El País, 10 de marzo de 1999
Base, Sociedad, pág. 34 - Opinión
FÍSICA - COMPUTACIÓN
ROGER PENROSE "Lo grande, lo pequeño y la mente humana" es el título del nuevo libro del matemático y cosmólogo británico Roger Penrose, que publica en España la editorial Cambridge University Press. En él, Penrose, entre otros temas, aborda las características comunes a la física, las matemáticas y la consciencia, discutiéndolas con colegas como Stephen Hawking, y llega a la conclusión de que el pensamiento es algo más que pura computación.

La explicación científica de la consciencia

Permítanme decir algo ahora sobre la *consciencia humana.* En particular, ¿es ésta una cuestión sobre la que deberíamos pensar en términos de explicación científica? Mi punto de vista personal es decididamente que sí deberíamos hacerlo. En particular, tomo muy en serio la flecha que une el mundo físico con el mundo mental. En otras palabras, estamos ante el reto de comprender el mundo mental en términos del mundo físico.

He resumido algunas características del mundo físico y el mundo mental en la figura. En el lado derecho tenemos aspectos del mundo físico: éste es percibido como si estuviera gobernado por leyes físicas y matemáticas precisas, como se discutió en los dos primeros capítulos. En el lado izquierdo situamos la consciencia, que pertenece al *mundo mental*, y palabras como
alma, espíritu, religión
y demás, aparecen con frecuencia. En la actualidad, se prefieren explicaciones científicas para las cosas. Más aún, se tiende a pensar que sería posible, en principio, introducir cualquier descripción científica en un ordenador. En consecuencia, si uno tiene un modelo matemático de algo, debería ser capaz, en principio, de programarlo en un ordenador. Esto es algo contra lo que argumentaré enérgicamente en este capítulo, a pesar de mi sesgo
fisicalista
.

Los términos utilizados para describir las leyes físicas en la figura son *predecible, calculable* -éstos tienen que ver con el hecho de si existe o no determinismo en nuestras leyes físicas y si podríamos o no utilizar un ordenador para simular la acción de estas leyes-. Por un lado, existe la opinión de que cualidades mentales como la emoción, la estética, la creatividad, la inspiración y el arte son ejemplos de cosas que serían difíciles de ver emergiendo de algún tipo de descripción computacional. En el extremo
científico
opuesto, algunas personas dirían: "nosotros somos simples ordenadores; quizá no sepamos cómo describir estas cosas aún, pero, de algún modo, si supiéramos el tipo correcto de cálculos que habría que realizar, seríamos capaces de describir todas las cualidades mentales enumeradas en la figura". A menudo se utiliza la palabra
emergencia
para describir este proceso. Estas cualidades
emergen

según estas personas, como resultado del tipo correcto de actividad computacional.

¿Qué es la *consciencia?* No sé cómo definirla. Pienso que no es éste el momento de intentar definir la consciencia, puesto que no sabemos qué es. Creo que es un concepto físicamente accesible; pese a todo, definirlo sería probablemente definir la cosa equivocada. Sin embargo, voy a describirla hasta cierto grado. Me parece que hay al menos dos aspectos diferentes en la consciencia. Por un lado, existen manifestaciones pasivas de la consciencia, que implican *conocimiento*
.
Utilizo esta categoría para incluir elementos como la percepción del color, de la armonía, el uso de la memoria, y así sucesivamente. Por otro lado, están sus manifestaciones *activas* que implican conceptos como el libre albedrío y la realización de acciones voluntarias y conscientes. La utilización de tales términos refleja aspectos diferentes de nuestra consciencia.

Aquí me concentraré principalmente en aquello que implica la consciencia de un modo esencial. Es diferente de los aspectos pasivo y activo de la consciencia, y quizá está en algún lugar entre ambos. Me refiero al uso del término *comprensión* o quizá *intuición directa* que a menudo es una palabra más adecuada. No voy a definir ninguno de estos términos; no sé lo que significan. Existen otros dos términos que tampoco entiendo:
conocimiento e inteligencia
. Entonces, ¿por qué estoy hablando sobre cosas que no sé lo que significan realmente? Probablemente se debe a que soy un matemático y los matemáticos no pensamos tanto acerca de estos conceptos. No necesitamos definiciones precisas de los elementos que estamos manejando, con tal de que podamos decir algo sobre las
relaciones
entre ellos.
El primer punto clave aquí es que opino que la inteligencia es algo que requiere comprensión. Utilizar el término inteligencia en un contexto en el que negamos que esté presente cualquier comprensión me parece poco razonable. De un modo análogo, comprender sin ningún conocimiento es también algo absurdo: comprender requiere algún tipo de conocimiento. Éste es el segundo punto clave. De ello se deduce que la inteligencia requiere conocimiento. Aunque no estoy definiendo ninguno de estos términos, me parece razonable insistir sobre estas relaciones entre ellos.

Se podrían adoptar diversos puntos de vista acerca de la relación entre pensamiento consciente y computación. He resumido (...) cuatro aproximaciones al conocimiento, que he designado A, B, C y D.

A) Todo pensamiento es computación; en particular, las sensaciones de conocimiento consciente son producidas por la mera ejecución de cálculos apropiados.

B) La consciencia es una característica de la acción física del cerebro y, mientras que cualquier acción física puede ser simulada computacionalmente, la simulación computacional no puede por sí misma producir consciencia.

C) La acción física apropiada del cerebro produce consciencia, pero esta acción física no

puede siquiera ser propiamente simulada computacionalmente.

D) La consciencia no puede ser explicada en términos físicos, computacionales o cualesquiera otros términos científicos.

El punto de vista que llamo A, a veces denominado inteligencia artificial fuerte (1A fuerte) o funcionalismo (computacional), afirma que todo pensamiento es simplemente la realización de ciertos cálculos y, en consecuencia, si uno realiza los cálculos apropiados, el resultado será el conocimiento.

He designado como B al segundo punto de vista y, según éste, en principio, sería posible simular la actividad de un cerebro cuando su propietario es conocedor de algo. La diferencia entre A y B es que, aunque dicha actividad puede ser simulada, esa mera simulación no podría en sí misma, según B, tener ningún sentimiento o ningún conocimiento -hay algo más, que quizá tiene que ver con la construcción física del objeto-. Por eso, un cerebro compuesto de neuronas y similares podría ser conocedor, mientras que una simulación de la actividad de dicho cerebro no sería conocedora. Éste es, hasta donde puedo alcanzar, el punto de vista defendido por John Searle.

A continuación está mi propia manera de ver las cosas, que he llamado C. Estoy de acuerdo con B, hay algo en la actividad física del cerebro que provoca conocimiento -en otras palabras, hay algo en la física a lo que tenemos que dirigirnos-, pero esta actividad física es algo que ni siquiera puede ser simulado computacionalmente. No hay ninguna simulación que pudiera realizarse de dicha actividad. Esto requiere que exista algún factor en la actividad física del cerebro que esté más allá de la computación.

Finalmente, siempre existe el punto de vista D, según el cual es un error considerar estas cuestiones en términos de ciencia. Quizá el conocimiento no pueda ser explicado en términos científicos.

Más que computación

Yo soy un defensor del punto de vista C. Existen, sin embargo, diversas variedades de C. Está lo que podría denominarse C *débil* y C *fuerte*. C *débil* propone que, de algún modo, con la física conocida, bastaría sólo con mirar con mucho cuidado para encontrar ciertos tipos de actividad que están más allá de la computación. Cuando digo "más allá de la computación", tengo que ser un poco más explícito, como lo seré en un instante. Según C
débil
no hay nada que necesitemos buscar fuera de la física conocida para encontrar la actividad no computacional apropiada. C
fuerte,
por el contrario, requiere que exista algo fuera de la física conocida; nuestra comprensión física es inadecuada para la descripción del conocimiento. Es incompleta, y, como ustedes habrán deducido del capítulo 2, yo sí creo que nuestra imagen física es incompleta, como indiqué en la figura 2.17. Desde el punto de vista de la C
fuerte
quizá la ciencia futura explicará la naturaleza de la consciencia, pero la ciencia actual no lo

hace.

Incluí algunas palabras en la figura 2.17 que no comenté en ese momento, en particular, el término *computable*. En la imagen estándar, uno se encuentra con física básicamente computable en el nivel cuántico, y el nivel clásico es, probablemente, computable, aunque hay cuestiones técnicas sobre cómo pasar desde sistemas discretos computables a sistemas continuos. Es un punto importante pero permítanme que no me ocupe aquí de ello. De hecho, me parece que los defensores de la C *débil* tendrán que encontrar algo con estas incertidumbres, algo que no pueda explicarse en términos de una descripción computable.

Para pasar del nivel cuántico al nivel clásico en la imagen convencional, introducimos el procedimiento que he llamado R, y que es una operación completamente probabilista. Lo que tenemos entonces es computabilidad junto con aleatoriedad. Voy a argumentar que esto no es suficiente: necesitamos algo diferente, y esta nueva teoría, que construye un puente entre estos dos niveles, tiene que ser una teoría no computable. Diré algo más sobre lo que entiendo por este término en un momento.

Ésta es mi versión de la C *fuerte*: buscamos la no computabilidad en la física que forma un puente entre los niveles cuántico y clásico, lo cual es pedir bastante. Estoy diciendo que no sólo necesitamos una nueva física, sino que también necesitamos una nueva física que sea relevante para la actividad del cerebro.