Capítulo 5

Amari.

5.1. Contexto.

No está por demás enfatizar el papel clave que jugó el año de 1972 para el área de las memorias asociativas, pues como se indicó en capítulos anteriores, este año fue testigo de una febril actividad de los investigadores pioneros en memorias asociativas.

Uno de los cuatro grandes de la generación 1972, además de James A. Anderson con su *Interactive Memory* (Anderson, 1972), Teuvo Kohonen y sus *Correlation Matrix Memories* (Kohonen, 1972) y Kaoru Nakano y su *Associatron* (Nakano, 1972), es, como ya se mencionó en el capítulo 4, Shun-ichi Amari, profesor de la *University of Tokyo* y uno de los más prolíficos escritores de artículos científicos hasta nuestros días.

En 1972 Amari publicó un trabajo teórico donde continuaba con sus investigaciones sobre las *Self-Organizing Nets of Threshold Elements* (Amari, 1972), tópico cuyo estudio había iniciado un par de años atrás. Este trabajo de Amari constituye una valiosa muestra del estilo teórico característico de este personaje y representa, históricamente, un importante antecedente para la creación de lo que una década más tarde se convertiría en el modelo de memoria asociativa por antonomasia: la memoria Hopfield.

Bosquejos de algunas de las ideas que aparecen en los trabajos de Amari son el tema central de este capítulo.

5.2. El trabajo de Amari.

Mientras que el *Linear Associator* resume la esencia de los trabajos pioneros de Steinbuch. Anderson, Kohonen y Nakano (ver capítulos 1 a 4), el trabajo de Amari apunta en otra dirección: propone métodos y enfoques teóricos novedosos para algunos conceptos conocidos, y pretende establecer nuevas teorías sobre cierto tipo de procesos. entre los que destaca el tema de nuestro interés, la existencia y funcionamiento de las memorias asociativas de patrones.

A diferencia de la *Lernmatrix*. del *correlograph* y del *Linear Associator*, en tanto que modelos factibles de simularse o implementarse casi de inmediato. las *Self-Organizing Nets of Threshold Elements* son estudiadas por Shun-ichi Amari desde un punto de vista teórico (Amari, 1972). Aún más. cinco años después volvió a la carga tomando como base estos resultados, para proponer una teoría neuronal de la asociación y la formación de conceptos (Amari, 1977).

El artículo de 1972 incluye un buen número de definiciones. tres lemas, ocho teoremas y un corolario, parte de los cuales tienen que ver con memorias asociativas. La información que se presenta a continuación está tomada libremente del contenido de este artículo, con las adaptaciones pertinentes y adecuadas en notación y estilo.

Con un enfoque teórico, el autor desarrolla los siguientes *temas*:

1. Una corta descripción de las ya conocidas (en 1972) redes de elementos de umbral. las transiciones de estados y estados de equilibrio

2. Definiciones de los números de estabilidad de la transición de estados y los números de estabilidad de los estados de equilibrio

3. Condiciones de convergencia para los números de estabilidad de los estados de equilibrio

4. Los fundamentos de la auto-organización de las redes de elementos de umbral

5. Las redes auto-organizativas de elementos de umbral como memorias asociativas

Parafraseando al autor, "una red auto-organizativa de elementos de umbral tiene la propiedad de que su estructura varía gradualmente, dependiendo de los patrones aplicados a la red, como estímulos externos. Cuando los patrones de estímulo se aplican repetidamente a la red, se espera que ésta aprenda de los estímulos y fije algunos de ellos como estados de equilibrio estables, por medio de auto-organización. Se puede decir que la red *recuerda* esos patrones. Una vez que un patrón es recordado como un estado de equilibrio estable, será recuperado y reproducido correctamente cuando un patrón vecino se da como estímulo."

La información para el *tema* **1**, cuyo resumen se presenta a continuación, procede de 12 fuentes a las que el autor hace referencia, entre ellas tres trabajos previos de él mismo, un artículo de Krohn & Rhodes de 1965 titulado precisamente *Nets of threshold elements*, y la edición original (1969) del famoso libro *Perceptrons* (Minsky & Papert, 1988).

Un elemento de umbral de n entradas se especifica por n cantidades $w_1, w_2, ..., w_n$ llamados pesos, y una cantidad h llamada umbral. El elemento se denota por $E(w_1, w_2, ..., w_n; h)$ o más brevemente $E(w_i; h)$. Consideremos n variables de entrada $x_1, x_2, ..., x_n$ cuyo dominio es el conjunto $\{-1, 1\}$. La salida y es 1 cuando la suma pesada $\sum_i w_i x_i$ excede h; de otro modo, la salida es -1. Por lo tanto, la relación de entrada-salida del elemento se describe por la siguiente expresión:

$$y = sgn\left(\sum_i w_i x_i - h\right)$$

donde

$$sgn(u) = \begin{cases} 1 & \text{si} \quad u > 0 \\ -1 & \text{si} \quad u < 0 \end{cases}$$

Amari asume, "con el fin de evitar discusiones innecesariamente complicadas", que el caso de igualdad nunca ocurre. Además, es posible hacer una simplificación adicional: si c es una constante positiva, los elementos de umbral $E(w_i; h)$ y $E(cw_i; ch)$ tienen las mismas características, por lo que es válido considerar que $|w_i| \leq 1$ y que además los pesos están normalizados de manera que el valor máximo de todos los $|w_i|$ es precisamente 1.

Sea una red compuesta de n elementos de umbral $E_1, E_2, \ldots, E_n$ donde $E_j = E(w_{j1}, w_{j2}, \ldots, w_{jn}; h)$. Estos elementos están conectados de manera tal, que la salida de E_i, denotada por y_i, está conectada con las i-ésimas entradas de todos los elementos con un retraso de una unidad de tiempo; es decir, la salida y_i se multiplica por w_{ji} cuando se opera con E_j.

La red trabaja de modo síncrono en intervalos fijos de tiempo y está especificada por la matriz de pesos $(w_{ji})_{n \times n}$ y el vector de umbrales $(h_i)_n$
El vector columna

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

donde y_i es la salida actual del elemento E_i, se llama el *estado actual* de la red.

El estado siguiente $\mathbf{y}'$ está determinado unívocamente por el estado actual $\mathbf{y}$, debido a que la salida y_i' del elemento E_i en el tiempo siguiente está determinada por

$$y_i' = sgn\left(\sum_j w_{ij} y_j - h_i\right)$$

Si se denota el operador de transición de estados por T, el estado siguiente $\mathbf{y}'$ se expresa así:

$$\mathbf{y}' = T\mathbf{y}$$

Un *estado de equilibrio* es un estado invariante bajo T; es decir, $\mathbf{y}_e$ es un estado de equilibrio si $\mathbf{y}_e = T\mathbf{y}_e$.

Un conjunto ordenado de m estados $B = \{\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_m\}$ para los cuales se cumple que $\mathbf{y}_{i+1} = T\mathbf{y}_i$ donde $i = 1, 2, ..., (m-1)$ se llama *secuencia de transiciones de estados* de longitud m.

Una secuencia de transiciones de estados $C = \{\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_r\}$ para la cual se cumple que $T\mathbf{y}_r = \mathbf{y}_1$, se llama *ciclo* de longitud r.

Dado que el número de estados en una red es finito, la secuencia de estados $\mathbf{y}_0, T\mathbf{y}_0, T^2\mathbf{y}_0, T^3\mathbf{y}_0, T^4\mathbf{y}_0, ...$ donde $\mathbf{y}_0$ es arbitrario, se comporta de dos posibles maneras:

· converge a un estado de equilibrio,
· o bien, cae en un ciclo dentro de un número finito de transiciones de estado

Los estados de equilibrio y los ciclos son patrones que la red puede *retener persistentemente* sin entradas externas.

Los *temas* **2 a 5** son las aportaciones originales de Amari.

En lo que sigue, y tomando como base la información de los párrafos anteriores, presentaremos algunos comentarios únicamente sobre los contenidos que están relacionados directamente con la capacidad que exhiben las redes de elementos de umbral para actuar como memorias asociativas. Es decir, no habrá comentarios sobre las definiciones, lemas y teoremas relacionados con la recuperación de secuencias y la formación de conceptos, tópicos importantes en el artículo, pero ajenos al tema de nuestro interés.

Con objeto de definir y usar los números de estabilidad mencionados en los *temas* **2 y 3**, Amari echó mano de dos conceptos: el de distancia de Hamming entre dos estados $\mathbf{x}$ y $\mathbf{y}$, expresada como $dis(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\sum_i |x_i - y_i|$, y el de k-vecindad de $\mathbf{x}$, cuya expresión es $N(\mathbf{x}, k) = \{\mathbf{y} \mid dis(\mathbf{x}, \mathbf{y}) \leq k\}$.

A partir de estos conceptos, definió las n funciones

$$u_i(\mathbf{x}, \mathbf{y}) = x_i \left(\sum_j w_{ij} y_j - h_i \right)$$

Es claro que $u_i(\mathbf{x}, \mathbf{y})$ es positiva cuando la i-ésima componente de

$$Ty = \sum_j w_{ij} y_j - h_i$$

coincide con la de **x**. y negativa de otra manera.

La combinación de estas funciones y los operadores $min(k)$, propuestos también por Amari, hace posible llegar a la definición de los números de estabilidad.

Para n números reales $r_1, r_2, ..., r_n$, el operador $min(k)$ denota el $(k+1)$-ésimo número más pequeño de todos los r_i; es decir, si los r_i se ordenan de acuerdo con su magnitud de manera que $r_{i_1} \leq r_{i_2} \leq r_{i_3} \leq \cdots \leq r_{i_n}$, el operador se define así:

$$min(k)\{r_i\} = \begin{cases} r_{i_{k+1}} \text{ si } 0 \leq k < n \\ r_{i_n} \text{ si } k \geq n \end{cases}$$

El primer resultado del artículo, que ilustra el papel de las funciones $u_i(\mathbf{x}, \mathbf{y})$ y del operador $min(k)$, afirma que una condición necesaria y suficiente para que un estado **y** caiga dentro de la k-vecindad de **x** después de una transición simple de estados, es decir $T\mathbf{y} \in N(\mathbf{x},k)$, es $min(k)\{u_i(\mathbf{x},\mathbf{y})\} > 0$.

Un caso particular del resultado anterior es de gran trascendencia; veamos: al hacer $k = 0$, $T\mathbf{y} \in N(\mathbf{x},k)$ se convierte en $T\mathbf{y} \in N(\mathbf{x},0)$ y esto significa ni más ni menos que $T\mathbf{y} = T\mathbf{x}$.

Si además de $k = 0$ hacemos $\mathbf{y} = \mathbf{x}$, entonces obtenemos el primer teorema:

$$\text{El estado } \mathbf{y} \text{ es de equilibrio} \leftrightarrow min_i\{u_i(\mathbf{y},\mathbf{y})\} > 0$$

Con la información y notación anteriores es posible presentar la primera definición central de este artículo, el grado de estabilidad de la transición de estados.

Sea k un número no negativo: el número de k-estabilidad $s(\mathbf{y},k)$ de la transición de estado $\mathbf{y} \rightarrow T\mathbf{y}$ se define como la parte entera de la cantidad $\frac{1}{2}min(k)\{u_i(\mathbf{y},T\mathbf{y})\}$: si $[n]$ representa la parte enera de n, la definción del número de k-estabilidad es:

$$s(\mathbf{y},k) = \left[\frac{1}{2}min(k)\{u_i(\mathbf{y},T\mathbf{y})\}\right]$$

El segundo resultado, que ilustra el significado del número de k-estabilidad $s(\mathbf{y},k)$, afirma que si $\mathbf{z}$ es un estado que pertenece a la $s(\mathbf{y},k)$-vecindad de $\mathbf{y}$, entonces $T\mathbf{z}$ pertenece a la k-vecindad de $T\mathbf{y}$; simbólicamente:

$$[\mathbf{z} \in \mathrm{N}(\mathbf{y},s(\mathbf{y},k))] \rightarrow [T\mathbf{z} \in \mathrm{N}(T\mathbf{y},k)]$$

En particular, $[\mathbf{z} \in \mathrm{N}(\mathbf{y},s(\mathbf{y},0))] \rightarrow [T\mathbf{z} = T\mathbf{y}]$.

Los números de estabilidad de un estado de equilibrio $\mathbf{y}$ se definen mediante el número de k-estabilidad $s(\mathbf{y},k)$; el primer número de estabilidad $s_1(\mathbf{y})$ de $\mathbf{y}$ es definido por $s_1(\mathbf{y}) = s(\mathbf{y},0) = \left[\frac{1}{2}min(0)\{u_i(\mathbf{y},\mathbf{y})\}\right]$, y los demás se definen recursivamente: $s_j(\mathbf{y}) = s(\mathbf{y},s_{j-1}(\mathbf{y}))$.

La secuencia $s_1(\mathbf{y}), s_2(\mathbf{y}), \ldots$ es monótona no decreciente y converge al límite $s(\mathbf{y})$ en un número finito de términos. Al límite $s(\mathbf{y})$ se le llama *número de estabilidad del estado de equilibrio* $\mathbf{y}$.

Un estado de equilibrio $\mathbf{y}$ es *estable* si $s(\mathbf{y}) > 0$ e *inestable* si $s(\mathbf{y}) = 0$. El estado de equilibrio $\mathbf{y}$ será inestable si y sólo si $s_1(\mathbf{y}) = 0$.

El k-ésimo *dominio de estabilidad* de un estado de equilibrio $\mathbf{y}$ es la $s_k(\mathbf{y})$-vecindad $N(\mathbf{y}, s_k(\mathbf{y}))$ y se denota por $D_k(\mathbf{y})$; es decir, $D_k(\mathbf{y}) = N(\mathbf{y}, s_k(\mathbf{y}$

Dos casos particulares importantes son $D_0(\mathbf{y}) = N(\mathbf{y}, 0) = \{\mathbf{y}\}$ y $D(\mathbf{y}) = N(\mathbf{y}, s(\mathbf{y}))$; a $D(\mathbf{y})$ se le llama simplemente *dominio de estabilidad* de $\mathbf{y}$.

Además, debido a que la secuencia $s_1(\mathbf{y}), s_2(\mathbf{y}), \ldots$ es monótona no decreciente y converge a $s(\mathbf{y})$, se cumple que

$$D_0(\mathbf{y}) \subset D_1(\mathbf{y}) \subset D_2(\mathbf{y}) \subset \cdots \subseteq D(\mathbf{y})$$

A continuación se presenta el enunciado completo del teorema 2 del artículo. La importancia de este teorema radica en que establece las condiciones para que se dé la convergencia de la red a los estados de equilibrio, en un número finito de pasos. Un resultado similar es la piedra angular del famoso artículo de J. Hopfield, que marcó el nacimiento de la memoria Hopfield, una década después.

Teorema 2.- Sea **y** un estado de equilibrio. Entonces, la red arriba al estado **y** después de un número finito de transiciones de estado, si su estado inicial pertenece a $D(\mathbf{y})$. Más específicamente, la red arriba al estado **y** en k transiciones si el estado inicial pertenece a $D_k(\mathbf{y})$.♦

En la presentación del *tema* 4, los fundamentos de la auto-organización de las redes de elementos de umbral, Amari asume que los patrones se aplican a la red como estímulos externos, y que un patrón se representa como un vector columna n-dimensional cuyas componentes son ± 1, donde n es el número de elementos de la red.

Se asume también que la red se fuerza a llegar al estado **x** cuando se aplica el patrón estímulo **x**. Sea $\mathbf{x}(t)$ el patrón estímulo aplicado en el tiempo t $(t = 1, 2, ...)$.

Al auto-organizarse, los pesos de la red cambian al recibir los patrones de estímulo, de acuerdo con la siguiente expresión para el incremento de los pesos en el tiempo t:

$$\Delta w_{ij}(t) = -aw_{ij}(t) + bx_i(t)x_j(t)$$

donde a y b son constantes positivas pequeñas.

Así llegamos al *tema* 5, las redes auto-organizativas de elementos de umbral como memorias asociativas.

Se considera que los pesos iniciales de la red auto-organizativa están dados por:

$$w_{ij}^0 = \begin{cases} 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases}$$

Sin pérdida de generalidad, se considera además que $h_i = 0, \forall i$.

El cumplimiento de estas condiciones conlleva al hecho de que todos los estados están inicialmente en equilibrio; esto se concluye de la definición de $u_i(\mathbf{x}, \mathbf{y})$ y del primer teorema:

$$u_i(\mathbf{x}, \mathbf{x}) = x_i \left(\sum_j w_{ij} x_j - h_i \right) = x_i \sum_j w_{ij}^0 x_j = 1$$

y de aquí resulta que $min_i \{u_i(\mathbf{x}, \mathbf{x})\} > 0$.

De acuerdo con el primer teorema, todos los estados $\mathbf{x}$ están inicialmente en equilibrio. Por consiguiente, la red puede retener persistentemente cualquier estado inicial, pero sus números de estabilidad son cero (porque $s_1(\mathbf{x}) = s(\mathbf{x}, 0) = \left[\frac{1}{2} min(0) \{u_i(\mathbf{x}, \mathbf{x})\}\right] = 0$) y todos los estados son inestables.

El asumir que $h_i = 0, \forall i$ tiene como consecuencia inmediata que si se cumple $T\mathbf{x} = \mathbf{y}$ entonces también se cumple $T(-\mathbf{x}) = (-\mathbf{y})$. Dado que el comportamiento de $(-\mathbf{x})$ está determinado por el comportamiento de $\mathbf{x}$, si $\mathbf{x}$ es estable, entonces $(-\mathbf{x})$ también será estable.

Se muestra la auto-organización de la red para el caso en que un solo patrón con ruido se aplica repetidamente. Sea $\tilde{\mathbf{x}}$ un patrón aleatorio cuyas componentes se obtienen cambiando el signo de cada componente de $\mathbf{x}$ en forma inependiente con probabilidad p, es decir:

$$\tilde{x}_i = \begin{cases} x_i & \text{con probabilidad } 1 - p \\ -x_i & \text{con probabilidad } p \end{cases}$$

Se propone el parámetro $\sigma^2 = 4p(1-p)$ como una cantidad que representa el ruido del patrón $\tilde{\mathbf{x}}$. Cuando $p = 0$, no hay ruido, y cuando $p = \frac{1}{2}$, $\sigma^2 = 1$. En este último caso, $\tilde{\mathbf{x}}$ no incluye información original del patrón $\mathbf{x}$.

Amari presenta el comportamiento de las redes auto-organizativas de elementos de umbral como memorias asociativas en los teoremas **4** y **5** de su artículo.

El teorema 4 exhibe el caso en que la red aprende un único patrón de estímulo, y el teorema 5 trata el caso más general, cuando m patrones se aplican repetidamente con cierta frecuencia a la red, y las condiciones para que la red pueda recuperar los patrones asociativamente.

Teorema 4.- Cuando un patrón $\mathbf{x}$ se aplica repetidamente bajo distorsiones de ruido de intensidad σ^2, la red se auto-organiza de manera que se cumple lo siguiente:

· El patrón $\mathbf{x}$ es almacenado en la red como un estado de equilibrio y su número de estabilidad es

$$s(\mathbf{x}) = \left[\frac{1-\sigma^2}{2}n + \frac{\sigma^2}{2}\right]$$

· Aquellos estados $\mathbf{y}$ que satisfacen la relación

$$|cos(\mathbf{x}, \mathbf{y})| < \frac{\sigma^2}{n(1-\sigma^2)}$$

permanecen como estados de equilibrio inestable con número de estabilidad 0. Todos los demás patrones cambian a $\mathbf{x}$ en una transición simple.♦

En la última expresión, $cos(\mathbf{x}, \mathbf{y})$ es el coseno del ángulo entre los patrones $\mathbf{x}$ y $\mathbf{y}$, el cual está dado por $cos(\mathbf{x}, \mathbf{y}) = \frac{1}{n}\mathbf{x} \cdot \mathbf{y}$, y $\mathbf{x} \cdot \mathbf{y}$ es el producto interno.

Se puede decir que la red recuerda a $\mathbf{x}$ correctamente como un estado de equilibrio, para valores de n suficientemente grandes. Consideremos el importante caso en el que se presenta a la red un patrón $\mathbf{y}$ que cumple con

$$|cos(\mathbf{x}, \mathbf{y})| \geq \frac{\sigma^2}{n(1-\sigma^2)}$$

después de que el aprendizaje es completado. La red es puesta en el estado $\mathbf{y}$ con la llegada del patrón externo, pero inmediatamente cambia al estado **x** y permanece en él. En otras palabras, la memoria asociativa recupera el patrón **x** a partir de un patrón ruidoso de entrada **y**.

Se omitirá el enunciado del teorema 5. debido a que su aportación no va más allá de lo exhibido por el *Linear Associator*: si los patrones aprendidos son ortogonales, la recuperación será perfecta.

En su artículo de 1977, Amari toma como base los resultados del artículo de 1972, y propone una teoría neuronal de la asociación y la formación de conceptos (Amari, 1977).

En un estilo teórico aún más riguroso, el autor expone su idea de que los componentes básicos de una red no son elementos simples. sino pequeños grupos de neuronas conectadas entre sí: Amari los llama *neuron pools*; muestra que un neuron pool puede actuar como un elemento de dos estados.

Respecto del tema de memorias asociativas, Amari toma las ideas del artículo de 1972 para definir los sistemas asociativos de manera tal que un patrón de entrada **x** da lugar a un patrón de salida **y** a través de una transformación T: es decir, $\mathbf{y} = T\mathbf{x}$.

Enfatiza la importancia de las relaciones de ortogonalidad para lograr salidas perfectas en la memoria asociativa, pues en el caso ortonormal desaparece el término de *cross-talk*, como ya se había mencionado al presentar el *Linear Associator*.

5.3. Consideraciones finales del capítulo.

En este capítulo hemos presentado el bosquejo de algunas de las ideas que aparecen en los trabajos del notable y prolífico teórico japonés, ingeniero matemático Shun-ichi Amari.

El Profesor Amari ha escrito más de 300 artículos internacionales en varias decenas de revistas de gran prestigio, ha participado en al menos 5 series de libros especializados incluyendo Neural Networks: Research and Applications (Pergamon Press), ha obtenido una docena de premios internacionales de investigación científica y es autor, entre otros muchos, del libro Mathematical Theory of Nerve Nets.

Actualmente es el Director of Brain Information Processing Group, RIKEN Frontier Research Program, en Japón (www.islab.brain.riken.go.jp/~amari/).