

SISTEMA DE DECISIÓN PARA DISPARAR CONDUCTAS MOTORAS DIFERENTES EN UN ROBOT AUTÓMATA MÓVIL

Alejandro Naranjo; Isabella Torres; José Londoño; Juan Camilo Tobar, Jhonatan Giraldo; Sergio Luis Castaño

alejandro.naranjo_z@uao.edu.co

isabella.torres@uao.edu.co

jose.londono_san@uao.edu.co

juan.tobar@uao.edu.co

jhonatan_Dav.giraldo@uao.edu.co

sergio.castano@uao.edu.co

Ingeniería Mecatrónica - Ingeniería Biomédica

Facultad de Ingeniería

Departamento de Electrónica y Automática.

Universidad Autónoma de Occidente

Calí, Colombia

Resumen — En el siguiente documento se encuentra el procedimiento llevado a cabo para poder dar solución al problema de diseño de una red neuronal biológica de decisión compleja para disparar conductas motoras de espera, exploración o huida en un robot autómata móvil, teniendo entradas de ganancia y riesgo provenientes idealmente del entorno. El resultado fue suficiente para cumplir con las expectativas de supremacía instantánea y por ventaja, más no se logró el objetivo del aprendizaje histórico.

Palabras clave— Inhibición, Excitación, Nacka Rushton, Neurona, Delay, Pesos sinápticos, Bias.

Abstract — In the following document you can find the procedure carried out in order to solve the problem of designing a complex biological decision neural network to trigger motor behaviors of waiting, exploration or escape in a mobile automaton robot, having gain and risk inputs ideally coming from the environment. The result was sufficient to meet the expectations of instantaneous supremacy and by advantage, but the objective of historical learning was not achieved.

Keywords – Inhibition, Excitation, Nacka Rushton, Neuron, Delay, Synaptic weights, Bias.

I. INTRODUCCIÓN

Dentro de la gran gama de aplicaciones y rompecabezas que pueden ser armados en la neurociencia computacional, es impresionante notar la complejidad que ganan ciertos sistemas. Pasando desde ciertos dilemas relativamente sencillos de resolver como la creación de una banda de valores de intensidad en espectro continuo, pasando por una red de detección de movimiento, hasta una red más compleja que las anteriores como lo es una que permita una decisión compleja. El último de estos, consiste en una red neuronal que funciona como un sistema de decisión para disparar conductas diferentes, como lo pueden ser explorar, huir o esperar a

determinadas señales de riesgo o ganancia que se pueden obtener del entorno. Todo esto en un robot autómata móvil.

El sistema de decisión tiene como entrada dos señales, E1 y E2, que compiten en el tiempo, dichas señales son valores numéricos en el intervalo [0,1] que representan riesgos y ganancias, respectivamente. Las señales de entrada progresan por cadenas de neuronas de proyección M separadas por neuronas de retardo Δ , de este modo se genera una ventana de tiempo donde se utilizan las señales registradas en 5 instantes de tiempo diferentes, de modo que no se tomen decisiones apresuradas. Las neuronas M proyectan sus estados a neuronas acumuladoras A, que producen una salida dependiendo de la suma de estados proveniente de las cadenas lineales. Las neuronas A excitan o inhiben a las neuronas U de acuerdo con el valor de la suma que almacenan y los umbrales que allí se configuren. Por su parte, las neuronas U son las responsables de disparar las conductas motoras: explorar, huir o esperar.

En el caso en que la suma de las señales de entrada de un canal sea mayor que la suma del otro, es decir, cuando la ganancia sea mucho mayor que el riesgo, se disparará la conducta motora de explorar, o por el contrario, cuando el riesgo sea mucho mayor que la ganancia la conducta motora que se disparará será la de huir. Finalmente, si no es posible para las entradas describir la necesidad de huir o la seguridad de explorar, se activará la conducta motora de esperar.

El grafo que ilustra la estructura de la red neuronal descrita se puede observar en la figura 1.

El objetivo de esta práctica es comprobar si la red neuronal propuesta cumple con dos propósitos propuestos por sus diseñadores.

- Propósito 1: El diseño de red neuronal permite que la decisión se tome considerando una ventana deslizante de observación que impida decisiones apresuradas por parte del autómata.
- Propósito 2: La operación en el tiempo le permite al autómata privilegiar una conducta sobre la otra, generando un sesgo que podría adaptarse con el tiempo llevando a decisiones más firmes y asertivas.

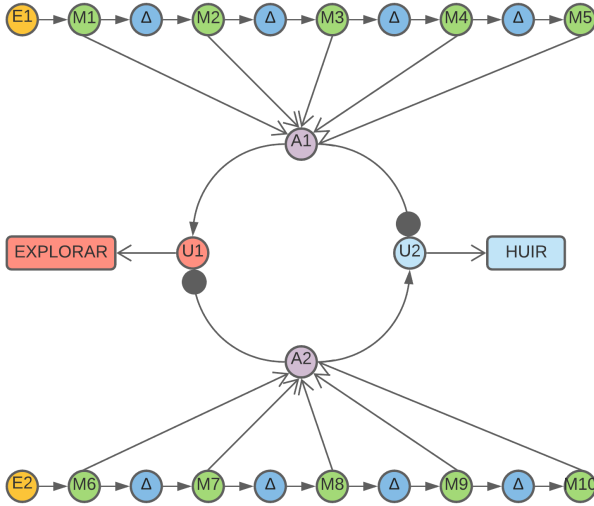


Fig. 1. Grafo red neuronal para sistema de decisión compleja.

II. MÉTODOS Y DESARROLLO

Para poder resolver el problema, y al ver la complejidad que este poseía, se decidió subdividirlo en varios pasos consecuentes para que en conjunto den solución al ejercicio planteado, entre ellos se encuentran: Sistema de delay, Neuronas acumuladoras y Decisión final.

El primer reto al que se enfrentó el equipo fue cómo generar el retardo que sugiere la estructura de la Figura 1. Para esto se propusieron diferentes modelos que pudieran transmitir o recrear los mismos valores (en este caso escalones) de las entradas en cada una de las neuronas que poseían un efecto de delay en el grafo inicial.

A continuación se mostrarán los modelos diseñados para intentar dar solución a la primera parte:

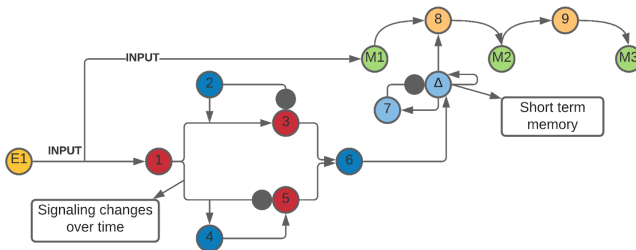


Fig. 2. Grafo con memoria a corto plazo y señalización de cambios en el tiempo.

El grafo que se observa en la Figura 2, es un modelo que contenía una memoria individual, que regulaba el paso de la información. Dicha memoria era accionada con un sistema de detección de cambio de señal visto en clase, utilizando la señal de subida para indicar un cambio en la señal de entrada. Aunque este sistema podía transmitir la información de una neurona de proyección a otra, ésta se perdía a medida que pasaba el tiempo, en el sentido de la reducción de su amplitud. Además, no había forma de actualizar el dato de la neurona, por lo que se terminó descartando.

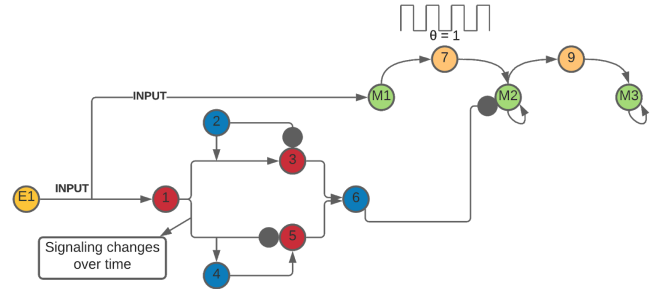


Fig. 3. Grafo con señalización de cambios en el tiempo, autoexcitación en neuronas M2, M3,...,Mn y señal de reloj como entrada a la neurona 7.

Luego de esto se propuso un sistema muy parecido pero con la diferencia que este vaciará cada neurona de derecha a izquierda para poder actualizarla nuevamente con la señal de entrada. Como se observa en la Figura 3, se genera una autoexcitación a la neurona para mantener el valor durante la ejecución, con el objetivo de limpiar la última neurona y que la anterior le pasara el dato a la siguiente. No obstante, este sistema era muy inestable y generaba demasiadas intermitencias como para poder cumplir el propósito 1 por lo que se descartó. Además, no se pudo crear la señal de reloj necesaria, pues no se encontró otra forma diferente a la de mantener una señal activa en el tiempo que dura la ejecución de todo el sistema, y esto no es biológicamente posible.

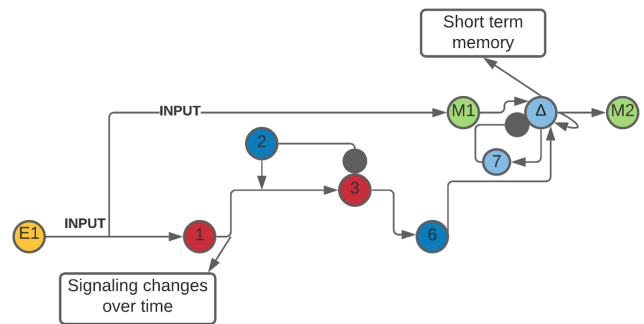


Fig. 4. Grafo con señalización de cambios en el tiempo y memoria a corto plazo.

El sistema que se observa en la Figura 4 recrea por completo la señal de la entrada utilizando una memoria de corto plazo. La idea fundamental era que la señal que mostraba la caída de la entrada activara la señal de memoria. Sin embargo, la magnitud de esta respuesta nunca pudo ser regulada utilizando

la entrada de manera pura, sino más bien utilizando pesos sinápticos, lo que no le permitía tener un comportamiento aceptable a cualquier entrada.

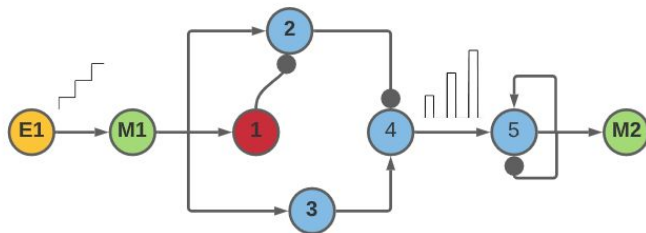


Fig. 5. Grafo de red para generar pulsos de activación retrasados con igual magnitud que la entrada.

Tras haber notado el efecto de una interneurona de relevo con una constante de tiempo superior a la neurona de proyección, se utilizó este concepto para generar retardos más grandes, empleando la idea de la red mostrada en la Figura 5. La intención con esta arquitectura fue inhibir a la neurona 4 por medio de la unidad 2, a su vez la unidad 2 tiene una función de activación Nacka-Rushton sin rectificación de valores negativos de la entrada neta, con ello se consigue que mientras la entrada que proviene de la neurona de proyección M1 y la salida de la unidad de relevo 1 no sean iguales, se va a generar una acción inhibitoria sobre la neurona 4. Por medio de la constante de tiempo de la unidad 1 se controla entonces el tiempo que estará inhibida la unidad 4.

Una vez se igualen los estados de M1 y la unidad 1, la unidad 4 con función de activación lineal, adquiere el estado que es proyectado a través de la unidad 3 generando pulsos cuya duración es controlable. El inconveniente con esta alternativa es que no se logró que la unidad 5, la cual es una unidad de memoria de corto plazo, mantuviese un estado igual o por lo menos proporcional a la magnitud de su entrada, incumpliendo el objetivo de una memoria. Por ende, también se descartó.

En este punto se tomó la decisión de optar por soluciones que algunos libros mostraban al respecto, en particular el de Spikes, Decisions and Actions de Hugh Wilson, donde explicaban el siguiente modelo:

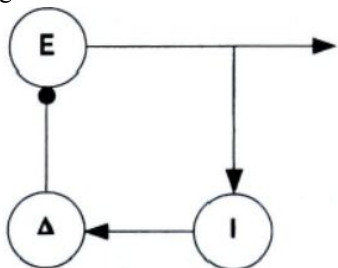


Fig. 6. Bucle de retroalimentación negativa con una etapa de retardo Δ introducida entre las neuronas I y E.

Este modelo se describe por medio de las ecuaciones representadas en la figura 7.

$$\begin{aligned}\frac{dE}{dt} &= \frac{1}{10}(-E - 2\Delta) \\ \frac{dI}{dt} &= \frac{1}{50}(-I + 8E) \\ \frac{d\Delta}{dt} &= \frac{1}{\delta}(-\Delta + I)\end{aligned}$$

Fig. 7. Ecuaciones del modelo. Encontrado en el libro de Spikes

Al leer y ensayar detenidamente se encontró que al excitar una neurona con una función de activación lineal y algunos valores de τ , el sistema empezaba a responder de una manera más retardada, controlando su efecto por el τ y por el número de neuronas de esta configuración que se conectaran en paralelo.

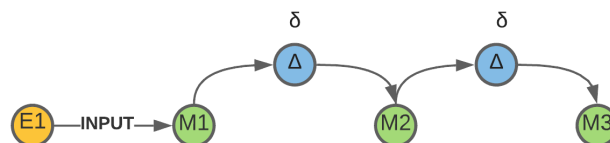


Fig. 8. Grafo de la última opción.

Parecía la opción más pausable, pues su delay no dependía entonces de la duración del escalón de la entrada como si pasaba con las soluciones anteriores, sin embargo, a la hora de implementarlo con los 5 retardos, la magnitud se perdió dañando el propósito, por lo que se acabó descartando.

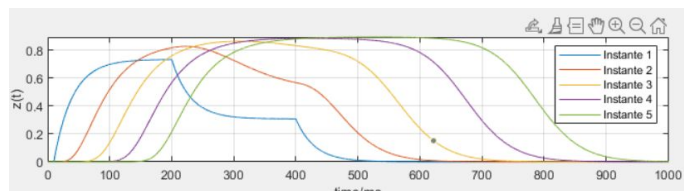


Fig. 9. Respuesta del último modelo.

En vista de que los modelos propuestos correspondientes a las cadenas lineales de neuronas de proyección M y de retardo Δ no tuvieron el desempeño esperado, se optó por implementar el retardo de manera matemática aun sabiendo que este comportamiento no es biológicamente posible. Por lo tanto, se utilizó un retardo de 1s entre cada componente de las señales de entrada.

A continuación se comienza a describir el proceso actual de la solución planteada:

Las señales de entrada son pulsos de amplitudes variables entre [0,1].

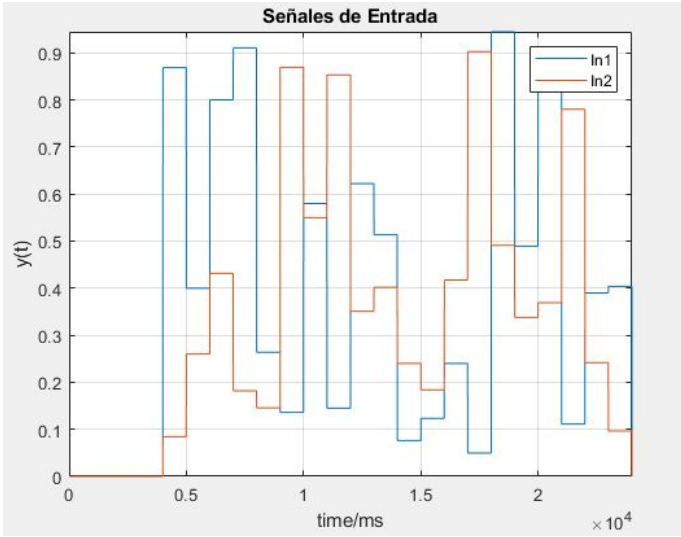


Fig. 10. Pulsos de entrada canal superior (In1) y canal inferior (In2).

Las ecuaciones que describen el modelo matemático del modelo se presentan a continuación.

Las ecuaciones (1)(2)(3)(4)(5) corresponden a las 5 neuronas de la cadena lineal superior.

$$\begin{aligned} \frac{dZ_1}{dt} &= \frac{1}{\tau} (-Z_1(n) + In_1) \\ (1) \\ \frac{dZ_2}{dt} &= \frac{1}{\tau} (-Z_2(n) + Z_1(n - 10000)) \\ (2) \\ \frac{dZ_3}{dt} &= \frac{1}{\tau} (-Z_3(n) + Z_1(n - 20000)) \\ (3) \\ \frac{dZ_4}{dt} &= \frac{1}{\tau} (-Z_4(n) + Z_1(n - 30000)) \\ (4) \\ \frac{dZ_5}{dt} &= \frac{1}{\tau} (-Z_5(n) + Z_1(n - 40000)) \\ (5) \end{aligned}$$

Las ecuaciones (6)(7)(8)(9)(10) corresponden a las 5 neuronas de la cadena lineal inferior.

$$\begin{aligned} \frac{dZ_6}{dt} &= \frac{1}{\tau} (-Z_6(n) + In_2) \\ (6) \\ \frac{dZ_7}{dt} &= \frac{1}{\tau} (-Z_7(n) + Z_6(n - 10000)) \\ (7) \\ \frac{dZ_8}{dt} &= \frac{1}{\tau} (-Z_8(n) + Z_6(n - 20000)) \\ (8) \\ \frac{dZ_9}{dt} &= \frac{1}{\tau} (-Z_9(n) + Z_6(n - 30000)) \\ (9) \\ \frac{dZ_{10}}{dt} &= \frac{1}{\tau} (-Z_{10}(n) + Z_6(n - 40000)) \\ (10) \end{aligned}$$

Donde las Z's equivalen a las neuronas utilizadas e In1 e In2 hacen referencia a los dos tipos de entradas que se utilizaron.

Las ecuaciones (11) y (12) corresponden a las neuronas tanque o acumuladoras A1 y A2, respectivamente.

$$\frac{dZ_{11}}{dt} = \frac{1}{\tau} (-Z_{11} + Z_1 + Z_2 + Z_3 + Z_4 + Z_5) \quad (11)$$

$$\frac{dZ_{12}}{dt} = \frac{1}{\tau} (-Z_{12} + Z_6 + Z_7 + Z_8 + Z_9 + Z_{10}) \quad (12)$$

Las ecuaciones (13) y (14) corresponden a las neuronas de decisión U1 y U2, respectivamente. En la ecuación (13) se observa un umbral de 0.5 que favorece a la conducta de explorar sobre la conducta de huir ya que el umbral que se observa en la ecuación (14) para esta conducta es de 0.2.

$$\frac{dZ_{13}}{dt} = \frac{1}{\tau} \left(-Z_{13}(n) + \frac{(Z_{11}(n) - Z_{12}(n) - 0.5)_+^2}{(0.05)^2 + (Z_{11}(n) - Z_{12}(n) - 0.5)_+^2} \right) \quad (13)$$

$$\frac{dZ_{14}}{dt} = \frac{1}{\tau} \left(-Z_{14}(n) + \frac{(Z_{12}(n) - Z_{11}(n) - 0.2)_+^2}{(0.05)^2 + (Z_{12}(n) - Z_{11}(n) - 0.2)_+^2} \right) \quad (14)$$

Se utilizaron funciones de activación lineal en las neuronas de proyección de ambos canales así como en los acumuladores. En las neuronas de decisión U1 y U2 se implementó la función de activación Naka Rushton.

III. RESULTADOS

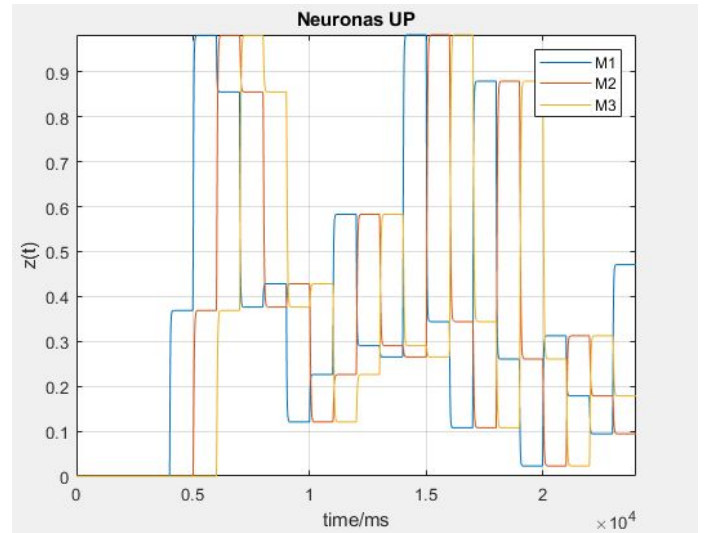


Fig. 11. Salida de tres neuronas de la cadena lineal superior.

En la Figura 11 se puede observar la salida de 3 neuronas únicamente con el objetivo de facilitar su observación. Así mismo, se observa que el delay se aplicó de manera correcta.

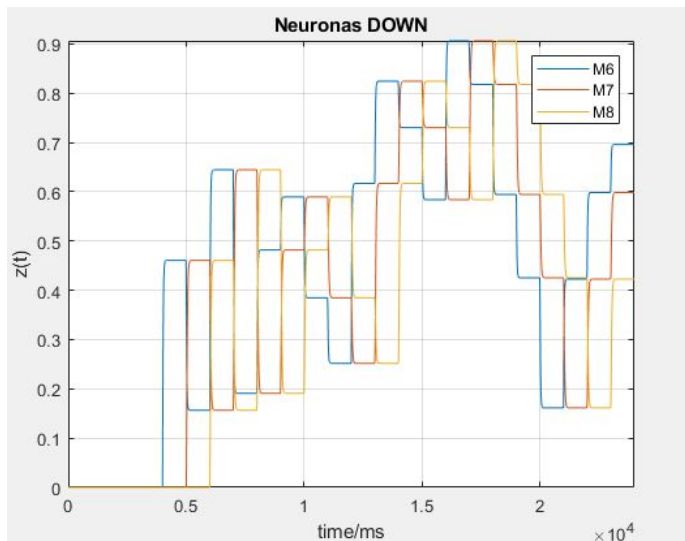


Fig. 12. Salida de tres neuronas de la cadena lineal inferior.

De la misma forma se puede ver en la Figura 12, que para el canal inferior el delay se aplica también de manera correcta. De nuevo, mostrando únicamente 3 neuronas para facilitar su observación.

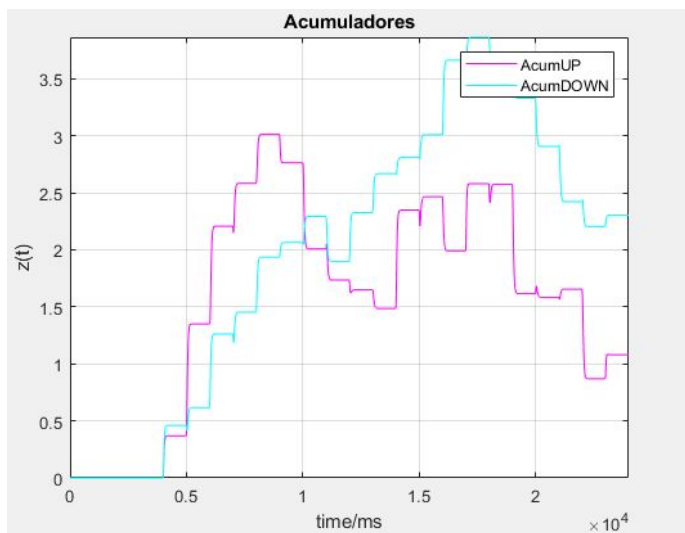


Fig. 13. Salida de las neuronas acumuladoras para cada uno de los canales (UP y DOWN).

Seguidamente se muestra el comportamiento de los acumuladores en la Figura 13, donde se puede ver los efectos acumulados de los 5 instantes de tiempo, mostrando la mayor importancia de cada una de las entradas con respecto a la otra.

Finalmente, se ve cómo se decide entre un comportamiento y otro, tomando en cuenta el umbral de relación de 0.5 y 0.2, el cual representa que el acumulado del riesgo debe superar al de ganancia al menos por 0.5 para generar el estímulo que dispare la conducta huir, mientras que basta con que la ganancia sea mayor al riesgo en 0.2 para que el autómatas explore.

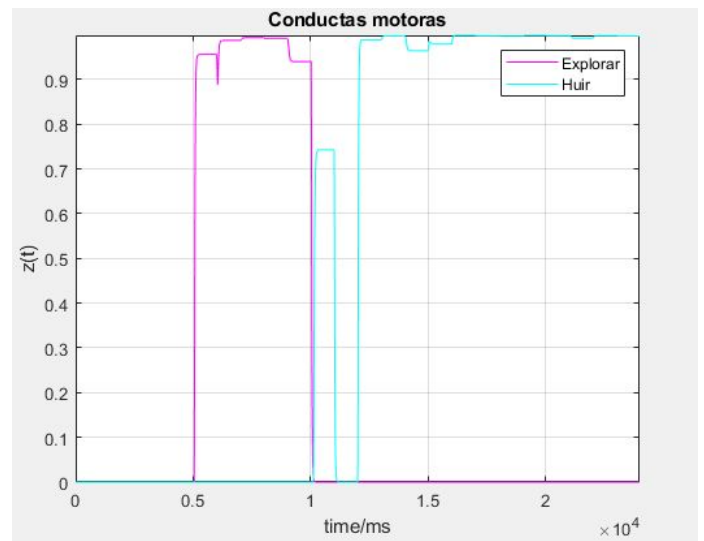


Fig. 14. Activación de las conductas motoras con base en el resultado de los acumuladores.

IV. DISCUSIÓN Y CONCLUSIONES

Al revisar el primer propósito, el cual plantea que la decisión se toma considerando una ventana deslizante de observación con el fin de impedir decisiones apresuradas por parte del autómatas, éste es logrado con la implementación de la memoria lineal, ya que con ella se logra un registro histórico de los valores de estimación de riesgos y ganancias, con lo que el autómatas no toma su decisión basado únicamente en la entrada instantánea, sino también considera la evolución de cada variable, evitando por ejemplo dejar de explorar ante una aparición repentina y momentánea de riesgo, solo hasta que el riesgo total, debido a la acumulación de estimaciones de riesgo significativo que progresan por las neuronas de proyección y se acumulan en las neuronas tanque, sea superior a los estímulos históricos de ganancia, el robot modifica su conducta y huye. Del mismo modo, el autómatas no dejará de huir hasta que las condiciones sean realmente seguras. El tamaño de la ventana de observación depende entonces del número de neuronas en la cadena de memoria lineal, entre mayor sea el tamaño de la ventana menos repentino será el cambio en la conducta del autómatas, lo cual, según la aplicación específica que se le vaya a dar al robot tendrá mayor o menor relevancia.

Es importante resaltar que partes del problema fueron atendidas correctamente, entre ellas podemos ver que la supremacía instantánea fue completada debido a que, si existe una diferencia muy grande entre ambas salidas del acumulador, referentes a las señales de ganancia y riesgo, de manera súbita se toma la decisión. Por otro lado, se puede ver que existe el concepto de ventaja al usar los umbrales, pues al poner un umbral más alto para una neurona U que para otra, se está condicionando a que la señal con desventaja tenga que tener un valor más alto para poder superarla. Resulta evidente entonces que por medio de la variación de este parámetro se puede conseguir el segundo propósito de la red, el cual hace referencia a la habilidad que tendría el autómatas de adaptarse

con el tiempo y poder generar un sesgo o dominancia reforzado de una conducta sobre la otra. Desafortunadamente, no se pudo implementar la dinámica de aprendizaje por falta de tiempo, debido principalmente a los esfuerzos depositados en la consecución de una forma biológicamente inspirada para generar retardos y una memoria lineal.

Con respecto a la dinámica de aprendizaje propuesta en la práctica para la modificación de los umbrales de las neuronas U1 y U2, de haber sido posible implementarla, se hubiera recurrido a un aprendizaje no supervisado, más específicamente, a un aprendizaje por refuerzo. Este aprendizaje consiste en dotar de información suficiente a un agente, en este caso el robot móvil, para que con base en ella pueda tomar decisiones más acertadas que le brinden mayor “recompensa”. El agente está pendiente de las señales de entrada provenientes del ambiente y, de acuerdo con las acciones o conductas que tenga disponible, responderá de una manera u otra. La mayor recompensa podrá ser definida en el caso en que se favorezca la conducta de exploración (ganancia) o por el contrario de huida (riesgo).

Los problemas correspondientes al campo de las neurociencias computacionales, como lo son las redes neuronales biológicamente inspiradas, son un campo de investigación actual en el que las soluciones presentadas pueden ser dadas de maneras muy distintas entre sí y hasta ahora no existe una forma reglamentada y estipulada para resolver ciertos tipos de problemáticas. Y, aunque existen redes neuronales básicas, la resolución de modelos más elaborados, como lo son problemas de decisión compleja, pueden ser resueltos con razonamientos muy dispersos entre sí, con lo cual es posible encontrar modelos más simplificados que otros o modelos que respondan de mejor manera a ciertas variaciones de los parámetros, y que de igual forma, tienen una respuesta asertiva frente a las problemáticas planteadas.

Con la ayuda de las neurociencias computacionales se pueden realizar aproximaciones matemáticas que nos permiten explicar de manera acertada las dinámicas y diversas funciones que cumplen las redes neuronales biológicas. Con ello es posible hallar principios universales que generan un entendimiento más amplio de sus estructuras, dinámicas y funciones que pueden ser predecidas y descritas a modelos matemáticos simulables.

REFERENCIAS

- [1] Wilson, H. R. (2004). *Spikes, Decisions, and Actions: The Dynamical Foundations of Neuroscience*. Oxford University Press, USA.
- [2] Ramirez, D.F. Hurtado, J. (2014). *MODELAMIENTO Y SIMULACIÓN DE CIRCUITOS SINÁPTICOS Y SENSORIOMOTORES: Introducción a la Neurobiología Computacional..* Universidad Autónoma de Occidente, Cali, Colombia.
- [2] Trappenberg, T.P (2010). *Fundamentals of Computational Neuroscience*. Oxford University Press, USA.
- [3] Hertz, J., Krogh, A., y Palmer, R. G. (1991). *Introduction to the theory of neural computation*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.
- [4] Gold, J. I. y Shadlen, M. N. (2007). *The Neural Basis of Decision Making*. *Annual Review of Neuroscience*, 30(1):535–574.
- [5] Eliasmith, C. y Anderson, C. H. (2003). *Neural Engineering (Computational Neuroscience Series): Computational, Representation, and Dynamics in Neurobiological Systems*. MIT Press, Cambridge, MA, USA.