

Red neural artificial del sistema mesolímbico-cortical que simula el aprendizaje discriminativo y de inversión

Guevara M A*
Hernández González M*
Olvera Cortés M E**
Robles Aguirre F A***

* Instituto de Neurociencias, Universidad de Guadalajara, Jalisco, México.

** Centro de Investigaciones Biomédicas de Michoacán – IMSS, México.

*** Centro Universitario del Norte de la Universidad de Guadalajara, Jalisco, México.

Correspondencia:

Francisco Abelardo Robles Aguirre
Centro Universitario del Norte de la Universidad de Guadalajara.
Carretera Federal No. 23, km 191, 46200, Colotlán, Jalisco, México.
Tel.: (52) (499) 992 13 33
E-mail: franciscoara@live.com

Artículo recibido: 08/julio/2011

Artículo aceptado: 21/mayo/2012

Este artículo puede ser consultado en versión completa en: <http://www.medigraphic.com/ingenieriabiomedica>

RESUMEN

En el presente estudio se desarrolló un modelo conexionista de aprendizaje no supervisado para simular una tarea de discriminación en un número reducido de iteraciones sin entrenamiento previo a la tarea. La red neural artificial consideró características neurofisiológicas de algunas de las estructuras del sistema dopaminérgico mesolímbico amígdala (AMG), corteza orbitofrontal (COF), área tegmental ventral (ATV) y núcleo accumbens (ACC). El modelo generó respuestas similares a las emitidas por ratas macho durante una tarea de discriminación e inversión en un laberinto T, usando como recompensa un incentivo sexual. En la actividad de las estructuras simuladas, se hallaron los fenómenos de preferencia de reforzador e inversión de la preferencia durante la inversión del incentivo en ACC y ATV; una codificación temprana en AMG; además de una codificación retardada y aumento en el reclutamiento de nodos neurales en la COF ante la inversión. Por último, las estructuras de salida mostraron una actividad de expectativa anterior a la entrega del reforzador.

Palabras clave: Aprendizaje sin entrenamiento previo, aprendizaje de inversión, redes neurales artificiales, sistema mesolímbico cortical, corteza orbitofrontal.

ABSTRACT

The present study develops a connectionist neural network with unsupervised learning rules to simulate a discrimination task in a reduced number of time steps without previous training. The design of the network took into account some neurophysiological findings of dopaminergic mesolimbic system from structures like amygdala (AMG), orbitofrontal cortex (COF), ventral tegmental area (ATV) and nucleus accumbens (ACC). The proposed model generated similar responses to those from male rats during a discrimination and reversal learning tasks in a T maze, using sex as reward. In the activity of simulated structures different phenomena were found, like reinforcement preference and its reversal during reversal learning phase in ACC and ATV. It was also found an early encode in AMG, besides a retarded encoding and an increase in recruitment of neural nodes in COF during reversal learning. All output structures showed an expectancy activity before reinforcer delivery.

Key words: Non previous-learning, reversal learning, artificial neural networks, mesolimbic system, orbitofrontal cortex.

INTRODUCCIÓN

En los últimos años, el desarrollo de redes neurales artificiales (RN) que utilizan un paradigma de aprendizaje no supervisado ha recibido particular atención¹⁻⁹, sobre todo buscando economía en el entrenamiento. Donahoe, Burgos y Palmer^{10,11}, por ejemplo, simularon el aprendizaje pavloviano y skinneriano en un número limitado de ensayos considerando las vías dopaminérgicas involucradas. Armony, Servan y LeDoux¹², en un modelo que consideró en mayor detalle las conexiones entre regiones cerebrales como el tálamo, la corteza auditiva y la AMG, consiguieron emular las respuestas de roedores durante una tarea de condicionamiento aversivo, reduciendo el número de ensayos requeridos a un solo dígito. Deco y Rolls¹³ fueron aún más lejos al lograr simular el cambio en la aplicación de una regla de respuesta en monos, con tan sólo un ensayo, al incorporar en su modelo características electrofisiológicas de la respuesta neuronal de la COF. Así, un número cada vez mayor de modelos de RN han incorporado características neurofisiológicas e incorporado datos del procesamiento de ciertas estructuras neuronales para emular conductas complejas como aprendizaje o toma de decisiones^{1,2,4,6,8,9}. De este modo, parece evidente que la inclusión de variables anatomofisiológicas en el diseño de la arquitectura, podría reducir las iteraciones requeridas comúnmente en una simulación computacional al incorporar recursos de un proceso de selección natural que ha moldeado filogenéticamente a un sistema determinado para cumplir con las tareas que le permitan resolver un problema concreto^{1,2}. Una de las tareas que tradicionalmente se ha considerado como paradigma de complejidad es el aprendizaje discriminativo, que involucra la asignación de valor incentivo (v) a estímulos novedosos para convertirlos en estímulos asociados a la obtención de recompensas (E^D) o a su no obtención (E^A); y su posterior inversión, en la cual el estímulo que antes era reforzado deja de serlo, mientras que aquel que en los ensayos anteriores no era reforzado, se torna ahora asociado con la entrega de reforzador (E^I)¹⁴⁻¹⁹. Para simular la solución de esta tarea se diseñó un modelo computacional conexionista que incorporó un conjunto de módulos de nodos neurales que consideraron algunas características de conectividad del sistema mesolímbico, el cual se ha descrito participa en los procesos motivoemocionales de las conductas dirigidas a una meta^{14,15,17-19}. De esta manera, tomando como base los datos conductuales de un

experimento previamente realizado²⁰, este estudio pretendió emular las respuestas de ratas macho durante el aprendizaje discriminativo inicial y el aprendizaje de inversión en un laberinto T usando como reforzador un incentivo sexual, intentando reducir el número de iteraciones requeridas para completar tareas de aprendizaje complejo al incorporar tales características anatomofisiológicas y evitando el costoso entrenamiento previo a la tarea por parte de las RNs.

METODOLOGÍA

Arquitectura de la red

Para Schoenbaum, Saddoris y Stalnaker²¹, la hipótesis más adecuada para dar cuenta del fenómeno de aprendizaje de inversión consiste en considerar que la COF facilita el aprendizaje asociativo en otras estructuras al participar en el mapeo entre una recompensa esperada y una recibida. De este modo, se ha comprobado que estructuras como la AMG, a través de sus densas conexiones con la COF es crítica para la asociación rápida propia del aprendizaje discriminativo²². Del mismo modo, se ha propuesto que esta facilitación por parte de la OFC es indispensable para la generación de señales directrices de tipo dopaminérgico (DA) que se propagan a otras estructuras como el estriado ventral²¹. De esta manera, partiendo de la evidencia experimental y a semejanza de otros modelos que han considerado la importancia de estructuras como la AMG^{12,13} y de los ganglios basales³ en tareas de aprendizaje asociativo, se construyó por computadora, en el programa Delphi 5.0 de Borland, un modelo de red neural artificial con una arquitectura como la mostrada en la Figura 1.

El modelo presentado aquí, contaba con cuatro unidades de entrada para la detección de estímulos con posibilidades de convertirse en E^I , representando vías de entrada sensorial de refuerzo positivo y negativo hacia el ATV. Dos nodos de ATV que permitían una representación del E^I distribuida hacia los módulos o conjuntos de nodos ocultos. Se programaron tres módulos con nodos ocultos, los cuales fueron AMG, que contó con cuatro nodos; COF, con seis nodos distribuidos en dos capas; y ACC, que consistía en dos capas, con dos nodos cada una, con vías de retroalimentación hacia ATV y a COF. Los dos nodos de la segunda capa de ACC, o de su capa de salida representaban además la activación que definiría la decisión de la red de dirigirse hacia la derecha o hacia la izquierda.

Por último, se colocaron 30 unidades de entrada para la detección de estímulos con probabilidades de convertirse en E^D o E^A . Las conexiones entre los nodos se clasificaron en fijas (con valor de 1.0), modificables (ajustables a la función de aprendizaje), dopaminérgicas (DA) (i.e. de efecto modulador o variable) e inhibitorias (valor de -0.25) (Figura 1).

Se introdujeron conexiones fijas únicamente del nodo E^I hacia los nodos de ATV que guardaban una relación topográfica con la decisión de dirigirse hacia la derecha o hacia la izquierda.

Por otra parte, se programaron conexiones anterógradas modificables de todos los nodos de entrada hacia los nodos de AMG, y de éstos hacia la primera capa de ACC y hacia COF. También de la capa de entrada de COF hacia su capa de salida, y de ésta a la primera capa de ACC, de acuerdo a lo descrito en la literatura²³. Finalmente, de la capa de entrada de ACC hacia la capa de salida.

Las conexiones DA, igualmente modificables, se establecieron de ATV hacia todos los nodos de AMG, COF y ACC considerando lo propuesto en la literatura²⁴⁻²⁸ (a la capa de salida de ACC, las conexiones guardaron relación 1:1 para sostener la relación topográfica con respecto a la decisión de ir a la derecha o a la izquierda; ver parte superior de Figura 1).

Además, se colocaron conexiones inhibitorias recíprocas entre los nodos de AMG (que se ha descrito como una propiedad de las neuronas de esta estructura^{26,27}), entre los del ATV así como entre los de la capa de salida de ACC^{29,30}.

Por último, se programaron conexiones retrógradas desde la salida de ACC hacia COF y ATV (aunque estas últimas serían inhibitorias, para permitir un control de retroalimentación de la señal DA a partir de las respuestas, como se ha reportado en diversos estudios^{17,24-27}), y de los nodos de la capa de salida de COF hacia ATV y AMG (también ampliamente descritas^{25,26}).

FUNCIONES DE APRENDIZAJE Y ACTIVACIÓN

La función de aprendizaje aplicada a las conexiones modificables de la red puede observarse en (1):

$$\omega_{ij}(t) = (\beta\omega_{ij}(t-1)) + [\delta(\alpha_j(t) \alpha_i(t))] \quad (1)$$

donde: ω_{ij} = es el peso de la conexión entre las unidades i y j ; β = es el coeficiente de sostenimiento de la actividad, igual a 0.95 (como en Donahoe,

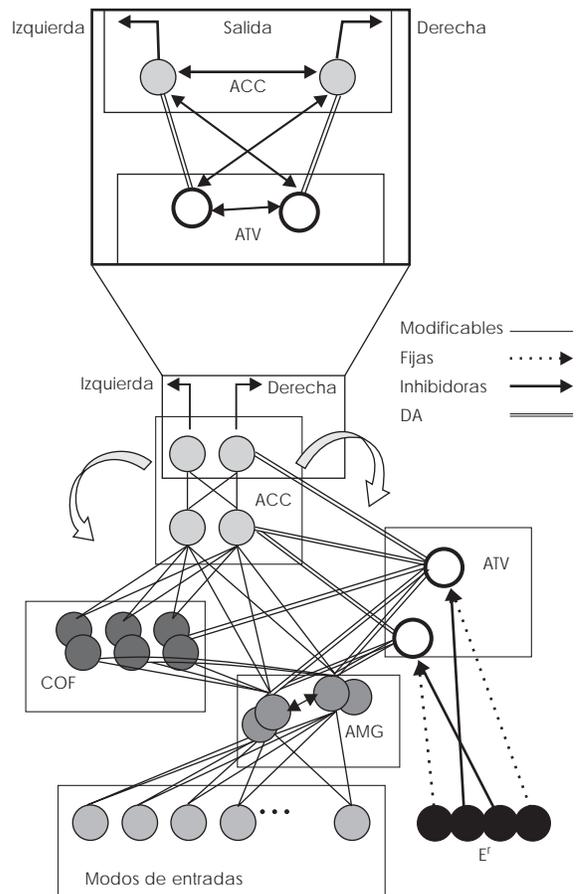


Figura 1. Esquema de la arquitectura de la red neural. No se presentan todos los nodos de entrada ni todas las conexiones entre los distintos módulos. Las conexiones retrógradas entre los nodos de salida de ACC hacia ATV y COF están indicadas por las flechas curvas en color blanco. En la parte superior puede verse un esquema más detallado de las conexiones inhibitorias y excitatorias entre ATV y ACC que permiten retroalimentar la respuesta de decisión para dirigirse hacia la derecha o a la izquierda en el laberinto. E^I : reforzador; AMG: amígdala; COF: orbitofrontal; ATV: área tegmental ventral; ACC: accumbens.

Burgos y Palmer¹⁰); δ = coeficiente de adquisición para las conexiones y será igual a 0.1; $\alpha_j(t)$ = activación del nodo j en un tiempo determinado t .

Los valores de activación de los nodos de entrada del modelo fueron de codificación binaria (0, 1) y su asignación se realizó con base en la posición de la rata virtual dentro del laberinto T , como se muestra en (2).

$$\alpha_x(t) \begin{cases} \text{Si } x = s, \text{ entonces } \alpha_x(t) = 1 \\ \text{Si } x \neq s, \text{ entonces } \alpha_x(t) = 0 \end{cases} \quad (2)$$

donde: $s = \{1..30\}$ es el valor indicador de la posición de la red neural en una de las 30 posiciones que puede ocupar en el laberinto; $x = \{1..30\}$ es el indicador del nodo de entrada que se corresponde con las posiciones en el laberinto.

Ahora bien, respecto a los nodos ocultos, la función de activación aplicada puede verse en (3), seguida de las ecuaciones que definen la función sigmoide de la entrada neta (4) y su valor bruto en cada nodo (5).

$$\alpha_{j(t)} \begin{cases} \text{Si } \text{net}_{j(t)} < \theta, \text{ entonces } \alpha_{j(t)} = \psi \\ \text{Si } \text{net}_{j(t)} \geq \theta, \text{ entonces } \alpha_{j(t)} = \\ \quad [\alpha_{j(t-1)}(1-a)] + a[f\text{net}_{j(t)}] \end{cases} \quad (3)$$

$$f\text{net}_{j(t)} = 1/[1 + \exp[(-1)(\text{net}_{j(t)}-0)]] \quad (4)$$

$$\text{net}_{j(t)} = \sum (\omega_{ij} \alpha_{i(t)}) \quad (5)$$

donde: θ = umbral del nodo neural asignado aleatoriamente a través de una curva Gaussiana con parámetros $\mu = 0.05$ y $\sigma = 0.025$ (parámetro semejante a Donahoe, Burgos y Palmer¹⁰); ψ = estado de activación basal de los nodos asignado aleatoriamente a través de una curva de probabilidad gaussiana con parámetros iguales a los de θ ; a = tasa de decremento de 0.75 (parámetro similar al utilizado en Donahoe, Burgos y Palmer¹⁰); ω_{ij} = peso de la conexión entre las unidades i y j .

Finalmente, para los nodos de salida en ACC, la función de activación aplicada se muestra en (6). En tanto que en (7) se presenta la ecuación que define el cálculo del valor de entrada DA neta (proveniente de ATV) a cada nodo de salida. Este valor se suma a los valores de entrada provenientes de otros nodos no DA, como se ve en (8).

$$\alpha_{j(t)} \begin{cases} \text{Si } \text{net}_{\text{TOT}j(t)} < \theta, \text{ entonces } \alpha_{j(t)} = (\psi)(\text{net}_{\text{jatv}(t)}) \\ \text{Si } \text{net}_{\text{TOT}j(t)} \geq \theta, \text{ entonces } \alpha_{j(t)} = [\alpha_{j(t-1)}(1-a)] + a[f\text{net}_{\text{TOT}j(t)}] \end{cases} \quad (6)$$

$$\text{net}_{\text{jatv}(t)} = \sum (\omega_{ij\text{atv}} \alpha_{i\text{atv}(t)}) \quad (7)$$

$$\text{net}_{\text{TOT}j(t)} = \text{net}_{\text{jatv}(t)} + \text{net}_{j(t)} \quad (8)$$

donde: $\alpha_{i\text{atv}(t)}$ = estado de activación del i -ésimo elemento de ATV; $\omega_{ij\text{atv}}$ = peso de la conexión en-

tre las unidades i y j , siendo i un elemento de ATV; $f\text{net}_{\text{TOT}j(t)}$ se define igual que en (4), utilizando $\text{net}_{\text{TOT}j}$ en lugar de net_{j} .

PARADIGMA EXPERIMENTAL

Las tareas básicas que la red neural realizó fueron la discriminación inicial y la inversión de ésta. Cada una de las fases de la tarea se compuso de siete ensayos. La discriminación debía ocurrir entre los dos brazos de un laberinto T a fin de obtener el E' que se encontraba en uno de ellos en cada ensayo. Al finalizar la discriminación, el E' se cambiaba al otro brazo del laberinto y se procedía a evaluar la inversión del aprendizaje. El laberinto se dividió en 30 segmentos, cada uno representaba una posible posición en el laberinto T, que la red podía ocupar. Los primeros 10 segmentos se ubicaron en el corredor principal del laberinto y el procesamiento de las señales de entrada de la red posibilitaba el que continuara avanzando o se detuviera en ese segmento, de acuerdo con (9):

$$s(t) \begin{cases} \text{Si } \alpha_{\text{ACC}(t-1)} < 0.40605, \text{ entonces } s(t) = s_{(t-1)} + 1 \\ \text{Si } \alpha_{\text{ACC}(t-1)} \geq 0.40605, \text{ entonces } s(t) = s_{(t-1)} \end{cases} \quad (9)$$

donde: $s(t)$ = segmento del laberinto a ocupar; $s_{(t-1)}$ = segmento del laberinto ocupado durante la iteración anterior y $\alpha_{\text{ACC}(t-1)}$ = estado de activación del nodo de ACC. Los valores asignados numéricos en (9) se asignaron tomando en cuenta la activación basal inicial descrita en (3) y el valor máximo de activación neural que estos nodos podían alcanzar (e. g. Figura 2). De este modo, el número asignado se encuentra muy por encima del umbral de disparo y de la activación inicial de los nodos, simulando el procesamiento iterativo que deben realizar las neuronas de esta estructura para cargar el programa motor necesario para realizar el avance.

Al llegar al segmento 10, la red adicionaba otra opción de respuesta: la elección del brazo. Los siguientes 20 segmentos eran consecuencia de la elección de brazo, del 11-20 para el brazo derecho, y del 21-30 para el izquierdo, en concordancia con (10):

$$s(t) \begin{cases} \text{Si } (\alpha_{\text{ACC}j(t-1)} < \alpha_{\text{ACC}i(t-1)}) \text{ y } (\alpha_{\text{ACC}i(t-1)} > 0.4061) \\ \text{entonces } s(t) = 11 \end{cases} \quad (10)$$

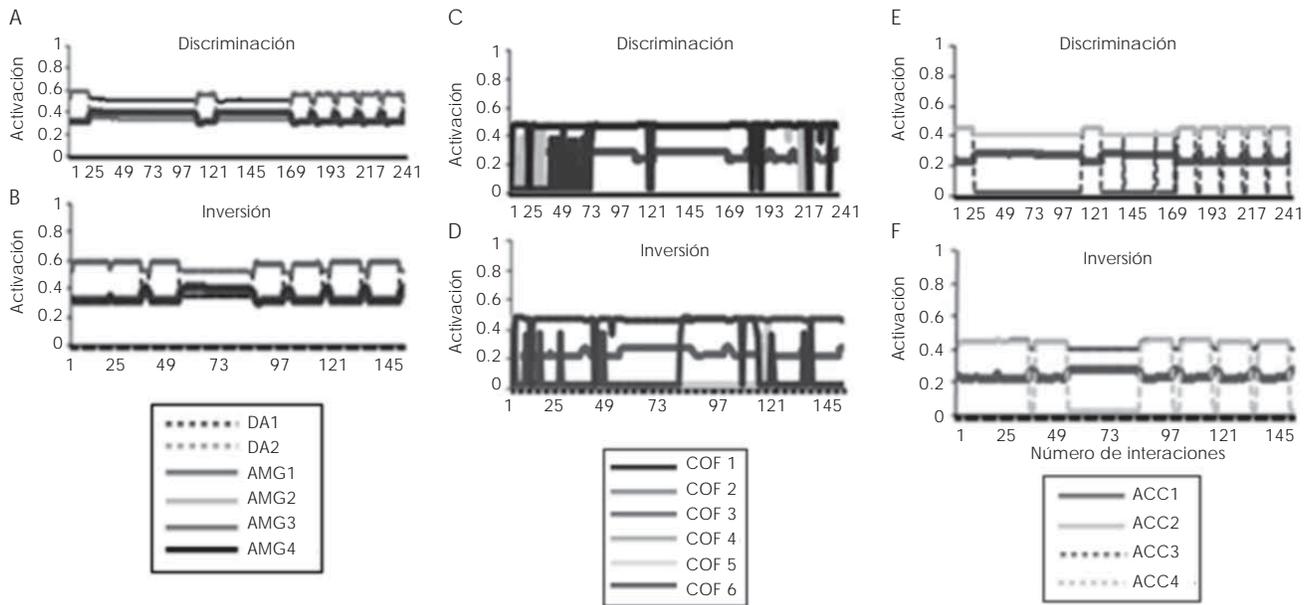


Figura 2. A y B, activación de nodos de ATV y AMG. C y D, activación de nodos de COF. E y F, activación de nodos de ACC.

$$\text{Si } (\alpha_{ACCj(t-1)} > \alpha_{ACCi(t-1)}) \text{ y } (\alpha_{ACC(t-1)} > 0.4061) \\ \text{entonces } s_{(t)} = 21$$

en donde: $\alpha_{ACCj(t-1)}$ = estado de activación del nodo del j-ésimo nodo en la capa de salida de ACC en la iteración anterior y $\alpha_{ACCi(t-1)}$ = estado de activación del i-ésimo nodo en la capa de salida de ACC durante la iteración anterior. El valor numérico que sirve como umbral para la toma de decisión se asignó considerando que debería ser ligeramente mayor que el valor en (9), dado que la decisión requiere un factor adicional de procesamiento a la puesta en marcha del programa motor.

PROCEDIMIENTO

Se realizaron 12 simulaciones de la red, suponiendo que cada simulación equivalía a un sujeto virtual desarrollando la tarea. Cada ensayo fue dividido en 100 momentos como máximo número de iteraciones para dar una solución. A la ubicación espacial de la rata en uno de los segmentos del laberinto T correspondía la activación del nodo de entrada pertinente (Figura 1). Cuando la elección del brazo era la correcta, los nodos de E' se activaban, de manera correspondiente, a partir de la posición 17 para el brazo derecho, o de la posición 27 para el izquierdo, y permanecían activos hasta finalizar el recorrido o completar las 100

iteraciones, lo que ocurriera primero. Si el brazo elegido era el incorrecto, el ensayo terminaba al llegar a la última posición del brazo incorrecto o hasta terminar las 100 iteraciones. Se registró el número de respuestas correctas, incorrectas y de ensayos sin respuesta de las 12 simulaciones, y se comparó con el número de respuestas correctas, incorrectas y ensayos sin respuesta de los 12 sujetos del experimento llevado a cabo por Guevara, Robles-Aguirre, Quirarte y Hernández-González²⁰, a través de un análisis de varianza para cada uno de los tipos de respuesta.

RESULTADOS

Fenómenos fisiológicos emulados

En lo concerniente a algunos de los fenómenos de la fisiología celular de las estructuras simuladas durante el aprendizaje discriminativo y su inversión, presentamos los efectos más significativos, acordes con la literatura, encontrados en los niveles de activación, primeramente, y los relativos a las modificaciones en el valor de los pesos, posteriormente, calculados en ambas fases.

En la Figura 2 se presentan los valores de activación de los diversos nodos de una de las redes intactas evaluadas. En las gráficas A y B se muestra la actividad de los nodos de AMG y ATV durante la discriminación y la inversión.

En ambas se evidencia el procesamiento de los estímulos que ingresan a la red durante las dos fases, con picos de actividad indicativos de la presentación del E^r , y bajas abruptas de actividad posteriores a su entrega. Mientras que en C y D observamos la activación de los nodos de COF. Su comportamiento tiende a mesetas de actividad creciente durante los primeros ensayos de la discriminación para los distintos nodos, situación que se ve alterada en la inversión, donde algunos nodos tienden a disminuir su activación. En ambas fases, su activación parece marcada por la aparición de picos de actividad coincidentes, de nuevo, con la expectativa de presentación del E^r y el flujo DA concomitante.

En la parte inferior, gráficas E y F, se muestra la actividad de los nodos de ACC. Podemos observar, nuevamente, un patrón regular oscilante que se instaura después de varios ensayos, tanto en la discriminación como en la inversión, y que parece codificar la expectativa de aparición del E^r asociada con la presentación del E^D y la baja de actividad propiciada por la entrega del E^r . Puede observarse la inversión de la actividad en los nodos de salida de ACC propia de la inversión en la respuesta hacia la derecha o hacia la izquierda. Esta actividad oscilante se mantiene durante breves iteraciones, pero se incrementa rápidamente una vez que inicia de nuevo el recorrido por el laberinto de las RN.

En lo referente a los valores de los pesos de las conexiones, se identificó un incremento más rápido de los valores de los pesos de las eferencias de AMG hacia ACC que hacia COF, tanto en la discriminación como en la inversión (Figura 3 A y B).

En la inversión del aprendizaje, se encontró un incremento en el reclutamiento de conexiones entre AMG y COF, similar a lo reportado en diversos estudios^{16,18}.

De manera parecida, las eferencias de COF hacia los nodos de salida de ACC y hacia la propia salida de COF se incrementaron mucho más dependiendo de su asociación con la entrega de E^r , pero únicamente durante la discriminación (Figura 3C).

Durante la inversión no se presentó esta disparidad en los valores de los pesos dependiente de su asociación con la entrega de E^r , evidenciando un sostenimiento de los valores alcanzados durante la discriminación anterior y un reclutamiento de nuevas conexiones para dar cuenta de la nueva discriminación implicada en la inversión^{16,18} (Figura 3D).

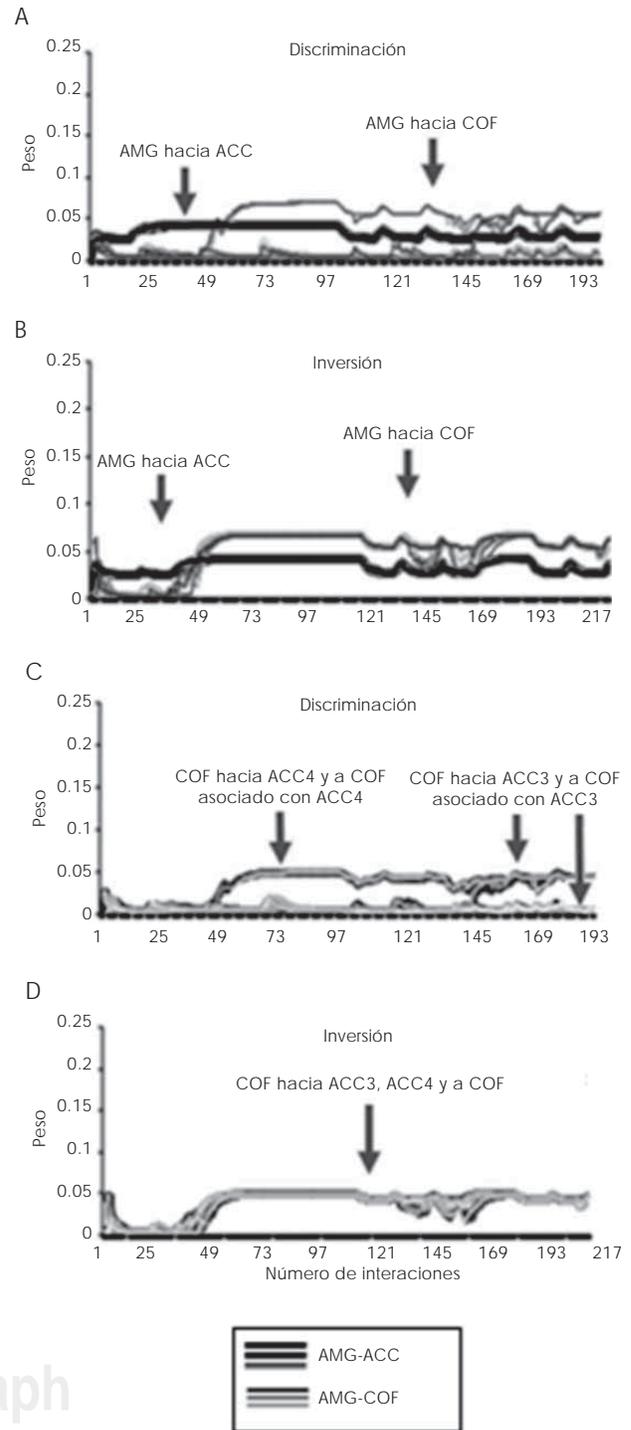


Figura 3. Valores de algunos pesos sinápticos modificables de las estructuras simuladas en el transcurso de las dos fases. En A y B se pueden ver los valores de los pesos sinápticos de las proyecciones desde el módulo de AMG hacia COF y hacia ACC durante la discriminación y la inversión. En C y D se muestran los valores de los pesos sinápticos de las proyecciones de COF hacia la salida de COF y hacia el ACC. ACC3: nodo que codifica la asociación de la elección izquierda/derecha con la entrega de E^r ; ACC4: nodo que codifica la elección de la respuesta incorrecta.

EMULACIÓN DE LA RESPUESTA

Finalmente, el modelo buscó la emulación de las respuestas de discriminación e inversión de ratas tratadas con solución salina (grupo SS, $n = 12$) de un estudio experimental realizado con anterioridad²⁰. Se encontró que los sujetos virtuales obtuvieron un número de respuestas correctas (discriminación, $\bar{x} = 4.8$; inversión, $\bar{x} = 5.25$) superior a las incorrectas (discriminación, $\bar{x} = 0.75$; inversión $\bar{x} = 1.75$) y un número bajo de ausencias de respuesta (discriminación $\bar{x} = 0.91$; inversión $\bar{x} = 0.66$).

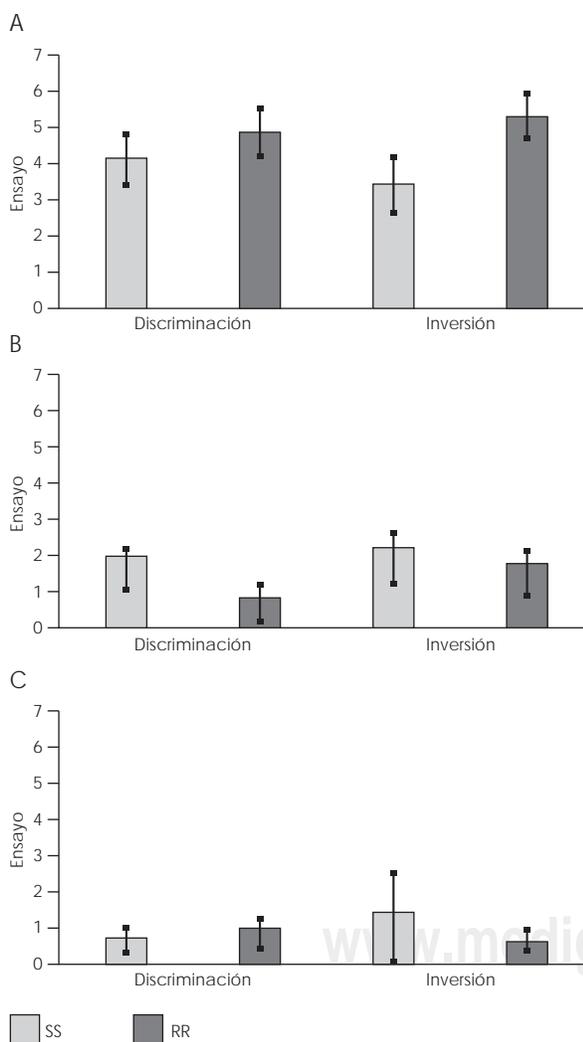


Figura 4. A. Media \pm 2 ES de ensayos correctos en las fases de discriminación e inversión de ratas tratadas con solución salina (SS) y de sujetos virtuales o redes neurales (RN). B. Media \pm 2 ES de ensayos incorrectos en discriminación e inversión de SS y de RN. C. Media \pm 2 ES de ensayos sin respuesta en discriminación e inversión, tanto de SS como de RN.

En la Figura 4 podemos observar la media del total de respuestas correctas, incorrectas y de no respuestas de las redes neurales (RN) en la ejecución de las tareas simuladas en comparación con el grupo de ratas registradas (SS).

El análisis estadístico, ANOVA de medidas repetidas, nos mostró que las respuestas de SS no difirieron significativamente de las de las RN, en respuestas correctas ($F_{AXB} = 2.88$; $p > .09$), ni en incorrectas ($F_{AXB} = 0.34$; $p = 0.56$), y tampoco en no respuestas ($F_{AXB} = 2.51$; $p = 0.12$), en ninguna de las dos fases. En general, todas las RN lograron tanto el aprendizaje discriminativo, como la inversión de éste, prescindiendo de la variable de no respuestas y reduciendo significativamente las respuestas incorrectas de un modo similar a como lo realizaron los sujetos del grupo SS.

DISCUSIÓN

La red neural artificial diseñada fue capaz de emular características del procesamiento neural que han sido reportadas en la literatura en relación al aprendizaje incentivo. Usando como base las respuestas emitidas por ratas durante el aprendizaje discriminativo y de inversión en un laberinto T, se logró la codificación de la aparición del E^I en AMG^{14,16,18,19,28}, en ACC^{29,30}, y además en COF^{14,18,31}, evidente en el mayor incremento de los valores de las proyecciones de AMG hacia COF^{21,28}.

El mayor incremento de los valores de los pesos de las conexiones eferentes de AMG hacia COF, observado en este modelo parece concordar con las hipótesis plasmadas en diversos estudios sobre el papel de estas estructuras en la asignación de V ^{14-16,21}. Evidencia adicional representa la activación diferenciada de sus nodos, oscilante con respecto a la entrega de E^I .

Por otra parte, el incremento en los valores de las eferencias de COF hacia la salida correcta de ACC y no hacia la incorrecta, es concordante con la hipótesis de su participación preponderante en la emisión de respuestas correctas^{19,21}. Finalmente, las oscilaciones de activación parecen señalar la participación de COF en la representación de E^D , como estímulo identificador de la entrega de E^I ^{15,16,18}.

Lo anterior, evidentemente, contrasta con los comportamientos de AMG, en donde las oscilaciones de la actividad de sus nodos codificaron de modo inverso a COF, a ATV y la capa de salida de ACC la presentación del E^I . Esta diferencia en las oscilaciones parece señalar que la capa de

salida del ACC, la COF y el ATV por un lado, y la AMG por el otro, podrían estar jugando un papel distinto con respecto a los E^D o E^A , los primeros, presentando una expectativa de aparición del E , tendencia más evidente aún en la actividad de los nodos de COF^{22,32}.

Los hallazgos anteriores permiten sugerir que el modelo desarrollado es biológicamente plausible al reproducir algunas de las características básicas de la neurofisiología de las estructuras integrantes del circuito. Es probable que precisamente las características básicas mencionadas, permitieran la gran similitud de ejecución al comparar las respuestas de las RN con las de las ratas tratadas con SS, sin presentarse diferencias estadísticamente significativas entre ambos grupos. Múltiples trabajos han sugerido que las características de las RN impactan directamente en su habilidad para ejecutar una determinada tarea^{1,2,8,9}. Este modelo, contrario a otro tipo de exploraciones con un tipo similar de aprendizaje no supervisado^{8,9,12}, incorpora en mayor detalle algunas de las características neurofisiológicas determinantes para la resolución de la tarea planteada. Es posible que ello permita que el modelo sea capaz de aprovechar las ventajas evolutivas, reduciendo significativamente el número de iteraciones requeridas para obtener una ejecución similar a aquella llevada a cabo por entidades biológicas.

Sin embargo, a diferencia de otros modelos utilizados en tareas similares, como el de Hasselmo³³, donde se simula el reforzamiento en la resolución de un laberinto, el expuesto en este trabajo no incorpora la señal propioceptora como elemento de retroalimentación, circunstancia que podría, dadas las características de optimización en el aprendizaje, resultar sumamente provechosa.

CONCLUSIONES

La red neural artificial que simula el sistema mesolímbico cortical permitió emular las respuestas conductuales en una tarea de aprendizaje discriminativo y de inversión a partir de una serie de fenómenos fisiológicos críticos para la ejecución adecuada de ratas en tareas de discriminación e inversión y que, parecen jugar un papel crítico para generar la respuesta instrumental en este tipo de tareas. La consideración de algunas características neurales suficientemente simples de activación, modificación de conexiones y arquitectura del sistema mesolímbico, permitió optimizar el número de iteraciones requeridas para la resolución de

tareas complejas aprovechando las ventajas del moldeamiento evolutivo de los circuitos cerebrales y evitando un entrenamiento previo a la tarea experimental de parte de las redes.

REFERENCIAS

1. Enquist M, Ghirlanda S. *Neural networks and animal behavior*. Princeton University Press (New Jersey), 2005.
2. Rolls ET, Romo RS. "Stochastic dynamics as a principle of brain function". *Progress in Neurobiology*, 2009; 88: 1-16.
3. Doya K. "What are the computations of the cerebellum, the basal ganglia and the cerebral cortex?". *Neural Networks*, 1999; 12: 961-974.
4. Grabenhorst F, Rolls ET. "Value, pleasure and choice in the ventral prefrontal cortex". *Trends in Cog. Sci.*, 2011; (2): 56-67.
5. Anderson J, Rosenfeld E. *Neurocomputing. Foundations of Research*. MIT Press (Cambridge), 1988.
6. Deco G, Rolls ET, Albantakis L, Romo R. "Brain mechanisms for perceptual and reward-related decision-making". *Progress in Neurobiology*, 2012; (en prensa).
7. Schmajuk NA, Gray J, Lam Y. "Latent inhibition: A neural network approach". *J Exp Psychol Anim Behav Proc*, 1996; 22(3): 321-349.
8. Balkenius C, Morén J. "Emotional learning: A computational model of the amygdala". *Cybernetics and Systems*, 2001; 32: 611-636.
9. Zhou W, Coggings R. "Computational models of the amygdala and the orbitofrontal cortex: A hierarchical reinforcement learning system for robotic control". In: McKey R, Slaney J, editors. *AI: 2002: Advances in artificial intelligence*, Springer-Verlag (Heidelberg), 2002: 419-430.
10. Donahoe J, Burgos JE, Palmer D. "A selectionist approach to reinforcement". *J Exp Anal Behav* 1993; 60(1): 17-40.
11. Donahoe J, Palmer D. *Learning and complex behavior*. Allyn and Bacon (Boston), 1994.
12. Armony J, Servan-Schreiber DJ, LeDoux J. "Computational modeling of emotion: explorations through the anatomy and physiology of fear conditioning". *Psychol Rev* 1997; 81(3): 199-213.
13. Deco G, Rolls ET. "Synaptic and spiking dynamics underlying reward reversal in the orbitofrontal cortex". *Cerebral Cortex*, 2005; 15: 15-30.
14. Rolls ET, Treeves A. *Neural networks and brain function*. Oxford University Press (Oxford), 2001.
15. Rolls ET. *Memory, attention and decision making*. Oxford University Press (Oxford), 2008.
16. Schoenbaum G, Chiba A, Gallagher M. "Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning". *J Neurosci* 1999; 19(5): 1876-1884.
17. Rolls ET. "The orbitofrontal cortex and reward". *Cerebral Cortex*, 2000; 10: 284-294.
18. Schoenbaum G, Chiba A, Gallagher M. "Changes in functional connectivity in orbitofrontal cortex and basolateral amygdala during learning and reversal training". *J Neurosci* 2000; 20(13): 5179-5189.
19. Schoenbaum G, Setlow B, Nugent S, Saddoris M, Gallagher M. "Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals". *Learn & Mem*, 2003; 10: 129-140.
20. Guevara MA, Robles-Aguirre FA, Quirarte GL, Hernández-González M. "Orbitofrontal cortex inactivation impairs early

- reversal learning in male rats during a sexually motivated task", *Int J Psychol & Psychol Ther* 2009; 9(2): 141-160.
21. Schoenbaum G, Saddoris M, Stalnaker T. "Reconciling the roles of orbitofrontal cortex in reversal learning and the encoding of outcome expectancies". In: Schoenbaum G, Gottfried J, Murray E, Ramus S, editors, *Linking affect to action*, New York Academy of Science (Boston), 2007: 320-335.
 22. Saddoris M, Gallagher M, Schoenbaum G. "Rapid associative encoding in basolateral amygdala depends on connections with orbitofrontal cortex. *Neuron* 2005; 46: 321-331.
 23. Fuster J. *The prefrontal cortex*. Raven Press (New York), 1989.
 24. Domesick V. "Neuroanatomical organization of dopamine neurons in the ventral tegmental area". In: Kalivas P, Nemeroff Ch, editors, *The mesocorticolimbic system*, New York Academy of Sciences (New York), 1988: 10-26.
 25. Öngür D, Price J. "The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans". *Cerebral Cortex*, 2000; 10: 206-219.
 26. McDonald, AJ. "Cortical pathways to the mammalian amygdala". *Prog Neurobiol* 1998; 55: 257-332.
 27. Pitkänen, A. "Connectivity of the rat amygdaloid complex". In: Aggleton JP, editor. *The amygdala*, Oxford University Press (Oxford), 2000: 31-115.
 28. Schoenbaum G, Setlow B, Saddoris M, Gallagher M. "Encoding predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from basolateral amygdala". *Neuron* 2003; 39: 855-867.
 29. Fernández-Espejo E. "¿Cómo funciona el *nucleus accumbens*?" *Rev Neurol* 2000; 30(9): 845-849.
 30. Parkinson JA, Olmsted MC, Burns L, Robbins TW, Everitt BJ. "Disassociation in effect of lesions of the nucleus accumbens core and shell on appetitive pavlovian approach behavior and the potentiation of conditioned reinforcement and locomotor activity by D-amphetamine". *J Neurosci* 1999; 19: 2401-2411.
 31. Zald DH, Won-Kim S. "The orbitofrontal cortex". In: Sallowey SP, editor. *The frontal lobes and neuropsychiatry illness*. American Psychiatric Publishing (New York), 2001: 33-69.
 32. Sokolowski JD, Salamone J. "The role of accumbens dopamine in lever pressing and response allocation: effects of 6-OHDA injected into core and dorsomedial shell". *Pharm Biochem Behav* 1996; 59: 557-566.
 33. Hasselmo M. "A model of prefrontal cortical mechanisms for goal directed behavior". *J Cog Neurosci* 2005; 17(7): 1115-1119.