



FILO:UBA
Facultad de Filosofía y Letras
Universidad de Buenos Aires

G

Métodos computacionales en Antropología

El abordaje conexionista

Autor:

Castro, Damián

Tutor:

Reynoso, Carlos

2003

Tesis presentada con el fin de cumplimentar con los requisitos finales para la obtención del título en Licenciatura de la Facultad de Filosofía y Letras de la Universidad de Buenos Aires en Antropología

Grado



FILO:UBA
Facultad de Filosofía y Letras

FILODIGITAL
Repositorio Institucional de la Facultad
de Filosofía y Letras, UBA

TESIS 10-3-4

Universidad de Buenos Aires
Facultad de Filosofía y Letras
Carrera de Ciencias Antropológicas

FACULTAD de FILOSOFIA Y LETRAS	
Nº 810122	MEC
24 OCT 2003	
Agr.	ENTRADA

Tesis de licenciatura en Ciencias Antropológicas

Métodos Computacionales en Antropología: el
abordaje conexionista

Autor:

Damián Castro

L.U.:22176/86

UNIVERSIDAD DE BUENOS AIRES
FACULTAD DE FILOSOFIA Y LETRAS
Dirección de Bibliotecas

Director:

Prof. Carlos Reynoso

Seattle, Septiembre de 2003.

1	Introducción: planteo de los objetivos, hipótesis y marco teórico	3
1.1	Objetivos	3
1.2	Hipótesis	4
1.3	Marco Teórico	5
1.4	La herramienta	8
2	La antropología cognitiva: la ciencia formal de lo aformal	16
2.1	Un problema acerca de la metodología formal de la antropología cognitiva ...	16
2.1.1	Significado y comportamiento humano	18
2.1.2	Significado y realidad psicológica	19
2.1.3	Cerebro humano y realidad psicológica	23
2.2	Mecanismos de abstracción en la antropología cognitiva	25
2.2.1	El método de la evocación semántica	25
2.2.2	Mecanismos formales de abstracción	28
2.2.3	Árboles, taxonomías y paradigmas	29
3	La revolución conexionista en la antropología cognitiva	35
3.1	De los términos focales a la semántica de prototipos	35
3.1.1	Mecanismos estadísticos	35
3.1.2	Rangos focal y extendido y semántica de prototipos	38
3.2	La teoría de los esquemas	40
3.2.1	Esquemas	41
3.2.2	Imágen-esquemas y proposición-esquemas	44
3.3	Conexionismo	47
3.3.1	Esquemas en la inteligencia Artificial	47
3.4	Esquemas y redes de satisfacción de constricciones	51
3.4.1	RU como redes de satisfacción de constricciones	51
3.4.2	Se cierra el círculo	56
4	Dos experiencias de RU aplicadas a la antropología	60
4.1	Límites metodológicos de la teoría de los esquemas y las RU	60
4.1.1	Modelos del mundo y RU	60
4.1.2	Dos obstáculos metodológicos	62
4.2	Un problema sobre direccionamiento interpersonal estadounidense	66
4.2.1	Modelo clásico	66
4.2.2	Modelo conexionista	68
4.3	Clasificando elementos del sitio de arte rupestre de Campo Morado	73
4.3.1	Entrenamiento supervisado	73
4.3.2	Entrenamiento no supervisado	80
5	Conclusiones	85
6	Bibliografía	91

358607

1 Introducción: planteo de los objetivos, hipótesis y marco teórico

1.1 Objetivos

El objetivo principal del trabajo que aquí comienza es el de determinar si la corriente de inteligencia artificial conexionista verdaderamente constituye un aporte metodológico de algún valor a las Ciencias Antropológicas. Para ello se analizará el proceso de incorporación de dos dispositivos conexionistas, las redes neuronales o redes de unidades (RU)¹ y la teoría cognitiva con que se relacionan, la de los esquemas. Se abordarán dos casos puntuales que se contraponen en el uso de supervisión en el entrenamiento de las RU, que es un aspecto central de la teoría conexionista. Los casos que se examinarán son una red que opera sobre los elementos que componen el sitio de arte rupestre de Campo Morado, Departamento de Tilcara, Jujuy, Argentina, estudiado por María Isabel Henandez Lloza y colaboradores (2001) y la red de direccionamiento interpersonal de Claudia Strauss y Naomi Quinn (1997).

Los objetivos de este trabajo tienen un importante componente metodológico en tanto que se explorará una herramienta y sus posibilidades de aplicación para resolver ciertos problemas de pertinencia antropológica. Se intentará desarrollar los objetivos sin apelar a

¹ En este trabajo se prefiere los términos red de unidades, abreviado RU, o simplemente redes, dado que no tienen las connotaciones biológicas que acarrea el término "neuronal". Por otro lado es el término más aceptado en la bibliografía utilizada para este trabajo.

las vicisitudes del desarrollo del software² que se usó en los experimentos, ni a las complejidades matemáticas del conexionismo en sí.

Por otro lado, gran parte de las cuestiones centrales que se discutirán, aunque relativas a esta herramienta y sus posibilidades metodológicas, son de raigambre teórica: a) el surgimiento la necesidad de herramientas del tipo de las RU en la antropología, específicamente en la antropología cognitiva; b) la teoría conexionista de los esquemas, que la antropología cognitiva retoma, como aglutinante de las conjeturas sobre la cognición humana y como marco a partir del cual se justifica el uso de las RU; c) se examinarán los fundamentos epistemológicos que harían deseable el uso de RU; d) de ser posible abordar exitosamente un problema usando como herramienta las RU, habría dos posibles situaciones: i) se usan las redes para simular una relación entre un conjunto de estímulos y otro de respuestas, ambos de pertinencia antropológica o ii) se puede crear un dispositivo solamente para efectuar operaciones de clasificación, búsqueda o inferencia sobre una población de patrones pertenecientes a un conjunto clases relevantes desde el punto de vista antropológico.

1.2 Hipótesis

De los objetivos de este trabajo pueden desprenderse dos hipótesis:

- a) Las RU pueden ser utilizadas en el marco de la antropología con éxito para simular ciertas conductas simples como las que devienen del sistema de direccionamiento interpersonal.

² El software usado fue desarrollado por ad-hoc, Damián castro, Diego Díaz, Eduardo Jezierski y Jorge Miceli, integrantes del grupo de trabajo del Profesor Carlos Reynoso

- b) Es posible utilizar redes de unidades con entrenamiento para la clasificación de imágenes de pertinencia antropológica, tanto para el desarrollo de herramientas de búsqueda a partir de categorizaciones conocidas como para inferir nuevas configuraciones de categorías.

Este trabajo se va a dividir en tres partes. Primero se intentará mostrar algunos problemas acerca de la realidad psicológica de los dispositivos de la antropología cognitiva, que dejan una brecha a ser ocupada por la teoría de los esquemas. En la segunda parte, se analizará la influencia del conexionismo en la antropología cognitiva en el contexto de la teoría de los esquemas, dado que ésta última ofrece la justificación teórica para el uso de RU. Se concluirá el trabajo con el análisis de los dos casos, mostrando cuáles son sus diferencias y cuáles sus limitaciones a partir de los resultados obtenidos.

1.3 Marco Teórico

La influencia de Kennet Pike fue -y es- muy grande en las ciencias antropológicas. Sin embargo, es realmente disímil. Esa disimilitud abarca un amplio rango de corrientes de investigación, en donde se podría distinguir dos extremos que se encuentran de algún modo yuxtapuestos desde el punto de vista metodológico. Por un lado las corrientes interpretativistas se despojan de toda intención científicista y se contentan con lograr atrapar lo que se da en llamar el punto de vista del actor, pero no como objetivo metodológico, en el sentido de hacer una etnografía que descubra el contenido de la mente de los sujetos estudiados como mecanismo de constitución del dato, el que luego se procesa con los instrumentos teóricos del caso, sino que se trata de un objetivo teórico en sí mismo. Los mecanismos de producción del trabajo científico están circunscriptos a

la particularidad del universo simbólico de los actores culturales. La generalización, luego, sólo es posible a través de mencionada particularidad (Geertz, 1976).

En el otro extremo se encuentra la corriente cognitivista, también llamada etnociencia, que, a pesar de compartir el mismo principio emic, desarrolla una serie de dispositivos con un relativamente alto nivel de formalización. Estas herramientas de relevo y análisis de la información son posibles en la medida que, aunque desde la óptica emic el contenido de una cultura es intraducible a los términos de otra –no pudiéndose establecer comparaciones o generalizaciones en la acepción etic de estos términos-, al menos cierta parte de la cultura posee un nivel de organización a partir del cual se pueden extraer conclusiones en términos formales. La parte de la cultura que más interesa a esta corriente antropológica es el corpus que organiza el conocimiento del entorno. En las palabras de Stephen Tyler:

... cognitive anthropology ... focuses on *discovering* how different peoples organize and use their cultures. This is not so much a search for some generalized unit of behavioral analysis as it is an attempt to understand the *organizing principles underlying* behavior. (Tyler, 1969:3).

Dentro de este corpus, los cognitivistas descubrirán que las clasificaciones zoológicas y botánicas y las relaciones de parentesco son especialmente susceptibles de ser abordadas con las herramientas formales que ellos proponen, conocidas en conjunto como análisis componencial. Por distintas razones, estas herramientas de análisis fracasan y gran parte de los cognitivistas renuncian luego a los principios metodológicos iniciales, siendo el mencionado Tyler uno de los ejemplos de esta conducta. Aunque de algún modo la ciencia cognitiva siguió activa, ya sea gracias a algunos tozudos ejemplos o al aporte de otros ámbitos científicos que estaban enfocados en cuestiones más específicas del

proceso humano de adquisición de conocimiento, a principios de la década del 1970 había encallado en las arenas de la imposibilidad de abarcar un vasto campo de análisis como es la totalidad de la cultura (Reynoso, 1998:31).

Una de los ámbitos científicos que se desarrollaron fuera de la antropología cognitiva, pero cuyo campo está estrechamente relacionado y va a ejercer influencia sobre ella, es el conexionismo. Esta escuela científica se desarrolla en el campo de la inteligencia artificial, el cual no es un yermo monótono e uniforme desde el punto de vista epistemológico, sino más bien un sinuoso territorio plagado de dobleces donde el conexionismo coexiste en pugna con otra escuela de pensamiento: los partidarios del "procesamiento simbólico". Si bien estos últimos tuvieron siempre una amplia supremacía en cuanto a subsidios, el conexionismo se ha desarrollado con bastante intensidad desde la década de 1980 a esta parte. Uno de los grupos de investigadores más importantes en el tema es el llamado "Parallel Distributed Processing" formado por, entre otros, David Rumelhart, James McClelland y Geoffrey Hinton. En este trabajo se utilizan distintos artículos presentados por integrantes de este grupo, puesto que los antropólogos cognitivistas incorporan las ideas conexionistas a través de ellos.

En el contexto de los diferentes proyectos de investigación asociados a la corriente conexionista se desarrollaron distintas herramientas que se agrupan bajo el nombre genérico de "redes de unidades", "redes neuronales" o simplemente "redes". Como se mencionó en los objetivos, en este trabajo se optará por la expresiones "redes de unidades", abreviándola como RU, o simplemente "redes". La posibilidad de utilizar RU como herramienta de representación o de análisis, aumenta enormemente con la

disponibilidad de computadoras capaces de procesar gran cantidad³ de cálculos en un segmento de tiempo pequeño, por lo que el nuevo hardware disponible en la mayoría de las universidades estadounidenses desde los '80 aceleró muchísimo el desarrollo del conexionismo.

1.4 La herramienta

Genéricamente, se puede definir RU como un conjunto de unidades cuya característica relevante es que ante un estímulo numérico, produce una respuesta numérica no aleatoria, regular. La combinatoria de relaciones numéricas entre las unidades –siempre expresadas por un valor o “peso”- es la que determina la respuesta final. Si se quiere una descripción un poco más precisa, en uno de los artículos escritos en conjunto por los tres autores mencionados más arriba, llamado “A General Framework for Parallel Distributed Processing” se encuentra una excelente enumeración de los principales aspectos del modelo de procesamiento paralelo distribuido (1986:49-54), que aquí se desarrollan brevemente:

- a. Un conjunto de unidades de procesamiento, la columna vertebral de las RU. Dependiendo del tipo de red, las unidades pueden ser abstractas desde las que emerge un patrón con alguna significación, o pueden representar ciertas características, como letras o palabras. Más adelante se discutirá la cuestión acerca de asignarle significación a las unidades en los términos de la propia red.

³ Luego veremos que la cantidad de cálculos no es tan grande si sólo disponemos de operaciones del tipo de intercambio de pesos entre neuronas. Como se ve en la figura 1, para una simple operación de tipo XOR, se requiere un grupo de 5 neuronas, una función para evaluar los pesos que se ejecuta en cada neurona, etc.

- b. Un estado de activación. Las unidades pueden aceptar distintos tipos de números como descriptores de su estado en cualquier tiempo t . Como mínimo, la red debe tener dos estados posibles, activo o inactivo (1 y 0 ó 1 y -1). También se puede usar una escala continua en vez de discreta.
- c. Valor de salida de las unidades. Como las unidades interactúan, cada unidad tiene un valor que afecta a las unidades hacia las que está conectada. Al decir “hacia”, se entiende que las conexiones entre unidades se dirigen en un sentido, y que el valor de salida de una unidad es el de entrada de todas las unidades hacia las que está conectada.
- d. Patrón de conectividad. Las unidades no se conectan entre sí necesariamente todas con todas o aleatoriamente. Es posible, entonces, que las mismas formen distintos tipos de “topologías”, donde un conjunto -también llamado nivel- de unidades se conecta con otro conjunto. Una red, por ejemplo, puede tener tres niveles de unidades. Un nivel puede ser compuesto por las unidades que reciben el estímulo externo. Estas unidades se conectan con las de otro nivel intermedio, o “escondido” ya que no interactúan con el exterior de la red. Este último a su vez se conecta con las unidades de un tercer nivel, el de “salida”, que son las que reciben alguna interpretación respecto de su relación con el estímulo inicial (si éste pertenece a una cierta clase, etc.).
- e. Regla de activación. Esta se calcula usando el vector de pesos de entrada que se propaga hacia la unidad en un tiempo $t+1$, combinándolo con el actual peso de la

unidad (es decir, el peso que tiene en un tiempo t)⁴. Así se produce un nuevo valor de activación en el tiempo $t+1$. Esta regla de activación se expresa en la forma de $a(t+1)=f(\text{net}(t))$, donde a es el valor de activación resultante, t es el tiempo actual y net es el vector con todos los pesos que recibe.

- f. Regla de propagación. Esta regla dirá como se convierten los pesos de entrada en un valor único de entrada. Los pesos pueden corresponder a conexiones inhibitorias o excitatorias y puede haber una inclinación o “bias”, que es un valor que se suma o multiplica al vector de pesos de entrada. Entonces, lo más simple sería sumar los pesos excitatorios, restar los inhibitorios y sumar o multiplicar la inclinación.
- g. Aprendizaje. En la mayoría redes se implementan mecanismos de modificación de los pesos entre unidades de modo de ir reduciendo paulatinamente el error en relación a un resultado esperado.

Un caso simple para ilustrar el funcionamiento de una red es la disyunción exclusiva, abreviada “XOR”, que consiste en 5 unidades, dos de entrada, dos internas y una de salida (figura 1).

⁴ Las RU funcionan a través de ciclos de actividad, en los que se aplican las reglas fluyendo en el sentido que establecen las conexiones. Es así que cuando se ingresa un valor a la red, primero se activan las unidades de entrada. Luego, se propagan los valores de salida de estas unidades a través de las conexiones que éstas tengan hacia el nivel siguiente de unidades. De esta forma se “secuencializa” el flujo de la red. La razón de esta secuencialización, puede que sea que, a pesar de todo, las máquinas con las que se simula las redes de unidades son computadoras secuenciales. Por otro lado, la lógica de propagación que tienen estos modelos, requieren que antes de que una unidad envíe su valor de salida hacia las unidades hacia las que están conectadas, debe recibir el peso de todas las conexiones que la tienen como destino.

Las dos primeras unidades desde el lado izquierdo son las de entrada y tienen cada una dos conexiones con las dos internas. Una conexión será excitativa (+1) y la otra inhibitoria (-1). La unidad de entrada sólo puede recibir los valores 1 ó 0. Si recibe el valor 1, se activa y envía a su vez un estímulo excitativo a una de las unidades internas y un inhibitorio a la otra. La segunda unidad de entrada hará lo propio, pero enviando la carga positiva hacia la unidad que recibe la negativa de la primera unidad de entrada, y la negativa a la que recibe la positiva. Por ello, cada una de las unidades internas recibe dos cargas al mismo tiempo que pueden ser una positiva y una negativa o las dos negativas.

A su tiempo, la unidad interna enviará una carga excitativa a la de salida si la suma de las cargas que recibe es mayor que 0. Entonces, para que la unidad de salida reciba excitación positiva, alcanza con que una de las unidades internas se "active" teniendo una carga positiva. Eso sólo ocurre cuando una sola unidad de entrada está activa pues, de otro modo, la carga positiva que se envía a la unidad interna se anularía con la carga negativa de la otra unidad de entrada. De ese modo, se activará con los patrones de entrada "10" y "01" y no con los "11" o "00".

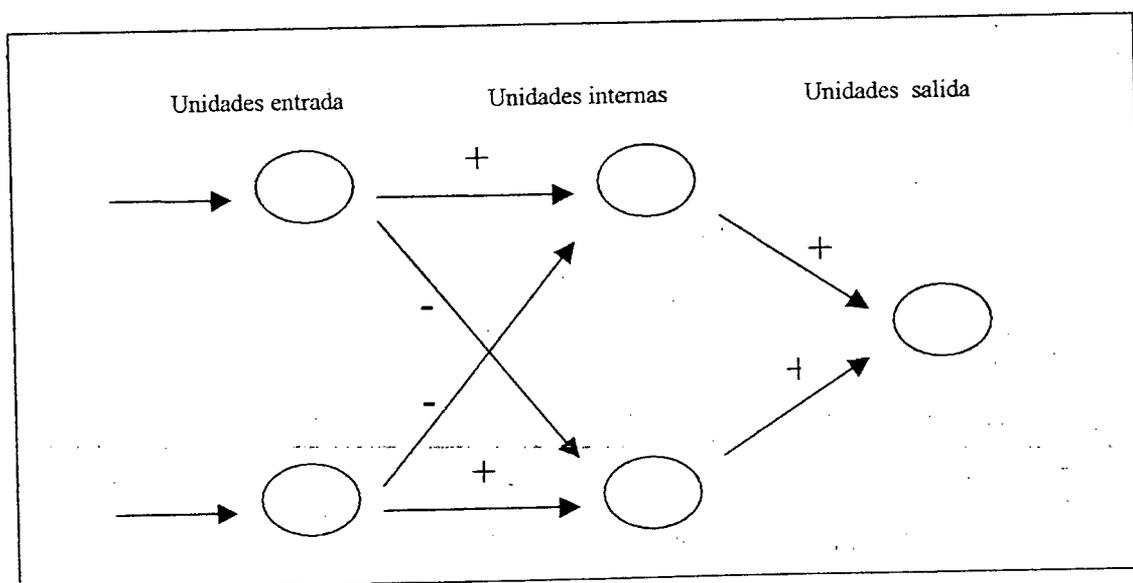


Figura 1. Una red que implementa la operación lógica XOR.

Una de las características notables de la RU que se acaba de describir, es que en ningún momento se explicita la regla lógica que la red simula: la misma es una propiedad emergente que no puede ser inferida por las propiedades de ninguna unidad o conexión en particular. Se puede explicar el proceso a partir del cual un estímulo se transforma en una respuesta luego de una serie de sumas y restas, pero la regla de disyunción exclusiva que la RU representa nunca es enunciada explícitamente con el uso de símbolo alguno, sino que es un emergente de la red.

Esta RU posee representados los 6 primeros aspectos que se enumeraron más arriba:

- a. Un conjunto de 5 unidades de procesamiento.
- b. De acuerdo al estímulo ingresado, cada una de estas 5 unidades adquiere un estado de activación que puede ser 1 o 0
- c. El valor de salida es igual al estado de activación.
- d. Esta RU posee un cierto patrón de conectividad, en el que, por ejemplo, se puede reconocer un nivel de entrada de 2 neuronas, un nivel intermedio de 2 y una salida de 1.
- e. La regla de propagación es simple: se suma las conexiones activas.
- f. La regla de activación es la siguiente función:

$$f = \begin{cases} \text{si } \sum_j w_{ij} o_j \geq 1 \rightarrow 1 \\ \text{si } \sum_j w_{ij} o_j < 1 \rightarrow 0 \end{cases}$$

en donde el valor de salida de la unidad μ_i es a) igual a 1 si la sumatoria de los pesos de las conexiones del nivel j con la unidad μ_i multiplicado por el valor de salida de cada unidad μ_j es igual o mayor que 1 o b) es igual a 0 si la sumatoria anterior es menor a 1.

Una particularidad de este tipo de red es que sus pesos son fijos, por lo que no aprende en el sentido del aspecto g) de la enumeración vista más arriba. En realidad, aprender para una red es sólo ajustar los pesos por un cierto algoritmo que disminuya gradualmente la diferencia entre la respuesta esperada y la obtenida. Según Rumelhart y Zipser (1986:160) existen cuatro paradigmas de aprendizaje en las RU:

1. Auto asociador. Aquí, un conjunto de patrones es presentado a la RU durante su entranamiento. Luego del mismo, si la RU recibe como ingreso un patrón o una parte de un patrón similar a alguno de los que se presentaron durante el entranamiento, la RU devuelve como resultado el patrón que se presentó al principio.
2. Asociador de patrones. Es en realidad una variante del anterior. La RU se va entrenar con pares de patrones y va a aprender a asociar uno con otro, de modo que si, por ejemplo, los patrones S_k y T_k son un par asociado, si se ingresa S_k , la RU devuelve T_k .
3. Paradigma de clasificación. En el paradigma de clasificación las RU aprenden a asignar un patrón a una clase a través del procesamiento de un conjunto de ejemplos de cada clase. El patrón que se ingresa puede ser tanto un patrón nuevo como uno que se le presentó como estímulo en la etapa de entranamiento, aunque distorsionado. Un caso de esto es una de las dos RU para clasificar elementos tomados de Campo Morado que se analizará con mayor detenimiento más adelante.
4. Detector de regularidades. En este paradigma, un conjunto de patrones se presentan a la RU. La red se supone que descubrirá las características salientes

o estadísticamente más frecuentes. En este caso, a diferencia de los dos anteriores, no existe una cantidad prefijada de clases, sino que la RU descubrirá o detectará los patrones que comparten las regularidades más destacables. También se verá un ejemplo de este paradigma más adelante, usando las imágenes correspondientes al sitio de Campo Morado.

Sea cual sea el paradigma, existe una variedad de algoritmos para determinar el valor óptimo de los pesos entre unidades y es realmente cierta la posibilidad de procesar estímulos mucho más complejos que el par de bits de la red XOR. Si se dispone de patrones que pueden ser representados en un vector y ser usados como ejemplos de clases, las RU pueden ser "incentivadas" a través de ajustar sus pesos de acuerdo a la distancia entre la respuesta actual y la correcta hasta lograr un equilibrio en el que los pesos son igualmente buenos para cualquiera de los ejemplos presentados. Para las RU cuyo ajuste no requiere incentivación (o supervisión, en la jerga conexionista) como las del paradigma d), la red podrá al menos decir que miembros de la población son más parecidos entre sí y formar una clase con ellos.

Esta herramienta presenta una serie de matices interesantes que la antropología intenta aprovechar en el marco de la corriente cognitivista. Sin embargo, como se ve más adelante, no se ha desarrollado una metodología que incluya a las RU como herramienta de uso aunque sea ocasional. En el grueso de los trabajos en que los conceptos conexionistas aparecen, las RU no se mencionan. Y en las únicas excepciones, como en la red de sistemas de direccionamiento de Strauss y Quinn, no se implementa ningún tipo de algoritmo de aprendizaje. En la red de estas autoras, los pesos entre unidades son asignados manualmente para determinar una respuesta en un caso puntual. Además, la

relación entre las unidades de entrada y salida y los complejos semánticos correspondientes (la persona direccionada y su apelativo correcto), es una hipótesis que no se discute dentro del trabajo. La red sería más bien como la que desarrollan Hinton, McClelland y Rumelhart al diseñar un sistema de representación distribuida que interrelaciona dos dominios, grafemas y sememas o unidades semánticas. (Hinton, McClelland y Rumelhart, 1986:96-104).

De cualquier modo, la mera aproximación a una herramienta tan compleja es meritoria y el hecho de que la red de direccionamiento realmente funciona en los parámetros definidos por las autoras es loable. Por otro lado, dado que el conexionismo durante los '70 y los '80 se desarrolla en el marco más genérico de las ciencias cognitivas, es de algún modo natural que los antropólogos cognitivos busquen en las redes las respuestas a los problemas que las deficiencias de sus herramientas de análisis son incapaces de ofrecer. Por esa razón también, en este trabajo se abordará el problema de la incorporación de esta herramienta al stock disponible de metodologías antropológicas a la luz de la corriente cognitivista.

2 La antropología cognitiva: la ciencia formal de lo aformal

2.1 Un problema acerca de la metodología formal de la antropología cognitiva

2.1.1 Cientificidad y punto de vista emic

A los ojos del novato lingüista del Summer Institute of Linguistics (SIL) que era Pike⁵ cuando uno de sus primeros trabajos le fue encomendado, trabajo que consistía en la traducción de la Biblia al mixteca, su labor no podía completarse sólo sabiendo utilizar fluidamente la lengua en cuestión, sino que debía ser capaz de transmitir la doctrina religiosa en términos tales que los receptores la interpretaran como si les fuera propia, es decir, en concordancia con su visión del mundo.

No obstante, para Pike la empresa de entender la visión del mundo en los ojos del otro no se haría pagando con la pérdida de la científicidad. Su explícita intención era que siguieran existiendo conceptos, categorías e incluso reglas y leyes que permitieran explicar la conducta humana a nivel general, tanto en la lingüística en particular como en todas las ciencias del comportamiento humano en general. A la luz de lo que sucedió después con el advenimiento del irracionalismo posmodernista, esta sana intención de seguir produciendo conocimiento científico a pesar de los mencionados preceptos teóricos relativistas sorprende positivamente. Hay dos ejemplos que pueden ilustrar claramente esta intención.

⁵ Sobre esta etapa de la vida de Kenneth Pike, es interesante el texto *Pike's Perspectives* (Pike, Kenneth and Hugh Steven 1989) cuyos capítulos 1 y 2 versan sobre las primeras ideas de Pike en la lingüística, empezando en 1935, cuando realiza su primer viaje al país mixteca.

El primero, en una carta del 4 de diciembre de 1968 dirigida a los miembros del SIL, Pike mientras habla de ciertos problemas con los pronombres de la lengua africana Bariba, sostiene que “tratando, más tarde, de dar sentido a los problemas en términos del inglés, me pareció que la teoría de grupos podría ser de ayuda” (Pike, 1989:4). En la misma carta discute brevemente si la lingüística es parte de la matemática, al ser la última el estudio del patrón abstracto y la primera la “abstracción del patrón del lenguaje”, o viceversa, pues la lingüística es el estudio de la comunicación y la matemática es una forma comunicable. De cualquier manera, ya para entonces él había llegado a la conclusión de que ambas están muy relacionadas. Inclusive, desde muy temprano el SIL intenta reclutar matemáticos en sus filas a la espera de que puedan contribuir a la lingüística, aunque este proyecto fracasa, como él mismo Pike reconoce, dado que se logra capacitarlos como lingüistas, aunque su previo conocimiento matemático no encuentra un lugar específico en la lingüística.

El otro ejemplo es su intento de utilizar los conceptos de onda y partícula, categorías tomadas directamente de la física, en el análisis de lenguaje y la conducta. Esta teoría comienza a plasmarse en el artículo de 1954 titulado “Language in relation to a unified theory of the structure of human behavior” y en otro de 1956 titulado “Towards a theory of the structure of human behavior”. En este último dice que la conducta humana existe en un continuum “que no es estático no-fluctuante sino que hay en él algunas ‘ondas’ de actividad”. Sin embargo, sigue, la gente reacciona al comportamiento humano como si fueran partículas, o más bien “una secuencia de partículas individuales de actividad” (Pike, 1972:107). Como sucedió con el intento de cerrar la brecha entre la matemática y la lingüística a través de incorporar al SIL matemáticos, este intento fue de algún modo

también un fracaso, pues no tuvo realmente un impacto en la metodología a través de la creación de nuevos dispositivos de análisis.

2.1.2 Significado y comportamiento humano

A pesar de compartir en cierta medida una visión de que es lo que constituye el objeto de estudio de las ciencias del comportamiento, Pike es raramente citado por los textos cognitivistas.⁶ Hay muchas razones que se pueden esgrimir para justificar este hecho, como los distintos programas en que ambos estudios se desarrollan, diferentes intereses en cuanto a la aplicabilidad, el foco puesto en partes disímiles del comportamiento humano, etcétera. Posiblemente estas razones sean todas válidas, pero hay una diferencia mucho más profunda y que es la llamativa escasez en la metodología de Pike de herramientas útiles que permitan expresar formalmente el significado. No es que Pike no estuviese preocupado por el significado; todo lo contrario, lo estaba y mucho, pero desde una perspectiva lingüística, más estrechamente relacionada con la estructura del lenguaje que con la relación del lenguaje con el entorno. La teoría de Pike de ondas y partículas en las que se constituye el comportamiento, era el producto de suponer que si el lenguaje es estructurado a partir de fronteras de significación trazadas en el continuum de sonidos y silencios, bien podría ser que toda la interacción humana se entendiese del mismo modo, en tanto las gentes están inmersas en un continuum de actividad, que lo humanos dominan e inflexionan a discreción para que en determinados puntos se convierta en unidades de comportamiento. El significado es un tipo de estas unidades que se separan

⁶ En la recopilación "Cognitive Anthropology" del Stephen Tyler, es sólo citado una vez, casi tangencialmente, en el artículo de Frake, el primero del libro, titulado "The ethnographic Study of cognitive Systems", (p. 30)

del continuun. Sin embargo, la teoría no explica porqué se establecen las fronteras de significación en uno u otro lugar. De ese modo, ese marco formal y cientificista dado al comportamiento humano deja la constitución del significado fuera del análisis.

El significado desde el punto de vista emic que Pike propicia para el estudio del comportamiento, sólo se analiza a través del pensamiento de los actores sociales. Por lo tanto, si el universo semántico de una cultura dada tiene reglas, leyes y estructura, no es posible conocerla de antemano o predecirla a partir de un conocimiento abstracto aplicable a nivel general. Son los propios sujetos culturales quienes lo darán a conocer durante el trabajo de campo. En ese sentido, el análisis formal del significado quedaba fuera de programa, al menos como contenido, pues, como se dijo, los propios actores sociales eran los únicos capaces de dar cuenta del significado.

Es en esta brecha donde la antropología cognitiva encuentra un campo fértil para desarrollar su empresa de encontrar una manera de expresar formalmente el significado, dar cuenta de los distintos componentes que lo conforman y analizar las reglas de combinación que operan en él. Pero, ¿cómo es posible abstraer reglas o leyes a partir de un análisis formal que excluya la comparación entre conjuntos de datos procedentes de distintas culturas, pero que a la vez las reglas les sean aplicables indistintamente? Dicho de otra forma, ¿se pueden abstraer reglas o conceptos universales a partir de universos en principio irreductibles entre sí?

2.1.3 Significado y realidad psicológica

Si se piensa la cuestión desde una óptica ecocultural, todas las sociedades tienen que satisfacer un conjunto más o menos similar de necesidades en un medio natural cuyas variaciones son conocidas con bastante detalle. Por lo tanto es posible que, si se analizan

ciertos significados en una cultura en relación a su entorno, se puedan encontrar algunos conceptos que se pueden utilizar en generalizaciones que crucen las fronteras de la cultura analizada. Al fin de cuentas, que un conjunto de símbolos no sea reductible a los términos de otro, no significa que ambos universos no compartan nada en un nivel más bajo –o más alto– de abstracción. La irreductibilidad de un universo semántico y otro no radica necesariamente en la presencia de elementos ontológicamente o cualitativamente diferentes. La irreductibilidad puede radicar en la manera que en ambos universos organizan componentes similares o como se aplican de manera diferente los mismos sistemas y reglas de organización a los mismos elementos. Nadie afirmaría, por otro lado, que las diferencias étnicas se relacionen con alguna diferencia fisiológica: el “hardware” humano es el mismo en todas partes. Entonces, los antropólogos cognitivistas llegan a la conclusión que, aún cuando se evite cualquier estrategia de análisis comparativista, a partir del estudio de una cultura es posible encontrar elementos comunes o reglas de amplia aplicabilidad.

El problema que se suscita en este punto es cómo se establece la realidad psicológica de esas formalizaciones. En ese sentido, una serie de preguntas comienzan a palpitarse cuando se piensa en este problema. ¿Cómo deben ser recabadas estas abstracciones? ¿Qué lenguaje debe utilizarse? ¿Qué estrategia de conformación de preguntas? ¿Qué conceptos nuevos/existentes en la cultura se usan y cómo se aprenden? Estas no son de ninguna manera preguntas menores, pues es posible que una mala estrategia de evocación de significados, introduzca los propios elementos abstractos que el antropólogo maneja, produciendo artificialmente en el relevamiento de campo aquello que se pretende encontrar. Esta cuestión puede ser caracterizada sintéticamente, ya que el problema es, en

última instancia, simple: ¿cómo se establece la validez fáctica de una abstracción de un componente semántico o cultural? Aún cuando el cuestionario sea lo suficientemente simple y sólo se usen universales del estilo “¿Qué tipo de árbol es aquel?” (Frake:1969), se corre el riesgo de que las abstracciones que se hagan no correspondan con la realidad psicológica de los nativos, entendiéndose por esto último que aunque sean válidas desde el punto lógico y tengan algún tipo de correlato en la conducta observable, no sea posible probar su existencia en las “mentes” de aquellos. Esto es lo que sucede con el número cero en el sistema de pisos que se usa en algunos países como EE.UU. Allí, es posible que se infiera la presencia lógica del concepto de cero a partir de los otros números y de su existencia en otros ámbitos de esa cultura. Incluso, se podría ubicarlo físicamente en el “primer subsuelo” y aducir que su uso es una conducta observable, pues la gente embarca y desembarca del ascensor en dicho piso. Sin embargo, un ascensor que marque el piso cero no tendría ningún sentido pues los nativos estadounidenses no sabrían exactamente a qué lugar físico se refiere.

La solución nihilista que niega la existencia de nada que se encuentre más allá de la textualidad del discurso del informante a la que uno no puede más que acoplarse con su propia textualidad, no sirve al programa científico que se pretende llevar a cabo. Si se tiene la intención un cierto nivel de abstracción, uno se tiene que separar del discurso del informante y desarrollar el lenguaje formal del caso, más allá de si es un enfoque emic o etic, del mismo modo que se deja en algún punto de mover peras y manzanas de un lado a otro y se empieza a usar signos abstractos para sumarlas o restarlas. Luego, el problema no es que el lenguaje formal esté presente o no en el discurso, ya que se descarta su

presencia de antemano, pues en su propia definición el lenguaje formal no repite el discurso nativo sino que abstrae de éste las propiedades relevantes a un problema dado.

Esta cuestión no estuvo muy en clara durante el cognitivismo de fines de los '50 y los '60. Como bien lo señala Carlos Reynoso en la obra "Corrientes en Antropología contemporánea" a propósito de la pretensión emic de poder abstraer el conocimiento antropológico directamente del discurso, "así como estar vivo no concede a nadie el conocimiento de biología, ser miembro de una cultura no habilita para arrojar una buena mirada antropológica sobre ella" (1998:31). Las abstracciones que la ciencia desarrolla, entonces, no son parte del campo semántico estudiado.

De todos modos, la relación entre las propiedades fácticas encontradas en el discurso y el lenguaje que las describe formalmente debe articularse de forma que se permita un ida y vuelta más o menos fluido entre lo que se afirma de manera abstracta y lo que sucede en la realidad analizada. Ida y vuelta en este caso quiere decir que las abstracciones tienen que tener consecuencias observacionales en la forma de predicciones, descubrimiento de nuevas propiedades o de regularidades que puedan predicarse universalmente o de una cierta clase en la que se agrupen los datos empíricos.

De ninguna manera, se quisiera aquí dar a entender que el problema de la realidad psicológica de las abstracciones sobre la cognición humana es trivial. Se trata de un tema que a pesar de los esfuerzos de varias disciplinas y más de cien años de investigación, no ha dejado de ser un terreno de especulación y conjetura. Pero justamente por ello es que el camino indirecto de las consecuencias observacionales adquiere interés. De todos modos, vale la pena detenerse algunas páginas sobre la cuestión, especialmente tomando

en cuenta que las redes de unidades o “redes neuronales” de alguna manera intentarían emular ciertas propiedades del cerebro humano.

2.1.4 Cerebro humano y realidad psicológica

¿Cómo es posible saber qué es lo hay, en la mente de un ser humano, en el sentido más rotundamente empírico? Recuérdese el gráfico de la página 11, en el que se establece un mecanismo a partir del cual es posible discernir entre los pares de números binarios, en cuáles uno es distinto de otro (10 ó 01). ¿Cómo se podría afirmar o descartar sin dudas que “eso” es lo que está sucediendo en el cerebro de los seres humano cuando se le pide a un individuo de dicha especie el discernir los pares que son XOR y los que no lo son? Desde el punto de vista fisiológico, esto es imposible de saber. En ese sentido, Jacob Schwartz en un artículo revela una serie de interesantes estadísticas acerca del cerebro humano:

“El cerebro humano consiste en aproximadamente 100 mil millones de neuronas, que posiblemente sean 10 veces más. Las neuronas se comunican habitualmente transmitiendo paquetes eléctricos discretos (potenciales de acción) a una población de neuronas vecinas. Por lo que se sabe, la amplitud precisa y la forma de esos paquetes y el tiempo preciso para su llegada en un intervalo de dos milisegundos o algo así, son detalles físicos que el sistema nervioso no es capaz de explotar. De allí que se pueda modelizar cada paquete como un bit singular... Esta forma de pensar lleva a un estimado 10 billones de bits por segundo, tomando un factor de cien, para el ancho de banda del cerebro.” (Schwartz, 1988:147)

Es decir que incluso si se lograra un método para registrar todas esas transmisiones químicas, la masa de datos a analizar sería de todos modos abrumadoramente imposible de manipular. No es que no haya a esta altura una máquina capaz de almacenar esa

cantidad, algo así como 1.250.000 gigabytes. En menos de 10 años, posiblemente cualquier universidad del primer mundo tenga algún que otro equipo con esa capacidad de almacenamiento. Sin embargo, no es el almacenamiento de esa cantidad de información lo que imposibilita este tipo de aproximaciones a la realidad física de las "ideas", sino el hecho que el cerebro procesa toda esa información en un lapso tan corto de tiempo, paralelizando su actividad, lo que convierte en infructuoso hasta el ridículo pensar en simular o analizar su funcionamiento con cualquier herramienta conocida. De ese modo, aunque tuvieran componentes directamente abstraídos del funcionamiento del cerebro, las RU no pueden resolver el problema de la realidad psicológica "simulando" un conjunto de cerebros interaccionado, como en una sociedad o medio cultural real.

No obstante, no es indispensable para representar un problema y elaborar una solución, conocer la totalidad de su universo causal. Sólo se requiere conocer las variables relevantes a la estructura del problema en cuestión, la que puede en muchos casos representarse formalmente. Por ejemplo, para contar manzanas en un cajón, no requiero saber su constitución molecular. La estructura del problema allí sólo requiere saber ciertas propiedades de la manzana como la de presentarse en unidades discretas, a diferencia de la harina, en cuyo caso la propiedad relevante podría ser el peso. Y si lo que tuviese que contar fueran parientes para hacer una lista para una fiesta o un ritual, las propiedades relevantes podrían ser las relaciones de sangre y alianzas nupciales.

Pero entonces, ¿en qué sentido falla la antropología cognitiva de los '60 en relación a representar la realidad psicológica? ¿Cómo esto le abre camino al conexionismo? Esas dos preguntas se contestarán en los próximos capítulos.

2.2 Mecanismos de abstracción en la antropología cognitiva

2.2.1 El método de la evocación semántica

En el artículo "Ethnographic description and the Study of Law" Black y Metsger describen el proceso de evocación (eliciting) de significados, en el que establecen una serie de pautas para construir cuestionarios confiables en lengua nativa. Allí sostienen que el investigador debe lograr un perfeccionamiento en sus preguntas de forma tal que posean tres propiedades: a) gramáticamente correctas; b) interpretables semánticamente; y, de serlo, c) carentes de ambigüedad (1969:143). Para lograr ello, desarrollan un método simple que prueba la adecuación semántica de las preguntas a través de

"... (1) observing the ease with which the informant responds, (2) direct questioning to the informant, (3) examining the linguistic form of the response, and (4) examining the semantic content of the response"(1969:143).

Este método, aparte de enfocarse en la evocación semántica exclusivamente, se caracteriza por tener cierto nivel de interactividad ya que el informante aprende el método y puede distinguir al poco tiempo si es cuestionado respecto de la adecuación de las preguntas o si se busca simplemente un significado específico. Se trata de usar a los propios nativos de jueces de la pertinencia de las preguntas en cuanto al objetivo de buscar significados concretos. La meta última es mejorar las preguntas hasta que las mismas carezcan de elementos ambiguos o desorientadores, de modo que la información obtenida sea lo más precisa posible.

Estos autores proponen descomponer a las preguntas en dos partes: el marco y el término (1969:143). El primero es el "esqueleto interrogativo", pues contiene la

referencia a una cualidad o propiedad sobre la que el investigador quiere saber si se puede predicar del segundo, que es el foco de la evocación de significado. Este nivel de especificación del método de cuestionamiento, puede que no sea compartido por todos los antropólogos cognitivistas, pero su fundamento último sí lo es. La idea que subyace a esta metodología es que los significados —o, como se verá más adelante, sus componentes— son susceptibles de ser evocados a través del lenguaje. Como dice Frake:

“Since the knowledge that enables one to behave appropriately is acquired from other people, it must be communicable in some symbolic system which can travel between one mind and another as code signal in a physical channel” (1969:124)

El hecho de que determinados significados se aprenden mucho antes que la capacidad de entender su definición invalida la idea de Frake, que excluye la posibilidad de que haya propiedades emergentes de la conducta y la interacción con las otras personas en la forma de significados. Inclusive, los significados que sólo se pueden adquirir con su uso son bastantes, pues se puede poner allí todos los adquiridos en la primera infancia. Sin embargo, como se vio en el capítulo anterior, los mecanismos formales requieren que se exprese en ellos únicamente las características relevantes de un determinado problema a resolver. Por lo tanto, a pesar de dejar afuera una buena parte del universo semántico al considerar sólo aquellos significados que se pueden transmitir a través de la comunicación simbólica, existe todavía la posibilidad de encontrar determinadas partes del campo de la cultura donde el modelo cognitivista se aplique, un subconjunto, si se quiere, de todos los significados posibles. Pero entonces ¿qué es significado para los antropólogos cognitivistas?

Donde se encuentra la respuesta a esta pregunta es en el trabajo de Charles Morris. El trabajo de este autor es usado como referente semiótico por varios antropólogos cognitivistas (Conklin, Black, Lounsbury, Wallace y Atkins y otros) los que emplean su terminología e ideas cuando tienen que referirse a la cuestión del significado. El pensamiento de Morris podría ser clasificado como "semiótica conductista" (behavioral semiotic), como él mismo lo afirma (1946:28). Este autor parte del supuesto que es mucho más fructífero que la ciencia de los signos se "desarrolle sobre fundamentos biológicos y específicamente en el marco de la ciencia de la conducta" (1946:2). Este supuesto, garantizaría cierto grado de cientificidad al estudio de los signos, y, si bien Morris reconoce que las categorías de la escuela "mentalistas" como "idea", "pensamiento", "conciencia" o "mente" no carecen de sentido, él cree que una aproximación conductista permitiría un mayor avance científico. La razón, por supuesto, es que semejantes cosas no son observables cuando la conducta sí lo es. En definitiva, signo se va a definir en términos de estímulo y respuesta:

"If anything, A, is a preparatory-stimulus which in the absence of stimulus objects initiating response-sequences of a certain behavior-family causes a disposition in some organism to respond under conditions by response-sequences of this behavior-family, then A is a sign" (1946:10)

Si el signo sólo puede ser abordado a través de hechos empíricamente observables, el significado debe ser estudiado a partir de, o bien la reacción conductual que el signo provoca en quien lo interpreta, o bien en el referente u objeto denotado por el signo. Es así que en la terminología básica de la semiótica de Morris, que los antropólogos cognitivistas seguirán, no se encuentran referencias a ningún proceso mental en tanto tal.

El análisis girará en torno a los siguientes conceptos: intérprete, interpretante, denotatum y significatum. Así se articulan en sus propias palabras.

“...Any organism for which something is a sign will be called an *interpreter*. The disposition in an interpreter to respond, because of the sign, by response-sequences of some behavior-family will be called an *interpretant*. Anything which would permit the completion of the response-sequences to which the interpreter is disposed because of a sign will be called *denotatum* of a sign. A sign will be said to denote a denotatum. Those conditions which are such that whatever fulfills them is a denotatum will be called a *significatum* of the sign. A sign will be said to signify a significatum;...”(1946:17)

2.2.2 Mecanismos formales de abstracción

Los mecanismos de abstracción de los antropólogos cognitivistas clasificarán palabras en distintas estructuras que organizarán dominios semánticos, separando en “componentes” las condiciones para que un objeto/sujeto pueda ser el denotatum de un término. Floyd G. Lounsbury llega a decir directamente que “la definición componencial de un término es la expresión de su significatum” (1969:193), lo que realmente sintetiza el objetivo último de su programa: expresar el significado de un término formalmente a través de un conjunto de condiciones que se denominan componentes.

Sin embargo, no son los componentes en sí lo que constituye la espina dorsal del análisis componencial, pues estos son tan sólo la expresión simbólica de las propiedades que definen la pertenencia de un objeto a una clase, atributos que se predicán de ésta. En la medida que se profundiza su análisis, sólo se entrará al ámbito de conocimiento específico al dominio de la clase, alejándose de la antropología cognitiva. Entonces, no es

la cantidad de componentes ni la especificidad de los mismos lo que se busca sino las relaciones que puedan establecerse entre componentes y términos.

En resumen, el análisis componencial comienza con la evocación del significado que ofrece un acercamiento descriptivo e inicial. De allí se extrae el "material" que llenará las estructuras formales que este análisis utiliza. En definitiva, la pregunta que intenta contestar la antropología cognitiva es, como decía Stephen Tyler, cómo los distintos pueblos organizan y usan sus culturas (1969:3). El análisis componencial intenta, entonces, descubrir la "forma" en que las culturas ordenan el mundo —o los distintos dominios que lo conforman— a través de la terminología utilizada.

2.2.3 Árboles, taxonomías y paradigmas

En el ámbito del análisis componencial se reconocen varios arreglos en que los componentes pueden ser ordenados, entre ellos los árboles, las taxonomías y los paradigmas⁷.

En los árboles, se organizan los términos a partir de la presencia o ausencia de un componente por vez, de modo que de cada uno parten como máximo dos ramas. Es importante no confundir, como señala Paul Kay (1969:82), árbol con "llave" (key), que es en realidad la forma en que se representa una determinada configuración semántica. En ese sentido, es tan sólo una manera de graficación donde los nodos se agrupan con un corchete bajo un nodo que los contiene. Según Kay, un árbol perfecto es una llave en la que cada dimensión es aplicada solamente en un nodo (1969:83), lo cual hace muy difícil

⁷ También podrían agregarse las matrices, que aquí no se verán.

su existencia en grandes dominios completos. Normalmente, ciertas dimensiones tienden a repetirse, como por ejemplo la forma o el color.

Siguiendo nuevamente a Kay, si cada nodo tiene asociado un lexema, el árbol es también una taxonomía (1969:83). La diferencia es muy sutil, pues el hecho que no haya un lexema para referirse a un determinado nodo no es en realidad inherente ni a la idea que se tiene de la clase que el nodo representa, ni a la realidad a la que se refiere. Inclusive, aunque no exista un lexema, puede que la idea o el sentido de clase sea muy fuerte en una cultura. Por ejemplo, "pariente político" - de uso en varias culturas de habla hispana- no es de ningún modo un lexema, aunque está agrupando una clase de términos bien definida en aquellas culturas. En la vida real, la mayoría de las taxonomías no son árboles, ya que suelen tener distintas dimensiones que se aplican a más de un nodo, como "el largo del pelo" que puede determinar distintas variedades de gato o perro o la "hoja de aguja", que está tanto en coníferas (pino) como en árboles de copa redondeada (taray).

De todos modos, ni siquiera en el ámbito de la clasificación del reino vegetal se presenta un conjunto arbolado de términos ordenados graciosamente en unívocas ramas y subramas. Por ejemplo, tanto entre los Yuroks como los Tolowa del Norte de California, donde abundan las coníferas como el palo colorado (del género sequoia), abetos (noble, blanco, de Douglas), cedro rojo occidental, etc.⁸, se utiliza los mismos términos para abeto y el resto de las coníferas -tepo en Yurok y tš'aamé√3 en Tolowa-, de acuerdo a su parecido con los abetos, aunque esto no puede ser establecido fehacientemente, según Jane y William Bright. De ese modo, los mencionados grupos étnicos, en vez de

⁸ La guía de campo sobre árboles de la Nacional Audubon Society ofrece información sobre la gran variedad de coníferas nativas de esa región.

organizar los árboles en una taxonomía, lo hacen en una suerte de “esferas de influencia” con un centro focal (1969:71).

Tanto árboles como taxonomías están ordenados por relaciones de dos tipos: inclusión y contraste. Las relaciones de inclusión son la que se manifiesten en la membresía de un elemento a una clase que lo incluye, en el sentido que abeto pertenece a la clase conífera. Contraste, es la relación que se da entre dos términos que pertenecen al mismo nivel de inclusión, pero que difieren en alguna dimensión. En los árboles, entonces, sólo puede haber contraste en una dimensión, pero en las taxonomías, puede haberlo en varias.

Si en cambio todas las dimensiones aparecen solamente en un renglón y ningún renglón tiene más de una dimensión, se trata de un paradigma. En este tipo de arreglo existe idealmente un término para cada intersección de dimensiones. Este arreglo fue muy utilizado para representar las terminologías de parentesco, ya que la misma casi nunca presenta relaciones de inclusión como las taxonomías —a lo sumo se encuentra plurales que agrupan parientes como “parentela” o “in-laws”— y en general hay una cobertura de todo el espacio paradigmático dado por las intersecciones de todas las dimensiones entre sí.

Un ejemplo de esto puede ser obtenido de Floyd Lounsbury, quien define paradigma como “(a) el significado de cada forma⁹ tiene un rasgo en común con el significado de todos los otros términos del conjunto y (b) el significado de cada forma difiere de aquel de cada otra forma del conjunto por uno o más rasgos” (1969:193). En el artículo donde está tomada esta definición, “The Structural Analysis of Kinships Semantics”, este autor analiza la terminología de parentesco de los Seneca (Iroquois) utilizando las distintas

⁹ Traduzco aquí “form” como forma, aunque podría entenderse aquí como “término”.

dimensiones en que diferencian un término de otros. Lounsboury define dimensión como un conjunto de rasgos mutuamente exclusivos, que comparten “los mismos privilegios de combinación” (1969:193).

Si se toman en cuenta solamente las variaciones que suceden en las dimensiones sexo y generación y se dejan las demás fijas en un valor, se obtiene una ejemplificación simple aunque representativa. En el primer caso hay dos rasgos, masculino y femenino. En el segundo, los Séneca distinguen cinco rasgos en relación a ego: misma generación, primera ascendente, primera descendente, segunda o más ascendente y segunda o más descendente. Las dimensiones que se fijan en este ejemplo son:

- a. Colateralidad: sólo se toma en cuenta parientes lineales
- b. sexo de la linealidad: sólo patriliniales
- c. sexo del pariente igual o distinto del primer eslabón desde ego: igual
- d. sexo del pariente igual o distinto del pariente: igual
- e. sexo del último eslabón antes que el pariente igual o distinto del de ego: igual

En la figura 3 se puede ver el paradigma representado a través de una llave –del tipo de las que propone Kay- y de un cuadro.

Dimensión “D”. d: todos los parientes patriliniales donde sexo es igual al de a) primer eslabón desde ego; b) del pariente con el que se tienen la relación; c) último eslabón antes que el pariente del que se predica la relación.

Dimensión “G”. g_0 = generación de ego (se omite); g_1 primer generación ascendente desde ego; g_2 = segunda o posterior generación desde ego; g_{-1} = primera generación descendente desde ego; g_{-2} = segunda o posterior generación desde ego.

Dimensión “S”. ♀ y ♂.

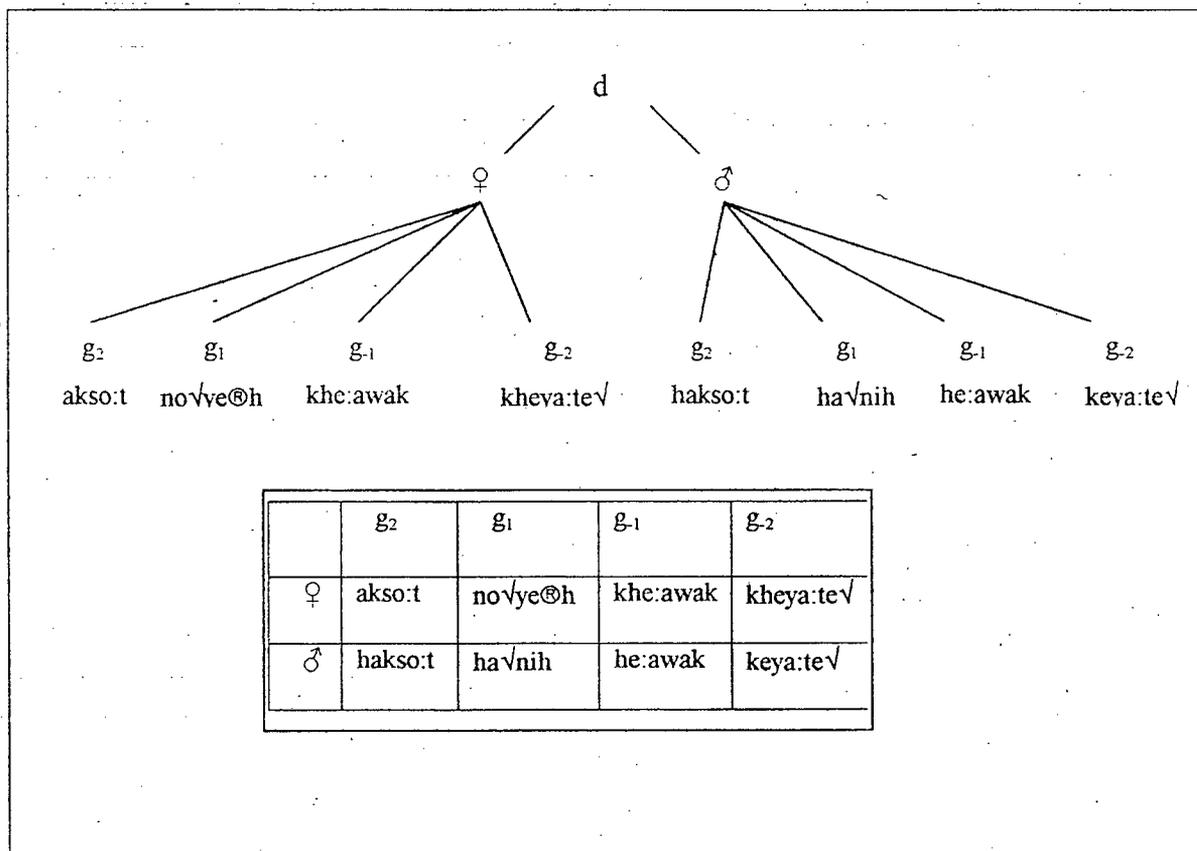


Figura 3: Paradigma de un subconjunto de las relaciones de parentesco Seneca, representado como árbol y en un cuadro

En este tipo de análisis se asumen que son todas las dimensiones de igual peso por las que los términos son discriminados por los hablantes. Es muy difícil que sea así. Por ejemplo, si bien “abuelo” y “prima” poseen valores distintos de la dimensión sexo, ésta no estaría entre las primeras con las que se diferencian. Linealidad y generación serían probablemente más factibles de ser utilizadas por los hablantes como factores primarios de diferenciación. En otras palabras, los paradigmas no contienen información del peso de una dimensión respecto de las demás en el significado de un término dado, ni, obviamente, tampoco en qué circunstancias esta importancia decrece o aumenta.

Este problema, en realidad, es endémico de todas las estructuras formales del análisis componencial, que padecen de la notable ausencia de conceptos cuantitativos, como

podría ser peso o grado. Ni siquiera se analiza la ordinalidad de las dimensiones, de manera de poder establecer aunque más no sea atributos de precedencia y procedencia. Tampoco se utiliza información estadística al punto que en un trabajo clásico de la antropología cognitiva, escrito por Ward Goodenough y llamado "Yankee Kinship Terminology: A Problem in Componential Analysis", no sólo no se utilizó estadística, sino que se contó con un único informante, ¡el propio antropólogo! Esto no le trae ningún resquemor, e incluso afirma que esta cuestión lo "ayudó a descubrir algunos de los criterios para el rechazo de aquello que se presenta siendo modelos componenciales mínimamente adecuados" (1969:256).

En este punto de esta argumentación, ya se expuso lo suficiente como para analizar por qué se hizo necesaria una revolución en la antropología cognitiva:

- a. Mientras se habla o se escribe, no se piensa al mismo tiempo en las dimensiones del significado de las palabras, ni en relaciones de inclusión, contraste e intersección. Los significados se usan de manera automática, por lo que es muy dudoso que inclusión, contraste e intersección formen parte de la realidad psicológica en relación al uso regular de significados.
- b. Entonces, los significados no son necesariamente enunciables a través del lenguaje. Al contrario de lo que creía Fraake, ciertos significados no se adquieren "de" otras personas, sino en el marco de la convivencia con otras personas y el hecho de que pueden en algún punto ser enunciados, no significa que ese sea el modo en que son aprendidos o utilizados.
- c. Cuestiones tales como la ausencia de cardinalidad o de gradaciones en las dimensiones hacen que sean muy bajas las posibilidades de encontrar ejemplos

empíricos que demuestren su existencia, pues en muy pocos casos los hablantes de una lengua o partícipes de una cultura van a clasificar la realidad de manera tan taxativa, sin la posibilidad de gradaciones, priorizaciones de unas dimensiones sobre otras, matices, etc. El caso de los Yurok es prototípico en ese sentido.

d. La cantidad de características salientes para diferenciar dos o más significados que las personas pueden utilizar es limitada. Incluso, por más que al ponerse en comparación distintos términos de parentesco se llegue al nivel de complejidad de las matrices desarrolladas por Goodenough en el mencionado trabajo sobre el parentesco yanqui, no se ha encontrado evidencia de que los usuarios manipulen realmente las 13 dimensiones allí presentadas. Un abordaje estadístico sobre casos hubiera podido iluminar la cuestión, aunque se sabe que para ese artículo, el mismo Goodenough había sido su único informante.

3 La revolución conexionista en la antropología cognitiva

3.1 De los términos focales a la semántica de prototipos

3.1.1 Mecanismos estadísticos

A partir de los años setentas, comenzaron a desarrollarse una serie de nuevas aproximaciones dentro de la antropología cognitiva para, por un lado, defenderse de la tremenda crítica a la que fuera sometida en la segunda mitad de la década del sesenta y, por otro, para poder abordar problemáticas que por distintas razones —entre las que se cuentan las arriba esbozadas— se resistían a ser tratadas con las herramientas del análisis componencial.

Uno de los virajes más interesantes es metodológico y se podría catalogar como la respuesta a la crítica de Marvin Harris, en la que éste señala la falta de mecanismos estadísticos para confirmar los datos (Reynoso, 1998:40). Por esta época comienzan a implementarse con mayor determinación mecanismos estadísticos para determinar la realidad psicológica de los arreglos cognitivistas. Por ejemplo en el estudio sobre la características salientes que conforman las dimensiones que las poblaciones americana y latinoamericanas usan para categorizar las enfermedades, realizado por el D'Andrade y colaboradores en 1972 (D'Andrade 1995). Aunque se sabía ya desde 1960 que no se podía recaer solamente en la destreza de los informantes para determinar la realidad psicológica de un arreglo (Wallace y Atkins 1969:367), solo en muy pocos estudios se apoyaban en mecanismos estadísticos para determinar si un arreglo era real desde el punto de vista psicológico.

Una de las razones para semejante subutilización de tan importante recurso metodológico, es el principio emic de no usar categorías que provengan del ámbito de la cultura del investigador. Lamentablemente, los mecanismos estadísticos fueron inventados en occidente y eso los hace tabú entre quienes quieren conservar su filiación a la ortodoxia emic. Para los no ortodoxos inclusive, esta herramienta puede ser al menos sospechosa. Desde una posición emic se tiene la certeza de que cualquier abordaje estadístico tiene como primer paso la reducción de la riqueza y complejidad del discurso nativo, homogeneizando las respuestas y, por lo tanto, borroneando particularidades y matices, desdibujando las identidades al fusionarlas unas con otras.

Esto es sólo cierto a medias, pues es verdad que algo que se trata estadísticamente ha sido de algún modo reducido a una expresión numérica, escondiendo rasgos y

diferencias. Sin embargo, no es en la metodología sino en el planteamiento del problema en donde se efectúan las operaciones de síntesis y reducción a una expresión tal que es susceptible de ser analizada estadísticamente. Entonces, por más emic que se sea, si se trata de ver cuál de dos arreglos se corresponde mejor con la realidad psicológica de los nativos, la reducción ya fue hecha al momento al formularse el problema. Entonces, ¿por qué no preguntarle a varios informantes y promediar las respuestas? Eso es lo que hicieron Romney y D'Andrade en 1964, cuando usaron a alumnos de secundario para determinar si el modelo de Wallace y Atkins o el de Romney era el que mejor se ajusta a la realidad psicológica de los nativos (Romney y D'Andrade, 1969)

Los mecanismos estadísticos empezaron a ser utilizados también para determinar el énfasis (saliency) puesto en determinadas dimensiones y no en otras, cosa que respondía una parte de la cuestión de la rigidez de los arreglos en el análisis componencial para representar el fenómeno de la importancia diferencial que unas dimensiones tienen en relación con otras, como se veía en el caso de los términos "abuelo" y "prima". Lo que se intenta hacer en este tipo de trabajos, es demostrar que característica se destaca sobre las demás a partir "juicios de similaridad". Por ejemplo, en un estudio de hecho en 1971 por Samuel Fillenbaun y Amnon Rapoport, los 26 entrevistados ranquearon el grado o nivel de disimilitud de todos los posibles pares entre 15 términos colores. La conclusión más importante de ese trabajo fue que, a diferencia de características como "masculino" y "femenino", que se presentan como dicotomías sin gradación, existen dimensiones en las que los distintos términos tienen una suerte de grado en una escala continua que se determina a partir de un análisis estadístico basado en el promedio del "score" de

similitud dado por los entrevistados. Esta técnica se llama "multidimensional scaling". (D'Andrade 1995:64).

Desde el punto de vista teórico, una de las consecuencias más importantes de esta nueva aproximación metodológica es que si ciertas características pueden representarse como ocupando un lugar en el continuum de una dimensión dada, se puede establecer la distancia relativa entre un término y otro, respondiendo a preguntas del tipo ¿"ardilla" se encuentra a la misma distancia de "león" que de "leopardo"? Justamente usando términos de animales, A. K. Romney intentó demostrar la validez psicológica de esta técnica usando preguntas del tipo:

GATO:LEOPARDO::MONO: ____

- A. ANTÍLOPE
- B. NUTRIA
- C. GORILA
- D. TIGRE

De esto da cuenta D'Andrade (D'Andrade 1995:66). Este último sostiene allí que en lo que respecta a la realidad psicológica de las dimensiones, la antropología cognitiva deja de buscar el "análisis componencial de características distintivas" cambiando por "análisis de matrices de marcos de creencia" que son el resultado de "un cambio en tratar de buscar características que podrían ser usadas para definir términos por tratar de buscar las características a las que la gente responde más fuertemente" (D'Andrade 1995:76).

3.1.2 Rangos focal y extendido y semántica de prototipos

Otros conceptos que comienzan a desarrollarse en esta etapa a partir del trabajo de Berlin, Breedlove y Raven (1974) son los de rango de términos básicos o focales y rango

de términos extendido. En el rango básico o focal se encuentran todos los referentes genuinos de una determinada clase, como “vaca” y “mono” en relación a la clase mamíferos. En el rango extendido se encuentran aquellos miembros de la clase que pueden no cumplir con todas las características propias de los miembros básicos, aunque son clasificados todavía en el rango extendido de la misma como, por ejemplo, “avestruz” está en el rango extendido de la clase “ave”. Estos conceptos fueron muy útiles una vez aplicados a la terminología de colores en un trabajo ya clásico de la antropología cognitiva. Allí, a partir de una serie de asunciones sobre qué es un término básico de color y usando un conjunto de 320 tarjetas de colores Munsell, descubrieron una serie de regularidades respecto del uso de la terminología de colores en distintas culturas, como la existencia de focos similares en todos los conjuntos estudiados y la famosa progresión en la que se incorporan los colores a la terminología, que parece seguir la misma secuencia más o menos universalmente.

El problema que resuelve este aporte se suscita cuando se intenta buscar qué es lo que hace que una determinada cosa sea el foco de una clase, en términos de las características distintivas que la componen. Una de las soluciones, propuesta desde la psicología, es la semántica de prototipos, que fue desarrollada a partir del trabajo de Eleanor Rosch (Reynoso, 1998). Allí, los exponentes más representativos de las clases no se definen en tanto un conjunto definido de características sino que, de manera cuasi-gestáldica, se definen a partir de sus ocurrencias más representativas. Estas ocurrencias, si bien son las que cumplen la mayor cantidad de características definitorias de la clase, no son evocadas por ello, sino en tanto ejemplos más claros de pertenencia. En ese sentido, como la misma

Rosch sostiene, las clases no se definen por la ocurrencia de casos marginales, sino por los ejemplos claros, los prototipos de la clase.

3.2 La teoría de los esquemas

La semántica de prototipos es muy rica y se ha omitido una vasta cantidad de cuestiones relativas a la acumulación de características de los términos de nivel superordinado (v.g. mueble), básico (v.g. silla) y subordinado (v.g. silla de cocina)¹⁰. De todos modos hay una serie de conceptos que a partir de aquí comenzaron a ser moneda corriente en la antropología cognitiva:

a. Como se dijo en el capítulo anterior, los significados de los términos no se evocan usando siempre las mismas características diferenciadoras. En algunos casos se usan unas y en otros otras e inclusive a veces no se usa ninguna en particular: la pertenencia a una clase simplemente “viene a la mente”.

b. Los significados de los conjuntos de terminologías para una gran cantidad de campos no se establecen arbitrariamente, sino que la estructura de la realidad es la que determina los focos y prototipos a partir de los cuales las terminologías evolucionan.

c. Se abandonan los arreglos del análisis componencial en cuanto no representan verdaderamente la realidad psicológica de quienes utilizan los términos y comienza una tendencia a pensar que es posible que, cuando se utiliza un término, no se involucren operaciones lógicas como las del tipo intersección, inclusión o contraste.

¹⁰ Dos buena síntesis de esta cuestión se encuentran en las citadas obras de Reynoso “Corrientes...” (1998:65-68) y D’Andrade, “The development of the cognitive Anthorpolgy”(1995:116-121)

O que si se lo hace, esta utilización se circunscribe a casos simples del tipo de la clasificación de plantas o del parentesco (D'Andrade 1995:123) En su lugar, se sostiene que los significados se evocan holísticamente, sin apelar a características específicas, en concordancia con un principio que se podría llamar de economía cognitiva: se tiende a utilizar la forma que requiera menos energía.

d. Como ya no se tiene el impedimento de los rígidos arreglos componenciales ni de una concepción conductista del significado, demasiado ortodoxa y arcaica, es posible abordar los problemas de nuevas perspectivas. El punto interesante aquí es que se abandona la postura antimentalista en la que se hace ciencia de la cultura como si no existiesen lazos causales entre ella y la mente humana.

La propia concepción gestáldica de los prototipos muestra la necesidad de una teoría que explique como es que los seres humanos poseen una tremenda agudeza en el manejo del conocimiento del mundo que los rodea, una tremenda capacidad clasificatoria y conceptual a través de complejos sistemas semánticos, sin utilizar para ello -de manera directa al menos- las características individuales de los objetos que componen la realidad clasificada. La teoría conexionista de los esquemas ofrece tanto los mecanismos como los conceptos necesarios para satisfacer esa necesidad.

3.2.1 Esquemas

Un esquema es una representación abstracta de un conjunto de experiencias que poseen algún tipo de regularidad. La entidad representada en un esquema puede ser simple o de una gran complejidad. La vida cotidiana de cualquier cultura está llena de situaciones en la que se requiere nuestro "conocimiento experto" a través del uso de esquemas: como subirse al colectivo, identificar una situación de peligro, navegar, conducir, buscar

trabajo, etcétera. Ese conocimiento es enunciable sólo parcialmente, pues gran parte de las destrezas que estas actividades requieren no pueden ser enseñadas sólo en teoría, sino más bien deben ser “transmitidas” por otros actores sociales (adultos más experimentados, instructores especializados, etcétera). Esta transmisión se desarrolla no en la forma de una cátedra donde el sujeto aprendiz debe entender y recordar un conjunto de reglas, sino en el contexto de una cierta práctica, donde el instructor hace pequeñas intervenciones –en comparación con el tiempo que el aprendiz practica- cuando considera que el aprendiz debe mejorar algo. Un esquema puede ser el resultado de experiencias mucho más complejas que el aprendizaje del significado de un determinado término o conjunto de términos.

D’Andrade, en general intenta darle una continuidad a los mecanismos de formalización que la antropología cognitiva fue desarrollando desde los ’50, considerando a los cambios que se fueron sucediendo como intentos de aplicar los dispositivos a campos más complejos. Así es que sostiene que los prototipos son instancias de esquemas, en donde los lugares vacíos son llenados por variables tomadas de un individuo “estándar” y representativo de toda la clase. (D’Andrade, 1995:124). El resultado de ese intento de D’Andrade es una antropología cognitiva donde coexisten paradigmas, taxonomías, prototipos y esquemas aplicándose a problemas de distintos niveles de complejidad (D’Andrade, 1995). Sin embargo como se vio más arriba, los primeros arreglos propuestos no solamente se aplican a problemas menos complejos sino que también a campos realmente muy acotados, y su realidad psicológica es dudosa en el mejor de los casos. Los problemas planteados hasta parecen de cierta manera triviales, sustentados por una metodología realmente pobre, donde la estadística juega un papel secundario y el

concepto conductista de significado no permite nutrirse del conocimiento psicológico proveniente de escuelas mentalistas.

El único arreglo que se podría decir que resistió fue la taxonomía, aunque con un gran esfuerzo de re-conceptualización, abandonando las características como atributos definitorios. Incluso la teoría de los niveles de profundidad de las taxonomías folk de los ya citados Berlin, Breedlove y Raven ha sido altamente criticada (D'Andrade, 1995:101), aunque ha subsistido la idea que la cantidad de niveles propuesta corresponde verdaderamente a la capacidad limitada de manipulación de memoria a corto plazo.

Los esquemas culturales tienen la particularidad de ser compartidos por un grupo de individuos en un determinado tiempo y lugar. Se pueden pensar también como la manera en que se codifica y almacena en nuestra mente el conocimiento cultural. Por ejemplo, si nunca se participó de un ritual y luego se va a uno, probablemente se recuerde luego este hecho como una situación única, llena de curiosidades. Incluso es posible que se recuerden cuestiones completamente intrascendentes, como algún tipo extraño de vestimenta, alguna cosa repulsiva o placentera, etcétera. Pero si se participa a menudo de rituales en la misma cultura, muy posiblemente se recuerde cosas completamente distintas, como aquello que no sucede normalmente en ese tipo de ritual, la presencia de extraños o gente nueva. Esto suena bastante coherente. Pero lo interesante es que el asistente habitual recuerda muchos más detalles que los que recuerda el primerizo. La razón, es que tiene un esquema formado de dicho evento y solo hace falta llenarlo con las particularidades del caso específico. A través de la experiencia cultural, el habitué posee algo así como una representación mental que le permite entender y manejarse en el

evento, al mismo tiempo que entender y recordar la instancia particular del evento con mayor facilidad.

Algo similar se puede decir del mito. Citando a Edwin Hutchins:

“The myth of the Baroweni contains a structure of relationships among things and actors and actions. If we ignore the specific identities of the things and actors and actions and concentrate in the relationship among them, we see a structure I call a schema.” (1987:279)

Los esquemas son una entidad dinámica por lo menos por dos razones. Por un lado, la realidad a la que se aplica es dinámica por lo que se puede inferir que, o bien la entidad representacional lo es o bien posee las operaciones para representar estáticamente ese dinamismo, siendo así potencialmente dinámica. Por otro lado, para llenar los espacios vacíos de la estructura abstracta del esquema con el estímulo de la realidad y transformarlo (en conocimiento útil sobre un hecho particular, en un relato, una respuesta, una decisión, etcétera) se requiere un proceso.

3.2.2 Imagen-esquemas y proposición-esquemas

Se pueden distinguir dos tipos de esquemas: los imagen-esquemas y proposición-esquemas, siendo estas las formas en que el conocimiento se puede moldear (Holland y Quinn, 1987). Se podría entender la proposición-esquema como un conjunto de ideas o conceptos y sus relaciones. Además, los esquemas poseen la capacidad de ser recursivos: un esquema puede estar compuesto por la relación entre dos conceptos, donde las dos entidades (relación y conceptos) son a su vez esquemas. Considérese la siguiente proposición-esquema sobre se conciben las emociones, investigados por Catherine Lutz en el atolón de Ifaluk situado en el océano Pacífico occidental (Lutz, 1987):

Si evento X, entonces experimento emoción Y

Si experimento emoción X, entonces puedo realizar acto Y

Si experimento emoción X, alguien debe estar experimentando emoción Y

Si experimento emoción X, puede que luego experimente emoción Y

Si se completan las variables X e Y, se obtienen instancias de las proposiciones-esquemas:

Si espíritu presente, entonces experimento miedo

Si experimento felicidad, entonces hablo mucho

Si experimento ira justificable, alguien debe estar experimentado miedo

Si experimento vergüenza, puede que luego experimente ira

Los imagen-esquemas, en cambio, son más simples y de alguna manera asociados a construcciones visuales. No son imágenes en sí ya que tienen también, en tanto esquemas, la capacidad de albergar huecos reservados para variables. Por otro lado, pueden contener información cinética, posicional, en relación a proporciones, estética, etcétera y de todas las propiedades que se puedan asignar a una imagen. Algunos autores (Quinn, 1987; Lakoff 1987) sostienen que hay también una estrecha relación entre los imagen-esquemas y la metáfora. Para ellos, la metáfora es una suerte de mecanismo para generar imagen-esquemas dado que cuando se representa la imagen de una cosa con la imagen de otra, se esta substrayendo un esquema de esta última. El esquema substraído contiene las propiedades de una imagen usadas para representar ciertas propiedades de la otra imagen que así es representada.

Lakoff y Kövecses desarrollaron un interesante análisis de la ira usando este concepto de imagen-esquema (Lakoff y Kövecses, 1987) como proveedor del marco conceptual

para entender ciertos usos metafóricos. Ellos encontraron en el idioma inglés una gran cantidad de formas de referirse a la ira tomando imágenes de otros dominios, que luego se transforman en distintas expresiones o giros idiomáticos. Por ejemplo, una imagen que analizaron es "Heat of fluid in a container". A partir de esta imagen-esquema y sus derivaciones, se construyen en inglés una gran cantidad de giros idiomáticos. Estos son algunos de ellos:

Derivación: When the intensity of the anger increases, the fluid raises.

Giros: His pent-up anger *welled up* inside him.

She could feel her *gorge rising*.

We got a *rise* out of him.

Derivación: When the anger of a person becomes too intense, the person explodes.

Giros: When I told him, he just *exploded*.

She *blew up* at me.

We won't tolerate any more of your *outburst*.

Hasta aquí se vio como la teoría de los esquemas ofrecen un marco para desarrollar una gran variedad de problemas. Sin embargo, hay una pregunta que el lector crítico puede hacerse: ¿cómo se podría representar formalmente estos esquemas? Este trabajo comenzó mostrando como la visión emic del estudio de la conducta humana es acompañada, tanto en Pike como en la antropología cognitiva, por la necesidad de implementar mecanismos formales. En el caso de Pike se vio que esto sólo se dio de manera conjetural, sin producir una metodología específica que considere la recolección de datos para satisfacer arreglos formales. En la antropología cognitiva de los '60 los arreglos formales fallaban en cuanto a su realidad psicológica, entre otras cosas.

Lo que resulta atractivo de la teoría de los esquemas es que podrían implementarse a través de RU, por lo que la respuesta que pregunta que inició este párrafo se encuentra en la corriente conexionista la inteligencia artificial, en última instancia generadora de la teoría de los esquemas.

3.3 Conexionismo

3.3.1 Esquemas en la inteligencia Artificial

El concepto de esquema es de larga data en las ciencias humanísticas y fue usado, entre otros pensadores, por Kant, Bartlett y Piaget, pero hasta los '70 fue en general rechazado en la psicología experimental por considerarse demasiado vago (Rhumelhart, Smolensky, McClelland y Hinton, 1986). A partir de los '70 comienza a aparecer la necesidad de una "estructura conceptual para representar las relaciones complejas implícitas en nuestra base de conocimiento" (1986:18). Al mismo tiempo, muchos científicos dedicados a la inteligencia artificial comienzan a tratar de idear algún concepto que ayude a explicar esas relaciones complejas. Por ejemplo Marvin Minsky, fundador de una de los principales laboratorios de inteligencia artificial, desarrolla la idea de "marco" (frame), el cual es "un formulario con muchos espacios en blanco para ser llenados. Llamamos a esos espacios *terminales*; los usamos como puntos de conexión a los que podemos adjuntar otras clases de información" (Minsky, 1985:245).

Sin embargo, la escuela de los partidarios del procesamiento simbólico –a la cual Minsky adscribe– se enfrentan siempre con una cuestión que es muy difícil de demostrar: que la mente humana posee algo así como un lenguaje interno, que incluye reglas y propiedades lógicas y que almacene la información codificada. Por un lado, está la cuestión de la capacidad de almacenamiento, que por lo que se sabe no sería suficiente.

Pero aún suponiendo que sea posible almacenar toda esa información, sería aún más difícil manipularla con la velocidad con que los seres humanos procesan información. Considérese las aseveraciones de Schwartz, que se expusieron más arriba. Allí se afirmaba que no es la velocidad la característica destacable, sino su capacidad de procesar en forma paralela. Por lo tanto, las estrategias de simulación del pensamiento a través de programas secuenciales estructurados por algún lenguaje (de ese se trata la programación simbólica), sea cual sea aquel, tienen pocas posibilidades de lograr la versatilidad que la teoría de los esquemas requiere.

Los dispositivos que se desarrolla en el marco de los proyectos conexionistas, en cambio, sí poseen ese nivel de versatilidad debido a que, a diferencia de la otra corriente, el estímulo no debe ser convertido a ningún lenguaje para ser almacenado/procesado. Como se vio en la red XOR de la introducción, la capacidad de discernir entre distintas combinaciones de "1" y "0", no es implementada a través de regla alguna sino que la misma es consecuencia de conectar las unidades de un cierto modo en especial. Es así que ninguna de las unidades posee un conocimiento específico, sino que éste último es un emergente de la red, en la que se encuentra distribuido más o menos uniformemente.

Probablemente sea esto último lo que sedujo a los antropólogos cognitivistas, la capacidad de las RU de ser una mejor herramienta para "simular como la gente construye esquemas culturales y los usa para extraer significado de ellos" (Strauss y Quinn, 1997). Durante los '70 fue más o menos aceptado dentro de la antropología cognitiva que gran parte del conocimiento sobre el mundo se desarrolla a partir de encontrar focos, prototipos o características destacadas ya existentes en la realidad, como si el aparato cognitivo humano se moldease con la realidad circundante. Las RU justamente se

moldean para resolver cierto problema específico. Los pesos entre unidades actúan como la "pasta" del molde, en tanto se modifican de acuerdo a las contricciones impuestas por el estímulo que se les envía.

También atrajo la atención de la antropología cognitiva el hecho de que, al ser un dispositivo que extrae conocimiento de las regularidades del entorno sin especificar ese conocimiento en ninguna parte, cuando las RU cumplen cierta tarea, lo hacen automáticamente: entre el estímulo y la respuesta no median reglas. Para Strauss y Quinn, a diferencia de los programas simbólicos, las RU tratan a las tareas como "...automáticas de compleción de patrones" (1997:58). Esta propiedad de distributividad de la información en la RU es de algún modo tomada del funcionamiento de cerebro. La investigación neurológica ha demostrado que no hay una parte del cortex que sea indispensable para el funcionamiento del cerebro, sino que "todas las partes funcionan juntas, influenciándose una a la otra, y cada región contribuye al desempeño general de la tarea y a la integración dentro de esta de ciertas clases de constricciones y de fuentes de información" (Rumelhart y Macclelland, 1986:134).

La idea de almacenamiento distribuido puede resultar antintuitiva en el sentido de que la gente está acostumbrada a mecanismos de búsquedas secuenciales, como los índices de los libros, los catálogos de las bibliotecas, los ficheros, etcétera. El significado de estas mismísimas palabras que se están leyendo, se encuentra almacenado en algún lugar, pero está claro que se aparecen de forma "automática". Para esclarecer esta cuestión, puede que sirva el siguiente ejemplo de almacenamiento distribuido desarrollado por Hinton, Macclelland y Rumelhart en el artículo "Distributed Representations" (1986).

Supóngase que existe un microlenguaje derivado del inglés en el que sólo existen palabras de tres letras que empiezan con w o l, siguen por i o e y terminan en g o r. En ese lenguaje wig y leg son válidas palabras y weg y lig no lo son. La forma que estos autores proponen para implementar una red que sepa qué palabras son válidas, es lo que ellos llaman “codificación conjuntiva” (conjunctive encoding), que consiste en tener una unidad para cada combinación de letras válida. Cada letra envía el valor 1 cuando es activada. Las unidades intermedias y las de salida, se activarán positivamente cuando la reciban 1.5 o más en total. Para eso es necesario 2 letras activadas, pues, como se dijo, cada una envía el valor 1.

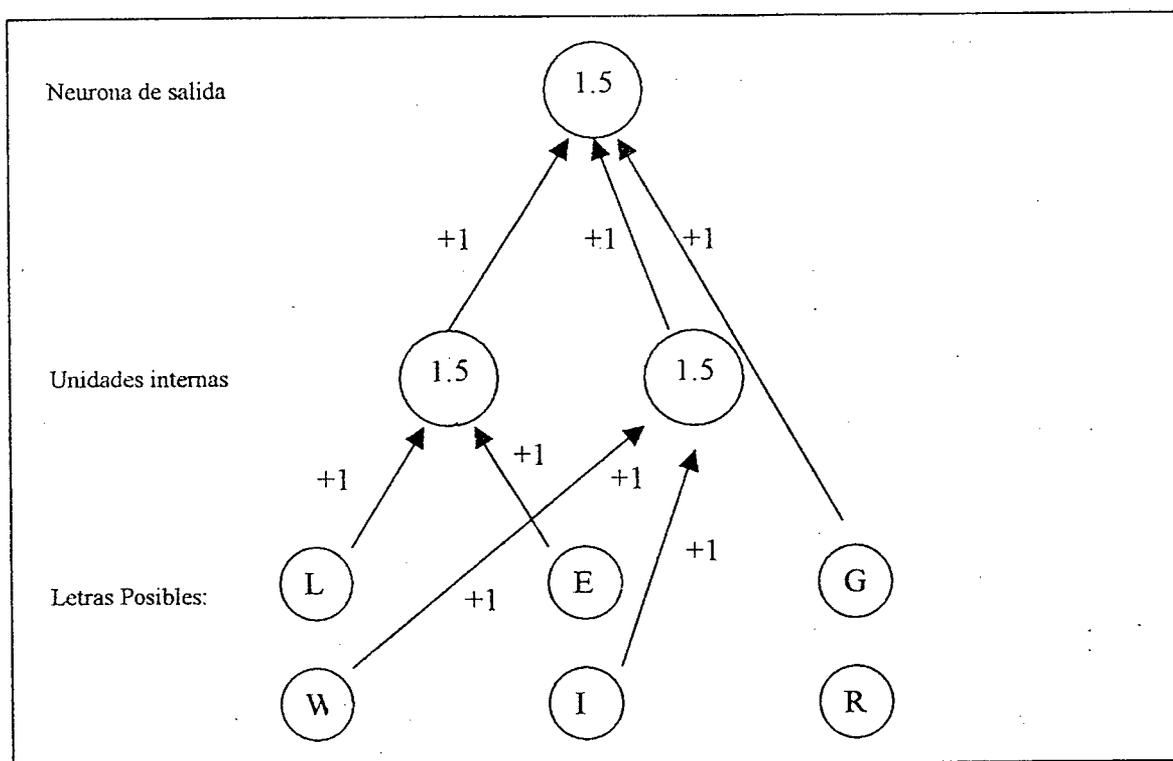


Figura 4: Representación de una RU que distribuye información a través de conjuntividad

La técnica de “codificación conjuntiva” consiste simplemente en tener una unidad por cada combinación posible. En este caso, como las palabras posibles son 2, hay dos

unidades intermedias que reciben estímulo de los pare “LE”y “WI”, que son los dos principios de palabra posibles en este lenguaje. Como ninguna palabra termina con r, esta no tiene conexión y como todas terminan con g, ésta se conecta directamente a la unidad de salida.

Como puede apreciarse, no hay una relación lógica entre la topología de la RU y el problema que se resuelve. Este es el tipo de cosas que impresionaron a los antropólogos cognitivistas en principio, dado que si es posible expandir el nivel de complejidad indefinidamente, sería factible representar formalmente a los esquemas de imágenes y proposicionales a través de este tipo de redes.

3.4 Esquemas y redes de satisfacción de constricciones

3.4.1 RU como redes de satisfacción de constricciones

Una de las maneras para caracterizar a la mayor parte de las RU que se proponen desde el grupo de PDP, es a través de considerarlas como mecanismos que satisfacen las constricciones que son impuestas por la interacción con el entorno. Esto se encuentra en el trabajo dedicado a los esquemas titulado “Schemata and sequential thought”. Allí, cada unidad “representa una hipótesis de algún tipo (por ejemplo, que cierta característica semántica, visual o acústica se encuentra presente en el estímulo) y en la cual cada conexión representa constricciones entre las hipótesis” (Rumelhart, Smolensky, McLelland y Hinton, 1986:9). Las redes desarrolladas en el contexto de alguno de los 4 paradigmas mencionados en la Introducción, tienden a asentarse a través del entrenamiento –supervisado o no- en un estado tal donde se satisfacen la mayor cantidad de constricciones posibles. Podría verse a los mecanismos de entrenamiento como la búsqueda de un “promedio” entre un vector de pesos que satisfacen una conjunto de

constricciones dados por los ejemplos pertenecientes a una clase y el vector de los pertenecientes a otra. Esta manera de verlo es superficial, ya que no es exactamente un promedio o una media estadística lo que el entrenamiento busca. No sólo eso, sino que ni siquiera se trata de un proceso lineal: varios resultados —es decir, configuraciones de pesos— pueden satisfacer el mismo conjunto dado de constricciones.

La cuestión de la linealidad fue fuente de polémicas y no fue hasta que comenzaron a desarrollarse redes con algún tipo de no-linealidad que el conexionismo comenzó a tomar impulso. Por otro lado, la debilidad de las RU de dos niveles, uno de entrada y otro de salida, es que no hay ningún problema que pueda ser aprendido en varios pasos de corrección de pesos, que no haya otra RU que no lo pueda hacer en un solo paso, por lo que la retroalimentación no sería indispensable (Rumelhart, Hinton y McClelland, 1986:62). La solución más simple que estos autores proponen es el uso "unidades de umbral lineal" (linear threshold units) que tiene en la red XOR vista en la Introducción a uno de los ejemplos más simples.

Otra debilidad de las redes de dos niveles, cuestión que fue señalada por Minsky y Pappert (1967), es que si el estímulo es muy distinto a la respuesta esperada, no es posible encontrar un conjunto de pesos desde el nivel de ingreso al de salida que represente la relación. Estos mismos autores demuestran también que si hay un nivel intermedio oculto de unidades, estas pueden efectuar el mapeo correspondiente para convertir cualquier patrón de ingreso en cualquiera de salida. Lo que sucede es que el nivel intermedio recodifica el patrón de ingreso a través de sus conexiones y de los valores que tienen las unidades, constituyendo lo que se llama "representación interna" del patrón ingresado. Minsky y Pappert muestran en el mismo lugar que existe una regla

simple de aprendizaje que puede utilizarse para todas las RU de dos niveles, que es $\Delta w_{ij} = \eta(t_i - a_i)a_j$ ¹¹, aunque no hay tal cosa para las RU de muchos niveles. Sin embargo, el hecho de que no exista una regla general, no impide que se hayan desarrollado distintos algoritmos o mecanismos, como señalan Rumelhart, Hinton y Williams (1986:321). Por ejemplo, el aprendizaje competitivo (Rumelhart y Zipser, 1986), o la asunción de una representación interna inicial que parezca a priori razonable.

Entonces, usando distintos algoritmos, las RU tienden a asentarse en un “estado local óptimo” (Rumelhart, Smolensky, McLelland y Hinton, 1986:9) en el cual se satisfacen la mayor cantidad de constricciones posibles, estado que se llama de “relajación”. También se suele decir que la RU, “se asentó en una interpretación” (settle in an interpretation).

El ejemplo que usan estos autores es una red de 16 unidades que sirve para llegar a una interpretación válida del cubo de Hecker. Como se sabe, en el cubo de Hecker cada vértice puede tener dos interpretaciones: o bien está en la cara delantera, o bien en la trasera. Las 16 unidades se dividen en dos grupos, cada uno representando la posición de cada vértice en una de las dos interpretaciones (ver figura 5). De esa manera, cada vértice es representado en una subred estando en la cara de adelante y en otra estando en la de atrás. Como cada cara tiene cuatro vértices, de ser cierta la hipótesis que cierto vértice está en la cara delantera, las hipótesis de los otros tres vértices que están en la misma cara

¹¹ Según Rumelhart, Hinton y McLelland (1986), esta fórmula se entiende del siguiente modo: 1) Los pesos se cambian sólo cuando el valor de input está prendido ($a_i = 1$); 2) si la RU es correcta en la unidad de entrada i ($t_i = a_i$), no se hacen cambios en los pesos; 3) si la unidad de salida j responde 0 cuando debiera ser 1, incrementar todos los pesos por el valor η y 4) si la unidad de salida j responde 1 cuando debiera ser 0, decrementar todos los valores por el valor η .

es reforzada. Por otro lado, si la hipótesis de alguna de las unidades de la otra subred que representan al mismo vértice y al opuesto es verdad, la unidad recibe una carga inhibitoria. Resumiendo, cada unidad tiene tres cargas positivas, dos negativas y el estímulo externo, que, de existir, es siempre positivo. También, las dos cargas negativas equivalen a las tres positivas de modo que si una unidad las recibe, el estímulo externo decide.

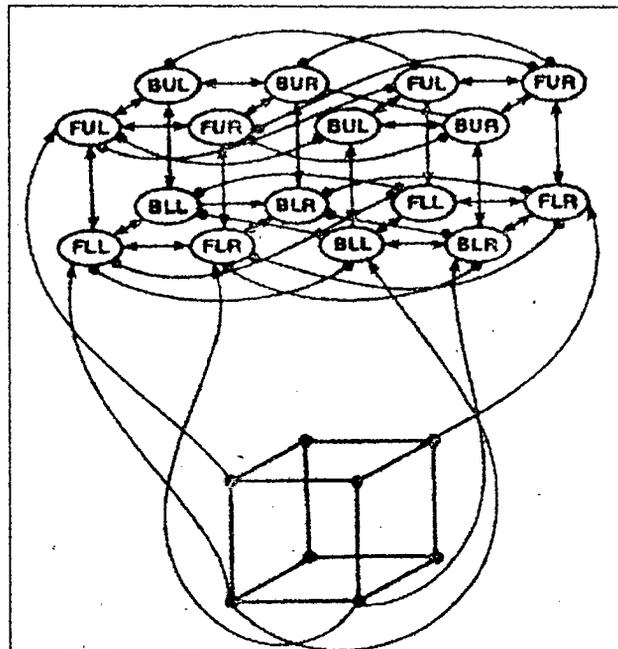


Figura 5: Cubo de Hecker con las dos subredes conectadas a sus vértices. F es forward, U upper, R right, L left y lower, así, BLL significa que esa unidad tiene la hipótesis que el vértice al que está conectada (abajo, izquierda) es el de atrás

Esta red se empieza a entrenar con todas las unidades en 0 y se van actualizando de a una unidad en cada paso. A medida que más unidades se van actualizando, el sistema tiende casi siempre a estabilizarse en una de las dos interpretaciones del cubo. El efecto a través del cual se llega a este estado se llama "trepada de colina" (hill climbing) y lo que los autores mencionados usan para graficarlo es un cubo y una superficie irregular, como se muestra en la figura 6. Como las hipótesis son a) que se interpreta con la cara izquierda adelante y la derecha atrás o b) la derecha adelante y la izquierda atrás, la mayor

correspondencia se da cuando todas las unidades de una de las dos subredes están activadas, que sería el caso de los dos picos que se pueden diferenciar a ambos lados del cubo. Cuando el sistema llega a alguno de los picos, el mismo se encuentra en un estado estable, y se ha asentado en una solución que sucesivos ciclos de entrenamiento no mejorarán ni empeorarán.

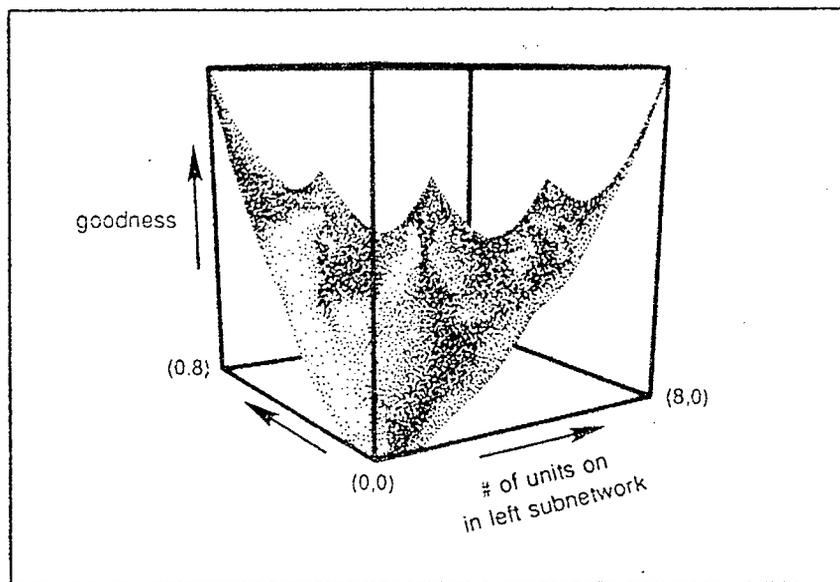


Figura 6: Los picos de la superficie representan las dos interpretaciones del cubo de Hecker. los ejes X y Z representan las unidades activadas; el eje Y representa la bondad de ajuste¹² de la hipótesis con los datos resultado de los ciclos de entrenamiento

Una cuestión cuya importancia Rumelhart, Smolensky, McClelland y Hinton (1986:14) no dejan de resaltar, es que los ajustes del valor de las unidades es un operación local, es decir, que cada unidad no requiere información del estado global del sistema para determinar su cambio de peso, sino que sólo se conecta, como ya se vio, con otras 5 unidades y un vértice del cubo, o siendo esto un conjunto acotado de constricciones que sólo actúan localmente.

¹² En el paper de los autores no se aclara que mecanismo de cálculo de bondad de ajuste se utiliza

3.4.2 Se cierra el círculo

Los mismos autores piensan que las definiciones de esquema disponibles hasta los setentas han sido demasiado vagas o con representaciones que no eran lo suficientemente activas (1986:19). Una de las características que atribuyen a los esquemas y que nunca fue incluida en las distintas implementaciones del concepto, es aquella que “involucra la naturaleza de las constricciones de variables y los valores predeterminados” (default values) (1986:19). Las constricciones, como se vio en el caso del cubo de Hecker, determinan que una cierta unidad adquiera un valor. Puede haber constricciones que no asignan valor, sino que determinan un rango de valores posibles. Inclusive, agregan los autores, puede haber interdependencias entre los distintos “slots”, en el sentido que si uno adquiere un valor a través del estímulo externo, influencia a su vez los valores posibles de otros slots. El problema de la falta de implementación de esta idea en las otras teorías de esquemas¹³, es que eran demasiado rígidas para concebir un nivel de flexibilidad que permitiera al concepto de esquema ajustarse a cualquier situación.

En este punto de la argumentación, los autores llegan a la conclusión que no existe una representación física de los esquemas, sino que estos son el emergente de un conjunto de elementos trabajando juntos. Y como son un emergente de una serie de constricciones locales, no existen en tanto entidad global hasta que no se produce la activación de las unidades. Se podría describir el proceso de existencia de los esquemas en los siguientes pasos:

¹³ A parte del mencionado Minsky, también discuten la idea de esquema de Kant, Piaget, Bartlett, Schank y Abelson, que acuñaron la noción de “script”.

1. Ingresar un estímulo al sistema. Se activan las unidades de "entrada" (la "retina" del sistema)
2. Estas unidades se conectan entre sí y con otras unidades internas formando la red de satisfacción de constricciones
3. El sistema se mueve hacia el estado estable
4. Cuando se llega a ese estado, el sistema tiende hacia otro punto de estabilidad, probablemente usando como estímulo el propio estado al que acaba de llegar, más información externa

Los esquemas más simples se combinan para formar esquemas más complejos. El esquema "Árbol", está compuesto por los esquemas "hoja", "tronco", "rama", "raíz", etc. Por otro lado, de acuerdo al tipo de hojas, es posible que se logre el estado estable activando las unidades de "conífera" o de "roble". Por supuesto, si se activa conífera, el valor predeterminado para forma de hoja es "aguja".

En definitiva, los esquemas no existen hasta que se producen a partir del efecto de las constricciones locales impuestas a un conjunto de unidades en un proceso que se podría llamar de instanciación. Si bien este término puede ser engañoso, en el sentido que una instancia podría suponer la existencia de una clase, los esquemas no están enunciados, ni especificados fuera de estas instancias. Por lo tanto, aunque se pueden identificar relaciones causales, estas son siempre locales y *no es posible inferir el esquema de las ellas*. El comportamiento de las RU es un emergente que no puede ser explicado por el análisis de sus partes separadamente. Prueba de ellos es que un mismo conjunto de unidades puede instanciar más de un esquema. La diferencia está en el valor con que se llenan las variables a partir de las constricciones definidas por los pesos y los estímulos

tanto externos como provenientes de otras unidades. Entonces, no es posible inferir esquemas a partir del conocimiento de la estructura interna de las RU por más profundo que éste sea.

Esta conclusión a la que llega el grupo del PDP, es poderosamente parecida a la semántica de prototipos, pues allí tampoco una clase puede ser enunciada en términos de una lista de características, sino a través de una instancia concreta y representativa. También, se podría conjeturar, en el universo de los esquemas no hay abstracciones porque no son necesarias. En ese universo, no hace falta saber que las coníferas son fanerógamas gimnospermas, que tienen las hojas en forma de aguja y forma cónica. Sólo hace falta almacenar un conjunto de unidades que sean afectadas por constricciones de modo que cuando se recibe el estímulo producido por un pino, se llega a un estado estable instanciando el esquema "conífera".

A partir de los '80, los esquemas se tornan moneda corriente en la antropología cognitiva y se comienzan a compartir otras ideas, como los modelos mentales, que para los conexionistas del PDP son redes de relajación que "toman como estímulo alguna especificación de las acciones que se intentan llevar a cabo y producen como una interpretación de 'qué pasaría si hacemos eso'" (Rumelhart, Smolensky, McLelland y Hinton, 1986:41). De algún modo permiten simular lo que sucede en el mundo exterior en nuestra mente para ayudarnos a tomar decisiones. La definición que ofrece D'Andrade es muy parecida, ya que para él, "uno usa un modelo para razonar con o calcular a través de manipular mentalmente las partes de un modelo en pos de resolver algún problema" (D'Andrade, 1995). Por cierto, en los modelos, tanto partes constitutivas como relaciones entre partes son esquemas.

En este punto, los objetivos de este trabajo expuestos en la Introducción, obligan a hacer una serie de preguntas:

- a. ¿cuáles son las consecuencias metodológicas concretas de teoría de los esquemas relativas al uso de las RU?
- b. ¿Se han podido verdaderamente implementar dispositivos conexionistas que tengan consecuencias observacionales aceptables?
- c. ¿Cuáles son los obstáculos para el uso antropológico de esta herramienta?

4 Dos experiencias de RU aplicadas a la antropología

4.1 Límites metodológicos de la teoría de los esquemas y las RU

4.1.1 Modelos del mundo y RU

Si bien alguien podría verse tentado a pensar que lo que se está haciendo con las RU son modelos del cerebro, o de la mente, o de los fenómenos físicos que subyacen a los procesos cognitivos, hay una gran cantidad de diferencias y, aunque no se discute que la mente funcione a través de mecanismos masivamente paralelos, lejos se está de entender a fondo el funcionamiento del cerebro. En ese sentido, las RU sólo pueden brindarnos una conjetura o hipótesis que no puede confirmarse por ahora. De cualquier manera, hay otra cuestión no resuelta: cómo los esquemas se interrelacionan entre sí para formar esquemas más complejos y modelos del tipo de los analizados en la ciencias cognitivas.

En el artículo ya citado "Schemata ..." Rumelhart y compañía sostienen que el modelo mental es en realidad otra red de relajación que actúa, por supuesto, en paralelo a la red de interpretación. La primera se nutre de la segunda mientras se actúa, pero a su vez la segunda sirve de estímulo a la primera a la hora de actuar. Ellos desarrollan un interesante ejemplo de una red que juega al tatetí, simulando las acciones del otro jugador. Esta simulación no es explícita, sino que se desarrolla a partir de constricciones hechas a la red que toma las decisiones del juego desde otra dedicada a representar las acciones del contrincante. Las personas usan estos modelos para resolver cuestiones complejas que de otro modo requerirían usar recursos lógicos muy difíciles, sólo solubles en lapsos de tiempo considerables para la mayoría.

En ese sentido, el fundamento último de sus ideas sobre esquemas y modelos son tres hechos que los autores creen que tornan necesaria una teoría como la que se está analizando. Primero, las personas son buenas en la búsqueda de patrones concordantes, incluso en relación a problemas lógicos complejos. Segundo, son buenas en modelar el mundo y anticiparse a los hechos, de manera natural y automática; y tercero, son buenas al manipular el entorno. Un caso que demuestra esto es el experimento desarrollado por Wason, un psicólogo citado por D'Andrade (1995:200). El experimento consiste en lo siguiente. Una fábrica de etiquetas produce unas que tienen impreso las letras "A" o "E" en el frente y los números 2 o 3 en el dorso. Hay una condición que se tienen que cumplir: si la etiqueta tiene una "E" en el frente, tiene que tener un 2 en el dorso. A los interrogados se les mostró 4 etiquetas de un solo lado, una "A", una "E", un "2" y un "3". Luego se les preguntó cuál darían vuelta para controlar si la regla se cumple. El 80% de los entrevistados falló y, en vez de elegir la "E" y la "3", elegían casi siempre la "E" y la "2".

Lo interesante de este caso, según descubrieron Jonson-Lairs, Legrenzi y Legrenzi en 1972 (D'Andrade, 1995:200) es que si en vez de letras abstractas se usan las etiquetas representando el mismo problema de manera realista, el 90% daba la respuesta correcta. La regla que usaron era del tipo: si el precio es mayor a 100, requiere a la aprobación del supervisor. Dos etiquetas tenían un precio (por ejemplo, 10 y 120), la tercera decía aprobado y tenía una firma y la cuarta también decía aprobado, pero sin firma. La respuesta correcta es que se da vuelta la que dice 120 y la que no tienen firma. D'Andrade sostiene que esto se debe a que las personas tienen problemas para representar de manera abstracta el modus tollens (Si p, entonces q; no q; luego, no p), e

imaginarse que pasaría si el consecuente no existiera. En cambio, si la relación antecedente-consecuente forma parte de un esquema conocido, los humanos la resuelven naturalmente, sin apelar a la compleja lógica subyacente.

Ambas concepciones de los modelos mentales, la de D'Andrade y del grupo de PDP, están afirmando que inclusive el pensamiento lógico, o al menos algunas partes de éste, opera a través de esquemas y modelos mentales.

En la antropología se ven proliferar una gran cantidad de trabajos buscando en distintos ámbitos imágenes-esquemas, proposiciones-esquemas, modelos mentales de distinto tipo. Sin embargo, no se va a pasar al desarrollo de redes masivamente.

4.1.2 Dos obstáculos metodológicos

Hay dos cuestiones que explican en gran parte la falta de ejemplos de ese tipo. Por un lado, no se han desarrollado todavía mecanismos para interrelacionar las RU implementado aquello que los esquemas prometen, que es la capacidad de incrustarse unos con otros para formar esquemas más complejos y, luego, modelos acerca de las cosas mundo. Si bien se puede vincular dos o más redes entre sí, esto sólo parece ser posible cuando el patrón de salida de una red no requiere la ayuda de "etiquetas". Las etiquetas son, por ejemplo, la manera que se asignaba letras a unidades en el ejemplo de formación de palabras con tres letras mostrado más arriba. En otras palabras, si para darle sentido a determinado esquema representado por una RU se requiere etiquetar, entonces, se pierde la capacidad de vincularlo con otra RU, dado que es imposible pasar la información semántica incluida en las etiquetas. Las RU sólo entienden patrones de entrada y salida, números en definitiva.

El segundo escollo está relacionado con las diferencias entre el marco epistemológico en que la herramienta se intenta utilizar y en el que está preparada para usarse. Tomando la tipificación que Reynoso sostiene en una obra inédita¹⁴, los modelos que usan las distintas ciencias para explicar o interpretar la realidad se dividen en cuatro: mecánicos, estadísticos, sistémicos e interpretativos. Los modelos mecánicos son aquellos que intentan explicar la lógica interna de un proceso a través de relaciones causales lineales. En ese sentido, la inferencia que desarrollan es analítico deductiva y su fin último, la explicación. Los modelos estadísticos, en cambio, no intentan explicar, sino correlacionar fenómenos cuya causalidad no se puede conocer o no es conocida en principio. Los modelos sistémicos, en las palabras de Reynoso, “procuran organizar la complejidad a través de un conjunto de ecuaciones que describen los diferentes aspectos de los sistemas” (ined:14). En los modelos de los sistemas complejos, en vez de establecerse correlaciones entre fenómenos o relaciones causales lineales, se desarrollan algoritmos, ecuaciones no lineales, grafos topológicos o diagramas de flujo.

Es fácil caer en la tentación de tratar a la representación interna de la RU como un mecanismo de “programación”, donde se asignan manualmente los pesos de la representación interna de los patrones a ingresar para obtener un resultado deseado, manejando en ese sentido la causalidad del modelo. Se podrá alegar que, hipotéticamente, es posible lograr a través del entrenamiento esa misma configuración de pesos, en tanto se logre el conjunto de ejemplos adecuado, por lo que es lícito manipular la representación interna de los patrones.

¹⁴ La obra se titula “Antropología de la complejidad” y se editaría en el 2004.

No obstante, las redes de relajación o redes de satisfacción de constricciones —que son la mayoría de las RU— se entrenan a través de operaciones estadísticas. ¿Por qué, entonces, se habría de pensar que la correlación lograda con ese mecanismo estadístico es la lógica con la que funciona la cognición? ¿No debería considerarse tan sólo como modelo estadístico de correlación entre una población de patrones y una respuesta que consiste en agruparlos de acuerdo a cierta hipótesis clasificatoria? A través de “programar” los pesos, se está implementando algo distinto, o por lo menos una forma diferente de causalidad, donde cierta distribución de pesos “causa” la respuesta esperada.

Si se toma en cuenta aquello que se dijo más arriba acerca de la falta de conocimiento respecto al funcionamiento del cerebro, ¿qué estaría diciendo sobre la cognición humana una RU diseñada así? Por ahora, no es posible establecer relaciones causales lineales entre el mundo y su cognición, por lo que o bien esa red solo estaría hablando de sí misma, o bien se trata de una hipótesis muy aventurada, una conjetura. Es lícito proceder en principio de esa manera, por ejemplo con fines pedagógicos o para determinar las posibilidades de cierta topología, o inclusive para simular algún fenómeno muy específico, pero sin la dinámica del entrenamiento, se estaría desarrollando un modelo mecánico con relaciones causales, condenado a la incontrastabilidad. De aquí, se puede concluir que las RU en la mayoría de los casos sólo pueden dar una aproximación estadística entre un conjunto de estímulos y un conjunto de respuestas.

Pero, como sostiene Reynoso, las RU iteran y su capacidad de clasificación es un emergente de la capacidad de almacenar información sobre la relación entre el patrón de estímulo y la respuesta clasificatoria. En ese sentido, pueden considerarse también como un modelo sistémico, en donde la causalidad no puede ser representada linealmente. Esto

estaría en línea con el hecho que, en realidad, el uso de los mecanismos estadísticos se aplica sólo a corregir los pesos entre unidades y la capacidad clasificatoria no es sólo consecuencia de su correcta aplicación, sino también, como se vio en la introducción, de las funciones de activación y de propagación, que no son estadísticas. También, hay redes que incorporan la competencia entre grupos de unidades (Rumelhart y Zipser 1986:151-193) y eso ya escaparía a la simple concomitancia estadística entre grupos de patrones. Lo que emerge de las RU consideradas como sistemas, es la estructura de la realidad "conocida" por la red, es decir, las regularidades que pueda haber en el conjunto de patrones con los que interactúa.

De cualquier manera, la decisión respecto a qué tipo de modelos se usará, no se toma en relación a qué tipo de herramienta se aplica, sino a qué tipo de problema se pretende resolver, que fenómeno describir o explicar, etcétera. Incluso, aun resolviendo un problema de acuerdo a un modelo es posible usar herramientas típicamente asociadas a otro en ciertas instancias de la investigación. Por ejemplo, un muestreo estadístico puede llevar a una hipótesis sobre una causa en un modelo mecánico, como la relación entre automovilistas ebrios y choques. También el problema podría ser susceptible de ser analizado con varios modelos, todos válidos, cada uno operando en distintos niveles, encargándose de una parte de la problemática. Por ejemplo, la medicina usa tanto modelos estadísticos como mecánicos: se buscan causas de enfermedades tanto en el laboratorio como en el muestreo estadístico, a tal punto que la hipótesis y la confirmación pueden provenir de ambos lados.

4.2 Un problema sobre direccionamiento interpersonal estadounidense

4.2.1 Modelo clásico

El problema del direccionamiento interpersonal puede ser resuelto por un modelo mecánico, de relativa complejidad, que se representa como un diagrama de flujo, donde se encadenan una serie de preguntas por sí o por no. En un modelo como éste, la decisión de llamar a alguien por su nombre de pila, por la relación de parentesco o por algún título puede entenderse, entonces, como la suma sobre sucesivas decisiones respecto a la posición social relativa del individuo, sus posibles vínculos afectivos, el contexto en que el se encuentra, etcétera.

Claudia Strauss y Naomi Quinn, contraponen a este tipo de modelo, uno conexionista basado en una RU de tres niveles. El modelo que toman de base para exponer su criticismo respecto a la metodología clásica es el de Susan Ervin-Tripp, de 1969 (Strauss y Quinn, 1997:61). Las críticas que le hacen al modelo clásico son dos. Por un lado, si se da alguna parte la información errónea, el modelo seguramente va a tomar una decisión equivocada. Por otro lado, si algún dato faltare, el proceso de decisión simplemente no podría continuar. Esto no sucede con las RU, que aún cuando ciertas constricciones que debieran actuar no lo hicieran, todavía es posible llegar a la decisión, en tanto haya otras constricciones sí presentes que puedan sostenerla. Otra crítica que hacen, aunque no al modelo de Ervin-Tripp específicamente, es que este tipo de aproximaciones no resuelven bien las situaciones nuevas (65-66).

Se podría agregar otra crítica a este tipo de modelos, aquella que fue vista más arriba en relación al aprendizaje de los significados, que se aprenden antes de saber enunciarlos.

Las personas no utilizan una serie de pasos lógicos para determinar la decisión a tomar respecto a usar tal o cual palabra para transmitir un significado. Esto se da de manera automática. Como se vio durante la discusión de los dispositivos de la antropología cognitiva de los '60, que fallaban al pensar que los nativos de una lengua podrían llegar a utilizarlos verdaderamente, el automatismo con que se manejan la asignación de un individuo a una clase hace imposible el uso de estructuras lógicas del tipo de las utilizadas en el modelo clásico.

De todos modos, se puede sostener que el modelo de Strauss y Quinn, como cualquier modelo conexionista, presupone operaciones entre unidades, sumas y multiplicaciones. Pero en este caso, las operaciones sí pueden ser ejecutadas por reacciones químicas de un dispositivo biológico masivamente paralelo como el cerebro humano. Aunque su naturaleza última y sus mecanismos de bajo nivel no sean aún conocidos, las herramientas conexionistas se asemejan un poco más a la naturaleza de lo que sí se sabe en este momento. La adición podría ser simplemente función de la cantidad de químicos durante los intercambios. La multiplicación, producto de algún cambio en el compuesto transmitido, que potencie la reacción en alguna proporción. Las ecuaciones del aprendizaje estadístico son mucho más difíciles de representar, casi tanto como las relaciones lógicas de los modelos mecánicos o las de inclusión y contraste del análisis componencial. Pero, como se vio en el capítulo anterior, los modelos estadísticos no describen lo que sucede "dentro" del fenómeno, sino relaciones o concomitancias entre datos, por lo tanto, desde ya está claro que no es eso lo que sucede en la mente humana, sino que eso tan sólo es una manera de representar la relación entre el entorno y su cognición.

4.2.2 Modelo conexionista

En el modelo de Strauss y Quinn, las 19 unidades del primer nivel, el de entrada, reciben la información necesaria para tomar la decisión acerca de cuál direccionamiento utilizar. La siguiente es una lista de ellas, con la cantidad de unidades de cada categoría entre paréntesis: parentesco (1), edad de ego (3), status local de ego (2), edad de alter (3), status local de alter (2), cercanía con alter (2), accesibilidad de alter (2), contexto mejorador de cercanía (2) y ego quiere estresar la cercanía (2). El nivel intermedio tiene 6 unidades y se conecta con todas las unidades del de entrada, a través de 114 conexiones. El nivel de salida tiene 6 unidades, cada una representando a 5 formas de direccionamiento: término de parentesco, nombre cariñoso, nombre, título más apellido, nombre genérico y ningún nombre o preguntar. Al igual que en el primer nivel, hay una conexión con cada unidad del nivel intermedio. En la figura 7 Strauss y Quinn muestran la red y sus conexiones (1997:63).

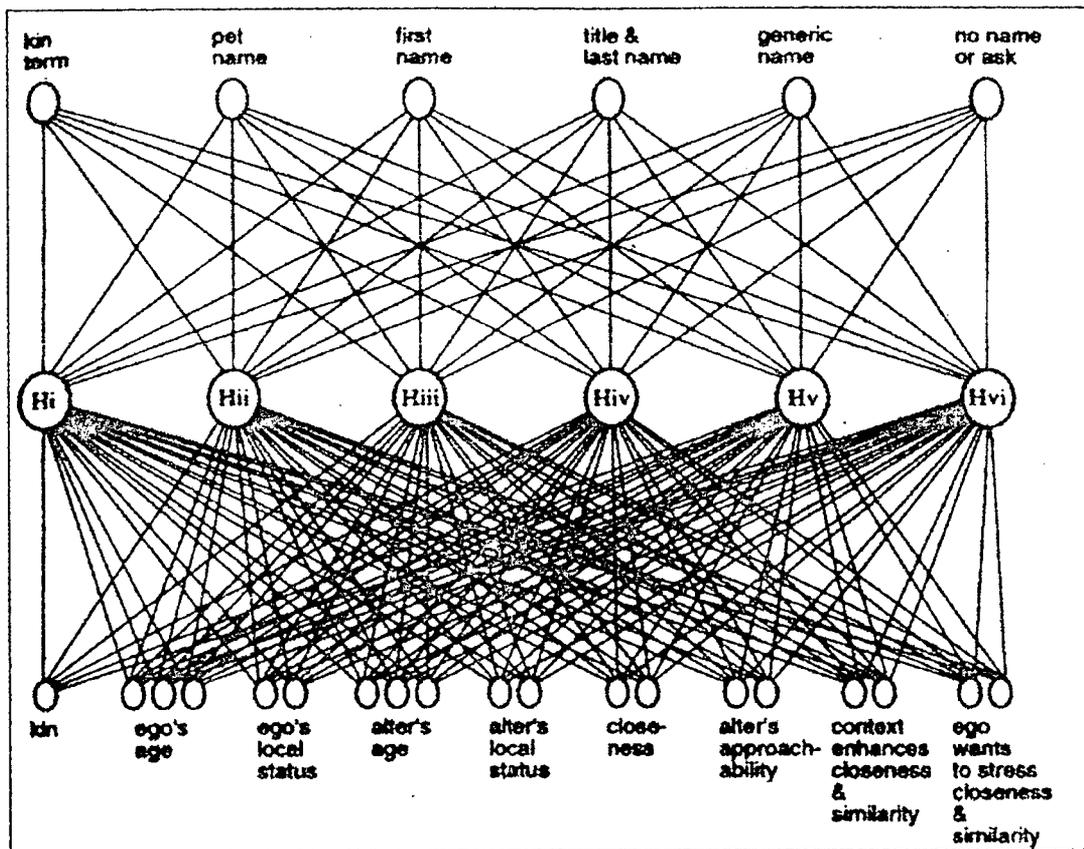


Figura 7: Red de direccionamiento interpersonal

Las autoras desarrollan el ejemplo de un estudiante pre-graduado de 20 años dirigiéndose a una profesora en sus treintas y se enfocan solamente en analizar el nivel de activación en las unidades de status, cercanía y accesibilidad. La primera unidad del status de ego (SE_i) tiene activación completa, la segunda (SE_{ii}) no tiene activación, representando el bajo status de ego en relación a alter. Para representar el status relativamente alto de alter no se activa la primera unidad (SA_i) y se activa la segunda (SA_{ii}). En el caso de cercanía se activa parcialmente la primera (C_i) y se activa totalmente la segunda (C_{ii}), denotando que se conocen pero no desarrollaron entre sí la confianza suficiente. Por último, de las unidades de accesibilidad se activan completamente la

primera (AC_i) y parcialmente la segunda (AC_{ii}) representando que para el alumno la profesora es una "persona amigable, informal" (1997:67).

Los pesos entre el nivel de entrada y el intermedio, son fijados de antemano y ningún algoritmo de aprendizaje es implementado, dado que, según las propias palabras de las autoras, ellas no son muy proficientes con las computadoras (1997:267). Los pesos pueden ser vistos en la siguiente tabla:

Valor de Activación	Unidades Entrada	H_i	H_{ii}	H_{iii}	H_{iv}	H_v	H_{vi}
1	ES_i	-0.1	0.1	-0.1	0.2	-0.4	0.6
0	ES_{ii}	-0.1	-0.1	0.2	0.2	0.4	-0.4
0	AS_i	-0.1	0.1	0.2	0.1	0.6	-0.7
1	AS_{ii}	-0.3	-0.1	-0.2	-0.2	-0.6	0.9
0.5	C_i	0.7	0.8	0.4	0.1	-0.2	-0.6
1	C_{ii}	-0.9	-0.9	-0.2	-0.2	0.5	0.6
1	AP_i	0.1	0.7	0.8	0.7	-0.2	-0.8
0.5	AO_{ii}	-0.9	-0.8	-0.2	-0.1	-0.1	0.7

La regla de propagación es la suma de los múltiplos de los pesos por los valores de activación de cada unidad, que se podría simbolizar como $a_j = \sum w_{ij} a_i$, donde a es una unidad, j representa al nivel intermedio e i al de entrada. La regla de activación es: si valor neto de entrada es menos que 0, el valor de activación es 0; si es entre 0 y 1, es el mismo valor; y si es mayor a 1, es igual a 1. Se puede representar de la siguiente forma:

$$f(j) = \begin{cases} \text{si } \sum_i w_{ij} a_i \geq 1 \rightarrow 1 \\ \text{si } \sum_i w_{ij} a_i \leq 0 \rightarrow 0 \\ \text{si } \sum_i w_{ij} a_i < 1 > 0 \rightarrow \sum_i w_{ij} a_i \end{cases}$$

Si se aplica esta función a las unidades del nivel intermedio, que toman los valores de activación de las unidades de entrada, aquellas tendrían los siguientes valores de activación:

H_i	H_{ii}	H_{iii}	H_{iv}	H_v	H_{vi}
-1.3	-0.2	0.4	0.5	-0.85	1.35

Como recién vimos, los valores de activación que son 0 o menos, se transforman en 0, por lo que H_i , H_{ii} y H_v , no están activadas. H_{iii} y H_{iv} , pasan con el mismo valor y H_{vi} tiene un valor de activación 0. Estos valores de activación se aplican usando la misma fórmula que en el nivel anterior, $a_j = \sum w_{ij} a_i$. La siguiente tabla muestra los pesos de estas unidades con el nivel de salida:

Valor de Activación	Unidades Intermedias	Nombre	Apellido	Término genérico	No usar nombre o preguntar
0	H_i	0.9	-0.9	-0.9	-0.9
0	H_{ii}	0.9	-0.9	-0.9	-0.9
0.4	H_{iii}	0.8	-0.6	-0.9	-0.4
0.5	H_{iv}	0.4	0.4	-0.6	0.2
0	H_v	0.2	-0.2	0.2	0.2
1	H_{vi}	-0.4	0.9	0.2	0.4

Finalmente, los pesos resultantes de las unidades de salida son los siguientes:

Nombre	Apellido	Término genérico	No usar nombre o preguntar
0.12	0.46	-0.46	0.34

Lo primero que infieren las autoras de la simulación hecha con esta RU es que lo único que hace esta red es descartar el uso del término genérico. En realidad eso es sólo parcialmente verdadero, porque tranquilamente se puede elegir el uso del apellido, que es la unidad con más valor de activación. De cualquier manera, ellas determinan la necesidad de modificar los umbrales de activación, de manera que se muestre las preferencias generales de estudiante, que se asume un poco tímido, por lo que tenderá a usar el apellido más frecuentemente. Entonces, el umbral de activación de apellido es menos que de nombre. Nombre genérico tendrá un umbral de 0.2 y no usar nombre, de 0.

Este modelo también requiere ciertas conexiones inhibitorias. Por ejemplo, si se usa el nombre cariñoso, no se usa el apellido. Si nombre y apellido son activados, inhibirán al resto. Si se usa título y apellido, no se usará genérico, etcétera. Lo que hacen estas relaciones inhibitorias es establecer constricciones entre las unidades de salida.

Las autoras ven que una de las ventajas más grandes de este modelo, a parte del hecho de ajustarse a la teorías de los esquemas y de estar más cerca de la realidad psicológica de este tipo de decisiones, es que es hipotéticamente posible representar en él el proceso de aprendizaje que va desde las primeras palabras de direccionamiento que se aprenden en la niñez a las formas más avanzadas de direccionamiento usado en edad adulta. Lamentablemente, dada su poca destreza en las computadoras, no desarrollan ningún esquema de aprendizaje. Lo cuál tampoco es trivial, ya que en general las RU no están diseñadas para aprender patrones simulando un proceso de desarrollo en el tiempo del tipo del aparato cognitivo humano. Como ya se dijo, los algoritmos de aprendizaje son de corrección del error estadístico y no contemplan formas de corrección "cualitativas". Estas formas cualitativas podrían estar relacionadas con umbrales a partir de los cuales se agregan clases o se amplía la red o se combinan con otras redes o cosas por el estilo.

Esta RU padece de los dos problemas ya señalados más arriba, y es que al hacer uso de etiquetas se pierde la posibilidad de que esta RU se vincule con otra pues, como se dijo, la relación entre símbolo y unidad es tan sólo una hipótesis no interpretable por la RU. Dicho en términos conexionistas, las etiquetas no generan ninguna constricción en particular.

El otro problema está relacionado con la cuestión de la asignación manual de los pesos. Esto, como ya se expuso, tiene como consecuencia la "mecanización" de la herramienta.

La única defensa posible es que este modelo sólo pretende mostrar las posibilidades de las RU para sortear los obstáculos de los modelos clásicos. También, se podría convenir, es posible que este modelo esté mucho más cerca de la realidad psicológica de los actores que los dispositivos basados en operaciones de lógica proposicional. Al igual que en la vida real, este modelo conexionista no requiere ninguna operación lógica o semántica compleja.

4.3 Clasificando elementos del sitio de arte rupestre de Campo Morado

4.3.1 Entrenamiento supervisado

Un uso diametralmente opuesto de las RU es la clasificación u ordenamiento de los patrones de estímulo sin etiquetar las unidades de entrada. Por supuesto que al no etiquetarse, los estímulos quedan limitados a aquello que pueda ser representado de manera completa con una matriz de números. En ese sentido, no sería posible analizar un problema como el de direccionamiento.

Esto limita los estímulos posibles y por ende los problemas a tratar. Por supuesto que queda descartada la posibilidad de usar el significado en cualquiera de sus formas conocidas. Aunque sí sería posible usar información del objeto denotado, como su imagen o sonido.

Las posibilidades metodológicas, entonces, son más acotadas. Por otro lado, habría aplicaciones más tangenciales a la investigación en sí, que están más relacionadas con la organización y búsqueda del material de estudio que con la formulación de hipótesis o construcción de modelos. En ese contexto se enmarca la aplicación presentada por Diaz y

Castro (Diaz y Castro, 2001) que contiene un módulo que implementa una red a través de un software de interfaz de programación de aplicación (llamado API, por Application Programmer Interface) provisto por terceros, disponible en aquel entonces en el mercado, que contiene preprogramados los procedimientos estándar de entrenamiento de redes.

En el mencionado trabajo, Diaz y Castro no ofrecen información sobre experimentos específicos dado que la aplicación se encontraba en desarrollo. Por ello, el entrenamiento de la red y la selección de imágenes se hicieron especialmente para este trabajo. Estos experimentos se desarrollaron usando la información provista por María Isabel Hernandez Llosas y colaboradores a través del trabajo que presentaron en el congreso de Tarija, Bolivia en el año 2001; titulado “Los grabados de Campo Morado –Departamento de Tilcara, Jujuy, Argentina”.

Hernandez y colaboradores utilizan como una unidad básica de relevamiento al “elemento”, entendiéndolo por “estas mínimas unidades de contenido gráfico que presentan unidad de sentido y de ejecución” (Hernandez et al, 2001:9). Por un lado se hizo así para evitar hacer inferencias respecto al agrupamiento de los elementos en esa etapa inicial de la investigación aunque también se tomó esa decisión para poder facilitar su procesamiento en la Base de Datos Computarizada de Diaz y Castro, en desarrollo en ese mismo proyecto. Los elementos se agrupan luego en motivos y estos a su vez en conjuntos y grupos estilísticos. Los experimentos aquí hechos usan solamente elementos.

Los elementos encontrados son clasificados en tres categorías: antropomorfos, zoomorfos y abstractos. En esta última categoría se subdivide en serpentiformes, puntiformes, círculos, circunferencias, cruciformes y trazos. De los 178 elementos

identificados, 107 son abstractos. De los figurativos, que son 28, 19 son antropomorfos y el resto zoomorfos.

La aplicación utilizada consiste en un "asistente" que guía al usuario por los pasos para definir la red. La RU puede poseer varios niveles intermedios. También se puede definir el tamaño de la matriz con que se trabajará. En estas experiencias se usó una de 64 x 64, es decir, se redujeron las imágenes a 4096 posiciones. Este tamaño permite cierta velocidad de procesamiento sin perder demasiada resolución. En la Figura 8 se muestra el ejemplo de un elemento y su representación en la matriz. En el asistente también se determinan los parámetros del entrenamiento en sí, que son cuántos ciclos como máximo se entrena a la red, cuál es el error máximo que se desea, el impulso (momentum) y la razón del aprendizaje. El error es la diferencia entre el estado real de los pesos y su estado óptimo, determinado por el resultado esperado. En otras palabras, cuán lejos está la red de asignar el patrón a la clase correcta. El impulso es que proporción de los pesos del último cambio que se aplicará en el nuevo cambio y tiene el objetivo de evitar que la red oscile entre los patrones de ejemplo sin reducir nunca el error a nivel general. La razón de aprendizaje controla cuán rápido aprende la red.

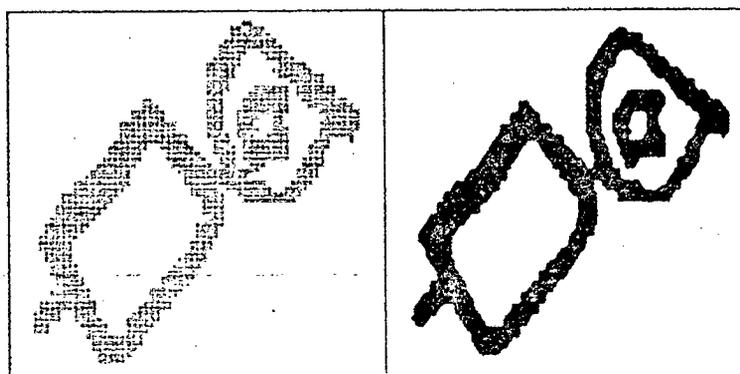


Figura 8: La primera imagen es la matriz de 64x64. La segunda es la fuente usada

EL primer experimento se realizó para determinar si se podía entrenar a la red para distinguir elementos de dos niveles de complejidad. Para ello se eligieron un grupo de tres elementos circulares y tres figuras compuestas por un tronco y una cabeza, posiblemente antropomorfas, las que por ser figurativas presentan un mayor nivel de complejidad. Ese nivel de complejidad consiste en el hecho que las figuras tienen más de un componente, como tronco y cabeza, o incluso tres, tronco cabeza y cola o piernas. Los tres ejemplos de cada clase elegidos pueden verse en la figura 9.

Al mismo tiempo, la idea era que la red pudiera identificar individuos no incluidos en el set de entrenamiento que tuvieran la particularidad de ser "liminales", que estuvieran cerca del límite entre una clase y otra. En el caso de los simples, el elemento elegido era un círculo con una línea que sale hacia fuera de su circunferencia, que se podría asemejar a una cola, pierna o algún tipo de extremidad. Para el caso del individuo complejo, se eligió a uno que parecía ser el más simple entre los complejos, que consiste en un tronco en forma de anillo, en cuyo centro hay pequeño círculo y con otra circunferencia menor a modo de cabeza. Ambos ejemplos pueden verse en la figura 10.

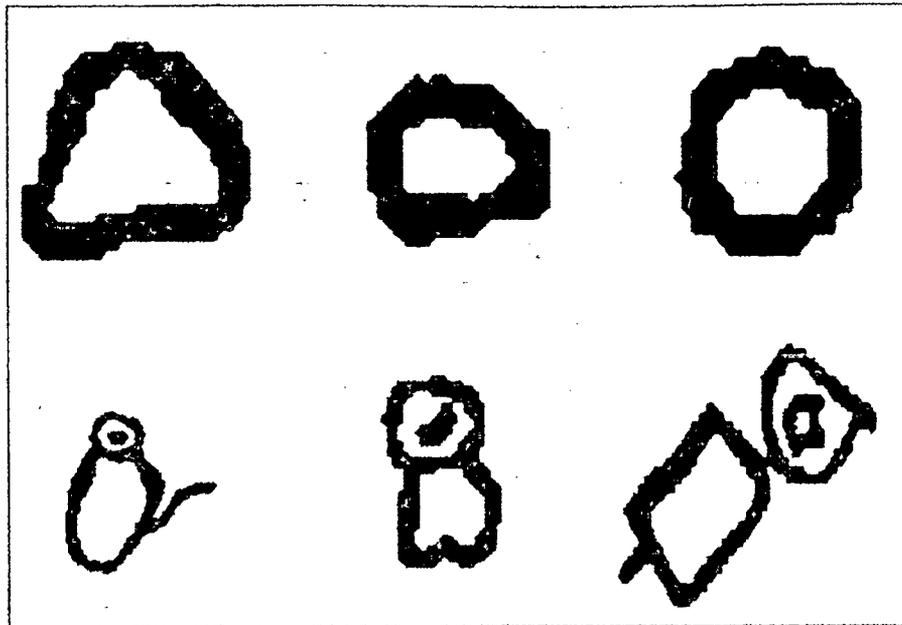


Figura 9: Ejemplos de circunferencias y motivos con cabeza y tronco (patrones a, b, c, d, e y f)

La red usada tiene tres niveles, el primero con el tamaño de la matriz, el segundo de 20 unidades y el último 2, que es a su vez la cantidad de clases. Se usó una razón de entrenamiento de 0.3 y un impulso de 0.1, poniendo como objetivo de error a 0.005. El entrenamiento duró sólo unos minutos en una máquina de doble procesador, llegando a identificaciones exitosas en los dos casos.

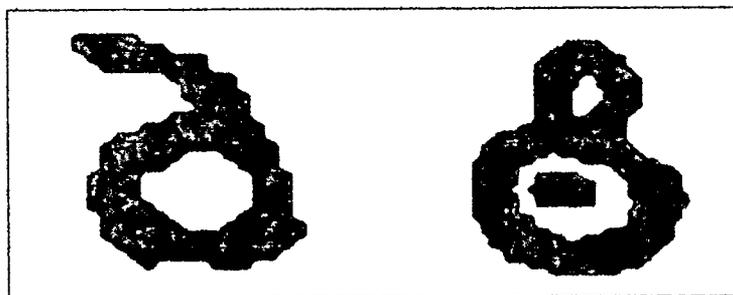


Figura 10: Dos figuras que se identificaron la primera (g) como simple y la segunda (h) como compleja

Sin embargo, cuando se le dio a la red algunos patrones muy simples, como un punto o un pequeño cuadrado en un lugar arbitrario de la matriz, los mismos fueron reconocidos como complejos. En cambio, cuando se le presentaba un cuadrado más o menos grande

(por ejemplo 40 x 30), lo que es aproximadamente las medidas promedio de los círculos presentados como ejemplo, aquel era identificado como un elemento simple. Aumentando la cantidad de unidades del nivel intermedio hasta 1000 unidades, probando la hipótesis de que la RU podría faltarle granularidad, empeoró la situación, ya que la red comenzó a “oscilar” entre valores de error altos en un cierto momento, aproximadamente a los 100 ciclos de entrenamiento.

Una manera de evitar la oscilación podría haber sido agregar un nivel intermedio más, de 20 unidades, dado que eso debiera bajar las oscilaciones al provocar ajustes de pesos menos abruptos¹⁵. En ese caso, el entrenamiento se completaba, pero la red dejó de reconocer como simple a la figura h.

En el segundo experimento fue una variación del primero, en el que se diversificaron los elementos simples, agregando una figura con forma de herradura y una con forma de flecha (Figura 11), usando los mismos parámetros de entrenamiento que en el caso anterior. La diferencia fue que aún en el caso de la configuración con un sólo nivel intermedio de 1000 unidades, la figura a) era reconocida como simple. Esto muestra un mayor poder de generalización, pero en vez de ser a costa de una mejor configuración de la red, se debe a las características de mayor variabilidad en los ejemplos presentados.

¹⁵ Tómese en cuenta que los ajustes de pesos de hacen de atrás para adelante, es decir del nivel de respuesta hacia el nivel de entrada. Al tener que cambiarse los pesos de intermedio de 1000 unidades a partir de la diferencia que se da en el de dos, no se pueden evitar los cambios de abruptos en los pesos.

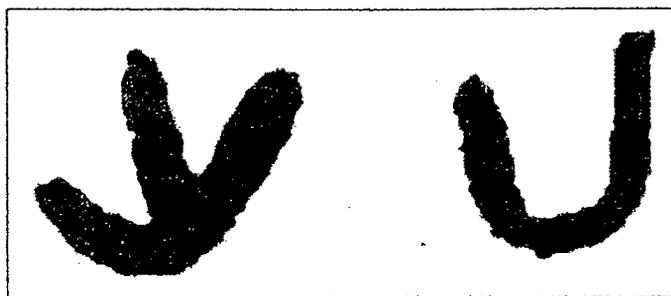


Figura 11: las dos figuras que se agregan en la clase de las circunferencias, reemplazando a dos de ellas (patrones i y j)

Para la siguiente experiencia, se le agregó a la red una tercera clase de los elementos que aparentan mayor complejidad. Estas figuras son en general antropomorfas y representan a veces dos individuos enfrentados (ver figura 12). Los juicios sobre complejidad, en realidad, son asunciones. Un elemento simple podría ser parte de un motivo más complejo o representar una idea compleja. De cualquier manera, complejas o no, se están eligiendo como muestras aquellas figuras que se parecen entre sí y, entonces, pueden ser asignadas a una misma clase. Los nombres que se le ponen a las clases, entonces, son irrelevantes desde punto de vista del algoritmo de aprendizaje.

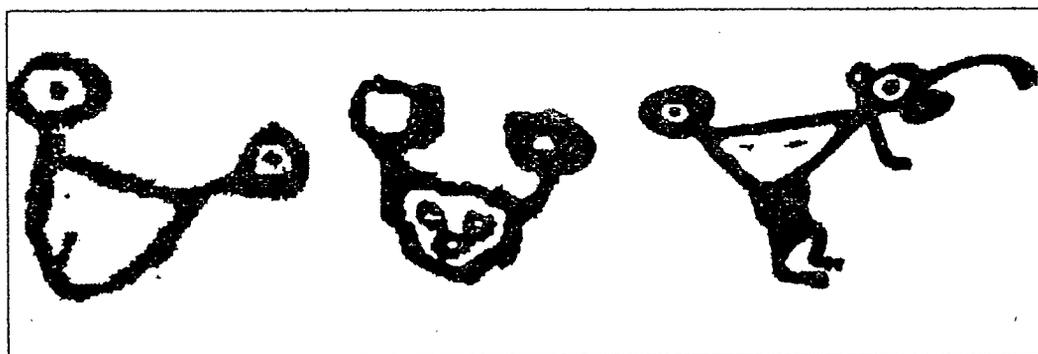


Figura 12: Elementos de la clase más compleja que se agregó en el tercer experimento (k, l y m).

Los individuos usados para el reconocimiento fueron los mismos que en la experiencia anterior, con el agregado de un elemento complejo con la particularidad de que está

incompleto (Figura 13). El mismo fue reconocido sin problemas después de 100 ciclos de entrenamiento.

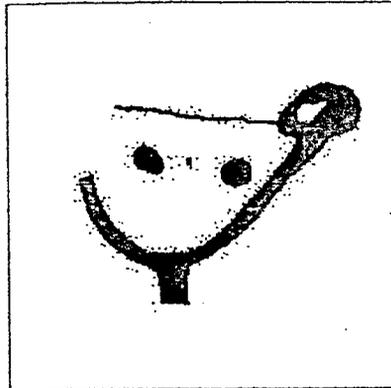


Figura 13: Elemento antropomorfo "incompleto" (patrón n)

A este punto se pueden vislumbrar interesantes usos a una red que aprenda a diferenciar entre clases de elementos. Por ejemplo, sería una útil herramienta de búsqueda, suponiendo que se tuviese una base de datos con los elementos cargados en algún campo, como lo sugieren Díaz y Castro en el trabajo mencionado. Allí se podrían hacer consultas del tipo de "todo los sitios en donde hubiere por lo menos un elemento de la clase A y otro de la B". Inclusive, se podría "indexar" esa base de datos a partir de distintos tipos de clasificaciones, usando para ellos distintas RU entrenadas con distintas configuraciones de clases. Estas confirmaciones formarían intersecciones donde los elementos pertenecerían a más de una. Por ejemplo, bien podría clasificarse por complejos y simple, humanos y animales, etcétera.

4.3.2 Entrenamiento no supervisado

En la introducción se mencionó que existen cuatro paradigmas de aprendizaje. Las experiencias que se acaban de ver corresponden todas al tercero, el paradigma de clasificación. El cuarto paradigma, el detector de regularidades, se caracteriza por la

ausencia de supervisión durante el entrenamiento, en el que entonces no hay información predefinida respecto de la pertenencia de los elementos a clases. Al no existir dicha información, la RU tiene como objetivo organizar en clases la población de patrones. Es en ese sentido que es un detector de regularidades, dado que aquellos patrones que pertenezcan a una misma clase lo hacen a partir de compartir ciertas características entre sí.

La RU que se usó para esta experiencia, es una del tipo de aprendizaje organizado inventada por Teuvo Kohonen. También se la conoce como "Mapeo Autoorganizado". Este tipo de RU tienen sólo dos niveles: uno de entrada del tamaño de los patrones a ingresar y uno de salida que determina tanto la cantidad máxima de clases a encontrar como el tamaño posible del "vecindario". El "vecindario" es la cantidad de neuronas adyacentes a cada lado de la unidad que representa a la clase ganadora y que se ven afectadas positivamente. Que una unidad "gane" significa que un patrón es asignado como perteneciente a la clase representada por esa unidad. Los pesos son ajustados como si todas las unidades del vecindario fuesen ganadoras de modo que se evita que una vez que una neurona empieza a ganar, sea imposible revertirlo y todos los patrones terminen perteneciendo a una misma clase. El otro parámetro que se utiliza es la razón de entrenamiento. La diferencia con la RU supervisada es que este parámetro se va reduciendo a medida que se sucede el entrenamiento.

Los patrones que se utilizaron son los mismos de los ejemplos anteriores aunque sin las asignaciones a clases que los acompañaban, con el agregado de ñ y o, que puede verse en la figura 14. También se eliminaron i y j, el patrón con forma de flecha y el de herradura.

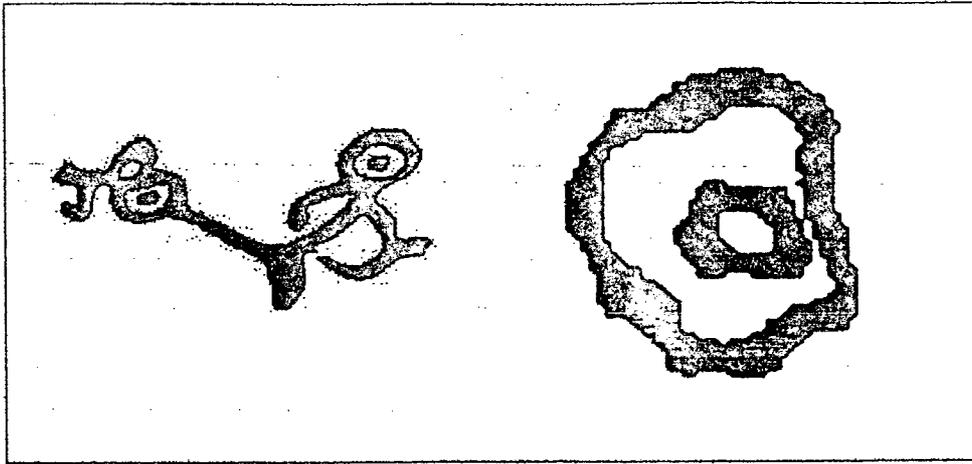


Figura 14: Patrones fi y o que se agrega para la experiencia de RU no supervisada.

Se experimentaron con varias combinaciones de parámetros. Una de las cuestiones que primero se presentó es la elección del tamaño del vecindario. Como no hay ninguna fórmula que pueda precisar cuál es el valor conveniente para una población y RU dadas, hay que buscar el valor corriendo el programa varias veces hasta determinar la configuración óptima. Cuando se le ponía un solo vecino de cada lado como parámetro, todos los elementos eran asignados a la misma clase. Si, en cambio, se asignaba la misma cantidad de vecinos que de unidades, el resultado consistía en tener a la mayoría asignados a la misma clase. En el primer caso el efecto era que aumentaban las probabilidades que una unidad empezara a ganar todo el tiempo y en el segundo, que aquella unidad que se encontraba en el vecindario de la mayor cantidad de otras unidades tendía a ganar, aunque algunos de los elementos eran asignados a otras clases. Por eso se trabajó con 5 y 8 vecinos, tomando como cantidad de unidades de salida a 10. La cantidad de unidades de salida, aparte de restringir la cantidad máxima de clases, determina la granularidad de la red, dado que a más unidades de salida, mayor cantidad de interconexiones. Más allá de 10 clases, no se vieron diferencias en la configuración de

clases lograda durante los experimentos realizados, por lo que se fijó así la cantidad de unidades de salida.

Se probaron varias combinaciones de razón de aprendizaje y tamaño de vecindario, aunque, excepto en los casos extremos mencionados, la asignación de clases fue aproximadamente la misma. La siguiente tabla muestra las unidades ganadoras en dos casos, con dos corridas del programa cada uno, en las que se usó tamaño de vecindario (v) igual a 5 y 8 y se mantuvo la razón de aprendizaje en 0.3.:

Figura	Clase en RU supervisada	v=8	v=8	v=5	v=5
D	Zoomorfos	10	1	1	1
E	Zoomorfos	10	1	1	1
F	Zoomorfos	1	10	10	10
H	Zoomorfos	10	1	1	1
O	n/d	6	1	1	1
C	Simples	6	10	6	6
B	Simples	6	10	6	6
A	Simples	6	10	6	6
G	Simples	6	1	2	2
K	Antropomorfos	4	9	10	10
L	Antropomorfos	1	10	10	10
M	Antropomorfos	1	10	10	10
N	Antropomorfos	1	10	10	10
N	n/d	1	9	10	10

Hay dos cosas que se puede concluir de estos experimentos y que pueden notarse en la tabla. Por un lado, dado que los pesos iniciales se establecen aleatoriamente, distintas corridas con los mismos parámetros pueden dar resultados levemente distintos. De cualquier manera, las RU no son lieneales, como ya se afirmó, por lo que inclusive con todo los pesos iniciados en 0, no se obtiene exactamente el mismo resultado. Estas diferencias atañen al número de unidad que asigna a cada clase y a los patrones que se podrían llamar "casos liminales", dado que las diferencias que se ven en dos corridas distintas se dan siempre en ellos. Sería aquellos que para la red, están más lejos del centro

focal o centroide de la clase. Esos casos serían f, o, g, k y ñ. El resto de los casos, todos pertenecen a la misma clase en todas las corridas, y están agrupados de la misma manera que se había definido para la experiencia de la RU supervisada.

Otra particularidad de la distribución lograda es que incluso cuando se aumentan la cantidad de unidades, la cantidad de clases en que son agrupados los elementos no varía y siempre es alrededor de 4 o 5. Si no se cuentan las clases de un solo miembro, que corresponden a los elementos que no pueden ser agrupados en las otras clases y, luego, tienen su propia clase, la RU clasificó consistentemente con la experiencia de RU supervisada.

En realidad, las anomalías no sólo deben considerarse como un atributo del elemento clasificado, como diciendo que hay un elemento anómalo en la población. Si se toma en cuenta que los pesos de la RU son modificados por cada patrón clasificado a medida que se sucede el auto-entrenamiento, es posible que un elemento o conjunto de elementos modifiquen los pesos de modo tal que afecten la manera en que se asigna algún otro elemento a su clase. De no estar aquel/aquellos, éste sería clasificado correctamente. Las asignaciones de clase, entonces, no son sólo atributo de las características de un elemento dado o de los elementos de una clase dada, sino de toda la población.

5 Conclusiones.

Se ha visto como los dispositivos desarrollados por la corriente cognitivista fallan al representar la realidad psicológica del uso de los significados, debido a que las relaciones de inclusión, contraste o intersección que utiliza el análisis componencial, no pueden manifestarse en un dispositivo que funcione con el automatismo de la mente humana. También se ha visto como una concepción holística, basada en focos y prototipos lleva a la antropología cognitiva a interesarse por los dispositivos conexionistas, teoría de los esquemas y las RU.

Las RU tienen la ventaja de poder representar mejor que los dispositivos propios del cognitivismo de los años 60 la forma en que se manifiestan ciertos aspectos de la cognición humana como por ejemplo la búsqueda de patrones concordantes o el acceso a los significados de manera automática. Sin embargo, a pesar del tremendo cambio en la teoría que esto significa, el uso de las RU no se ha extendido demasiado.

De todos modos, de acuerdo a lo visto en los capítulos anteriores, es posible determinar la función que la herramienta analizada puede desempeñar en la investigación antropológica:

- a. Se pueden desarrollar con las RU modelos de aprendizaje estadístico, de pertinencia antropológica, con objetivos clasificatorios en base a una distribución de una población de patrones.
- b. Si es que se quiere incrustarlas unas con otras, no se pueden usar RU que incluyan etiquetas. Ahora bien, sólo es posible inferir la presencia de los esquemas a partir de la conducta emergente de RU y el uso de etiquetas. Entonces, los esquemas sólo se pueden representar como entidades aisladas que

no es posible interrelacionar. Para la antropología eso significa que solo es posible abordar problemas simples como el de direccionamiento interpersonal de Strauss y Quinn.

- c. La red de direccionamiento interpersonal sirve para confirmar la hipótesis a) esbozada al comienzo de este trabajo. En ella se sostenía que es posible abordar exitosamente un problema simple de pertinencia antropológica con una RU. La conclusión es que sí es posible hacerlo y que se es exitoso allí donde los modelos clásicos fracasan.
- d. Lo interesante de la RU de direccionamiento, es que se podría generalizar y armar un modelo que prediga como se direccionan las personas en un determinado medio cultural, no sólo con los pesos diagramados para la situación de un individuo. Incluso, podría investigarse los cambios necesarios para que funcione en otros medios culturales.
- e. Las RU de entrenamiento supervisado no sirven como mecanismo de inferencia y, aunque útil, no es posible sacar de él consecuencias observacionales directas que modifiquen o sugieran alguna nueva hipótesis ni tampoco que confirme o refuten alguna existente. La razón es que para el entrenamiento ya se debió haber contrastado y confirmado las hipótesis respecto a qué clases existen y cuáles son los ejemplos prototípicos de cada clase definida.
- f. Sin embargo, existe la posibilidad de usar las RU supervisada como mecanismo indirecto de inferencia. Por ejemplo, en caso de configuraciones de clases confusas o de difícil decisión, se podría explorar qué configuraciones de clases ofrecen una mejor capacidad de ser aprendidas por una RU.

g. Por otro lado, se puede concluir que las RU de Kohonen tienen capacidad de inferir clasificaciones de cierta coherencia a partir de una población de elementos. Por lo que es posible confirmar la segunda hipótesis presentada en la introducción: es posible tanto desarrollar mecanismos que pueden asociar elementos a clases a partir de los ejemplos ya existentes y también es posible crear un mecanismo de inferencia de clases que posea un cierto nivel de éxito.

Sería muy aventurado afirmar categóricamente que lo aquí mostrado es una prueba indiscutida de las posibilidades de la herramienta analizada. Si bien quedó demostrado que la herramienta ofrece potencial a distintos niveles, existen variadas limitaciones que podrían llevar al fracaso en la resolución de problemas específicos.

En primer lugar, la estrategia de implementación debería determinar si es posible abordar el problema haciendo algunas experiencias a través de un "prototipo" de resolución. La razón de esto es que hay preguntas que sólo se contestan con la prototipación. Por ejemplo, ¿son mis patrones lo suficiente parecidos para que se produzca algún nivel de generalización en la RU?; ¿son lo suficientemente distintos para que la RU no "memorice"? También hay una cuestión de índole logística, que es el nivel de detalle que se necesita para diferenciar un conjunto de patrones. Si es muy alto, puede que no haya equipos para procesar los algoritmos de aprendizaje.

Sin embargo, los problemas mencionados no atañen a la esencia de la teoría en que se basa la herramienta ni a sus principales postulados. Atañen a cuestiones de índole práctica y el hecho que no sean solubles en este momento no significa que no lo sean en unos años, con la disponibilidad de mejor equipamiento, el uso de RU implementadas en chips para aumentar su velocidad, el descubrimiento de nuevos algoritmos de aprendizaje

o el desarrollo de mejores topologías de interconexión. Inclusive, la solución a algunas de las cuestiones puede que suceda a partir de alguna otra tecnología que se combine con las RU.

Finalmente, si bien se puede concluir que el uso de las RU ha sido fructífero y puede seguir asistiendo en la modelación de soluciones de pertinencia antropológica, hay dos obstáculos que no han podido ser resueltos y que deben considerarse cuando se tiene que tomar la decisión de usar RU en algún proyecto de investigación.

El problema de la relación entre redes, es quizá el que posee más impacto negativo. Si bien está asociado al uso de etiquetas, en realidad parte desde una cuestión más de fondo. Uno de los postulados básicos del PDP es que no hay algo así como un lenguaje interno de representación del conocimiento en la mente humana y las RU se modelaron tomando en cuenta eso. Si el paso de una RU a otra, al poseer etiquetamiento requiere un proceso de codificación-recodificación, no queda otra posibilidad que asumir la presencia de un código, por más primitivo y simple que sea. Por esa razón, sólo se pueden abordar problemas simples que no requieran recodificación.

El otro problema es el de la asignación de los pesos de las redes, en el sentido de diseñar una configuración de unidades y de pesos como en el caso de la red de direccionamiento interpersonal. Exceptuando las redes de dos niveles como la de Kohonen, las redes siempre pueden llegar a una configuración de pesos deseada a través del entrenamiento. Aunque como señalan Minsky y Papert (1967), no existe un algoritmo universal para ello, es posible desarrollar reglas locales de cambio de pesos, como así también partir de una configuración de pesos que no sea aleatoria. Esta decisión podría tener una

justificación biológica, pues es improbable que todas cargas químicas de las células nerviosas se iniciasen con valores aleatorios o todos en cero.

También es posible diseñar la topología de la red como cuando, por ejemplo, se asignan tres unidades para representar la edad de ego en la red de direccionamiento interpersonal. Cualquier respuesta coherente que se de a preguntas del tipo "¿por qué n unidades para representar la característica x ?", remitirá seguramente a alguna relación causal y, por ello, al de un orden simbólico irrepresentable en la RU.

La solución a este problema puede estar en desarrollar mecanismos no simbólicos que generen las configuraciones de pesos y de las topologías convenientes, cuando esto no es posible de hacer a través del entrenamiento. Esto se podría lograr con algoritmos genéticos o programación evolutiva en tanto que la RU óptima puede considerarse como un individuo de una población de RU. Esto es posible siempre y cuando el criterio de evaluación de éxito y las variables que determinan los distintos tipos de soluciones a comparar son conocidos. En la redes ambos lo son.

En última instancia, esto remite a la cuestión del tipo de modelo que es posible desarrollar con la herramienta en cuestión. Si se conviene que las redes son en la mayor parte de los casos un mecanismo para desarrollar modelos estadísticos, no se podrá representar con ella las relaciones causales ni simular comportamientos emergentes. Si una nueva tecnología o un desarrollo ulterior de las actuales RU permitieran en el futuro representar procesos más complejos, entonces el poder de esta herramienta aumentaría sustancialmente y se estaría frente en la posibilidad de desarrollar modelos que simulen las complejidades emergentes de la relación entre mente y entorno. Mientras tanto, los antropólogos tendrán que contentarse con usar las redes para problemas más bien

simples. O con desarrollar modelos estadísticos que tengan el objetivo de encontrar o confirmar regularidades en poblaciones de patrones. Lo cual no es poco, pues eso significa levantar levemente el velo de una pequeña parte de la estructura de la realidad.

6 Bibliografía

- BERLIN Brent, Dennis BREEDLOVE y Peter RAVEN, "Covert Categories and Folk Taxonomies" en *American Anthropologist*, 70 290:299, 1968.
- BERLIN Brent, Paul KAY, *Basic Color Terms*, CSLI Publications, Stanford, USA, (1969) 1999.
- BLACK, Mary y Duane METZGER, "Ethnographic Description and the Study of Law" en Stephen Tyler (ed.), *Cognitive Anthropology*, Holt Rinehart and Winston, New York, (1964) 1969.
- BRIGHT, Jane O. y William BRIGHT "Semantic Structures in Northwestern California and the Sapir - Whorf Hypothesis" en Stephen Tyler (ed.), *Cognitive Anthropology*, Holt Rinehart and Winston, New York, (1965) 1969.
- COLLINS, Alan y Dedre GENTNER, "How people construct mental models" en Dorothy Holland y Naomi Quinn, *Cultural Models in Language and Thought*, Cambridge University Press, Cambridge, UK, 1987.
- D'ANDRADE, Roy, *The development of Cognitive Anthropology*, Cambridge University Press, Cambridge, UK, 1995.
- DÍAZ, Diego y Damián CASTRO, "Pattern recognition applied to rock art" pp 463-468 Göran Burenhult, editor *Archaeological informatics: Pushing the Envelope CAA 2001 29th conference*, Gotland, abril 2001 BAR International Series 1016, Oxford, 2002
- FRAKE, Charles, "Notes on Queries in Ethnography" en Stephen Tyler (ed.), *Cognitive Anthropology*, Holt Rinehart and Winston, New York, (1964) 1969 .
- GEERTZ, Clifford. *The interpretation of cultures*, Basic Books, New York, 1976.

- GOODENOUGH, Ward, "Yankee Kinship Terminology: A Problem in Componential Análisis" en Stephen Tyler (ed.), *Cognitive Anthropology*, Holt Rinehart and Winston, New York (1965) 1969.
- GRAUBARD, Stephen R. *El Nuevo debate sobre la inteligencia artificial*, Gedisa, Barcelona, 1993.
- HERNANDEZ LLOSA, María Isabel, Judith CHARLIN, Antonela DI VRUNO, Mariel LOPEZ, Laura MARCHESE, Fabián VALLADARES y Silvina VIGLIANI "Los grabados de Campo Morado Departamento Tilcara, Jujuy, Argentina" en Alicia Fernández Distel (ed.) *Arte Rupestre y Región. Arte Rupestre y Menhires en el Sur de Bolivia, Noroeste de Argentina y Norte de Chile*, Anuario del Centro de Estudios Indígenas y Coloniales CEIC N° 2, paginas 109 -144. Editorial Universidad Nacional de Jujuy, San Salvador de Jujuy, 2001.
- HINTON, Geoffrey y James MCCLELLAND y David RUMELHART, "Distributed Representations" en Rumelhart, David, James McClelland (Ed.), *Parallel Distributed Processing I*, MIT Press, Cambridge, Massachusetts, USA, 1986.
- HOLLAND, Dorothy y Naomi QUINN, "Culture and Cognition" en Dorothy Holland y Naomi Quinn, *Cultural Models in Language and Thoughts*, Cambridge University Press, Cambridge, UK, 1987.
- HUTCHINS, Edwin "Myth and Experience in the Trobriand Islands" en Dorothy Holland y Naomi Quinn, *Cultural Models in Language and Thoughts*, Cambridge University Press, Cambridge, UK, (1983) 1987.
- KAY, Paul "Comments on Colby" en Stephen Tyler (ed.), *Cognitive Anthropology*, Holt Rinehart and Winston, New York, (1967) 1969.

- KEMPTON, Willet, "The theories of home heat control" en Dorothy Holland y Naomi Quinn, Cultural Models in Language and Thoughts, Cambridge University Press, Cambridge, UK, (1986) 1987.
- LAKOFF, George y Zoltán KÖVECSES, "The cognitive model of anger inherent in American English" en Dorothy Holland y Naomi Quinn, Cultural Models in Language and Thoughts, Cambridge University Press, Cambridge, UK, 1987
- LITTLE, Elbert, National Audubon Society – Field Guide to Trees Western Region, Alfred A. Knopf, New York, 1980
- LOUNSBURY, Floyd G., "The Structural Analysis of Kinship Semantics" en Stephen Tyler (ed.), Cognitive Anthropology, Holt Rinehart and Winston, New York, (1964) 1969.
- LUTZ, Catherine "Goals, events and understanding in Ifaluk emotion theory" en Dorothy Holland y Naomi Quinn, Cultural Models in Language and Thoughts, Cambridge University Press, Cambridge, UK, (1983) 1987
- MISNKY, Marvin L. y Seymour A. PAPERT, Perceptrons, MIT Press, Cambridge, Massachusetts, USA, (1969) 1990.
- MISNKY, Marvin, The society of mind, Simon and Schuster, New York, 1986.
- MORRIS, Charles, Signs, Language and Behavior, Prentice-Hall, New York, USA, 1946.
- PIKE, Kenneth, "Towards a theory of the structure of human behavior" en Estudios Antropológicos Publicados en Homenaje al Doctor Manuel Gamio, Mexico, 1956
Selected readings p 107-116, Mouton, La Haya, 1972
- PIKE, Kenneth and Hugh STEVEN, Pike's Perspectives, Credo Publishing Corporation, Langley, British Columbia, Canada, 1989.

- KEMPTON, Willet, "The theories of home heat control" en Dorothy Holland y Naomi Quinn, Cultural Models in Language and Thoughts, Cambridge University Press, Cambridge, UK, (1986) 1987.
- LAKOFF, George y Zoltán KÖVECSES, "The cognitive model of anger inherent in American English" en Dorothy Holland y Naomi Quinn, Cultural Models in Language and Thoughts, Cambridge University Press, Cambridge, UK, 1987
- LITTLE, Elbert, National Audubon Society – Field Guide to Trees Western Region, Alfred A. Knopf, New York, 1980
- LOUNSBURY, Floyd G., "The Structural Analysis of Kinship Semantics" en Stephen Tyler (ed.), Cognitive Anthropology, Holt Rinehart and Winston, New York, (1964) 1969.
- LUTZ, Catherine "Goals, events and understanding in Ifaluk emotion theory" en Dorothy Holland y Naomi Quinn, Cultural Models in Language and Thoughts, Cambridge University Press, Cambridge, UK, (1983) 1987
- MISNKY, Marvin L. y Seymour A. PAPERT, Perceptrons, MIT Press, Cambridge, Massachusetts, USA, (1969) 1990.
- MISNKY, Marvin, The society of mind, Simon and Schuster, New York, 1986.
- MORRIS, Charles, Signs, Language and Behavior, Prentice-Hall, New York, USA, 1946.
- PIKE, Kenneth, "Towards a theory of the structure of human behavior" en Estudios Antropológicos Publicados en Homenaje al Doctor Manuel Gamio, Mexico, 1956
Selected readings p 107-116, Mouton, La Haya, 1972
- PIKE, Kenneth and Hugh STEVEN, Pike's Perspectives, Credo Publishing Corporation, Langley, British Columbia, Canada, 1989.

Rumelhart, David E., James L McLelland Parallel Distributed Processing II, MIT Press, Cambridge, Massachusetts, USA, 1986.

SCHWARTZ, Jacob T., "El nuevo conexionismo: desarrollando relaciones entre la neurociencia y la inteligencia artificial" en Graubard, Stephen R (ed.), El Nuevo debate sobre la inteligencia artificial, Gedisa, Barcelona, (1988) 1993.

STRAUSS, Claudia y Naomi QUINN , A Cognitive Theory of Cultural Meaning, Cambridge University Press, Cambridge, UK, 1997.

TYLER , Stephen Cognitive Anthropology, Holt Rinehart and Winston, New York, 1969

WALLACE, Anthony y John ATKINS, "The Meaning of Kinship Terms" en Stephen Tyler (ed.), Cognitive Anthropology, Holt Rinehart and Winston, New York, (1960) 1969

UNIVERSIDAD DE BUENOS AIRES
FACULTAD DE FILOSOFÍA Y LETRAS
Dirección de Bibliotecas