Helo Guzmán, José E. *Redes neuronales y autómatas finitos.*
Tecnología en Marcha. Vol. 13 no. 2.

Redes neuronales y autómatas finitos

*José E. Helo Guzmán**

Los autómatas finitos tienen propiedades claramente estudiadas, sus capacidades de reconocimiento de patrones han servido para la construcción de programas de análisis léxico tales como los compiladores.

Al tratar de implantar la inteligencia artificial en un computador es importante garantizar que el nuevo modelo que se pretende utilizar sea al menos tan poderoso como los modelos ya existentes; de ahí la necesidad de encontrar sus vínculos con los modelos conocidos.

En los últimos años las redes neuronales han tomado gran auge como un modelo; el presente artículo tratará de vincular las redes neuronales con los autómatas finitos al demostrar que cualquier autómata finito puede ser sustituido por una red neuronal; como consecuencia de esta sustitución debemos deducir que el modelo de redes neuronales tendrá entonces como mínimo las capacidades de reconocimiento de patrones que poseen los autómatas finitos.

Para el desarrollo de este vínculo se utilizará la teoría clásica de computabilidad y las redes neuronales McCulloch Pitts.

* *Departamento de Computación. Instituto Tecnológico de Costa Rica.*

Redes neuronales y sus vínculos con la neurofisiología

La hipótesis fundamental que sustenta el modelo de redes neuronales se basa en que el funcionamiento del sistema nervioso está mediatizado por el paso de impulsos eléctricos a través de células especializadas llamadas neuronas (2).

Es innegable la existencia de otros elementos además de las neuronas, tal es el caso de las células gliales, las cuales se pensaba hasta hace poco tiempo tenían solo funciones de nutrir a las neuronas; sin embargo se ha descubierto que las células gliales cumplen además funciones tales como las de memorizar; igualmente, la influencia hormonal en las neuronas es otro factor importante. No obstante lo anterior podemos centrar el funcionamiento de nuestro modelo en los impulsos neuronales únicamente.

El funcionamiento de los impulsos neuronales puede visualizarse como un sistema de tres componentes: receptores, red neuronal y efectores. Las neuronas receptoras son las encargadas de recoger los impulsos del medio (impulsos que provienen de los conos de los ojos, nervios del cuerpo, etc.), las neuronas de la red procesan estos impulsos y finalmente las neuronas efectoras producen una salida o efecto (1).

El elemento principal de la formulación es entonces la red neuronal que procesa las entradas y produce salidas; estas neuronas poseen un cuerpo llamado*soma* que es el elemento procesador.

Para efectos de comunicación, la neurona tiene un gran número de filamentos llamados *dendritas* los cuales reciben información, y un único *axón* que es otro filamento por el cual pueden transmitir información. Las uniones entre neuronas del tipo *axón-dendrita* se llaman *sinapsis* y cada sinapsis tiene una fuerza diferente caracterizada por la energía de su conexión (en la naturaleza existen también otros tipos de contactos tales como axón-axón o dendrita-dendrita los cuales aparentemente tienen funciones reguladoras, y que deberemos dejar fuera de nuestra formulación). Una neurona entra en actividad cuando el flujo total de sus entradas (o sea los impulsos comunicados por sus dendritas) superan un cierto umbral, activando el soma y provocando un cambio en su estado hasta la siguiente comunicación; de lo contrario si el número de impulsos totales no supera el umbral, la neurona no se activa y se dice que no llega a su nivel de excitación.

Una vez efectuada una comunicación entre dos neuronas, quedan en el llamado *período refractario* durante el cual la dendrita o axón no son capaces de transmitir nuevamente; debe transcurrir un cierto intervalo de tiempo antes de la nueva transmisión.

El Modelo de McCulloch Pitts

De las consideraciones anteriores podemos extraer el siguiente modelo matemático:

a. Cada neurona o módulo es un elemento que puede tener entradas (x_1, x_2, ..., x_m) y una sola salida "d". Además cada entrada debe tener asignado un peso (W_1, W_2, ..., W_m); finalmente el umbral de excitación de la neurona es "u". La neurona opera sobre una escala discreta de tiempo t=0,1,2,3,... y la activación de la salida "d" en el instante t+1 dependerá de que la suma de las entradas contra el peso en el instante t sea mayor que el umbral; dicho de otra manera:

si en el instante t ($\sum 1$ wi xi) > u entonces en el instante t+1 se activa d y cambia de estado, en otro caso no se excita y queda en el mismo estado.

Si introducimos la notación:

m (t) = 1 si m se dispara en el instante t

m (t) = 0 si m no se dispara en el instante t

entonces la activación de la salida "d" se dará en t+1 si la suma de las entradas en t contra los pesos es mayor que el umbral:

d (t+1) si y solo si $\sum_i w_i x_i (t) > u$

Aquí los pesos positivos $w_i > 0$ son valores que pueden provocar excitación neuronal; mientras que los $w_i < 0$ tienden a producir inhibición neuronal.

b. Una red modular será un grupo de módulos interconectados entre sí, de tal manera que la salida "d" de un módulo se puede ramificar para ser la entrada de varios módulos más.

Una restricción importante es que todos los módulos actuarán bajo la misma escala de tiempos, de tal manera que la red adquiere una armonía poco natural.

Restricciones del modelo McCulloch Pitts

Algunas de las restricciones del modelo con respecto al funcionamiento biológico real son:

a. Suponemos que todas las neuronas trabajan bajo la misma escala de tiempos, es decir que tienen iguales períodos refractarios e iguales capacidades de transmisión. Sabemos que en la naturaleza los períodos refractarios de las neuronas son diferentes, y que, debido a la acción de ciertas sustancias como el potasio, las velocidades de transmisión de las dendritas no son siempre iguales, por lo que la red no tiene una sincronización tan marcada.

b. Los umbrales y pesos de cada neurona no cambian en el tiempo. En las redes verdaderas, los umbrales y pesos dependen del tiempo y de sustancias exógenas a la neurona.

c. Se han desestimado los efectos de hormonas y de productos químicos generados por el cuerpo o bien administrados artificialmente.

d. Excepto por las sinapsis, se han desestimado las otras interacciones entre neuronas. Señalamos que ocurrían interacciones tales como las de axón-axón que no modelaríamos.

e. No se han tomado en cuenta las células gliales.

Autómatas finitos determinísticos

Un autómata finito determinístico se puede definir de la siguiente manera (3):

$M = (K, \Sigma, G, O, s, F)$

donde:

K: es un conjunto finito de estados

Σ: es un alfabeto

$s \in K$: s es elemento de K, es el estado inicial

F está contenido en K: es el conjunto de estados finales

G: es la función de transición, $G: K \times \Sigma \rightarrow K$, toma un par ordenado, un estado y un símbolo del alfabeto para generar otro estado

O: es la función de salida, $O: K \times \Sigma \rightarrow \Sigma$, toma un par ordenado, un estado y un símbolo del alfabeto para generar otro símbolo del alfabeto; en los autómatas de reconocimiento la función de salida es siempre la función identidad (el mismo símbolo de entrada es el de salida), se incluye esta función para esbozar de mejor manera la conexión del autómata finito con las redes neurales.

Con esta definición queda claro que un autómata M cambia de un estado al otro gracias a la función de transición, así si M se encuentra en el estado q y recibe el símbolo a cambiará al estado $G(q, a) \in K$.

Ahora bien, la función del autómata es tomar un símbolo de una lista de símbolos, hasta que acabe con la lista; al terminar el estado final en el que haya quedado el autómata determinará si la lista de símbolos puede ser identificada o no. Como el autómata no puede devolverse sobre la lista de símbolos, el que reconozca o no un caso particular depende de dos factores: el estado en el que se encuentra y los símbolos que faltan por ser procesados. En otras palabras, la configuración de un autómata finito está dada por el par ordenado $K \times \Sigma^*$ (Σ^* significa un grupo de elementos de Σ), por ejemplo: ("q", "ababa") significa que el autómata está en el estado "q" y le falta por procesar la lista de símbolos "ababa".

La relación binaria :- se mantiene si y solo si el autómata M puede pasar de una configuración a otra en un solo movimiento. Así si (q', W') son dos configuraciones diferentes de M, (q,W) :- (q', W') se mantiene si y solo si W= eW' con $e \,\varepsilon\, \Sigma$ y G (q,e)= q'. Definido de esta manera :- es una función de $(K \times \Sigma)$ -> $(K \times \Sigma)$.

Definimos la cláusula transitiva de :- como :-*; que significa la aplicación de :-

una o más veces. Así que decimos que un autómata reconoce una lista de símbolos w Σ^* si y solo si existe una cláusula transitiva tal que lleve (s,w) :-* (q,e) con q Σ F, e lista vacía.

El lenguaje aceptado por M se denomina L (M) y corresponde a todas las posibles listas de símbolos de Σ^* que puede aceptar M.

Reemplazo de un autómata finito por una red neuronal

Reemplazo de un autómata finito determinístico

Debemos establecer primero que cualquier autómata finito debe trabajar sobre una escala discreta de tiempos, de tal forma que si se encuentra en el estado q en el instante t y recibe la entrada "a", cambiará al estado G (q,"a") en el tiempo t +1, y producirá la salida O (q, "a"). Es por lo tanto evidente el hecho de que tanto el autómata finito como la red neuronal (debido a nuestras simplificaciones del modelo biológico) trabajan con escalas discretas de tiempos.

Si nuestro autómata tiene m entradas de i_o, ..., i_{m-1} (en este caso las entradas coinciden con el número de símbolos del alfabeto Σ) y r salidas de O_o, ..., O_{r-1} (las salidas corresponden a la función de salida o, recordamos que en el caso de un autómata finito de reconocimiento las salidas siempre son iguales a las entradas, o sea la función de salida es la identidad o (q,x)=x, y por lo tanto el número de entradas coincide con el número de salidas), entonces, para construir una red modular N a partir de un autómata finito A, N deberá tener también m entradas h_o, ..., h_{m-1} y r salidas p_o, ..., p_{r-1}. De tal forma que debemos asociar la entrada i_j de A con la entrada h_j de N; igualmente definimos la salida o_j de A con la salida p_j de N. La red N deberá tener n*m módulos (donde n corresponde al número de estados que puede tener A, o sea la cardinalidad de k, y r el número de entradas) donde cada módulo se etiquetará (k,j) lo que significa que es el módulo del estado k y la entrada i_j. Igualmente deberá haber m módulos etiquetados (k) que corresponden a la salida o_k.

Lo importante en la construcción de N es la sincronización de los tiempos de tal forma que reproduzcan acertadamente la sincronización de A. Para ello debemos disponer las conexiones de forma tal que el módulo (k,j) se encuentre "on" en el instante t+1, lo cual es posible si y solo si el autómata A se encontraba en el estado k y recibió la entrada i, en el instante t; asimismo el módulo k está "on" en el instante t+1 si y solo si A emite la señal o_k en el instante t; notamos que N producirá la misma salida que A pero una unidad de tiempo después que A, o sea si A produce una salida o en t, N la producirá en t+1; lo anterior debido a que A puede, en la misma unidad de tiempo, cambiar de estado y producir la salida, mientras que N debe efectuar una operación.

Los módulos (k,j) pueden estar entonces en dos posibles estados, "on" y "off" (activados y desactivados), y su funcionamiento es el siguiente al recibir una entrada que supera su umbral y lo excita:

Si el módulo (k,j)= on en el instante t

entonces

produce salida d en t

cambia de estado a "off"

en otro caso

cambia de estado a "on"

Los pares estado-entrada, son aquellos módulos que envían a A a un estado k:

$\{k_1, k_2, \dots k_{n(k)}\} = \{(i,j) \ / \ G\ (i,j) = k\}$

Estos pares estado-entrada se disparan en (t+1) si y solo si el estado de A en el instante (t) es k y la entrada a A es e. Se activan en (t+1) si algún par estado-entrada de A se activa en (t) y recibió una entrada en ese mismo instante o sea:

$h_e(t) \wedge [\, k_1\,(t) \vee k_2(t) \vee \ldots \vee k_{n(n)}(t)]$

Los pares estado-salida que obligan al autómata A a emitir la salida k, se definen como:

$\{k_1, k_2, \ldots, k_{m(k)}\} = \{(i,j) \,/\, 0\,(i,j) = k\}$

En el caso en que 0 sea la función identidad:

$\{k_1, k_2, \ldots, k_{m(k)}\} = \{(i,j) \,/\, 0\,(i,j) = j\}$

Luego, estos módulos de estado-salida producirán una salida en N en el instante (t+1) si y solo si se producía una salida en A en el instante (t), o sea en A:

$K_1\,(t) \vee k_2\,(t) \vee \ldots \vee k_{m(k)}\,(t)$

El anterior razonamiento lleva al esbozo del siguiente teorema (2):

Sea A = (k, Σ, G , O , s , F) con

$\Sigma = \{\, i_o, \ldots, i_{m-1}\}$

o con salidas $\{o_o, \ldots, o_{r-1}\}$

$k = \{s_o, \ldots, s_{q-1}\}$

Entonces existe una red modular N, con subconjuntos

$\{h_o, \ldots, h_{m-1}\}$ de entradas,

$\{p_o, \ldots, p_{r-1}\}$ de salidas,

$\{n_o, \ldots, n_{q*m-1}\}$ de neuronas estado-entrada

tal que si la entrada $\{i_{j1}, i_{j2}, \ldots, i_{jn}\}$ a A que se encuentra inicialmente en el estado s_j, produce la salida $\{o_{k1}, o_{k2}, \ldots, o_{km}\}$; entonces la entrada $\{h_{j1}, h_{j2}, \ldots, h_{jn}\}$ a N, con las neuronas de estado s_j encendidas producirá la salida $\{p_{k1}, p_{k2}, \ldots, p_{km}\}$ con un retardo de una unidad como máximo.

La demostración del teorema anterior es simplemente reproducir el proceso que hemos descrito. Hasta aquí hemos esbozado teóricamente que el comportamiento de un autómata finito determinístico puede siempre reproducirse por medio de una red neuronal; o sea que puede sustituirse cualquier autómata finito por una red neuronal.

Reemplazo de un autómata finito no determinístico

Los autómatas finitos no determinísticos tienen la habilidad de cambiar de estados de forma tal que el estado siguiente es determinado parcialmente por el estado actual y por el símbolo leído. Se permite un grupo posible de estados siguientes para una combinación de un estado dado y símbolo leído. A pesar de que los autómatas no determinísticos son una extensión interesante de la teoría, en la práctica no aportan realmente más poder computacional.

Puede probarse matemáticamente que, para cualquier autómata no determinístico, existe un autómata determinístico equivalente; donde equivalente significa que ambos autómatas son capaces de reconocer el mismo lenguaje; o sea que si M1 es un autómata no determinístico y M2 es su equivalente determinístico L(M1) = L (M2). Está fuera del alcance de este artículo esbozar dicha prueba por lo que referimos a Lewis (3), para hallar la demostración de la misma.

Por lo anterior se puede deducir que cualquier autómata no determinístico M1, puede convertirse a un autómata determinístico M2, y M2 puede ser reemplazado por una red neuronal como la descrita en el capítulo anterior.

Ejemplos de reemplazos de autómatas finitos a redes neuronales

Ejemplo 1

Consideremos el siguiente autómata finito M, que reconoce el lenguaje

L (M) = {w/w Σ {a,b} * y tiene un número par de b' es:

M = (k, Σ, G, O, s, F) donde

K = (q0, q1)

Σ = {a,b}

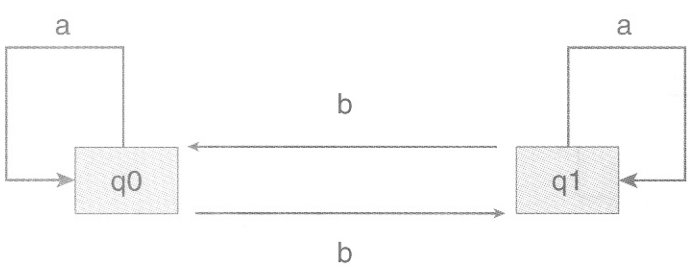
s = $\{q^0\}$

F = $\{q^0\}$

Las funciones de transición y salida se expresan en el cuadro siguiente:

q	e	G (q,e)	O (q,e)
q^0	a	q^0	a
q^0	b	q^1	b
q^1	a	q^1	a
q^1	b	q^o	b

Gráficamente podemos visualizar el autómata como:



De acuerdo con lo expuesto en el teorema anterior, notamos que M tiene dos posibles entradas y dos posibles salidas (definidas en Σ, el alfabeto del autómata); de igual manera nuestra Red Neuronal N, tendrá dos entradas y dos salidas.

El número de módulos es n*r (n número de estados, r número de entradas) o sea 2*2 = 4. Cada uno de ellos lleva la tiqueta (k,j) (estará "on" en t+1 si recibe entrada ij en t); así los cuatro módulos que constituyen la red serán: (qo, a), (qo, b), (q1, a) (q1, b). Los módulos de salida serán dos, pues la función de salida es de dos estados.

Los pares estado-entrada nos mostrarán cómo deben ir cambiando en el tiempo los estados de "on" y "off" de los módulos:

q0 = {(q0, a), (q1, b)}

q1 = {(q1, a), (q0, b)}

Con las anteriores consideraciones la red neuronal que representa el autómata M será la que se muestra en la Figura 1.

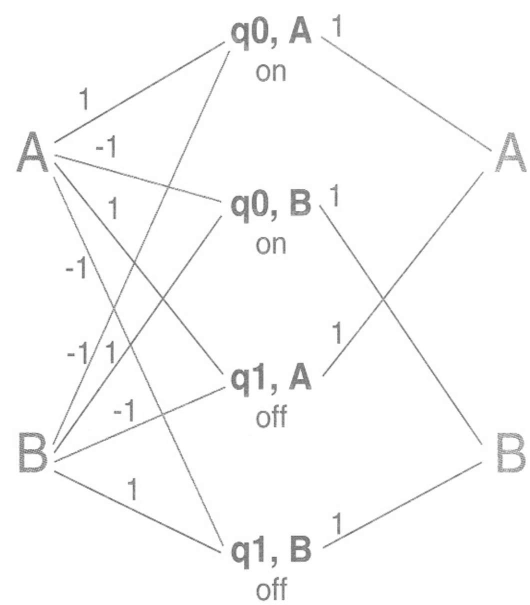


Figura 1. *Red neuronal que representa el autómata M.*

Ejemplo 2

Consideremos ahora el autómata finito M que reconoce el lenguaje:

L (M) = {w/w Σ {a,b} * y w no tiene 3 b' es seguidas}

Tenemos:

M = (k, Σ, G, O, s, F) donde

K = [q0, q1, q2, q3]

Σ = {a, b}

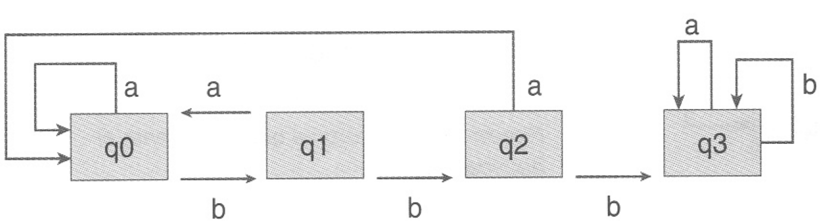
s = q0

F = [q0, q1, q2]

Las funciones de transición y salida son:

q	e	G (q, e)	O (q,e)
q0	a	q0	a
q0	b	q1	b
q1	a	q0	a
q1	b	q2	b
q2	a	q0	a
q2	b	q3	b
q3	a	q3	a
q3	b	q3	b

Gráficamente podemos visualizar el autómata como:



De acuerdo con lo expuesto en el teorema, notamos que M tiene dos posibles entradas y dos posibles salidas (definidas en Σ, el alfabeto del autómata), de igual manera nuestra Red Neuronal N, tendrá dos entradas y dos salidas.

El número de módulos será n*r (n número de estados, r número de entradas) o sea 4*2=8. Cada uno de ellos llevará la etiqueta (k,j) (estará "on" en t+1 si recibe entrada 1j en t); así los cuatro módulos que constituyen la red serán: (q0, a), (q0, b), (q1, a), (q1, b), (q2, a), (q2, b), (q3, a), (q3, b). Los módulos de salida serán dos, pues la función de salida es de dos estados.

Los pares estado-entrada nos mostrarán cómo deben ir cambiando en el tiempo los estados de "on" y "off" de los módulos:

q0 = {(q0, a), (q1, a), (q2, a)}

q1 = (q0, b)

q2 = {(q1, b)}

q3 = {(q2, b), (q3, a), (q3, b)}

La red neuronal que representa el autómata M se muestra como Figura 2.

Por claridad en el gráfico se han omitido las conexiones de B; en el caso de los nodos de entrada hacia las neuronas, las conexiones son inversas a las de A, o sea donde A se conecta con 1, B se conecta con -1; en el caso de las conexiones que van de las neuronas a los nodos de salida, solamente tienen conexiones las neuronas de la forma (qx,B) con un peso de 1 dirigidos hacia el nodo de salida B.

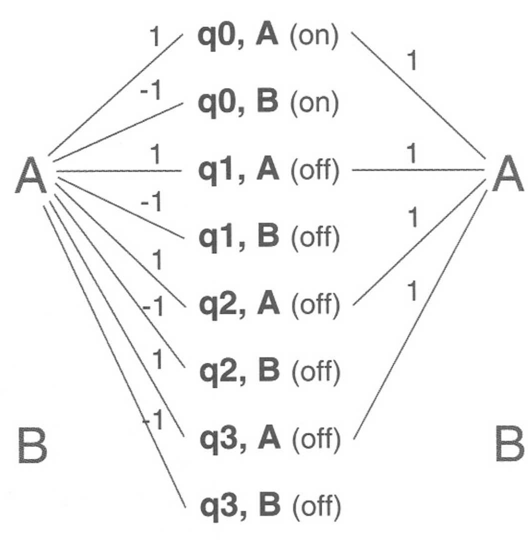


Figura 2. *Red neuronal N.*

Conclusiones

El modelo McCulloch Pitts se basa en una simplificación del modelo biológico de las neuronas.

Los autómatas finitos tienen características bien conocidas y demostradas que los convierten en excelentes modelos de reconocimiento de expresiones regulares.

El artículo ha querido demostrar el vínculo existente entre estos dos modelos, de tal forma que cualquier autómata finito M puede ser reemplazado por una Red Neuronal N, teniendo la red N iguales capacidades de reconocimiento de M, es decir N es capaz de reconocer el lenguaje de M, L (M) con igual precisión.

De lo anterior se puede deducir que los modelos de redes neuronales tienen por lo menos capacidades de reconocimiento de patrones iguales a los autómatas de estado finito.

La deducción anterior permite situar a las redes neuronales como modelos de alto interés para el reconocimiento de patrones, y vincula esta nueva tendencia con la teoría clásica de la computación.

Literatura citada

1. Wolf, Gerald. ***Neurobiología: estudio básico de psicofisiología.*** 1978.
2. Arbib, Michael A. ***Cerebros, máquinas y matemáticas.*** 1976.
3. Lewis, Harri R.; Papadimitriou, Christos H. ***Elements of the theory of computation.*** 1982.
4. Solano, Yadira. ***Neural networks.*** 1986.
5. Buhman, J. Schulten, K. ***Associative recognition and storage in a model network of physiological neurons.*** IEEE, 1986.
6. Wasserman, Phillip D. *Neural networks.* ***Communications*** IEEE, 1987.