

# El papel de la función sucesor en la arquitectura cognitiva

*Mario Moro Hernández*

Universidad Autónoma de Madrid

---

## Resumen

El presente trabajo pretende explorar el papel que desempeña la función sucesor en el sistema cognitivo humano. Es Zenon Pylyshyn quien sugiere el papel fundamental de esta función en la arquitectura del sistema cognitivo humano (1984). Concretamente, este autor sostiene que si existe alguna función primitiva en nuestra arquitectura ésta ha de ser, necesariamente, la función sucesor. Así pues, el objetivo de este trabajo consiste en explicar el papel que puede desempeñar la función sucesor en una arquitectura funcional y justificar su naturaleza.

*Palabras clave:* psicología, psicología cognitiva, arquitectura cognitiva.

---

## Abstract

This work tries to explore the role of *Successor Function* into Human Cognitive System. It is Zenon Pylyshyn who suggests which this function is central for the architecture of Human Cognitive System (1984). Moreover, this author claims that if there is any primitive function on Cognitive Architecture, it should be Successor Function necessarily. So, the aim of this work is to explain the role which Successor Function would play in a functional architecture and justify its nature.

*Keywords:* Psychology, Cognitive Psychology, Cognitive Architecture.

NOTA: El autor quiere expresar su agradecimiento a las siguientes personas: Eduardo Moro, María Dolores Hernández, David Travieso, Florentino Blanco, Tomás Sánchez, Rubén Gómez, Marta Linares y, muy especialmente, a Mònica Balltandre Pla. Sin su apoyo este trabajo nunca hubiera llegado a buen puerto.

## 1. INTRODUCCIÓN

Los psicólogos cognitivos contemplan una concepción monista del sujeto psicológico. Esta afirmación encierra dentro de sí una serie de restricciones acerca de la ontología del ser humano. La más destacable sería que el origen de la conducta es físico. Por tanto, la mejor manera de describir las conductas sería a través de la descripción de los cambios en la estructura del sistema nervioso. Sin embargo, debido a la complejidad de sus estructuras y al desconocimiento que tenemos del mismo, la postura que adopta la psicología cognitiva es la siguiente: para realizar descripciones informativas acerca de la conducta de los sujetos, lo mejor es atender al plano intencional de las mismas. Esto evita, además, el problema de demostrar que la conducta se debe a patrones específicos y universales de activación neurológica.

El problema está en integrar las conductas descritas a nivel intencional-simbólico en el plano físico utilizando entidades no-físicas –representaciones– como origen de las conductas. La psicología cognitiva asume que el origen de la conducta son los estados intencionales del sujeto. Y he aquí la paradoja: ¿cómo se conjuga que el origen de la conducta sea un estado intencional y un estado físico? La respuesta clásica (por ejemplo, Johnson-Laird, 1993) consiste en afirmar que las conductas tienen un origen físico pero que para su estudio y comprensión es mejor suponer un origen de tipo simbólico. Esto es lo que se denomina «dualismo metodológico»: los componentes simbólicos de la conducta están implementados en estructuras simbólicas que, a su vez, están implementadas en estructuras físicas. Es decir, existe una *arquitectura cognitiva* que está implementada en la arquitectura biológica del sistema nervioso. Ahora, esta postura no elimina el problema, puesto que no se elimina la concepción dualista del sujeto. Por tanto, no nos vamos a llamar a engaño: la psicología cognitiva es una psicología dualista *sensu stricto*.

## 2. EL CONCEPTO DE ARQUITECTURA

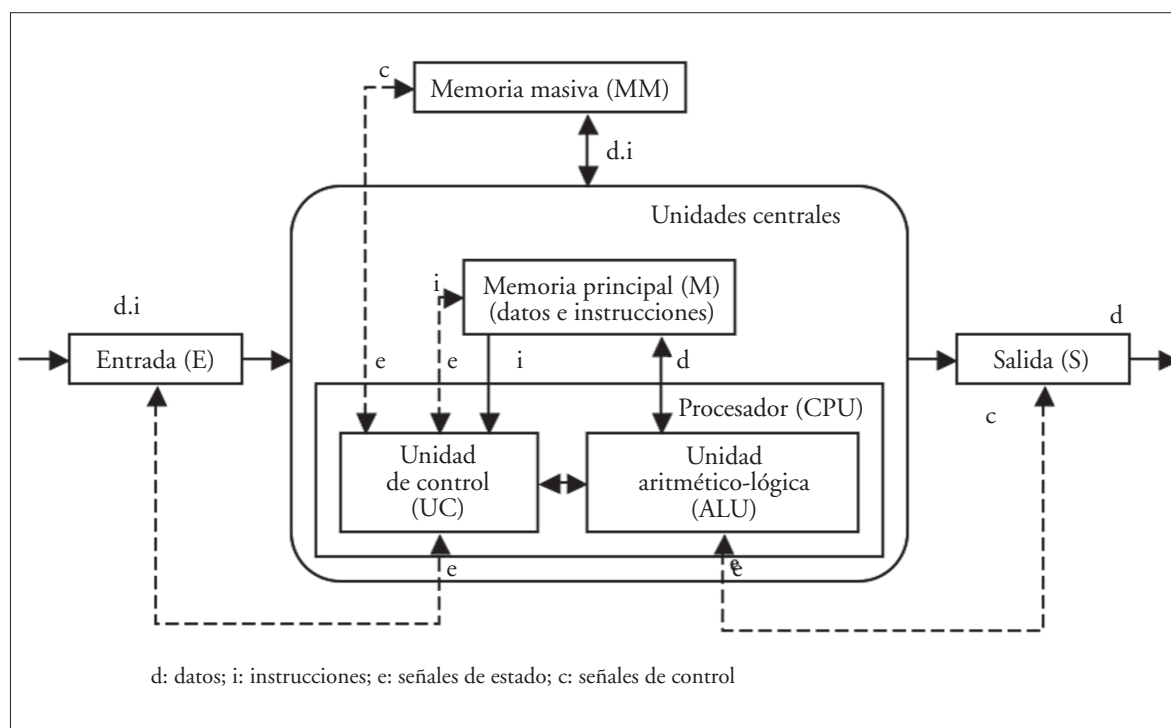


Figura 1. Arquitectura von Neumann

El concepto de arquitectura nace en el seno de la ingeniería y hace referencia al diseño de un dispositivo de cálculo. Al igual que ocurre con las personas, las computadoras pueden describirse utilizando varios niveles de análisis.

Cuando pensamos en los niveles de descripción de un ordenador, lo más habitual es establecer la división *hardware/software*. Esta es la distinción de niveles más básica que cabe esperar. Sin embargo, se pueden establecer hasta seis niveles distintos para el análisis de un computador (cf. Prieto-Espinosa *et al.*, 2002). En la figura 2 aparece un esquema de estos niveles de descripción. Podemos ver cómo el nivel *arquitectura* se sitúa en medio de la división *software-hardware*. Así, el nivel *arquitectura* nos proporciona descripciones «monistas» de la conducta del ordenador. Por este motivo es tan importante para la ciencia cognitiva establecer una arquitectura cognitiva. Si se lograra este objetivo, se resolvería el problema del dualismo inherente a la postura cognitiva.

La psicología cognitiva importó el concepto de arquitectura de la ingeniería. Así, los libros y manuales introductorios al enfoque suelen dedicar algún apartado a explicar qué es una arquitectura cognitiva (por ejemplo, Pylyshyn, 1984; Johnson-Laird, 1993). Según la enciclopedia MIT, el término *arquitectura cognitiva* se referiría al diseño y a la organización de la mente (Sloman, 1999).

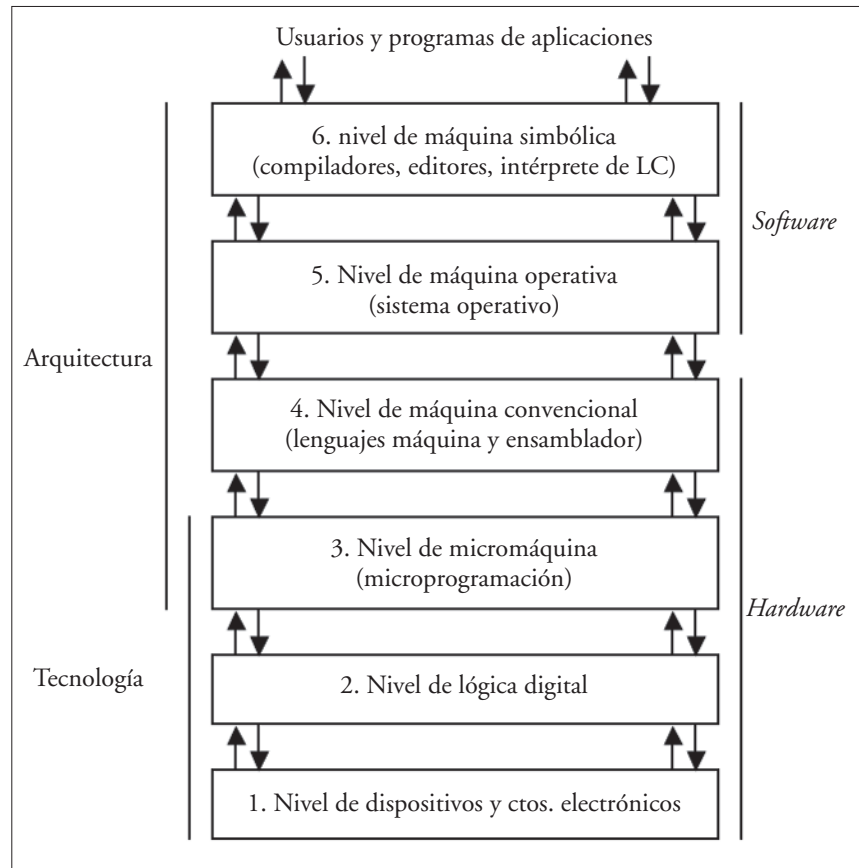


Figura 2. Niveles conceptuales de descripción de una computadora

Dentro de las distintas arquitecturas cognitivas habría que distinguir dos categorías fundamentales: seriales y de procesamiento paralelo.

### 3. SISTEMA NERVIOSO Y SISTEMAS DIGITALES

Desde un punto de vista físico, el funcionamiento de una neurona es similar al de un transistor. Los transistores están fabricados con un material semiconductor y tienen  $n$  entradas y una salida. Sólo cuando la corriente que llega a las entradas es lo suficientemente potente ésta seguirá su camino a través del circuito. Las neuronas emiten un potencial de acción sólo cuando el potencial de acción que reciben de las neuronas aferentes es lo suficientemente fuerte como para producir una despolarización. Si se examina con detenimiento una puerta lógica, se ve cómo sus componentes fundamentales son, precisamente, transistores. El caso más simple es el de la puerta lógica «Y» de dos entradas. Dicha puerta la compone un transistor únicamente:

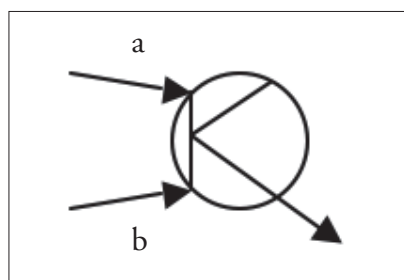


Figura 3. Transistor

El esquema de una puerta «Y» neuronal sería el siguiente:

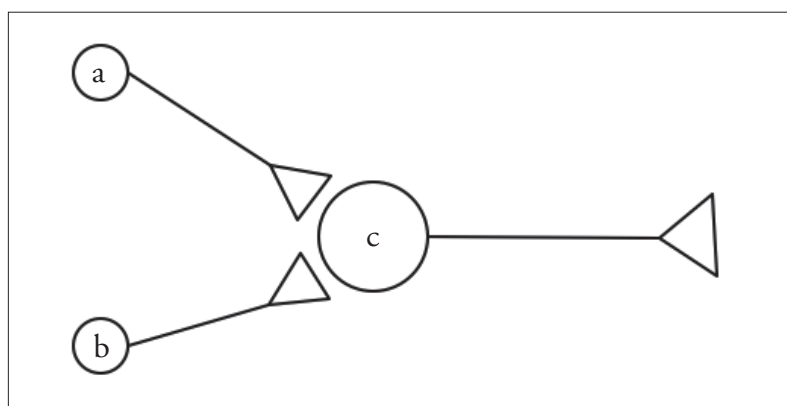


Figura 4. Puerta lógica «Y»

La neurona  $c$  se comportará como una puerta «Y» siempre y cuando las neuronas  $a$  y  $b$  se disparen simultáneamente y no se inhiban entre sí. En el caso de que el disparo de una inhibiera a la otra y la corriente fuera suficiente para producir la despolarización de la neurona  $c$ , estaríamos ante una puerta «O».

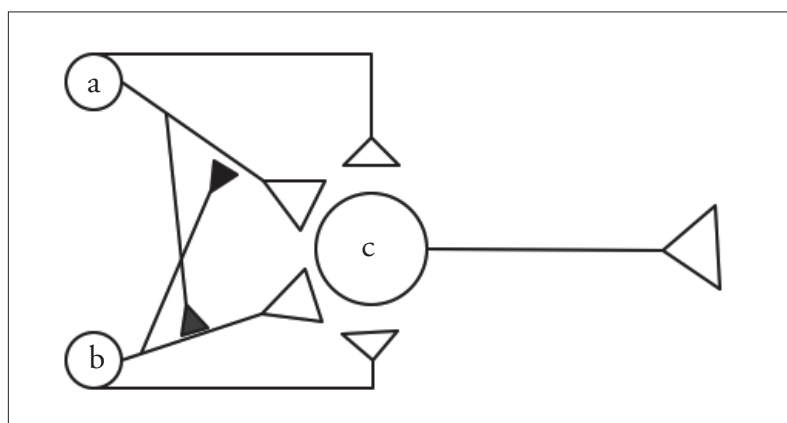


Figura 5. Puerta lógica «O»

La evidencia empírica es más que suficiente para poder afirmar que en el sistema nervioso se dan este tipo de combinaciones neuronales. Por tanto, resulta plausible la asunción de un funcionamiento digital del sistema nervioso de acuerdo con el álgebra de conmutación. Así, en el trabajo clásico de Warren McCulloch y Walter Pitts quedaba demostrado cómo las neuronas tenían un comportamiento digital. Según estos autores: «Un *módulo* (o *neurona formal*) es un elemento con  $m$  entradas  $x_1, \dots, x_m$  ( $m \neq 1$ ) y una sola salida  $d$ » (Arbib, 1969; p. 20 de la v. e.).

Una neurona McCulloch-Pitts no puede sintetizar muchas de las funciones booleanas básicas. Sin embargo, se pueden combinar para formar una red modular.<sup>1</sup> Si además añadimos lo que Arbib denomina «Interacción de Aferentes»,<sup>2</sup> tendremos la «potencia» neuronal suficiente para poder sintetizar cualquier circuito análogo a los circuitos que se encuentran en un ordenador.

Todo esto nos lleva a poder asumir que la información viaja a través de los distintos puntos del sistema nervioso en forma de señal eléctrica. Además, podemos asumir que existen estructuras neuronales análogas a las estructuras combinatorias de los ordenadores. También podemos considerar los nervios como buses que conducen la información a través del sistema nervioso. En definitiva, podemos asumir que tanto los dispositivos de cálculo digitales, como los sistemas nerviosos son equivalentes y que la información que viaja a través de estas estructuras tiene las mismas propiedades: el mensaje, una entrada de control y una serie de «bits para la corrección de errores».

#### 4. LA ARQUITECTURA COGNITIVA, EL CONTROL Y LA FUNCIÓN SUCESOR

Ya tenemos la estructura física: el sistema nervioso. Ahora falta implementar «el aparato lógico». El punto relevante en esta discusión es el siguiente: cómo es posible que el sistema nervioso regule la conducta de un organismo de acuerdo con las condiciones de contorno que impone el medio en el que está insertado.

La respuesta tradicional es que la conducta se regula a través de la ejecución de algoritmos implementados en el sistema cognitivo (por ejemplo, Gardner, 1985; John-

1. Una red modular es una colección de módulos, todos ellos con la misma escala de tiempos, interconectados a base de ramificar la salida de cada módulo en un cierto número de líneas y conectar algunas de ellas, o todas, las entradas de otros módulos. Así pues, una salida puede conducir a varias entradas, pero una entrada sólo puede proceder de una salida como máximo (Arbib, 1969, p. 21 de la v. e.)
2. La interacción de aferentes consiste en conectar el axón de una neurona al axón de una neurona aferente al módulo o «soma» de nuestra neurona. La acción del axón sobre el axón actúa como una puerta de interrupción. Así se puede sintetizar cualquier función booleana  $f(x_1, \dots, x_n)$ .

son-Laird, 1993; Pylyshyn, 1984). Esto se hace de manera «muy sencilla»: el estímulo llega a los receptores en forma de energía física, los receptores traducen dicha energía física a información, dicha información llega al cerebro por las vías aferentes y allí es procesada de tal manera que se emite un nuevo mensaje a través de las vías eferentes y se produce una respuesta.

¿Pero recoge este esquema E-O-R el concepto de arquitectura? Para que la respuesta sea afirmativa, se han de dar una serie de condiciones. En primer lugar, han de existir circuitos neuronales que permitan la implementación física de una arquitectura de propósito general. No parece que esto sea un problema. En segundo lugar, tiene que haber implementados una serie de algoritmos que son necesarios para el procesamiento de la información.

Ahora bien, una arquitectura de propósito general ha de ser lo suficientemente flexible como para poder implementar cualquier procedimiento efectivo: si los algoritmos necesarios para el procesamiento de la información estuvieran implementados de modo innato e inamovible, obtendríamos un autómata finito. ¿Cómo conseguir una arquitectura de propósito general? Implementando físicamente un álgebra que contenga (*a*) los elementos necesarios para la asignación de valores a las variables, (*b*) una serie de axiomas y (*c*) una serie de reglas de producción que permitan la derivación de teoremas.

La última condición es que se ha de producir una biyección entre los estados físicos y los estados mentales. Esto es factible en la medida en que se cumplan las dos condiciones anteriores.

En el caso de una arquitectura von Neumann, es bastante inmediato comprobar que se dan estas condiciones. La existencia de la unidad aritmético-lógica garantiza la implementación del álgebra de computación y las funciones aritméticas básicas. A partir de estos elementos se puede derivar cualquier función computable. Por otro lado, la unidad de control garantiza la biyección de los estados lógicos y los estados físicos del sistema, ya que la ejecución de las instrucciones depende de (y modifica) los estados físicos de la máquina.

Pero qué ocurre en el sistema cognitivo humano. Por un lado, encontramos los conceptos. Éstos desempeñan un papel central en la obtención del álgebra que postulábamos con anterioridad, ya que serían los elementos que completan las funciones. ¿Y dónde se implementan los conceptos? en la memoria. Por esto en todos los modelos de arquitectura cognitiva la memoria y sus procesos son cruciales.

Por otro lado, es innegable el papel que desempeñan los sistemas de producción. Los ejemplos más destacables se encuentran en las teorías del sistema de símbolos físicos (Newell, 1980/1986) y el ACT (Anderson, 1983). Los sistemas de producción no son más que reglas condicionales «si *p*, entonces *q*» que modulan la conducta en función de las condiciones de contorno activadas en el sistema. Es decir, dadas unas condiciones, el sujeto elicitó o recupera una serie de conceptos y en función de los mismos ejecuta una conducta.

Vemos que la «implicación material» es necesaria para la ejecución de las conductas. Esta operación no es sino una «aplicación» en el sentido lógico del término. Su existencia depende de la existencia de una serie de axiomas y de teoremas que necesariamente han de estar implementados en el sistema y que, a su vez, permitan definir funciones de este tipo. En otras palabras: puesto que el sistema opera mediante la utilización de pares condición-acción, las operaciones necesarias para la derivación de estas funciones han de estar implementadas en el sistema.

Dichas operaciones son «Y», «O» y «NO».<sup>3</sup> Como ya hemos visto, estas operaciones se pueden implementar físicamente de un modo muy sencillo. Igualmente se pueden implementar el resto de axiomas de Boole (c. f. Prieto-Espinosa *et al.*, 2002, pp. 132-133). Derivar los teoremas fundamentales de Boole a partir de sus axiomas es inmediato. Con el conjunto de conceptos y los axiomas y teoremas de Boole ya tenemos un álgebra lo suficientemente potente para derivar cualquier función computable. Y lo mejor de todo: implementable en el sistema nervioso.<sup>4</sup>

En este punto se podría pensar que tenemos todo lo necesario para que el esquema E-O-R recoja el concepto de arquitectura cognitiva. Sin embargo, falta un elemento que resulta crucial: el control. Y aquí es donde según Pylyshyn (1984; cap. 5) entra la función sucesor.

Decidir qué significa *control* en este contexto es complicado. Por un lado, puede hacer referencia al orden de ejecución de las instrucciones de un algoritmo. Sin embargo, también se puede entender que hace referencia a la dirección que ha de seguir el pensamiento o la acción –i. e., hacia dónde debe orientarse. Nótese cómo en el primer caso estamos hablando del funcionamiento interno del sistema, mientras que en el segundo caso nos referimos a la conducta que despliega el sistema.

Respecto al control de la conducta, éste viene determinado (*a*) por las condiciones del ambiente en el que está insertado el sistema, así como (*b*) por las características de construcción del propio sistema. Otra cuestión bien distinta es la regulación interna del agente.

Ya se ha comentado que la psicología cognitiva explica la ejecución de las conductas a partir de la aplicación de algoritmos que procesan la información que el sistema recoge del entorno. Podemos decir que un algoritmo es una descripción detallada de un procedimiento.<sup>5</sup> Independientemente de su complejidad, la aplicación de un

3. Las dos primeras son las leyes de composición interna del álgebra de Boole y la tercera es la que define el elemento complementario único.

4. El modo más inmediato de ver esto es fijarse en el álgebra de conmutación de los circuitos digitales (*vid.* Prieto-Espinosa *et. al.*, 2002, pp. 132 y ss.)

5. *Sensu stricto*, un algoritmo es el procedimiento en sí. Lo que se afirma en el texto que es un algoritmo en realidad es la definición del algoritmo.



algoritmo implica un procedimiento secuencial en el que cada etapa se corresponde con un estado del sistema. En un estado determinado, si el algoritmo está definido de manera correcta, el sistema sólo podrá acceder a un conjunto finito de estados y no a otros. Además, una vez que accede a uno de los estados posibles de la siguiente etapa, no puede acceder a ningún otro.<sup>6</sup>

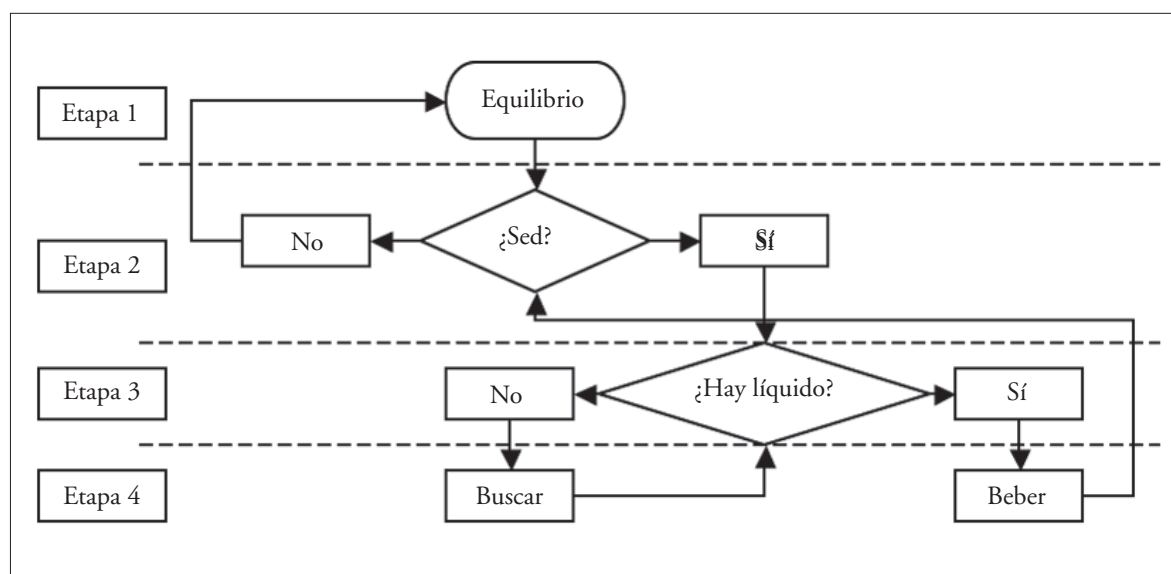


Figura 6. Posible algoritmo «conducta de saciar la sed»

Podríamos establecer un código para el algoritmo anterior del modo que sigue:

Etapa1-Equilibrio	→	Etapa2
Etapa2- Sed_No	→	Etapa1
Etapa2-Sed_Sí	→	Etapa3
Etapa3-Líquido_No	→	Etapa4
Etapa3-Líquido_Sí	→	Etapa2
Etapa4-Buscar	→	Etapa3

La ejecución de un algoritmo necesita que el sistema identifique en cada momento el estado en el que se encuentra y el modo más sencillo de conseguir esto es utilizar un código que identifique cada estado.

Un algoritmo es un conjunto finito de instrucciones. Ahora bien, dichas instrucciones no tienen por qué estar dispuestas en el orden «adecuado». Por ello es necesario establecer un principio por el que las instrucciones aparezcan en el orden correcto. Y aquí es donde entra en juego la función sucesor.

6. Puesto que los estados han de estar definidos de forma exclusiva y exhaustiva.

La función sucesor es aquella que ordena los elementos del conjunto de los números naturales. Siguiendo la notación estándar:  $S(x) = x + 1$ ;  $\forall x \in \mathbb{N}$ , si hacemos una aplicación de un subconjunto de los números naturales en el conjunto de las instrucciones de un algoritmo, tenemos que a cada una de las instrucciones le corresponde un número natural y sólo uno:

- 1  $\rightarrow$  [Etapa1-Equilibrio  $\rightarrow$  Etapa2]
- 2  $\rightarrow$  [Etapa2-Sed\_No  $\rightarrow$  Etapa1]
- 3  $\rightarrow$  [Etapa2-Sed\_Sí  $\rightarrow$  Etapa3]
- 4  $\rightarrow$  [Etapa3-Líquido\_No  $\rightarrow$  Etapa4]
- 5  $\rightarrow$  [Etapa3-Líquido\_Sí  $\rightarrow$  Etapa2]
- 6  $\rightarrow$  [Etapa4-Buscar  $\rightarrow$  Etapa3]

De este modo, el sistema puede identificar cada una de las instrucciones que tiene que ejecutar en cada momento dependiendo del estado en el que se encuentre:

- 1: Equilibrio  $\rightarrow$  2
- 2: Si sed = no  $\rightarrow$  1  
en caso contrario  $\rightarrow$  3
- 3: Si líquido = no  $\rightarrow$  5  
en caso contrario  $\rightarrow$  4
- 4: Beber  $\rightarrow$  2
- 5: Buscar  $\rightarrow$  3

Pero, además de identificar qué instrucción es la que tiene que ejecutar, el sistema también puede identificar en qué estado se encuentra –puesto que la ejecución de cada una de las instrucciones se produce en un determinado estado (y sólo uno).

Así las cosas, es evidente que la arquitectura cognitiva ha de poseer algún mecanismo de control de ejecución de los algoritmos. Además, ha de tener implementada la función sucesor para que el sistema pueda funcionar de modo correcto. El ejemplo más notable de un dispositivo de este tipo serían los «procesos centrales» de la arquitectura modular de Fodor (1983).

Una vez incluido el elemento de control, sí que podemos afirmar que el esquema E-O-R recoge el concepto de arquitectura.

## 5. RECAPITULACIÓN FINAL

En la medida en que la supervivencia del sistema depende de la ejecución correcta de los algoritmos que utiliza para procesar la información, parece más que necesaria la aplicación de un método que controle la ejecución de los mismos. Y dentro de este

proceso de control, la función sucesor parece que desempeña un papel fundamental. Así pues, tanto este dispositivo como la función sucesor han de estar implementados desde el primer momento en el sistema.

Esta es una parte del argumento que presenta Zenon Pylyshyn en *Computation and Cognition* (1984). Por otro lado, este autor defiende que la función sucesor es la única función de la que se puede afirmar que es una primitiva del sistema. Tiene razón. Anteriormente se han establecido unas primitivas previas: los teoremas del álgebra de Boole. Sin embargo, no hay nada que haga suponer que el sistema cognitivo funcione de acuerdo con los principios de esta álgebra. De hecho, suponiendo que el sistema funcione utilizando una lógica de primer orden, no tiene por qué ser la lógica booleana. Empero, puesto que la función sucesor es una función computable, ésta puede caracterizarse utilizando cualquier formalismo lógico de primer orden (por la tesis de Church-Turing). Por eso, la única función de la que podemos asegurar su carácter primitivo es la función sucesor.

Ahora bien, ¿qué nos dice esto acerca de la posibilidad de una descripción simultánea del sistema cognitivo en todos los niveles? Muy poco. Por simple probabilística puede que existan circuitos neuronales que sintetizen puertas lógicas. Incluso, que se den las interacciones de aferentes necesarias para sintetizar circuitos más complejos. Pero la biyección de niveles implica algo más: que dichos circuitos estén perfectamente descritos tanto a nivel físico, como a nivel lógico. En principio esta tarea resulta posible, pero no muy probable de realizarse. Quizá el *boom* actual de la neurociencia con un espíritu netamente localizacionista tenga una estrecha vinculación con este objetivo. Sin embargo, aunque se consiga una descripción perfecta del sistema nervioso, ésta no garantiza una concepción monista del sujeto cognitivo, puesto que faltará la integración del nivel lógico. Por tanto, sin una buena descripción integrada de ambos niveles y de su funcionamiento, nunca se podrá hablar de una arquitectura cognitiva unitaria. Así, el sujeto de la psicología cognitiva estará condenado a ser por siempre un sujeto dualista.

### *Referencias bibliográficas*

- ANDERSON, J. R. (1983): *The Architecture of Cognition*. Cambridge, MA, Harvard University Press.
- ARBIB, M. A. (1969): *Brains, Machines and Mathematics*. Nueva York, NY, McGraw-Hill.
- FODOR, J. A. (1983): *The Modularity of Mind*. Cambridge, MA, The MIT Press.
- GARDNER, H. (1985): *The Mind's New Science. A history of the Cognitive Revolution*. Nueva York, NY, Basic Books.

- JOHNSON-LAIRD, P. N. (1993): *The Computer and the Mind: An Introduction to Cognitive Science*. Londres, Fontana Press.
- NEWELL, A. (1980/1986): «Sistemas de Símbolos Físicos», en D. Norman (ed.) (1986), *Perspectivas de la Ciencia Cognitiva*. Barcelona, Paidós.
- PRIETO-ESPINOSA, A., A. LLORIS y J. C. TORRES-CANTERO (2002): *Introducción a la Informática*. Madrid, McGraw-Hill (3.<sup>a</sup> edición).
- PYLYSHYN, Z. W. (1984): *Computation and cognition. Towards a Foundation of a Cognitive Science*. Cambridge, MA, The MIT Press.
- SLOMAN, S. (1999): «Cognitive Architecture», en R. A. Wilson y F. C. Keil (eds.), *The MIT Encyclopedia of the Cognitive Sciences*, pp. 124-126.