cuales refiere a todas las que se encuentran por delante, directa e indirectamente. Luego, es deseable una noción de q-referencia transitiva. Pero tambien es posible la referencia a secas, que podr a darse tal vez, v.g., entre dos oraciones tales que la primera m-refiere a una tercera, que a su vez q-refiere a la segunda. De no, primero, la referencia directa como la union de la

m-referencia y la q-referencia y, luego, la referencia indirecta como la clausura transitiva de esta ultima.

De nicion 7.6 (Referencia directa a oraciones). Dadas dos oraciones $s, t \in L$, s re ere directamente a t sii m-re ere directamente a t o q-re ere directamente a t .

De nicion 7.7 (Cadenas de referencia directa). Una secuencia de oraciones $\theta_1, \dots, \theta_n \in L$, $n \geq 1$, es una cadena de referencia directa sii, para cada $i < n$, θ_i re ere directamente a θ_{i+1} .

De nicion 7.8 (Referencia a oraciones). Sean s y oraciones de L . s re ere a t sii existe una cadena de referencia directa $\theta_1, \dots, \theta_n$ tal que $\theta_1 = s$ y $\theta_n = t$.

Esta ultima de nicion rescata las intuiciones alrededor de los bucles y las cade-nas. Por ejemplo, en el caso del ciclo de Henkin de longitud 2 en (7.1), como $Bew_{pa^+}(h_2)$ contiene un termino que denota $Bew_{pa^+}(h_1)$ y viceversa, si bien ninguna de las oraciones m-re ere directamente a s misma, cada una (m-)re ere indirectamente a la otra mismas, de acuerdo con la de nicion 7.8.

Otro tanto se puede decir de las cadenas, como la cadena de oraciones de Henkin en (7.3). Se sabe, por la de nicion 7.4 de m-referencia, que, para cada $i \geq 1$, $Bew_{pa^+}(h_i)$ re ere directamente a $Bew_{pa^+}(h_{i+1})$. Luego, en virtud de la de nicion 7.8, $Bew_{pa^+}(h_i)$ tambien re ere a $Bew_{pa^+}(h_{i+2})$ a traves de $Bew_{pa^+}(h_{i+1})$. Pero entonces tambien re ere a $Bew_{pa^+}(h_{i+3})$ mediante $Bew_{pa^+}(h_{i+2})$, y as siguiendo. Cada oracion en la secuencia termina re riendo a todas las que se encuentran por debajo, como se quer a.

No obstante, la transitividad de la referencia no parece ser siempre algo deseable; hay predicados para los cuales resulta mas natural que otros. Mientras que para predicados como T o Bew parece necesaria, no tanto as para otros como $Sent$ o U .⁶

Considerese la siguiente formula:

$$Sent_{L_{pa^+}}(pU(t; s; p0 = 0q)q) \tag{7.8}$$

Esta oracion m-re ere a $U(t; s; p0 = 0q)$, que a su vez m-re ere a $0 = 0$. Luego, por la de nicion 7.8, se debe concluir que (7.8) tambien re ere a $0 = 0$. Esto puede resultar un tanto contraintuitivo, en la medida en que (7.8) hace una afirmacion sobre la sintaxis de $U(t; s; p0 = 0q)$ que, a diferencia de enunciados que predicen verdad o teoremicidad, de ninguna manera se puede traducir en una afirmacion acerca de los objetos de los cuales habla esta ultima oracion, i.e., $0 = 0$. Distinguir entre predicados para los cuales

⁶ $U(v_1; v_2; v_3) \in L_{pa^+}^*$ representa en q^* la relacion recursiva entre los codigos de una oracion v_1 , una secuencia v_2 de codigos de variables de individuo y una formula v_3 que se da siempre y cuando v_1 conste de v_3 precedida por una secuencia de los cuantificadores universales sobre las variables en v_2 en ese orden. Vease el apartado 1.2.3 para refrescar la notacion.

la transitividad es razonable y aquellos para los que no es, al menos, complejo. Tras una serie de intentos he desistido. Si debe considerarse o no la referencia indirecta dependera, en cada caso, del uso que se quiera dar a las nociones de referencia a oraciones que ofrezco.

A partir de los conceptos introducidos es posible dar nociones de autorreferencia y buena fundacion (well-foundedness), que no debe confundirse con la fundacion a secas, v.g., de Kripke o Leitgeb (veanse los apartados [3.1](#) y [3.2](#)).

Definicion 7.9 (Autorreferencia). Una oracion $\phi \in L$ es autorreferente sii refiere a si misma.

Definicion 7.10 (Buena fundacion). Una oracion $\phi \in L$ esta bien fundada sii ninguna cadena de referencia que empiece con ϕ es indefinidamente extensible.

Las oraciones de Gödel y Henkin tanto en sus versiones fuertes como debiles resultan autorreferentes y, por tanto, tambien no bien fundadas o mal fundadas, como evidencian las cadenas de referencia $\vdash Bew_{pa^+}(g)$; $\vdash Bew_{pa^+}(g)$; $\vdash \vdots$ y $Bew_{pa^+}(h)$; $Bew_{pa^+}(h)$; $\vdash \vdots$, indefinidamente extensibles. Asimismo, los ciclos son autorreferentes y no bien fundados. Por ejemplo, para el ciclo de Gödel de longitud 3 en [\(7.2\)](#) existe la cadena de referencia indefinidamente extensible

$$\vdash Bew_{pa^+}(g_2); \vdash Bew_{pa^+}(g_3); \vdash Bew_{pa^+}(g_1); \vdash Bew_{pa^+}(g_2); \vdash Bew_{pa^+}(g_3); \vdash \vdots$$

Las cadenas de oraciones dadas por la proposicion [1.11](#) resultan ser todas mal fundadas aunque lamentablemente tambien autorreferentes, por el modo en el que estan contruidos los terminos. Cada termino t_n es tal que $t_n = p'(t_{n+1})q = p'(g.(num(p'(t_n)q)))q$.

Luego, t_{n+1} o, lo que es lo mismo, $t(g.(num(p'(t_n)q)))$, m-refiere a $t(t_n)$, que a su vez m-refiere a $t(t_{n+1})$. Algo parecido sucede tambien con las secuencias yablezcas_de oraciones como en [\(7.7\)](#). El codigo de cada oracion $\delta_w \delta_u > n (w = t(u=pvq) \wedge !'(w)) (= t(n))$ de la secuencia es claramente mayor a n , porque la codificacion es monotona. Pero, a la vez, esta oracion q-refiere a toda oracion cuyo codigo sea mayor a n , porque, como $t(u=pvq)$ representa una funcion, para cada $k > n$ se puede probar que

$$q \wedge \delta_w (k > n \wedge w = t(k=pvq))$$

y, a fortiori, que

$$q \wedge \delta_w (t(n) > n \wedge w = t(num(t(n)=pvq)))$$

Luego, por la clausula 1.b) de la definicion de q-referencia, todas las oraciones de la secuencia yablezca son autorreferentes.

Estas consecuencias no deben achacarse a los conceptos de referencia que introduce en esta sección, sino más bien a la carencia de L_{pa}^+ de constantes de individuo y predicado, nombres simples, para oraciones y sus propiedades correspondientes.

Claramente, el concepto de buena fundación de nido en [7.10](#) está muy lejos, v.g., del concepto de fundación de [Kripke \(1975\)](#) o de [Leitgeb \(2005\)](#), donde todas las oraciones de L_{pa}^+ están fundadas porque expresan 'hechos no semánticos', aun cuando, como en el caso de [\(7.3\)](#), no estén bien fundadas de acuerdo con la definición [7.10](#). Las nociones de buena fundación alética que introduzco en la siguiente sección tampoco respetan las intuiciones de un concepto de fundación pero son mucho más cercanas a este.

7.3. Referencia alética

Esta investigación se centra principalmente en el estudio de las oraciones que dan lugar a paradojas semánticas. Si los patrones de referencia que subyacen a estas expresiones son su característica distintiva, el foco de atención debe ponerse principalmente en L_T , el lenguaje que resulta de expandir L_{pa}^+ con el símbolo monádico de predicado T para la verdad. Lamentablemente, las nociones de referencia dadas en el apartado anterior no pueden emplearse para investigar la referencia de las expresiones de L_T , porque están únicamente de nidas para oraciones de un lenguaje L para el cual existe ya un modelo pretendido N^L . Pero este modelo es precisamente el resultado de la búsqueda de una teoría de la verdad, lo que todavía no se tiene.

Como lo que se busca es una teoría desentremalladora de la verdad y para formularla hay que saber primero que oraciones pueden dar lugar a instancias 'seguras' del [Esquema-T](#), lo cual, por el momento conjeturo, puede hacerse en términos de sus patrones referenciales, es necesario averiguar en primera instancia cuáles son estos patrones para recién después emplearlos en la construcción de una teoría de la verdad. Considerese, por ejemplo, la siguiente oración:

$$\exists x(Tx \rightarrow \text{Bew}_{pa^+}(x))$$

>¿A que se refiere de acuerdo con la definición [7.8](#) de referencia del apartado anterior? Para saberlo, debería averiguarse que oraciones satisfacen el predicado T , para lo cual sería necesario saber primero cuáles enunciados dan lugar a instancias del [Esquema-T](#), lo cual no es posible.

En consecuencia, esta sección está dedicada a una familia de nociones especiales que giran en torno a la referencia relativa al predicado de verdad, a través del predicado

de verdad, que llamo, por esta razón, 'referencia aletica'. La referencia aletica es, a grandes rasgos, una restricción de la referencia a oraciones estudiada en el apartado anterior al predicado de verdad, con algunas salvedades.

Como anticipo, en [7.3.1](#) doy una noción de referencia aletica semantica que, al igual que los conceptos de referencia a números y a oraciones, esta definida en términos de N^+ . Esta noción, si bien es materialmente adecuada y rescata las intuiciones detrás de la noción técnica de autorreferencia de Priest, es indistinguible en el lenguaje de la aritmética, al igual que aquellas. En consecuencia, restringir el [Esquema-T](#) utilizando criterios formulados en base a esta noción semantica introducir a nociones que exceden la sintaxis en los principios para la verdad y, al mismo tiempo, dificultar a su axiomatización, lo cual da lugar a incompatibilidades con el deflacionismo y el proyecto minimalista de Horwich (véase el apartado [5.3](#)), no alcanzando el objetivo central de la tesis.

Consecuentemente, en [7.3.2](#) doy una noción más simple de referencia aletica, en la cual se preserva buena parte de las virtudes del concepto semantico pero la idea de verdad en el modelo estándar es reemplazada por la de demostrabilidad en una teoría que axiomatice, al menos parcialmente, ese modelo; las nociones semanticas son reemplazadas por nociones de teoría de la prueba, dando lugar a la referencia aletica sintactica o referencia aletica, simpliciter. Esta será la noción que jugará un rol central en la construcción de teorías axiomáticas desentramadoras de la verdad en el capítulo siguiente, sistemas no solo correctos sino deductivamente poderosos y, lo que es más importante, capaces de ser adoptados por el deflacionismo y el minimalismo.

En ambos casos, muestro que la tesis referencialista ortodoxa es falsa, no toda paradoja involucra autorreferencia y, a la vez, doy una buena cantidad de evidencia a favor de la tesis referencialista general, mostrando que todas las paradojas conocidas son mal fundadas.

7.3.1. Referencia aletica semantica

Al tratarse de referencia a través del predicado de verdad, solo oraciones que contengan T podrán referir aleticamente a otras, pero hay muchas más sutilezas. Por ejemplo, $p \rightarrow q \leftrightarrow T p \rightarrow q$ será intuitivamente aletica, a y a, pero solo a esta última a través del predicado de verdad; y $\exists x(x = h \wedge T :x)$ será prima facie directamente aletica a la oración de Henkin y a su negación, pero solo será directamente aletica a su negación a través del predicado de verdad, porque 'dice' que $\text{Bew}_{pa^+}(h)$ es verdadera, pero no dice nada de la verdad de $\text{Bew}_{pa^+}(h)$. En consecuencia, además de la cuestión central, de los condicionales relativizados a expresiones que contienen el predicado veritativo, algunos detalles de las definiciones de m y q -referencia a oraciones de la sección anterior

deben ser modificados para dar conceptos adecuados de referencia alética, lo cual ocupa el siguiente apartado.

7.3.1.1. Nociones básicas y su ejemplificación

Al trabajar con un predicado primitivo como único 'generador de referencia', los problemas que los términos complejos de L_{pa}^+ generaban a la hora de definir la m-referencia, ya sea a números o a oraciones, quedan atrás. Ahora es posible considerar únicamente los términos 'enteros' que ocurren en las oraciones e ignorar sus subtérminos propios. También los problemas alrededor de la presunta transitividad de la referencia dejan de existir, porque T es claramente un predicado para el cual la referencia indirecta es deseable.

Notese que las nociones de m-referencia a números y a oraciones en las definiciones 7.1 y 7.4 no recurren a ninguna noción semántica, no hablan de la verdad en el modelo estándar del lenguaje de ninguna oración. Por ende, su 'restricción' al predicado de verdad no da lugar a una noción semántica de m-referencia alética. En efecto, la definición que sigue servirá para definir tanto la referencia alética semántica como la sintáctica. El carácter semántico de las nociones de referencia que introduzco proviene siempre de la q-referencia.

Definición 7.11 (M-referencia alética). Sean ϕ y ψ oraciones de L_T . ϕ m-referirá aléticamente a ψ si y sólo si $T\phi$ es una suboración de ψ y $q^+ \vdash \phi \rightarrow \psi$.

Esta definición cubre las intuiciones con respecto a la oración fuerte del mentiroso (reforzado):

$$q^+ \vdash \neg T\phi$$

y muchas otras. En general, cualquier oración que se obtenga diagonalizando fuertemente fórmulas $\phi(v) \in L_T$ sobre la variable v donde Tv ocurre en ϕ como una subfórmula m-referirá a sí misma. Dos casos paradigmáticos aparte del mentiroso son una versión formal fuerte de la oración del honesto, esto es,

$$t = pTtq$$

una oración que dice de sí misma que es verdadera, y una oración fuerte de Curry,

$$c = (pTcq \rightarrow c) \rightarrow c$$

que dice de sí misma que implica una falsedad.

Sin embargo, versiones alternativas del mentiroso y Curry, como

$$p \rightarrow (T : p) \rightarrow \neg p \quad (7.9)$$

que dice de si misma que es falsa, y

$$c \rightarrow (T(c) \rightarrow 0 = 0) \quad (7.10)$$

que dice de si misma que es verdad que implica una falsedad, que se obtienen aplicando el lema de diagonalizacion fuerte a las formulas $T : p \rightarrow \neg p$ y $T : p \rightarrow 0 = 0$, respectivamente, no resultan autorreferentes, porque, por ejemplo, en $T : p \rightarrow \neg p$, no se predica verdad de $T : p \rightarrow \neg p$ sino de $T : p$. Esto no va a ser fuente de problemas, como se vera mas adelante. Ademas, tiene cierto atractivo intuitivo porque, en lo que al predicado de verdad T concierne, (7.9) no dice nada acerca de la verdad de si misma sino de su negacion, al igual que (7.10), donde solo $T(c) \rightarrow 0 = 0$ cae bajo el alcance de T . Por otro lado, el proceso de diagonalizacion si arroja oraciones autorreferentes, aunque no las que se espera, i.e., $T : p \rightarrow \neg p$ y $T(c) \rightarrow 0 = 0$, en cada caso.

La definicion de m-referencia aleatica es muy semejante a la m-referencia a oraciones del apartado anterior. Para la q-referencia aleatica, en cambio, hay que hacer una serie de modificaciones adicionales porque, como ya se menciono, todavia no puede saberse la extension de T y las clausulas de la definicion 7.5 no ayudan. Lo unico que se mantiene invariable es que la q-referencia aleatica tambien esta cerrada bajo conectivas proposicionales.

En primer lugar, los cuantificadores superfluos ya no podran ser fuente de q-referencia, porque esta no se realiza a traves del predicado veritativo. Luego, una ocurrencia de un cuantificador $\forall v$ en una oracion ϕ generara instancias de q-referencia aleatica solo si v esta libre en la subformula de ϕ que le sigue inmediatamente. Si v no esta libre en ϕ , $\forall v \phi$ y ϕ referiran aleaticamente a lo mismo. No sera sensato ir mas all y pedir que $\forall v \phi$ est bajo el alcance de T en ϕ , por casos como $\forall x(x = 1 \rightarrow \forall y(y = x \rightarrow Ty))$, donde la variable x esta realizando cierto trabajo para que la oracion referiera al mentiroso, pero no cae bajo el alcance de T .

Segundo, oraciones de la forma $\forall v_1 \dots \forall v_n T \phi$ no referiran ya a todas las oraciones de L_T sino solo a aquellas que 'caigan bajo el alcance' de T , esto es, a todas las oraciones denotadas por terminos $t[k_1 = v_1] \dots [k_n = v_n]$, donde $k_1 \dots k_n \in L_T$. $\forall x T x$, por ejemplo, q-referira aleaticamente a todo, pero $\forall x T \neg x$ solo a negaciones. Al igual que las oraciones de \forall dadas por

$$q^+(n) = \forall x(x > n \rightarrow T(x)) \quad (7.11)$$

$$q^+(n) = p8y(9z(y = (z)^z > n) ! :T y)q; \quad (7.12)$$

expresiones como $8xT :x$ pueden pensarse como $8x(9y x = :y ! T x)$, donde los cuantificadores estan relativizados a un predicado.

Lo mismo vale para oraciones de la forma $8v_1 : : v_n$; solo que para estas es un poco mas complejo. $[k_1=v_1] : : [k_n=v_n]$ puede contener un termino 'nuevo', que no estaba ya en ϕ , que denote una oracion, pero este termino puede no estar bajo el alcance de T , con lo cual no cuenta. Por ejemplo, no sera correcto decir que $8x(:x = :x \wedge T(x!x))$ q-re ere aleticamente a todas las negaciones sino solo a los condicionales, aunque toda oracion ϕ es denotada por un termino $:p'q$ que ocurre en $(:x = :x \wedge T(x!x))[p'q=x]$.

Ademas, puede ocurrir que $[k_1=v_1] : : [k_n=v_n]$ re era aleticamente a una oracion, no por mencion sino por cuantificacion. Por ejemplo, en $8x:8y(T(x!y)), :8y(T(p'q!y))$ no m-re ere a 'ni a nada pero q-re ere a todas las oraciones condicionales. Luego, lo mejor es adoptar una clausula recursiva para expresiones dadas por cuantificadores seguidos por negaciones. Grosso modo, $8v_1 : : v_n$ q-referira aleticamente a todo aquello que $8v_1 : : v_{i-1}v_{i+1}v_n: [k=v_i]$, donde $k \geq 2$!, re era. Esta es la version aletica de la cuantificacion irrestricta.

Por ultimo, con respecto a oraciones de la forma $8v_1 : : v_n(!)$, solo se puede hablar de cuantificacion restringida mediante formulas que no contengan T , porque para las que lo contienen no se sabe en principio todavia que objetos las satisfacen. Luego, para evitar riesgos, si el predicado veritativo aparece tanto en ϕ como en ψ , la expresion sera tratada como si fuera una negacion, con cuantificacion irrestricta; pero, si solo aparece en el antecedente o solo en el consecuente, hay posibilidad de restringir la referencia. Si T ocurre solo en ψ , en principio $8v_1 : : v_n(!)$ referir a unicamente a aquellas oraciones que satisficieran ψ , viceversa, si T ocurre solo en ϕ , la oracion q-referir a en principio unicamente a aquellas oraciones que satisficieran la negacion del consecuente.

Esto funciona muy bien para expresiones de la forma $8v(!Tv)$, donde T no ocurre en ψ y se aplica a la misma variable sobre la cual se esta cuantificando. Sin embargo, en el caso de expresiones de la forma $8v(!Tf(v))$, donde tampoco T ocurre en ψ y f es un simbolo de funcion del lenguaje, debera decirse, como se hizo en la definicion 7.5, que tambien referen a las oraciones codificadas por imagenes a traves de la funcion representada por f de los numeros que satisfacen ψ . De hecho, si se habla unicamente de referencia aletica, en analogia con lo dicho en dos parrafos atras, solo estas ultimas relaciones de referencia son relevantes, y no aquellas que ligan $8v(!Tf(v))$ con oraciones meramente porque sus codigos satisfacen ψ . Luego, se dira

que, si T ocurre en ϕ y no en ψ , $\exists v_1 \dots v_n (\phi \supset \psi)$ q-re ere aleticamente a lo que m-re ere cada $[k_1 = v_1] \dots [k_n = v_n]$, donde $k_1, \dots, k_n \in \{!, \neg\}$ y N^+ $[k_1 = v_1] \dots [k_n = v_n]$. Pero, al igual que en el caso de universales seguidos de negaciones en el parrafo anterior, se

deber a considerar que, para estos k_1, \dots, k_n , $[k_1 = v_1] \dots [k_n = v_n]$ puede, no solo m-referir aleticamente a otras oraciones sino tambien q-referir. Por eso, nuevamente, se recurrira a una clausula recursiva. De hecho, doy una de nicion mas general, cuanti cada y menos esquematica, en terminos de cuanti cadores existenciales, como en la de nicion 7.5 de q-referencia del apartado anterior.

Definicion 7.12 (Q-referencia aletica semantica) . Dadas dos oraciones $\phi, \psi \in \mathcal{L}_T$, ϕ q-re ere aletica y semanticamente a ψ sii T ocurre en ϕ y se cumple una de las condiciones 1-3:

$$1. \phi := \exists v_1 \dots v_n \psi$$

a) $\psi := \exists t \phi$ o $\psi := \forall t \phi$, para algunos $k_1, \dots, k_n \in \{!, \neg\}$, $[k_1 = v_1] \dots [k_n = v_n]$ q-re ere aletica y semanticamente a ψ , o contiene una ocurrencia de Ts como suboracion que no ocurre en ψ tal que $\phi^+ \setminus s = \psi$; o

$$b) \psi := \neg \phi$$

i. ϕ y ψ contienen T y, para algunos $k_1, \dots, k_n \in \{!, \neg\}$, $[k_1 = v_1] \dots [k_n = v_n]$ q-re ere aletica y semanticamente a ψ , o contiene una ocurrencia de Tt como suboracion que no ocurre en ψ tal que $\phi^+ \setminus t = \psi$, o

ii. solo contiene T y existen $k_2 \in \{!, \neg\}$ i n tales que

$$N^+$$

$$\exists v_1 \dots v_i \neg v_{i+1} \dots v_n [k = v_i]$$

y $\exists v_1 \dots v_i \neg v_{i+1} \dots v_n (\neg \phi)$ $[k = v_i]$ q-re ere aletica y semanticamente a ψ , o contiene una ocurrencia de Tt como suboracion que no ocurre en ψ tal que $\phi^+ \setminus t = \psi$, o

iii. solo contiene T y existen $k_2 \in \{!, \neg\}$ i n tales que

$$N^+$$

$$\exists v_1 \dots v_i \neg v_{i+1} \dots v_n: [k = v_i]$$

y $\exists v_1 \dots v_i \neg v_{i+1} \dots v_n (\neg \phi)$ $[k = v_i]$ q-re ere aletica y semanticamente a ψ o contiene una ocurrencia de Tt como suboracion que no ocurre en ψ tal que $\phi^+ \setminus t = \psi$.

$$2. \phi := \exists \psi \text{ q-re ere aletica y semanticamente a } \psi.$$

$$3. \phi := \neg \psi \text{ o q-re ere aletica y semanticamente a } \psi.$$

Luego, si en $\exists v_1 : \dots : v_n \ T$ no ocurre aplicado a ningún término abierto, toda

oración de la forma $T \ t$ que ocurra en $[k_1 \dots = v_1] : \dots : [k_n = v_n]$ ya habrá ocurrido en t y no dará lugar, por ende, a una relación de q-referencia; $\exists v_1 : \dots : v_n$ podrá m-referir pero no q-referirá aleticamente a nada. Esto garantiza la inocuidad de los cuantificadores superfluos.

La oración del mentiroso que se obtiene en pat aplicando el lema de diagonalización débil a $\exists T v$, i.e.,

$$\text{pat} \quad \exists :T p q;$$

dada por

$$\exists u (u = p \exists v (\text{Diag}(u; v) ! :T v) q ! \exists v (\text{Diag}(u; v) ! :T v))$$

q-referirá aleticamente a sí misma, porque, por la cláusula 1.b)ii de la definición [7.12](#), q-referirá a todo lo que

$$\exists v (\text{Diag}(p \exists v (\text{Diag}(u; v) ! :T v) q; v) ! :T v)$$

q-referirá, que, por la misma cláusula y la definición del predicado Diag , q-referirá a todo lo que

$$\exists :T p q$$

q-referirá. Y, en general, las oraciones que se obtienen al diagonalizar débilmente sobre un predicado que contiene $\exists T v$ como una subfórmula resultan q-autorreferenciales.

En cambio, aquellas en las cuales esto no sucede, pueden no ser autorreferenciales. Basta con considerar casos análogos a las oraciones [\(7.9\)](#), [\(7.10\)](#) donde el lema de diagonalización débil se aplica en lugar del fuerte. Nuevamente, esto no conlleva grandes inconvenientes y es coherente con la noción de referencia a través del predicado de verdad.

Definición 7.13 (Referencia directa aletica semantica) . Dadas dos oraciones $'_1, ' _2 \in L_T$, $'_1$ q-referirá directa, aletica y semanticamente a $'_2$ si m-referirá o q-referirá directa, aletica y semanticamente a $'_2$.

Definición 7.14 (Cadenas de referencia directa aletica semantica). Una secuencia de oraciones $\sigma; : \dots ; \sigma_n \in L_T, n \geq 1$, es una cadena de referencia directa aletica semantica si, para cada $i < n$, σ_i q-referirá directa, aletica y semanticamente a σ_{i+1} .

Definición 7.15 (Referencia aletica semantica) . Sean $'_1$ y $'_n$ oraciones de L_T . $'_1$ q-referirá aletica y semanticamente a $'_n$ si existe una cadena de referencia semantica aletica $\sigma; : \dots ; \sigma_n$ tal que $\sigma_1 = ' _1$ y $\sigma_n = ' _n$.

A partir de las nociones dadas se pueden definir autorreferencia y buena fundación aleticas.

Definición 7.16 (Autorreferencia aletica semantica). Sea ϕ una oracion de L_T . ϕ es autorreferente aletica y semanticamente si y solo si ϕ es aletica y semanticamente a si misma.

Definición 7.17 (Buena fundacion aletica semantica). Sea ϕ una oracion de L_T . ϕ es aletica y semanticamente bien fundada si y solo si ninguna cadena de referencia aletica y semantica que empiece en ϕ es indefinidamente extensible.

Las oraciones de la secuencia de Yablo, tanto en su formulacion original en (7.11) como en su formulacion alternativa (7.12), resultan aleticamente mal fundadas pero no autorreferentes. En el primer caso, cada termino (n) denota una oracion de la forma

$$\exists x(x > n \rightarrow \neg T(x))$$

que, por la definicion 7.11, no es aleticamente a nada, y por la definicion 7.12, es a todas

aquellas oraciones denotadas por (k) tales que $k > n$, es decir, cada oracion de la lista es directamente solo a las oraciones que se encuentran mas adelante en la lista. Luego, por la definicion 7.15, es simpliciter a estas oraciones exclusivamente. Y por la definicion 7.16, ninguna de las oraciones de la formulacion original de la lista de Yablo es aleticamente autorreferente. Por otro lado, como todo segmento inicial de la lista de Yablo (7.11) es una cadena de referencia aletica, todos sus miembros son mal fundados de acuerdo con la definicion 7.17.

Con respecto a la formulacion alternativa en (7.12) se puede decir lo mismo: cada termino (n) denota la oracion

$$\exists y(\exists z(y = (z) \wedge z > n) \rightarrow \neg T(y))$$

que no es aleticamente a nada y es solo a aquellas oraciones a las cuales cada $T(k)$ es aleticamente, donde $k \in \mathbb{N}^+$ $\exists z(k = (z) \wedge z > n)$, i.e., a todas aquellas denotadas por k , donde $k = (z) \wedge z > n$.

Las nociones presentadas tambien dan un veredicto intuitivamente correcto sobre otras secuencias de oraciones semejantes a la de Yablo, que tambien dan lugar a inconsistencias junto con sus correspondientes instancias del Esquema-T. Por ejemplo, las siguientes oraciones conforman lo que se conoce como 'lista dual de Yablo' (cf. Cook

(2014a):

$$:\exists x(x > 0 \wedge \neg T(x)) \tag{7.13}$$

$$:\exists x(x > 1 \wedge \neg T(x))$$

$$:\exists x(x > 2 \wedge \neg T(x))$$

...

donde cada (n) con $n \geq 2$

(n) denota la n-esima oracion en la secuencia. Estas pueden obtenerse en q^+ mediante un argumento diagonal al igual que las oraciones de Yablo. Intuitivamente cada una de las oraciones en la lista dice que alguna de las expresiones que se encuentran mas adelante no es verdadera. Siendo que el patron de referencia es el mismo que el que subyace a la lista de Yablo, cada una de estas oraciones es mal fundada pero no autorreferencial.

Tambien es posible formular secuencias de oraciones alternativas, con patrones de referencia diferentes, dando lugar a \neg -inconsistencias, como la siguiente:

$$:\exists x(x > 0 \wedge \neg \exists y(y > x \wedge T(y))) \tag{7.14}$$

$$:\exists x(x > 1 \wedge \neg \exists y(y > x \wedge T(y)))$$

$$:\exists x(x > 2 \wedge \neg \exists y(y > x \wedge T(y)))$$

...

2

donde cada (n) con $n \geq 1$ denota la n-esima oracion de la lista. Esta se refiere directamente a lo que cada $(y > k \wedge T(y))$ donde $k > n$ se refiere, esto es, a todas las oraciones $T(m)$ donde $m > k > n$. Luego, cada oracion de la lista se refiere a todas las que se encuentran dos o mas lugares adelante, pero no a la que le sigue inmediatamente. Por ende, son tambien mal fundadas y no autorreferenciales.

Apelando a la proposicion [1.11](#) es posible tambien generar \neg -cadenas de oraciones que contengan T tales que cada una m-refiere aleticamente a la que le sigue en la cadena. Por ejemplo, la siguiente es una cadena de honestos:

$$t_0 = pT t_1q \tag{7.15}$$

$$t_1 = pT t_2q$$

$$t_2 = pT t_3q$$

...

Por la definicion [7.15](#) de referencia aleetica semantica, cada oracion $T t_n$ se refiere, no solo a

la que le sigue, sino a todas las que estan por debajo suyo y, como todo segmento inicial

de esta cadena es una cadena de referencia por la definicion 7.14, todas sus oraciones son mal fundadas. Pero, a diferencia del veredicto dado por la nocion de referencia directa no aletica de la seccion anterior, como se hizo para la cadena de oraciones (7.3) de L_{pa}^+ , ninguna de estas expresiones es autorreferente, porque solo importan los terminos $t_i, i \geq 1$, no sus subterminos.

Para otras oraciones como $\exists x T f(x; p_0 = 0q)$, donde $f(0; p'q) := p'q$ y $f(S(x); p'q) := p T \text{ num}(f(x; p'q)) q$ (f antepone x veces el predicado de verdad a ') y la oracion de McGee, la referencia aletica tambien se corresponde con lo que dicta la intuicion. $\exists x T f(x; p_0 = 0q)$ dice que todas las iteraciones de T sobre $0 = 0$ son verdaderas. De acuerdo con la clausula 1.a) de la definicion 7.12 de q-referencia, refiere a todas aquellas oraciones denotadas por $f(k; p_0 = 0q)$ para cada $k \geq 1$, i.e., a $0 = 0$, denotada por $f(0; p_0 = 0q)$, a $T p_0 = 0q$, denotada por $f(1; p_0 = 0q)$, a $T p T p_0 = 0qq$, denotada por $f(2; p_0 = 0q)$, etc. Cada una de estas oraciones refiere directamente a un numero finito de oraciones, v.g., $T p T p_0 = 0qq$ refiere a $T p_0 = 0q$ y, por transitividad, a $p_0 = 0q$. Luego, no existe una cadena de referencia indecididamente extensible que comience en $\exists x T f(x; p_0 = 0q)$: es una oracion bien fundada.

Por su parte, la oracion $\exists x T f(x; m)$ de McGee, i.e.,

$$q^+ \text{ ` } m = p : \exists x T f(x; m)q \tag{7.16}$$

que bajo ciertas condiciones da lugar a ω -paradojas, como se vio en el apartado 3.3, refiere tambien a las oraciones denotadas por $f(k; m)$ para cada $k \geq 1$ y, por ende, refiere a si misma, es mal fundada.

La nocion de buena fundacion dada por la definicion 7.17 no coincide con las nociones de fundacion de Kripke (1975) o Leitgeb (2005) (veanse los apartados 3.1 y 3.2), aunque no esta demasiado alejada. Por ejemplo, $\exists x (T x \rightarrow T x)$ refiere aleticamente a toda oracion, pero esta fundada en el sentido de Leitgeb, porque es verdadera en toda extension de N^+ . Asimismo, si $t = p t = t T t q$ se obtiene en q^+ diagonalizando $x = x T x$, refiere aleticamente a si misma y, por tanto, esta mal fundada, pero esta fundada tanto para Kripke como para Leitgeb. En general, y a diferencia de muchos otros abordajes, ac todas las oraciones autorreferentes, directa o indirectamente, resultan mal fundadas, porque es posible producir para ellas cadenas de referencia aletica repetitivas indecididamente extensibles. La autorreferencia es, desde este punto de vista, un tipo especial de mala fundacion.

Observacion 7.18. Si ' es una oracion aleticamente autorreferente de L_T , ' no esta aleticamente bien fundada.

7.3.1.2. Indenibilidad y axiomatización

Al estar de nida en terminos del predicado de verdad-en- N^+ , la q-referencia aletica semantica no es de nible en L_{pa}^+ y, por tanto, tampoco la referencia aletica semantica directa, como muestro en este apartado. Luego, estas nociones no puede emplearse di-rectamente en la formulacion de una teor a axiomatica de la verdad en L_T . Es preciso introducir en el lenguaje un predicado primitivo binario $R_q^S(u; v)$ para expresarla y luego incorporar axiomas a pat que gobiernen su comportamiento. En este apartado ofrezco una axiomatización con estas caracter sticas.

Dada una codi cación de las expresiones de L_T con numeros naturales, la nocion de m-referencia aletica de u a v puede de nirse directamente en L_{pa}^+ mediante la siguiente expresion:⁷

$$\text{Sent}_{L_T}(u \rightarrow v) \wedge \exists t < u (\text{SubF orm}(u; T t) \wedge t = v) \quad (7.17)$$

p . q

En castellano, tanto u como v son oraciones de L_T |porque el condicional que tiene a u como antecedente y a v como consecuente lo es| y existe un termino cerrado t que denota v tal que $T t$ es una subformula de u. La noto $R_m(u; v)$. Al ser una expresion τ (vease el apartado [1.2.5](#)), R_m representa la m-referencia aletica directa en q^+ y esta es, por ende, una nocion recursiva.

Si la referencia aletica semantica directa fuera de nible en L_{pa}^+ mediante una formula $R^S(u; v)$, la q-referencia tambien lo ser a, por medio del predicado $R_q^S(u; v) := R^S(u; v) \wedge \exists R_m(u; v)$. En vista de la de nición [7.12](#) de q-referencia aletica semantica, este predicado deber a satisfacer los siguientes principios:

$$\text{qrs1 } \exists x; y (\text{Sent}_{L_{pa}}(x) \rightarrow R_q^S(x; y))$$

$$\text{qrs2 } \exists x; y (\text{At}_{L_T}(x) \rightarrow R_q^S(x; y))$$

$$\text{qrs3 } \exists x; y; z; u (\text{U}(x; z; T, u) \wedge \text{Sent}_{L_T}(y) \rightarrow (R_q^S(x; y) \rightarrow \exists w (\text{lg}(w) = \text{lg}(z) \wedge \exists t (\text{SubF orm}(\sim s.(T, u(p0q=t); w); T t) \wedge t = y))))))$$

$$\text{qrs4 } \exists x; y; z; w (\text{U}(x; z; \cdot, u) \wedge \text{Sent}_{L_T}(y) \rightarrow (R_q^S(x; y) \rightarrow \exists w (\text{lg}(w) = \text{lg}(z) \wedge (R_q^S(\sim s.(\cdot, u; w); y) \rightarrow \exists t (\text{SubF orm}(\sim s.(\cdot, u(p0q=t); w); T t) \wedge t = y))))))$$

$$\text{qrs5 } \exists x; y; z; u; v (\text{U}(x; z; u!v) \wedge \text{Sent}_{L_T}(y) \rightarrow (\text{F orm}_T(u) \wedge \text{F orm}_{L_{pa}}(v) \rightarrow (R_q^S(x; y) \rightarrow \exists w (\text{lg}(w) = \text{lg}(z) \wedge (R_q^S(\sim s.(u!v; w); y) \rightarrow \exists t (\text{SubF orm}(\sim s.(u!v.(p0q=t); w); T t) \wedge t = y))))))$$

⁷Vease el apartado [1.2.3](#) para detalles de notacion.

$$\begin{aligned}
 \text{qrs6 } & \exists y (F \text{orm}_{L_{pa^+}}(pq) \wedge :F \text{orm}_{L_{pa^+}}(p \ q) \wedge \text{Sent}_{L_T}(y) \wedge (R_q^S(p \delta v_1 \dots v_n(\cdot)q); y) \wedge \\
 & \exists x_{i=1}^n (\exists v_1 \dots v_i \exists v_{i+1} \dots v_n [x=v_i] \wedge (R_q^S(p \delta v_1 \dots v_i \exists v_{i+1} \dots v_n(\cdot)q(x=p \delta v_i q); y) \wedge \\
 & \exists t (\text{SubF orm}(\sim s(x(\cdot) \delta t); u); T t) \wedge t = y)) \\
 \text{qrs7 } & \exists x; y; z; w; u; v (U(x; z; u; \cdot; v) \wedge U(w; z; \cdot; v; \cdot; u) \wedge (R_q^S(x; y) \wedge R_q^S(w; y))) \\
 \text{qrs8 } & \exists x; y (R_q^S(\cdot; x; y) \wedge R_q^S(x; y)) \\
 \text{qrs9 } & \exists x; y; z (R_q^S(x; \cdot; z; y) \wedge R_q^S(x; y) \wedge R_q^S(z; y))
 \end{aligned}$$

donde δ \in 2^{L_T} tiene como variables libres a $v_1; \dots; v_n$.

'qrs' es un acrónimo de 'Q-Referencia Semántica'. qrs1 indica que las oraciones del lenguaje de la aritmética no q-referen, mientras que qrs2 establece lo mismo para fórmulas atómicas. Los axiomas qrs3 y qrs4 se corresponden con la cláusula 1.a) de la definición 7.12 de q-referencia, qrs5 y qrs6 con 1.b)i y ii, respectivamente, qrs7, con la ayuda de qrs6, garantiza la cláusula 1.b)ii y qrs8 y qrs9 formalizan las condiciones 2 y 3, correspondientemente. A modo de aclaración, el predicado $t(\text{SubF orm}(\sim s(x(\cdot) \delta t); u); T t) \wedge t = y)$ establece que hay un término cerrado t tal que sustituir t por δ en (la fórmula) x y luego sustituir todos los argumentos de la secuencia u por cada una de las variables libres en esta fórmula da por resultado una fórmula que tiene a $T t$ como subfórmula y t denota (la oración) y . En otras palabras, t ocurre en la sustitución de las variables libres en x por los miembros de u como resultado de esta sustitución, no estaba antes. Asimismo, las fórmulas de la forma $\exists x_{i=1}^n (\exists v_1 \dots v_i \exists v_{i+1} \dots v_n [x=v_i] \wedge \dots)$ son abreviaturas de $\exists x((\exists v_2 \dots v_n [x=v_1] \wedge \dots) \wedge (\exists v_1 \dots v_{i-1} [x=v_n] \wedge \dots))$

No obstante, si una teoría formulada en L_T satisface estas cláusulas, será inconsistente.

Proposición 7.19 (Indenibilidad de la q-referencia aritmética semántica). Si $th \ q^+$ es una teoría formulada en L_T y este lenguaje contiene un predicado $R_q^S(u; v)$ que satisface qrs6 en th , th es inconsistente.

Demostración. Aplicando el lema de diagonalización fuerte (teorema 1.8) a la fórmula $:R_q^S(\delta.pvq(u; \cdot; pT \neg q); p'q)$, donde δ es una oración de L_{pa^+} , se obtiene un término t tal que

$$q^+ \cdot t = p : R_q^S(\delta.pvq(t; \cdot; pT \neg q); p'q) \tag{7.18}$$

Intuitivamente, $\text{R}_q^S(8.pvq(t!.pT.vq); p'q)$ es una oración que dice de sí misma que 'no la satisface'; en una suerte de oración del mentiroso. Instanciando qrs6 en

$$8v(:\text{R}_q^S(8.pvq(t!.pT.vq); p'q) ! T v)$$

se obtiene en th

$$\text{R}_q^S(p8v(:\text{R}_q^S(8.pvq(t!.pT.vq); p'q) ! T v)q; y) \$ 9x(:\text{R}_q^S(8.pvq(t!.pT.vq); p'q) ^ 9t(\text{SubF orm}(pT.xq; T t) ^ t = y))$$

e, instanciando y con ',

$$\begin{aligned} \text{R}_q^S(p8v(:\text{R}_q^S(8.pvq(t!.pT.vq); p'q) ! T v)q; p'q) \$ 9x(:\text{R}_q^S(8.pvq(t!.pT.vq); p'q) ^ 9t(\text{SubF orm}(pT.xq; T t) ^ t = p'q)) \\ \$:\text{R}_q^S(8.pvq(t!.pT.vq); p'q) \\ ^ 9t(t = p'q) \\ \$:\text{R}_q^S(8.pvq(t!.pT.vq); p'q) \end{aligned}$$

lo cual es una contradicción, en virtud de (7.18). □

Luego, como anticipo, no existe una fórmula de L_{pa}^+ que de na la q-referencia aletica semantica y, por ende, tampoco la referencia aletica semantica directa. Mediante un razonamiento analogo es sencillo probar resultados semejantes para las nociones de q-referencia a numeros y q-referencia a oraciones dadas por las de niciones 7.2 y 7.5, respectivamente.

Luego, si se desea un abordaje axiomático de la referencia aletica semantica, sera preciso incorporar el símbolo de predicado R_q^S como un primitivo junto con axiomas para que regulen su comportamiento. Sea $L_R = L_T [f\text{R}_q^S g$. Los candidatos obvios a axiomas son los mismo principios qrs1-qrs9, que ya no daran lugar a inconsistencias porque el esquema qrs6 solo admite instancias dadas por formulas de L_T , que no contienen el nuevo símbolo de predicado R_q^S . La siguiente es una axiomatización esquemática de la q-referencia aletica semantica.⁸

Definición 7.20 (qrs). qrs L_R es la teoría que extiende el resultado de formular la aritmetica de primer orden en L_R , con todas las instancias del esquema de axioma de inducción dadas por formulas de este lenguaje, mediante los axiomas qrs1-qrs9.

⁸Como el predicado 'N' no es traducible a ningún lenguaje de primer orden, existen dos modos de expresar las cláusulas de la definición 7.2 en una extensión de L_{pa}^+ : o bien se recurre a esquemas, como en qrs6, o bien al predicado de verdad T. Dado que el objetivo es axiomatizar T en función de R_q^S y no a la inversa, como todavía no se sabe nada de T, tomo el primer camino. La desventaja principal es que la teoría resultante agrega infinitos axiomas para R_q^S .

Proposición 7.21. qrs es !-consistente.

Demostración. Siendo los axiomas de qrs formalizaciones directas de las cláusulas de la definición recursiva de q-referencia alética semántica, el modelo $\langle \mathcal{N}^+, ?; R_i \rangle$, que extiende \mathcal{N}^+ a todo L_R asignando el conjunto $\mathcal{?}$ a T y R a R_q^S como sus interpretaciones, tal que $R = fh\#('); \#()i : ';$ $\forall L_T$ y ' q-re ere alética y semánticamente a \mathcal{g} es un modelo de qrs. □

A partir de la definición de m-referencia en (7.17) y la axiomatización de la q-referencia alética semántica dada por la teoría qrs, puede definirse la referencia alética semántica directa en L_R mediante la fórmula

$$R_d^S(u; v) := R_m(u; v) _ R_q^S(u; v)$$

$R_d^S(u; v)$ se aplica a (los códigos de) dos oraciones $u; v$ |i.e., u re ere alética, semántica y directamente a v | si y solo si u m-re ere o q-re ere alética, semántica y directamente a v .

Siendo la relación de referencia semántica la unión de la m-referencia y la q-referencia y dado que la primera es recursiva, la referencia hereda la complejidad de la q-referencia. Bajo el modelo estándar de qrs de nido en la prueba de la proposición 7.21 pueden armarse muchos más casos de referencia que los que esta teoría informa. Por ejemplo, la oración $\exists x(\text{Bew}_{qrs}(x) ! Tx)$ q-re ere semánticamente a $0 \neq 0$, porque qrs es una teoría consistente y, por ende, no prueba $0 \neq 0$, pero, como indica el segundo resultado de incompletitud de Gödel (teorema 1.13), qrs no tiene modo de saber esto último, precisamente porque es consistente.

Notese que la noción de q-referencia que he dado esta definida exclusivamente para oraciones de L_T , so pena de trivialidad. Sin embargo, una vez axiomatizada, parece natural extenderla a oraciones del lenguaje de la nueva teoría, esto es, L_R . Después de todo, las oraciones que contienen el predicado R_q^S también pueden q-referir. Intuitivamente, v.g., $\exists x(R_q^S(p \exists x T : xq; x) ! Tx)$ q-re ere a todo aquello a lo que $\exists x T : x$ q-re ere, esto es, a todas las oraciones que son negaciones.

Como muestra el teorema de indecibilidad, extender la axiomatización dada en 7.20 a oraciones del lenguaje completo de la teoría no es sencillo. Si bien puede fácilmente extenderse qrs1 a fórmulas de L_R $\forall T \mathcal{g}$, qrs2 a fórmulas atómicas de la forma $R_q^S(t; s)$ y relativizar los axiomas qrs3-qrs5 y qrs7-qrs9 a oraciones de L_R en lugar de L_T , el esquema de axioma qrs6 requiere más trabajo; extenderlo simplemente a fórmulas y pertenecientes a todo L_R lleva a contradicción.

Un modo relativamente simple de extender los axiomas de qrs a oraciones de L_R consiste en establecer una jerarquía de lenguajes y teorías. En qrs6 es preciso que las oraciones a la derecha del bicondicional no contengan el predicado de referencia que se encuentra a la izquierda. Luego, se precisa una secuencia de predicados $R_q^0; R_q^1; \dots; R_q^n; \dots$, que incluso podría extenderse transitivamente, postulando un predicado R_q para cada ordinal α . Un segundo camino, que evita las jerarquías, consiste en restringir el axioma qrs6 a oraciones 'no problemáticas', como pretendo hacer con el [Esquema-T](#) en el capítulo 8. Pero se supone que la misma noción de referencia era la que iba a hacer este trabajo. Este camino parece llevar a un regreso infinito.

En definitiva, el problema de la referencia semántica es al menos tan complejo como el de la verdad y, en alguna medida, uno se reduce al otro. La noción es permeable a paradojas prácticamente idénticas a la del mentiroso, como se vio en la proposición [7.19](#). Por tanto, aunque tiene interés en sí mismo, no parece tener demasiado sentido investigar la referencia alética semántica para dar una solución al problema de la verdad.

Por otro lado, como argumenté en la sección [5.3](#), si lo que se desea es emplear una noción de referencia en la construcción de teorías de acionistas de la verdad, la noción es preferentemente de nivel en el lenguaje de la aritmética. Pero la proposición [7.19](#) establece lo contrario; la noción de referencia alética semántica es demasiado compleja, como lo son la fundación de [Kripke \(1975\)](#) y la dependencia de [Leitgeb \(2005\)](#) (veanse los apartados [3.1](#) y [3.2](#)).

En la sección que sigue ofrezco nociones análogas a la q-referencia y referencia alética semántica pero reemplazando la noción semántica de verdad-en- N^+ por la noción sintáctica de teoremicidad. Naturalmente, el resultado son nociones mucho más simples, de niveles en el lenguaje de la aritmética pero, a la vez, menos generales, dependientes de la teoría a la cual la noción de teoremicidad que se está utilizando corresponda.

7.3.2. Referencia alética sintáctica

Una familia de conceptos similares a los de la sección anterior pero más simples, que llamo 'sintácticos', puede darse reemplazando de la noción de verdad en el modelo estándar en las cláusulas 1.b)ii-iii de la definición [7.12](#) de q-referencia alética semántica por una noción de prueba en una teoría. Por supuesto, las nociones resultantes van a depender, ya no solo del lenguaje para el cual están formuladas, sino también de esta teoría a cuya noción de prueba se ha utilizado en la definición.

Prima facie, se estaría ignorando una inmensa cantidad de casos de referencia que 'suceden' en el modelo estándar de la teoría sobre la cual se trabaja, pero que esta

es incapaz de probar; más adelante se verán algunos ejemplos. Sin embargo, si se piensa desde el punto de vista de la teoría, que es lo que importa a la hora de analizar paradojas, porque estas, para serlo, deben dar lugar a una inconsistencia o a afirmaciones incorrectas en la teoría, no queda claro que sea correcto afirmar esas relaciones de referencia que solo tienen lugar en algunos pero no en todos sus modelos. En cualquier caso, espero, se juzgar a los conceptos que aquí introduzco por su utilidad más que por los detalles de su formulación.

En el apartado siguiente introduzco nociones de q-referencia, referencia directa, referencia, autorreferencia y buena fundación aritméticas sintácticas y muestro, respectivamente, que todas ellas son definibles en el lenguaje de la aritmética de primer orden. En [7.3.2.2](#) señalo ciertos inconvenientes que estos conceptos sintácticos traen consigo y ofrezco una solución posible. Finalmente, en [7.3.2.3](#) introduzco una noción que permite aislar los inconvenientes señalados en el apartado precedente.

7.3.2.1. Nociones básicas, de definibilidad y ejemplos

Sea \mathcal{T} cualquier teoría formulada en $L_{\mathcal{T}}$ que extienda Pat con axiomas pertenecientes a L_{pa}^+ que sean verdaderos en \mathbb{N}^+ . Notese que la m-referencia aritmética no precisa una definición en términos de teoría de la prueba porque, tal como se la presento en la definición [7.11](#), ya es un concepto plenamente no semántico. Las nociones sintácticas (en el sentido de 'teoría de la prueba', 'proof-theoretic') que presento a continuación deberán ser acompañadas por este calificativo, que omito por cuestiones de legibilidad. Por ejemplo, en lugar de hablar de q-referencia aritmética sintáctica hablo de q-referencia aritmética a secas para referirme a ella.

Definición 7.22 (Q-referencia aritmética). Dadas dos oraciones $\phi, \psi \in L_{\mathcal{T}}$, ϕ q-referencia

aritméticamente a ψ en \mathcal{T} si y solo si ocurre en ϕ y se cumple una de las condiciones 1-3:

$$1. \phi := \exists v_1 \dots \exists v_n \psi \text{ y}$$

a) $\phi := \forall t \psi$ o $\phi := \exists t \psi$, para algunos $k_1, \dots, k_n \in \mathbb{N}^+$, $[k_1=v_1] \dots [k_n=v_n]$ q-referencia aritméticamente a ψ , o contiene una ocurrencia de \forall o \exists como suboración que no ocurre en ψ tal que $q^+ \text{ ` } s = p \text{ ` } q; \psi$

$$b) \phi := \neg \psi \text{ y}$$

i. tanto ϕ como ψ contienen \forall y, para algunos $k_1, \dots, k_n \in \mathbb{N}^+$, $[k_1=v_1] \dots [k_n=v_n]$ q-referencia aritméticamente a ψ o contiene una ocurrencia de \forall como suboración que no ocurre en ψ tal que $q^+ \text{ ` } t = p \text{ ` } q, \psi$

ii. solo contiene T y existen $k \geq 1$ índices tales que

$$\text{th} \vdash \exists v_1 \dots v_i \exists v_{i+1} \dots v_n [k=v_i]$$

o contiene

y $\exists v_1 \dots v_i \exists v_{i+1} \dots v_n (\neg [k=v_i])$ q-re ere aleticamente a

una ocurrencia de Tt que no ocurre en tal que $q^+ \vdash t = p, q$, o

iii. solo contiene T y existen $k \geq 1$ índices tales que

$$\text{th} \vdash \exists v_1 \dots v_i \exists v_{i+1} \dots v_n : [k=v_i]$$

o contiene

y $\exists v_1 \dots v_i \exists v_{i+1} \dots v_n (\neg [k=v_i])$ q-re ere aleticamente a

una ocurrencia de Tt que no ocurre en tal que $q^+ \vdash t = p, q$.

2.' := : y q-re ere aleticamente a .

3.' := ! y o q-re ere aleticamente a .

Al igual que la q-referencia semantica, la q-referencia aletica sintactica puede

extenderse naturalmente a un lenguaje con conjunciones y existenciales como símbolos primitivos. Asimismo, es capaz de dar cuenta de la autorreferencialidad de todas las formulas que se obtienen aplicando el lema diagonal debil (cf. teorema [1.7](#)) a una formula que contiene como subformula a Tv. Aplicando el lema diagonal debil a una formula '(v) con estas características, se obtiene que una oración tal que

$$\text{pat} \vdash \$(p, q);$$

cuya verdadera forma es

$$\exists u (u = p \exists v (\text{Diag}(u; v) \vdash '(v)) \wedge \exists v (\text{Diag}(u; v) \vdash '(v)))$$

Como q^+ prueba todas las identidades verdaderas, por la clausula 1.b)ii de la definición [7.22](#), esta oración q-re ere aleticamente a todo lo que

$$\exists v (\text{Diag}(p \exists v (\text{Diag}(u; v) \vdash '(v)) \wedge \exists v (\text{Diag}(u; v) \vdash '(v)))$$

re ere, que, por la misma clausula y el hecho de que el predicado Diag representa la relacion de diagonalizacion en q^+ (vease el apartado [1.2.3](#)), q-re ere a todo lo que '(p, q) re ere que, como contiene una ocurrencia de T p, q, re ere a .

En cambio, si Tv no ocurre en ', al igual que la noción semantica, la de noción [7.22](#) no garantiza que su diagonalizacion debil sea autorreferente, lo cual, como argument para la noción semantica, es intuitivamente correcto.

A diferencia de lo que ocurre en la definición 7.12 de q-referencia semántica,

ya no es preciso saber si un número x satisface en el modelo estándar de la teoría el antecedente de un condicional precedido por una tira de cuantificadores en una oración para decidir si la oración verdadera o no a un enunciado dado, sino que meramente hay que averiguar si la teoría prueba o no que el número satisface el antecedente. Dado que el conjunto de teoremas de una teoría axiomatizable es siempre semirecursivo, la q-referencia será asimismo una relación semirecursiva y, por tanto definible en L^+ para L_T y débilmente representable en q^+ mediante una fórmula φ . Porque es posible programar una máquina que tome dos argumentos x e y y responda 's' cuando tanto x como y codifiquen oraciones y x q-referencia aleticamente a y , y se cuelgue o responda 'no' cuando este no sea el caso.

La máquina recibe como entradas dos números naturales, x e y , y averigua primero si ambos codifican oraciones de L_T y si en x ocurre un término abierto bajo el alcance de T . En caso negativo, responde 'no' y para. En caso afirmativo, remueve los cuantificadores superfluos en cada componente proposicional de x en el cual ocurre un término abierto bajo el alcance de T y les asigna una etiqueta $[0]; [1]; \dots$ a cada una de las fórmulas resultantes, de acuerdo con el orden en el que aparecen en x .⁹ De ahora en más la máquina opera con las siguientes tres fases, comenzando por la fase 1 y la fórmula etiquetada $[0]$:

Fase 1: La máquina recibe una oración etiquetada $[j]$. Si $[j] := \forall v_1 \dots \forall v_n$, donde es una oración atómica, una negación o un condicional en el cual T ocurre tanto en el antecedente como en el consecuente, la máquina pasa a la fase 2 con (k_1, \dots, k_n) la primera n -tupla $(k_1, \dots, k_n) \in \mathbb{N}^n$ que no ha sido utilizada todavía para esta etiqueta (de acuerdo con una codificación dada de tuplas mediante números naturales), las variables v_1, \dots, v_n y la oración y . Si $[j] := \forall v_1 \dots \forall v_n(\varphi)$ donde T ocurre solo en (φ) , la máquina chequea si el primer par (k_1, k_2) que no ha sido aun considerado para esta etiqueta (nuevamente, de acuerdo con una enumeración $\langle k_i \rangle$) es tal que, para algún $1 \leq i \leq n$, m es el código

de una prueba de $\forall v_1 \dots \forall v_n \varphi$ [$k=v_i$]

$(\forall v_1 \dots \forall v_n \varphi : [k=v_i])$ en \mathcal{M} . Si no lo es, reinicia la fase 1 con la fórmula etiquetada $[j + 1]$, de existir; de lo contrario, con $[0]$. Si lo es, inicia la fase 2 con $\forall v_1 \dots \forall v_n \varphi$, m , v_i e y .

⁹ Los componentes proposicionales $\varphi_1, \dots, \varphi_n$ de una fórmula φ de L son las subfórmulas más cortas de φ tales que, si p es una letra proposicional, el resultado de sustituir en φ las fórmulas $\varphi_1, \dots, \varphi_n$ por p es una fórmula del lenguaje de la lógica proposicional. Intuitivamente, los componentes proposicionales de una fórmula φ son las subfórmulas que aparecen en una rama del árbol constructivo de φ en el primer nodo a partir del cual solo se agregan conectivas proposicionales. Son siempre fórmulas atómicas o fórmulas cuya conectiva principal es un cuantificador. Por ejemplo, los componentes de $(\exists x(0 =$

$0 \wedge (x = 1) \wedge (2 + 0 = 5))$ son $\exists x(0 \neq 0$

$\wedge (x = 1) \wedge (2 + 0 = 5)$. Listar los componentes proposicionales

de una fórmula dada en su orden de aparición es claramente un procedimiento computable.

Fase 2: La maquina recibe como entradas una formula ϕ , numeros k_1, \dots, k_n , variables

v_1, \dots, v_n y una oracion ψ , y chequea si $\exists [k_1=v_1 \dots k_n=v_n] \psi$ contiene una nueva ocurrencia de T , donde t es un termino cerrado, tal que $t = p \rightarrow q$. En caso negativo,

inicia la fase 3 con $\exists [k_1=v_1 \dots k_n=v_n] \psi$ como entradas. En caso afirmativo, responde "si" y para.

Fase 3: La maquina toma como entradas dos oraciones ϕ y ψ , remueve todos los cuan-

tificadores superfluos de cada componente proposicional de ϕ en el cual ocurre un termino abierto bajo el alcance de T y etiqueta los enunciados universales resultantes con $[n]; [n+1]; \dots$ de acuerdo con el orden de aparicion en ϕ , comenzando por el minimo numero no utilizado ya como una etiqueta y vuelve a la fase 1.

En las tres fases solo hay procesos recursivos involucrados. Si ϕ de hecho q-re ere a ψ , la maquina lo dira en algun momento. Si no lo hace puede deberse a razones diversas. Si es porque ϕ no contiene cuantificadores, la maquina dira en efecto que ϕ no q-re ere a ψ . Si es porque $\phi := \exists v_1 \dots v_n (\psi)$ tal que T ocurre solo en ψ y no existen k_2, \dots, k_n

en ψ para los cuales puede probarse que $\exists [k=v_i] \psi$ y $\exists v_1 \dots v_n \exists v_{i+1} \dots v_n (\psi)[k=v_i]$ q-re ere a ψ , la maquina seguira para siempre sin dar una respuesta, recorriendo todos los numeros para ver si alguno satisface ψ . Sea $R_q(u; v)$ la formula $\exists [k=v_i]$ que define la q-referencia aletica en L_{pa}^+ y la representa debilmente en q^+ .

Definicion 7.23 (Referencia directa aletica). Dadas dos oraciones $\phi; \psi \in L_T$, ϕ re ere aletica y directamente a ψ sii m-re ere o q-re ere aleticamente a ψ en \mathcal{M} .

Esta relacion puede tambien definirse en L_{pa}^+ y debilmente representarse en q^+ mediante la formula

$$R_d(u; v) := R_q(u; v) \wedge R_m(u; v) \tag{7.19}$$

Como R_m es un predicado \exists (cf. (7.17)), mientras que R_q es solo \exists , R_d es tambien \exists .

El siguiente resultado sera util para establecer conexiones entre la nocion de referencia aletica directa (sintactica) y la nocion de dependencia de Leitgeb en el capitulo 8. Estas conexiones son lo que va a posibilitar las demostraciones de consistencia y !-consistencia de las teor as que proveo mas adelante en ese capitulo.

Proposicion 7.24. Sean $\phi; \psi$ formulas de L_T .

1. Si $\phi \in L_{pa}^+$, ϕ no re ere aletica y directamente a ψ .
2. ϕ re ere aletica y directamente a ψ sii ϕ re ere aletica y directamente a ψ .

3. '!' re ere aletica y directamente a si tanto ' como son oraciones de L_T

y alguna de ellas re ere aletica y directamente a .

Demostracion. El punto 1 se sigue del hecho de que tanto la de nicion [7.11](#) como la [7.22](#) requieren que ' contenga T para que re era a una oracion. Los puntos 2 y 3 se deben a que ni la m-referencia ni la q-referencia se ven afectadas por la presencia o ausencia de conectivas, como indican la de nicion [7.11](#) y los puntos 3 y 4 de la de nicion [7.22](#). □

En consecuencia, la referencia aletica directa es composicional, a diferencia de las nociones de Putnam, Goodman y Urbaniak y, por ende, no lleva a los inconvenientes que estas nociones padecen, como se~nale en el apartado [6.3](#).

Definicion 7.25 (Cadenas de referencia aletica directa). Una secuencia de oraciones $\theta; \dots; \theta_n \in L_T, n \in \mathbb{N}$, es una cadena de referencia aletica directa en th sii, para cada $i < n$, θ_i re ere aletica y directamente a θ_{i+1} en th.

El conjunto de secuencias que son cadenas de referencia aletica directa es asimismo semirecursivo, pues es de nible en L_{pa}^+ a partir del predicado R_d mediante la siguiente expresion:

$$Cref(v) := Seq(v) \wedge \exists i < lg(v) R_d((v)_i; (v)_{i+1})$$

de acuerdo con la cual v es una cadena de referencia si codifica una secuencia de oraciones tales que cada una re ere directamente a la que sigue en la secuencia. La formula es Δ_1 porque Seq es Δ_0 , por ser un predicado recursivo (vease el apartado [1.2.3](#)), R_d es Δ_1 y el cuantificador $\exists i < lg(x)$ esta acotado.

Definicion 7.26 (Referencia aletica). Sean ' y ' oraciones de L_T . ' re ere aleticamente a en th sii existe una cadena de referencia aletica $\theta; \dots; \theta_n$ en th tal que $\theta := ' \wedge \theta_n := '$.

Dado que la nocion de referencia esta definida en terminos de cadenas de referencia, que son de nibles en L_{pa}^+ , esta tambien lo es, mediante la expresion

$$R(u; v) := \exists x (Cref(x) \wedge u = (x)_0 \wedge v = (x)_{lg(x)}) \tag{7.20}$$

que establece que u re ere a v si son los extremos de una secuencia que es una cadena de referencia. Su complejidad es nuevamente Δ_1 , porque $Cref(x) \wedge u = (x)_0 \wedge v = (x)_{lg(x)}$ es Δ_1 por lo dicho anteriormente y el unico operador logico que la antecede en la formula es un cuantificador existencial. Luego, la referencia aletica sintactica es una relacion semirecursiva, aunque no recursiva.

Al igual que en los apartados anteriores, a partir de las nociones dadas se pueden definir autorreferencia y buena fundación aleticas que serán, en consecuencia, también definibles en L_{pa}^+ .

Definición 7.27 (Autorreferencia aletica). Una oración $\phi \in L_T$ es autorreferente en \mathcal{M} si y sólo si se refiere aleticamente a sí misma en \mathcal{M} .

Naturalmente, la siguiente fórmula define autorreferencia aletica en L_{pa}^+ :

$$SR(v) := R(v; v) \tag{7.21}$$

Definición 7.28 (Buena fundación aletica). Una oración $\phi \in L_T$ está aleticamente bien fundada en \mathcal{M} si y sólo si no hay una cadena de referencia en \mathcal{M} indefinidamente extensible que empiece en ϕ .

Si bien todavía es definible en L_{pa}^+ por medio de la fórmula

$$Wf(v) := \text{Sent}_{L_T}(v) \wedge \neg \exists x (\text{Cref}(x) \wedge (x)_0 = v) \tag{7.22}$$

$$\exists y > \text{lg}(x) \exists z (\text{Seq}(z) \wedge \text{lg}(z) = y \wedge \exists i \text{ lg}(x)(x)_i = (z)_i \wedge \neg \text{Cref}(z))$$

esta noción es levemente más compleja que las anteriores. De acuerdo con (7.22), para toda cadena de referencia x que comience en v hay un y tal que toda secuencia de longitud y que extienda x no es una cadena de referencia, esto es, bloquea la extensibilidad indefinida. Si (7.22) es la fórmula más sencilla que define la buena fundación aletica, siendo \mathcal{L} , el conjunto de oraciones bien fundadas no es siquiera semirecursivo.

Al igual que las nociones semánticas presentadas en el apartado 7.3.1, las oraciones de la secuencia de Yablo en (7.11) y, su versión alternativa, en (7.12), su versión dual en (7.13) y otras listas como (7.14) resultan aleticamente mal fundadas pero no autorreferentes desde el punto de vista sintáctico, lo cual puede verse por medio de razonamientos análogos a los dados en ese apartado. Naturalmente, lo mismo sucede con las cadenas, como (7.15), pero también con la oración de McGee (7.16). Asimismo, vale el siguiente resultado.

Proposición 7.29. Si ϕ es una oración aleticamente autorreferente de L_T , ϕ no está aleticamente bien fundada.

Sin embargo, existen ciertas oraciones cuyos patrones referenciales son capturados por las nociones semánticas pero no por las sintácticas, como muestro en el siguiente apartado.

7.3.2.2. Incompletitud y jerarquías

Al ser una relación semirecursiva, la q -referencia alética sintáctica es débilmente representable en pa^+ , esto es, pa^+ prueba todos los casos positivos de q -referencia. A diferencia de cualquier axiomatización de la q -referencia semántica, en el modelo estándar de L_{pa}^+ no hay más casos de q -referencia en th que lo que pa^+ th informa, porque pa^+ prueba todas las oraciones γ de L_{pa}^+ que son verdaderas en el modelo estándar es decir, pa^+ es γ completa.

No obstante, esto trae algunas consecuencias prima facie contraintuitivas.¹⁰ Considerese, por ejemplo, la oración

$$\exists x(\text{Bew}_{th}(x) \rightarrow \neg T x) \quad (7.23)$$

Es claro que th no prueba este enunciado (recuérdese que th solo extiende pat con oraciones verdaderas de L_{pa}^+ en N^+), razón por la cual este satisface su propio antecedente y, en consecuencia, refiere a sí mismo, de acuerdo con la noción de q -referencia alética semántica introducida en la definición 7.12. Sin embargo, que (7.23) satisfaga su propio antecedente no es suficiente para concluir que refiere aléticamente a sí misma en th ; es preciso que th pruebe que lo hace, i.e., que

$$th \vdash \text{Bew}_{th}(p \rightarrow \text{Bew}_{th}(p \rightarrow \exists x(\text{Bew}_{th}(x) \rightarrow \neg T x)))$$

o, lo que es lo mismo, que

$$th \vdash \exists x(\text{Bew}_{th}(p \rightarrow \exists x(\text{Bew}_{th}(x) \rightarrow \neg T x)))$$

lo cual es imposible por el segundo resultado de incompletitud de Gödel (teorema 1.13). (7.23) no q -refiere a sí misma en th , sino solo en una teoría más poderosa, que sea capaz de probar que el enunciado satisface su propio antecedente. Esta es una parte de lo que llamo 'incompletitud de la referencia alética sintáctica'.

Si bien parece que la referencia alética semántica da el veredicto correcto cuando su contraparte sintáctica falla, mientras que esta última permite un tratamiento axiomático sin costos, ninguna axiomatización de la referencia semántica será capaz de dar cuenta de todas las relaciones de referencia de oraciones de L_T ; en particular, ninguna axiomatización th tal podrá probar que (7.23) es autorreferencial, que es lo relevante para esta investigación.

¹⁰Este ejemplo me fue sugerido hace tiempo por Halbach como un contraejemplo o caso contraintuitivo para la referencia sintáctica. Hoy lo veo como una simple consecuencia o, si se quiere, 'muerdo la bala'.

Al mismo tiempo, buena parte de los casos negativos de q-referencia sintáctica en th no son demostrables en pa^+ ni en $th\ pa^+$. Por ejemplo, th no es capaz de probar que la oración

$$\exists x(x = p0 = 0q \ ! \ T \ x) \quad (7.24)$$

no q-re ere en th, v.g., al mentiroso $\!:\!T \!$, a s misma o a cualquier otra oración porque, nuevamente, por el segundo resultado de incompletitud de Gödel, al ser $\!-\!$ consistente th no puede probar $\!:\!Bew_{th}(p1 = p0 = 0qq)$, $\!:\!Bew_{th}(p\exists x(x = p0 = 0q \ ! \ T \ x)q = p0 = 0qq)$ ni, en general, $\!:\!Bew_{th}(p'q)$ para ninguna oración ' $\!$ ', i.e., su propia consistencia. Esta es la otra parte de la incompletitud de la referencia aletica sintáctica.

Naturalmente, al ser inde nible, ninguna axiomatización de la referencia aletica semántica será capaz de probar todos los casos negativos de referencia, pero una teoría como qrs s logra demostrar muchos más que cualquier teoría a th con respecto a la referencia aletica sintáctica. Notese que, v.g., para probar en qrs que enunciados de la forma $\exists v(\ ! \ T \ v)$ donde T no ocurre en ' $\!$ ' no q-re eren a una oración basta con probar $\!:\![p \ q=v]$, lo cual es en principio posible, mientras que en th es preciso demostrar $\!:\!Bew(p'[p \ q=v]q)$, que es ciertamente imposible en th.

La razón principal por la cual desarrollo nociones de referencia aletica para oraciones de L_T es el estudio de los patrones de referencia que subyacen a las paradojas semánticas y otras expresiones patológicas, para la posterior formulación de teorías axiomáticas de la verdad ($\!-\!$)consistentes y, a la vez, abarcativas y poderosas, basadas en una caracterización de las patologías en términos de sus patrones referenciales. Como se ha visto, en muchos casos la autorreferencia y la mala fundación resultan ser patrones peligrosos, en tanto ciertas oraciones autorreferentes o mal fundadas, como $\!:\!T \!$ y las oraciones de Yablo, dan lugar a paradojas u $\!-\!$ paradojas. Luego, por un lado, si la teoría a de la referencia en la cual trabajo no logra identificar ciertas expresiones a todas luces autorreferentes o mal fundadas, como (7.23), las teorías de la verdad que formule basándose en la noción de referencia provista correrán el riesgo de ser inconsistentes u $\!-\!$ inconsistentes. Y, por otro, si la teoría a de la referencia no es capaz de señalar que oraciones obviamente 'inocentes', como (7.24), no son autorreferenciales ni mal fundadas, no hay esperanzas de contar con una teoría a lo sucientemente abarcativa y poderosa, como quiere Horwich (1998b) (vease el apartado 5.3.1).

La incompletitud de la referencia aletica sintáctica, la escasa información sobre la presencia o ausencia de ciertas relaciones de referencia aletica pone en juego el proyecto. Pero hay una solución evidente, esto es, agregar axiomas que permitan probar más relaciones de referencia y refutar más relaciones de no referencia aletica. A diferencia de las axiomatizaciones de la noción semántica de referencia aletica, estas no necesitan de un símbolo de predicado primitivo, ni un metalenguaje, ni esquemas ni el predicado de

verdad, sino meramente enunciados de L_{pa}^+ que son verdaderos en N^+ , como muestro a continuacion. Esto pone a la nocion sintactica en clara ventaja con respecto a la nocion semantica, al menos en lo que concierne a la posibilidad de formular teor as de la verdad axiomaticas de acionistas y minimalistas en L_T .

El axioma mas natural parece ser el que establece que si th refuta, esto es, prueba la negacion de todas las clausuras existenciales del antecedente θ (o negacion del consecuente ϕ) de un condicional $\theta \rightarrow \phi$ en el cual solo θ contiene T , reemplazando cada variable libre $v_1; \dots; v_n$ en θ por k , con k no libre en ϕ , solo $\exists v_1 \dots v_n (\theta \rightarrow \phi)$ para cada $1 \leq i \leq n$, solo $\exists v_1 \dots v_n (\theta \rightarrow \phi)$ no q-re ere a en th . La formulacion de este principio en L_T dar a lugar a una formula demasiado complicada. En su lugar, recorro a una formulacion equivalente mucho mas sencilla:

$$\exists x (Bew_{th}(\theta(x)) \rightarrow Bew_{th}(x)) \tag{qr}$$

Supongase que $\exists v(\theta)$ es una oracion de L_T tal que θ no contiene T , s ,

$\theta \rightarrow \exists v(\theta)$, $k \neq v$, y que, si $m \neq v$, $\theta \rightarrow \exists v(\theta)$ no re ere directamente a en th . Luego, $\exists x (Bew_{th}(\theta(x)) \rightarrow Bew_{th}(x))$ implica que $\theta \rightarrow \exists v(\theta)$ y, por ende, $\exists v(\theta)$ no q-re ere a en th . Y, a la inversa, si $\exists v(\theta)$ es una oracion de L_T tal que θ no contiene T , s , $\theta \rightarrow \exists v(\theta)$, $k \neq v$, y, si $m \neq v$, $\theta \rightarrow \exists v(\theta)$ no re ere directamente a en th , $Bew_{th}(\theta(p(k)q))$ implica que $\exists v(\theta)$.

$Bew_{th}(p(k)q)$. Y lo mismo puede extenderse a una cantidad arbitraria de variables de individuo $v_1; \dots; v_n$.

Definicion 7.30 (qr(th)). $qr(th)$ es L_T que extiende th por medio del axioma qr.

Como el principio qr establece la consistencia de th , que por estipulacion es verdadera en toda extension de N^+ , se tiene que

Proposicion 7.31. $qr(th)$ es ω -consistente.

Siendo la identidad un predicado recursivo, $qr(th)$ permite saber ahora que (7.24) q-re ere unicamente a $0 = 0$, y tambien que, por ejemplo,

$$\exists x (Bew_{th}(x) \rightarrow Tx) \tag{7.25}$$

no q-re ere a $0 \neq 0$, porque $\theta \rightarrow Bew_{th}(p(0 \neq 0)q)$ y, por qr, $qr(th) \rightarrow Bew_{th}(p(0 \neq 0)q)$. Asimismo, si bien no es posible establecer en th que las oraciones de Yablo no son autorreferenciales en th , s lo es en $qr(th)$.

Naturalmente, $qr(th)$ no es suficiente para convertir a th en una teoría completa con respecto a la negación, porque es una teoría axiomática consistente que extiende pa^+ y, por ende, está sujeta a los resultados de incompletitud de Gödel. Por ejemplo, $qr(th)$ no permite concluir que (7.25) no refiere a la oración $:T I$ del mentiroso en th .

Si bien $R_q(x; y)$ es una fórmula de L_{pa}^+ y, por ende, por medio de fórmulas como (7.19), (7.20), (7.21), etc. se puede contar con predicados R_d, R, SR , etc. que den las nociones de la familia de la referencia aleática aplicables a oraciones que contienen ellas mismas estos predicados, i.e., no tipados, los conceptos referencia sintáctica están en alguna medida estratificados. Por ejemplo, $qr(th)$ genera cierta estratificación, porque da lugar a nuevos casos de referencia dados por lo que se puede probar en $qr(th)$ y no en th . Considerese la oración

$$\exists x(:R_q(p\exists x(x = 0 \wedge T x)q; x) \wedge T x) \quad (7.26)$$

Intuitivamente, (7.26) q -refiere a toda oración a la cual $\exists x(x = 0 \wedge T x)$ no q -refiere. Como $qr(th)$ implica que esta última no q -refiere a, v.g., $:T I$, (7.26) debería referir a $:T I$ en th . Pero no lo hace, porque th no es capaz de probar que $:T I$ satisface el antecedente de (7.26). En cambio, como $qr(th)$ sí lo es, (7.26) q -refiere a $:T I$ en $qr(th)$.

Lo mismo sucede con (7.23). Esta oración es autorreferencial, no en th , sino en $qr(th)$. En ninguno de estos casos se da que hay una relación de q -referencia en th que th no puede probar, porque la noción es \perp . Es simplemente falso que (7.26) refiera al mentiroso y que (7.23) es autorreferente en th . Las nociones de referencia en th y $qr(th)$ no coinciden sino que la primera es un subconjunto de la segunda, porque ambas están formuladas en términos de su correspondiente predicado de prueba. Ante la nueva noción de q -referencia que $qr(th)$ conlleva, surge, en aras de la completitud, la necesidad/posibilidad/utilidad de incorporar axiomas análogos a qr pero ahora relativos al predicado de prueba de $qr(th)$, que a su vez dar lugar a una nueva teoría con su propia noción de q -referencia, etc. y que, tal vez, permitan probar, v.g., la autorreferencialidad de (7.23).

Para evitar esta estratificación, la distinción entre teoría y metateoría, sería ideal una teoría $qr(th)$, idéntica a $qr(th)$ excepto en que las apariciones de Bew_{th} en (7.22) y qr son reemplazadas por $Bew_{qr(th)}$. Lamentablemente, cualquier teoría que cumpla con estos requisitos es trivialmente trivial. Notese que $qr(th)$ prueba la consistencia de th . Si fuera posible una teoría como $qr(th)$, esta probaría su propia consistencia y, por el segundo teorema de incompletitud de Gödel, sería inconsistente.

Trabaja entonces con una jerarquía de teorías $pat := qr_0; qr_1; qr_2; \dots$ y sus nociones de referencia aleática correspondientes, definibles en L_{pa}^+ por las fórmulas $R^0(u; v)$,

$R^1(u; v), R^2(u; v); 2 \uparrow$, siguiendo la definición 7.22.¹¹ Es posible incluso extender esta jerarquía a niveles transnitos. Esto permitiría saber más acerca de las relaciones de referencia aleática que se dan entre las oraciones de L_T pero también sobre las que no se dan.

Definición 7.32. Sea $qr_\theta := \text{pat}$ y, si $0 < \theta$, qr extiende pat con el siguiente esquema de axioma, para cada ordinal $\alpha < \theta$:

$$qr \vdash \exists x (\text{Bew}_{qr}(\cdot; x) \rightarrow \text{Bew}_{qr}(x))$$

A diferencia de la m-referencia aleática, para cada ordinal hay una noción de q-referencia aleática relativa a qr , como muestra la definición 7.22, que es definible en L^+_{pa} y débilmente representable en pat^+ por una fórmula que notamos R_q . Luego, existen también jerarquías para cada una de las nociones que están definidas en términos de la q-referencia aleática sintáctica, esto es, la referencia directa, la referencia simpliciter, la autorreferencia y la buena fundación aleáticas. Si $\alpha < \theta$, las notamos $R_d(u; v)$, $R(u; v)$, $SR(v)$ y $Wf(v)$, respectivamente.

Proposición 7.33. Si $\theta < \omega$, qr_θ es ω -consistente.

Demostración. A partir del hecho de que pat es verdadera en toda extensión de N^+ a L_T , por inducción transnita sobre $\alpha < \theta$. □

Hasta donde es correcto avanzar en la jerarquía quedar determinado por el ordinal prueba-teórico de la teoría base, pat , puesto que los axiomas de cada qr_α son ω y agregar oraciones ω verdaderas en el modelo estándar a th no da lugar a nuevas instancias de inducción transnita.¹² Luego, la teoría más poderosa de la jerarquía a que es razonable aceptar es qr_θ .

Observación 7.34. Sean $\alpha < \theta$ y t y s dos términos cualesquiera de L^+_{pa} . Si $qr_\alpha \vdash R_q(t; s)$, entonces, para todo $\beta > \alpha$, $qr_\beta \vdash R_q(t; s)$.

Esta proposición establece que, si una oración q-referencia aleáticamente a otra relativamente a una teoría en la jerarquía, lo hará en todas las teorías que se encuentren por encima de ella, esto es, establece la acumulatividad de las relaciones positivas de referencia en la jerarquía. En cambio, no se da que, para todas las oraciones $\varphi \in L_T$, si una teoría en la jerarquía prueba que φ no q-referencia aleáticamente a otra oración ψ relativamente a

¹¹Notese que en la definición 7.22 th puede ser cualquier teoría que extienda pat con oraciones ω de L^+_{pa} tales que N^+ , requisito que satisfacen todas las teorías que extienden pat con axiomas de la forma de qr , si el subíndice th que ahí ocurre cumple también la condición mencionada.

¹²Dado que los predicados de prueba estándar $\text{Bew}_{\text{th}}(x)$ son fórmulas ω . Cf. Michael Rathjen (1999).

otra teoría de la jerarquía, entonces no q -referirá relativamente a ninguna de las teorías que aparecen posteriormente a esta última; la jerarquía de teorías qr no es acumulativa con respecto a los casos negativos de q -referencia.

Notese que esto puede representar un grave inconveniente. Si acaso se concluye, relativamente a una teoría qr de la jerarquía, que, por ejemplo, 're ere únicamente a $0 = 0$ y, por ende, esta bien fundada, sería deseable poder afirmar que es una instancia segura del [Esquema-T](#). No obstante, por la no acumulatividad, podría ocurrir que en un nivel superior $>$, 're era no solo a $0 = 0$ sino, v.g., a sí misma, tornándose

una instancia 'peligrosa' del [Esquema-T](#). Luego, que una teoría qr asigne patrones seguros a una expresión no es garantía alguna de que su instancia correspondiente del [Esquema-T](#) no generar problemas y, por ende, las nociones de referencia aleática parecen ser inútiles.

Sin embargo, este razonamiento es demasiado apresurado. Si bien para algunas oraciones la no referencia no es estable, para muchas sí lo es. Para estas últimas, es posible dar veredictos correctos con respecto a sus patrones referenciales subyacentes y, por tanto, con respecto a su 'peligrosidad' que instancias del [Esquema-T](#). En el apartado siguiente identifico la clase de estas oraciones cuya referencia es estable a partir de un punto en la jerarquía, que denomino 'r-decidibles', y pruebo su estabilidad en la jerarquía de teorías qr , con θ .

7.3.2.3. Referencia estable

En este apartado caracterizo las oraciones cuya q -referencia es estable o decidible en alguna teoría de la jerarquía, esto es, aquellas para las cuales, si para algún θ , qr prueba que estas oraciones no q -referen a ciertas otras relativamente a una teoría de la jerarquía, no existe una teoría qr , $>$ que pruebe que sí q -referen relativamente a ninguna otra teoría. Primero es preciso definir algunas nociones auxiliares.

Definición 7.35 (Decidibilidad). Una fórmula $\varphi \in L_{pa}^+$ con $v_1; \dots; v_n$ como sus únicas variables libres es decidible en th_{sii} , para cualesquiera $k_1; \dots; k_n \in \mathbb{N}$, $th_{sii}(\varphi; k_1; \dots; k_n)$ o $th_{sii}(\neg\varphi; k_1; \dots; k_n)$.

Las fórmulas decidibles en th_{sii} son precisamente aquellas que representan en th_{sii} la relación entre números naturales que dependen. Lamentablemente, conjeturo que esta noción no es siquiera recursivamente enumerable; la fórmula de menor complejidad que, he encontrado, la expresa en L_{pa}^+ es Σ_1^1 :

Form_L^{pa} (v) ^ 8x(Seq(x) ^ lg(x) = free(v) ! Bew_{th}(~s.(v; x)) _ Bew_{th}(:~s.(v; x)))

La nota $\text{Dec}^{\text{th}}(v)$. De acuerdo con este predicado, si v es una formula de L_{pa}^+ y x (el codigo de) una secuencia de (los codigos de) las variables libres en v , th es capaz de probar, o bien el resultado de reemplazar en v cada variable libre por cada miembro de x , o bien su negacion. Si bien esta formula es demasiado compleja, existen muchos casos de formulas decidibles para las cuales se puede probar que lo son en pa^+ . Por ejemplo, para todas las formulas equivalentes a formulas \mathcal{L} , que son decidibles.

Proposicion 7.36. Si $\varphi \in L_{pa}^+$ es equivalente en th a una formula \mathcal{L} , entonces $\text{th} \vdash \text{Dec}^{\text{th}}(p'q)$.

Demostracion. Sea $\sim v$ una abreviatura de $v_1; \dots; v_n$. Sean $\varphi(\sim v); (\sim v) \in L_{pa}^+$ tales que $\text{th} \vdash \varphi$ y φ es \mathcal{L} . Luego, existen formulas $\varphi(\sim v) \in \mathcal{L}$ y $\varphi(\sim v) \in \mathcal{L}$, ambas logicas y, por tanto, demostrablemente equivalentes a φ , por ende, a φ en th . Notese que φ es asimismo \mathcal{L} . Como pa^+ sabe de su propia completitud \mathcal{L} y, a fortiori, de la de $\text{th}(\mathcal{L} \in L_{pa}^+)$, si $\varphi(\sim v) \in \mathcal{L}$,¹³

$$pa^+ \vdash (\sim v) \vdash \text{Bew}_{\text{th}}(p(\sim v)q) \quad (7.27)$$

Ademas, de acuerdo con la tercera condicion de $L \cdot \text{ob}$ para los predicados de prueba estandares Bew (vease el apartado 1.2.3),

$$pa^+ \vdash \text{Bew}_{\text{th}}(p(x),y) \vdash (\text{Bew}_{\text{th}}(x) \vdash \text{Bew}_{\text{th}}(y)) \quad (7.28)$$

Luego, razono en th . Por un lado,

$$\begin{aligned} & \vdash (\sim v) \vdash \text{Bew}_{\text{th}}(p(\sim v)q) && \text{por (7.27)} \\ & \vdash \text{Bew}_{\text{th}}(p(\sim v)q) && \text{por (7.28)} \\ & \vdash \text{Bew}_{\text{th}}(p'(\sim v)q) \end{aligned}$$

y, por otro,

$$\begin{aligned} & \vdash (\sim v) \vdash \text{Bew}_{\text{th}}(p(\sim v)q) && \text{por (7.27)} \\ & \vdash \text{Bew}_{\text{th}}(p'(\sim v)q) && \text{por (7.28)} \end{aligned}$$

Luego, como $\vdash \vdash$ implica $\vdash \vdash$,

$$\text{th} \vdash \text{Bew}_{\text{th}}(p'(\sim v)q) \vdash \text{Bew}_{\text{th}}(p'(\sim v)q)$$

¹³Vease [Craig Smorinsky \(1985, p. 61, teorema 6.22\)](#) y [Stephen G. Simpson \(2009, p. 370-371, de nición IX.3.4, lema IX.3.5 y teorema IX.3.6\)](#). La prueba puede facilmente adaptarse a pa^+ , extendiendo la interpretacion canonica de L_{pa} en el lenguaje de pa a L_{pa}^+ , porque el lenguaje de pa contiene terminos para cada función p.r.

o, lo que es lo mismo,

$$\text{th} \vdash \exists v_1 \dots \exists v_n (\text{Bew}_{\text{th}}(p'(v_1, \dots, v_n)q) \leftrightarrow \text{Bew}_{\text{th}}(p'(v_1, \dots, v_n)q)):$$

□

Notese que, si bien todas las formulas que son logicamente equivalentes a una formula φ son ellas mismas φ , este no es necesariamente el caso para formulas que no son equivalentes logicamente sino solo en pa^+ o en th . Luego, la proposición [7.36](#) implica que mas formulas son decidibles que aquellas que son φ .

Definición 7.37 (Rd-decidibilidad). Una oración $\varphi \in L_T$ es rd-decidible en th sii en todas sus subformulas de la forma $\exists x \varphi(x)$, en las cuales una variable libre ocurre bajo el alcance de T y solo φ o solo $\neg \varphi$ contienen T , son tales que la que no contiene T es decidible.

Intuitivamente, las oraciones rd-decidibles en th son aquellas cuya referencia directa queda completamente determinada por la teoría; no es posible que en una teoría consistente que incluya th la oración referida directamente a mas cosas que en th , como muestra la proposición [7.39](#). Por ejemplo, $\exists x(x = \ulcorner \neg Tx \urcorner)$ es rd-decidible, porque la identidad es un predicado recursivo. Además, cada una de las oraciones de la lista de Yablo en [\(7.11\)](#) son rd-decidibles en th , porque $u > v$ es una formula decidible en cualquier $\text{th} \text{ q}^+$, y lo mismo para sus formulaciones alternativas. Y $\exists x(\text{Bew}_{\text{th}}(x) \wedge \neg Tx)$ es rd-decidible porque ninguna variable libre aparece bajo el alcance de T . En cambio, $\exists x(\text{Bew}_{\text{th}}(x) \wedge Tx)$ no es rd-decidible en th , porque $\text{Bew}_{\text{th}}(x)$ no es decidible en th .

La rd-decidibilidad es claramente definible en L_T en terminos de $\text{Dec}^{\text{th}}(v)$, digamos mediante la formula $\text{RdDec}^{\text{th}}(v)$. Dado que Dec^{th} es Σ_1 y no aparece bajo el alcance de una negación en la definición [7.37](#), RdDec^{th} es tambien Σ_1 .

Definición 7.38 (R-decidibilidad). Una oración $\varphi \in L_T$ es r-decidible en th sii es rd-decidible y solo referida aleticamente a oraciones rd-decidibles.

Las oraciones r-decidibles son aquellas para las cuales la referencia aletica simpli-citer queda fijada en una teoría, no se modifica en teorías consistentes con mayor poder de prueba, como indica el corolario [7.40](#). El conjunto de oraciones r-decidibles en th es definible en L_{pa}^+ en terminos de RdDec^{th} mediante la formula

$$\text{RdDec}^{\text{th}}(v) \wedge \exists x(R(v; x) \wedge \text{RdDec}^{\text{th}}(x))$$

que abrevio $\text{RDec}^{\text{th}}(v)$. Como $R(u; v) \in \Sigma_1$, RDec^{th} es tambien Σ_1 .

Si ϕ , sea \mathcal{N}^+ ; i la interpretacion en que extiende \mathcal{N}^+ a todo L_T , asignando a T como su extension.

Proposicion 7.39. Sean ϕ y oraciones de L_T tales que ϕ es rd-decidible en \mathcal{N} , sean \mathcal{N}^+ ; i \mathcal{N}^0 \mathcal{N} para algun \mathcal{N} y sean R_d y R_d^0 las formulas que debilmente representan la referencia aletica directa de cada una de estas teor as, respectivamente, en L_T . Si \mathcal{N}^+ ; $i : R_d(p;q; p q)$, entonces \mathcal{N}^+ ; $i : R_d^0(p;q; p q)$.

Demostracion. Como la referencia aletica directa es la union de la m y la q -referencia, basta probar el resultado para cada una de estas por separado. Para la m -referencia es obvio. Pruebo el resultado para la q -referencia por induccion sobre la complejidad de ϕ . Asumo que ϕ contiene T ; de lo contrario la prueba es trivial. Si ϕ es atomica, entonces no q -re ere ni en \mathcal{N} ni en \mathcal{N}^0 , por la de nicion 7.22.

Supongase que para toda oracion de menor complejidad que ϕ el enunciado de la proposicion es verdadero. Si $\phi := \neg \psi$ y ψ no q -re ere a en \mathcal{N} , entonces ψ tampoco, por la clausula 3 de la de nicion 7.22. Por hipotesis inductiva, ψ no q -re ere a en \mathcal{N}^0 y, por ende, tampoco ϕ . El caso para el condicional es analogo.

Sea $\phi := \exists v_1 \dots \exists v_n \psi$ tal que no q -re ere a en \mathcal{N} . Si $\psi := \forall t \phi$, para cada $k_1, \dots, k_n \in \mathcal{N}$, $[\mathcal{N}, v_1 = k_1, \dots, v_n = k_n]$ no contiene un termino que denote ψ , por ende, no q -re ere a en \mathcal{N}^0 tampoco. Luego, ψ no re ere a en \mathcal{N}^0 . Si $\psi := \exists t \phi$, para cada $k_1, \dots, k_n \in \mathcal{N}$, $[\mathcal{N}, v_1 = k_1, \dots, v_n = k_n]$ no q -re ere a en \mathcal{N} ni contiene un termino 'nuevo' que denote ψ . Por hipotesis inductiva, $[\mathcal{N}, v_1 = k_1, \dots, v_n = k_n]$ no q -re ere a en \mathcal{N}^0 ni, obviamente, contiene un termino 'nuevo' que denote ψ , para ningun $k_1, \dots, k_n \in \mathcal{N}$. Luego, $v_1 \dots v_n$ no q -re ere a en \mathcal{N}^0 .

Finalmente, si $\phi := \exists t \psi$, caben tres opciones: que tanto ψ como ϕ contengan T , que ψ no contenga T y ϕ s y, viceversa, que ψ contenga T y ϕ no. El primer caso es analogo al de la negacion, y los dos ultimos son semejantes entre s. Lo pruebo solo para el segundo. Como ψ no q -re ere a en \mathcal{N} , para todo $k \in \mathcal{N}$ y todo $1 \leq i \leq n$, si $\exists v_1 \dots \exists v_i \neg v_{i+1} \dots \exists v_n (k=v_i)$ q -re ere a en \mathcal{N} o contiene una ocurrencia de T que no esta en ψ tal que $q^+ \neg t = p q$, th $\exists v_1 \dots \exists v_i \neg v_{i+1} \dots \exists v_n [k=v_i]$. Como ψ es rd-decidible, debe ser decidible, con lo cual $\mathcal{N} \models \exists v_1 \dots \exists v_i \neg v_{i+1} \dots \exists v_n [k=v_i]$.

Y, como $\mathcal{N} \models \exists t \psi$, th $\exists v_1 \dots \exists v_i \neg v_{i+1} \dots \exists v_n [k=v_i]$. Adem as, como es consistente, $\mathcal{N} \models \exists t \exists v_1 \dots \exists v_i \neg v_{i+1} \dots \exists v_n [k=v_i]$ tambien. Luego, si $v_1 \dots v_i \neg v_{i+1} \dots v_n (k=v_i)$ q -re ere a en \mathcal{N} o contiene una ocurrencia de T que no esta en ψ tal que $q^+ \neg t = p q$, th $\exists v_1 \dots \exists v_i \neg v_{i+1} \dots \exists v_n [k=v_i]$. Por otro lado, por hipotesis inductiva, si $v_1 \dots v_i \neg v_{i+1} \dots v_n (k=v_i)$ q -re ere a en \mathcal{N} (o contiene una ocurrencia de T que no esta en ψ tal que $q^+ \neg t = p q$), tambien lo hace en \mathcal{N} . En consecuencia, si $v_1 \dots v_i \neg v_{i+1} \dots v_n (k=v_i)$ q -re ere a en \mathcal{N} o contiene una ocurrencia de T

que no está en \mathcal{L} tal que $q \vdash p \rightarrow q$, $\text{th } \mathcal{L} \vdash v_1 \dots v_n [k=v_i]$, con lo cual, por la definición [7.22](#), \mathcal{L} no q -representa a $\text{th } \mathcal{L}^0$. \square

Corolario 7.40. Sean \mathcal{L} y oraciones de L_T tales que \mathcal{L} es r -decidible en $\text{th } \mathcal{L}$, sean $\text{th } \mathcal{L}^+$; i $\text{th } \mathcal{L}^0$ th para algún i y sean R y R^0 las fórmulas que débilmente representan la referencia débil de cada una de estas teorías, respectivamente, en L_T . Si $\text{th } \mathcal{L}^+$; i $:R(p \rightarrow q; p \rightarrow q)$, entonces $\text{th } \mathcal{L}^+$; i $:R^0(p \rightarrow q; p \rightarrow q)$.

Corolario 7.41. Si \mathcal{L} y \mathcal{L}' son oraciones de L_T , si $\text{th } \mathcal{L}^+ :R_q(p \rightarrow q; p \rightarrow q)$ y \mathcal{L}' es rd -decidible en qr , entonces, para todo \mathcal{L} , $\text{th } \mathcal{L}^+ :R_q(p \rightarrow q; p \rightarrow q)$.

Los resultados de este apartado serán de suma importancia en el capítulo [8](#), donde formulo teorías de la verdad en las cuales el [Esquema-T](#) está restringido a ciertas oraciones según los patrones de referencia que las subyacen. Grosso modo, las oraciones que generen instancias del [Esquema-T](#) deberán ser r -decidibles, porque solo de ellas se puede tener certeza de que sus patrones referenciales son seguros.

La r -indecidibilidad guarda cierta analogía con las adscripciones ciegas de verdad, como "Lo último que dijo Tarski es verdadero" dicho por alguien que no sabe que es lo último que Tarski dijo.¹⁴ Del mismo modo en el que estas adscripciones de verdad podrán resultar autorreferenciales o mal fundadas [v.g., en el caso en que lo último que Tarski haya dicho fuera precisamente "Lo último que dijo Tarski es falso"] también podrán resultar autorreferenciales o mal fundadas ciertas oraciones r -indecidibles en una teoría dada, aunque al momento parezcan inofensivas. En lo que a la teoría para la cual se define referencia respecta, las oraciones r -indecidibles contienen adscripciones de verdad relativizadas a predicados sobre cuya extensión no se tiene conocimiento completo. Luego, podrán, en principio, resultar autorreferenciales o mal fundadas en sistemas más poderosos, donde se tenga más información sobre los predicados relativizantes.

§

A lo largo de este capítulo he provisto cuatro nociones de referencia diferentes, dos para L_{pa}^+ y dos para L_T , todas ellas adecuadas con respecto a la noción intuitiva esbozada por [Leitgeb \(2002\)](#) y, por ende, hiperintensionales, es decir, fallan la condición de equivalencia (vease el apartado [6.1.2](#)). Dejando la hiperintensionalidad de lado, los nuevos conceptos han logrado superar las dificultades que afectan los abordajes de Putnam, Goodman y Urbaniak al contenido referencial, así como el desafío de Leitgeb.

Mientras que tanto la referencia a números (definición [7.1](#)) como la referencia a oraciones (definición [7.8](#)) son compleciones de la noción intuitiva de Leitgeb, esta última

¹⁴Estoy agradecida con Halbach por hacerme ver este punto.

rescata también la intuición detrás de la noción técnica de autorreferencia que Leitgeb adjudica a Priest, en tanto explica como los lemas de diagonalización son capaces de establecer autorreferencia. Como consecuencia, da cuenta del carácter autorreferencial de las oraciones de Gödel y Henkin, así como de la circularidad de los ciclos y la mala fundación de las cadenas.

Por su parte, las nociones de referencia alética que presento en [7.3](#) no solo capturan las nociones intuitiva y técnica, aunque únicamente a través del predicado de verdad, sino que dan veredictos intuitivamente correctos sobre todas las expresiones que dan lugar a paradojas semánticas consideradas. Consecuentemente, la siguiente observación ha perdido validez:

The relation of dependence as given by [definición [3.26](#)] may thus at best be regarded as a formal substitute for the pre-theoretic relation of aboutness concerning semantic properties. Such a substitution, which certainly exceeds the boundaries of an explicatory definition, seems to be admissible in view of the difficulties of grasping the proper notion of aboutness since the latter is perhaps even inherently unclear [. . .]. (Leitgeb, 2005, p. 176, sus *italicas*)

Las nociones de referencia alética han permitido superar el estado de precariedad de la referencia que not Leitgeb en su momento.

Esto vuelve plausible la formulación de la visión referencialista sobre las paradojas semánticas en términos de estas nuevas nociones. Dado que la lista de Yablo y sus formulaciones alternativas resultan mal fundadas pero no autorreferenciales de acuerdo con ambas nociones de referencia alética, la versión ortodoxa de la tesis referencialista queda refutada.

Habiendo mostrado que la noción de referencia alética semántica es indeseable en L_{pa}^+ y, por ende, poco útil para la formulación de teorías axiomáticas de la verdad que restrinjan el [Esquema-T](#) a instancias cuyos patrones de referencia sean 'seguros', en [7.3.2](#) exploré un concepto semejante en el cual las nociones semánticas, las ocurrencias del predicado de verdad-en- N^+ , son reemplazadas por nociones sintácticas, por ocurrencias del predicado de prueba de la teoría para la cual se quiera definir referencia. El resultado, la referencia alética sintáctica, es una relación sintáctico-aritmética, deseable en L_{pa}^+ , como he mostrado, aunque presenta ciertos inconvenientes originados en los fenómenos de incompletitud de Gödel, que se ven en el apartado [7.3.2.2](#). Tomando algunos recaudos indicados en [7.3.2.3](#), la referencia alética sintáctica está lista para ser utilizada en la formulación de restricciones sintácticas sobre instancias del [Esquema-T](#) que den lugar a teorías axiomáticas de la verdad de acionistas y minimalistas, lo cual va a confirmar la tesis referencialista general en el próximo capítulo.

8

Verdad minimalista

\. . . assume every concept to be significant everywhere except for certain 'singular points' or 'limiting points', so that the paradoxes appear as something analogous to dividing by zero."

{ Kurt Gödel, Russell's mathematical logic

En el capítulo [7](#) introduje nociones correctas de referencia para expresiones del lenguaje de la verdad, adecuadas para el estudio de patrones de referencia que subyacen las expresiones paradójicas. El paso siguiente, que ocupa este capítulo, es identificar las notas comunes que estos patrones tienen entre sí y utilizarlas para restringir el conjunto de oraciones que generan instancias legítimas del [Esquema-T](#). Dado que la noción de referencia que empleo es de corte sintáctico, los criterios de restricción de instancias de este principio también lo serán y, por ende, darán lugar a teorías desentramadoras axiomáticas de la verdad que cuadrarán con las demandas del proyecto minimalista de Horwich y, aun más, servirán de plataforma para la construcción de teorías de acionistas.

Asimismo, los sistemas resultantes serán evidencia a favor de la visión referencialista de las paradojas, i.e., la idea según la cual las paradojas semánticas y otras expresiones patológicas pueden ser identificadas a través de sus patrones de referencia. Si bien popular y mayormente incuestionada, esta tesis jamás ha sido propiamente evaluada hasta hoy, principalmente por la falta de nociones precisas y correctas, o al menos decentes, de referencia de oraciones a oraciones en el lenguaje de la verdad (vease el capítulo [6](#)), falta que he reparado en el apartado [7.3](#).

En [8.1](#) ofrezco una serie de resultados que conectan la noción de referencia alética sintáctica con el concepto de dependencia de Leitgeb (veanse los apartados [7.3.2](#) y [3.2](#)). Estos puentes son los que van a permitir establecer la consistencia u !-consistencia, según

el caso, de las teorías de la verdad que introduzco más adelante en el capítulo 8.2 provee cuatro sistemas consistentes desentrecorilladores, examina sus propiedades y los evalúa en función de las cláusulas minimalistas (a)-(d) y las cláusulas de acionistas (1)-(3) en los apartados 8.2.1-8.2.4, correspondientemente. Mientras que todos ellos resultan adecuados desde el punto de vista minimalista, solo los tres primeros satisfarán las demandas del de acionismo. En cada caso, las propiedades de los sistemas provistos permitirán extraer conclusiones acerca de los patrones de referencia subyacentes a las expresiones que dan lugar a paradojas semánticas y conrmarán la tesis referencialista.

8.1. Conexiones entre referencia aletica y dependencia

Los siguientes resultados son la clave para establecer la consistencia u !-consistencia, según el caso, de las teorías que presento más adelante en el capítulo. Prueban ciertas conexiones entre la referencia aletica sintactica y la dependencia de Leitgeb que van a permitir importar resultados de satisfacibilidad del apartado 3.2 sobre esta última a resultados acerca de la referencia aletica.

Sea $\text{th} \quad L_T$ una extensión de pat mediante oraciones de L_{pa}^+ verdaderas en N^+ .

Proposición 8.1. Sean ϕ y ψ oraciones de L_T tales que la primera es rd-decidible en th . Si ϕ no refiere aletica y directamente a ψ en th , hay un α del cual ϕ depende tal que $\alpha = \psi$.

Demostración. Por inducción sobre la complejidad de ϕ . Si ϕ es una oración atómica, o bien $\phi \in L_{pa}^+$ o bien $\phi := \neg \psi$, donde $\psi \in L_{pa}^+$, porque ϕ no refiere a . Luego, por los puntos 1 y 2 de la proposición 3.31, ϕ depende de un α tal que $\alpha = \psi$.

Supongase ahora que para toda oración con menos operadores lógicos que ϕ que no refiere a ψ hay un α del cual depende tal que $\alpha = \psi$. Si $\phi := \psi \wedge \chi$, como ϕ no refiere a ψ , por el punto 2 de la proposición 7.24, tampoco refiere a ψ . Por hipótesis inductiva, hay un α del cual ψ depende tal que $\alpha = \psi$. Y por el punto 3 de la proposición 3.31, ϕ también depende de α . El caso para oraciones de forma condicional es análogo.

Sea $\phi := \forall v_1 \dots \forall v_n \psi$. Si ψ es una fórmula atómica, una negación o un condicional donde el antecedente contiene T si el consecuente también, como ϕ no refiere a ψ en th , para ningún $k \in \{1, \dots, n\}$ sucede que $\forall v_1 \dots \forall v_i \neg \forall v_{i+1} \dots \forall v_n [k=v_i]$ refiere a ψ . Por hipótesis inductiva, para cada $k \in \{1, \dots, n\}$, hay un α_k del cual $\forall v_1 \dots \forall v_i \neg \forall v_{i+1} \dots \forall v_n [k=v_i]$ depende tal que $\alpha_k = \psi$. Luego, por el punto 5 de la proposición 3.31, hay un α del cual $\forall v_1 \dots \forall v_n \psi$ depende tal que $\alpha = \psi$.

Finalmente, sea $\phi := \exists v_1 \dots \exists v_n (\psi)$, donde ψ contiene T si y sólo si no. Suponga-

se sin pérdida de generalidad que ψ no contiene T . Como ϕ no refiere a ω , para todo $k \in \mathbb{N}$ y todo $1 \leq i \leq n$, si $\exists v_1 \dots \exists v_i \neg v_{i+1} \dots \exists v_n (\psi)[k=v_i]$ refiere a ω , entonces

9

$\exists v_1 \dots \exists v_i \neg v_{i+1} \dots \exists v_n [k=v_i]$. Como ϕ es rd-decidible, por la definición 7.37, ϕ es decidible, esto es, $\text{th} \vdash \exists v_1 \dots \exists v_i \neg v_{i+1} \dots \exists v_n [k=v_i]$. Como th es verdadera en toda extensión de \mathbb{N}^+ a L_T , $\mathbb{N}^+ \models \exists v_1 \dots \exists v_i \neg v_{i+1} \dots \exists v_n [k=v_i]$. Luego, para todo $k \in \mathbb{N}$ y todo $1 \leq i \leq n$, si $\mathbb{N}^+ \models \exists v_1 \dots \exists v_i \neg v_{i+1} \dots \exists v_n [k=v_i]$, $\exists v_1 \dots \exists v_i \neg v_{i+1} \dots \exists v_n (\psi)[k=v_i]$ no refiere a ω . Por hipótesis inductiva, para cada una de estas instancias existe un ω del cual depende tal que $\omega \models \psi$. Y, por el punto 5 de la proposición 3.31, hay un ω del cual $\exists v_1 \dots \exists v_n (\psi)$ depende tal que $\omega \models \psi$. \square

Idealmente no habría que restringirse a fórmulas rd-decidibles para que la dependencia implique la referencia alética directa. Lamentablemente, un enunciado como 8.1 donde no se especifica que ϕ es rd-decidible sería falso. Ya he puesto ejemplos de oraciones que, si bien no refieren a otras en una teoría th , si lo hacen en otras más abarcativas (vease el apartado 7.3.2.2). Por ejemplo, $\exists x (\text{Bew}_{\text{th}}(x) \rightarrow Tx)$, que contiene la oración de Gödel de th en todo conjunto del cual depende, no refiere en th a la oración de Gödel de th , aunque sólo hace en $\text{qr}(\text{th})$. Que una oración no refiera directamente a otra en una teoría no es prueba suficiente de que no depende de ella; a menos que sea rd-decidible en la teoría, como indica la proposición 7.39: si ϕ es rd-decidible en una teoría th sobre la cual no refiere directamente a otra oración, tampoco lo hará sobre teorías que expandan th con oraciones verdaderas en \mathbb{N}^+ , sean o no axiomatizables.

La proposición 8.1 conecta la dependencia de Leitgeb con la referencia alética y permite probar las siguientes conexiones entre la buena fundación alética y el conjunto de oraciones d-fundadas de Leitgeb, pf (veanse las definiciones 3.34, 3.33 y 7.28), por un lado, y la autorreferencia alética y la autodependencia (veanse las definiciones 3.32 y 7.27), por otro.

Corolario 8.2. Sea ϕ una oración de L_T . Si ϕ es r-decidible y está aléticamente bien fundada en th , $\phi \in \text{pf}$.

Demostración. De acuerdo con la definición 7.28, una oración ϕ de L_T está aléticamente bien fundada en th si y sólo si no es posible una cadena de referencia alética indefinidamente extensible que comience con ella en esta teoría. Sea $m := \max\{n : \text{existe una cadena de referencia alética directa en } \text{th} \text{ que comienza en } \phi \text{ de longitud } n\}$. $m > 0$ puede ser un ordinal transnito, como en el caso, v.g. de la oración $\exists x T^x p_0 = 0$. Pruebo por inducción transnita sobre $m > 0$ que toda ϕ r-decidible y aléticamente bien fundada pertenece a pf , donde pf es como en la definición 3.33.

Si $m = 1$, en virtud de la definicion [7.25](#), ϕ no es aleticamente a nada en th . Luego, por la proposicion [8.1](#), depende de ω y, por ende, $\omega \leq 1$.

Supongase que, para toda oracion r -decidible y aleticamente bien fundada en th , si $m < \omega$, $\omega \leq m$. Si ϕ tal que $m = \omega$ es r -decidible y aleticamente bien fundada en th , es unicamente a oraciones rd -decidibles en th tales que $m < \omega$ que, por la hipotesis inductiva, pertenecen a m . Luego, por la proposicion [8.1](#), ϕ depende de $S_{i < \omega}$ y, por ende, pertenece a m . □

Como puede verse en el paso inductivo de la demostracion de este corolario, es necesario que ω sea r -decidible, no basta con que sea rd -decidible. Efectivamente, si solo se pide esto ultimo, el enunciado es falso, hay contraejemplos. Considerese el ciclo de mentirosos dado por las identidades

$$t_1 = p \wedge t_2$$

$$t_2 = p \wedge (\text{pat} \wedge x = t_1 \rightarrow \neg Tx)$$

donde pat es la oracion de Godel de pat que se obtiene aplicando el lema de diagonalizacion debil (teorema [1.7](#)) a $\neg \text{Bew}_{\text{pat}}(v)$. t_2 es rd -decidible y es solo a $\neg (\text{pat} \wedge x = t_1 \rightarrow \neg Tx)$, que, en pat esta bien fundada, porque nada satisface el antecedente. Pero no se puede concluir que $t_2 \leq \omega$.

Los siguientes enunciados tambien se siguen de la proposicion [8.1](#).

Corolario 8.3. Sea ϕ una oracion de L_T . Si ϕ es r -decidible y no aleticamente auto-referente en th , ϕ no es autodependiente.

Corolario 8.4. Sea ϕ una oracion de L_T . Si ϕ es rd -decidible y depende esencialmente de un conjunto finito de oraciones Σ , entonces es aleetica y directamente solo a los miembros de Σ .

Estos tres corolarios, junto con los resultados obtenidos en el apartado [3.2](#), son suficientes para establecer la consistencia y, en algunos casos, la ω -consistencia, de las teorías de la verdad a las cuales dan lugar los criterios de seleccion de instancias del [Esquema-T](#) basados en la nocion de referencia aleetica sintactica que introduce en el capitulo anterior, como se vera mas adelante.

8.2. Referencia y desentrecorrelacion

Como se vio en el cap tulo [5](#), el [Esquema-T](#) no es estrictamente necesario ni estrictamente suficiente para garantizar la utilidad logico-expresiva del predicado veritativo. Por

ejemplo, en contextos clásicos la dirección de izquierda a derecha del bicondicional, esto es,

$$\vdash p \leftrightarrow q \quad \text{(T-Out)}$$

que no engendra contradicción alguna per se, basta para ese propósito. En consecuencia, las paradojas semánticas que el [Esquema-T](#) irrestricto engendra (cf. teorema [2.9](#)) no son razón suficiente para abandonar la lógica clásica si se adopta una postura de acionista, de acuerdo con la cual el predicado de verdad solo existe para cumplir estos roles lógico-expresivos y, más aun, solo tiene sentido en el plano formal por sus usos estrictamente lógicos. No obstante, como también se señaló en el capítulo [5](#), contar adicionalmente con ciertas instancias del [Esquema-T](#) puede resultar útil para llevar a cabo otras funciones secundarias y son, en general, principios deseables en aras de la descriptividad y la completitud, en tanto no engendren inconsistencias o tengan consecuencias contraintuitivas o incorrectas.

La selección de instancias no es sencilla, en particular cuando se desea obtener una teoría correcta, abarcativa, axiomatizable, explicativa y bien motivada, i.e., cuando se busca satisfacer las condiciones (a)-(d) del proyecto minimalista, enunciadas y examinadas en el apartado [5.3.1](#). Por un lado, el teorema de McGee (teorema [3.25](#)) muestra que la mera consistencia no es un criterio viable, por ser demasiado amplio y dar lugar a conjuntos de instancias del [Esquema-T](#) demasiado complejos. Por otro, los criterios alternativos propuestos hasta el momento no resultan del todo satisfactorios tampoco. Mientras que la distinción tarskiana entre lenguaje y metalenguaje en teorías como tb (vease el apartado [2.3.3](#)) es demasiado restrictiva, los criterios basados en la idea de fundación dan lugar a teorías demasiado complejas, como la teoría semántica de Leitgeb dada por fhN^+ ; pf ig (vease el apartado [3.2](#)) y la restricción del [Esquema-T](#) a instancias donde el predicado de verdad ocurre positivamente (vease el apartado [3.1.3.2](#)) resulta demasiado arbitraria y poco explicativa.

Tanto como la fundación o la dependencia, la idea de utilizar un criterio de restricción basado en los patrones referenciales que subyacen a las oraciones de L_T tiene ciertamente gran atractivo desde el punto de vista de la condición (d), esto es, tiene poder explicativo, en tanto se cree comúnmente que la causa de las paradojas yace en ciertos patrones referenciales 'peligrosos' como la autorreferencia y, desde que la paradoja de Yablo entro en escena, la mala fundación. Como se señaló en el apartado [6.3](#), la v a no ha sido explorada aun por la falta de nociones adecuadas de referencia, en particular para lenguajes formales como L_T . Dado que he reparado esta falta en [7.3](#), el proyecto toma un nuevo impulso. Aun más, como una de las nociones de referencia aleática que introduzco, la sintáctica, es aritmética, i.e., de nivel en L_{pa}^+ , las teorías desentremilladoras

resultantes, a diferencia de aquellas basadas en las nociones de fundación, satisfarán la condición (c), es decir, recurrirán únicamente a restricciones sintáctico-aritméticas.

Habiendo ya introducido nociones apropiadas para estudiar los patrones de referencia que subyacen a las expresiones que dan lugar a paradojas semánticas, el siguiente paso es detectar las notas que estos patrones tienen en común y utilizarlas para restringir el conjunto de oraciones de L_T que generan instancias legítimas del [Esquema-T](#).

Naturalmente, toda inferencia que se haga sobre los patrones de referencia comunes a las paradojas será inductiva e informal, porque solo se tiene una muestra limitada: por un lado, oraciones autorreferenciales como las distintas versiones del mentiroso, las oraciones de Curry y de McGee, y sus correspondientes ciclos. Por otro, listas de oraciones en las cuales cada miembro refiere directamente solo a los que se encuentran más adelante. Todas estas expresiones tienen en común el ser alejadas de la fundación, ya sean autorreferenciales, como las primeras, o no, como las segundas.

No obstante, este análisis podría resultar poco relevante. Quizás no todos los patrones de mala fundación sean problemáticos, sino solo aquellos que comparten ciertas características. Más adelante se verá que este es efectivamente el caso. En consecuencia, no existe un único modo de restringir el [Esquema-T](#) de acuerdo con los patrones referenciales de las oraciones que dan lugar a sus instancias. En los cuatro apartados que siguen exploro cuatro criterios de restricción diferentes, cada uno de los cuales da lugar a una teoría desentramadora para la cual demuestro que es consistente y al menos deductivamente tan poderosa como $RT_{\leq \omega}$; $PUTB$ y KF (cf. de nociones [2.14](#), [3.19](#) y [3.13](#)), las teorías axiomáticas clásicas mejor consideradas de la literatura.

8.2.1. Verdad bien fundada

Como señalé, un criterio evidente es el que restringe las instancias del [Esquema-T](#) a oraciones bien fundadas, de acuerdo con la definición [7.28](#). De hecho, el mismo Horwich parece favorecer un abordaje por el estilo en el siguiente pasaje, ya citado en el capítulo [5](#):

The intuitive idea is that an instance of the equivalence [T]-schema will be acceptable, even if it governs a proposition concerning truth (e.g. "What John said is true"), as long as that proposition (or its negation) is grounded|i.e. is entailed either by the non-truth-theoretic facts, or by those facts together with whichever truth-theoretic facts are 'immediately' entailed by them (via the already legitimized instances of the equivalence schema), or . . . and so on. ([Horwich, 2005](#), p. 81)

Como indique entonces, si por 'grounded' se entiende fundación, ya sea en el sentido de [Kripke \(1975\)](#) o de [Leitgeb \(2005\)](#) (d-fundación, veanse los apartados [3.13.2](#)), se viola la condición (c) del proyecto minimalista. Pero si, en lugar de fundación, se entiende algo más bien sintáctico, como la buena fundación alética (de [nición 7.28](#)), lo cual es claramente una interpretación caritativa de la cita de Horwich, una teoría minimalista en esas líneas es posible.

En lugar de $RDec^{qr}(v)$, esto es, la fórmula que define la noción de r-decidibilidad en qr (cf. de [nición 7.32](#)) dada por la de [nición 7.38](#), como en el capítulo anterior, escribo directamente $RDec(v)$, por cuestiones de legibilidad. Del mismo modo, en lugar de $Bew_{qr}(v)$ escribo Bew . Los predicados $Wf(v)$, con $\langle \theta \rangle$, definen la buena fundación alética en cada qr, como indique en el apartado [7.3.2.2](#). Dados (códigos de) términos $t_1; \dots; t_n$ y una fórmula $\phi(v) \in L_T$, $\phi(v(t_1; \dots; t_n))$ abrevia la fórmula

$$\exists x (\sim s.(v; x) \wedge \lg(x) = n \wedge (x)_1 = t_1 \wedge \dots \wedge (x)_n = t_n)$$

esto es, $v(t_1; \dots; t_n)$ es el resultado de reemplazar en v la primera variable libre por lo que t_1 denota, la segunda por lo que t_2 denota, ..., la n-ésima por lo que t_n denota.

La siguiente teoría restringe las instancias del [Esquema-T Uniforme](#), esto es,

$$\exists t_1 \dots t_n (T(\phi(t_1; \dots; t_n)) \leftrightarrow \phi(t_1; \dots; t_n)) \quad (\text{Esquema-T Uniforme})$$

donde $\phi \in L_T$ tiene n variables libres, ϕ es una expresión r-decidible y bien fundada o, más precisamente, ϕ es una expresión r-decidible y bien fundada en qr θ a un enunciado r-decidible y bien fundado.¹

De [nición 8.5](#) (wfutb). wfutb L_T extiende qr θ con el siguiente esquema de axioma, para cada $\langle \theta \rangle$ y cada $\phi \in L_T$ con exactamente n variables libres:

$$wfutb \exists t_1 \dots t_n \exists x (RDec(x(t_1; \dots; t_n)) \wedge Wf(x(t_1; \dots; t_n)) \wedge$$

$$Bew_\theta(\phi(t_1; \dots; t_n) \leftrightarrow \phi(t_1; \dots; t_n)) \rightarrow (T(\phi(t_1; \dots; t_n)) \leftrightarrow \phi(t_1; \dots; t_n)))$$

wfutb (por 'Well-founded Uniform Tarski Biconditionals') se limita a oraciones r-decidibles porque así lo requiere la prueba del [corolario 8.2](#), que ayuda a probar la ω -consistencia de la teoría a continuación. wfutb establece que, si $\langle \theta \rangle$ y el resultado de sustituir en una fórmula ϕ sus (posibles) n variables libres por los términos $t_1; \dots; t_n$ es una oración r-decidible y bien fundada en qr, entonces toda oración ϕ a la cual es equivalente en qr θ tiene una instancia correspondiente del [Esquema-T](#). Esto incluye,

¹Vease el apartado [1.2.3](#) para detalles de notación.

además de las oraciones r -decidibles bien fundadas, oraciones como

$$\exists x((T \mid \mid T \mid) \wedge x = p_0 = 0q \mid T x)$$

y

$$\exists x(T x \mid T x)$$

que, de acuerdo con la definición [7.22](#), pertenecen a todo y , por tanto, no están alejadas bien fundadas. Por otro lado, excluye inevitablemente muchas oraciones intuitivamente seguras, v.g., $\exists x(\text{Bew}_{pa^+}(x) \mid T x)$, aunque esto no debe confundirse con los teoremas de pa^+ , que sí tienen instancias correspondientes del [Esquema-T](#) en wf_{tb} .

Proposición 8.6. wf_{tb} es \exists -consistente.

Demostración. Por la proposición [3.35](#) y el corolario [8.2](#), dado que pf está cerrado bajo equivalencia, como indica la proposición [3.29](#). \square

De hecho, vale también el siguiente resultado.

Observación 8.7. $wf_{tb} + \text{T-Out}$ es \exists -consistente.

Demostración. Por el corolario [3.38](#) y los corolarios [8.2](#), [8.3](#) y [8.4](#). \square

wf_{tb} y las teorías que presento más adelante, que son extensiones de wf_{tb} , son deductivamente poderosas. Por ejemplo, prueban los mismos teoremas de L_{pa}^+ que $rt_{< \theta}$, porque son capaces de definir, para cada ordinal $< \theta$ predicados de verdad que satisfagan los axiomas de esta teoría.

Proposición 8.8. $rt_{< \theta}$ es relativamente interpretable en wf_{tb} .²

Demostración. Para probarlo, muestro que los predicados T con $< \theta$ de $rt_{< \theta}$ son definibles en wf_{tb} (definición [1.1](#)). Si $\nu(v)$ es una fórmula de L_T que contiene T , sea $L' \subseteq L_T$ el lenguaje que resulta de extender L_{pa}^+ con $\nu(v)$ como un símbolo primitivo y cerrarlo bajo operadores lógicos, como es usual. T aparece en L' solo en el contexto de $\nu(t)$ para cualquier término t . La relación entre (los códigos de) una fórmula v y una oración u de L_T que se da cuando u es una oración de L_v es claramente recursiva y, por tanto, representable en pa^+ por una fórmula $\text{Sent}(u; v)$.² Considerese la siguiente

²Vease la definición [1.1](#).

formula de L_T :

$$\begin{aligned}
 & \exists x_1 \exists x_2 (x_1 = x_2 \wedge x_2 = x_1) \quad (8.1) \\
 & \exists u ((\text{Sent}_{L_{\text{pa}}}(u) \wedge \exists y \text{Sent}(u; w(\neg=pyq))) \wedge x = :.u \wedge :T w(y=\bar{p}yq)(u=\bar{p}xq)) \wedge \\
 & \exists u; v((\text{Sent}_{L_{\text{pa}}}(u!.v) \wedge \exists y (\text{Sent}(u; w(\neg=pyq)) \wedge \text{Sent}(v; w(\neg=pyq)))) \wedge \\
 & x = (u!.v) \wedge (T w(y=\bar{p}yq)(u=\bar{p}xq) \wedge T w(y=\bar{p}yq)(v=\bar{p}xq))) \wedge \\
 & \exists u; v((\text{Sent}_{L_{\text{pa}}}(8.vu) \wedge \exists y \text{Sent}(8.vu; w(\neg=pyq))) \wedge x = 8.vu \wedge \\
 & \exists t T w(\bar{y}=pyq)(\bar{p}u(t=\bar{v})q=pxq)) \wedge \\
 & \exists t \exists y ((\text{Sent}_{L_{\text{pa}}}(t) \wedge \exists \text{Sent}(t; w(\neg=pyq))) \wedge x = w(\neg=pyq)(t=pxq) \wedge \\
 & T w(\neg=pyq)(t=pxq))
 \end{aligned}$$

Aplicando el lema de diagonalización sobre w , se obtiene un predicado $\#(x; y)$, que noto $\#_y(x)$, tal que

$$\text{pat } \#_y(x) \text{ } \exists t \exists t^1 ; t(x = (t = t^2) \wedge t^1 = t) \quad (8.2)$$

$$\begin{aligned}
 & \exists u (\text{Sent}_{<y}(x) \wedge x = :.u \wedge :T p\#_y(\bar{u}\bar{q}) \wedge \\
 & \exists u; v (\text{Sent}_{<y}(u) \wedge \text{Sent}_{<y}(v) \wedge x = (u!.v) \wedge (T p\#_y(u\bar{q}) \wedge T p\#_y(v\bar{q}))) \wedge \\
 & \exists u; v (\text{Sent}_{<y}(8.vu) \wedge x = 8.vu \wedge \exists t T p\#_y(\bar{u}(t=\bar{v}))q) \wedge \\
 & \exists t \exists y (\text{Sent}_{<}(t) \wedge x = p\#_{\bar{y}}(t)q \wedge T p\#_{\bar{y}}(t)q)
 \end{aligned}$$

donde $\text{Sent}_{<y}(u)$ es un modo abreviado de escribir $\text{Sent}_{L_{\text{pa}}}(u) \wedge \exists v \text{Sent}(u; p\#_{\bar{y}}(u)q)$ para cualquier variable sobre ordinales .

A continuación muestro que, si $1 < \alpha < \omega$, existen predicados $\#(x)$ para cada α que satisfacen los axiomas de $rt_{<}$. Para hacerlo, pruebo que

$$\exists p \exists q \exists t (T \#_{\alpha}(t) \wedge \#_{\alpha}(t)) \quad (8.3)$$

es un teorema de wf_{utb} , que a su vez, en virtud de su axioma, se sigue de

$$\exists t (RDec(p\#_{\bar{y}}(t)q) \wedge Wf(p\#_{\bar{y}}(t)q)) \quad (8.4)$$

para algún $\alpha < \omega$, que demuestro a continuación.

El lado derecho del bicondicional en (8.3) es una variante notacional de

$$\begin{aligned} & :(\exists t_1; t_2(x = (t_1 \neq t_2) \wedge t_1 \neq t_2) \wedge \exists u : \\ & (\text{Sent}_{<y}(u) \wedge x \neq u) \wedge \exists p \#_y(u)q)) \wedge \dots \\ & (\exists u; v : (\exists (\text{Sent}_{<y}(u) \wedge \text{Sent}_{<y}(v)) \wedge x \neq (u \vee v)) \wedge \exists (T p \#_y(u)q \wedge T p \#_y(v)q)) \wedge \exists u; \bar{v} : \\ & (\text{Sent}_{<y}(\exists v u) \wedge x \neq \exists v u) \wedge \exists t T p \#_y(u(t \neq \bar{v}))q)) \wedge \dots \\ & \exists t \exists y : (\text{Sent}_{<}(t) \wedge x \neq p \#_y(t)q) \wedge \exists T p \#_y(t)q \end{aligned}$$

con la cual directamente identifico $\#_y(x)$ (habiendo aplicado el lema diagonal fuerte a (8.1) sin abreviaturas). Luego, para todo termino t y todo $\#, \#(t)$ es claramente r-decidible (y se puede demostrar en pa^+ , como indica la proposicion 7.36) y pertenece directamente solo a formulas de la forma $\#(s)$ con $\#,$ esto es, $\#(t)$ es r-decidible.

Falta probar que, para todo t y todo $\#, \#(t)$ esta aleticamente bien fundada en qr para algun θ . Como todas las formulas de esta forma son r-decidibles, si estan bien fundadas en qr tambien lo estaran en qr para todo $>$. Pruebo el resultado mediante dos inducciones internas anidadas. La principal, una induccion transnita sobre los $\#$ lo cual es posible por lo dicho en el apartado 1.4, (IT θ); la secundaria, sobre la complejidad de la oracion codificada por t . Si $\text{Sent}_{<}(t)$, el resultado es trivial porque $\#(t)$ no pertenece a nada. Si t es una formula atomica [i.e., una identidad o una formula de la forma $\#(s)$, donde $\#$ su complejidad es 0. Cada simbolo logico que se agregue a t aumenta su complejidad en 1.

Supongase que, si $\#, \#(t)$ esta bien fundada en alguna teoria de la jerarquia para todo t . Hay que ver que, para cada $t, \#(t)$ est bien fundada. Si $t = (t_1 \neq t_2)$, el resultado es trivial. Si $t = p \#(s)q$, donde $\wedge \text{Sent}_{<}(s), \#(t)$ pertenece exclusivamente a $\#(s)$ que, por hipotesis inductiva, esta bien fundada. Luego, $\#(t)$ tambien. Supongase ahora que para todo s que denota una formula de menor complejidad que $t \#(s)$ esta aleticamente bien fundada. Si $t = \exists v u, \#(t)$ pertenece unicamente a las oraciones de la forma $\#(u(s=v))$. Como la complejidad de $u(s=v)$ es menor a la de t pues contiene un cuantificador universal menos, para todo $s \#(u(s=v))$ esta bien fundada y, por lo tanto, tambien lo esta $\#(t)$. Los casos para la negacion y el condicional son analogos.

Ahora se esta en condiciones de probar en wftb todos los axiomas de $\text{rt}_{<}$. Lo muestro solo para $\text{rt}_{<4}, \text{rt}_{<5}$ y $\text{rt}_{<6}$. Los otros casos son analogos al de $\text{rt}_{<4}$. Sean $<$

$$\begin{aligned} \text{Sent}_{<}(\exists v u) \wedge (\#(\exists v u) \wedge \exists t T p \#(u(t \neq \bar{v}))q) & \text{ por (8.2), 4ta linea, y =} \\ \wedge \exists t \#(u(t=v)) & \text{ por (8.3)} \end{aligned}$$

Esto completa la prueba de $rt < 4$.

$$\text{Sent} \quad (t) ! (\# (p\# (t.)q) \$ T p\# (t.)q) \quad \text{por (8.2), 5ta l nea, y =}$$

$$< \quad \$ \# (t)) \quad \text{por (8.3)}$$

Esto completa la prueba de $rt < 5$.

Pruebo el axioma $rt < 6$ mediante una induccion interna sobre la complejidad de la formula denotada por t , donde $\text{Sent} < (t)$. Si $t = (t_1 = t_2)$,

$$\# (p\# _ (t.)q) \$ T p\# _ (t.)q \quad \text{por (8.2), 5ta l nea, y =}$$

$$\$ \# (t_1 = t_2) \$ \quad \text{por (8.3) por (8.2), 1ra l}$$

$$t_1 = t_2 \$ \# \quad \text{nea, y =}$$

$$(t_1 = t_2) \quad \text{por (8.2), 1ra l nea, y =}$$

Si $t = p\# (s)q$, donde

$$\# (p\# _ (t.)q) \$ T p\# _ (t.)q \quad \text{por (8.2), 5ta l nea, y =}$$

$$\$ \# (p\# (s)q) \quad \text{por (8.3) por (8.2), 5ta l}$$

$$\$ \# (s) \quad \text{nea, y =}$$

$$\$ \# (p\# (s)q) \quad \text{por (8.2), 5ta l nea, y =}$$

Supongase ahora que, para todo s que denota una formula de menor complejidad que t , $\# (\# _ (s)) \$ \# (s)$. Si $t = (s)$,

$$\# (\# _ (s)) \$ \# (s) \quad \text{por (8.2), 5ta l nea; y =}$$

$$\# (\# _ (s)) \$ \# (s) \quad \text{por (8.3) por (8.2), 2da l}$$

$$\# (\# _ (s)) \$ \# (s) \quad \text{nea, y =}$$

$$\# (\# _ (s)) \$ \# (s) \quad \text{por (8.2), 5ta l nea, y =}$$

$$\# (\# _ (s)) \$ \# (s) \quad \text{Hipotesis inductiva por}$$

$$\# (\# _ (s)) \$ \# (s) \quad \text{(8.2), 2da l nea, y =}$$

Los otros casos son analogos. Esto completa la prueba de $rt < 6$.

Consiguientemente, para cada $1 < \theta$, wf_{θ} contiene predicados que satisfacen los axiomas de $rt < \theta$ y, por ende, esta teoria puede definir los predicados de verdad de $rt < \theta$.

Estrictamente hablando, los predicados $\# (x)$ no son definiciones de los predicados $T(x)$ de $rt < \theta$ porque las oraciones a las cuales se aplican son sintacticamente diferentes:

mientras que los primeros se aplican a ciertas oraciones de $L_{< \theta}$, los segundos a oraciones de $L_{\leq \theta}$. Para circunvenir este problema y completar la prueba de interpretabilidad relativa es preciso traducir las oraciones de $L_{< \theta}$ a L_T reemplazando las ocurrencias de los predicados $T(t)$ por $\#(t)$ y estipular que una oración de $L_{< \theta}$ es verdadera si su traducción lo es. Es preciso reemplazar, no solo los predicados de verdad que ocurren en cada oración de $L_{< \theta}$ sino los que ocurren en las expresiones a las cuales estas oraciones se refieren. Por ejemplo, $T!(pT_0(p0 = 0q)q)$ no puede traducirse como $\#!(pT_0(p0 = 0q)q)$ sino que es necesario también reemplazar T_0 por $\#_0$ dentro del término cerrado $pT_0(p0 = 0q)q$. Para eso es posible aplicar el teorema de recursión de Kleene para obtener la siguiente función de traducción: $L_{< \theta} \rightarrow L_T$.³

$$\begin{array}{l}
 \text{si } \varphi \in L_{pa}^+ \text{ entonces } \varphi \in L_{< \theta} \\
 \text{si } \varphi := \varphi \\
 \text{si } \varphi := ! \text{ si } \varphi := \\
 \text{si } \varphi := T t \\
 \text{si } \varphi := \# (\cdot (t))
 \end{array}$$

donde \cdot es un término que representa en pa^+ (asumo que L_{pa}^+ tiene de hecho símbolos de función suficientes como para construir este término). Bajo esta traducción, es sencillo probar que, si $rt_{< \theta} \vdash \varphi$, $wf_{utb} \vdash (\varphi)$. □

Corolario 8.9. kf y $putb$ son relativamente interpretables en wf_{utb} .

Demostración. Directamente, por los teoremas 3.17 y 3.23, la proposición 8.8 y la transitividad de la interpretabilidad relativa. □

Luego, wf_{utb} es (a) $!$ -consistente, (b) lo suficientemente abarcativa como para interpretar relativamente kf , i.e., una teoría composicional de la verdad, a pesar de ser wf_{utb} una teoría desentremilladora, (c) que emplea criterios de restricción lo suficientemente simples como para ser definibles en L_{pa}^+ , esto es, criterios sintactico-aritméticos, y (d) lo suficientemente bien motivada tanto por la cita de Horwich a comienzos del capítulo como por la tesis referencialista y el análisis de ciertas paradojas como el mentiroso, Curry, Yablo y diversas variaciones de esta última, en el apartado 7.3. Luego, puede considerarse una muy buena teoría minimalista. Asimismo, al ser consistente con T-Out, $wf_{utb} + \text{T-Out}$ es una buena teoría de acciónista de la verdad.

Si bien wf_{utb} es capaz de probar los axiomas composicionales de $rt_{< \theta}$ para fórmulas $\#(v)$ de L_T para cada θ , eso no quiere decir que ella misma sea una teoría a

³Vease [Halbach \(2011\)](#), p. 37, lema 5.2).

composicional de la verdad, porque puede que no pruebe ningún principio composicional para su propio predicado de verdad, T . De hecho, al ser compatible con [T-Out](#), es sencillo ver que w_{futb} no prueba el principio composicional que establece la conmutatividad de T con la negación,

$$\exists x(\text{Sent}_{LT}(x) \wedge (T : x \wedge \neg T x)) \tag{8.5}$$

ni el que expresa la validez interna de las reglas de introducción y eliminación del condicional:

$$\exists x(\text{Sent}_{LT}(x!y) \wedge (T(x!y) \wedge (T x \wedge T y))) \tag{8.6}$$

porque [T-Out](#) es inconsistente con ambos. Por un lado, la instancia de [T-Out](#) generada por $\neg T I$,

$$T I \wedge \neg T I$$

es equivalente a $\neg(T : \neg T I) \wedge \neg T I$, incompatible con

$$\exists x(\text{Sent}_{LT}(x) \wedge (T : x \wedge \neg T x)) \tag{8.7}$$

Por otro, la instancia de [T-Out](#) generada por $T(c^0 \wedge \neg p \wedge q \wedge 0 = 0)$ es

$$T(c^0 \wedge \neg p \wedge q \wedge 0 = 0) \wedge \neg T(c^0 \wedge \neg p \wedge q \wedge 0 = 0)$$

que, junto con (8.6) implica $T(c^0 \wedge \neg p \wedge q \wedge 0 = 0) \wedge (T(c^0 \wedge \neg p \wedge q \wedge 0 = 0) \wedge \neg T(c^0 \wedge \neg p \wedge q \wedge 0 = 0))$. Como $0 = 0$ también tiene una instancia asociada del [Esquema-T](#) en w_{futb} , que implica que $\neg T p \wedge 0 = 0$, se tiene que $\neg T c^0$ y, por tanto, que $T c^0 \wedge \neg T 0 = 0$. Luego, nuevamente por (8.6), esto implicaría a $T c^0$, lo cual es un absurdo.

Nada de esto permite descartar la posibilidad de que w_{futb} sea capaz de probar principios composicionales tipeados, como

$$\exists x(\text{Sent}_{L+pa}(x) \wedge (T : x \wedge \neg T x))$$

y

$$\exists x(\text{Sent}_{L+pa}(x!y) \wedge (T(x!y) \wedge (T x \wedge T y)))$$

aunque conjeturo que no es el caso, basándome en un resultado análogo para putb en [Halbach \(2009, x6\)](#).

w_{futb} es el primer fruto del trabajo argumentativo y propositivo que he realizado a lo largo de toda la tesis. Por un lado, el análisis conceptual sobre la posición de acionista acerca de la verdad, que desembocó en la necesidad de elaborar teorías formales de la verdad que cumplan las condiciones (1)-(3) en el apartado 5.3. Por otro, la exploración de diferentes nociones de referencia, que dio por resultado una noción aletica

y sintáctica en el apartado 7.3.2, no solo adecuada para el estudio de las paradojas que afectan a las teorías formales de la verdad, sino para la formulación de criterios restrictivos sintácticos del Esquema-T que se adapten a condiciones de acionistas. Ambos dieron lugar a este sistema formal desentrecorillador que no solo cuadra perfectamente con las condiciones (1)-(3) al expandirse con T-Out sino que, por sí solo, es un perfecto candidato para el proyecto minimalista de Horwich presentado en el apartado 5.3.1, a la vez que permite verificar la tesis referencialista de las paradojas semánticas. En los apartados restantes continuo en esta dirección, puliendo algunas aristas de este primer resultado y proveyendo nueva evidencia en favor de la tesis referencialista.

8.2.2. Verdad y referencia nita

Una mirada un poco más de cerca a las expresiones que dan lugar a paradojas y !-paradojas revela que en el caso de oraciones no autorreferenciales estas siempre refieren directamente a un número nito de otras oraciones, como las oraciones de Yablo en

$$(n) = \text{p}8x > n : T(x)q \tag{8.8}$$

y sus variantes, como

$$(n) = \text{p} : 8x(x > n ! T(x))q$$

y

$$(n) = \text{p}8(x > n ! : 8y(y > x ! T(y)))q$$

Luego, un segundo criterio un poco más refinado pero arriesgado podría consistir en incluir todas las oraciones fundadas y todas las mal fundadas que no sean autorreferenciales y solo refieren directamente a un número nito de otras oraciones.

La siguiente fórmula de L_T expresa que la oración v refiere directamente y alelicamente en qr solo a un número nito de oraciones:

$$Rf \text{ in } (v) := \text{Sent}_{L_T}(v) \wedge \exists x(\text{Seq}(x) \wedge \exists y(R_d(v; y) \wedge \exists i \text{ lg}(x)(x)_i = y))$$

Literalmente, v es una oración de L_T y hay una secuencia nita de números que contiene todos los códigos de cada una de las oraciones a las cuales v refiere alelicamente y directamente en qr .

Definición 8.10 (frutb). frutb L_T extiende wfutb con el siguiente esquema de axioma, para cada $\langle \sigma \rangle$ y cada σ' con exactamente n variables libres:

frutb $\exists t_1 \dots \exists t_n (RDec(p'(t_1, \dots, t_n)q) \wedge SR(p'(t_1, \dots, t_n)q) \wedge Rf \text{ in } (p'(t_1, \dots, t_n)q) !$
 $(T p'(t_1, \dots, t_n)q \text{ } \$ (t_1, \dots, t_n)))$

frutb (por Finite Reference Uniform Tarski Biconditionals) materializa una version debilitada del criterio de seleccion propuesto; en tanto no es posible dejar de lado el requisito de la r-decidibilidad, por las mismas razones que antes. El axioma frutb establece que, si $\langle \theta \rangle$ y el resultado de sustituir en ' sus n variables libres con los terminos t_1, \dots, t_n es una oracion r-decidible, no aleticamente autorreferente y re ere aletica y directamente a un numero nito de oraciones en qr , entonces esta oracion tiene una instancia correspondiente del [Esquema-T](#). Esto incluye, por ejemplo, las cade-nas de oraciones mal fundadas que se obtienen aplicando la proposicion [1.11](#) a formulas que contienen una variable libre bajo el alcance de T , como la cadena de honestos

$$\begin{aligned} t_0 &= pT t_1q \\ t_1 &= pT t_2q \\ t_2 &= pT t_3q \\ &\vdots \end{aligned}$$

y tambien

$$\begin{aligned} l_0 &= p:T l_1q \\ l_1 &= p:T l_2q \\ l_2 &= p:T l_3q \\ &\vdots \end{aligned}$$

pero no las secuencias yablezcas, donde cada oracion es mal fundada y re ere directamente a un numero in nito de otras expresiones.

Proposicion 8.11. frutb es !-consistente.

Demostracion. Por la proposicion [3.37](#) y los corolarios [8.2](#), [8.3](#) y [8.4](#). □

Este resultado puede reforzarse aun mas.

Corolario 8.12. frutb + T-Out es !-consistente.

Demostracion. Por los corolarios [3.38](#), [8.2](#), [8.3](#) y [8.4](#). □

Corolario 8.13. $rt_{\langle \theta \rangle}$, kf y putb son relativamente interpretables en frutb.

Demostración. Directamente, por la proposición [8.8](#), el corolario [8.9](#) y el hecho de que $\text{frutb} \models \text{wfutb}$. \square

No obstante, quizás frutb permite instancias del [Esquema-T](#) que no son del todo deseables. Por ejemplo, para oraciones que, si bien no son autorreferenciales, se refieren a otras que sí lo son. La oración del mentiroso (reforzado) !T , donde

$$q^+ \cdot \text{!} = p:\text{T} \text{!} q$$

es claramente autorreferente y, por ende, no tiene una instancia asociada del [Esquema-T](#) en frutb . Si la tuviera, la teoría se haría inconsistente, como muestra el teorema de Tarski (teorema [2.9](#)). Sin embargo,

$$\exists x(x = \text{!} \text{!} : \text{T} x)$$

que es lógicamente equivalente a !T , sí tiene una instancia del [Esquema-T](#) en la teoría, porque es rd -decidible, no autorreferente y solo se refiere a !T . Si bien no da lugar a inconsistencias, casos extraños de este tipo implican que la teoría es incompatible con un axioma que cierre la aplicación del predicado veritativo bajo equivalencia. No es posible decir, so pena de caer en contradicción, que si una oración es equivalente a otra que satisface el predicado $\text{RDec}(v) \wedge \text{!SR}(v) \wedge \text{Rf in}(v)$ entonces hay una instancia del [Esquema-T](#) para ella.

Otro caso extraño está dado por el mentiroso alternativo

$$q^+ \cdot \text{!}^0 = p:\text{T} : \text{!}^0 q$$

Como $:\text{!}^0$ no denota $\text{T} : \text{!}^0$ sino $:\text{T} : \text{!}^0$, $\text{T} : \text{!}^0$ no es autorreferente, como se vio en el apartado [7.3.1.1](#); frutb contiene la instancia del [Esquema-T](#) dada por $\text{T} : \text{!}^0$. No obstante,

$:\text{T} : \text{!}^0$ sí resulta autorreferente y, por ende, no hay una instancia de este principio para ella en la teoría. Mientras que la instancia del [Esquema-T](#) para esta última oración,

$$\text{T} p:\text{T} : \text{!}^0 q \text{ \$ } :\text{T} : \text{!}^0$$

da lugar inmediatamente a una inconsistencia, porque $p:\text{T} : \text{!}^0 q = :\text{!}^0$, la instancia para $\text{T} : \text{!}^0$,

$$\text{T} p:\text{T} : \text{!}^0 q \text{ \$ } \text{T} : \text{!}^0;$$

no genera inconsistencias sino simplemente establece que frutb es incompatible con el principio composicional ([8.5](#)), al igual que kf y putb (veanse los apartados [3.1.3.1](#) y [3.1.3.2](#)). Mientras que en kf y putb esto puede estar motivado por la lógica interna

paracompleta, no parece que en este contexto esté disponible una motivación semejante para frutb .

Asimismo, frutb es incompatible con [8.6](#), porque la oración de Curry $T(c^0!_p0 \neq 0q)$ es rd -decidible y refiere únicamente a $T(c^0!_p0 \neq 0q) \neq 0 \neq 0$, que es también rd -decidible y solo refiere a sí misma. Luego, $T(c^0!_p0 \neq 0q)$ no es autorreferente y, por ende, tiene una instancia correspondiente del [Esquema-T](#) en frutb , que implica

$$T c^0 \ \$ \ T (c^0!_p0 \neq 0q):$$

Luego, junto con [\(8.6\)](#), $T c^0 \ \$ \ (T c^0 \ ! \ T p0 \neq 0q)$. Como $0 \neq 0$ también tiene una instancia asociada del [Esquema-T](#) en frutb , que implica que $:T p0 \neq 0q$, se tiene que $T c^0 \ \$ \ :T c^0$.

Si se desea contar con ciertos principios composicionales para T , frutb resulta inaceptable. Sin embargo, dado que la prioridad desde el punto de vista de accionista es contar con [T-Out](#), que ya bloquea la posibilidad de adoptar principios como [\(8.5\)](#) y [\(8.6\)](#), la incompatibilidad de frutb con ellos no parece ser un problema real.

No obstante, desde una perspectiva filosófica, el criterio de restricción de frutb puede resultar demasiado permisivo, porque ciertas oraciones mal fundadas, a pesar de no ser autorreferenciales y referir directamente a un número finito de otras expresiones, exhiben patrones de referencia claramente 'peligrosos'. Por ejemplo, $T I$ no es autorreferente y refiere directamente solo a la oración del mentiroso, pero parece a pesar de eso contener algún tipo de circularidad. Luego, frutb podría estar violando la condición (d) del proyecto minimalista de Horwich.

8.2.3. Verdad Equivalente

Si contar con la posibilidad de cerrar el predicado veritativo bajo equivalencia o agregar al menos las versiones parciales de los principios composicionales en [\(8.7\)](#) es un requisito, o resulta contraintuitivo incluir instancias del [Esquema-T](#) generadas por oraciones como $\exists x(x = I \ ! \ :T x)$, $T I$ o $T :I^0$, se puede elaborar un tercer criterio, de alguna manera más restrictivo que el anterior, que permita instancias del [Esquema-T](#) generadas solo por oraciones o bien fundadas o que refieren directamente solo a un número finito de expresiones, ninguna de las cuales sea autorreferencial. Naturalmente, se perderán por un lado ciertas instancias del [Esquema-T](#), pero a la vez será posible extender los axiomas resultantes a oraciones equivalentes a las permitidas, ganando nuevas instancias por este otro lado.

Definición 8.14 (erutb). erutb L_T extiende wfutb con el siguiente esquema de axioma, para cada θ y cada φ con exactamente n variables libres:

$$\text{erutb } \exists t_1 \dots t_n (R \text{Dec}(x(t_1; \dots; t_n)) \wedge \exists y (SR(y) \wedge R(x(t_1; \dots; t_n); y)) \wedge R \text{f} \\ \text{in}(x(t_1; \dots; t_n)) \wedge \text{Bew}_\theta(p'(t_1; \dots; t_n) \rightarrow x(t_1; \dots; t_n)) \wedge \\ (T(p'(t_1; \dots; t_n) \rightarrow \varphi'(t_1; \dots; t_n))))$$

erutb (por Equivalent Reference Uniform Tarski Biconditionals) establece que, si $\theta < \theta$ y el resultado de sustituir en una fórmula x sus (posibles) n variables libres por los términos $t_1; \dots; t_n$ es una oración r -decidible, que no refiere a ninguna oración autorreferencial y solo refiere directamente a un número finito de expresiones en qr_θ , entonces toda oración a la cual es equivalente en qr_θ tiene una instancia correspondiente del [Esquema-T](#).

Proposición 8.15. erutb es ω -consistente.

Demostración. Sea L_T como en las hipótesis de la proposición [3.37](#) y, además, tal que, si $\varphi \in \Sigma$ y pertenece al conjunto del cual φ depende esencialmente (de definición [3.30](#)), entonces $\varphi \in \Sigma$. Sea el conjunto de todas las oraciones de L_T que son equivalentes a algún miembro de Σ . Luego,

1. Σ , por ende, es finito.
2. Todas las oraciones de Σ dependen esencialmente de un conjunto finito, porque son equivalentes en N^+ a una oración en Σ y oraciones equivalentes dependen de los mismos conjuntos, como indica la proposición [3.29](#).
3. Ninguna de las oraciones en Σ es autorreferencial. Si alguna $\varphi \in \Sigma$ lo fuera, se tendría, por un lado, que $\varphi \in \Sigma$ y que existen oraciones $\varphi_1; \dots; \varphi_n$ de L_T tales que, para todo ψ del cual φ depende, $\varphi_1 \in \Sigma$; si $1 < i < n$, para todo ψ del cual φ_i depende, $\varphi_{i+1} \in \Sigma$; y, para todo ψ del cual φ_n depende, $\psi \in \Sigma$. Por otro lado, habrá una oración $\psi \in \Sigma$ equivalente a φ en N^+ que, por la proposición [3.29](#), depende de los mismos conjuntos que φ , con lo cual, φ_1 pertenece al conjunto del que depende esencialmente. Y como está cerrado bajo dependencia, $\varphi_1 \in \Sigma$ y, por ende, $\varphi_2 \in \Sigma$, etc., lo que implica que $\varphi \in \Sigma$, una contradicción.

En consecuencia, también satisface las hipótesis de la proposición [3.37](#). Por los corolarios [8.2](#), [8.3](#) y [8.4](#), erutb es ω -consistente. \square

Nuevamente, este resultado puede reforzarse.

Corolario 8.16. $er_{\text{tb}} + \text{T-Out}$ es !-consistente.

Demostración. El modelo de la aplicación de la proposición [3.37](#) en la prueba de la proposición anterior es un modelo de [T-Out](#). \square

Corolario 8.17. $rt_{< \theta}$, k_f y put_{b} son relativamente interpretables en er_{tb} .

Demostración. Directamente, por la proposición [8.8](#), el corolario [8.9](#) y el hecho de que $er_{\text{tb}} \text{ wfut}_{\text{b}}$. \square

Luego, er_{tb} es (a) !-consistente, (b) aun más abarcativa que wfut_{b} y, a fortiori, lo sucientemente abarcativa como para interpretar relativamente las teorías más populares en la literatura, (c) que emplea criterios sintáctico-aritméticos, y (d) lógicamente bien motivada, basada en el análisis de las paradojas semánticas conocidas y la tesis referencialista. En consecuencia, er_{tb} también es una excelente candidata a teoría minimalista. Además, por ser consistente con [T-Out](#), $er_{\text{tb}} + \text{T-Out}$ es asimismo una buena teoría de acciónista de la verdad.

8.2.4. Verdad y autorreferencia

Finalmente, también es posible dar una teoría consistente de la verdad dada por instancias del [Esquema-T](#) generadas por oraciones que simplemente no sean aleléricamente autorreferentes [o, mejor dicho, no autorreferentes y r-decidibles, por razones semejantes a las dadas en las teorías que introduje previamente] además de aquellas dadas por oraciones bien fundadas.

Notese que las paradojas semánticas estudiadas en las cuales no hay oraciones autorreferentes involucradas no dan lugar a inconsistencias sino solo a !-inconsistencias. Si lo único que importa es evitar contradicciones explícitas, si que la teoría carezca de modelo estándar y tenga consecuencias cuanto menos contraintuitivas no resulta ser un requisito imprescindible en ciertos casos, la siguiente teoría podrá ser una opción.

Definición 8.18 (srt_{b}). $\text{srt}_{\text{b}} \perp_{\text{T}}$ extiende wfut_{b} con el siguiente esquema de axioma, para cada θ y cada oración $\phi \in \perp_{\text{T}}$:

$$\text{srt}_{\text{b}} \text{ RDec } (\phi \rightarrow \theta) \wedge \text{SR } (\phi \rightarrow \theta) \rightarrow (\text{T } \phi \rightarrow \theta)$$

srt_{b} (por Self-Reference Tarski Biconditionals) establece simplemente que si ϕ es r-decidible y no autorreferencial entonces tiene una instancia correspondiente del

[Esquema-T](#). Si bien frtb no es una subteoría de srtb porque adopta un axioma uniforme, todas las instancias de frtb se siguen evidentemente de srtb . En consecuencia, srtb también es incompatible con un principio que cierre el axioma srtb bajo equivalencia, y con los principios composicionales (8.5) y (8.6).

Corolario 8.19. srtb es consistente.

Demostración. Por compacidad, la proposición 3.36 y el lema 8.3. □

Corolario 8.20. $\text{srtb} + \text{T-Out}$ es consistente.

Demostración. Nuevamente, por compacidad, la proposición 3.36 y los corolarios 3.38 y 8.3. □

Corolario 8.21. $\text{rt}_{< \omega}$, kf y putb son relativamente interpretables en srtb .

Demostración. Directamente, por la proposición 8.8, el corolario 8.9 y el hecho de que srtb wfutb . □

No obstante srtb no tiene modelo estándar y, por ende, tampoco $\text{srtb} + \text{T-Out}$.

Proposición 8.22. srtb es ω -inconsistente.

Demostración. Es posible probar en qr_{ω} que cada una de las oraciones de Yablo en (8.8) es r -decidible y no alelicamente autorreferente en pat . Luego, srtb prueba el principio LYDP, i.e.,

$$\exists y > n: \text{T}(y) \leftrightarrow \forall y > n: \text{T}(y) \quad (\text{LYDP})$$

para cada $n \in \mathbb{N}$, que, por la proposición 6.1, es ω -inconsistente sobre pat srtb . □

Si la posibilidad de interpretar la teoría mediante una extensión de \mathbb{N}^+ no es un requisito, srtb es una teoría atractiva por aplicar un criterio de selección de instancias del [Esquema-T](#) muy inclusivo y relativamente sencillo. Existen en la literatura otras teorías de la verdad ω -inconsistentes, como fs (vease el apartado 3.3), e incluso teorías semánticas de la verdad, como la teoría de la revisión $t^{\#}$ de Gupta y Belnap (cf. [Gupta & Belnap \(1993\)](#), cap. 6)), ambas con principios prima facie intuitivos pero usualmente criticadas precisamente por carecer de modelo estándar (cf. [Leitgeb \(2007\)](#), [Barrio \(2010\)](#) y [Barrio & Picollo \(2013\)](#)). En esta tesis no me pronuncio al respecto. En cualquier caso, la elección de la teoría de la verdad que se haga dependerá de los objetivos que con ella quieran alcanzarse. Desde una perspectiva de acionista, srtb resulta inaceptable, porque, como argumenté en el apartado 3.1.3.1, prueba enunciados falsos que involucran

el predicado veritativo y, por ende, excede de alguna manera la función básica de la verdad: no satisface la condición (2) del de acionismo.

A diferencia de las teorías frtb y erutb , srtb no implica instancias del [Esquema-T Uniforme](#) sino solo del [Esquema-T](#) para oraciones que no son equivalentes a expresiones r -decidibles y bien fundadas. Si en su lugar se adoptaran instancias uniformes, la teoría resultante sería simplemente inconsistente.

Definición 8.23 (srtb). srtb extiende wfutb con el siguiente esquema de axioma, para cada $n < \omega$ y cada τ con exactamente n variables libres:

$$\text{srtb} \vdash \forall t_1, \dots, t_n (\text{RDec}(p'(t_1, \dots, t_n)q) \wedge \text{SR}(p'(t_1, \dots, t_n)q)) \rightarrow \neg (T p'(t_1, \dots, t_n)q \leftrightarrow p'(t_1, \dots, t_n)q))$$

Proposición 8.24. srtb es inconsistente.

Demostración. Es posible probar en qr_ω que

$$\forall t (\text{RDec}(p_8 y > t : T(y)q) \wedge \text{SR}(p_8 y > t : T(y)q))$$

esto es, que todas las oraciones de la secuencia de Yablo en (8.8) son r -decidibles y no aleticamente autorreferentes en pat . Luego, se puede probar en srtb el [UYDP](#), i.e.,

$$\forall q (\exists y (T p_8 y > t : T(y)q \wedge y > t : T(y)q) \rightarrow \neg (T p_8 y > t : T(y)q)) \quad (\text{UYDP})$$

que, por la proposición 6.2, es inconsistente sobre pat srtb . □

Luego, srtb (por Self-Reference Uniform Tarski Biconditionals) es ciertamente inviable. La interpretación que se debe hacer de este resultado de inconsistencia no es obvia. Por un lado, podría pensarse que entonces la autorreferencia no es el único patrón que estrictamente da lugar a paradojas, esto es, contradicciones y no meramente \neg -inconsistencias, contrario a la evidencia que provee el análisis de los patrones referenciales expresiones paradójicas en el apartado 7.3.

Otro modo de interpretar el resultado de la proposición 8.24 es siguiendo a [Priest \(1997\)](#), i.e., entendiendo que la circularidad de las oraciones de Yablo está en el predicado y es este el que está dando lugar a instancias inconsistentes, ya no del [Esquema-T](#), porque este no permite la aplicación de T a predicados, sino del [Esquema-T Uniforme](#). Esto último parece bastante acertado. De hecho, es posible extender las nociones de referencia aletica y sus familiares a fórmulas y ver que, por ejemplo, predicados como el de Yablo resultan autorreferenciales. Con estos conceptos extendidos tal vez podría

obtenerse nuevas teorías desentremilladoras consistentes de la verdad; quizás el objeto de un trabajo futuro.

Sin embargo, por otro lado, habiendo mostrado que teorías dadas por principios uniformes como w_{fub} , f_{ub} y e_{ub} son \neg -consistentes, donde no se ha bloqueado la posible autorreferencialidad de ningún predicado sino solo de oraciones, la objeción de Priest pierde fuerza. Lo que en estas teorías se ha prohibido, a diferencia de s_{ub} , son ciertas oraciones mal fundadas y, por tanto, podría concluirse que esta es una nota común a las paradojas, en lugar de la autorreferencialidad de oraciones y predicados. Quizás la autorreferencia de predicados se traduzca en mala fundación de oraciones, como parece suceder en la secuencia de Yablo y otras semejantes. Nuevamente, dejo esto para una futura investigación. En cualquier caso, queda claro que todas las expresiones que dan lugar a paradojas semánticas son, no solo mal fundadas sino también, o bien autorreferentes, o bien refieren directamente a un número infinito de otras expresiones.

§

Seguramente sea posible encontrar criterios más refinados, que permitan una caracterización más detallada de las expresiones que dan lugar a paradojas semánticas y otras patologías. No pretendo haber agotado los usos de los conceptos de referencia alética dados para formular teorías de la verdad desentremilladoras, minimalistas y de acionistas en este capítulo, sino más bien haber ilustrado como esto es posible y en gran medida exitoso. Idealmente, este capítulo fomenta la formulación de nuevas y mejores teorías de acionistas y minimalistas de la verdad basadas en el concepto de referencia alética.

Conclusion

A lo largo de esta investigación, primeramente, he argumentado que el predicado veri-tativo cumple unicamente roles logico-expresivos en el lenguaje, todos los cuales pueden explicarse por medio de su funcion basica o naturaleza (des)entrecomilladora, i.e., el hecho de que una oracion A y su predicacion de verdad "A es verdadera" sean equivalentes, y que la segunda tenga a A por su objeto. Luego, he inferido que esta naturaleza desentrecomilladora constituye o encapsula el significado de la verdad. Ademas, he defendido la idea de que este significado puede expresarse por medio de principios de transparencia, i.e., el [Esquema T](#), las inferencias de [Introduccion T](#), [Eliminacion T](#) o el principio de Intersustitutividad. Estas tres tesis, sostuve, conforman el nucleo del de acionismo.

Asimismo argument que el de acionismo y cualquiera que crea que la verdad cumple los roles logico-expresivos mencionados debe tener especial interes en hallar quasi logicas o teor as formales de la verdad que garanticen estos usos. He mostrado que esto se traduce lograr que el predicado de verdad de la teor a satisfaga lo que llame 'propiedad de eliminacion' de la verdad.

Creo haber logrado establecer tambien que, en el plano formal, el [Esquema-T](#) u otros principios de transparencia no son ni necesarios ni suficientes para otorgar al predicado de verdad la propiedad de eliminacion. Por un lado, sin ciertos principios logicos clasicos los principios de equivalencia son incapaces de hacerlo. Por otro, dados estos principios logicos, [T-Out](#), la direccion de izquierda a derecha de los bicondicionales del [Esquema-T](#), es no solo necesario sino tambien suficiente. Como en contextos clasicos [T-Out](#) no engendra paradojas ceteris paribus, he concluido que querer contar con un predicado que permita los usos logicos de la verdad no es razon suficiente para abandonar la logica clasica, porque esta es capaz de otorgar a la verdad las propiedades necesarias para llevar a cabo sus roles expresivos, i.e., la propiedad de eliminacion, sin problemas.

Adicionalmente, defend la idea segun la cual una teor a de la verdad dise~nada para permitir los usos logicos de la verdad debe no solo satisfacer la propiedad de eliminacion sino tambien contar al menos con una axiomatizacion parcial razonable que

garantice esta propiedad, para poder ser efectivamente utilizada como una logica, aplicada al razonamiento. Ademas, sostuve que, si bien puede resultar deseable contar con mas principios para el predicado veritativo que las instancias de [T-Out](#) por cuestiones practicas o descriptivas, para cuadrar con las tesis centrales del de acionismo los principios de verdad que la teoria valide no deben exceder los principios de transparencia porque, de lo contrario, la tesis segun la cual el estos principios capturan el significado de la verdad ser a falsa.

En este sentido, argumente que el programa minimalista de Horwich entendido como la busqueda de una teoria axiomatica dada por instancias consistentes del [Esquema-T](#) que utilice un criterio de restriccion abarcativo, sintactico y explicativo es un proyecto digno de exploracion, si bien hasta el momento no ha habido muchos avances en esta direccion. Propuse como un camino posible a explorar aquel basado en la tesis referencialista sobre las paradojas semanticas, de acuerdo con la cual las expresiones paradójicas, que han de ser excluidas del [Esquema-T](#), pueden identificarse a traves de los patrones referenciales que la subyacen. Esta tesis, si bien considerada de 'sentido comun', jamas ha sido evaluada correctamente debido a la falta de nociones adecuadas y precisas de referencia capaces de dar cuenta de los patrones referenciales detras de las expresiones del lenguajes como L_T , como mostre en detalle.

He argumentado que una nocion materialmente adecuada de referencia aplicable a lenguajes formales no debe ser extensional sino hiperintensional, esto es, no debe ser el caso de que dos oraciones logicamente equivalentes referan necesariamente a los mismos objetos, y he dado condiciones generales que cualquier nocion de referencia por el estilo debe satisfacer, dando argumentos cuando fueron necesarios.

Luego, he mostrado que es posible dar una nocion materialmente adecuada de referencia para ciertos lenguajes formales de primer orden, i.e., el lenguaje de la aritmetica, as como de referencia aleatica para el lenguaje de la verdad, que superen las dificultades que otros abordajes previos han encontrado, satisfaciendo las condiciones mencionadas en el parrafo anterior. Las nociones que di son intuitivamente correctas y explican, entre otras cosas, como los lemas diagonales debil y fuerte son capaces de dar lugar a oraciones autorreferenciales siendo hiperintensionales. He ofrecido nociones de referencia tanto semanticas como sintacticas. Para las primeras demostré un resultado de independencia, mientras que para las segundas, al contrario, mostre que son dependientes en el lenguaje de la aritmetica.

Debido a la paradoja de Yablo y otras antinomias infinitarias, empleando las nociones de referencia aleatica que introduce, refutó la tesis referencialista ortodoxa sobre las paradojas semanticas: no todas las paradojas semanticas son

autorreferentes; el patron de referencia comun a todas las paradojas no es circular.

Mostre tambien que es posible recurrir a la nocion de referencia aletica sintactica para formular criterios de restriccion de instancias del [Esquema-T](#). Mostre primero que, si el criterio incluye, grosso modo, las oraciones equivalentes a expresiones aleticamente bien fundadas, el resultado es una teoria desentrecorolladora de la verdad ω -consistente y deductivamente poderosa, lo cual constituye de algun modo una verificacion de la vision referencialista general sobre las paradojas semanticas: el poder deductivo de la teoria muestra que el criterio de seleccion es poco restrictivo y que, por ende, la mala fundacion esta plausiblemente detras de las paradojas.

En segundo lugar, señale como el criterio puede refinarse al menos en dos direcciones, caracterizando los patrones referenciales que subyacen a las expresiones paradójicas con más precisión. El criterio de selección de instancias del [Esquema-T](#) que incluye todas las oraciones que son o bien bien fundadas o bien no autorreferenciales y refieren directamente solo a un número finito de otras expresiones da lugar a una teoría desentrecorolladora que también es ω -consistente y deductivamente poderosa, al igual que aquel que incluye las oraciones que son equivalentes a expresiones que son o bien bien fundadas o bien no refieren a expresiones autorreferentes ni directamente a un número infinito de enunciados.

Finalmente, he mostrado que empleando la nocion de referencia aletica sintactica es posible construir teorías de la verdad que cuadran con el proyecto minimalista de Horwich. Los tres criterios de selección mencionados en los dos párrafos anteriores no solo son explicativos y abarcativos sino que, al estar formulados en términos de la referencia aletica sintactica, una nocion de nible en el lenguaje de la aritmetica de primer orden, son sintacticos y lo suficientemente simples como para dar lugar a una teoria parcialmente axiomatizable. Además, las teorías resultantes no implican principios que exceden el [Esquema-T](#) y, por último pero no menos importante, son todas compatibles con [T-Out](#). En consecuencia, sus extensiones mediante este principio tienen la propiedad de eliminacion y son, por ende, buenas candidatas para el decidacionismo.

Bibliografía

- Asenjo, Florencio G. A Calculus of Antinomies. *Notre Dame Journal of Formal Logic*, 7:103{105, 1966.
- Austin, John L. Truth. *Proceedings of the Aristotelian Society*, 24:111{172, 1950.
- Barrio, Eduardo A. *La Verdad Desestructurada*. Eudeba, Buenos Aires, 1998.
- Barrio, Eduardo A. Theories of truth without standard models and Yablo's sequences. *Studia Logica*, 96:375{391, 2010.
- Barrio, Eduardo A. Definiciones tarskianas de la verdad. En Barrio, Eduardo A. (Ed.), *La Logica de la Verdad*, cap. 1, pp. 25{73. Eudeba, Buenos Aires, 2014a.
- Barrio, Eduardo A. (Ed.). *La Logica de la Verdad*. Eudeba, Buenos Aires, 2014b.
- Barrio, Eduardo A. (Ed.). *Paradojas, Paradojas y mas Paradojas*. College Publications, Londres, 2014c.
- Barrio, Eduardo A. & Picollo, Lavinia. Notes on ω -inconsistent Theories of Truth in Second-Order Languages. *Review of Symbolic Logic*, 6:733{741, 2013.
- Barrio, Eduardo A., Rosenblatt, Lucas & Tajer, Diego. The Logics of Strict-Tolerant Logic. *Journal of Philosophical Logic*, pp. 1{21, forthcoming.
- Beall, J.C. Is Yablo's Paradox non-circular? *Analysis*, 61:176{187, 2001.
- Beall, J.C. Transparent disquotationalism. En Beall, J.C. & Armour-Garb, Bradley (Eds.), *Deflationism and Paradox*, cap. 1, pp. 7{22. Oxford University Press, Oxford, 2005.
- Beall, J.C. *Spandrels of Truth*. Oxford University Press, Oxford, 2009.
- Beall, J.C. A simple approach towards recapturing consistent theories in paraconsistent settings. *Review of Symbolic Logic*, 6:755{764, 2013.
- Beall, J.C. & Armour-Garb, Bradley (Eds.). *Deflationism and Paradox*. Oxford University Press, Oxford, 2005.

- Beall, J.C. & Armour-Garb, Bradley (Eds.). *Revenge of the liar*. Oxford University Press, Oxford, 2007.
- Bell, John L. In *Intuitionistic Logic*. En Edward N. Zalta, (Ed.), *The Stanford Encyclopedia of Philosophy*. Primavera edición, 2012.
- Belnap, Nuel D. Tonk, Plonk, and Plink. *Analysis*, 22:130{134, 1962.
- Belnap, Nuel D. Gupta's rule of revision theory of truth. *Journal of Philosophical Logic*, 11:110{116, 1982.
- Benacerraf, Paul & Putnam, Hilary (Eds.). *Philosophy of Mathematics*. Cambridge University Press, Cambridge, segunda edición, 1983.
- Beringer, Timo & Schindler, Thomas. *Reference-graphs, games for truth, and paradox*. manuscrito, 2014.
- Blackburn, S. & Simmons, K. (Eds.). *Truth*. Oxford University Press, Oxford, 1999.
- Bolander, Thomas, Hendricks, Vincent F. & Pedersen, Stig A. (Eds.). *Self-Reference*. CSLI Publications, Stanford, 2004.
- Boolos, George. To Be Is To Be a Value of a Variable (or to Be Some Values of Some Variables). *Journal of Philosophy*, 81:430{450, 1984.
- Boolos, George. Nominalist Platonism. *Philosophical Review*, 94:327{344, 1985.
- Boolos, George, Burgess, John P. & Jeffrey, Richard C. *Computability and Logic*. Cambridge University Press, Cambridge, quinta edición, 2007.
- Bringsjord, Selmer & van Heuveln, Bram. The 'mental eye' defence of an intuitionized version of Yablo's Paradox. *Analysis*, 63:61{70, 2003.
- Cantini, Andrea. Notes on formal theories of truth. *Zeitschr. f. math. Logik und Grundlagen d. Math.*, 35:97{130, 1989.
- Cantini, Andrea. *Logical frameworks for truth and abstraction. An axiomatic study*, volumen 135 de *Studies in Logic and the Foundations of Mathematics*. Elsevier, Amsterdam, 1996.
- Carnap, Rudolf. *Logical Syntax of Language*. Routledge, Londres, 1937.
- Carnielli, Walter, Coniglio, Marcelo & Marcos, Joao. *Logics of Formal Inconsistency*. En Gabbay, Dov M. & Guenther, Franz (Eds.), *Handbook of Philosophical Logic*, volumen 14, pp. 1{93. Springer, Nueva York, segunda edición, 2007.

- Chang, Chen Chung & Keisler, H. Jerome. *Model Theory*. North Holland, Amsterdam, tercera edición, 1990.
- Cieslinski, Cezary. De ationism, conservativeness and maximality. *Journal of Philosophical Logic*, 36:695{705, 2007.
- Cintula, Petr, Fermüller, Christian, Godo, Lluís & Hajek, Petr (Eds.). *Reasoning Under Vagueness: Logical, Philosophical, and Linguistic Perspectives*. College Publications, Londres, 2011.
- Cobreros, Pablo, Egge, Paul, Ripley, David & Rooij, Robert van . Reaching Transparent Truth. *Mind*, 122:841{866, 2013.
- Cook, Roy T. Patterns of paradox. *Journal of Symbolic Logic*, 69(3):767{774, 2004.
- Cook, Roy T. What's wrong with tonk(?). *Journal of Philosophical Logic*, 34:217{226, 2005.
- Cook, Roy T. There are non-circular paradoxes (but Yablo's Isn't One of Them!). *The Monist*, 89:118{149, 2006.
- Cook, Roy T. *The Yablo Paradox. An Essay on Circularity*. Oxford University Press, Oxford, 2014a.
- Cook, Roy T. There is no paradox of logical validity! *Logica Universalis*, 8:447{467, 2014b.
- Coquand, Thierry. Type Theory. En Edward N. Zalta, (Ed.), *The Stanford Encyclopedia of Philosophy*. Verano edición, 2015.
- Craig, William & Vaught, Robert L. Finite axiomatizability using additional predicates. *Journal of Symbolic Logic*, 23:289{308, 1958.
- Curry, Haskell B. The inconsistency of certain formal logics. *Journal of Symbolic Logic*, 7:115{117, 1942.
- da Costa, Newton. On the Theory of Inconsistent Formal Systems. *Notre Dame Journal of Formal Logic*, 15:497{510, 1974.
- Davidson, Donald. Truth and Meaning. *Synthese*, 17:304{323, 1967.
- Davidson, Donald. True to the Facts. *Journal of Philosophy*, 66:748{764, 1969.
- Davidson, Donald. The Structure and Content of Truth. *Journal of Philosophy*, 87: 279{328, 1990.

- Dowden, Bradley. Accepting inconsistencies from the paradoxes. *Journal of Philosophical Logic*, 13:125{130, 1984.
- Dummett, Michael. Truth. *Proceedings of the Aristotelian Society*, 59:141{162, 1958-1959.
- Dummett, Michael. *Truth and Other Enigmas*. Harvard University Press, Cambridge, Mass., 1978.
- Enderton, Herbert B. *Elements of Set Theory*. Academic Press, Londres, 1977.
- Etchemendy, John. Tarski on Truth and Logical Consequence. *Journal of Symbolic Logic*, 53:51{79, 1988.
- Etchemendy, John. *The Concept of Logical Consequence*. Harvard University Press, Cambridge, Mass., 1990.
- Feferman, Solomon. Reflecting on incompleteness. *Journal of Symbolic Logic*, 56:1{49, 1991.
- Feferman, Solomon. Axioms for determinateness and truth. *Review of Symbolic Logic*, 1:204{217, 2008a.
- Feferman, Solomon. Tarski's Conceptual Analysis of Semantical Notions. En Patterson, D. (Ed.), *New Essays on Tarski and Philosophy*, pp. 72{93. Oxford University Press, Nueva York, 2008b.
- Field, Hartry. The Deflationary Conception of Truth. En MacDonald, G. & Wright, C. (Eds.), *Fact, Science and Morality*. Blackwell, Oxford, 1986.
- Field, Hartry. Critical Notice: Paul Horwich's "Truth". *Philosophy of Science*, 59:321{330, 1992.
- Field, Hartry. Deflationist views of meaning and content. *Mind*, 103:249{284, 1994.
- Field, Hartry. Defeating the Conservativeness Argument. *Journal of Philosophy*, 96:533{540, 1999.
- Field, Hartry. Saving the Truth Schema from Paradox. *Journal of Philosophical Logic*, 31:1{27, 2002.
- Field, Hartry. A revenge-immune solution to the semantic paradoxes. *Journal of Philosophical Logic*, 32:139{177, 2003.

Field, Hartry. Variations on a Theme by Yablo. En Beall, J.C. & Armour-Garb, Bradley (Eds.), De ationism and Paradox, pp. 53{74. Oxford University Press, Oxford, 2005.

- Field, Hartry. Solving the paradoxes, escaping revenge. En Beall, J.C. (Ed.), *Revenge of the liar*, pp. 78{144. Oxford University Press, Oxford, 2007.
- Field, Hartry. *Saving Truth from Paradox*. Oxford University Press, Nueva York, 2008.
- Fischer, Martin, Halbach, Volker, Stern, Johannes & Kreiner, J•onne. Axiomatizing semantic theories of truth? *Review of Symbolic Logic*, forthcoming:22, 2015. doi: <http://dx.doi.org/10.1017/S1755020314000379>.
- Forster, Thomas. The signi cance of Yablo's Paradox without self-reference. *Logique et Analyse*, 47:461{462, 2004.
- Frege, Gottlob. *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Louis Nebert, Halle, 1879. Traducido al ingles como 'Concept Script, a formal language of pure thought modelled upon that of arithmetic' en [van Heijenoort \(1967\)](#).
- Frege, Gottlob. *Der Gedanke: Eine logische Untersuchung*. *Beitr•age zur Philosophie des Deutschen Idealismus*, 1:58{77, 1918-19. Traducido al ingles como 'The Thought: A Logical Inquiry' en [Frege \(1956\)](#) y reimpresso en [Blackburn & Simmons \(1999\)](#), pp. 85-105).
- Frege, Gottlob. *The Thought: A Logical Inquiry*. *Mind*, 65:289{311, 1956.
- Friedman, Harvey & Sheard, Michael. An axiomatic approach to self{referential truth. *Annals of Pure and Applied Logic*, 33:1{21, 1987.
- Gabbay, Dov M. & G•unthner, Franz (Eds.). *Handbook of Philosophical Logic*, volumen IV. Reidel, Dordrecht, 1989.
- Gabbay, Dov M. & G•unthner, Franz (Eds.). *Handbook of Philosophical Logic*, volumen 6. Kluwer, Dordrecht, 2002.
- Gabbay, Dov M. & G•unthner, Franz (Eds.). *Handbook of Philosophical Logic*, volumen 14. Springer, Nueva York, 2007.
- Garc a-Carpintero, Manuel. The Grounds for the Model-Theoretic Account of the Logical Properties. *Notre Dame Journal of Formal Logic*, 34:107{131, 1993.
- Gentzen, Gerhard. *Untersuchungen uber• das Logische Schlie•en*. *Math. Zeitschrift*, 39: 176{210, 1934. Traducido al ingles como 'Investigations into Logical Deduction' en ?, pp. 68-131.
- Glanzberg, Michael. Minimalism and Paradoxes. *Synthese*, 135:13{36, 2003.

- Glanzberg, Michael. Minimalism, Deflationism and Paradoxes. En Beall, J.C. & Armour-Garb, Bradley (Eds.), *Deflationism and Paradox*, cap. 8, pp. 107-132. Oxford University Press, 2005.
- Gödel, Kurt. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, 38:173-198, 1931. Traducido al inglés como 'On formally undecidable propositions of Principia Mathematica and related systems I' en ?, pp. 145-195.
- Gödel, Kurt. The present situation in the foundations of mathematics. En ?, pp. 45-53, 1933. Manuscrito para una conferencia de la Asociación Matemática de América en Cambridge, Mass., 29 y 30 de diciembre.
- Gödel, Kurt. Russell's mathematical logic. En Schilpp, Paul A. (Ed.), *The Philosophy of Bertrand Russell*, volumen III de *The Library of Living Philosophers*, pp. 125-153. Evanston & Chicago, Northwestern University, 1944. Reimpreso en [Benacerraf & Putnam \(1983\)](#), pp. 447-469.
- Gomez-Torrente, Mario. Logical Truth and Tarskian Logical Truth. *Synthese*, 117: 375-408, 1998/1999.
- Gomez-Torrente, Mario. The indefinability of truth in the 'Wahrheitsbegriff'. *Annals of Pure and Applied Logic*, 126:27-37, 2004.
- Goodman, Nelson. *About*. *Mind*, 70:1-24, 1961.
- Grover, Dorothy L., Camp, Joseph L. & Belnap, Nuel D. A prosentential Theory of Truth. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 27:73-125, 1975.
- Gupta, Anil. Truth and Paradox. *Journal of Philosophical Logic*, 11:1-60, 1982.
- Gupta, Anil. A critique of deflationism. *Philosophical Topics*, 21:57-81, 1993.
- Gupta, Anil & Belnap, Nuel D. *The Revision Theory of Truth*. MIT Press, Cambridge, Mass., 1993.
- Halbach, Volker. A System of Complete and Consistent Truth. *Notre Dame Journal of Formal Logic*, 35:311-327, 1994.
- Halbach, Volker. Tarski Hierarchies. *Erkenntnis*, 43:339-367, 1995.
- Halbach, Volker. Disquotationalism and infinite conjunctions. *Mind*, 108:1-22, 1999.
- Halbach, Volker. How Innocent is Deflationism? *Synthese*, 126:167-194, 2001.

- Halbach, Volker. *Understanding Shapiro's Guru*. 2003.
- Halbach, Volker. Reducing compositional to disquotational truth. *Review of Symbolic Logic*, 2:786{798, 2009.
- Halbach, Volker. *Axiomatic Theories of Truth*. Cambridge University Press, Cambridge, 2011.
- Halbach, Volker & Horsten, Leon. The De ationist's Axioms for Truth. En Beall, J.C. & Armour-Garb, Bradley (Eds.), *De ationism and Paradox*, cap. 12, pp. 203{217. Oxford University Press, 2005.
- Halbach, Volker & Horsten, Leon. Axiomatizing Kripke's Theory of Truth. *Journal of Symbolic Logic*, 71:677{712, 2006.
- Halbach, Volker & Visser, Albert. Self-reference in Arithmetic I. *Review of Symbolic Logic*, 7:671{691, 2014a.
- Halbach, Volker & Visser, Albert. Self-reference in Arithmetic II. *Review of Symbolic Logic*, 7:692{712, 2014b.
- Hardy, James. Is Yablo's Paradox Liar-Like? *Analysis*, 55:197{198, 1995.
- Heck, Richard Jr. Se -reference and the Languages of Arithmetic. *Philosophia Mathematica*, III:1{29, 2007.
- Henkin, Leon. A problem concerning provability. *Journal of Symbolic Logic*, 17:160, 1952.
- Herzberger, Hans G. Paradoxes of grounding in semantics. *Journal of Philosophical Logic*, 67:145{167, 1970.
- Herzberger, Hans G. Notes on Naive Semantics. *Journal of Philosophical Logic*, 11: 61{102, 1982a.
- Herzberger, Hans G. Naive Semantics and the Liar Paradox. *Journal of Philosophy*, 79: 479{497, 1982b.
- Horsten, Leon. The semantical paradoxes, the neutrality of truth, and the neutrality of the minimalist theory of truth. En Cortois, P. (Ed.), *The Many Problems of Realism*, pp. 173{187. Tilburg University Press, Tilburgo, 1995.
- Horsten, Leon. *The Tarskian Turn: De ationism and Axiomatic Truth*. MIT Press, Cambridge, Mass., 2011.
- Horwich, Paul. De ationary Truth and the Problem of Aboutness. *Philosophical Issues*, 8, Truth:95{106, 1997.

- Horwich, Paul. *Meaning*. Clarendon Press, Oxford, 1998a.
- Horwich, Paul. *Truth*. Blackwell, Oxford, segunda edición, 1998b.
- Horwich, Paul. A minimalist critique of Tarski on truth. En Beall, J.C. & Armour-Garb, Bradley (Eds.), *Deflationism and Paradox*, pp. 75{84. Oxford University Press, Oxford, 2005.
- Jespersen, Bjørn & Düzi, Marie. Introduction to a special issue of *Synthese* on Hyperintensionality. *Synthese*, en prensa.
- Ketland, Jeffrey. Deflationism and Tarski's paradise. *Mind*, 108:69{94, 1999.
- Ketland, Jeffrey. Bueno and Colyvan on Yablo's Paradox. *Analysis*, 64:165{172, 2004.
- Ketland, Jeffrey. Yablo's Paradox and ω -inconsistency. *Synthese*, 145:295{307, 2005.
- Kleene, Stephen C. Finite axiomatizability of theories in the predicate calculus using additional predicate symbols. *Memoirs of the American Mathematical Society*, 10: 27{68, 1952.
- Kreisel, Georg. On a problem of Henkin's. *Indagationes Mathematicae*, 15:405{406, 1953.
- Kremer, Michael. Kripke and the Logic of Truth. *Journal of Philosophical Logic*, 17: 225{278, 1988.
- Kripke, Saul. *Outline of a Theory of Truth*. *Journal of Philosophy*, 72:690{716, 1975.
- Kripke, Saul. *Naming and Necessity*. Harvard University Press, Cambridge, Mass., 1980.
- Leeds, Stephen. Theories of reference and truth. *Erkenntnis*, 13:111{129, 1978.
- Leitgeb, Hannes. What is a Self-Referential Sentence? Critical remarks on the alleged (non)-circularity of Yablo's Paradox. *Logique et Analyse*, 177-178:3{14, 2002.
- Leitgeb, Hannes. What Truth Depends On. *Journal of Philosophical Logic*, 34:155{192, 2005.
- Leitgeb, Hannes. What Theories of Truth Should be Like (but Cannot be). *Philosophy Compass*, 2:276{290, 2007.
- Levy, Azriel. *Basic Set Theory*. Springer-Verlag, Berlin, 1979. Reimpreso por Dover Publications, 2003.
- Lewis, David. *Papers in Philosophical Logic*. Cambridge University Press, Cambridge, 1988a.

- Lewis, David. Statements partly about observation. En *Papers in Philosophical Logic*, pp. 125{155. Cambridge University Press, Cambridge, 1988b.
- Linnebo, ystein & Rayo, Agust n. Hierarchies Ontological and Ideological. *Mind*, 121: 269{308, 2012.
- L•ob, Martin H. Solution of a Problem of Leon Henkin. *Journal of Symbolic Logic*, 20: 115{118, 1955.
- Martin, Robert L. & Woodru , Peter W. On Representing True-in-L in L. *Philosophia*, 5:213{217, 1975.
- Mart nez, Jose. La Paradoja del Mentiroso. En Barrio, Eduardo A. (Ed.), *Paradojas, Paradojas y mas Paradojas*, pp. 11{26. College Publications, Londres, 2014.
- McGee, Vann. How Truth-Like can a Predicate Be? A Negative Result. *Journal of Philosophical Logic*, 14:399{410, 1985.
- McGee, Vann. *Truth, Vagueness, and Paradox. An Essay on the Logic of Truth.* Hackett, Indiana, 1991.
- McGee, Vann. Maximal consistent sets of instances of Tarski's schema. *Journal of Philosophical Logic*, 21:235{241, 1992a.
- McGee, Vann. Two Problems with Tarski's Theory of Consequence. *Proceedings of the Aristotelian Society*, 92:273{292, 1992b.
- Montague, Richard. Syntactical Treatments of Modality, with Corollaries on Re ection Principles and Finite Axiomatizability. *Acta Philosophica Fennica*, 16:153{167, 1963. Reimpreso en [Montague \(1974\)](#), pp. 286-302).
- Montague, Richard. *Formal Philosophy. Selected Papers of Richard Montague.* Yale University Press, New Haven, 1974.
- Moretti, Alberto & Hurtado, Guillermo (Eds.). *La Paradoja de Orayen.* Eudeba, Buenos Aires, 2003.
- Moschovakis, Yiannis. *Notes on Set Theory.* Springer, Nueva York, segunda edicion, 2006.
- Nelson, Michael. Existence. En Edward N. Zalta, (Ed.), *The Stanford Encyclopedia of Philosophy.* Invierno edicion, 2012.
- Orayen, Raul. Una paradoja en la semantica de la teoria de conjuntos. En Moretti, Alberto & Hurtado, Guillermo (Eds.), *La Paradoja de Orayen*, pp. 35{59. Eudeba, Buenos Aires, 2003.

- Pailos, Federico. La Paradoja de Curry. En Barrio, Eduardo A. (Ed.), *Paradojas, Paradojas y mas Paradojas*, pp. 41{52. College Publications, Londres, 2014.
- Pailos, Federico & Rosenblatt, Lucas. Non-deterministic Conditionals and Transparent Truth. *Studia Logica*, 2014. doi: 10.1007/s11225-014-9580-1.
- Paoli, Francesco. *Substructural Logics: A Primer*. Kluwer, Dordrecht, 2002.
- Patterson, Douglas (Ed.). *New Essays on Tarski and Philosophy*. Oxford University Press, Nueva York, 2008.
- Picollo, Lavinia. La Inocencia del De acionismo. *Manuscrito*, 33:425{443, 2010.
- Picollo, Lavinia. The Old-Fashioned Yablo Paradox. *Analisis Filosófico*, 32:21{29, 2012.
- Picollo, Lavinia. Yablo's Paradox in Second-Order Languages: Consistency and Unsatisfiability. *Studia Logica*, 101:601{617, 2013.
- Picollo, Lavinia & Buacar, Natalia. La Teoría Revisionista de la Verdad. En Barrio, Eduardo A. (Ed.), *La Lógica de la Verdad*, cap. 3, pp. 117{186. Eudeba, Buenos Aires, 2014.
- Picollo, Lavinia & Schindler, Thomas. Disquotation and Infinite Conjunctions. 2015. En preparación, 26 páginas.
- Pohlers, Wolfram. *Proof Theory: The First Step into Impredicativity*. Springer, Berlin/Heidelberg, 2009.
- Poincaré, Henri. Les Mathématiques et la logique. *Revue de Métaphysique et de Morale*, 14:294{317, 1906.
- Priest, Graham. The logic of paradox. *Journal of Philosophical Logic*, 9:219{241, 1979.
- Priest, Graham. In *Contradiction*. Martinus Nijhoff, Dordrecht, primera edición, 1987.
- Priest, Graham. Yablo's Paradox. *Analysis*, 57:236{242, 1997.
- Priest, Graham. Paraconsistent logics. En Gabbay, Dov M. & Gunthner, Franz (Eds.), *Handbook of Philosophical Logic*, volumen 6, pp. 287{393. Kluwer, Dordrecht, 2002.
- Priest, Graham. *Doubt Truth To Be A Liar*. Oxford University Press, Oxford, 2006a.
- Priest, Graham. In *Contradiction*. Oxford University Press, Nueva York, segunda edición, 2006b.

- Priest, Graham. Paraconsistency and Dialetheism. En Gabbay, Dov M. & Woods, John (Eds.), *Handbook of the History of Logic*, volumen 8: The Nonmonotonic Turn in Logic, pp. 129{204. North Holland, Amsterdam, 2007.
- Priest, Graham. *An Introduction to Non-Classical Logic*. Cambridge University Press, Cambridge, segunda edicion, 2008.
- Prior, Arthur N. Runabout Inference Ticket. *Analysis*, 21:38{39, 1960/61.
- Prior, Arthur N. Conjunction and Contonktion Revisited. *Analysis*, 24:191{195, 1964.
- Prior, Arthur N. *Objects of Thought*. Clarendon Press, Oxford, 1971.
- Putnam, Hilary. Formalization of the Concept of "About". *Philosophy of Science*, 25: 125{130, 1958.
- Putnam, Hilary. A Comparison of Something with Something Else. *New Literary History*, 17(1, Philosophy of Science and Literary Theory), 1985.
- Putnam, Hilary. *Representation and Reality*. MIT Press, Cambridge, Mass., 1988.
- Quine, Willard Van Orman. *Philosophy of Logic*. Harvard University Press, Cambridge, Mass., 1970.
- Quine, Willard Van Orman. *Pursuit of truth*. Harvard University Press, Cambridge, Mass., revisada edicion, 1992.
- Rabern, Landon, Rabern, Brian & Macauley, Matthew. Dangerous reference graphs and semantic paradoxes. *Journal of Philosophical Logic*, 42(5):727{765, 2013.
- Ramsey, Frank P. Facts and Propositions I. *Proceedings of the Aristotelian Society*, 7: 153{170, 1927.
- Rathjen, Michael. The realm of ordinal analysis. En Cooper, S. Barry & Truss, John K. (Eds.), *Sets and Proofs*, volumen 258 de London Mathematical Society Lecture Note Series, pp. 219{278. Cambridge University Press, Cambridge, 1999.
- Rayo, Agust n & Uzquiano, Gabriel. Toward a Theory of Second-Order Consequence. *Notre Dame Journal of Formal Logic*, 40:315{325, 1999.
- Rayo, Agust n & Uzquiano, Gabriel. Introduction. En Rayo, Agust n & Uzquiano, Gabriel (Eds.), *Absolute Generality*, pp. 1{19. Oxford University Press, Oxford, 2006a.
- Rayo, Agust n & Uzquiano, Gabriel (Eds.). *Absolute Generality*. Oxford University Press, Oxford, 2006b.

- Reinhardt, William. Some remarks on extending and interpreting theories with a partial predicate for truth. *Journal of Philosophical Logic*, 15:219{251, 1986.
- Resnik, Michael. *Mathematics as a Science of Patterns*. Clarendon Press, Oxford, 1997.
- Restall, Greg. *An Introduction to Substructural Logics*. Routledge, Londres, 2000.
- Ripley, David. Paradoxes and Failures of Cut. *Australasian Journal of Philosophy*, 91: 139{164, 2012a.
- Ripley, David. Conservatively extending classical logic with transparent truth. *Review of Symbolic Logic*, 5:354{378, 2012b.
- Rosenblatt, Lucas & Pailos, Federico. Paracompletitud So sticada. En Barrio, Eduardo A. (Ed.), *La Logica de la Verdad*, cap. 4. Eudeba, 2014.
- Russell, Bertrand. *The Principles of Mathematics*. Cambridge University Press, Cambridge, 1903.
- Russell, Bertrand. *Mathematical Logic as Based on the Theory of Types*. *American Journal of Mathematics*, 30:222{262, 1908.
- Russell, Bertrand & Whitehead, Alfred N. *Principia Mathematica*. Cambridge University Press, Cambridge, 1910.
- Ryle, Gilbert. 'About'. *Analysis*, 1:10{12, 1933a.
- Ryle, Gilbert. *Imaginary Objects*. *Proceedings of the Aristotelian Society*, volumen suplementario 12:18{43, 1933b.
- Schindler, Thomas. Axioms for grounded truth. *Review of Symbolic Logic*, 7:73{83, 2014a.
- Schindler, Thomas. A disquotational theory of truth as strong as Z_2 . *Journal of Philosophical Logic*, p. 18, 2014b. doi: 10.1007/s10992-014-9327-5.
- Shapiro, Stewart. *Foundations without Foundationalism: A Case for Second-Order Logic*. Oxford University Press, Nueva York, 1991.
- Shapiro, Stewart. *Philosophy of Mathematics: Structure and Ontology*. Oxford University Press, Nueva York, 1997.
- Shapiro, Stewart. Proof and Truth: Through Thick and Thin. *Journal of Philosophy*, 95:493{521, 1998.
- Shapiro, Stewart. The Guru, the Logician and the De ationist: Truth and Logical Consequence. *Nous*, 37:113{132, 2003.

- Simpson, Stephen G. *Subsystems of Second Order Arithmetic*. Cambridge University Press, Cambridge, segunda edición, 2009.
- Smith, Peter. *An Introduction to Gödel's Theorems*. Cambridge University Press, Cambridge, 2007.
- Smorinsky, Craig. *Self-Reference and Modal Logic*. Springer-Verlag, Nueva York, 1985.
- Soames, Scott. *Understanding Truth*. Oxford University Press, Nueva York/Oxford, 1999.
- Sorensen, Roy A. Yablo's Paradox and kindred in nite liars. *Mind*, 107:137{155, 1998.
- Strawson, Peter F. Truth. *Analysis*, 9:83{97, 1949.
- Tajer, Diego. Dialete smo: una teor a contradictoria de la verdad. En Barrio, Eduardo A. (Ed.), *La Logica de la Verdad*, cap. 5, pp. 249{292. Eudeba, Buenos Aires, 2014.
- Tarski, Alfred. Prace Towarzystwa Naukowego Warszawskiego. Wydział III Nauk Matematyczno-Fizycznych, 34, 1933. Traducido al ingles como 'The concept of truth in the languages of the deductive sciences' y expandido en ?, pp. 152-278.
- Tarski, Alfred. The Semantic Conception of Truth: and the Foundations of Semantics. *Philosophy and Phenomenological Research*, 4:341{376, 1944.
- Tarski, Alfred. Some notions and methods on the borderline of algebra and metamathematics. En *Proceedings of the International Congress of Mathematicians*, Cambridge, Mass. 1950, volumen 1, pp. 705{720, Providence, 1952. American Mathematical Society.
- Tarski, Alfred & Vaught, Robert L. Arithmetical extensions of relational systems. *Compositio Mathematica*, 13:81{102, 1956.
- Tarski, Alfred, Mostowski, Andrzej & Robinson, Raphael M. *Undecidable theories*. North Holland, Amsterdam, 1953.
- Teijeiro, Paula & Szmuc, Damian. Teor a de Puntos Fijos de Kripke. En Barrio, Eduardo A. (Ed.), *La Logica de la Verdad*, cap. 2, pp. 73{114. Eudeba, 2014.
- Urbaniak, Rafal. Leitgeb, 'About,' Yablo. *Logique et Analyse*, 207:239{254, 2009.
- Uzquiano, Gabriel. Plural Quantification and Classes. *Philosophia Mathematica*, 11: 67{81, 2003.
- van Heijenoort, Jean (Ed.). *From Frege to Gödel: A Source Book in Mathematical Logic, 1879{1931*. Harvard University Press, Cambridge, Mass., 1967.

- van Rooij, Robert. Vagueness, tolerance and non-transitive entailment. En Cintula, Petr, Fermüller, Christian, Godo, Lluís & Hajek, Petr (Eds.), *Reasoning Under Vagueness: Logical, Philosophical, and Linguistic Perspectives*, pp. 205{223. College Publications, Londres, 2011.
- Visser, Albert. Semantics and the liar paradox. En Gabbay, Dov M. & Günthner, Franz (Eds.), *Handbook of Philosophical Logic*, volumen IV, pp. 617{706. Reidel, Dordrecht, 1989.
- Weir, Alan. Naïve truth and sophisticated logic. En Beall, J.C. & Armour-Garb, Bradley (Eds.), *Deflationism and Paradox*, pp. 218{249. Oxford University Press, Oxford, 2005.
- Welch, Philip. *The Complexity of the Dependence Operator*, 2014.
- Williams, Michael. Epistemological Realism and the Basis of Skepticism. *Mind*, 97: 415{439, 1988.
- Yablo, Stephen. Grounding, Dependence, and Paradox. *Journal of Philosophical Logic*, 11:117{137, 1982.
- Yablo, Stephen. Truth and reflexion. *Journal of Philosophical Logic*, 14:297{349, 1985.
- Yablo, Stephen. Paradox without Self-Reference. *Analysis*, 53:251{252, 1993.
- Yablo, Stephen. Circularity and Paradox. En Bolander, Thomas, Hendricks, Vincent F. & Pedersen, Stig A. (Eds.), *Self-Reference*, pp. 139{157. CSLI Publications, Stanford, 2004.
- Yablo, Stephen. *Aboutness*. Princeton University Press, Princeton, 2014.
- Zardini, Elia. Truth Without Contra(d)iction. *Review of Symbolic Logic*, 4:498{535, 2011.

Indice alfabetico

adscripciones de

verdad ciegas, [51](#),
[107](#) expl citas, [98](#),
[102](#) generales, [98](#)
impl citas, [99](#)
singulares, [98](#), [102](#)

aritmetica

de Peano, vease pa, pa⁺, [3](#)
de Robinson, vease q, q⁺, [3](#)

Asenjo, Florencio G., [81](#),
[82](#) asignaciones, [30](#), [31](#),
[36](#) autodependencia, [74](#)

autorreferencia

aletica

semantica, [232](#)
sintactica, [245](#)

de oraciones, [224](#)

Ayer, Alfred J., [96](#)

Barrio, Eduardo A., [38](#), [65](#), [79](#), [91](#), [119](#), [123](#),
[124](#), [183](#), [276](#)

Beall, J.C., [62](#), [81](#), [83](#){[85](#), [89](#), [100](#), [140](#), [182](#),
[184](#)

Bell, John L., [141](#)

Belnap, Nuel D., [63](#), [97](#), [99](#), [106](#), [114](#), [115](#),
[127](#), [128](#), [133](#), [134](#), [276](#)

Beringer, Timo, [188](#)

Boolos, George, [4](#), [5](#), [27](#), [42](#), [44](#), [93](#)

Brandom, Robert, [100](#)

Bringsjord, Selmer,
[182](#) b3, [56](#), [57](#)

Buekens, Filip, [103](#)

buena fundacion

aletica

semantica, [232](#)
sintactica, [245](#)

de oraciones, [224](#)

Burgess, John P., [4](#), [5](#), [27](#), [93](#)

c rculo de Vienna,

[18](#) cadena

de Henkin, [219](#)
de honestos, [233](#)

de referencia directa
a oraciones, [223](#)

Camp, Joseph L., [97](#), [99](#), [106](#)

Cantini, Andrea, [65](#){[67](#)

cardinales de von Neumann, [14](#),
[16](#) Carnap, Rudolf, [189](#)

Carnielli, Walter, [81](#)

Chang, Chen Chung, [35](#),

[45](#) ciclo

de G•odel, [218](#) de
Henkin, [218](#) de
mentirosos, [70](#)

Cieslinski, Cezary, [168](#)

clausura de un modelo parcial,

[64](#) Cobreros, Pablo, [93](#), [140](#) codi
cacion de G•odel, [5](#)

monotona, [5](#)

completitud [1](#), [9](#)

componentes proposicionales de una
formu-la, [242](#)

Compositional Truth, vease ct, [47](#)

condicion de equivalencia, [186](#)

condicional de Priest, [86](#){[88](#)

- condiciones de teoremicidad de Lob, [6](#)
- conservatividad, [127](#), [145](#)
- Convencion T, [25](#)
- Cook, Roy T., [122](#), [128](#), [129](#), [187](#), [188](#), [233](#)
- Coquand, Thierry, [29](#)
- correspondentismo, [19](#), [20](#), [22](#)
- Craig, William, [109](#)
- criterio de Kreisel-Henkin, [191](#), [217](#)
- ct, [46](#){[48](#)
- cuantificacion proposicional, [27](#), [99](#)
- Curry, Haskell B., [85](#)
- d-fundacion, [74](#), [75](#)
- da Costa, Newton, [81](#)
- Davidson, Donald, [21](#), [110](#), [116](#), [123](#)
- de nibilidad, [3](#)
- de niciones tarskianas de verdad
- absoluta, [31](#), [32](#), [34](#)
 - relativa a un modelo, [35](#), [38](#)
- de acionismo, [26](#), [95](#){[97](#), [99](#), [100](#), [105](#), [111](#), [112](#), [115](#){[120](#), [123](#), [125](#){[128](#), [134](#), [136](#){[139](#), [143](#), [163](#), [166](#)
- pf , vease d-fundacion, [74](#)
- dependencia, [71](#), [72](#)
- esencial, [73](#)
- dialectismo, [81](#), [85](#)
- dispositivo desentramador, [101](#), [112](#), [117](#)
- Dowden, Bradley, [84](#)
- Duzi, Marie, [188](#)
- Dummett, Michael, [110](#), [116](#), [123](#)
- ω , [16](#), [17](#)
- Egre, Paul, [93](#), [140](#)
- Eliminacion T, [120](#)
- Enderton, Herbert B., [14](#)
- Equivalent Reference Uniform Tarski Biconditionals, vease erutb, [274](#)
- erutb, [274](#)
- esquema de valuacion
- ep, de Priest, [87](#)
 - sk, Kleene fuerte, [54](#)
 - wk, Kleene debil, [56](#)
- Esquema T, [21](#)
- Tipeado, [25](#)
- Esquema-T, [40](#)
- de Priest, [88](#)
 - Uniforme, [45](#)
- Etchemendy, John, [39](#), [47](#)
- ev-fundacion, [60](#)
- formula decidible, [251](#)
- Feferman, Solomon, [25](#), [34](#), [63](#), [64](#)
- Field, Hartry, [61](#), [62](#), [98](#), [100](#), [113](#), [114](#), [116](#), [118](#), [120](#), [123](#), [124](#), [128](#), [129](#), [140](#), [160](#)
- Fine, Arthur, [100](#)
- Finite Reference Uniform Tarski Biconditionals, vease frutb, [271](#)
- Fischer, Martin, [130](#), [139](#), [165](#), [170](#)
- forma normal de Cantor, [16](#)
- Forster, Thomas, [180](#)
- Frege, Gottlob, [96](#), [126](#)
- Friedman, Harvey, [78](#), [79](#)
- frutb, [270](#), [271](#)
- fs, [78](#), [79](#)
- funcion basica de la verdad, [101](#), [112](#)
- funcion de Veblen, [16](#)
- fundacion, [59](#), [60](#), [71](#)
- ω , [17](#)
- Gödel, Kurt, [9](#), [11](#), [13](#), [14](#), [30](#)
- Gomez -Torrente, Mario, [29](#), [39](#)
- Garc a-Carpintero, Manuel, [39](#)
- generalidad absoluta, [43](#)
- Gentzen, Gerhard, [126](#)
- Glanzberg, Michael, [166](#), [167](#)
- Goodman, Nelson, [189](#), [197](#){[199](#), [201](#)
- Grover, Dorothy L., [97](#), [99](#), [100](#), [106](#)
- Gupta, Anil, [63](#), [97](#), [114](#), [115](#), [117](#), [124](#), [133](#), [134](#), [276](#)

- Halbach, Volker, [5](#), [43](#), [45](#){[47](#), [63](#), [66](#), [68](#), [78](#){[80](#), [114](#), [120](#), [125](#), [127](#){[130](#), [137](#){
[139](#), [142](#){[144](#), [161](#), [165](#), [168](#), [170](#),
[172](#), [187](#), [190](#), [191](#), [255](#), [268](#), [269](#)
- Hardy, James, [182](#)
- Heck, Richard Jr, [5](#)
- Henkin, Leon, [191](#)
- Herzberger, Hans G., [59](#), [63](#), [184](#), [190](#)
- hiperintensionalidad, [188](#)
- Horsten, Leon, [63](#), [66](#), [100](#), [117](#){[120](#), [125](#),
[127](#), [128](#), [137](#){[140](#), [164](#), [172](#)
- Horwich, Paul, [100](#), [101](#), [111](#), [116](#), [118](#), [120](#),
[121](#), [123](#), [125](#), [126](#), [131](#), [132](#), [140](#),
[149](#), [158](#), [160](#), [165](#){[167](#), [169](#), [170](#),
[262](#)
- inferencialismo lógico, [126](#){[128](#)
- interpretaciones relativas, [2](#)
- Introducción T, [120](#)
- IS, [84](#)
- Je rey, Richard C., [4](#), [5](#), [27](#), [93](#)
- jerarquía
analítica, [9](#)
aritmética, [8](#)
- Jespersen, Bjørn, [188](#)
- Keisler, Jerome, [35](#), [45](#)
- Ketland, Je rey, [127](#), [182](#), [183](#), [212](#)
- kf, [63](#), [64](#)
- Kleene, Stephen C., [109](#)
- Kreisel, Georg, [191](#)
- Kremer, Michael, [60](#){[62](#)
- Kriener, Jönnne, [130](#), [139](#), [165](#), [170](#)
- Kripke, Saul, [51](#){[53](#), [55](#), [57](#){[61](#), [68](#), [71](#), [123](#),
[179](#)
- Kripke-Feferman, vease kf, [53](#)
- k3, [55](#)
- L
- Löb, Martin H., [6](#), [191](#)
- de las paradojas, vease lp, [81](#)
- de Priest, vease pl, [86](#)
- externa, [65](#)
- interna, [65](#)
- paracompleta, [52](#), [54](#)
- paraconsistente, [81](#), [92](#)
- subestructural, [90](#)
- Leeds, Stephen, [100](#), [111](#), [120](#), [123](#)
- Leigh, Graham, [137](#)
- Leitgeb, Hannes, [65](#), [67](#), [69](#){[75](#), [78](#), [79](#), [178](#),
[180](#), [183](#){[190](#), [208](#), [209](#), [221](#), [256](#),
[276](#)
- lema
de König, [77](#)
diagonal
cadenas, [12](#)
ciclos, [11](#), [12](#)
débil, [9](#)
fuerte, [10](#)
- lenguaje
de la aritmética, vease L_{pa} , L_{pa}^+ , [3](#)
de la teoría de conjuntos, vease L_{zfc} ,
 L_{zfc}^+ , [14](#)
de primer orden, [1](#)
objeto, [24](#)
- lenguajes semanticamente cerrados, [23](#)
- Lewis, David, [189](#)
- Linnebo, ystein, [30](#)
- lp, [81](#){[83](#)
- lp Transparent Truth, vease lp_{tt}, [83](#)
- L_{pa} , [3](#)
- L_{pa}^+ , [7](#)
- lp_{tt}, [83](#), [84](#)
- L_R , [237](#)
- L_T , [8](#)
- L_{zfc} , [14](#)

- L_{zfc}^+ , [15](#)
 L_{zfc}^{+n} , [15](#)
 L_{zfc}^{+u} , [15](#)
 m-referencia
 a números, [213](#)
 a oraciones, [217](#)
 aletica, [227](#)
 Macauley, Matthew, [188](#)
 Martínez, Jose, [22](#)
 Martin, Robert L., [52](#)
 McGee, Vann, [39](#), [65](#), [66](#), [70](#), [80](#)
 metalenguaje, [24](#)
 minimalismo, [120](#), [124](#), [158](#)
 modelos, [2](#)
 clásicos, [36](#)
 de Priest, [86](#)
 de punto jo, [58](#), [61](#)
 m nimo, [59](#)
 no clásicos, [54](#)
 paracompletos/parciales, [54](#)
 paraconsistentes, [81](#), [82](#)
 Montague, Richard, [80](#)
 Moschovakis, Yiannis, [14](#)
 Mostowski, Andrzej, [14](#)
 N, [3](#)
 N^+ , [8](#)
 N^+ -categoricidad, [65](#)
 N , [14](#), [15](#)
 Nelson, Michael, [119](#)
 nombres transparentes, [21](#)
 comillas, [21](#), [22](#), [28](#)
 descripciones estructurales, [21](#), [28](#)
 notaciones para ordinales, [16](#)
 numeral, [3](#)
 operador salto, [57](#), [58](#)
 oración
 de Curry, [86](#), [227](#)
 de Gödel, [14](#)
 de Henkin, [191](#)
 de McGee, [79](#)
 de Yablo, [179](#), [181](#), [233](#)
 del honesto, [59](#), [209](#), [227](#)
 del mentiroso, [22](#), [41](#), [59](#)
 reforzado, [169](#)
 r-decidible, [253](#)
 rd-decidible, [253](#)
 Orayen, Raul, [42](#)
 ordinal
 de Feferman-Schütte, vease [o](#), [17](#)
 prueba-teórico, [17](#)
 pa, [3](#)
 Pailos, Federico, [62](#), [85](#)
 Paoli, Francesco, [90](#)
 pa^+ , [8](#)
 paradoja
 de Curry, [85](#), [86](#)
 de la postal, [70](#)
 de McGee, [80](#)
 de Montague, [80](#)
 de Russell, [40](#)
 de Yablo, [179](#)
 del mentiroso, [22](#), [41](#), [60](#)
 pat, [8](#)
 $pat<$, [48](#)
 pesimismo semántico, [41](#)
 Picollo, Lavinia, [63](#), [65](#), [79](#), [128](#), [129](#), [183](#),
 [276](#)
 pl, [87](#), [88](#)
 pltt, [88](#)
 pluralidades, [42](#)
 Pohlers, Wolfram, [16](#)
 Poincaré, Henri, [178](#)
 portadores de verdad, [21](#), [22](#)
 Positive Uniform Tarski Biconditionals, vease
 putb, [53](#)

- Priest Logic Transparent Truth, vease pltt, [89](#)
- Priest, Graham, [81](#), [82](#), [85](#){[89](#), [110](#), [113](#){
[116](#), [129](#), [133](#), [134](#), [140](#), [182](#){[184](#)
- principio de
 indiscernibilidad de los identicos, [128](#)
 induccion trans nita, [17](#)
 intersustitutividad, [62](#), [84](#), [120](#)
 sustitutividad, [129](#)
 transparencia/equivalencia, [53](#), [62](#), [119](#)
- Prior, Arthur N., [99](#), [127](#)
- propiedad
 de eliminacion de la verdad, [149](#)
 de introduccion de la verdad, [153](#)
 debil, [155](#)
- proyecto minimalista, [166](#), [168](#)
- putb, [66](#), [67](#)
- Putnam, Hilary, [122](#), [193](#), [194](#), [196](#)
- q, [3](#)
- q-referencia
 a numeros, [214](#)
 a oraciones, [220](#)
 aletica
 semantica, [230](#)
 sintactica, [240](#)
- q^+ , [8](#)
- qr, [248](#)
- qr, [250](#)
- qrs, [237](#)
- Quine, Willard Van Orman, [100](#), [101](#), [103](#),
[107](#), [108](#), [111](#), [143](#)
- Rabern, Brian, [188](#)
- Rabern, Landon, [188](#)
- Ramsey, Frank P., [96](#){[99](#), [103](#), [106](#)
- Rathjen, Michael, [250](#)
- Rayo, Agust n, [30](#), [42](#), [43](#)
- redundantismo, [96](#),
[97](#) referencia
- a numeros, [215](#)
- a oraciones, [223](#)
- aletica, [226](#)
- semantica, [231](#)
- sintactica, [244](#)
- referencia directa
 a oraciones, [223](#)
- aletica
 sintactica, [243](#)
- regla
 !, [131](#)
 de induccion in nita, [131](#), [132](#)
- Reinhardt, William, [63](#), [64](#)
- representabilidad
 debil, [4](#)
 de funciones, [4](#)
 de relaciones, [4](#)
- Resnik, Michael, [112](#)
- Restall, Greg, [83](#), [90](#)
- resultados de incompletitud de G•odel
 primero, [13](#)
 segundo, [14](#)
- Ripley, David, [81](#), [91](#), [93](#), [94](#), [128](#), [140](#)
- riqueza esencial, [25](#), [41](#), [42](#)
- Robinson, Raphael M., [14](#)
- Rorty, Richard, [100](#)
- Rosenblatt, Lucas, [62](#), [85](#), [91](#)
- $rt<$, [48](#)
- Russell, Bertrand, [25](#), [99](#), [168](#), [178](#)
- Ryle, Gilbert, [192](#)
- satisfaccion, [29](#){[32](#), [37](#)
- Schindler, Thomas, [76](#), [78](#), [135](#), [169](#), [170](#),
[188](#)
- Self-Reference Tarski Biconditionals, vease
 srtb, [275](#)
- Self-Reference Uniform Tarski Biconditio-
 nals, vease srutb, [277](#)

- Shapiro, Stewart, [9](#), [36](#), [100](#), [112](#){[114](#), [127](#),
[198](#)
 Sheard, Michael, [78](#), [79](#)
 Simpson, Stephen G., [252](#)
 Smith, Peter, [4](#)
 Smorinsky, Craig, [252](#)
 Soames, Scott, [65](#)
 Sorensen, Roy A., [179](#), [184](#)
 sr**tb**, [275](#)
 sr**utb**, [277](#)
 Stern, Johannes, [130](#), [139](#), [165](#), [170](#)
 st, [90](#){[93](#)
 Strawson, Peter F., [97](#)
 Strict Tolerant, vease st, [90](#)
 Strict Tolerant Transparent Truth, vease sttt, [93](#)
 sttt, [93](#)
 Szmuc, Damian, [54](#)
 T In, [121](#)
 T Out, [121](#)
 T-Elim, [52](#)
 T-In, [154](#)
 T-Intro, [52](#)
 T-Out, [65](#)
 T -positividad, [67](#)
 Tajer, Diego, [81](#), [91](#)
 Tarski, Alfred, [14](#), [18](#){[32](#), [34](#){[38](#), [40](#){[42](#), [44](#){
[47](#), [50](#), [118](#), [120](#), [121](#), [125](#), [129](#), [130](#),
[132](#), [168](#)
 Tarski-biconditionals, vease tb, [44](#)
 tb, [44](#)
 tc, [42](#), [43](#)
 Teijeiro, Paula, [54](#)
 teor a
 axiomatica, [2](#)
 de clases, [42](#)
 hiperclases, [43](#)
 de conjuntos con urelementos, vease zfc_U^+ ,
 zfc_n^+ , [15](#)
 de conjuntos Zermelo-Fraenkel, vease
 zfc , zfc^+ , [14](#)
 simple de tipos, [29](#)
 teor as de la verdad
 axiomaticas, [53](#)
 composicionales, [46](#)
 no tipeadas, [53](#)
 semanticas, [53](#)
 tipeadas, [25](#), [49](#)
 transparente, [58](#)
 teorema
 de Beth, [45](#)
 de L•ob, [14](#)
 de la inde nibilidad de la q-referencia
 aletica semantica, [236](#)
 de Tarski de la inde nibilidad de la ver-
 dad, [23](#), [24](#), [40](#), [41](#)
 tonk, [127](#)
 Uniform Tarski-biconditionals, vease utb,
[45](#)
 Urbaniak, Rafal, [198](#), [201](#){[206](#)
 urelementos, [15](#)
 utb, [45](#)
 Uzquiano, Gabriel, [42](#), [43](#)
 van Heuveln, Bram, [182](#)
 van Rooij, Robert, [91](#), [93](#), [140](#)
 Vaught, Robert L., [25](#), [35](#), [36](#), [38](#), [109](#)
 Verdad Rami cada, vease $rt_<$, [48](#)
 Visser, Albert, [5](#), [180](#), [182](#), [187](#), [190](#), [191](#)
 Weir, Alan, [81](#)
 Welch, Philip, [78](#)
 Well-founded Uniform Tarski Biconditio-
 nals, vease wfutb, [263](#)
 wfutb, [263](#), [264](#)
 Whitehead, Alfred N., [99](#)

Williams, Michael, [100](#), [120](#)

Woodruff, Peter W., [52](#)

Yablo, Stephen, [71](#), [178](#){[180](#), [189](#)

Zardini, Elia, [90](#)

zfc, [14](#)

zfc⁺, [15](#)

zfc⁺_n, [15](#)

zfc⁺_u, [15](#)

!, [3](#), [14](#), [16](#) !-

consistencia,

[3](#) !-paradojas,

[183](#)