
BIOLOGÍA DE LA VIDA
Y LA CONCIENCIA:
A PROPÓSITO DE DARWIN

ANDRÉS MOYA

ABSTRACT. One important feature of the theory of complexity is the explanation of emergent properties and common features that occur, no matter the trait under study. In the present paper I give support to the thesis that in order to reproduce materialistically any emergent property; we need to know in detail the corresponding physical, chemical and biological properties of the corresponding complex feature, properties that cannot be abstracted. For instance, to synthesize life or to build a brain we need to know the molecules and chemical reactions of prebiotic chemistry or the neurons and brain organization. This knowledge should be added to the theoretical one that shows us the commonalities that are present in the generation of emergent properties, in this case two complex features, life and brain. The theory of evolution as a framework constitutes a good example of integration between the theory of complexity and the nature of the material components of complex traits in the explanation of how emergent properties may have arisen.

KEY WORDS. Origin of life, consciousness, qualia, emergence, complexity, computer simulation, artificial intelligence, algorithm.

INTRODUCCIÓN

El de la vida y la conciencia son dos orígenes que nos interesan particularmente. De hecho, junto al del universo, constituyen la base sobre la que se han construido sistemas filosóficos o desarrollado teorías científicas de gran calado. En este estudio voy a centrarme en dos de ellos, el de la vida y el de la conciencia, porque entiendo existe una cuestión que cada vez se nos hace más aparente y requiere reflexión: nuestra capacidad de sintetizar o reproducir fenómenos complejos, particularmente, como digo, la de sintetizar vida en el laboratorio o la capacidad de construir entidades que tengan, como nosotros, comportamiento consciente. Desde la consideración ontológica fundamental, que podemos tomar como supuesto básico,

Instituto Cavanilles de Biodiversidad y Biología Evolutiva, Universitat de València, Centro Superior de Investigación en Salud Pública (CSISP), València, y CIBER en Epidemiología y Salud Pública (CIBEResp), España. / andres.moya@uv.es

de que ambas son propiedades emergentes de procesos físico-químico-biológicos que han acontecido en determinado momento de la dinámica del planeta o de la evolución biológica, según toque, la vida emerge en el contexto de la evolución química, y la conciencia emerge en el de la evolución biológica del cerebro.

EL PROBLEMA: LA CONCIENCIA COMO MODELO

Son las tesis de Searle (1985) en torno a la irreductibilidad de lo que denomina estados, propiedades o características de la conciencia, o *qualia*, lo que me lleva a plantear su eventual similitud con otros fenómenos complejos. La conciencia es producto de procesos cerebrales, y no existe razón objetiva para sustentar la existencia de dualismo alguno mente/cerebro en la medida que los estados mentales son producto de procesos cerebrales. Cuestión otra es, primero, el que hayamos sido capaces, todavía no, de dar cuenta de todos y cada uno de los factores que promueven los diferentes estados del cerebro, particularmente el de la conciencia humana y, segundo, la naturaleza particular de alguno de esos estados que nos compelen a decir que son estados autónomos, independientes, irreductibles, que tienen una existencia específica (Searle hace referencia explícita a, por ejemplo, el 'dolor'). Searle (2000) considera que las tesis de Crick (1994) o Edelman (1989) van en su línea de aproximar el estudio de los procesos cerebrales y descubrir, en su biología, cómo pueden ir gestándose *qualia* particulares. Reconoce que es el camino correcto para ir avanzando en el terreno de la comprensión y reproducción de las propiedades del cerebro. Critica severamente a Dennett (1995), con quien mantiene una disputa filosófica de primer nivel, por cuanto para este filósofo fenómenos como la conciencia o el dolor no tienen estatus ontológico, sino que son manifestaciones de procesos biológicos estrictamente materiales. La esencia de su disputa, a mi juicio muy pertinente, tiene que ver, como luego examinaré, con la posibilidad de que puedan aparecer propiedades emergentes y que tales propiedades tengan estatuto ontológico, sin necesidad de pensar en una reducción de las mismas a los elementos que componen el proceso de aparición de la citada propiedad. Y, finalmente, debate las tesis de Penrose (1996) sobre la supuesta necesidad de abordar la comprensión de los procesos cerebrales recurriendo a nuevas leyes o ciencias físicas, en la medida que, a juicio de Penrose, y contrariamente a la posición de Searle, algunos de los procesos cerebrales, por ejemplo los que tienen que ver con la generación del estado de conciencia o los procesos implicados en las resolución de teoremas matemáticos no pueden ser simulados por algoritmos computacionales. En otras palabras, los procesos cerebrales, en general, no pueden ser reproducidos con ninguna máquina de Turing. Penrose, es bien conocido, no sólo critica el programa de la inteligencia

artificial fuerte, sino también la débil. La primera nos indica que el cerebro es un computador, y la segunda que, aunque no lo sea, existen procesos cerebrales que se pueden simular como si fueran algoritmos. Searle (2000) desarrolla extensamente su tesis para sostener, tras criticar con algún detalle las tesis de Penrose (1996), que el cerebro puede investigarse bajo la aproximación de una inteligencia artificial débil. Searle se declara partidario de seguir con la actual investigación neurobiológica del cerebro, y no descarta el que podamos simular procesos cerebrales, aunque los algoritmos correspondientes sean de una naturaleza totalmente distinta a los que Penrose propone. Penrose considera que el cerebro en su conjunto, o procesos particulares de él, no pueden ser simulados porque los algoritmos correspondientes que pudieran desarrollarse van a entrar en contradicción o no podrán resolver ciertas cuestiones que al proceso cerebral correspondiente se le aparecen como obvias o verdaderas. Es decir, el algoritmo no puede decidir, como verdadero o falso, algo que el proceso cerebral sí hace. Penrose se basa para sostener estas tesis en los teoremas de Gödel, y es por ello por lo que recurre a plantear un nivel nuevo o nueva ciencia para el estudio del cerebro, para el que todavía no disponemos de la física adecuada. Searle ciertamente desenvuelve sus tesis en un mundo menos abstracto y argumenta que, aunque los algoritmos que Penrose tiene en mente son algoritmos relacionados con la demostración de teoremas matemáticos, la naturaleza de los procesos cerebrales no necesariamente requiere un tipo de simulación de ese tipo y que, llegado el momento, podríamos dar con uno que, en efecto, pudiera dar cuenta de la misma verdad que al proceso cerebral. Vuelvo sobre este importante debate más adelante.

PARALELISMO ENTRE FENÓMENOS COMPLEJOS

Searle (2000) se concentra en el tema de la conciencia como una manifestación genuina del cerebro humano. Dejemos por el momento de un lado la posibilidad de una evolución cerebral que permita estados conscientes en otras especies filogenéticamente próximas, cosa harto más que factible. Lo que deseo sostener aquí es la similitud o paralelismo fundamental con la naturaleza de otros fenómenos complejos, como el del origen de la vida o cualquier otro fenómeno complejo que haya podido aparecer en la historia de la vida en el planeta. No obstante, consideremos como ejemplo paradigmático de 'cualquier otro fenómeno complejo' el del origen de la vida. La vida tiene un estatus similar respecto de sus componentes como la conciencia respecto de los suyos. Searle (2000) sólo menciona una sola vez la palabra 'emergencia'. Lo hace, precisamente, cuando nos muestra el estatuto ontológico del agua como molécula respecto de sus componentes atómicos, oxígeno e hidrógeno. El agua muestra propiedades de líquido

que no son predecibles a partir de los componentes atómicos, con propiedades de gas cada uno de ellos. Pues bien, la conciencia es una propiedad del cerebro humano, una propiedad física, que ha evolucionado en el mundo animal a partir de un conjunto más o menos amplio de componentes que interactúan en forma que no nos es comprensible todavía, pero que, en todo caso, constituye o se constituye en un nuevo estado con respecto a los elementos implicados en el proceso. La conciencia adquiere entonces estatuto ontológico respecto de los elementos implicados en el proceso cerebral que los genera, y de hecho se caracteriza por *qualia* particulares no reducibles a las propiedades de los componentes particulares del proceso que la origina. Recurre Searle a un muy sugerente concepto de 'umbral' del sistema en evolución, cuya superación dispara la aparición de tales estados nuevos, físicamente causados. De darse esas condiciones de umbral que permitan el salto al nuevo estado, el estado cristaliza, hace acto de presencia, cobra realidad ontológica propia, tal como el agua la tiene con respecto a sus componentes.

El origen de la vida es una emergencia con sus *qualia* y la propia evolución de la vida en el planeta está salpicada de nuevas emergencias y, por lo tanto, llevan asociados sus *qualia* respectivos. Lo que sorprende es que Searle no llame a la conciencia propiedad emergente en la evolución del cerebro de aquel o aquellos procesos cerebrales que lo generan. Por lo tanto, la emergencia es una propiedad fundamental, física y fundamental de la vida. La vida en sí representa un estado nuevo con sus *qualia* respecto de sus componentes y toda la teoría del origen de la vida es un intento por comprender las reglas, condiciones, factores, que han permitido una emergencia tal. El programa de estudio del origen de la vida, por lo tanto, tiene todo su vigor conceptual, aunque evolucione lentamente en sus avances empíricos. Lo mismo ocurre, si no es un reto todavía mayor, con el estudio de la conciencia en el caso de la evolución del cerebro. No todos los organismos han adquirido un estado de tal nivel, simplemente han desarrollado evoluciones y han logrado determinados estados, pero no han llegado a un supuesto estado umbral que permita la emergencia de esa particular propiedad. Pero al igual que ocurre con el campo de estudio del origen de la vida, hemos de desarrollar el campo de estudio del origen de la conciencia.

TEORÍA DE LA COMPLEJIDAD: ¿SUFICIENTE?

Es aquí donde deseo enfatizar que aunque existan paralelismos entre las emergencias, probablemente no sea muy adecuada la aproximación que se nos presenta desde el campo de la teoría de los sistemas complejos, que hacen abstracción de los componentes para sugerirnos que las propiedades de los sistemas son las mismas con independencia de ellos. Los

sistemas complejos pueden presentar regularidades, ciertamente, pero tal observación no es fundamental para reproducirlos, simularlos o sintetizarlos.

Searle (2000) abunda en que para entender la emergencia de la conciencia hay que estudiar 'la biología de la conciencia', al igual que para estudiar el origen de la vida hay que estudiar 'la biología de la vida'. Y para estudiar la emergencia de cualquier otra fenomenología biológica compleja hemos de estudiar la biología correspondiente. No quiero con ello sostener que no podamos utilizar ciertas regularidades observadas en unas emergencias para predecir que aparecerán en otras. Esto es válido. Pero dentro del contexto de reproducir fenómenos, por ejemplo, sintetizar vida o lograr simular que un ente tenga conciencia, probablemente debamos seguir investigando las biologías respectivas, porque todavía nos queda un largo camino y las regularidades o abstracciones de la teoría de la complejidad no son suficientes.

MÁS CIENCIA CUANTA MÁS CIENCIA

Por lo tanto, y por inducción, para el estudio de cualquier sistema vivo que muestre *qualia* emergentes hay que estudiar su 'biología'. Esta no es otra cosa, en todo caso, que estudiar los procesos particulares que caracterizan su evolución. Vuelvo de nuevo a las consideraciones de Searle (2000) en su polémica con Pearson (1996) en torno a considerar al cerebro como un computador. El programa computacional o de inteligencia artificial fuerte viene a decir que el cerebro es un algoritmo computacional complejo, en última instancia un programa. La crítica de Searle a esta visión es muy importante porque comenta que la dinámica del cerebro no se puede reproducir con un programa, porque un algoritmo es una secuencia de instrucciones sintácticas, y el cerebro está hecho de múltiples componentes biológicos que son, en esencia, de naturaleza semántica, esto es, dicen algo. Por lo tanto, aunque podamos reproducir comportamientos del cerebro difícilmente, a su juicio, vamos a poder sostener que el programa 'llegará' a tener conciencia, a 'ser' consciente. Podrá 'dar esa impresión' pero no 'será consciente'. Hasta aquí no hay desacuerdo entre Searle y Pearson. Pero contrariamente a este último, Searle no excluye el que podamos, en algún momento, llegar a montar una entidad con capacidad de ser consciente. Es sólo cuestión de que nuestro conocimiento biológico sobre el cerebro avance para que realmente podamos reproducir o lograr una entidad con las mismas capacidades que tenemos nosotros. Eso lo considera posible, y no detecta problema metafísico u ontológico mayor que nos permita avanzar en ese campo. Es cuestión solamente de añadir más ciencia a la ciencia. Tengo la impresión de que lo mismo ocurre en otras fenomenologías biológicas complejas, donde solamente nueva investigación empírica adicional en cada una de las áreas puede dar pistas

fundamentales para emular o reproducir las complejidades de los sistemas que nos gustaría simular (Moya, et al. 2009).

Si el concepto de simulación de Searle no equivale a la simulación computacional, en el sentido de inteligencia artificial fuerte, a mi juicio tampoco el hecho de encontrar regularidades o similitudes en el comportamiento de sistemas complejos, tal y como nos muestra la teoría de la complejidad, nos asegura que tengamos capacidad en determinado momento de reproducir, simular o sintetizar esos fenómenos y observar los *qualia* asociados.

Pearson ha desarrollado la tesis que sostiene la imposibilidad de computar la conciencia o, dicho de otro modo, que ni en un sentido de inteligencia artificial fuerte o débil se pueda hacer tal cosa. La inteligencia artificial fuerte manifiesta que el cerebro es un computador, mientras que la débil indica que los procesos mentales, si se conocen suficientemente bien, pueden ser simulados, lo que no quiere decir que sean asociables a un algoritmo. Pearson niega estas dos formas de inteligencia artificial aplicadas al cerebro, y manifiesta que la conciencia como un proceso cerebral requiere del desarrollo de una nueva física, que los estados cerebrales que promueven la conciencia tienen naturaleza mecano-cuántica y que son determinadas estructuras moleculares de las neuronas las que a través de tal comportamiento pueden generar el proceso de conciencia.

Interesa en todo caso examinar los motivos que le llevan a Pearson a no dar valor al programa de inteligencia débil, que es precisamente el que sí considera Searle que puede proporcionar una adecuada comprensión de la emergencia de la conciencia como proceso cerebral específico. Y esto tiene interés porque de ser válida la generalización que estoy presentado aquí, aquellas reflexiones que son válidas para los *qualia* de la conciencia como manifestación de procesos cerebrales, también lo son para otros fenómenos complejos, origen de la vida y otros de la historia de vida.

La diferencia fundamental entre Searle y Pearson radica en el alcance de los teoremas de Gödel, particularmente en su aplicación a determinados procesos cerebrales. Pearson recurre a esos teoremas para argumentar en contra de la capacidad de explicar computacionalmente la conciencia, criterio que utiliza para restar valor al programa de inteligencia artificial débil. Pero Searle argumenta, como ya comenté al principio de este escrito, que esos teoremas pueden ser tenidos en cuenta para determinado tipo de algoritmos, los que tienen que ver con la resolución de teoremas matemáticos, pero no para otros. No tenemos evidencia *a priori* de que la forma en cómo proceden los procesos cerebrales sea mediante algoritmos de ese tipo, pues bien pudieran ser otros que, en todo caso, podrían llegar a mostrar que es verdadero o asumible aquello que al proceso cerebral y físico correspondiente se le muestra como verdadero. Es decir, que no habría diferencia entre el proceso cerebral y la simulación. Pero claro, esto

requiere un nivel de comprensión de los procesos cerebrales, los que permitirían su eventual simulación, que todavía no disponemos. Searle no descarta, por lo tanto, que llegado un nivel de conocimiento en nuestra investigación del cerebro seamos capaces, entiendo, de formular un algoritmo que, primero, no sea del mismo tipo que los que Pearson indica son los que operan en nuestro cerebro y, segundo, que pueda llegar a reproducir las mismas propiedades o *qualia* que el proceso cerebral que trata de simular. ¿Es esto suficiente para sostener que hemos llegado al fin de la historia? No, porque el algoritmo sigue siendo una secuencia sintáctica de símbolos. Por lo tanto, y tomando el ejemplo de la conciencia, aunque reproduzcamos las propiedades o *qualia* de la conciencia no podremos decir que tenemos un algoritmo o una entidad que 'tiene conciencia'. Nos sigue faltando algo. Y ese algo es el conocimiento detallado del sistema en cuestión, de su semántica.

Si ahora procedo con la generalización y remplazo conciencia por propiedad emergente tengo dos enunciados que son muy sugerentes dentro del programa de la inteligencia artificial débil:

- Uno: los procesos biológicos son capaces de generar determinadas emergencias (la vida, alguna propiedad clave de la evolución, la conciencia), y esos procesos son susceptibles de ser simulados por un computador.
- Dos: pero la simulación, por sí misma, no garantiza la emergencia de esas determinadas propiedades o, dicho de otro modo, que las simulaciones tengan exactamente las mismas características que las entidades simuladas.

TEOREMAS DE GÖDEL Y SÍSIFO

Es aquí donde creo que tienen una mejor interpretación los teoremas de Gödel, precisamente a nuestro continuo incremento del conocimiento y al hecho de que, una vez pensado que tenemos un sistema totalmente explicado, emergente, por ejemplo, pueden aparecer propiedades en él que no son explicables a partir de los supuestos de los que nos valemos para su comprensión: componentes elementales y reglas de funcionamiento.

Aplicados al mundo de los fenómenos físicos, los teoremas sostienen que aunque dispongamos del conocimiento de todas las reglas o leyes que regulan los procesos, incluidas determinadas emergencias, no podemos excluir el que puedan aparecer nuevas propiedades, nuevas conductas, nuevas emergencias, que no sean explicables en los términos de las reglas, leyes y componentes del sistema. Visto en perspectiva evolutiva, cabe la observación de que a lo largo de la historia de la vida han aparecido o emergido algunas, pocas o muchas, características que no serían explicables a partir de supuestas leyes ya disponibles. Tales características son ciertas o verdaderas porque existen, pero podrían no ser explicables a

partir de sistemas explícitos. Pero, y esto es lo importante, somos capaces de modificar el sistema, ampliar las leyes y, por lo tanto, incorporar tales novedades y hacer que lo que era inexplicable o indecible en el sistema previo deje de serlo en el nuevo. Esto supone un avance importante, pero no resuelve el mito de Sísifo, porque volveremos a encontrarnos con una nueva propiedad inexplicable en nuestro nuevo sistema que nos obligará a reconsiderar de nuevo el estado de nuestra ciencia particular. La ciencia tiene mucho del mito de Sísifo, de volver a empezar, aunque con la variante añadida de que la piedra que subimos a la montaña no la arroja-mos al mismo valle. Aquí estamos subiendo o ascendiendo en la cadena montañosa del conocimiento. Los teoremas de Gödel, y variantes poste-riores, no creo que sean incompatibles con el ejercicio de la ciencia. En la medida en que seamos capaces de generar un conjunto de reglas y prescribir los elementos que componen un sistema determinado estare-mos en condiciones de simular los procesos correspondientes. Llegado determinado momento de nuestro conocimiento, los procesos conducen-tes a determinadas emergencias serán suficientemente detallados; dispon-dremos del arsenal de reglas y componentes para poder decir que 'podemos simularlo' y tal emergencia será explicable dentro del sistema, porque además conoceremos con suficiente detalle la física y biología del proceso.

Las tesis de Searle sobre la conciencia son muy apropiadas, al igual que la forma en cómo he presentado las formulaciones de Gödel, para enten-der la dinámica de la vida, tanto su origen como su evolución, en la medida que ellas y otros fenómenos complejos tienen características similares al proceso mental que genera la conciencia. Por lo tanto, las emergencias evolutivas, algunas de ellas al menos, entran perfectamente en el marco de lo que Searle denomina programa de inteligencia artificial débil. A mi juicio, ese programa es el más próximo a la forma en cómo se hace biología, donde siempre se deja abierta la puerta a la incertidumbre, a ciertos fenómenos que no son predecibles o a la posibilidad de que puedan aparecer sin que tengamos capacidad alguna, desde un sistema de reglas y componentes determinados, de poder predecirlos (Moya, et al. 2009). Y lo que es más importante, este no es un programa incompatible con la teoría evolutiva. Probablemente esta forma de reflexión no aboga por una nueva teoría del cambio biológico. A ciencia cierta, todos los elementos necesarios para comprender el proceso evolutivo están encima de la mesa, y es cuestión de tiempo el que vayamos componiendo el *puzzle* de forma más acabada y dando cumplida cuenta de cómo han aparecido emergen-cias, aunque no podamos hacer predicciones de evaluación futura si dejamos que el sistema 'siga su curso natural' sin intervenirlo. La gran fuente de cambio o novedad procede de la génesis interna de variación genética, pero tampoco podemos decir que el curso de la historia de la vida

depende de esta emergencia, porque existen condiciones que limitan sustancialmente el que puedan tener éxito, algo que viene dictado por el juego de la selección natural.

AGRADECIMIENTOS

Este trabajo ha sido financiado por el proyecto europeo *Tarpol*.

LITERATURA CITADA

- Crick, F. (1994), *La búsqueda científica del alma*. Barcelona: Círculo de Lectores.
- Edelman, G. (1989), *Remembered Present: a Biological Theory of Consciousness*. New York: Basic Books.
- Moya, A., Krasnogor, N., Peretó, J. and Latorre, A. (2009), "Goethe's dream. Challenges and opportunities for synthetic biology", *EMBO Reports* 10: S28-S32.
- Penrose, R. (1996), *Las sombras de la mente*. Barcelona: Grijalbo.
- Dennett, D. (1995), *La conciencia explicada*. Barcelona: Paidós.
- Searle, J.R. (1985), *Mentes, cerebros y ciencia*. Madrid: Cátedra.
- Searle, J.R. (2000), *El misterio de la conciencia*. Barcelona: Paidós.