



FACULTADE DE INFORMÁTICA DE A CORUÑA

TESIS DOCTORAL

**Título: El Problema de la Emergencia
de Propiedades desde la perspectiva
de la IA (Ciencia Cognitiva).**

**DEPARTAMENTO DE TECNOLOXÍAS DA INFORMACIÓN E AS
COMUNICACIONES**

Autor: Cesareo Barreiro Sorrivas
Director/es: Juan Ares Casal
Aurora Martínez Rey

2012

Resumen:

Esta tesis trata sobre la conciencia dentro de la ciencia cognoscitiva y la Inteligencia Artificial (IA).

A lo largo de la historia ha habido distintas interpretaciones para explicar la mente y sus atributos como la conciencia.

Algunas teorías hacen de la conciencia una propiedad emergente del cerebro.

En este trabajo se expone la emergencia de propiedades en el marco de una nueva ontología del universo (la teoría EMI) y mediante unos constructos (holones e informones) proporcionándose ejemplos de casos computacionales.

Se concluye que la emergencia funcional de propiedades, algo necesario para entender la mente, es asequible mediante métodos computacionales.

Abstract:

This thesis is about consciousness within cognitive science and artificial intelligence (AI). Throughout history there have been different interpretations to explain the mind and its attributes as consciousness.

Some theories make consciousness an emergent property of the brain.

This work exposes the emergence of properties under a new ontology of the universe (EMI theory) and by some constructs (holons and informons) being provided examples of computing.

It is concluded that the emergence of functional properties, which is necessary to understand the mind, is affordable by computational methods.

Resume:

Esta tese é sobre a conciencia na ciencia cognitiva e Intelixencia Artificial (AI).

Ao longo da historia houbo diferentes interpretacións para explicar a mente e os seus atributos como a conciencia.

Algunhas teorías fan da conciencia unha propiedade emerxente do cerebro.

Este traballo expón a emerxencia de propiedades desde unha nova ontoloxía do universo (teoría EMI) e mediante uns constructos (holóns e informóns) sendo proporcionados exemplos de casos computacionais.

Conclúese que a emerxencia de propiedades funcionais, algo que é necesario para entender a mente, é accesible por métodos computacionais.

AGRADECIMIENTOS

A mis padres quiero dedicarle el fruto de este trabajo, con un especial recuerdo a mi padre, fallecido recientemente, que no ha podido llegar a ver el final de este proyecto.

Mi mujer, Isabel, y mis hijos, Andrés y Xaquín, saben que, debido a la realización de este trabajo, he tenido que alejarme de ellos más de lo que quisiera y, a pesar de ello, no me lo han tenido en cuenta.

Mi hermano Chema ha conseguido que me atreva con algo que inicialmente no me sentía capaz de hacer; gracias a su ánimo y al empujón, en momentos de incertidumbre, el proyecto está finalizado.

También quiero agradecer a mis directores de tesis profesores Doctores Juan Ares Casal y Aurora Martínez Rey, por su paciencia para conmigo y todo el apoyo y asesoramiento recibido; siempre con buen ánimo y sin una palabra descortés, a pesar de la lata que les he ido dando; sus consejos y orientaciones han servido para que no me perdiese y pudiese ver la luz al final del túnel de este proceso.

De igual forma quisiera reconocer la inestimable ayuda de los grupos de investigación de la Facultad de Informática de la Universidad de A Coruña y de la Facultad de Informática de la Universidad Politécnica de Madrid, por su asesoramiento, y el uso de material y aportaciones que resultaron fundamentales en la elaboración de esta tesis.

Un especial agradecimiento va dirigido al profesor Juan Pazos Sierra quien, con sus amenas y doctas explicaciones que transmiten su sabiduría, ha conseguido que aumente mi interés por el tema y que haya comprendido mejor, no solo el tema de la conciencia sino que además aspectos como la educación, la cultura la neurociencia y la ciencia en general.

ÍNDICE DE CONTENIDOS

1. INTRODUCCIÓN	7
1.1. El problema mente-cuerpo:	10
1.1.1. Antecedentes	10
1.1.2. El enfoque filosófico	11
1.2. El lugar del ser humano en el Universo	14
1.3. El problema	19
1.4. Conciencia y física cuántica.....	49
1.5. La única prueba “objetiva” de la conciencia.....	53
2. ESTADO DE LA CUESTIÓN.....	57
2.1. Evolución de la cuestión	57
2.2. Enfoques.....	63
2.2.1. Filosófico.....	63
2.2.2. Científico	72
2.3. Reduccionismo versus Emergencia.....	96
2.3.1. Reduccionismo	96
2.3.2. Emergencia.....	107
2.4. Teoría del caos y conciencia	110
2.5. Mente y Computador.....	118
2.6. Características del estado consciente	126
2.6.1. Problemas fáciles.....	132
2.6.2. El Problema Arduo o Duro.....	133
2.7. Resultados obtenidos a partir del estado de la cuestión	137
2.7.1. Tipología de los investigadores de la conciencia	137
3. EL PROBLEMA ABORDADO Y LA SOLUCIÓN	147
3.1. El marco conceptual: La ontología del universo. El Modelo EMI.	147
3.1.1. La información el tercer constructo del Universo.....	147
3.1.2. El Modelo EMI.....	152

3.1.3. Equivalencia entre Materia-Energía-Información	158
3.1.4. Ejemplos de la Relación Energía-Información.....	161
3.2. Los Constructos: Holones e Informones	168
3.2.1. Informón.....	168
3.2.1.1. Definición.....	168
3.2.1.2. Niveles y elementos de información	169
3.2.1.3. Aspectos de la información	170
3.2.2. El “Holón”	173
3.2.2.1. La parábola de los relojeros. Definición de holón	173
3.2.2.2. Características de los holones	178
3.2.2.3. Niveles holónicos	181
3.2.2.4. Dominios de cooperación y colaboración holónica	183
3.2.2.5. Holones versus Agentes	187
3.2.3. Postulados de la Teoría	189
3.2.4. Experimentación.....	190
3.3. Emergencia de propiedades.....	193
3.3.1. Emergencia o <i>fulguratio</i>	193
3.3.2. Clasificación de fenómenos emergentes: <i>Simplejidad y Complicidad</i>	199
3.3.3. Sistemas Constitutivos.	205
3.3.4. Emergencia de Propiedades: Casos de estudio Computacionales.....	209
4. RESULTADOS, CONCLUSIONES Y FUTURAS LÍNEAS DE TRABAJO.....	219
4.1. Resultados	219
4.2. Conclusiones	219
4.3. Futuras líneas de investigación	220
5. BIBLIOGRAFÍA.....	221

1. INTRODUCCIÓN

La cuestión mente-cerebro fue planteada por primera vez el siglo V a. C., por Hipócrates y Platón. Hipócrates, el primer médico, además de desechar la superstición, basó su pensamiento en observaciones clínicas y afirmaba que todos los procesos mentales emanan del cerebro. Platón, que rechazaba observaciones y experimentos, creía que la única razón por la que los humanos piensan acerca de ellos mismos y su cuerpo mortal es que tienen un alma inmaterial e inmortal. Esta idea fue incorporada en el pensamiento cristiano el siglo XIII por Santo Tomás de Aquino quien elaboró, con otros pensadores, la idea de que el alma, el generador de la conciencia, no es sólo distinta del cuerpo, sino que es de origen divino.

Esta dualidad fue hecha explícita en el siglo XVII por Descartes cuando desarrolló la idea de que los seres humanos tiene una naturaleza dual: cuerpo, hecho de sustancia material y mente, que deriva de la naturaleza espiritual del alma. El alma recibe señales del cuerpo y puede influenciar sus acciones. El pensamiento de Descartes da origen a la visión de que ciertas acciones: comer, pasear, percepción sensorial, apetitos, pasiones e incluso algunas formas de aprendizaje, están mediatizadas por el cerebro y pueden estudiarse científicamente. Sin embargo, la mente es sagrada y como tal no es un sujeto adecuado para la ciencia. Es notable que estas ideas del siglo XVII estuvieran aún vigentes en 1980. Karl Popper y John Eccles se adhirieron al dualismo. Ellos coincidían con Santo Tomás en que el alma es inmortal e independiente del cerebro. Gilbert Ryle, el filósofo de la ciencia británico, se refería al alma como “el fantasma en la máquina”.

Para Freud, “lo que se entiende por conciencia no requiere discusión; está más allá de toda duda”. Nada más lejos de la realidad. De hecho, no existe consenso alguno en lo referente a la conciencia. Más aún, las posiciones defendidas pueden ser diametralmente opuestas.

Hasta los años 60 del siglo pasado, la psicología dominada por el conductismo evitaba el término “conciencia” en cualquier discusión que se preciara de científica. ¿Qué fue, entonces, lo que causó la explosión del interés por la conciencia en las décadas posteriores? Algunos atribuyen dicho interés a los llamativos avances en la imaginaria

cerebral que permitieron observar las partes del cerebro que se activan en respuesta a estímulos particulares. Pero, de acuerdo con un editor de “Journal of Consciousness Studies”, “es más probable que el resurgimiento de los estudios de la conciencia se debiera a razones sociológicas: los estudiantes de los años 60, que disfrutaron de una rica aproximación extraacadémica a los “estudios de la conciencia”, están ahora dirigiendo departamentos científicos”. Por otra parte, el interés en los fundamentos de la Mecánica Cuántica y la conexión con la conciencia ha florecido al mismo tiempo que los estudios de la conciencia.

Aunque realmente el verdadero principio del problema de la conciencia hay que remontarlo, según William James, en su obra de 1890 “Principios de Psicología”, (James, 1890), a la innegable experiencia del “monólogo interior”. El Yo es sólo una ilusión; es decir, algo que existe pero que no es lo que parece.

Hoy en día la mayoría de los filósofos de la mente coinciden en que lo que se denomina conciencia se deriva del cerebro físico, pero algunos discrepan en si puede abordarse científicamente. Así, Searle mantiene que el cerebro es la causa de la mente, insistiendo en que los estados mentales gozan de existencia independiente y no pueden ser “reducidos” a estados cerebrales. Unos pocos, tal como Colin McGinn, creen que la conciencia simplemente no puede ser estudiada, porque la arquitectura del cerebro establece limitaciones acerca de las capacidades cognitivas humanas. Con la visión de McGinn, la mente humana simplemente es incapaz de resolver ciertos problemas. En el otro extremo, filósofos tales como Daniel Dennett niega que exista cualquier problema en absoluto. Dennett argumenta, de un modo parecido al neurólogo John Hughlings Jackson, hace un siglo, que la conciencia no es una operación distinta del cerebro; más bien, es el resultado del trabajo computacional de áreas del más alto orden del cerebro concernidas con los últimos estados de procesamiento de la información. Finalmente, filósofos como John Searle y Thomas Nagel toman una posición intermedia, sosteniendo que la conciencia es un conjunto discreto de procesos biológicos. Los procesos son accesibles al análisis pero se han hecho pocos avances en entenderlos, porque son muy complejos y representan más que la suma de sus partes. Por lo tanto, la conciencia es mucho más complicada que cualquier propiedad del cerebro que se entiende.

En muchas ocasiones la mejor manera de resolver un problema es simplemente planteando las preguntas adecuadas. Y a veces la mejor manera de plantear las preguntas adecuadas es con un ejemplo que haga explícita la naturaleza del problema. Para James la conciencia es un proceso: privado, selectivo y continuo pero en constante cambio.

“Supongamos”, decía Leibniz (Leibniz, 1981), que existe una máquina cuya estructura produce pensamientos, sentimientos y percepción; imaginemos esa máquina ampliada, pero conservando las mismas proporciones, de manera que podamos entrar en ella como si fuera un molino... ¿qué veríamos allí? Nada más que unas piezas que se empujan y mueven unas a otras, pero nunca nada que pueda explicar la percepción. El fin de Leibniz, con este experimento mental, es mostrar que el pensamiento, los sentimientos y la percepción no pueden explicarse mediante mecanismos, por meras partes y movimiento de partes, como proclaman los materialistas. En otras palabras, como aseveran los dualistas, la mente es más que el cerebro. Pero hoy se sabe que los estados mentales tienen correlaciones físicas, eléctricas y bioquímicas, en el cerebro. Sin embargo, un cerebro parece que sólo puede indicar que piensa, no qué se está pensando. Además, la experiencia real del pensamiento no es en absoluto el estado psicológico del estado del cerebro. El experimento mental sólo prueba que el pensamiento, sentimiento y percepción no son procesos visibles a los demás, y no se sabe cómo serían *“visibles”* a uno mismo. La mayor crítica a este experimento es que supone que el todo es más que la suma de sus partes. David Cole (Cole, 1984), propuso un contraejemplo para falsar el experimento mental de Leibniz. Imagínese una gota de agua ampliada de tamaño de modo que cada molécula sea del tamaño de una rueda de molino. Si uno entra en tal gota, podrá ver cosas maravillosas pero nada *“húmedo”*. Es decir, el agua no necesita constar de algo adicional a moléculas de H₂O para tener humedad, justo como la máquina de Leibniz, el cerebro, no necesita algo adicional a las neuronas, los neurotransmisores y la glía para producir pensamientos.

Samuel Johnson decía (Minsky, 1987): *La materia sólo puede diferir de la materia en forma, masa, densidad, movimiento y dirección del movimiento. ¿A cuál de estas cosas, por variadas o combinadas que estén, se puede asociar la conciencia? Ser redondo o cuadrado, sólido o fluido, grande o pequeño, moverse despacio o deprisa, en un sentido u otro, son modos de existencia material, todos igualmente ajenos a la naturaleza de la cognición.*

1.1. El problema mente-cuerpo

1.1.1. Antecedentes

La generación de seres artificiales dotados de inteligencia, se remonta a los griegos y chinos. Entre los primeros, Homero, en *La Iliada* (Homero, 1996), cuenta como los “esclavos” de Hefastos o Hefesto “tenían inteligencia en sus corazones, hablaban y habían aprendido de los dioses inmortales cómo hacer las cosas”. Historias similares aparecen en la China de aquella época (Needham, 1981). King Mu, de la dinastía Chou, tenía un robot inmortal, con carne y sangre artificiales, mostrando la gran fascinación china con la vida generada por medios químicos.

Esta idea encuentra nueva vida, nueve o diez siglos después, entre los árabes en “*Ilmaltakwin*”, “la ciencia de la generación artificial”, que eventualmente conducirá al “*Al-iksir*” o elixir de la vida de los alquimistas europeos medievales. Estos árabes pueden haber sido los primeros en establecer formalmente la distinción entre sustancias naturales y artificiales. Esta distinción, hecha por los árabes, no implicaba ninguna superioridad de lo natural sobre lo artificial, sino sólo que era diferente. Sin embargo, aseveraban que no tan diferente, por lo cual preconizaban que un excelente medio de conocer lo natural era mediante el estudio de lo artificial.

Aunque esta premisa subyace a la idea de modelización común en la ciencia occidental, la misma no tenía sentido alguno en la China de aquel entonces, ya que en esta cultura la realidad era apreciada como un todo, animado o inanimado, natural y artificial, como aspectos indisolublemente conectados. *Esto se ve en la religión de Min Tao (taoísmo) según la cual, y de acuerdo al "Ying-Yang", la realidad tiene dos aspectos contrarios pero complementarios, tales como: día-noche, masculino-femenino, natural-artificial; etc. Esta síntesis del taoísmo, que tanta influencia tuvo y tiene sobre el pensamiento chino, está en los antípodas del enfoque analítico occidental.*

La idea árabe se trasladó a la Europa medieval. El papa Silvestre II, Alberto Magno, Roger Bacon, etc., afirmaban que habían construido cabezas de cobre o bronce que predecían el futuro y resolvían problemas difíciles. Raimundo Lull, a partir de la máquina pensante árabe, denominada “*zairja*” la reelaboró en su *Ars Magna* (Lull, 1305). Ambas,

fueron modeladas en el supuesto de que el pensamiento podría ser llevado a efecto fuera del cerebro y de la mente humana; esto es, verdaderamente mecanizado.

1.1.2. El enfoque filosófico

A) La causalidad mental: Eliminacionistas y Realistas

Existe una propuesta de explicación exclusivamente materialista (los griegos) que dio origen al conocido debate animismo-mecanicismo. Como ocurre en otros casos, en Descartes (ej: cogito ergo sum) se postula lo que se pretende demostrar.

Si la mente es una propiedad o proceso fundamental de los seres humanos, su naturaleza es elusiva. Las disputas del siglo XVII en torno al cuerpo y la mente son prueba fehaciente de ello. Descartes (Descartes, 1642) concluyó que eran muy diferentes. El filósofo judeo-español, afincado en Holanda, Spinoza (Spinoza, 1951), rechazaba el dualismo creyendo que la mente y el cuerpo humanos son aspectos de la misma cosa, dos atributos de Dios. *Aunque Spinoza no sostenía la creencia en un Dios personal, sino justamente, un Dios que se manifiesta en la naturaleza, siendo todas esas manifestaciones las que definen su esencia.* Leibniz también ponderó el asunto mente-cuerpo y eventualmente postuló que eran esencialmente distintos, y no obstante exactamente equiparables, mónadas. Filósofos como Locke y Hume, influenciados por Newton, se dedicaron a buscar un entendimiento racional de la mente, intentando adivinar leyes análogas a las de Newton no relativas a la materia sino a la mente. Julien La Mettrie se burlaba de estos ambiciosos pero fútiles intentos calificándolos de palabras “sin sustancia” y “especulación sin conocimiento”. En resumen, nueva retórica. En su libro “L'Homme Machine” (La Mettrie, 1987), basándose en su práctica como médico, propuso una teoría comprensiva de la mente y ponía énfasis en un argumento simple y eficaz: la evidencia de que, circunstancias físicas, como dieta, preñez, drogas, fatiga, enfermedad, etc., afectaban inobjetablemente la mente. Indudablemente, la contribución de La Mettrie al debate mente-cuerpo es sustantiva dado que fue el primero en citar evidencia empírica en apoyo de su teoría. Puede decirse con propiedad que su trabajo marcó el principio del fin del “amateurismo” en el entendimiento del pensamiento humano.

La complementariedad, la describió Oppenheimer (Dubois, 1986), como “aquel orden que supone que las partes encajan en un todo y que el todo requiere de las partes”. El enfoque filosófico al problema mente-cuerpo se centra en la causalidad mental. El problema filosófico estriba en explicar el lugar de la mente en el Universo o, lo que es lo mismo, dar una explicación de la naturaleza de la mente que sea compatible con una concepción materialista del mundo. En efecto, si se acepta que los elementos últimos constituyentes de los objetos del mundo son materiales, la idea misma de algo cuyas propiedades son radicalmente distintas de los objetos materiales resulta, *prima facie*, problemática. Entre esas propiedades están las dos siguientes:

- a) Capacidad de representación del mundo, como en la creencia de que hay un traje azul en el armario, o el deseo de que lo haya.
- b) La capacidad de tener determinadas sensaciones, como cuando alguien nota un dolor de muelas o le gusta la música.

Lo crucial en este punto es, en primer lugar, dilucidar si o no ambas propiedades son exclusivamente, no específicamente, mentales. Y si así es, si todo lo que existe es material, esto es, está constituido por elementos materiales, ¿cómo puede explicarse la existencia de propiedades tan peculiares? Los antecedentes del problema, o mejor, la “primera” propuesta de solución se remonta a Descartes (Descartes, 1644): Lo mental -*res cogitans*- y lo material -*res extensa*- son sustancias con propiedades mutuamente excluyentes; respectivamente el pensamiento y la extensión que, no obstante, se relacionan causalmente entre sí: lo mental afecta causalmente a lo corporal y viceversa. Esta solución ha sido rechazada al considerar que la postulación de relaciones causales entre sustancias esencialmente distintas es más un recurso a la magia que una explicación genuina. Sin embargo, paradójicamente, con ello acabaron por aceptar el planteamiento cartesiano: explicar el lugar de la mente en el mundo es explicar cómo lo mental puede actuar causalmente sobre lo material.

La cuestión que se plantea es: ¿por qué hay un problema filosófico, especialmente acerca de la causalidad mental? La respuesta estriba en, por una parte, la aceptación de la sobredeterminación causal en la explicación de la conducta humana y, por otra parte, la negación de la clausura causal del mundo físico. En efecto, supóngase que alguien va al frutero porque tiene hambre y cree que allí hay una manzana. Obviamente, ésta es una

explicación de la conducta en virtud de causas mentales; en este caso concreto, la sensación de hambre y la creencia de que hay algo comestible, una manzana, en el frutero. Pero evidentemente, no es la única explicación posible, puesto que podría darse una explicación causal de la misma conducta que no apelara a excitaciones neuronales, secreciones hormonales, etc., del cuerpo humano del sujeto en cuestión. Siendo así, habría dos explicaciones distintas, completas e independientes de la misma conducta, a saber, una en virtud de causas mentales y otra en virtud de causas neurofisiológicas. En este caso, la explicación de la conducta humana estaría sobredeterminada causalmente; es decir, en general, cualquier acción humana tendría más causas de las que son necesarias. Esto, filosóficamente al menos, es inaceptable pues atenta contra la regla generalmente aceptada, según la cual un fenómeno cualquiera salvo situaciones excepcionales e insólitas, tiene sólo una causa, *(aunque esta causa puede ser compuesta de varias causas, pero continúa siendo conceptualmente única)*.

Por el contrario, en filosofía, el problema de la subdeterminación, opuesto al anterior, es equivalente a lo que en ciencia es establecer que “casi siempre hay modelos alternativos para cualquier conjunto de supuestos” (Paulos, 1999). Por ello, uno de los errores más comunes en la ciencia y por supuesto en la filosofía es tomar el resultado de un acontecimiento como una consecuencia del mismo, tal y como afirmó Gastón de Levis en sus “Pensamientos escogidos” (Gaston de Levis, 1812)

Por otra parte, los filósofos de la mente actuales, han aceptado generalmente la clausura causal del mundo físico; esto es, han aceptado que en la explicación de cualquier fenómeno físico no hace falta apelar más que a causas físicas. O sea, si se quiere explicar causalmente el conjunto de eventos neurofisiológicos y hormonales de “ir al frutero”, sólo son necesarias causas físicas como las que tienen que ver con conexiones sinápticas excitatorias, posterior secreción hormonal, actividad del aparato locomotriz, etc. Pero en este caso, es fácil ver que la idea misma de causalidad mental se vuelve en sí misma problemática. Efectivamente, como regla general, cualquier fenómeno tiene una sola causa y la causa de cualquier fenómeno físico es estrictamente física, “ergo” lo mental carece de poder causal sobre el mundo físico. Con esto, parece diluirse, no sólo la posibilidad de una explicación de la conducta humana en virtud de creencias, deseos, sensaciones, etc., sino también el lugar mismo que la mente ocupa en el mundo.

Algunos autores, denominados “eliminacionistas”, como Paul y Patricia Smith Churchland, *han llegado a asumir una postura aún más radical, consistente en la eliminación de lo mental, en la medida en que creencias, deseos o sensaciones son sólo términos teóricos de malas explicaciones de la conducta humana.* Otros autores, denominados “instrumentalistas” como Daniel Dennett, han considerado que las creencias, sensaciones o deseos son sólo instrumentos útiles en la explicación de la conducta humana. Sin embargo, la mayoría de los filósofos de la mente, a los que se les acostumbra llamar “realistas”, buscan preservar la doble intuición de que lo mental existe y es causalmente eficaz.

1.2. El lugar del ser humano en el Universo

El matemático y filósofo Federico Enriques (1871-1946), al indagar las formas más dispares de una reacción ante la ciencia (Enriques, 1912), escribió que el rasgo común de estos movimientos consiste en el malestar provocado cuando *algunas doctrinas bastantes generales de la ciencia semejan invalidar la base de algún valor supuestamente duradero.* Los ejemplos abundan: Anaxágoras destruyó el mito de Apolo al ver en el Sol una simple *masa incandescente mayor que el Peloponeso.* Copérnico y Galileo, al negar *la inmovilidad y la posición central de la Tierra,* abrieron camino *a la duda sobre la estabilidad del mundo habitado.* La biología y la economía también han hecho de las suyas: La evolución de Wallace y Darwin *ofendió el orgullo de la estirpe humana;* mientras que a las tristes previsiones de la ciencia económica no les inquieta *herir los intereses de las masas o de las clases dominantes.* En suma, Enriques fue el primero en hablar del triple descentramiento que la ciencia hizo del papel del ser humano. Estas ideas fueron “nostrificadas” por Freud (Freud, 1975), al hablar del drama o de las tres grandes “heridas” que el conocimiento científico moderno infligió al narcisismo y la megalomanía humana como de otros tantos descentramientos de las imágenes a través de las cuales los seres humanos pensaron su papel en el Cosmos.

El primero, debido al heliocentrismo de Copérnico (Copérnico, 1507-1977), que despojó a la Tierra de su estatus de centro del Universo, y transformó a los seres humanos, antes habitantes del centro del Cosmos, en moradores de un pequeño o insignificante planeta periférico de una galaxia de la periferia. El segundo, debido a la “Teoría de la Evolución” de Alfred Russell Wallace (Wallace, 1858) y Charles Darwin (Darwin, 1859), despojó a

los seres humanos de su condición de reyes de la creación, y convirtió a la especie humana, antes un punto “singular” del designio del Creador, en la salida contingente, prescindible y reciente de una evolución con múltiples ramificaciones. El tercero, provocado por los “descubrimientos” de Sigmund Freud, sustrae a la conciencia del hombre el control y el poder sobre el conjunto de su propia vida psíquica: el hombre ni siquiera es dueño de su casa.

En efecto, durante muchos siglos, y siguiendo el aforismo del sofista griego, paisano de Demócrito, Protágoras (Protágoras, 1997) de que el hombre era la medida de todas las cosas, el ser humano se consideraba el ombligo del mundo. Esta posición privilegiada que el mismo se otorgó, recibió el primer envite de Aristarco. Posteriormente, cuando en 1507, Nicolás Copérnico difundió en copias manuscritas el primer esbozo de su sistema heliocéntrico, conocido con el nombre de “*Commentariolus*” (*De Hypothesibus Motuum Coelestium a se Constitutis Commentariolus*) remató la faena, a mediados del siglo XVI, concretamente cuando hizo perder a la Tierra su posición central en el Universo (Copérnico, 1997). Y la cosa fue a peor, puesto que no sólo la Tierra no era el centro del Universo, sino que sólo es un suburbio del Sistema Solar. Situado éste en un brazo de una galaxia (La Vía Láctea), que contiene otros cientos de millones de sistemas solares. Galaxia que, a su vez, ocupa poco más de una billonésima parte del Universo conocido que, para más inri, no es sino uno entre una infinidad de universos.

Más precisamente, Aristarco, Copérnico, Kepler y Galileo, apartaron a la humanidad del centro del Sistema Solar. El hombre no es el ombligo del Universo. El “hogar” de la humanidad, la Tierra, no es el punto central en torno al cual giran los demás cuerpos celestes. Antes bien, resulta ser un pedrusco irrelevante en la inmensidad sideral, hasta tal punto, que las historias de visitas de extraterrestres no son más que los últimos coletazos de antropocentrismo residual. En palabras de Stephen Hawking (Deutsch, 1997): *The human race is just a chemical scum on a moderate-sized planet, or biting round a very average star in outer suburb of one among a hundred billion galaxies*. Es decir, lentamente se había ido elaborando una imagen lucreciana del Universo, que durante un siglo por lo menos, hasta el barón d’Holbach, constituirá una alternativa al deísmo y la imagen newtoniana del Universo.

Es preciso que los hombres, había dicho Pierre Borel (Borel, 1657), aprendan a dejar de ser como esos campesinos que no han visto nunca una ciudad y durante toda su vida siguen pensando que no puede haber nada más grande ni más bello que su pequeño pueblo. La Tierra entera se configura ahora únicamente como una provincia o un pueblo del Universo. De manera parecida a lo que había ocurrido, en el Mediterráneo y el Occidente ante los descubrimientos geográficos y los viajes a continentes desconocidos y pueblos lejanos.

Ya, previamente, Christian Huygens había proporcionado una visión parecida. Huygens, murió en 1695 dejando inédito el manuscrito “*Cosmotheoros sive de terris coelestibus earumque ornatu conjecturae*”, publicado en 1698. Según Huygens (Huygens, 1850-1888), *ni Nicolás de Cusa, ni Bruno ni Fontenelle habían llevado a cabo una investigación seria acerca de los habitantes de los otros mundos. Y, sin embargo, las vías que conducen al conocimiento de cosas muy lejanas no están cortadas, y hay materia para elaborar una serie de conjeturas verosímiles.* Luego da como razones para ello la de no imponer límites a la curiosidad humana y que la búsqueda de teorías verosímiles es la esencia misma de la física. Huygens opone a las extrañas teorías de Kepler la tesis “bruniana”, de la identidad de la naturaleza entre el Sol y las estrellas. Huygens, para apoyar su posición, se sitúa mentalmente en un punto del Universo equidistante del Sol y de las estrellas fijas más cercanas y, desde este punto de vista, considera al Sol y a la Tierra, que se ha vuelto invisible. Es éste un tipo de “experimento mental” que no pertenece a la misma clase que los experimentos mentales habituales en la filosofía natural de Galileo. Supone distanciarse de un punto de vista terrestre o heliocéntrico en la consideración del cosmos, una especie de relativismo cosmológico, que se va abriendo paso en las mismas fechas del nacimiento del relativismo cultural. Esto decía Huygens: *Conviene que nos consideremos como si estuviéramos fuera de la Tierra y fuéramos capaces de observarla desde lejos. Podemos entonces preguntarnos si es cierto que la naturaleza ha otorgado sólo a ella todos sus ornamentos. Así podremos entender mejor qué es la Tierra y cómo debemos considerarla. Del mismo modo que quienes realizan grandes viajes son mejores jueces de las cosas de su patria que los que no la han abandonado nunca.*

Las largas disputas sobre la infinitud del Universo y sobre la pluralidad y habitabilidad de los mundos contribuyeron, en un contexto cultural más amplio, no sólo a poner en solfa

todos las concepciones antropocéntricas y terrestres del Universo, sino también a vaciar de contenido el tradicional discurso de los humanistas sobre la naturaleza y dignidad del hombre. En suma, todo eso constituyó y contribuyó a la crisis y fin del antropocentrismo. Y, por si esto fuera poco, Zwicky mostró que la materia de que están hechos los humanos no es del tipo dominante. En efecto, más del 90% de la materia del Universo es hidrógeno, elemento que en los seres humanos está en un porcentaje muy inferior. Ante eso el hombre se rebela y grita a los cuatro vientos que si él no es el centro del Universo, que nadie lo sea o que lo sea todo el mundo, declarando, como lo hizo Einstein, equivalentes a efectos globales todos los puntos de observación del Universo. Esa es la esencia del Principio Cosmológico (PC).

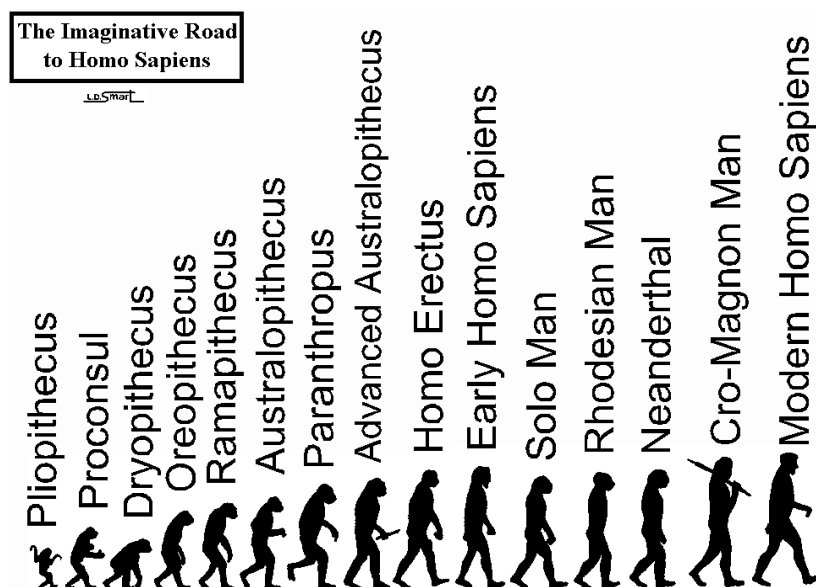


Figura 1.1. Evolución de la humanidad.

El segundo golpe al narcisismo humano provino de la mano de Wallace (Wallace, 1858) y Darwin (Darwin, 1859), codescubridores de la selección natural, al producir el descentramiento evolutivo del ser humano. Aunque Wallace, posteriormente vacilara en sus afirmaciones sobre la evolución lo que llevó a Darwin a escribirle, en la primavera de 1769, una carta en la que le decía: “Espero que no hayas matado del todo a tu hijo y al mío”. Wallace y Darwin hicieron del hombre un mero eslabón en el árbol evolutivo de las especies. Es decir, la humanidad, no está hecha a imagen y semejanza de ningún demiurgo. No es producto de la mano divina, al menos “prima facie”. Más bien, la humanidad es el humilde resultado de una evolución lenta e imparable que, como lo

muestra la figura 1.1, viene de muy antiguo y procede de lo más elemental de la naturaleza. Tampoco existe certeza de que dicha evolución se haya completado. Sin embargo, las predicciones de seres futuros con enorme y complejísimo cerebro e inteligencia superior, de nuevo no son más que delirios de grandeza residuales.

La violenta reacción emotiva que suscitaron los trabajos de Wallace y Darwin, al menos aparentemente, se han calmado a costa, eso sí, de que el hombre sólo se haya reconciliado parcialmente con su origen. En efecto, ha aceptado, aunque a regañadientes y no todos, su parentesco con otros seres vivos a condición de mantener una supremacía absoluta sobre ellos, tanto material como, sobre todo, moral. La primera, es el resultado concreto e indiscutible de su capacidad de aprendizaje genérico, de raciocinio; en fin, de su superioridad intelectual. La segunda, más polémica, como consecuencia de su inmensa capacidad de proyección en el mundo abstracto del pensamiento. Sus facultades creativas le envanecieron hasta tal punto que llegó a considerarse hecho a imagen y semejanza del mismísimo Dios todopoderoso, omnisciente y creador a partir de la nada. En una fase posterior, esa convicción se reforzó hasta el extremo de hacerle cambiar radicalmente los términos de la ecuación, invirtiendo las posiciones respectivas, representando al creador a imagen y semejanza del hombre, tal y como se muestra en el “Juicio Final” de Miguel Ángel, que puede contemplarse en la Capilla Sixtina en el Vaticano. Naturalmente, haciendo caso omiso de lo que decía al respecto el filósofo presocrático Jenófanes (Tales, 1995):

*Los mortales creen que los dioses han nacido
y que tienen vestido, voz y figura como ellos.
Pero si los bueyes, (caballos) y leones tuvieran manos
o pudieran dibujar con ellas
y realizar obras como los hombres,
dibujarían los aspectos de los dioses y harían sus cuerpos,
los caballos semejantes a los caballos, los bueyes a bueyes,
tal como estuvieran la figura correspondiente (a cada uno).*

El tercero comenzó cuando Freud le asestó al egocentrismo humano la tercera puñalada de la historia. Claro que Platón lo había anunciado más de veinte siglos antes. El ser humano no es dueño de sus actos. Su voluntad no obedece a su mente privilegiada. De hecho, los humanos están a merced de instancias ingobernables incluso incognoscibles

que ni siquiera tienen procedencia sobrenatural, sino que están ahí dentro en algún lugar, ignoto e ignaro, de la mente humana. Realmente, el ser humano está más desamparado frente a sí mismo que los animales frente a la naturaleza. Al menos ellos cuentan con un bagaje instintivo completo que cubre todas las actuaciones mientras que los humanos desde la infancia son ya perversos polimorfos dependientes y dominantes. Y ésta es su servidumbre y su gloria.

Y el definitivo palo se lo está dando el hombre a sí mismo como consecuencia de su ilimitada confianza en su capacidad intelectual. Esta confianza y los éxitos conseguidos gracias a unas estrategias cada vez más elaboradas, le animaron a lanzarse no sólo al conocimiento del Universo y de sus propios orígenes, sino a la tarea, mucho más ardua, de desvelar el misterio de la consciencia, la inteligencia o, más genéricamente, de la mente; es decir, a alcanzar la anhelada meta de conseguir el conocimiento de su propio conocimiento. En esta partida, que se juega en un tablero difuso y cambiante, el hombre ha movido habitualmente las “piezas” que tiene para conseguir un resultado victorioso, pero el contrincante es un adversario formidable: la naturaleza, por lo que las probabilidades de éxito son mínimas o nulas, si se le hace caso a Dante Alighieri (Alighieri, 1992) cuando advertía:

Loco es quien espera que nuestra razón

Puede recorrer el infinito camino.

Curiosamente, y desde un punto de vista cronológico, lo que muestra esta “deshumanización” es, según lo expuso Bertrand Russell, en su ensayo “Ciencia y Religión” (Russell, 1935), que: *las ciencias han progresado en orden inverso al que cabía esperar. Se sometió al dictado de la ley primero lo que se hallaba más lejos de nosotros y luego, de manera gradual, lo que estaba más cerca: los cielos, la tierra, la vida animal y vegetal, el cuerpo humano y, por último, aunque de manera aún imperfecta, la mente.*

1.3. El problema

El catedrático de química de la universidad de Oxford Peter Atkins (Atkins, 2003) estableció explícitamente los dos problemas verdaderamente profundos que le quedan por resolver a la ciencia actual; a saber: Uno, el origen del Universo, su desarrollo y futuro.

Dos, la naturaleza de la conciencia, la propiedad más desconcertante de la materia. Yendo más allá, el físico y codescubridor del modelo en espiral del ADN y del código genético, Francis Harry Comptom Crick y la médica Rita Levi Montalcini, ambos premios Nobel de medicina, señalan que la investigación del cerebro y por ende la mente y la conciencia, es aún más importante que la del Universo.

Se habla de “conciencia” pero nunca se la define. Y las definiciones de ella que aparecen en los diccionarios confunden más que aclaran. El término “conciencia” se usa como equivalente, más o menos, a “percatación”, o quizá a la sensación de “percatación”. El concepto, ciertamente, incluye la percepción del libre albedrío. En física cuántica, el término “conciencia” es el estándar en el tratamiento del problema de la medida. En todo caso, como le dijo Humpty Dumpty a Alicia (Carroll, 1872-1982): “cuando yo uso una palabra... quiere decir lo que yo quiero que diga”, afirmación con la que estaría bastante de acuerdo Wittgenstein.

El psicoanálisis introdujo en el dominio de atención científico el inconsciente en sus varias formas. En efecto, Freud propuso una “teoría” estructural de la mente en forma de “ego”, “id” o “it” y el “superego”. De ahí, que muchos científicos, que trabajaban en el cerebro, se sintieran implicados acerca de la más grande cuestión acerca del cerebro: la naturaleza de la conciencia, esto es, cómo se relacionan varios procesos psicológicos inconscientes con el pensamiento consciente.

Para conseguir avances en el entendimiento de la conciencia, se ha establecido una definición de trabajo sobre ella como un estado “awareness” perceptual o de atención selectiva muy rara y general. En su núcleo, la conciencia en la gente es una conciencia de sí mismo, una conciencia de ser “perspicaz”. En este sentido, la conciencia se refiere a la capacidad del ser humano de, no simplemente experimentar placer o dolor, sino de prestarles atención y ocuparse de ellos y reflexionar sobre esas experiencias, y hacerlo en el contexto de la vida cotidiana y sus antecedentes históricos. La atención consciente le permite a los humanos excluir experiencias extrañas y centrarse en los eventos críticos que confronta, sean placenteros o penosos.

Sin embargo, fue Crick, quien al abordar el más elusivo e ineludible de los fenómenos; o sea, la conciencia, la convirtió en un tema legítimo para la ciencia. En 1988, en un libro

autobiográfico (Crick, 1988) él ya abordó, aunque de pasada, el problema del cerebro. Allí y entonces se preguntaba ¿Qué había acerca de la teoría del cerebro? Y respondía que era fácil ver que una teoría de algún tipo es esencial ya que cualquier explicación del cerebro tiene que implicar grandes números de neuronas que interaccionan de formas complicadas. Además el sistema es claramente no lineal y de este modo es fácil adivinar exactamente como funcionaría cualquier tipo de modelo complejo. A continuación, hace una crítica acertada de los modelos teóricos basados en redes de neuronas afirmando que dichos modelos no son realmente teorías sino más bien “demostraciones” (las comillas son de Crick). Son pruebas de que unidades que se parecen en algo a las neuronas pueden hacer en realidad cosas muy sorprendentes, pero hay muy poco que sugiera concluyentemente que el cerebro se comporte en realidad de la manera que dicen los conexionistas.

Después considera que el problema con la neurociencia teórica es que está situada de alguna forma entre otros tres campos. En un extremo están los que trabajan directamente sobre el cerebro. Esto es ciencia. Están tratando de descubrir qué mecanismo utiliza realmente la naturaleza. En el otro extremo se encuentra la IA (Inteligencia Artificial), esto es ingeniería. Su objeto es “producir una máquina que trabaje de la forma deseada”. El tercer campo son las matemáticas, que no se preocupan ni por la ciencia ni por la ingeniería (excepto como una fuente de problemas), sino únicamente de relaciones entre entidades abstractas. Luego, pasa a hacer una crítica corrosiva de la psicología, a la que tilda de ciencia “blanda”, *la cual raramente produce resultados definitivos, sino que va saltando de una manía teórica a la siguiente. A nadie le gusta averiguar si un modelo es correcto ya que si lo hiciera, una gran parte del trabajo se detendría visiblemente.*

Por último, Crick afirma que son necesarias tres aproximaciones principales para resolver un sistema complicado como es el cerebro. Una, de análisis o estructural, establece el orden de componentes y cómo pueden deshacerse y caracterizar todas las partes aisladas: de qué están hechas y cómo funcionan. Otra, de síntesis o funcional, define el orden de procesos y puede encontrar exactamente dónde está localizada cada parte del sistema en relación con todas las demás partes. Es improbable que estas dos aproximaciones revelen, sólo por sí mismas, cómo funciona exactamente un sistema. Para hacer esto se necesita una tercera, comportamental, que establezca el orden de acciones; es decir, debe también estudiar el comportamiento del sistema y de sus componentes, cuando se interfiere de

forma muy delicada en sus diferentes partes para ver qué efecto producen las alteraciones en el comportamiento en todos los niveles. Si se pudiese hacer esto con el cerebro, se encontraría de forma inmediata cómo funciona. Y es esencial entender el cerebro humano con algún detalle si se quiere comprender correctamente el lugar del ser humano en el vasto y complicado Universo que lo rodea.

En la línea de Crick, Capra, definió tres conceptos clave en los sistemas vivos (Capra, 1996): (a) El “patrón organizativo” o funcionalidad. Es la configuración de relaciones entre los componentes del sistema. Que determina las características esenciales. (b) Estructura. Es la plasmación física del “patrón organizativo”. Esto implica un “principio” o aplicación abstracta de relaciones, mientras que una descripción de la estructura implica describir las componentes físicas reales, sus formas, composición, etc. En las organizaciones la estructura también apunta al propósito del sistema, porque existe en primer lugar. (c) Proceso o comportamiento. Es el enlace entre el patrón organizativo y la estructura.

A estos tres conceptos clave para entender todo sistema vivo cognoscitivo hay que añadirle dos propiedades más: a saber, (d) Autopoiesis (Maturana, 1982). Un sistema autopoietico es aquel que se está produciendo a si mismo continuamente, de modo que en un sistema de este tipo, ser y actuar son inseparables. Este proceso continuo de producir es de naturaleza cognoscitiva. En teoría todos los sistemas vivos y el cerebro en concreto, son cognoscitivos y la cognición siempre implica la existencia de una red autopoietica. (e) Disipación. Los sistemas vivos son estructuras disipativas abiertos a los flujos de materia y energía. De hecho, como lo señaló Schrödinger, los seres vivos se alimentan de negentropía (Brillouin, 1964); es decir, de entropía negativa.

Francisco Varela, estudió lo que él denominó “**yoes emergentes**” o “**identidades virtuales**”. La suya es una visión inmanente de la realidad basada en metáforas derivadas de la autoinformación y una epistemología de inspiración oriental más que de la ingeniería y la informática. Si la concepción tradicional de la IA es la de que el mundo existe independientemente del organismo, cuya tarea es elaborar un modelo adecuado de ese mundo; es decir, “consultar” antes de actuar, la de Varela es la de un mundo no representacionista, o quizás un “mundo como experiencia” que no tiene existencia independiente, sino que es en sí mismo un producto de las interacciones entre los

organismo y el entorno. Su teoría de la autopoiesis, o “autoproducción”, tiene que ver con el automantenimiento activo de los sistemas vivos cuyas identidades se mantienen constantes, mientras sus componentes cambian continuamente, Varela es un crítico radical de la ciencia cognoscitiva porque cree en la **emergencia**, no como el vitalismo de los años veinte del siglo pasado; es decir, la idea de una propiedad mágica que surge inexplicablemente de procesos mecánicos inferiores, sino de que todo es resultado de la dinámica de las partes componentes. Piensa que la ciencia cognoscitiva clásica, inspirada en la informática, es demasiado mecanicista y simple (Varela, 1996). Esta cuestión de la **emergencia** es el aspecto clave que aborda este trabajo.

A partir de los trabajos de Maturana y Varela, el acto básico de la cognición o “knowing”, no es un simple espejo de la realidad objetiva “fuera de él”. En realidad, la cognición es un proceso activo enraizado en la estructura biológica de los seres conscientes por el cual éstos crean realmente su mundo de experiencia; esto es, crean conocimiento. Por lo tanto, el proceso entonces puede identificarse con la cognición. De acuerdo con esto, la mente no es una cosa sino un proceso de interacción con el entorno a través de acciones cognoscitivas. En consecuencia, entender un sistema vivo y, más en concreto, la mente y su sustrato físico el cerebro, implica pues: (a) identificar su patrón. (b) Describir su estructura. (c) Especificar sus procesos cognoscitivos en forma de redes autopoieticas y (d) Establecer los intercambios de materia y energía. A esto el autor de este trabajo, propone, siguiendo a John Archibald Wheeler, Tom Stonier, Yingxu Wang, Juan Pazos y colaboradores, un innovador marco de tratamiento de la conciencia, el modelo EMI, que también engloba a la información.

Años después, en 1990, Crick y un colaborador suyo Christof Koch proclamaron en “Seminars in the Neurosciences” (Crick, 1990) que ya era hora de estudiar el, tal vez, más elusivo e ineludible de los fenómenos: la conciencia, convirtiéndose así en un tema legítimo de investigación empírica. Sostenían, asimismo, que ni se podía esperar alcanzar una verdadera comprensión de la conciencia ni de cualquier otro fenómeno mental tratando el cerebro como una caja negra; es decir, como un objeto cuya estructura interna es desconocida y hasta irrelevante. Sólo examinando las neuronas y las distintas interacciones entre ellas pueden acumularse los conocimientos inequívocos necesarios para crear modelos verdaderamente científicos de la conciencia, análogos a los que explicaban la herencia en términos del ADN. Crick y Koch rechazaban con rotundidad la

creencia, que albergaban muchos, de que la conciencia no se puede definir, y menos aún estudiar. Para ellos, la conciencia o consciencia, y todas sus formas, dirigidas hacia objetos, ya del mundo sensible, ya altamente abstractos o internos, parecen necesitar el mismo mecanismo subyacente que combine, de acuerdo con William James, la atención con la memoria a corto plazo. Para Crick la atención, crucial en su definición de conciencia, implicaba algo más que un simple proceso de información. Crick y Koch habían logrado lo aparentemente imposible, cual fue el que la conciencia dejará de ser un misterio filosófico y se convirtiera en un problema empírico. De este modo, una teoría de la conciencia representaría el apogeo o culminación de la neurociencia.

Cuatro años después, en 1994, Crick expuso detalladamente sus opiniones sobre la conciencia en un libro que él tituló: “The Astonishing Hipotesis”, a la que su editora añadió el subtítulo “The Scientific Search for the Soul” (Crick, 1994), lo que resulta incongruente con la postura reduccionista y materialista de Crick. El materialismo craso de Crick evoca el del filósofo, británico y oxfordiano, Gilbert Ryle (Ryle, 1949), quien, en los años treinta, para mofarse del dualismo, acuñó la ya citada frase *ghost in the machine*. Es decir, Ryle señalaba que no hay fantasmas en la máquina. Ryle afirmó que el dualismo, que sostenía que la mente era un fenómeno separado e independiente de su sustrato físico y capaz de ejercer influjo sobre él, viola el principio de conservación de la energía y, por tanto, de toda la física. Según Ryle, la mente es una propiedad de la materia y sólo trazando detalladamente los intrincados meandros de la materia en el cerebro se puede “explicar” la conciencia. Ryle, desde luego, no fue el primero en proponer un paradigma materialista. Hace más de cuatro siglos, Francis Bacon (Bacon, 1834) instó a los filósofos de su época a que dejaran de empeñarse en mostrar cómo evolucionaba el Universo a partir del pensamiento y empezaran a considerar cómo evolucionaba el pensamiento a partir del Universo.

Bacon, siguiendo a Platón, había definido lo que ha de tener necesariamente un filósofo como *una mente suficientemente móvil para reconocer (que es lo más necesario de todo) la analogía entre las cosas, y suficientemente fija y atenta como para observar las sutilezas de las diferencias; es decir, alguien capaz de percibir la unidad en la multitud y de discernir la diversidad en la aparente igualdad*. La consciencia se resolvió, para los materialistas, cuando alguien decidió que era un mero epifenómeno del mundo material.

Entender la consciencia es, con mucho, el desafío más importante y trascendental que afronta la ciencia. La verdad de esta aseveración puede plasmarse en la carrera de Crick, sin duda el biólogo más creativo e influyente (Kandel, 2006) de la segunda mitad del siglo XX. En efecto, cuando Crick, de formación físico, entró en el campo de la biología después de la Segunda Guerra Mundial, había dos grandes cuestiones que se creía estaban más allá de la capacidad de la ciencia para responderlas: ¿Qué distingue el mundo de la vida de la no vida? y ¿qué es la naturaleza biológica de la consciencia? Crick dirigió, en primer lugar, sus esfuerzos al problema más fácil, el de diferenciar la materia animada de la inanimada y exploró la naturaleza del gen. En 1953, tras dos años de colaboración entre él y Watson (Watson, 1968) éste contó que: “al almuerzo Francis irrumpió en el pub Eagle para decir a todo el mundo que quisiera oírle que habíamos encontrado el secreto de la vida”. En las dos décadas siguientes, Crick ayudó a descifrar el código genético: cómo el ADN hace el ARN y éste a las proteínas. Posteriormente, en 1976, a la edad de 60 años, Crick se dedicó al misterio científico que permanecía incólume: la naturaleza biológica de la consciencia. Asunto al que dedicó el resto de su vida con la colaboración de Christof Koch, pero sin realizar avances sustanciales sobre el mismo. Pero, eso sí, hizo que la comunidad científica se centrara en él. Realmente algunos científicos y muchos filósofos de la mente continúan encontrando la consciencia tan inexplicable que temen que nunca pueda explicarse en términos físicos y se preguntan: ¿cómo puede un sistema biológico sentir algo? incluso más ¿cómo puede pensar sobre sí mismo?.

Crick (Crick, 1994) decía: *Para el hombre no existe una investigación científica más importante que la de su cerebro. Nuestra visión del Universo está íntimamente unida a ella.* Este enfoque lo comparte la premio Nobel de medicina de 1986, por el descubrimiento e identificación del NGF “Nerve Grow Factor”, proteína estimuladora del crecimiento de las fibras nerviosas, Rita Levi Montalcini cuando dijo (Levi, 2000): *más aún que el Universo es fundamental que lleguemos a conocer nuestro cerebro, pues no sólo es la clave para entender el propio Universo, sino también el funcionamiento de la mente humana.*

Por su parte, el filósofo y lógico Bertrand Arthur William Russell (Russell, 1927) y el fundador de la neurofisiología y gran neurofisiólogo Sir Charles Scott Sherrington (Sherrington, 1951) proponían dos ejemplos muy parecidos (el segundo, usando la cúpula celeste y el primero, empleando un objeto cualquiera) de cuál es el verdadero problema

de la conciencia y su dificultad de solucionarlo. En su forma más abstracta, el problema puede plantearse del modo siguiente: Supóngase que un proceso físico P_f da comienzo en un objeto sensible, O_s . Dicho proceso va hasta el correspondiente órgano sensorial (vista, oído, olfato, gusto, tacto,...) donde se convierte en otro proceso físico, que provoca aún otro proceso físico, en este caso sensorial, P_s , en el(los) nervio(s) sensorial(es) correspondientes. Finalmente, dicho(s) nervio(s) sensorial(es) llevan las señales al cerebro donde se produce un efecto mental, E_f , al mismo tiempo que se “sensorializa” (ve, oye, etc.) el objeto donde se inició el proceso. Ahora bien, E_f ; esto es, el proceso mental, es algo de naturaleza totalmente distinta a la de los procesos físicos, eléctricos y químicos, que le preceden y acompañan. Es decir, la cadena entera de eventos es física, cada uno de los pasos es una reacción foto-eléctrica-química bien conocida y determinada. Pero al llegar a la correspondiente zona cerebral, se produce un cambio completamente distinto a los que le precedieron y completamente inexplicable e inesperado. Ante la mente, se presenta una escena visual, una melodía, etcétera. Éste es, de forma abstracta y resumida, el problema que, habitual y comúnmente, se discute y denomina como el problema de la conciencia. Sherrington consideraba que el problema de la mente estaba fuera del alcance del ser humano (Sherrington, 1906) y, por lo tanto, carecía de solución; es decir, estaba convencido de que el problema de la conciencia era científicamente inexplicable.

Aunque, como lo señaló la editora ejecutiva del *Journal of Consciousness Studies*, Keith Sutherland (Sutherland, 2003), en realidad se trata de tres problemas distintos que, según distintas concepciones, se superponen en diverso grado, a saber:

A. El de la generación o Qualia. El problema difícil o arduo.

La mayoría de las objeciones filosóficas a la explicación científica de la conciencia se refieren a los qualia, al igual que el escepticismo de muchos científicos. La implicación es, a menudo, que el misterio de la conciencia, el nudo del mundo, no podrá desatarse sin desentrañar primero el misterio de los qualia. Los qualia prototípicos que discuten los filósofos son sensaciones simples como la cualidad rojo del color rojo o la cualidad de dolor del dolor.

En un célebre experimento mental fisiológico, Mary, una neurocientífica del futuro, “daltónica”, lo sabe todo acerca del sistema visual y del cerebro y, en particular, de la fisiología de la discriminación del color. Sin embargo, cuando por fin logra recuperar la visión del color, todo aquel conocimiento se le revela insuficiente comparado con la auténtica experiencia del color. Es decir, ninguna teoría, científica o no, ninguna descripción, bastará nunca para que una persona daltónica consiga experimentar el color. John Locke (Locke, 1980) se percató del problema hace mucho tiempo, concretamente en 1690: *Un hombre ciego y estudioso que se había exprimido los sesos para comprender los objetos visibles, y que había utilizado todas las explicaciones de sus libros y de sus amigos para comprender el significado de la luz y de los nombres de los colores que tan a menudo encontraba, se jactaba un día de que por fin había logrado comprender el significado del color escarlata. A lo que un amigo le preguntó: ¿Qué es el color escarlata? Y él responde: “Es como el sonido de una trompeta”.*

Se han erigido todo tipo de argumentos filosóficos sobre la presunta irreductibilidad de la qualia. ¿Por qué se siente el rojo cómo se siente? ¿Podría existir un zombi filosófico? Es decir, una réplica del ser humano que diera exactamente las mismas réplicas conductuales y tuviera exactamente los mismos mecanismos neuronales, pero que no tuviera ningún tipo de experiencia subjetiva, que no tuviera qualia asociados a su experiencia. Y si así fuera ¿cómo se sabría? La razón por la que resulta tan difícil explicar los qualia en términos neuronales no exige grandes esfuerzos esotéricos. Puede ilustrarse con la paradoja del ser humano y el fotodiodo.

El término técnico para el primero de los problemas de la conciencia, el de los qualia, es el que el filósofo australiano David J. Chalmers denominó *problema arduo* y para los pensadores victorianos *el gran abismo* o *el abismo insondable*, o, en términos de Arthur Schopenhauer, *el nudo del mundo*. Es decir, el problema de cómo los procesos o configuraciones materiales producen experiencia consciente. “*Problema fácil versus arduo*”: En efecto, a principios de los 90 del siglo pasado David Chalmers, un joven filósofo australiano de formación físico y matemático, revolucionó el estudio de la conciencia al identificar lo que denominó “problema difícil”. En síntesis, el problema difícil es el de explicar, la generación, por el cerebro biológico, del mundo interno de la experiencia subjetiva. Los problemas fáciles de Chalmers incluyen cosas tales como: la reacción a estímulos, o la comunicabilidad de estados mentales, y el resto de estudios de

la conciencia. Chalmers clasificó estos problemas de fáciles, no porque lo sean en sí mismos, sino en comparación con el problema difícil. El interés actual en el problema difícil de la conciencia emana de su aparente similitud, ¿y conexión?, con el problema difícil de la Mecánica Cuántica, en adelante MC, el de la observación.

Para Chalmers, la conciencia plantea los problemas más desconcertantes en la ciencia de la mente. No hay nada que se conozca más íntimamente que la experiencia consciente, pero no hay nada más difícil de explicar. Aunque no ocupa un lugar central en su argumentación, Chalmers considera que la MC probablemente tiene relevancia para el problema de la conciencia. El último capítulo de su gran obra “La Mente Consciente” (Chalmers, 1996) se titula “*La interpretación de la MC*”.

Para Chalmers descubrir el proceso biológico de la conciencia en el cerebro; esto es, resolver el problema fácil es sólo cuestión de tiempo. Tarde o temprano, se averiguará. Otra cosa es el, así denominado, “problema arduo”; es decir, el de la generación de la propia conciencia: ¿Cómo los procesos o configuraciones materiales producen experiencia consciente? ¿Cómo puede surgir conciencia de la estimulación del sistema nervioso? ¿Cómo es posible que cosas objetivas como las neuronas cerebrales produzcan experiencias subjetivas como el sentimiento de, por ejemplo, yo camino sobre el césped? Es la versión actual del antiguo problema del cuerpo y la mente, sólo que, paradójicamente, parece hacerse más difícil, cuanto más se sabe sobre el cerebro. En efecto, las neurociencias explican cada vez mejor cómo el cerebro discrimina colores, resuelve problemas y planifica y organiza acciones, pero el problema arduo persiste. El mundo objetivo que rodea a los seres dotados de mente y las experiencias subjetivas de su interés, parecen ser de naturaleza distinta. Preguntarse el modo en que uno produce las otras parece un sinsentido.

Los problemas fáciles de Chalmers tienen que ver a menudo con la correlación entre la actividad cerebral y aspectos de la conciencia, los *correlatos cerebrales de la conciencia*. La tecnología de imágenes cerebrales permite la visualización detallada de la actividad del cerebro, cuando piensa y siente, y ha estimulado estudios fascinantes de los procesos mentales. La electroencefalografía (EEG), la tomografía por emisión de positrones (TAC) y la resonancia magnética (RM), son claves para localizar la actividad neuronal, incluso la RM funcional casi puede localizar la región del cerebro que está consumiendo más

oxígeno durante una función cerebral particular. La RM funcional permite, pues, correlacionar regiones cerebrales concretas con los procesos neuronales implicados en la memoria, el habla, la visión, y la conciencia, entre otras. Sus inconvenientes son, por una parte, el que, como cualquier técnica basada en la actividad metabólica, no es rápida y, por otra y más importante, que excluye la glía.

Actualmente, los datos disponibles que relacionan la actividad electroquímica cerebral con la conciencia son fragmentarios. Pero supóngase, lo que acabará consiguiéndose, que la mejora de las técnicas actuales o una nueva técnica futura pudiera identificar “completamente” las activaciones cerebrales particulares ligadas a ciertas experiencias conscientes. Eso permitiría correlacionar todas las sensaciones conscientes reportadas, con la actividad metabólica cerebral y quizás incluso con los fenómenos electroquímicos subyacentes. Semejante catálogo de correlatos cerebrales de la conciencia es la meta de buena parte de la investigación actual del cerebro consciente. Si se cumple este objetivo, dicen algunos, se habría conseguido todo lo que se “puede” conseguir. La conciencia, afirman, quedaría explicada, porque en ella no hay nada más allá de la actividad cerebral que se correlaciona con las “sensaciones” de conciencia experimentadas y reportadas. Si se desmonta un viejo reloj de péndulo y se ve el mecanismo de muelles y engranajes que convierte la oscilación del peso en el giro de las manecillas, se puede saber todo lo que hay que saber sobre el funcionamiento del reloj. Hay quienes afirman que la conciencia quedará igualmente explicada por el conocimiento de lo que hacen las neuronas que componen el cerebro, ¿y la glía qué?.

Chalmers, por su parte, principal portavoz de un punto de vista diametralmente opuesto al de Crick, ve imposible explicar la conciencia sólo a partir de sus correlatos neuronales. A lo sumo, sostiene Chalmers, “esas teorías nos dicen algo acerca del papel físico que puede desempeñar la conciencia, pero no nos dicen cómo surge. Para cualquier proceso físico que especifiquemos habrá una pregunta sin respuesta: ¿por qué este proceso debería dar lugar a la experiencia [consciente]?. Dado uno cualquiera de tales procesos, es conceptualmente coherente que pudiera... [existir] en ausencia de experiencia. De lo que se sigue que ninguna mera descripción de procesos físicos nos dirá por qué surge la experiencia. La **emergencia** de la experiencia va más allá de lo que puede derivarse de la teoría física. Si en la teoría atómica puede explicar de forma reduccionista la humedad del agua y porque se pega a los dedos, esto está muy lejos de explicar la *sensación* de humedad”. Chalmers

niega la posibilidad de cualquier explicación reduccionista de la conciencia y sugiere que toda teoría de la conciencia debería tomar la experiencia como entidad primaria, junto con la masa, la carga y el espacio-tiempo. Además, esta nueva propiedad fundamental conllevaría nuevas leyes fundamentales, que él llama “principios psicofísicos”.

Chalmers no se detiene aquí. Un principio que considera básico, y el más interesante para el doctorando, es una “hipótesis natural: que **la información (al menos cierta información) tiene dos aspectos básicos, uno físico y otro fenoménico**”. Esto recuerda la situación de la MC donde la función de onda también presenta dos aspectos: por un lado, es la realidad física total de un objeto, y, por otro, conjeturan algunos, esa realidad es pura información (¡signifique eso lo que signifique!).

Fue David J. Chalmers quien, en el congreso de Tucson de 1994, expuso el punto de vista misterioso en términos rigurosos, del modo siguiente: *El estudio de las neuronas no puede revelar por qué la incidencia de, por ejemplo, ondas acústicas en el oído humano da origen a la experiencia subjetiva de la quinta sinfonía de Beethoven*. Para Chalmers todas las teorías físicas describen sólo funciones, verbigracia: la memoria, la percepción, la atención, la intención, la introspección, la toma de decisiones, etc., relacionadas con procesos físicos específicos del cerebro. Pero ninguna de esas teorías puede explicar por qué el ejercicio de esas funciones cognitivas va acompañado de la experiencia subjetiva. Tanto las hipótesis de la conciencia cuántica como las hipótesis neuronales son rechazadas por Chalmers quien afirmó que las teorías físicas sólo podrían explicar las distintas “funciones” del cerebro. Es decir, en ningún caso, explicarán por qué esas funciones cognitivas eran desempeñadas por sensaciones conscientes, los “qualia”. Y aportaba lo que él creía la solución al problema duro, siguiente: *Así como la física suponía la existencia de propiedades fundamentales de la naturaleza, como el espacio, el tiempo, la energía, la masa, etc., así una teoría de la conciencia en regla debía postular la existencia de una nueva propiedad fundamental: la información. Ésta, además de tener una plasmación física, era más que eso, pues también presenta un aspecto fenoménico; esto es, subjetivo o experimental*. Según esto, cualquier “ente” que procesara información debía tener alguna experiencia consciente. Hasta un termostato es consciente. Después de todo, es posible imaginar perfectamente un mundo de zombis, en todo iguales a los humanos, salvo en que no tendrían experiencia consciente del mundo. Según Chalmers, por más que se sepa del cerebro, las neurociencias jamás podrán colmar

ese abismo explicativo *que existe entre los ámbitos físico y subjetivo con una teoría estrictamente física.*

Hasta aquí Chalmers no hacía sino expresar el mismo punto de vista misteriano básico de Nagel y McGinn. Pero más adelante pasó a decir que aunque la ciencia no pudiera resolver el problema mente-cuerpo, la filosofía podría hacerlo, y la solución que propuso fue que los científicos asumirían que **la información es una propiedad de la realidad tan esencial como la materia y la energía**. Su teoría era parecida al “it from bit” de Wheeler, de quien Chalmers se reconocía deudor y, naturalmente, presentaba el mismo talón de Aquiles. La información no tiene sentido mientras no haya un procesador de la misma que la recoja y actúe sobre ella. La materia y la energía estuvieron presentes en el alba de la creación, no así la vida. Entonces ¿cómo puede ser la información tan fundamental como la materia y la energía? (Chalmers, 1995).

En una concepción adecuada de la mente, no sólo ha de explicarse la capacidad de representación del mundo, sino también sus sensaciones, esas propiedades fenoménicas que reciben el nombre de *qualia*. El argumento del murciélago de Thomas Nagel (Nagel, 1979) es relevante al respecto. Lo que señala Nagel es que las propiedades fenoménicas implican la adopción de un punto de vista o perspectiva, con lo que no es posible explicarla desde fuera de ese punto de vista. Por eso, las explicaciones fisicalista y funcionalista, que se proponen como explicaciones objetivas, esto es, trascendiendo cualquier punto de vista específico, son incapaces de explicar “el carácter subjetivo de la experiencia”. Este argumento, parece envolver en un aura de misterio la naturaleza misma de los *qualia*. Ahora bien, partiendo de una ontología materialista, el misterio no puede ser ontológico. Frank Jackson (Jackson, 1986) ha caracterizado el misterio en términos epistémicos como sigue: Incluso conociendo todos los hechos funcionales y neurofisiológicos acerca de la visión policromática del mundo, alguien que ve por primera vez algo rojo adquiere algún conocimiento nuevo. Esto lo expresó muy gráficamente Mark Twain cuando dijo “El que agarra a un toro por el rabo sabe seis o siete cosas más que el que sólo lo vio agarrar”. Desde los ochenta se ha intentado dar una respuesta convincente al reto de los *qualia*; sin embargo, personas como Colin McGinn y David J. Chalmers, simplemente han aceptado el carácter misterioso de los “qualia” bien como cuestión epistémica: nunca podrían explicarse, o como cuestión ontológica: son propiedades no físicas del mundo. Esta última posición implica un neodualismo que se

corresponde bien con la concepción emergentista de lo mental. Sea como fuere, la cuestión acerca de la naturaleza de los *qualia* está lejos de estar resuelta e incluso, o tal vez a consecuencia de ello, bien planteada.

Platón en el Fedón (Platón, 1997), al describir las últimas horas de Sócrates, ridiculiza la noción de que su conducta pudiera explicarse en términos físicos. Ésa resulta ser la alusión más antigua de la que se tiene noticia de lo que en la actualidad se conoce como “**laguna explicativa**”, expresión acuñada por Joseph Levine (Levine, 1983) para expresar la desconcertante incapacidad de las teorías fisiológicas para explicar los fenómenos psicológicos y en particular la conciencia o “**qualia**”; es decir, las sensaciones subjetivas del mundo. También, naturalmente, otras funciones mentales como: la percepción, la memoria, el razonamiento, las emociones, así como la conducta humana.

En efecto, tal vez, el problema más desalentador de la conciencia sea el de los “qualia”. Se define por “quale”, en singular, y “qualia” en plural, la cualidad específica de la experiencia subjetiva. Pues bien, los qualia, tales como el color, el calor, el dolor, el ruido, etc., se han sustraído siempre a una explicación científica. Al respecto, la posición del Premio Nobel Edelman y Tononi es la siguiente (Edelman, 2002):

- a) Para experimentar qualia es necesario poseer cuerpo y cerebro que sustenten procesos neuronales. En ningún caso, una descripción o una teoría, por muy correcta que sea en la descripción de los mecanismos subyacentes, podría sustituir la experiencia individual de un quale.
- b) Cada experiencia consciente diferenciable representa un quale distinto, tanto si es primariamente una sensación, una imagen, un pensamiento o incluso un humor, y tanto si retrospectivamente parece simple o compuesto.
- c) Cada quale corresponde a un estado distinto que puede diferenciarse de miles de millones de otros estados dentro de un espacio neuronal de múltiples dimensiones. Las dimensiones relevantes vienen dadas por el número de grupos neuronales cuya actividad constituye un núcleo dinámico de gran complejidad. Los qualia son, por tanto, discriminaciones de una alta dimensionalidad.
- d) El desarrollo de los primeros “qualia” se produce fundamentalmente sobre la base de discriminaciones multimodales centradas en el cuerpo y realizadas por los sistemas

proprioceptivo, quinestésico y autónomo que poseen el embrión y el cerebro del recién nacido, especialmente en el tronco cerebral. Todos los qualia subsiguientes pueden ser referidos a este conjunto inicial de discriminaciones, que constituye la base del sentido más primitivo de la identidad.

Llegar a comprender de qué modo surgen los qualia a partir de cierto tipo de actividad neuronal es lo que tanto a Sherrington como a Russell les parecía insuperable. Sobre todo ¿qué procesos neuronales pueden explicar esas discriminaciones subjetivas que se denominan qualia? Por ejemplo, tal y como se muestra en la figura 1.2, se puede definir un espacio neuronal tridimensional para el color que puede explicar muchos aspectos de la discriminación de los colores en los humanos. En la figura, cada eje corresponde a un conjunto de grupos neuronales cuya descarga varía de 0 hercios (Hz), completamente inhibido, hasta 100 Hz, descarga máxima, pasando por 10 Hz, descarga neutra o espontánea. Las tasas de descarga de estos tres conjuntos de grupos neuronales definen un espacio tridimensional dentro del cual cada color que un ser humano puede diferenciar corresponde a un punto distinto. El rojo puro, verbigracia, correspondería a un punto en este espacio, que se representa con un círculo con aspa, cuyas coordenadas son 100 en el eje rojo-verde, y 10 en los ejes azul-amarillo y blanco-negro. Las neuronas que tienen estas propiedades se encuentran en las áreas de color de la corteza cerebral. Paul M. Churchland (Churchland, 1995) sugirió y argumentó que el cerebro podría utilizar representaciones vectoriales para especificar colores, olores y caras de forma económica.

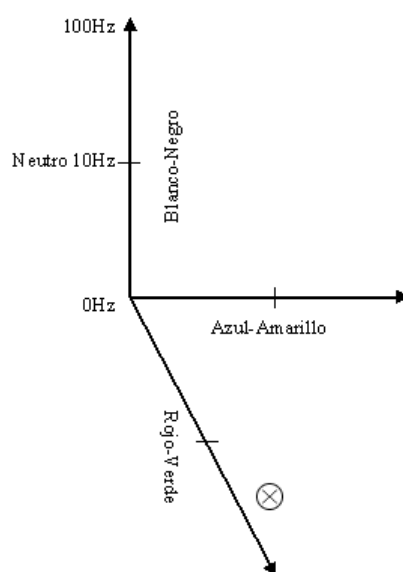


Figura 1.2. Espacio del Color

La forma más simple de acometer el problema de los qualia es suponiendo que, para cada quale, todo lo que se necesita es una o un grupo de neuronas, que, al dispararse, represente explícitamente ese aspecto particular de la conciencia o quale. Éste es el enfoque sugerido por Crick y Koch (Crick, 1995). Con esta hipótesis bastaría con estar seguros de que el grupo neuronal en cuestión se dispara cada vez que la persona dice percibir, verbigracia, un color, y que no se dispara cada vez que dice no percibir el color. Una manera específica de eludir la cuestión de los qualia es la vieja idea de las líneas etiquetadas; o sea, la ley de las energías nerviosas específicas de Müller y colegas (Müller, 1991): El sentimiento subjetivo que determina cómo se siente la descarga de esta o aquella neurona viene especificado de algún modo en los órganos sensoriales a los que está conectada la neurona. Si está conectada a un ojo, dará lugar a una sensación visual, si a la piel, una táctil, etcétera. Naturalmente, no es una explicación válida habida cuenta de la existencia de neuronas baroreceptoras que no originan quale alguno.

En este sentido, a finales del siglo XIX, el psicólogo Edward B. Titchener (Titchener, 1901) intentó identificar los átomos de la conciencia, queriendo abarcar la fenomenología completa de la conciencia en términos de sensaciones elementales, pero su intento resultó ser infructuoso. Además, la hipótesis atomista se enfrenta al problema, mucho más grave, de que no dice nada sobre el mismísimo núcleo del problema. ¿Cómo se produce la transformación que tan incomprensible le parecía a Sherrington y Russell? A Arthur Schopenhauer (Schopenhauer, 1996), con referencia al nudo del mundo, le hubiera gustado ver cómo se ha acabado en el punto de partida de la cuestión. Este garrafal “*petitio principii*”, diría, habría provocado un descomunal ataque de risa a los moradores del Olimpo. Por eso, Edelman y Tononi proponen un espacio de qualia N-dimensional, tal y como se muestra en la figura 1.3. Donde N es el número, muy grande, de grupos neuronales. En la figura 1.3, sólo se muestra un conjunto mínimo de dimensiones. Algunas de estas dimensiones corresponden a grupos neuronales selectivos al color y, exactamente igual que en la figura 1.2, exhiben constancia del color. Otras, corresponden a impresiones auditivas, propioceptores, etcétera. En esta hipótesis, cada punto discriminable en el espacio N-dimensional, identifica un estado consciente, mientras que una trayectoria que una varios puntos de este espacio corresponde a una secuencia temporal de los estados conscientes.

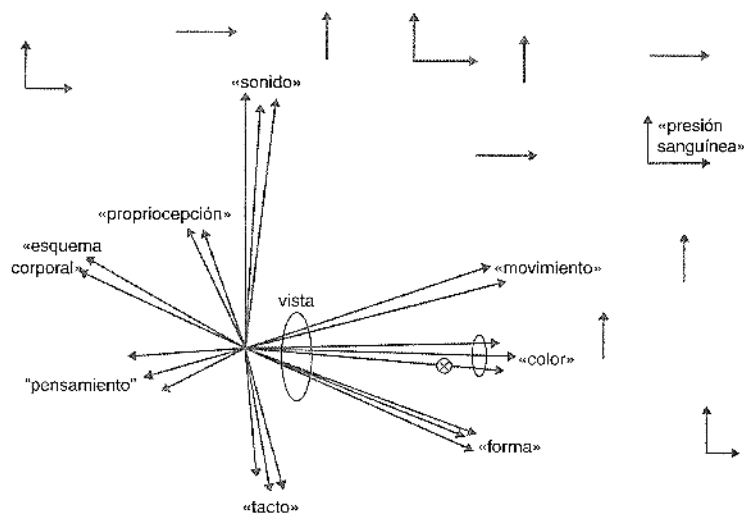


Figura 1.3. Espacio Qualia

En contraposición a la mayoría de los científicos y filósofos, de los que la excepción es John R. Searle (Searle, 1998), Edelman y Tononi sugieren que, por estar especificado de forma única y no ser reducible a componentes independientes, *cada estado consciente distinto merece ser llamado un quale*. Como William James previó correctamente, no existen sensaciones puras atomistas (James, 1890): *Nadie ha tenido nunca una sensación simple por sí sola. Nuestra conciencia, desde el día en que nacemos, es la de una abigarrada multiplicidad de objetos y relaciones, y lo que llamamos sensaciones simples son el resultado de una atención discriminadora, a menudo en un alto grado.*

Otro aspecto interesante es establecer la relación entre procesos neuronales inconscientes y otros conscientes; por ejemplo, en el caso del habla. Como lo señaló M. V. Egger (Egger, 1881): *Antes de hablar, uno apenas sabe lo que pretende decir, pues tras hacerlo uno se llena de admiración y sorpresa por haberlo dicho y pensado tan bien.* En la novela “Howards End” de E. M. Forster (Edelman, 2002) una de sus personajes lo expresa de forma memorable cuando dice: *¿Cómo puedo saber lo que pienso hasta oír lo que digo?*

Francis Crick, buscó “la neurona de la conciencia”. Para él, las experiencias subjetivas, la conciencia, no es más que la actividad de esas neuronas. En su libro “La búsqueda científica del alma” (Crick, 1994a) defiende esta hipótesis: “Tú, tus alegrías y tristezas, tus recuerdos y tus ambiciones, tú sentido de la identidad personal y del libre albedrío, no

son en realidad sino el comportamiento de un vasto entramado de células nerviosas y sus moléculas asociadas”.

Si es así, la sensación en los seres humanos de que la conciencia y el libre albedrío están más allá del mero funcionamiento de electrones y moléculas es una ilusión. En consecuencia, la conciencia debería tener una explicación en última instancia reduccionista; esto es, debería poder describirse completamente en términos de entidades más simples, los correlatos cerebrales de la conciencia. Así, supuestamente, **emergen** las sensaciones objetivas de la de las neuronas ¿Y de la glía qué? Esto es parecido a la idea, más fácil y asequible, de aceptar que la humedad del agua emerge de la interacción entre los átomos de hidrógeno y oxígeno que forman las moléculas de H₂O. **Justamente esta emergencia constituye la “hipótesis asombrosa” de Crick** (Crick, 1994b). Pero, ¿es realmente tan asombrosa? Desde luego a la mayoría de los físicos, químicos e informáticos, especializados en IA les parecería la conjetura más natural.

Christof Koch, durante largo tiempo colaborador de Crick, adopta una postura más matizada: *Dada la centralidad de las sensaciones subjetivas en la vida diaria, hacía falta una extraordinaria prueba fáctica antes de concluir que los qualia y las sensaciones son ilusorias. El enfoque provisional que adopto consiste en considerar las experiencias en primera persona como hechos primarios de la vida que requieren explicación. Y en un contexto algo diferente, añadía: “Aunque no puedo descartar que la explicación de la conciencia pueda requerir leyes fundamentalmente nuevas, ahora mismo no veo una necesidad apremiante de dar tal paso. E igual que Crick, añadió: [...] doy por sentado que la parte física de la conciencia es una propiedad emergente de interacciones específicas entre neuronas y sus elementos [...]. [...] los caracteres de los estados cerebrales y de los estados fenoménicos [de las neuronas] parecen demasiado diferentes para que los unos sean completamente reducibles a los otros. Sospecho que su relación es más compleja de lo que tradicionalmente se ha imaginado.*

El filósofo Daniel Dennett en su libro “La conciencia explicada” (Dennett, 1991) describe el tratamiento de la información por el cerebro como un proceso donde “múltiples borradores” son corregidos constantemente, “conglomerándose” de vez en cuando para producir la experiencia. Dennett niega la existencia de un “problema difícil” que ve como una especie de dualismo mente-cerebro algo que él afirma refutar con el argumento

siguiente: “Ninguna energía ni masa física se asocia con ellas [las señales de la mente al cerebro] ¿cómo entonces, tiene algún efecto sobre lo que pasa en las células cerebrales que debe afectar, si es que la mente debe tener alguna influencia sobre el cuerpo? [...] esta confrontación entre la física convencional y el dualismo [...] es contemplada por todo el mundo como el ineludible y fallo fatal del dualismo”.

Dado que Chalmers argumenta que la conciencia obedece a principios que “están fuera” de la física convencional, no está nada claro que un argumento “basado” en ella puede ser una refutación válida de Chalmers. Es más, en el argumento de Dennett hay un vacío cuántico: no necesariamente se requiere masa o energía para determinar en cuál de sus estados posibles colapsará una función de onda como consecuencia de una observación.

B. El “Yo”.

Otra cuestión compleja es la del “Yo” o voz interior que construye el diálogo interno compuesto de episodios y acontecimientos sobre la propia vida de uno y lleva a interrogantes como los siguientes: ¿Por qué se tienen experiencias subjetivas? ¿Por qué uno se siente uno y no otro? Es éste, el Yo, el segundo problema de la conciencia que naturalmente está relacionado con el anterior.

Este problema puede remontarse a Aristóteles (Aristóteles, 1998). Para el filósofo griego el “Yo” está estrechamente ligado a la existencia corpórea del individuo de modo que mente y cuerpo conforman una unidad y la mente participa en funciones vitales que van desde la locomoción a la contemplación filosófica. El punto de vista opuesto, defendido por platónicos y pitagóricos, separa la mente del cuerpo y conduce directamente a la teoría cartesiana de un Yo descorporeizado. Locke, en su “Ensayo Sobre el Entendimiento Humano” de 1690 (Locke, 1980), abordó el problema de la identidad personal proponiendo la memoria como mecanismo de la continuidad del “Yo” en el tiempo. Cuando Hume (Hume, 1939) dirigió su atención a ese “Yo”, no encontró más que un hatajo de impresiones sin evidencia introspectiva de que realmente existiera, y despachó, consecuentemente, al “Yo”, como una ficción de la imaginación. Sus seguidores, los filósofos Anthony Kenny y Robert Nozick opinan parecido. El primero sostiene que *el Yo no es más que un sinsentido de los filósofos derivado de un malentendido del pronombre reflexivo*. El segundo, por su parte, conviene en que el “Yo”

se sintetiza en el acto de la autoreferencia reflexiva; en otros términos, el “Yo” no es más que un producto de la reflexión. Por su parte, la concepción del “Yo” de Daniel Dennett resulta ser una ficción teórica que, sin embargo, desempeña un útil papel explicativo. Finalmente, la premio Nobel Rita Levi Montalcini (Levi, 2002) dice que, la conciencia relaciona el Yo de la persona con las experiencias de los sucesos, dado que permite comprender la existencia de la persona como entidad pensante, que reconoce la identidad personal a través de la sucesión temporal de las vivencias del pasado, el presente y un presagiado futuro.

Para Joseph Le Doux (Le Doux, 1999) explicar la conciencia no es tan importante como saber de qué manera se sirve el cerebro a la vez de los genes y de la experiencia para crear un Yo, una identidad personal en cada individuo. Para él, la cuestión principal es saber cómo el cerebro de las personas les hace ser quienes son. Y añade que explicar la conciencia no bastaría para explicar eso.

Una vez comienza a **emerger** la conciencia de orden superior, es posible construir un Yo a partir de relaciones sociales y afectivas. Este Yo, que supone el desarrollo de un agente autoconsciente, un sujeto, va mucho más allá de la individualidad de base biológica que posee un animal con conciencia primaria. La **emergencia** del Yo conduce a un refinamiento de la experiencia fenomenológica, a enlazar los sentimientos con el pensamiento, la cultura y las creencias. Libera la imaginación y abre el pensamiento a los vastos dominios de la metáfora. Con una visión combinada de la conciencia primaria y la de orden superior es posible, cuando menos, clasificar tres grandes misterios; a saber, el de la conciencia continua, el del Yo y el de la construcción de historias, planes y ficciones.

Los puntos de vista sobre la relación entre la subjetividad y el desarrollo de la identidad difieren enormemente, y pueden denominarse las dos posiciones extremas sobre lo subjetivo “internalismo” y el “externalismo”. De acuerdo con el internalismo, primero se produce una experiencia subjetiva y, a continuación, la creciente diferenciación del Yo, de modo que la autoconciencia sobreviene como resultado de interacciones tanto sociales como lingüísticas. Aunque no hay ninguna manera directa de conocer la naturaleza de cada estado subjetivo, se considera que es la base necesaria para que posteriormente surja un yo verdadero. Conforme a esta visión, es ontogénicamente posible alguna forma de

pensamiento antes de la adquisición del lenguaje. De acuerdo con el punto de vista alternativo, no tiene sentido hablar de respuestas subjetivas o estados interiores antes de que se haya adquirido el lenguaje, lo que se consigue por medio de relaciones interpersonales que son sociales. Sólo cuando se ha desarrollado suficiente lenguaje, **emergen** las bases conceptuales del Yo y sólo entonces puede considerarse que el individuo es consciente y, sobre todo, autoconsciente. De acuerdo con esto, la subjetividad previa es indeterminada, y no tiene sentido preguntarse, como hizo Thomas Nagel (Nagel, 1979), qué se siente al ser X, tanto si X es un murciélago o cualquier otra cosa sin lenguaje. Con la **emergencia** de la conciencia de orden superior a través del lenguaje, se produce un acoplamiento consciente y explícito de sentimientos y valores que da como resultado las emociones como componentes cognoscitivos que experimenta una persona, un sujeto, un Yo, cuyos pensamientos son necesariamente conscientes.

Las formas sofisticadas de intercambio de información desarrollados por los humanos no pueden considerarse sin la presencia de la consciencia. Con la aparición del “Homo Sapiens” y de la conciencia de orden superior fue posible crear sistemas de símbolos sintácticamente ricos, crear códigos e incluso la(s) lógica(s). Con el tiempo se crearon métodos de análisis científico que llevaron a la formulación de leyes naturales. Para muchos, entre ellos el doctorando, esas leyes son informaciones, en forma de conocimientos. Para la naturaleza, externa a los seres humanos ¿qué es lo que se intercambia, energía o información codificada? ¿viene codificada o se codifica? (Wheeler, 1994) ¿qué fue primero, la biología o la lógica?.

La conciencia surge de ciertos arreglos del orden material del cerebro. El pensamiento consciente es un conjunto de relaciones con un significado que va más allá que la simple energía o materia, aunque implique a ambas. Y la mente que dio origen a ese pensamiento es material y tiene significado. La mente se sostiene sobre una base material en la forma de un conjunto de relaciones: La acción del cerebro de una persona y todos sus mecanismos, de abajo arriba, de los átomos a las conductas, da como resultado una mente que puede ocuparse de procesos de significado. Mientras genera estas relaciones innaturales que serán reconocidas por ésta y por otras mentes, esta mente se sustenta y depende completamente de los procesos físicos que ocurren en su propio funcionamiento, dos dominios completamente separados de mente y materia; no hay razón para el dualismo. Son las extraordinariamente complejas estructuras materiales del sistema

nervioso y hormonal del cuerpo humano lo que da origen a los procesos mentales y al significado. No hace falta suponer nada más, ni otros mundos, ni espíritus, ni extrañas fuerzas todavía por desvelar, como la gravedad cuántica de Penrose.

C. Agente. El libre albedrío

Finalmente está el “problema del agente” que puede plantearse del modo siguiente: ¿Quién es el autor de los actos volitivos conscientes? Existe, por otra parte, otro tipo de conciencia trascendental que va más allá de los límites espacio-temporales, como la que aparece durante el sueño, la meditación o los estados alterados por efecto de sustancias químicas, etc. Este problema está estrechamente relacionado con el “Yo”. Es decir, si el “Yo” es una ficción narrativa ¿quién es el autor de los actos volitivos conscientes? Y, lo que es más importante, ¿cómo se puede reconciliar el sentido de los humanos como agentes libres con la clara evidencia de que el agente debe estar sujeto a las leyes, deterministas, de la física?

Crick, como le dijo a Horgan (Horgan, 1998), no aceptaba la vana esperanza de que los humanos pudieran tener libre albedrío, y señalaba al respecto: *De lo que somos conscientes es de que tomamos una decisión. A nosotros nos parece un acto libre, pero en realidad es el resultado de cosas de las que no somos conscientes.* Esto ya lo había sugerido Sigmund Freud en 1949 cuando, reflejando su inquietud sobre el inconsciente que ya se había desarrollado en el siglo XIX, planteó la hipótesis de que la mayoría de los actos humanos estaban determinados por impulsos inconscientes. Hoy la neurociencia sabe que eso es cierto.

A continuación se van a describir dos experimentos, llevados a cabo con sujetos humanos y que tiene importantes implicaciones acerca de retardo de la conciencia; es decir, que tienen que ver con el tiempo que necesita la conciencia para actuar y ser activada. El primero, concierne al papel activo de la conciencia, y el segundo a su papel pasivo. Esas implicaciones, consideradas juntas, resultan aún más sorprendentes.

El primero, fue realizado por L. Decke, B Grotzinger y H. H. Kornhuber (Decke, 1976) y consistió en lo siguiente: en cierto número de sujetos se registraron las señales eléctricas en un punto de sus cabezas (EEG), y se les pidió que flexionaran varias veces, y

respectivamente, el dedo índice de sus manos derechas, “totalmente a su libre elección”. La idea era que los registros de los EEG indicarían algo de la actividad mental que estaba teniendo lugar dentro del cráneo y que está involucrada en la decisión real consciente de flexionar el índice.

Para obtener una señal relevante y significativa a partir de las trazas del EEG es necesario promediar las trazas de varias series diferentes y la señal resultante no es muy definida. Sin embargo, se encontró algo curioso: hay un aumento gradual del potencial eléctrico registrado un “segundo entero”; o quizás incluso hasta un segundo y medio, “antes” de que el dedo sea flexionado. Esto parece indicar que el proceso de decisión consciente; necesita un segundo o más para actuar. Esto puede ser contrastado con el tiempo, mucho más corto, que lleva responder a una señal externa si el modo de respuesta ha sido establecido por adelantado. Por ejemplo, la flexión del dedo podría ser la respuesta al destello de una señal luminosa en lugar de ser “libremente voluntaria”. En tal caso, es normal un tiempo de reacción de aproximadamente un quinto de segundo y que es unas cinco veces más rápido que la acción “voluntaria” que, tal y como se muestra en la figura 1.4, era lo que se ponía a prueba en los datos de Kornhuber y colegas. Como se ve en la figura, la decisión de flexionar el dedo parece hacerse en el tiempo cero, pese a que la señal precursora, promediada sobre muchos ensayos, sugiere un “preconocimiento” de la intención de flexionar el dedo.

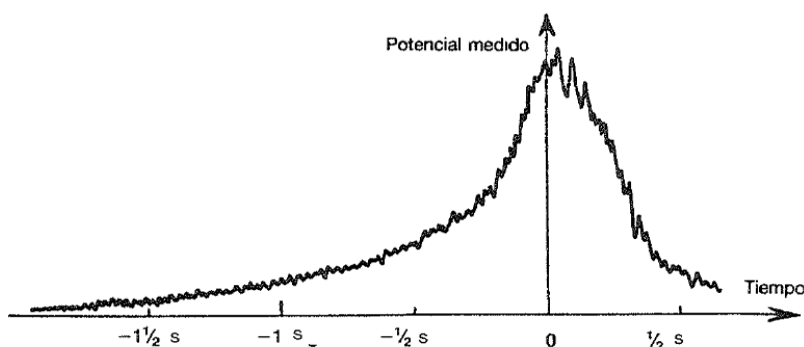


Figura 1.4. Tiempo de reacción de flexionar el dedo.

Curiosamente, estos resultados fueron anticipados por el padre de la mecánica cuántica Niels Bohr al interpretar los desenlaces habituales de las películas del Oeste. Según lo contó su discípulo George Gamow (Gamow, 1971): «Al atardecer,..., aparecía Bohr

diciendo estar muy cansado y que le gustaría *hacer algo*. Hacer algo significaba indefectiblemente ir al cine, y las únicas películas que le gustaban eran las tituladas: *Lucha a tiros en el rancho Lazy Gee* o *El jinete solitario y una muchacha India* [...]. La afición de Bohr a las películas del Oeste se tradujo en una teoría desconocida para todos excepto para sus compañeros de cine de aquel tiempo. Todo el mundo sabe que en todas las películas del Oeste, al menos en el estilo de Hollywood, el pillo siempre se dispone a sacar el revólver con ventaja, pero el héroe es más rápido y siempre acaba matando al bribón. Niels Bohr atribuyó este fenómeno a la diferencia entre las acciones deliberadas y acciones condicionadas. El bribón ha de decidir cuándo ha de echar mano a la pistola, lo que retrasa su acción, mientras que el héroe actúa más rápidamente porque lo hace, instintivamente, por acto reflejo, sin pensar, cuando ve al bribón coger la pistola. Todos discrepábamos de la teoría y a la mañana siguiente el autor se fue a una tienda de juguetes para comprar un par de pistolas de *cowboy*. Nosotros disparábamos sobre Bohr qué hacía de héroe pero él nos mató a todos». Sólo añadir que Bohr era de los de más edad lo que supone que sus reflejos eran algo más lentos.

El segundo experimento, es el de Benjamín Libet, de la universidad de California, en colaboración con Bertram Feinstein, del Instituto Neurológico Monte Sinaí de San Francisco y colegas (Libet, 1979). En la primavera de 1994, la Universidad de Arizona en Tucson, organizó un simposio interdisciplinario sobre la consciencia. Allí y entonces, Benjamín Libet (Norretranders, 1998), psicólogo de la Universidad de California en San Francisco, realizó, tal y como se muestra en la figura 1.5, el siguiente experimento: pidió a unos voluntarios que movieran una mano cuando ellos quisieran y que al mismo tiempo determinaran con ayuda de un reloj especial, tal y como el que se muestra en la figura 1.5(a), el momento en el que habían tomado la decisión. Un electrodo colocado en la sien registraba simultáneamente el “potencial preparatorio”; es decir, el momento en que el cerebro ordenaba a la mano que se preparase para la acción, figura 1.5(b). Pidió a los participantes que movieran la mano, doblaran un dedo en el momento de haberlo decidido mientras anotaban el instante de su decisión tal y como indicaba el reloj. Los participantes tardaron 0,2 segundos de media en doblar su dedo tras haber decidido hacerlo; pero la electroencefalografía, que registró sus ondas cerebrales, reveló que la máxima actividad cerebral tenía lugar 0,55 segundos “antes” de haber tomado conscientemente la decisión de doblar el dedo. Es decir, el cerebro se ponía manos a la obra antes de que se fuese consciente de la decisión. El resultado fue sorprendente. Los instrumentos de medición

indicaron que el potencial de disponibilidad aparecía 550 milisegundos antes de que se produjese el movimiento, una acción que los propios voluntarios aseguraban haber decidido realizar hacía sólo 200 milisegundos. Es decir, el cerebro había decidido, por su cuenta, efectuar el movimiento 350 milisegundos antes que el “yo consciente”. El psicólogo Wolfgang Prinz ha encontrado una forma muy sencilla de formular la conclusión de este experimento: *No hacemos lo que queremos, sino que queremos lo que hacemos.*

Los participantes en el experimento determinan, con un pulsador, el momento exacto en el que deciden ejecutar el movimiento. Lo hacen con ayuda del reloj especial dividido en 60 falsos segundos. El objetivo es que ninguno de ellos esté condicionado por el tiempo real. El gráfico de la figura 1.5(c), muestra el proceso de la toma de la decisión en una secuencia temporal. Los electrodos registran en primer lugar que el cerebro se prepara para ejecutar la acción (potencial preparatorio). La decisión de realizar la acción sólo se hace consciente 200 milisegundos después. A continuación, en el momento marcado como 0 se produce el movimiento.

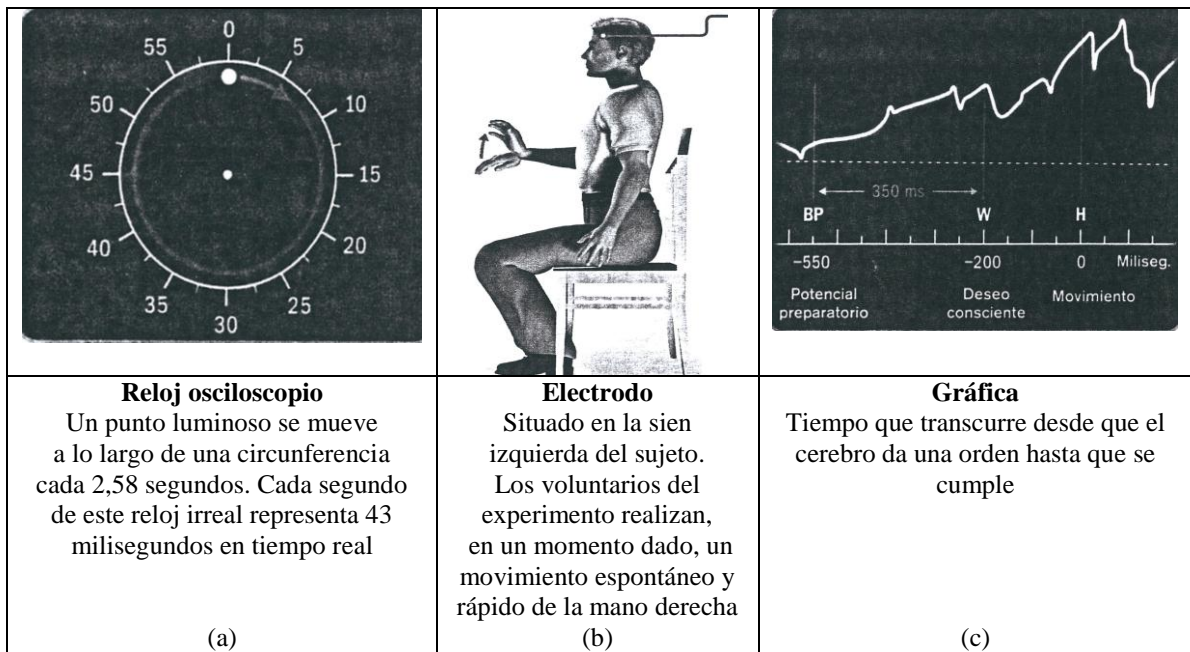


Figura 1.5. Elementos del experimento de Benjamín Libet

En el experimento se ponían a prueba sujetos que tenían que sufrir una operación quirúrgica en el cerebro por alguna razón que no tuviera relación con el experimento y qué consistió en tener electrodos colocados en puntos del córtex somatosensorial del

cerebro. El resultado del experimento fue el siguiente: Cuando se aplicaba un estímulo en la piel de esos pacientes transcurría aproximadamente medio segundo antes de que fueran conscientes de dicho estímulo, pese al hecho de que el propio cerebro había recibido la señal del estímulo en sólo una centésima de segundo y podría lograrse una respuesta preprogramada “refleja” del cerebro a dichos estímulos en, tal y como se muestra en la figura 1.6, aproximadamente una décima de segundo. En dicha figura, en el apartado (a) el estímulo en la piel “parece” percibirse aproximadamente en el momento real del estímulo. Como se muestra en (b), un estímulo cortical de menos de medio segundo no es percibido. En tanto que, como se ve en (c), un estímulo cortical de más de medio segundo, se percibe a partir de medio segundo en adelante. Ahora bien, como prueba (d), tal estímulo cortical puede “enmascarar retroactivamente” un estímulo anterior en la piel, lo que indica que la conciencia del estímulo en la piel “no ha tenido lugar aún” en el momento del estímulo cortical. Finalmente, (e) muestra que si un estímulo en la piel se aplica inmediatamente “después” del estímulo cortical, la percepción de la piel es “remitida hacia atrás”, pero la percepción cortical no lo es. Además, a pesar del retardo de medio segundo antes de que el estímulo alcanzase la consciencia, existía la impresión subjetiva de los propios pacientes de que no había habido ningún retraso en su toma de conciencia del estímulo. Algunos de los experimentos de Libet y colegas implicaban estimulación del tálamo con resultados similares a los del córtex somatosensorial.

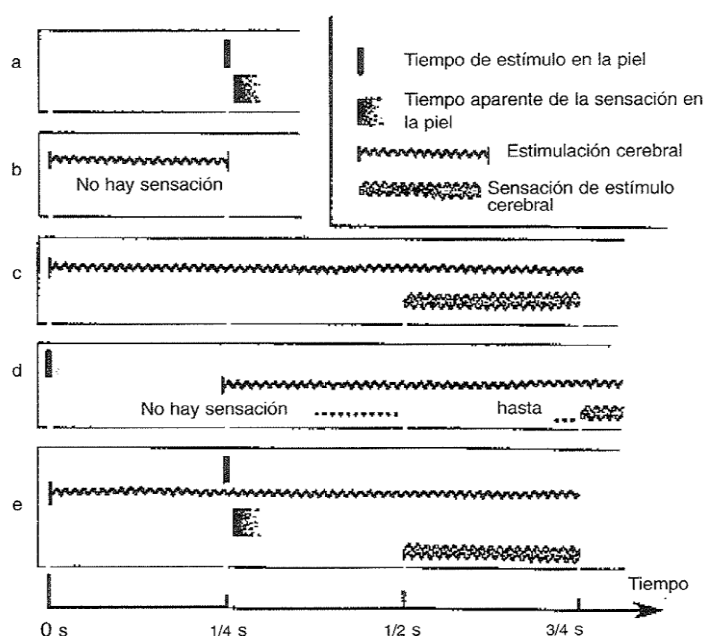


Figura 1.6. Experimento de Libet. Resultados

Recuérdese que el córtex somatosensorial es la región del cerebro por donde entran las señales nerviosas. Por consiguiente, la estimulación eléctrica de un punto del córtex somatosensorial, correspondiente a un punto particular de la piel, sería para el sujeto igual que si algo hubiera tocado realmente la piel en el punto correspondiente. Sin embargo, resulta que si la estimulación eléctrica es demasiado breve, durante menos de aproximadamente medio segundo, entonces el sujeto no llega a ser consciente de ninguna sensación. Esto debe contrastarse con una estimulación directa del punto de la misma piel, puesto que un contacto momentáneo en la piel si puede ser sentido. Supóngase ahora que se toca primero la piel y luego se estimula eléctricamente el punto del córtex somatosensorial. ¿Qué siente el paciente? Si la estimulación eléctrica se inicia aproximadamente un cuarto de segundo después de tocar la piel, entonces ¡el contacto en la piel no se siente en absoluto! Éste es un efecto conocido como “enmascaramiento retroactivo”. De algún modo la estimulación del córtex sirve para impedir que la sensación normal de contacto en la piel sea sentida conscientemente.

La percepción consciente puede ser inhibida, “enmascarada”, por un suceso posterior siempre que dicho suceso ocurra antes de medio segundo. Esto indica que la percepción consciente de tal sensación ocurre algo así como ¡medio segundo después del suceso real que produce dicha sensación!. Sin embargo, no parece que uno sea “consciente” de un retardo temporal tan largo en las percepciones. Una manera de hacer comprensible este curioso descubrimiento podría consistir, para Roger Penrose (Penrose, 1989), en imaginar que el “tiempo” de todas las percepciones humanas está retrasado efectivamente alrededor de medio segundo respecto al “tiempo real”, como si el reloj interno de cada uno simplemente estuviese “equivocado” en, aproximadamente, medio segundo.

Quizá algo de esa naturaleza queda corroborado en una segunda parte del experimento de Libet, en donde “primero” se iniciaba una estimulación eléctrica del córtex, continuando está estimulación durante un intervalo mucho mayor de medio segundo y tocando también la piel mientras seguía esta estimulación, aunque menos de medio segundo después de su inicio. Tanto la estimulación cortical como el contacto con la piel se percibían por separado, y resultaba claro para el sujeto cual era cada uno. Sin embargo, cuando se le preguntaba que estímulo ocurrió “primero”, el sujeto respondía que fue el contacto con la piel, ¡pese al hecho de que la estimulación cortical se inició mucho antes! Por lo tanto, el sujeto parece remitir la percepción del contacto en la piel “hacia atrás en el

tiempo” en, tal y como se muestra en la figura 1.6, aproximadamente medio segundo. Sin embargo, parece que esto no es simplemente un error “global” en el tiempo internamente percibido, si no una readaptación de la percepción temporal de los sucesos, pues la estimulación cortical, suponiendo que se perciba realmente no más de medio segundo después de su inicio, “no” parece estar remitida hacia atrás de esta forma.

A partir del primero de los experimentos anteriores, parece deducirse que la acción consciente necesita algo así como un segundo o un segundo y medio antes de poder llevarse a cabo, mientras que, según el segundo experimento, la consciencia de un suceso externo no parece ocurrir hasta medio segundo después de que haya tenido lugar el suceso. Imagínese lo que sucede cuando se responde a algún acontecer externo imprevisto. Supóngase que la respuesta es algo que requiere un momento de contemplación consciente. Sobre la base de los experimentos de Libet y colegas, parecería que debiera transcurrir medio segundo antes de que la consciencia sea llamada a escena; y luego, como parece indicar los datos Kornhuber y colegas, se necesita bastante más de un segundo antes de que pueda tener efecto una respuesta “voluntaria”. El proceso total, desde la entrada sensorial hasta la salida motora, parecería necesitar un tiempo ¡del orden de dos segundos! La explicación aparente de estos dos experimentos considerados juntos es que la consciencia no puede siquiera entrar en juego “en absoluto” en respuesta a un suceso externo, ¡si dicha respuesta tiene que tener lugar en menos de un par de segundos aproximadamente!

Si estos experimentos se toman al pie de la letra, se está conducido a llegar a la conclusión de que los seres humanos actúan completamente como “autómatas” cuando se lleva a cabo cualquier acción que necesite menos de un segundo o dos para modificar una respuesta. Sin duda, la consciencia es de acción lenta comparada con otros mecanismos del sistema nervioso como lo señaló Penrose, quizá la consciencia es, después de todo, un espectador que no experimenta más que una acción repetida “del drama entero”. Y después de cuestionar el hecho de que la consciencia no juega papel alguno en ciertas decisiones como jugar al tenis o al hablar, propone que no se apliquen las leyes físicas habituales para el tiempo cuando se considera la consciencia. Pero eso es, de momento, demasiado especulativo.

De los experimentos de Kornhuber y colegas parece concluirse lo siguiente:

- (a) El acto consciente de “libre albedrío” es una pura ilusión y, en cierto sentido, estaba ya programado de antemano en la actividad inconsciente precedente del cerebro.
- (b) Existe una posible función “de último instante” para la voluntad de modo que a veces, pero no normalmente, puede invertir la decisión que había sido formada y consciente que durante un segundo o así antes.
- (c) En realidad el sujeto decide conscientemente la flexión del dedo en un tiempo anterior en un segundo más o menos a que la flexión tenga lugar, pero percibe erróneamente, de forma sistemática, que la acción consciente ocurre en un instante muy posterior, justo antes de que el dedo sea flexionado realmente.

En experimentos más recientes Benjamín Libet y colaboradores (Libet, 1990) (Libet, 1992) han repetido estos experimentos, pero con refinamientos añadidos, como se muestra en la figura 1.6, diseñados para cronometrar más directamente el acto real de desear la flexión del dedo, pidiendo al sujeto que anote la posición de un reloj manual en el momento en que se tomó la decisión. Las conclusiones parecen confirmar los resultados anteriores pero también hablan en contra de (c), y el propio Libet parece favorecer (b).

Quizá uno de los experimentos anteriores, por si mismo, no sería paradójico aunque sí algo perturbador. Quizás las aparentes decisiones conscientes de uno eran realmente hechas “inconscientemente” en algún tiempo anterior, al menos un segundo antes. Quizás las sensaciones conscientes necesiten algo así como medio segundo de actividad cerebral antes de que puedan ser producidas realmente. Pero si se juntan ambos experimentos, entonces parece que uno se ve llevado a la conclusión de que cualquier acción en la que un estímulo externo conduce a una respuesta conscientemente controlada, parecería necesitar un retardo temporal de entre un segundo y un segundo y medio antes de que pueda ocurrir la respuesta. De hecho, la consciencia ni siquiera tendría lugar hasta que ha pasado medio segundo, y si dicha consciencia va a ponerse en acción, entonces la maquinaria aparentemente lenta del libre albedrío tendría que entrar en escena, quizás con otro segundo de retraso.

¿Realmente son tan lentas las respuestas conscientes? En la conversación habitual, por ejemplo, no parece que sea así. Aceptar (b) llevaría a concluir que la mayoría de los actos de respuesta son completamente inconscientes, aunque de vez en cuando uno podría ser

capaz de reemplazar una respuesta semejante de aproximadamente un segundo, por una consciente. Pero si la respuesta es normalmente inconsciente, entonces a menos que sea tan lenta como una consciente no hay oportunidad para que la consciencia la reemplace. De otro modo, cuando el actor consciente entra en juego, la respuesta inconsciente ya se ha producido y es ahora demasiado tarde para que la consciencia le afecte. Así pues, a menos que los actos conscientes puedan “a veces” ser rápidos, la “propia” respuesta inconsciente necesitaría alrededor de un segundo. Una reacción inconsciente “preprogramada” sólo tarda un quinto de segundo. Por supuesto, se podría tener aún una respuesta inconsciente rápida, de un quinto de segundo, junto con (a) que podría ser compatible con la ignorancia total, por parte del sistema de respuesta inconsciente, de cualquier actividad consciente, sensorial, que pudieran venir más tarde. En este caso, y la situación (c) es mucho peor, el único papel para la conciencia en una conversación razonablemente rápida sería el de espectador, siendo conocedora sólo de una “acción repetida” del drama completo. No hay en ello ninguna contradicción real. Es plausible y desde luego posible, que la selección natural haya producido la consciencia sólo para su función de pensamiento deliberado, mientras que en cualquier actividad razonablemente rápida la consciencia sea sólo un observador. Quizá la facultad de la consciencia “ha” evolucionado sólo con el propósito de una actividad mental tan lenta y contemplativa, mientras que las respuestas rápidas son enteramente inconscientes, pese a estar acompañadas por una percepción consciente retardada de las mismas que no desempeñan ningún papel activo.

Al preguntársele a Libet si sus hallazgos tenían relación con el problema del libre albedrío su respuesta fue: *Yo siempre he conseguido esquivar esa cuestión.* Y, cautelosamente, sugirió que se podría practicar el libre albedrío no iniciando intenciones, sino vetando, accediendo o también reaccionando a ellas una vez que han brotado del inconsciente de la persona.

Estos tres problemas se superponen en cierto grado. Epicuro creía que átomos del “alma”, distribuidos por todo el cuerpo, eran los responsables de la consciencia y la vitalidad. William James y Henri Bergson, sostenían que la conciencia proviene del exterior y se transmite a través del cerebro, que sólo permite filtrar cierto tipo de conciencia. Sin embargo, para el filósofo italiano Andrea Vaccaro, las funciones atribuidas a la conciencia: humores, sentimientos, autoconciencia, intuiciones, etc., son, desde siempre,

propias y exclusivas del alma humana. Por su parte, para Chalmers, la consciencia de las sensaciones es una propiedad de todos los sistemas funcionales, incluso, verbigracia, los termostatos, mientras que la identidad personal y la libertad de actuación quedan restringidas a sistemas sofisticados dotados de lenguaje. En cambio, Searle sostiene que la consciencia es una propiedad natural de los sistemas biológicos complejos.

1.4. Conciencia y física cuántica

Por supuesto, el interés de los físicos cuánticos en el problema difícil viene de que la física se ha encontrado con la consciencia en el enigma cuántico, que se conoce como “el problema de la medida”, donde hay aspectos de la observación física que se acercan a los de la experiencia consciente. En ambos casos, la solución parece requerir algo más allá del tratamiento normal de la física y la psicología. Una manera de explorar la naturaleza de la consciencia, y su existencia, preguntarse e indagar quién o qué puede poseerla y aquí entra en juego la IA fuerte. La consciencia es un misterio que actualmente se explora porque el encuentro de la física con ella pone a los científicos ante el enigma cuántico. La consciencia y el enigma cuántico no son sólo dos misterios: Son los dos misterios. El primero, la demostración física del enigma cuántico, pone a la ciencia ante un misterio fundamental del mundo objetivo “ahí fuera”. El segundo, la percepción consciente pone a la ciencia ante el misterio fundamental de lo subjetivo, el mundo mental “interior”. La MC parece conectar ambos mundos.

Como lo señaló Chalmers, cuando hay dos misterios es tentador suponer que tienen la misma fuente. Esta tentación se magnifica por el hecho de que los problemas de la MC parecen estar profundamente ligados a la noción de observación que implica de manera crucial la relación entre la experiencia de un sujeto y el resto del mundo.

¿Puede afirmarse que el experimento cuántico es una manifestación de la consciencia proyectándose y haciendo algo físico? Los físicos lo niegan... pero no todos. Eugene Wigner, uno de los fundadores de la MC y premio Nobel de Física señaló (Wigner, 1977): *el respaldo de la existencia de una influencia de la consciencia en el mundo físico se basa en la observación de que no sabemos de ningún fenómeno en el que un sujeto recibe la influencia de otro sin ejercer una consiguiente influencia. Esto le parece convincente a quien esto escribe. Es verdad, que en las condiciones usuales de la física o*

la biología experimentales, la influencia de cualquier conciencia es ciertamente pequeña. “No se necesita suponer que tal efecto existe”. Sin embargo, es bueno recordar que lo mismo puede decirse de la relación de la luz con los objetos mecánicos [...]. Es improbable que el [pequeño] efecto se hubiera detectado si consideraciones teóricas no hubieran sugerido su existencia. Cuando el dominio de la teoría física, señaló Eugene Wigner, se amplió para abarcar los fenómenos microscópicos a través de la creación de la MC, el concepto de conciencia volvió a saltar a la palestra: no era posible formular las leyes de la MC de manera plenamente consistente sin ninguna referencia a la conciencia”.

A Eugene Paul Wigner le interesaba sobre todo el proceso de medición cuántica que le lleva a la cuestión de la conciencia y su papel en la reflexión del físico. De acuerdo con von Neumann, London y Bauer, autores de las formulaciones más compactas y explícitas de la estructura conceptual de la MC, cada medición es una interacción entre el objeto y el observador. Wigner apostilla que, en efecto, el estado del objeto puede describirse mediante un vector de estado, una función de onda; ahora bien el objeto no puede ser el Universo entero porque, al menos, el observador no está incluido. El objeto obedece las leyes del movimiento, ecuación de Schrödinger, en tanto en cuanto sea un sistema cerrado; es decir, separado del mundo. Se puede suponer que tal es el caso durante los intervalos entre las mediciones. No lo es evidentemente cuando se procede a una medición, porque ésta es una interacción entre el objeto y el observador. Para Wigner la cadena de transmisión de la información del objeto a la conciencia del observador podría constar de varias etapas. Éstas podrían reconstruirse con mayor o menor detenimiento. Pero no se puede llevar la transmisión de información hasta sus últimas consecuencias; esto es, hasta la conciencia del observador, porque la física actual no abarca la conciencia. Existen pues dos tipos de realidad: la conciencia de cada uno y todo lo demás. Dicha dualidad se presenta cuando el resultado de una observación entra en la conciencia del observador; más precisamente, es la propia conciencia de uno que observa puesto que es el que observa el único observador, las demás personas son objeto de las observaciones del observador.

En su trabajo de 1932, “The Mathematical Foundations of Quantum Mechanics”, John von Neumann (Neumann, 1932) dejó claro que la teoría cuántica hace inevitable el encuentro de la física con la conciencia. Considérese, como hizo von Neumann, un aparato de medida tal como un contador Geiger. Está aislado del resto del mundo pero entra en

contacto con un sistema cuántico; verbigracia, un átomo simultáneamente presente en un par de cajas. El contador Geiger se dispara si el átomo está en la caja de arriba y no lo hace si está en la de abajo. Von Neumann mostró que si el contador Geiger es un sistema físico gobernado por la MC, estaría en un estado de superposición con el átomo tal que estaría simultáneamente disparado y no disparado, de un modo similar a lo que sucede con el gato de Schrödinger.

Si un segundo aparato de medida aislado entra en contacto con el contador Geiger; por ejemplo, un dispositivo electrónico que registre si el contador Geiger se ha disparado, se unirá al estado de superposición y registraría ambas situaciones simultáneamente. Esta cadena de von Neumann puede alargarse indefinidamente. Von Neumann demostró que “ningún” sistema físico que obedezca la MC podría incluir el colapso de una función de onda en estado de superposición para dar un resultado particular. Ahora bien cuando se mira el contador Geiger siempre se ve un resultado particular, no una superposición. Von Neumann concluyó que sólo un observador consciente, haciendo algo no contemplado por la física actual, puede inducir el colapso de una función de onda. Aunque a todos los efectos prácticos, uno puede considerar colapsada la función de onda en cualquier eslabón macroscópico de la cadena de von Neumann, sólo un observador consciente puede llevar a cabo una auténtica observación.

Un par de años después Schrödinger ideó su gato, que se basaba esencialmente en la conclusión de von Neumann de que requería un observador consciente para colapsar un estado de superposición. Así pues, la física, la ciencia más empírica, parece basarse en última instancia en la consciencia.

Esto lleva a que en la pirámide reduccionista, figura 1.7(a), se modifique de modo que el fundamento empírico primario se sustituya por la consciencia figura 1.7(b) dado el colapso de la función de onda por la observación.

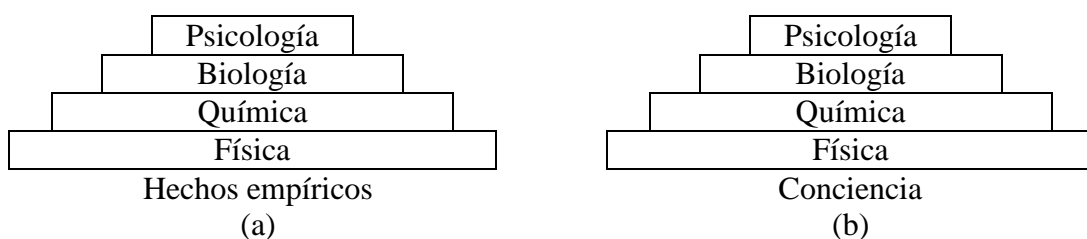


Figura 1.7. Reduccionismo científico clásico (a) y cuántico (b).

Retomando el átomo en un par de cajas, se va a plantear una extraña pregunta sobre la conciencia. Se puede demostrar que un fotón que atraviesa una de las cajas no observa si el átomo está o no en esa caja. En vez de eso se suma a un estado de superposición con el átomo. Si el fotón colisiona con un contador Geiger “aislado”; ese contador simplemente se suma al estado de superposición átomo-fotón, como en el análisis de von Neumann. El contador Geiger no observado está simultáneamente disparado y no disparado y el átomo sigue estando a la vez en las dos cajas. Considérese ahora una situación diferente. Supóngase que el contador Geiger golpeado por un fotón está sobre una mesa que se apoya en el suelo. Este contador no aislado interacciona con la mesa, al rebotar sus átomos contra los de la mesa. Por lo tanto, está entrelazado con la mesa y, por ende, con el resto del mundo, lo que incluye las personas. El átomo, entrelazado con el fotón, entrelazado con el contador, está ahora entrelazado con las conciencias individuales. Pero, si nadie mira, nadie “sabe” en que caja está el átomo. He aquí la extraña cuestión ¿es este entrelazamiento directo del átomo con el resto del mundo y las conciencias individuales lo que lo colapsa en una sola caja, o el colapso requiere la constatación consciente de la presencia del átomo en una caja mediante una observación efectiva del átomo o el contador Geiger? ¿Cómo se podría discernir si el átomo ha colapsado en una caja o todavía está presente simultáneamente en ambas? En rigor, sin “mirar” el átomo o el contador Geiger, no hay manera de saberlo. Siendo rigurosos, a menos que se invoque algo más allá de la MC, el átomo quizá siga estando en ambas cajas.

Una cosa más entre conciencia y entrelazamiento. Todo el mundo se entrelaza instantáneamente con el fotón tan pronto como éste incide en el contador Geiger “no” aislado. El entrelazamiento viaja a velocidad infinita. Pero para que una persona en la distancia “se percate” de la condición del contador o del átomo, tiene que comunicarse por algún medio físico que no podría superar la velocidad de la luz. Así pues, la “percatación” no puede ser más rápida que la luz. ¡Y se necesita un observador consciente!.

A menudo se puntualiza que la “única” manera que hay de conocer la existencia de la conciencia es a través de la propia sensación en primera persona o del testimonio en segunda persona de otros. Aunque esto último, presenta un desafío que la MC, pretende superar como limitación. Lo que, aquí, ahora, preocupa de la “conciencia”, es que tiene un papel central en el enigma cuántico, es que parece afectar a la realidad física. Esto no

implica necesariamente que la conciencia sea una “cosa física” que hace algo físico. Se está, pues describiendo un enigma, no proponiendo una solución al mismo. El ejemplo simple en MC es que la observación de la totalidad de un objeto dentro de la caja “causaba” su presencia allí. Y se dice “causaba” porque, presumiblemente, uno “podría” haber optado por observar una interferencia y establecer una situación contradictoria con la anterior en la que el objeto había sido una onda presente a la vez en ambas cajas. Tal vez, siguiendo a Bohr, se debe emplear el término “influyó en” vez de “causaba”.

Una demostración semejante ¿requiere necesariamente un observador consciente? Después de todo ¿no podría efectuar la observación un robot no consciente o incluso un contador Geiger? Como se verá, no se dice que la conciencia sea un requerimiento obligado. Recuérdese, de momento, que si ese robot contador no estuviera en contacto con el resto del mundo y obedeciera la MC, simplemente se entrelazaría convirtiéndose en parte de un estado de superposición total, como el gato de Schrödinger. En este sentido no “observaría”. La demostración del encuentro con la conciencia mediante el montaje de cajas pareadas descansa en el supuesto de que se “podría” haber optado por otro experimento distinto del que se hizo; es decir, de que se tiene libre albedrío. De hecho, la existencia el enigma cuántico depende crucialmente del libre albedrío.

1.5. La única prueba “objetiva” de la conciencia

Se entiende por “prueba objetiva” una prueba en tercera persona esencialmente apreciable por cualquiera. En este sentido, la prueba objetiva es el requerimiento para establecer la fiabilidad de una teoría científica. Cada uno de los seres humanos sabe que es consciente; eso es prueba en primera persona. Los otros dicen que son conscientes; esto es prueba en segunda persona. Sin prueba en tercera persona, la prueba objetiva de que la conciencia misma puede implicar directamente algo físicamente observable, su misma existencia es cuestionable. Como ya se vio, hay quien afirma que la “conciencia” no es más que un nombre para el comportamiento electroquímico de un vasto entramado de neuronas y sus moléculas asociadas en el cerebro. Por “implicación directa” sólo quiere decirse algún papel para la conciencia más allá del aspecto electroquímico confinado en el cuerpo humano. Y recuérdese que sólo se está hablando de “prueba” y no de demostración.

Roseblum y Kuttner (Roseblum, 2006) dieron una prueba objetiva mediante el experimento del par de cajas. Ciertamente, dicha prueba ofrecida por el experimento cuántico es circunstancial. Es decir, se valieron de un hecho, la interferencia, para establecer un segundo hecho, que el objeto estaba distribuido entre ambas cajas. La prueba circunstancial puede ser convincente más allá de toda duda razonable, incluso judicialmente. Pero la lógica de la prueba circunstancial puede ser tortuosa por eso presentaron una alegoría válida que proporcionó una prueba directa de la “conciencia”.

El experimento cuántico arquetípico, el de las dos rendijas, o el de las cajas pareadas de Roseblum y Kuttner exponen que la elección consciente del tipo de experimento, inspección de cajas o interferencia, creaba “cualquiera” de dos situaciones físicas contradictorias en las cajas pareadas. Así pues, el experimento cuántico es una prueba objetiva de la conciencia, más allá de sus correlatos cerebrales. Por supuesto, prueba no equivale a demostración. Pero aunque el experimento cuántico, al requerir la interferencia, no pasa de ser una prueba “circunstancial”, es la “única” prueba objetiva de la conciencia. Es decir, se ha aportado una huella en la “escena del crimen”, sin señalar un culpable.

Penrose y Hameroff, basándose en la gravedad cuántica e incorporando ideas de la lógica matemática y biología neuronal, presentaron una teoría basada en los tres pilares siguientes: la no computabilidad, la gravedad cuántica y el papel de la tubulina, proteínas celulares presentes en las neuronas que podían manifestar “reducción objetiva” de la función de onda sin observador. De esta teoría se ha dicho que tiene el poder explicativo de unos polvos mágicos en la sinapsis. No obstante, es falsable.

Otra teoría es la de Henry Stapp. Como se ha visto, la física clásica determinista sólo permite el libre albedrío a costa de excluir la mente del ámbito de la física. Stapp señala que la extensión de la física clásica al binomio cerebro-mente impondrá unos pensamientos controlados “de abajo arriba” por el movimiento determinista de partículas y campos. No habría ningún mecanismo para una influencia consciente “de arriba abajo”. Stapp, basándose en la formulación “vonneumaniana”, postula dos realidades una física y una mental. La primera, incluye al cerebro. Quizá en un estado de superposición cuántico particular. La segunda, incluye a la conciencia y, en particular, las intenciones. Lo mental puede actuar de manera intencionada sobre el cerebro físico para escoger un estado de

superposición que luego colapsa en una situación concreta, En dicha teoría, la conciencia no se “proyecta” al mundo exterior directamente, pero esta selección mental determina en parte el carácter del mundo físico externo al cuerpo; por ejemplo, si un objeto está en una caja de su par o simultáneamente en ambas. El aspecto aleatorio final de la opción, que el objeto está en la caja A y no en la B, verbigracia, es obra de la Naturaleza.

La cuestión va de suyo ¿como un cerebro grande y tibio puede permanecer en un estado cuántico particular el tiempo suficiente para que las intenciones de una persona influyan en él? Stapp respondió a esto con el “efecto Zenón cuántico”, así llamado por una sentencia estilo zenoniano: “un cazo no hierve si no deja de mirarse”. Cuando un sistema cuántico cae de un estado a otro inferior la caída comienza muy lentamente. Si se observa muy poco después de que la caída haya comenzado, es casi seguro que se encontrará en el estado original. Luego la caída vuelve a comenzar desde el principio. Si el sistema es observado de manera casi continuada, casi nunca se degrada. Stapp aplica esto a las intenciones mentales que “observan” el cerebro y lo mantienen en un estado cuántico dado durante un tiempo suficiente. Stapp citó diversos hallazgos psicológicos en apoyo de su teoría.

En 1866, el biólogo Thomas Henry Huxley, el defensor más efectivo de la evolución en el siglo XIX, al leer por primera vez “El Origen de las Especies” de Darwin y dominar su idea central, exclamó y escribió: *¡Que increíblemente estúpido por mi parte no haber pensado en ello!* (Huxley, 1957). Él expresaba el problema de la conciencia como sigue: *Que algo tan extraordinario como un estado de conciencia sea el resultado de estimular el tejido nervioso es tan incomprensible como la aparición del genio cuando Aladino frota la lámpara.* Desde entonces, a juicio del filósofo Jerry Fodor, no se han hecho demasiados progresos: *Nadie tiene la más mínima idea de cómo algo material puede ser consciente.* Y añade: *Nadie sabe siquiera cómo sería tener la más mínima idea de cómo algo material puede ser consciente. Así anda la filosofía de la consciencia.* Los neurofilósofos Paul y Patricia Churchland (Churchland, 1990) afirman que el problema de la conciencia no se verá de forma muy distinta al problema de la vida. Para ellos los “misterios” de la conciencia son el equivalente actual de los vitalistas.

El neurofisiólogo español José Manuel Rodríguez Delgado, que obtuvo extraordinarios éxitos en el control de toros bravos poniéndoles electrodos en su cerebro, decía al

respecto (Rodríguez, 1969): *La mente puede definirse como la “elaboración intracerebral de la información extracerebral. Y en otra ocasión Rodríguez Delgado decía (Rodríguez, 1979): La mente surge cuando, por medios materiales, los símbolos dan forma a las neuronas”. Y, finalmente, añadió: La mente aparece y se desarrolla cuando elementos extracerebrales (símbolos, relaciones temporales entre informaciones que llegan y códigos de conducta) apartados por medios materiales penetran en el cerebro dando forma a la estructura material de las neuronas.*

La cuestión es si, o no, los eventos mentales y de comportamiento y los fisiológicos pueden describirse con los mismos modelos. Si sí, eso significaría que existe un morfismo entre ellos; es decir, un modelo abstracto, psicofísicamente neutral, puede aplicarse en los dos sentidos (Bertalanffy, 1974). Esto lleva a una teoría unitaria (White, 1960) en la que “cuerpo” y “mente”, en sus aspectos formales o de estructura, queden comprendidos por un sistema conceptual “neutro”. White, no lo consiguió, pero aquí se propone un sistema conceptual abstracto que sí da cuenta de el cerebro y la mente.

Buscar una explicación molecular a la conciencia, decía Gunther Stent (Stent, 1969), es una pérdida de tiempo, pues se verá que los procesos fisiológicos responsables de esta experiencia completamente privada degeneran en reacciones completamente corrientes y de lo más prosaicas, ni más ni menos fascinantes que las que se producen, por ejemplo, en el hígado.

2. ESTADO DE LA CUESTIÓN

2.1. Evolución de la cuestión

El estudio de la conciencia retornó a la agenda académica en la década de los 90, cuando, por una parte, la Sociedad para la Neurociencia, que tiene más de 30.000 afiliados, convenció al Congreso de los Estados Unidos para que declarara los años noventa la “Década del Cerebro”, el premio Nobel de 1981, por sus trabajos en neurociencia, de origen sueco, Trosten Wiesel se opuso. Su argumento era que la idea era absurda pues se necesitaría al menos un siglo, tal vez incluso un milenio para comprender el cerebro. Y por otra, el presidente George H. W. Bush bautizó los 90 como la Década del Cerebro, después de un parón de casi un siglo.

A juicio del filósofo John Searle, dicho parón era legado del conductismo, que sostenía que la psicología científica debía limitarse a los fenómenos observables, y no a los estados mentales subjetivos. Esto provocó que estuviera mal visto plantear el problema de la conciencia en las discusiones sobre ciencia cognoscitiva. Todo cambió, cuando dos premios Nobel en Medicina, el físico Francis Harry Compton Crick, fallecido en el 2004, codescubridor de la estructura en hélice del ADN y el inmunólogo Gerald Edelman, quien, en 1972, recibió el premio por el trabajo que había realizado sobre las inmunoglobulinas, comenzaron a trabajar en este campo en la década de los 80. Sin embargo, esta opinión no es compartida por Semir Zeki del University College London, uno de los pioneros de la ciencia cognoscitiva, quien atribuye la vuelta al interés por la conciencia a los avances en las técnicas iconográficas de generación de imágenes cerebrales (TAC, PET, RMF) que permiten a los psicólogos estudiar los correlatos neuronales de la conciencia con toda la gloria del technicolor. Sea como fuere, el caso es que es un tópico de máxima actualidad.

La teoría científica de la conciencia fue a menudo calificada como la última frontera del conocimiento. John Horgan (Horgan, 1998) lo expresó como sigue: La mente, que no el espacio, es la última frontera de la ciencia.

En un artículo de 1947, que apareció dentro de una recopilación denominada “*Las tareas del intelecto*”, una disertación sobre la naturaleza de la actividad intelectual en el campo

de las matemáticas titulada “El matemático”, John Von Neuman (Neuman, 1947, 1961-1963) trata sobre la “*pan-matemática*”; es decir, como todo puede “*reducirse*” a las matemáticas; que liga todo. Esta “*profesión de fe*”, culminaba con la exposición del proyecto científico más ambicioso: el estudio de la mente humana, el misterio inextricable que subyace en la exposición formal de cualquier teoría de la naturaleza, de la sociedad o del hombre. Si este proyecto se podría afrontar, afirmaba, la clave debían de ser las matemáticas y continuaba: *“Otra cosa sobre la que no podemos decir hoy tanto como nos gustaría, pero sobre la que sabemos bastante, es que sería muy razonable esperarse un círculo vicioso si se intenta analizar el substrato que produce la ciencia, la función de la inteligencia humana. Todos los indicios que la exploración de esta área nos ofrece apuntan a que el sistema que está presente en la actividad intelectual, en otras palabras el sistema nervioso humano, puede ser investigado con métodos físicos y matemáticos. Pero, a pesar de ello, podemos imaginar que el hecho de que, en cualquier instante dado, un individuo deba estar informado completamente sobre el estado de su aparato nervioso en dicho instante conllevaría algún tipo de contradicción. La esperanza es que las limitaciones absolutas que existen aquí puedan ser a su vez expresadas en términos matemáticos y sólo en términos matemáticos”*.

Sin duda, una de las propiedades más asombrosas y fascinantes del cerebro humano es ser consciente de su propia conciencia. Entendiendo por tal el estado de conocimiento de la existencia humana como entidad individual, en la que cada uno es responsable de sus actos. El etólogo W. H. Thorpe, en varias ocasiones, expresó un pensamiento parecido (Thorpe, 1967) (Thorpe, 1969), al tratar la cuestión de la mente como **emergente**, planteando su origen desde varios aspectos. En primer término, está el aspecto epistemológico, según el cual, la mente es algo más que un mecanismo de cálculo o representación en cuanto que implica autoconciencia. Y, citando a Karl Popper, afirma que las calculadoras son incapaces de responder a todas las preguntas porque les es imposible poseer información inicial completamente al corriente acerca de sí mismas. *En otras palabras, ningún sistema de información puede incorporar en sí una representación detallada y al corriente de sí mismo incluida esa representación*. De este modo, si la autoconciencia es, en cualquier sentido, un hecho, entonces la mente tiene una cualidad fundamental que la distingue de las máquinas calculadoras. No obstante, para Thorpe el argumento más poderoso en cuanto a la autoconciencia es el siguiente: *El hecho de que la conciencia se haya desarrollado es por sí mismo prueba de que no es un producto*

marginal accidental. Sugiere que al menos ese aspecto de la vida mental puede cumplir algo para lo cual un mecanismo neuronal solo, por complejo y elaborado que sea, es inadecuado. Hasta qué punto sea útil ese concepto de autoconciencia, nadie puede decirlo.

Thorpe define la conciencia como el resultado de tres propiedades: Percepción interna, conciencia de sí mismo y de la propia vida. Concepto de unidad que resulta de la fusión de los pensamientos y los sentimientos en un todo único. Según lo expresó Mackay (Mackay, 1961): *En el caso de los seres humanos no se dice que piensa el “cerebro”, sino que lo hacen las personas.* Posteriormente, Mackay (Mackay, 1966), desarrolló más este argumento al señalar que el punto clave es saber si lo que una persona cree afecta correspondientemente al estado de su sistema organizador y si sí, ningún estado de cuentas completo “al corriente” podría ser creído por el aludido sistema sin que “ipso facto” dejase de estar “al corriente”. Por lo mismo, aun dados los datos más completos, no podría deducirse ninguna predicción completa del estado futuro de su sistema organizador, en la que pudiesen estar correctamente de acuerdo el agente y el observador. Una predicción hecha “secretamente” por un observador completamente desinteresado podría ser válida para él y sus compañeros de observación, pero sobre el propio agente no puede tener fuerza lógica obligante.

La comprensión del problema más misterioso y fascinante de todos es entender la relación entre las bases morfológicas-funcionales y las actividades mentales. El neurobiólogo David Hubel parafraseando, aunque sin citarlo, al Barón Munchausen (Raspe, 1985) ha comparado esta pretensión con la de elevarse en el aire tirándose hacia arriba de los tirantes. Este problema puede plantearse como sigue: ¿cuál es la naturaleza de los procesos mentales típicamente expresados por el pensamiento humano, y cómo se relacionan con los fenómenos físicos?

Arthur Schopenhauer (Schopenhauer, 1981) denominó “**el nudo del mundo**”, al enigma por antonomasia; es decir: establecer de qué modo se relaciona la experiencia subjetiva con ciertos eventos descriptibles objetivamente; en otros términos, al problema de la conciencia. La cuestión es si, o no, es posible, al igual que hizo Alejandro Magno, deshacer el nudo gordiano que le permitió conquistar el mundo, y de qué modo. En efecto, según contó Plutarco (Plutarco, 7821, 1822, 1830), Alejandro después de poner

bajo su obediencia la Frigia vio aquel celebrado carro atado con corteza de serbal, y oyó la relación allí creída por aquellos bárbaros, según la cual el hado ofrecía al que desatase aquel nudo, el ser rey de toda la tierra. Los más refieren que este nudo tenía ciegos los cabos enredados unos con otros con muchas vueltas, y que desesperado Alejandro de desatarlo, lo cortó con la espada por medio, apareciendo muchos cabos después de cortado. Aristóbulo, sin embargo, dice que le fue muy fácil el desatarlo, porque quitó del timón la clavija que une con éste el yugo, y después fácilmente quitó el yugo mismo.

Opiniones acerca de este arduo, y a la vez fascinante problema, hay muchas y tan variadas como las alas de las mariposas. Véanse alguna de ellas:

Thomas Reid (1710-1756) señalaba en sus “*Essays on the intellectual powers of man*”: *There is no prejudice more natural to man, than to conceive of the mind as having, some similitude to body in its operations. Hence, men have been prone to imagine, that as bodies are put in motion by some impulse or impression made upon them by contiguous bodies; so the mind is made to think and to perceive by some impression made upon it, or some impulse given to it by continuous bodies.*

Sir William Hamilton (1788-1856) en sus “*Lectures on Methaphysics*” dijo: *On Earth there is nothing great but man; in man there is nothing great but mind.*

Fernández, Macedonio (1874-1952) poéticamente señaló que lo único que existe es la mente:

*[...] en nuestra mente todo lo que eres, está
pues nunca estuviste sino en nuestra mente
y nuestra mente es lo único que jamás existió....*

Richard Rorty decía en “*Philosophy and the Mirror of Nature*”: *The mind-body problem, we can now say, was merely a result of Locke’s infortunate mistake about how words get meaning, combined whit his and Plato’s muddled attempt to talk about adjectives as if the were nouns.*

Charles M. Schulz, en *It's show time, Snoopy*, por “bocadillo” de Charli Brown, afirmaba: *My Mind and my body hate each other.*

Tengo un alma: no me la niegues. Descuartízame y no la hallarás; desarma una máquina de vapor y verás como tampoco encontrarás este fluido. Pero como me llamo George: es lo que la hace funcionar (reflexión de Tarleton en Misallience).

El gran matemático David Hilbert, en su “Lecture delivered before the International Congress of Mathematicians at Paris in 1900” (Hilbert, 1900) dijo: *We hear within us the perpetual call: There is the problem. Seek its solution. You can find it by pure reason, for in mathematics there is no ignorabimus.* Pues bien, si a la pura razón se le añade la experimentación, y se cambia matemáticas por cualquier dominio o rama del saber, en este caso, la mente, el doctorando suscribe lo dicho por Hilbert, incluso conociendo las limitaciones de la ciencia, incluida la propuesta de Hilbert y al no aceptar el “*ignorabimus*”, sólo queda el investigar para culminar en un se sabrá.

Martin Gardner (Gardner, 1996), el famoso escritor divulgador de ciencia y matemáticas, duda que la IA pueda alguna vez construir máquinas conscientes e inteligentes. Gardner se considera un “misteriano”; es decir, cree que el libre albedrío, la conciencia y otros aspectos de la mente son misterios que no puede explicar la ciencia.

Jerry Fodor, en otro tiempo exponente destacado de la teoría computacional de la mente, así como de la noción de que ésta no es un mecanismo de aprendizaje para uso general, sino que está dividida en módulos especializados en tareas específicas, ahora es más escéptico. Así, a partir de 1998 (Fodor, 1998), señaló que ciertas tareas cognoscitivas, como detectar el color o analizar sintácticamente una frase, se pueden reducir a computación sin grandes problemas. Sin embargo, aunque se divida la mente en muchos computadores o módulos especializados, sigue en pie la pregunta de cómo se integran los resultados de todas esas computadoras.

Stuart Sutherland en “The International Dictionary of Psychology” (Crick, 1994), ofrece la siguiente definición de conciencia: *La capacidad para tener percepciones, pensamientos y sentimientos. Palabra imposible de definir salvo en términos que resulten ininteligibles sin el menor atisbo de lo que significa realmente. Son muchos los que caen en la trampa*

de equiparar conciencia y autoconciencia, para ser consciente sólo se necesita serlo del mundo exterior. La conciencia es un fenómeno fascinante, pero escurridizo. Es imposible especificar qué es, qué hace o por qué ha evolucionado. Sobre ella no se ha escrito nada que merezca la pena leer.

La definición más razonable de conciencia es la que emplean Koch y colegas (Koch, 1994) cuando afirman que tener conciencia es simplemente sentir o percatarse, fenómeno que, probablemente, se dé también en muchos animales. La conciencia es también el corolario de determinados procesos físicos que tienen lugar en determinados tipos de materia, sin los cuales la conciencia no puede existir. Esta visión desemboca en un materialismo extremo que rechaza cualquier esquema que sitúe a la mente y a la materia en el mismo plano de igualdad y, menos aún, cualquiera que privilegie a la mente. Dicho llanamente, todo el mundo puede ver “zombis”; esto es, cuerpos sin mente, pero sólo los místicos, los videntes y los psicóticos han visto mentes sin cuerpo.

En un apunte de su diario de 1846, Soren Kierkegaard (Kierkegaard, 1980) escribió: *Es totalmente natural que un hombre diga, con sencillez y seriedad que no puede entender cómo aparece la conciencia. Pero que el hombre pegue el ojo a un microscopio y mire y mire y siga mirando y a pesar de ello sea incapaz de comprender lo que ocurre es ridículo. Y es particularmente ridículo cuando se supone que es serio... Si en tiempos de Sócrates las ciencias naturales, hubieran tenido el desarrollo que tienen ahora, todos los sofistas habrían sido científicos. Alguno habría colgado un microscopio en la puerta de su tienda para atraer clientes, con un cartel que dijera: ¡Aprenda y descubra por medio de un microscopio gigante cómo piensa el hombre! (y al leer el anuncio Sócrates habría dicho: “Así se comportan los hombres que no piensan”).*

Una fuente de confusión en el debate sobre la conciencia lo constituye el hecho de que la conciencia se defina de forma “ad hoc”; es decir, de acuerdo con la escuela de pensamiento a la que pertenece quien la define. Por ejemplo, para Danah Zohar (Zohar, 1990), conciencia significa, autoconciencia o incluso un tipo especial de hiperconciencia mística, pero ésta se manifiesta sólo ocasionalmente y en unos pocos humanos, toda vez que es simplemente un tipo especial de conciencia en la que el objeto es uno mismo. En suma, el actor y el objeto de la física cuántica. Por su parte, Roger Penrose (Penrose, 1989) cuando habla de conciencia la entiende en términos de capacidad para resolver

problemas extremadamente complejos. La mayoría de la gente a eso le llamaría inteligencia y bastante especial.

Karl Pribram (Pribram, 1991), neurocientífico puso de actualidad una teoría, en otros tiempos muy popular, que él había desarrollado en los años sesenta de la memoria como holografía. El rasgo más sorprendente de una holografía, técnica en la que se produce una imagen tridimensional por superposición de dos rayos láser, es que cada parte de la imagen incorpora la totalidad de ésta, aunque con una resolución inferior. De este modo, se puede reconstruir toda la imagen a partir de un componente diminuto. Pribram afirmó que unas claves mnemotécnicas diminutas podrían también reproducir recuerdos enteros. El punto débil del modelo holográfico era que no parecía haber ningún equivalente neuronal de los rayos láser. Posteriormente, se descubrió que grandes grupos de neuronas se disparan a menudo repetidas veces escalonadamente y a la misma frecuencia, al igual que las ondas de luz de un láser. Pribram sugirió que estas neuronas oscilantes generaban unos campos electromagnéticos diminutos análogos a las ondas de luz de un holograma. La interferencia y la resonancia entre los campos superpuestos generaban algunas, sino todas, las propiedades mentales. De opinión parecida es, el ya citado, Danah Zohar (Zohar, 1990), quien afirma que había llegado el momento de superar el dualismo y de aceptar que tanto la materia como la mente eran productos de una fuente más profunda, “el quantum”. Y aseguraba que los pensamientos humanos, eran fluctuaciones cuánticas de la energía vacía del universo, que *es realmente Dios*. Stuart Hameroff principal organizador de la Primera conferencia científica interdisciplinaria americana sobre la conciencia, organizada por la Universidad de Arizona, Tucson, en Abril de 1994 (Hameroff, 1996) es partidario de la conciencia cuántica.

2.2. Enfoques

2.2.1. Filosófico

Hasta donde se sabe, el primero en tratar los procesos mentales, fue Demócrito, el filósofo de Abdera y padre, con Leucipo, del atomismo. Allí y entonces, Demócrito afirmaba que los procesos mentales sólo son la consecuencia, dentro del cráneo, de la actividad caótica de los átomos en continuo movimiento (Leucipo, 1996). Y añadió: *Pobre mente, tú que recibes de nosotros tus pruebas, ¿tratas de demolerlos? Nuestra*

derrota será tu propia ruina. Este concepto fue recogido por Thomas Hobbes y la escuela empírica inglesa que entendía la actividad cognoscitiva como una compleja forma de cálculo, independiente de la realidad física (Hobbes, 2001).

Platón (Platón, 1997), tenía una idea diametralmente opuesta a la de Demócrito. De hecho, creía que el pensamiento no es de naturaleza física, sino una entidad muy sensible e independiente de la corporeidad que lo contiene. Para Platón el gran descubrimiento de Anaxágoras de Clazómenas, maestro de Demócrito, fue el afirmar que es la mente la que lo ordena todo y es la causa de todo. Anaxágoras (Empédocles, 1996) efectivamente, afirmaba que la mente es autónoma, no se mezcla con nada y ordena todas las cosas que tienen vida tanto las más grandes como las más pequeñas marchando a través de ellas. De hecho, fue el primero que dio primacía a la mente sobre las cosas, aun cuando no atendió con cuidado a la causa eficiente, ya que describió algunos torbellinos sin-sentido aparejados a la inactividad o insensatez de la mente. Hasta tal punto Anaxágoras promocionó la mente que, según Diógenes Laercio (Laercio, 1985) y Plutarco (Plutarco, 1821, 1822, 1830), de él dice Timón en sus sátiras:

*Se cuenta por allí que existe Anaxágoras, valiente héroe,
[llamado] Mente, pues tenía una mente que despertó súbitamente,
y combinó todas las cosas, que, juntas, habían estado confundidas.*

Para Anaxágoras el intelecto es una cosa, aunque esté incontaminado y sea autónomo, siendo *la más sutil y pura de todas las cosas*. Y continuaba: *Todas las cosas tienen una porción de todo, pero la mente (nous) es infinita, autónoma y no está mezclada con ninguna, sino que ella sola es por sí misma.*

El cerebro – es mayor que el cielo –

Pues – ponlos juntos –

Y el uno al otro contendrá

Fácilmente – y a ti – además.

Platón estaba convencido de que el cerebro era el asiento de la conciencia. Aristóteles, sin embargo, apostaba por el corazón. Se ha realizado una ingente cantidad de trabajo para proporcionar una disección detallada de las estructuras cerebrales desconocidas que son importantes para la conciencia que van desde, sólo por citar a los pioneros, los trabajos de Wilder Penfield (Penfield, 1975), hasta el trabajo de Crick (Crick, 1994). Sin embargo,

los neurocientíficos actuales debaten acaloradamente qué regiones de la corteza cerebral, o incluso qué neuronas concretas, contribuyen o no a la experiencia consciente.

La idea de la mente de Nicolás de Cusa (1401-1464) (Cusa, 2005), iniciador de la filosofía moderna, y el distinguir los diversos modos de conocimiento, aporta una originalísima perspectiva de interpretarla como ligada a la “mensura”, esto es la medición.

René Descartes planteó el problema desde la perspectiva del conocimiento o pensamiento dubitativo con su ya citado famoso entimema: *Cogito ergo sum*. El entimema de René Descartes (Descartes, 1642), tomado literalmente, conduce al solipsismo filosófico; es decir, a la creencia de que nada existe fuera de la propia conciencia. Sin embargo, considerado de modo más realista, conduce a las posturas filosóficas idealistas, que son las que hacen más hincapié en la mente que en la materia; esto es, en las ideas, la percepción, el pensamiento; en suma, en la conciencia. Pero al tomar la mente como punto de partida, estas filosofías se las ven y se las desean para explicar la materia, lo que no es, por otra parte, necesariamente menos peliagudo que partir de la materia para explicar la mente. Descartes sostenía en su libro “Las Pasiones del Alma” (Descartes, 1947) que había una distinción absoluta entre la sustancia mental, “res cogitans”, y la material, “res extensa”. Ésta estaba caracterizada por su extensión; o sea, su ocupación del espacio, lo que, en consecuencia, la hace susceptible de explicaciones físicas. Aquella, por su parte, se caracteriza por estar consciente o, en sentido amplio del término, por pensar. Desde este enfoque, la sustancia mental sólo existe en forma de mentes individuales. Esta postura filosófica, intuitivamente sencilla y atractiva, es, científica y epistemológicamente, insatisfactoria al menos para explicar la conexión entre mente y cuerpo que, según Descartes, se hacía en la glándula pineal.

Descartes, en efecto, fundamentó toda su física en una definición esencialista o aristotélica del cuerpo o de la materia: Un cuerpo, en su esencia o sustancia, es extenso, y la materia en su esencia o sustancia es “extensión”. De este modo, la materia es sustancia extensa en oposición a la mente que, como sustancia pensante o de experiencia, es, en esencia, “intención”. Dado que el cuerpo o materia es idéntico a la extensión, toda extensión, todo espacio es cuerpo o materia: el mundo es pleno, no hay vacío. Es la teoría de Parménides, tal como la entendía Descartes. Descartes preparó el terreno para las

mónadas no extensas de Leibniz, al afirmar la división infinita de la materia. Un punto no tiene extensión y consecuentemente es inmaterial. Leibniz aceptó la ecuación cartesiana fundamental: cuerpo-extensión. Pero mientras Descartes pensaba que era irreductible; es decir, evidente, clara y distinta, Leibniz la cuestionó. Esto le indujo a las mónadas. Pero esa es otra historia, así que, retomando la cuestión se dice que Descartes dio la primera propuesta de solución.

Es decir, Descartes con su aseveración hacía un reconocimiento explícito del papel central que representa la conciencia, tanto con respecto a la ontología; o sea, qué es, como a la epistemología; esto es, qué se conoce y cómo se conoce. De hecho, para el Nobel y padre de la Etología Konrad Lorenz (Lorenz, 1974) el descubrimiento del propio yo que implica este asunto, es el mayor de cuantos hiciere el hombre a lo largo de su historia natural. Ahora bien, la cita cartesiana se ha popularizado erróneamente al traducirla como “Pienso, luego existo”, pues realmente significa “dudo, luego existo”. Esto es, el saber sobre la propia existencia humana es inmediato e indiscutible, pero de ello no cabe inferir que el saber mediato y distal sea incierto. Sólo la unión de ambos saberes resolverá el nudo del mundo. Dicho de otro modo, siguiendo a Percy Bridgman (Bridgman, 1958) cuando comenta la actitud epistemológica del padre de la Física Cuántica Niels Bohr: *The object of knowledge and the instrument of knowledge cannot legitimately be separated, but must be taken together as a whole*. Esta postura que, de acuerdo con Campbell (Campbell, 1966), se denomina **realismo hipotético**, es la que, aunque con matices, toma el doctorando. Cuando Descartes planteó su entimema, colocó el pensamiento sobre un pedestal donde la ciencia no podía alcanzarlo, al dividir el Universo de tal manera que todo lo material se podía medir, contar y pesar, mientras que los pensamientos pertenecían a otro dominio distinto del tiempo y el espacio cuantificable.

Como lo indicó Edelman, un buen reflejo de la arrogancia humana es el hecho de que se hayan erigido sistemas filosóficos completos sobre la base de una fenomenología subjetiva, como es la experiencia consciente de un solo individuo con inclinaciones filosóficas, de las que el “cogito ergo sum” de Descartes es el ejemplo paradigmático. Ahora bien, tal y como el propio Descartes sugirió, esta arrogancia está justificada dado que la propia experiencia consciente es la única ontología sobre la que se tienen evidencia directa. Sin embargo, como lo señaló Schopenhauer, esa afirmación genera una curiosa paradoja. La inmensa riqueza del mundo fenomenológico que experimenta el ser humano;

es decir, la experiencia consciente como tal, parece sustentarse y, por lo tanto, depender, de una nimiedad del mensaje de este mundo: un trazo gelatinoso del interior del cráneo. El cerebro humano, un actor secundario y fugaz que casi ninguno llega a ver sobre el escenario de la conciencia, parece ser el guardián de todo el teatro. De la fugacidad y vulnerabilidad de cerebro es prueba evidente y contundente que la más mínima agresión al mismo, bien sea física, como química (anestésico, tóxicos, etc.), puede modificar permanentemente el mundo exterior e incluso atribular a la persona.

Ludwing von Bertalanffy (Bertalanffy, 1964) dice que el dualismo-cartesiano entre mente “res cogitans” y materia “res extensa”, cerebro y conciencia, yo interior y yo exterior, son incorrectos pues es una conceptualización procedente de la física del XVII. La ciencia no debe hacer afirmaciones metafísicas, ya sean de la variedad materialista, de la idealista o de la positivista, de datos sensoriales. Es una construcción conceptual que reproduce aspectos limitados de la experiencia en su estructura formal. Esto debe llevar a una “teoría unificada” de la cual puede derivarse, tanto aspectos materiales como mentales, conscientes o inconscientes. Mente y materia son conceptos que impiden el progreso (Bertalanffy, 1974).

Otras versiones del dualismo son la teoría del doble aspecto de Spinoza, el ocasionalismo de Malebranche, el paralelismo de Leibniz y su doctrina de la armonía preestablecida (Leibniz, 1926) (Leibniz, 1981) y el epifenomenalismo que llegó hasta la actualidad.

Se considera un epifenómeno como algo que acompaña a un fenómeno pero no forma parte de él; por ejemplo, la sombra y el objeto del que procede. El epifenomenalismo, afirma que los eventos mentales y los físicos son diferentes, pero sostiene que las únicas causas verdaderas de las experiencias mentales son eventos físicos, en tanto que la mente es un subproducto incapaz de actuar como causa. En palabras de Thomas Huxley (Huxley, 1901): *La consciencia no parece estar relacionada con el mecanismo del cuerpo más que como un producto secundario de su funcionamiento, y no parece que tenga más poder de modificar su funcionamiento que el que tiene el pitido del vapor de influenciar la maquinaria de una locomotora. De qué modo algo tan destacable como un estado de conciencia resulta de la irritación del tejido nervioso, es algo tan inexplicable como lo era la aparición del genio cuando Aladino frotaba la lámpara.*

El filósofo Owen Flanagan, abundó en la línea de Koch y Freeman afirmando que probablemente había un mecanismo capaz de resolver el enigma de la conciencia. Flanagan (Flanagan, 1991) (Flanagan, 1992) se declaraba a favor del naturalismo constructivo, filosofía según la cual la conciencia es un fenómeno biológico corriente que se da no sólo entre los humanos, sino también en otros animales. La comprensión plena de la conciencia sería resultado de lo que él denominaba proceso de triangulación; es decir, una combinación de informes de los humanos sobre sus propias experiencias subjetivas y datos objetivos tanto de la psicología como de la neurociencia.

Para Locke (Locke, 1980), la cuestión central de la filosofía consiste en investigar la multiplicidad de procesos de la mente. Es decir, lo que le interesa no es únicamente el conocimiento científico sino, más generalmente, todo lo que se refiere al entendimiento humano.

Feigl (Feigl, 1934) sugiere que el problema mente-cerebro surge únicamente por el empleo de dos lenguajes distintos, el psicológico de las experiencias introspectivas inmediatas (datos) y el físico de las abstracciones (eventos espacio-temporales). Afirma que la introspección solamente puede usar el lenguaje de los datos, que es muy universal pero simplista y que el lenguaje inter-subjetivo es el físico, que expresa exclusivamente formas y estructuras lógicas. Cree que no hay razón para negar que las investigaciones neurofisiológicas serán en principio capaces de expresar en el lenguaje físico los correlatos neurofisiológicos de los procesos mentales, y que, si esto se lograra, toda la psicología introspectiva que se expresa en el lenguaje de los datos podría entonces ser considerada como una parte del lenguaje físico. Feigl concluye que el problema cerebromente es un pseudo-problema, creado exclusivamente por el lenguaje. Skinner y Köhler (Skinner, 1953) (Köhler, 1938) concuerdan en este enfoque.

Russell, en 1927 (Russell, 1927), defendió el “monismo neutro” según el cual “tanto la mente como la materia son estructuras constituidas por una sustancia más primitiva, que no es ni mental ni material”. En 1948, Russell (Russell, 1948) ya no se refirió al “monismo neutro”, sino que definió los eventos materiales como aquellos que sólo pueden ser conocidos, sin necesidad de ser inferidos. Russell consideraba que los eventos mentales y los procesos neurofisiológicos con los que están correlacionados son simultáneos.

Hace años William James alarmó al mundo con un artículo titulado “¿Existe la conciencia?” que reimprimió en su libro titulado “Ensayos sobre Empirismo Radical” (James, 1912). En él, exponía su opinión de que *existe solamente una materia prima en el mundo*, y que el término “conciencia” representa una función no una entidad. James sostenía que hay “pensamientos” que efectúan la función de “conocer”, pero que estos pensamientos no están constituidos por ninguna “materia” diferente de la que integra los demás objetos materiales. De esta forma, estableció los fundamentos de lo que se ha denominado “monismo neutral” que, con Russell (Russell, 1975), defiende la mayoría de los filósofos realistas americanos.

Existen dos significados muy distintos que se atribuyen a la palabra “conciencia”. Por una parte, se dice que alguien es “consciente de” algo. En este sentido, “conciencia” representa una relación. William James (James, 1890) definía la conciencia como la facultad que surge cuando el sistema nervioso se ha vuelto demasiado complejo para regularse a sí mismo. Dicho de otra forma, **la conciencia es como una propiedad emergente del cerebro que asume el control de su propio sistema nervioso**. Sin embargo, nadie es capaz de decir, cuándo ni cómo surgió esa propiedad. James, consideraba a la conciencia como un proceso, privado, selectivo y continuo, pero en constante cambio. Por otra parte, la “conciencia” puede considerarse como una cualidad de los hechos mentales, que no estriba en su relación con otros objetos. Si se quiere saber lo que quiere decirse comúnmente por “conciencia” será preciso previamente establecer lo que se entiende por hecho “mental” relacionado con un “objeto”; por ejemplo, un acto habitual de percepción. Así los sucesos mentales serán aquellos que tienen lugar en el cerebro y producen una reacción del género que puede denominarse conocimiento.

Searle combina el fisicalismo, cuyo lema es “el cerebro es la causa de la mente”, con la insistencia en que los estados mentales gozan de existencia independiente y no pueden ser “reducidos” a estados cerebrales. Queda así enfrentado tanto a aquellos que identifican los estados mentales con estados físicos del cerebro, como a los funcionalistas, como Jerry A. Fodor (Fodor, 1983), que afirman que los estados mentales **emergen** de los patrones o relaciones entre los componentes de un sistema físico. Para los funcionalistas, el material del que está hecho un sistema es irrelevante con respecto a los estados mentales que genera. Los componentes pueden ser neuronas, circuitos impresos, o latas viejas, siempre y cuando se encuentren adecuadamente relacionados, y la información

pueda circular apropiadamente entre ellos. Muchos, entre los que se encuentran Searle y Penrose, aunque con argumentos distintos, afirman que la mera organización de este tipo; es decir, la “sintaxis”, no puede dar cuenta del significado o entendimiento; esto es, la semántica, esenciales al proceso consciente del pensamiento. **Y el doctorando añada, y menos aún al comportamiento mental, el tercer aspecto de cualquier sistema suficientemente complejo como para ser susceptible de conciencia.**

Por el contrario, el agente se equivocaría si creyese así, incluso aunque los observadores exteriores hiciesen bien en creerlo. Es ésta una clase curiosa y muy importante de “principio de relatividad lógica”. Para que los observadores puedan creer válidamente lo que creen acerca del cerebro del agente, es necesario que el agente creyera alguna otra cosa; porque si cree lo mismo que los observadores, entonces lo creído no sería válido para él o para ellos. Para decirlo de otro modo: “ninguna predicción de una elección típica podría ser creída por el agente sin afectar a su propia validez”.

Si se está en lo cierto al pensar que la comprensión de la mente no se logrará por poner en claro algunos detalles no explicados del análisis físico, sino por una nueva comprensión de la significación de la actividad del cerebro y el cuerpo, entonces la mente tiene un contacto de trabajo con la materia más íntimo que el de una forma de energía sobre otra; porque mente y materia son verdaderamente complementarias, la misteriosa unidad es conocida como “sujeto personal”.

Algunos, como Dennett, afirmaron que si uno se va a referir a la “mente” como si fuera simplemente otra clase de sustancia, aunque diferente de la materia y satisfaciendo los principios de un tipo muy diferente, se está cometiendo un error categorial.

Para subrayar una vez más el punto principal, ninguna suma de conocimiento físico del estado de un mecanismo (cerebral) puede dar origen a una predicción de una acción típica que tiene fuerza lógica necesitante sobre el sujeto cuyo cerebro es. En resumen, la creencia de que el ser humano es normalmente libre al hacer una elección, lejos de ser “contradecible”, no tiene alternativa válida, desde el punto de vista de la física. En ese sentido, al menos, el ser humano es ineludiblemente responsable del control consciente de la acción. Aparte de esto puede señalarse que las leyes de la física están inferidas a partir de la experiencia, en tanto que la existencia del pensamiento es una experiencia

inmediata. Bertrand Russell acertó cuando dijo: *Sostengo que todo cuanto es conocido sin inferencia es mental.*

Hume (Hume, 1991) se inventó una teoría de valencias químicas para explicar cómo una idea podría llevar aparejada la siguiente. Lo que Hume pretendía era que las ideas pensarán por sí mismas. Quería librarse de la figura del pensador, porque comprendía que era un callejón sin salida. Si se mantiene a ese intermediario interno no puede haber progreso. La idea de Hume consistía en introducir enlaces covalentes entre las ideas, de manera que cada una pudiera pensar por sí misma y luego hacer pensar a la siguiente y así sucesivamente. De este modo, se eliminaba al intermediario, pero la cosa no funcionó. La única idea que de forma demostrable elimina el intermediario son los computadores. Los homúnculos ya no son un problema, porque se sabe cómo tratarlos. Se puede coger un homúnculo y dividirlo en homúnculos cada vez más pequeños, hasta llegar a un homúnculo fácilmente reemplazable por una máquina. Dennett pretendió profundizar la teoría de Hume.

Colin McGinn, cree que la conciencia no puede ser estudiada porque la arquitectura de los procesos cerebrales pone limitaciones a las capacidades cognoscitivas humanas. En breve, la mente humana simplemente es incapaz de resolver ciertos problemas. Otros, como Daniel Dennett niega que haya un problema en absoluto. Dennett argumenta, siguiendo al neurólogo John Hughlings Jackson, en el siglo XIX, que la conciencia no es una operación distinta del cerebro; antes bien, es el resultado combinado del trabajo computacional de las áreas del más alto orden del cerebro concernidas con las últimas etapas del procesamiento de la información. Finalmente, John Searle y Thomas Nagel creen que la conciencia es un conjunto discreto de procesos biológicos. Los procesos son accesibles a análisis pero son difíciles de entender por ser muy complejos y de naturaleza holística.

El debate filosófico **mente-cuerpo** es actualmente muy sofisticado y hoy existen las teorías siguientes: De la identidad, del estado central, el monismo neutral, el conductismo lógico, el fisicalismo de tipos y de casos, el epifenomenalismo de tipos y de casos, el monismo anómalo, el materialismo emergente, varias formas de funcionalismo, etcétera (Block, 1997), (Shear, 1997), (Warner, 1994), (Humphrey, 1993), (Flanagan, 1992), (Searle, 1992), (Dennett, 1991). Pero la profusión de posturas filosóficas, no parece

probable que puedan proporcionar, al menos por si solas, una solución satisfactoria al problema mente-cuerpo. En palabras de Colin McGinn (McGinn, 1989): *Llevamos mucho tiempo intentando resolver el problema cuerpo-mente, pero se nos ha resistido obstinadamente y el misterio persiste. Creo que ha llegado el momento de decir con franqueza de que no seremos capaces de resolver el misterio. Seguimos sin tener idea de cómo el agua del cerebro se convierte en el vino de la conciencia.* De hecho, los filósofos no han destacado tanto por su capacidad para proponer soluciones al problema, cuanto por su habilidad para resaltar hasta qué punto el problema es intratable. Su postura, más o menos, puede plantearse como sigue: Hagan lo que hagan los científicos, será imposible reconciliar las perspectivas de primera y tercera personas, el yo y el ello, de Freud, de individuos conscientes, imposible tender un puente entre las explicaciones enfrentadas, e imposible resolver el problema fundamental, cual es: la generación de sensaciones, de estados fenoménicos o experienciales, a partir de, como dijo Chalmers (Chalmers, 1995), “el zumbido de las neuronas”.

2.2.2. Científico

El neurocientífico cognitivo Francisco Varela (Varela, 1996), sugirió que la neurociencia, la fenomenología y la ciencia conductista habían de unirse para resolver el problema de la conciencia. Sin embargo, para la premio Nobel Rita Levi Montalcini (Levi, 2000) las tres líneas más importantes de investigación del cerebro y la mente, que han dado resultados más novedosos son: las neurociencias, las ciencias cognoscitivas y la IA. No se va a entrar aquí en la polémica entre la IA fuerte y la débil, pero si señalar que Searle (Searle, 1990) niega el principio básico de la IA de que la mente es al cerebro lo que el programa es al computador. La IA fuerte cae en el error cartesiano de que las facultades mentales pueden separarse del substrato estructural de las circunvoluciones neocorticales. Según la IA fuerte, la mente es independiente del cerebro: es un programa de computador y como tal no hay ningún circuito que sea su soporte. Es éste el cuarto gran desafío que se le plantea al ser humano, es el de la posibilidad, o no, de desarrollar inteligencia en otros entes distintos de los que hasta ahora tenían la exclusividad de la misma: los humanos. Y en éstas anda la humanidad; es decir, desarrollando la IA y de paso ayudando a entender mejor la mente. Ahora bien, en el caso mejor, la inteligencia es una de las funcionalidades mentales; en consecuencia, incluso obteniéndola fuera del cerebro, seguiría sin resolverse el problema cabal de la mente.

En este sentido, Maxwell (Mackay, 1992) decía: *Así, la ciencia deshace, una tras otra, las materializaciones más o menos groseras con que intentamos formar una imagen objetiva del alma. Hasta los científicos, cuando especulan como cualquier ser humano en sus ratos no profesionales sobre dónde nos puede conducir la ciencia, han profetizado que muy pronto tendremos que confesar que el alma no es más que una función de ciertos sistemas materiales complejos.* En este punto es necesario recordar a un desesperado Maxwell quien se planteó una cuestión en principio indecidible; a saber, cómo explicar el *pensamiento ad ovo*, esto es, a partir de las leyes que gobiernan el movimiento de las moléculas en el cerebro. Maxwell despachó la cuestión con la memorable frase: *but does not the way to it lie through the very den of the metaphysician, strewn with the bones of former explorers and abhorred by every man of science?*.

Actualmente, se han adoptado posturas más materialistas si cabe y, desde luego, extremadamente reduccionistas que llevan a afirmar que mente y conciencia son lo mismo que el funcionamiento del cerebro o al menos que ciertas funciones del funcionamiento del mismo. Algunas posiciones son tan radicales que niegan a la conciencia cualquier valor ontológico o epistémico, hasta el punto de insistir en que no hay nada más allá del funcionamiento de los circuitos del cerebro o, cuando menos, que no hay nada que necesite ser explicado. De hecho, arguyen que cuando se comprenda suficientemente el funcionamiento del cerebro, el concepto de conciencia se disipará del mismo modo que se disipó el concepto de flogisto (el hipotético constituyente volátil de todas las sustancias combustibles que, según se creía, se liberaba en forma de llama durante la combustión), una vez se comprendió el proceso de la oxidación. El ejemplo de Raymond Kurzweil, retomado de Turing (Turing, 1950) sin citarlo, es muy descriptivo al respecto (Kurzweil, 1999): *Comprender la inteligencia es como deshacer una cebolla. Cada capa descubre una nueva cebolla y al final del proceso tenemos muchas capas, pero no tenemos ya la cebolla.* De este modo, se consigue obviar el problema cuerpo-mente negando o desestimando el problema de la conciencia.

El padre de la física cuántica y premio Nobel de física de 1918, Max Karl Ernst Ludwig Planck, analizó muy atinadamente el problema, mente-cerebro, que él denominó “cuerpo-alma” concluyendo que “los procesos corporales y los procesos psíquicos no son en absoluto distintos unos de otros. Se trata de los mismos procesos sólo que considerados desde perspectivas contrapuestas”. Y termina diciendo que “el problema alma cuerpo

resulta ser un problema aparente”. T.H. Key (Key, 1986) señaló esto en un ingenioso juego de palabras: “What is mind? No matter. What is the matter never mind”

Max Planck (Planck, 2000) afirmó que *nunca se captará este misterio del Universo (la mente) porque nosotros mismos somos el misterio*. En efecto, Planck en una conferencia impartida en Gotinga el 17 de junio de 1946, definió los problemas aparentes, como aquellos en los que las premisas esconden algún error, tanto explícito como tácito, y los que están, por ambigüedad o vaguedad, formulados de manera insatisfactoria. Después de poner varios ejemplos de ambos y de mostrar como en algún caso lo que hoy es un problema aparente puede convertirse en el futuro en uno real, como sucedió en el sueño en el caso de convertir un metal común, el mercurio en oro, abordó el problema “cuerpo-alma” o “mente-cuerpo” y lo enfoca como sigue: los procesos mentales no necesitan ir acompañados en modo alguno, por procesos corporales sino que pueden producirse con total independencia de estos últimos. Entonces, el problema mente-cuerpo se divide en dos problemas distintos. El de la mente y el del cuerpo, con lo cual pierde su sentido y degenera en un problema aparente. Después de analizar la volición y los procesos psíquicos en general, señala que hay dos puntos de vista para analizarlos, a saber, el interno, que perturba dichos procesos, y el externo, que no los perturba y por eso es el adecuado. Ambos puntos de vista se excluyen mutuamente y como consecuencia no tiene sentido implicarlos simultáneamente. Con este presupuesto, afirma que la voluntad considerada desde fuera, está causalmente determinada, en tanto que considerada desde dentro, es libre. Esto, según Planck, resolvería el problema del libre albedrío. Es decir, su aparición como problema se debió inicialmente a que no se tuvo cuidado de explicitar el punto de vista desde el que se pretendía abordarlo ni de atenerse luego a él. Es, por tanto, un ejemplo típico de problema aparente.

Por su parte, Eddington, acerca del problema mente-cerebro, presentó opiniones diversas y contradictorias que iban desde el monadismo al triadismo, para explicar las voliciones, pasando por el dualismo que es el más sensato, al menos para Rosenblueth. Éste, uno de los padres de la cibernética y fisiólogo (Rosenblueth, 1971), al respecto, dice que para resolver el problema, a los cuatro postulados científicos agregó tres más. Son estos siete postulados los siguientes:

- 1) Aun cuando los únicos eventos de los cuales me puedo percatar directamente son mis propios procesos mentales, otros organismos afines, humanos o animales, tienen experiencias similares, de las cuales, a su vez, se percatan directamente.
- 2) Nuestras sensaciones están relacionadas causalmente con eventos que ocurren en un universo material del cual forman parte nuestros organismos.
- 3) Existen regularidades en la sucesión temporal de eventos que ocurren en un material, y en algunas características mensurables de los sistemas materiales relativamente aislados cuando están en equilibrio.
- 4) En la elaboración de las teorías científicas se aceptan como válidas los siguientes métodos de inferencia: la deducción, la inducción probabilística, la analogía y la abducción o retroducción.
- 5) Los procesos físicos correlacionados con nuestros eventos mentales son fenómenos neurofisiológicos que ocurren en nuestros cerebros.
- 6) Cada evento mental específico está correlacionado con un diseño espaciotemporal específico de actividad neuronal.
- 7) Las leyes de la física son aplicables a todo el universo material, incluyendo nuestros organismos y sistemas nerviosos; además, ni los procesos neuronales, ni los mentales, están sujetos a cualquier determinismo, causal o probabilístico, distinto del que es aplicable a los fenómenos físicos.

Bohr, a instancias de Heisenberg (Heisenberg, 1996), trató el problema de la conciencia. Así, después de plantearle Heisenberg una cuestión acerca de la teoría darwiniana le dice: Un argumento más que se aduce a veces a favor de la necesidad de ampliar la teoría cuántica es la existencia de la conciencia humana. Es evidente que el concepto de “conciencia” no aparece ni en la física ni en la química, y no se puede comprender en realidad cómo deba darse algo parecido en el campo de la mecánica cuántica. En una ciencia de la naturaleza que abarque también los organismos vivos la conciencia debe tener su puesto porque también ella pertenece a la realidad. A lo que Bohr contestó: *Este argumento parece, a primera vista, muy convincente. Entre los conceptos fisicoquímicos no podemos encontrar nada que ni de muy lejos tenga que ver con la conciencia. Sabemos que la conciencia existe solamente porque nosotros mismos la poseemos. “La conciencia es, por tanto, también una parte de la naturaleza, o, en términos más*

generales, de la realidad; y debemos poder describir y entender al lado de la física y de la química, cuyas leyes están abarcadas en la teoría cuántica, o mediante otras leyes, de esencia totalmente distintas. Pero no sé, a este respecto, si se precisa más libertad que la ya dada para el concepto de complementariedad. También a mi me parece que es poca la diferencia que existe entre un método que, como el de la interpretación estadística de la termología, incorpora nuevos conceptos a la mecánica cuántica y formula con ellos nuevas leyes, y un procedimiento que, como sucedió con la ampliación de la física clásica que llevó a la teoría cuántica, debe ampliar la misma teoría cuántica con un formalismo más general para abarcar también la existencia de la conciencia. El verdadero problema es el siguiente ¿cómo puede concordar la parte de la realidad que comienza con la conciencia con aquella otra que se describe con la física y la química? ¿cómo hacer para que no entre en conflicto las leyes propias de estas dos partes? Aquí se trata, evidentemente, de una estricta situación de complementariedad, que debiera analizarse naturalmente, con mucha mayor exactitud cuánto más adelante sean más amplios nuestros conocimientos biológicos”.

En las famosas charlas que impartió el físico de formación Max Delbrück, premio Nobel de Medicina y Fisiología en 1969, en 1974-1975 en el Caltech, acerca del problema mente-cerebro o mente y materia, y que definió como *una investigación de la capacidad cognoscitiva de la especie humana, tal como se manifiesta en diversas ciencias*, Delbrück, convirtió la epistemología de Bohr en la infraestructura molecular de la biología molecular. Esto sirvió de carta de navegación, brújula y sextante, o mejor de guía filosófica para navegar entre la Escila del puro reduccionismo bioquímico y el Caribdis del vitalismo oscurantista. Delbrück tenía un interés particular en examinar si hay algún área de estudio para la cual el concepto de complementariedad pudiera ser una ayuda decisiva (Stent, 1989). El objetivo de Delbrück (Delbrück, 1989) era plantear, examinar y, si fuera posible, contestar las tres cuestiones siguientes:

1. ¿Cómo es posible que en un mundo que inicialmente ni contenía vida, ni, por lo tanto, mente alguna, apareciera la inteligencia?
2. Si para contestar a la pregunta anterior se usa la teoría darwiniana, ¿cómo es posible que el proceso haya dado lugar a una mente capaz de ahondar en las más complejas abstracciones matemáticas, en la estructura de la materia o en la naturaleza de la vida misma, todo lo cual era de dudosa o nula utilidad para el hombre primitivo?

3. ¿Cómo pueden aparecer el entendimiento y la capacidad de discernir la verdad a partir de la materia inerte?

Delbrück, ni en sus conferencias ni en sus escritos, da la más mínima indicación explícita acerca de lo que él entiende por “mente” o acerca de la relación entre mente y cerebro y cómo los sentimientos, los pensamientos y la consciencia aparecen en el cerebro humano y que el origen evolutivo de la mente debe ser sinónimo del desarrollo del cerebro humano. En un primer momento, mientras adopta la postura epistemológica del realismo ingenuo, parece aceptar gratuitamente que los fenómenos mentales son productos directos o epifenómenos del funcionamiento del cerebro humano. Posteriormente, abandona dicha postura y considera a la mente como un fenómeno unitario con dos facetas. Una, que se alcanza por “extrospección”, desde fuera, y da cuenta del cerebro como una red formada por millones de células nerviosas. Otra, introspectiva, desde dentro; es decir, directamente en uno mismo e indirectamente por proyección en los demás, que presenta a la mente como un complejo de pensamientos, sentimientos y conciencia. Delbrück, basándose en el principio de complementariedad de Bohr, destaca la improcedencia de todas las versiones del realismo que tienen su origen en una visión errónea de la relación entre mente y materia. Visión ejemplarizada en la “reducción cartesiana” que, como se ha reiterado, escinde la existencia entre un mundo interior de pensamientos, volición y emociones: la “res cogitans” o mente, y un mundo exterior de cosas: la realidad física o “res extensa”. Ahora bien, si los fenómenos de la mente andan parejos con los de la materia; es decir, si la reducción cartesiana no es más que una ilusión, ¿entonces la mente no es más que una máquina material, equipada con una compleja red de circuitos cibernéticos? Para él así es.

Sin embargo, últimamente, la cosa ha cambiado y muchos neurocientíficos, como es el caso de John Eccles (Eccles, 1990), han pasado a estudiar este asunto considerando a la conciencia como un parámetro experimental más. Si bien ciertas hipótesis científicas recientes no abarcan un abanico tan amplio como el que ofrecen los filósofos, en cierto sentido son más exóticas y extremas. Por ejemplo, algunos neurocientíficos han tomado posiciones dualistas, según las cuales la mente consciente interactúa con el cerebro gracias a “psicones” que se comunican con los “dendrones” o dendritas en determinadas regiones del hemisferio cerebral izquierdo. Otros científicos, como Penrose, han llegado a la conclusión de que la física convencional no es fundamento suficiente para las teorías

sobre la conciencia, e invocan conceptos físicos esotéricos como la gravedad cuántica, para explicar la conciencia.

Un ferviente abanderado de la teoría cuántica de la conciencia es, el acabado de citar, John Eccles (Eccles, 1992), neurocientífico británico, premio Nobel de 1963 por sus estudios sobre la transmisión neuronal. Eccles es, sin duda, el científico moderno más eminente que defiende el dualismo, doctrina según la cual la mente existe independiente de su sustrato físico. En su libro, con el filósofo de la ciencia y padre del falsacionismo Karl Raymond Popper (Eccles, 1977) defienden paladinamente el dualismo y rechazan el determinismo físico a favor del libre albedrío: la mente podría escoger entre diferentes pensamientos y acciones emprendidas por el cerebro y el cuerpo (Eccles, 1994). La objeción más corriente al dualismo es que viola el principio de conservación de la energía. En efecto, ¿cómo, si no tiene existencia física, puede iniciar la mente cambios físicos en el cerebro? En colaboración con el físico alemán Friedrich Beck, Eccles ha ofrecido la siguiente respuesta: las células nerviosas del cerebro se excitan cuando moléculas cargadas, o iones, se acumulan en una sinapsis, haciéndole liberar neurotransmisores. Esto presenta un problema puesto que la presencia de un determinado número de iones en una sinapsis no siempre desencadena la excitación de una neurona. A esta objeción contesta Eccles que la razón es que, al menos durante un instante, los iones existen en una superposición cuántica de estados, de modo que en determinados estados la neurona descarga y en otros no. La mente ejerce su influjo sobre el cerebro “decidiendo” que neuronas se excitarán y cuáles no. Mientras se conserva la probabilidad a través del cerebro, este ejercicio de libre determinación no viola la conservación de la energía. La mala noticia es que Eccles y Beck no tienen ninguna prueba al respecto.

John Taylor, físico y especialista de redes neuronales, criticó, en la conferencia de Tucson, el enfoque cuántico pues, como aseveró, la no localidad y los demás efectos cuánticos exóticos que se supone cruciales para la conciencia suelen observarse sólo a temperaturas cercanas al cero absoluto o, en cualquier caso, muy por debajo de la temperatura ambiente de la mayor parte de los cerebros vivos. A eso añadía objeciones a la hipótesis cuántica basándose en pruebas pragmáticas, pues, hasta la fecha, la física nuclear no había tenido ninguna importancia para la biología.

El aparente no “determinismo” del Universo, es decir de las leyes físicas, propuesto por el principio de indeterminación de Heisenberg, y de los fenómenos mentales es debido a fenómenos no computables, tales como la acción a distancia que se ignora en las actuales leyes físicas. Características no computables deben añadirse a la cuenta para el aparente no determinismo de la teoría cuántica, la aparente impredecibilidad de la intuición (“insight”) y creatividad, y la aparente libertad de la elección individual.

Ahora bien, Bohr, Heisenberg y seguidores interpretan el principio de indeterminación como una propiedad ontológica; es decir, una propiedad inherente del mundo real. En tanto que Penrose, siguiendo a Einstein, la considera sólo como una cuestión epistemológica; esto es, la indeterminación sólo es debida las inherentes limitaciones de los mecanismos de observación. Tales limitaciones son ilustradas por el experimento sobre la red de pesca de Eddington donde un pescador descubre la “*ley de la naturaleza*” de que todo pez tiene al menos dos pulgadas de diámetro habida cuenta de que el tamaño de la malla de su red de pesca es justamente, de dos pulgadas.

Pero a los neurocientíficos actuales les interesa más saber cómo están estructuradas y funcionan las mentes actualmente. La conciencia, y el sentido subjetivo de la misma, siempre ha parecido ser un enigma metafísico. El legado del conductismo pervivió, hasta mediados del siglo XX, en la renuencia de los científicos a estudiar el más elusivo e ineludible de los fenómenos: la conciencia. Esta actitud cambió cuando el premio Nobel de física Erwin Schrödinger, en 1958, escribió un ensayo cuyo título lo decía todo: “Mind and Matter” (Schrödinger, 1958), cuyo primer capítulo trataba sobre las bases físicas de la conciencia. Para él la conciencia es un fenómeno del área de la evolución. Y añade que la conciencia y la discordia con el propio Yo están inseparablemente unidas, incluso deben ser proporcionales entre sí. Además creía que el cerebro humano no era el supremo “ne plus ultra” de un órgano pensante en el cual se refleja el mundo. Para él, resulta más verosímil suponer que una especie puede llegar a adquirir un mecanismo similar al humano, pero cuyas imágenes sean a las humanas lo que éstas son a las del perro, o las del perro a las del caracol, etcétera. Para Schrödinger, las sensaciones y los pensamientos no pertenecen al mundo de la energía y, en consecuencia, como lo señalaron Spinoza y Sherrington, no pueden producir el menor cambio en este mundo. En palabras de éste último: *La Física... nos plantea el “impasse” de que la mente no puede “per se” tocar el piano (la mente no puede “per se” mover un dedo de la mano). [...]. Frente al “cómo” la*

mente influencia la materia, el vacío. Schrödinger concluyó diciendo: Mi mente y el mundo están compuestos por los mismos elementos. Lo mismo ocurre para todas las mentes y los respectivos mundos, a pesar de la insondable abundancia de interacciones mutuas. El mundo me es dado de una sola vez: no uno existente y otro percibido. Sujeto y objeto son una sola cosa. Y no podemos decir que la barrera que los separa se ha roto como consecuencia de la experiencia reciente en la Física, porque esa barrera no existe. Según Schrödinger, la razón por la que no podemos encontrar nuestro ego sensible perceptor y pensante en lugar alguno de nuestra imagen científica del mundo, puede expresarse fácilmente con siete palabras: porque esta imagen es la mente misma. Todo el conocimiento científico se basa en los sentidos y, a pesar de todo, las descripciones científicas de los procesos naturales así elaborados carecen de todas las cualidades sensoriales por lo que no pueden dar cuenta de ellas, no pueden explicarlas.

Sir Charles Scott Sherrington en su libro “Man on his Nature” (Sherrington, 1951) intentó extraer evidencia objetiva de la interacción entre mente y materia. Así lo expresó: *Para todo aquello que alcanza la percepción, la mente interviene en nuestro mundo espacial como un verdadero fantasma invisible, intangible, es algo que incluso carece de contorno; no es una “cosa”. Se queda sin confirmación sensorial y así se queda para siempre.* Sherrington, decía al respecto: *If it is for mind that we are searching the brain, then we are supposing the brain to be much more than a telephone-exchange. We are supposing it a telephone-exchange along with the subscribers as well.* Realmente Sherrington consideraba que el problema de la mente estaba fuera del alcance del ser humano y, por lo tanto, carecía de solución; es decir, estaba convencido de que el problema de la conciencia era científicamente inexplicable (Sherrington, 1906).

Para Schrödinger, la interacción que buscaba Sherrington no existe. Más aún, la mente ha construido el mundo objetivo exterior fuera de su propia sustancia. La mente no ha podido abordar esta gigantesca tarea sin el recurso simplificador de excluirse a sí misma, de omitirse en su creación conceptual. De dónde se deduce que tal creación no contiene a su creador. De hecho Sir Charles Sherrington en el citado libro dice: *El mundo material se ha construido sólo a costa de extraer de él el yo; esto es, la mente.* Spinoza (Ethics. Pt. III. Prop2) (Spinoza, 1951): *Nec corpus mentem ad cogitandum nec mens corpus ad motum neque ad quietem nec ad aliquid (si quid est) aliud determinare potest.* Es decir: *El cuerpo no determina que la mente piense, ni la mente determina que el cuerpo se mueva,*

o esté en reposo, o cualquier otra cosa (si ello así sucediera). Jung, en un artículo publicado en Cuadernos de Eranos (Jung, 1946), va más allá y dice que: *Toda ciencia es, sin embargo, una función del alma en la que se arraiga todo conocimiento. El alma es el más grande de los milagros cósmicos, es la “conditio sine qua non” del mundo considerado como un objeto.* Para ello, Schrödinger hace una analogía entre un cuadro de Durero, la “Adoración de los Reyes”, y la mente. La mente es, por un lado, el artista que ha producido el todo; sin embargo, en la obra terminada, no es sino un accesorio insignificante que puede omitirse sin que por ello el efecto total pierda el menor mérito.

Antonio Damasio, neurólogo de la Universidad de Iowa (EE.UU.), ideó un ingenioso experimento para evaluar la intensidad de las sensaciones. Pidió a unos voluntarios que cogieran naipes de cuatro montones diferentes. Lo que no sabían es que cada elección implicaba perder o ganar dinero en función del montón elegido. Dos de ellos estaban dotados con 100 dólares y los otros dos con 50. Damasio colocó sensores en su piel para medir, mediante el sudor, su grado de implicación en el juego. La sudoración aumentaba poco antes de cada elección, y lo hacía más cuando el montón elegido era el de 100 dólares. Lo curioso es que la reacción se producía incluso cuando los participantes no eran conscientes del valor de cada montón.

Tras el experimento, Damasio desarrolló la hipótesis de los “marcadores somáticos”, según la cual el ser humano guarda todas sus vivencias en el cerebro y las clasifica como buenas o malas. Con este archivo forma una “memoria de vivencias” que le hace perseguir las positivas y evitar las negativas. Este mecanismo es, según Damasio, el que determina la toma de decisiones intuitivas. La intuición parece ser una capacidad que viene determinada por nuestros genes, pero que mejora continuamente gracias a las vivencias personales y a las experiencias del día a día. Por eso, es más habitual que la reacción espontánea de un veterano sea más correcta que la de un joven.

Las tomografías cerebrales han demostrado que, a medida que un aprendiz se convierte en experto, el lugar donde se procesa la información va cambiando de sitio. Un principiante, ya sea un músico, un taxista o un ornitólogo, utiliza su hemisferio cerebral derecho; un experto, el izquierdo.

Además, la sabiduría que muestran las intuiciones tiene mucho que ver con la gratificación y el placer que se obtiene al hacer o conseguir algo. Ese sistema de gratificación del cuerpo siempre se encuentra activado. El sistema límbico hace que uno se sienta bien cuando come golosinas, lee un verso hermoso o resuelve un problema. En ese momento, las células nerviosas liberan dopamina, la hormona de la felicidad, que produce una sensación de placer y que hace desear volver a experimentar esa sensación. Y el cerebro, al sentir ese placer, toma nota de que es bueno.

Para Kirsten Volz, del Instituto Max Planck de Neurología y Psicología Cognitiva de Leipzig, la sorprendente conclusión de su experimento fue que el cerebro hace sus cálculos sin que uno sea consciente de ello. Cuando los voluntarios participantes en el trabajo de laboratorio tuvieron que contestar qué ciudad de las distintas parejas que se les ofrecían tenían mayor población, no sólo activaron las regiones cerebrales donde reside la capacidad de reconocer nombres, sino también aquellas que procesan las informaciones abstractas. En otras palabras, cuando el cerebro se enfrenta a cuestiones peliagudas no sólo echa mano de las simples reglas de conducta aprendidas que le ayudan a desenvolverse, sino de aquellas que él ha ido creando poco a poco para tomar decisiones rápidas en situaciones cruciales.

Albert Einstein, según contó él mismo, alumbró la teoría de la relatividad de una forma intuitiva y calificó como “regalo de los cielos” esa capacidad que le había permitido hacer uno de los descubrimientos más destacados de la historia. Decía el físico *la intuición es la única cosa realmente valiosa*.

Cuando se toma una decisión, entra en acción un complejo mecanismo situado en el centro de la materia gris. En la figura 3.2 se muestran las estructuras implicadas.

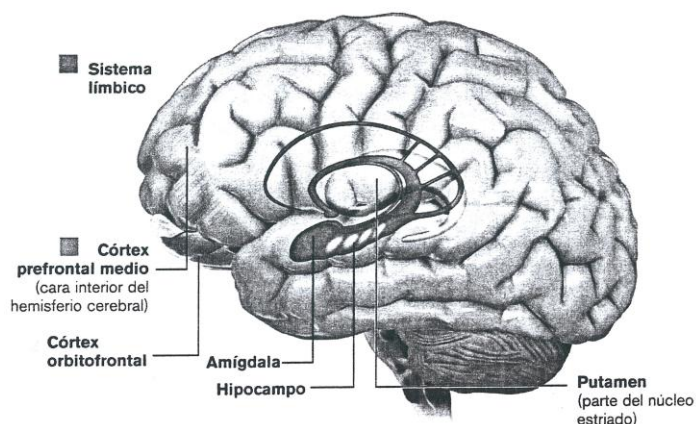


Figura 3.2. Centros donde se toman las decisiones.

El córtex prefrontal medio se activa en el mismo momento en que es necesario tomar una decisión, como por ejemplo resolver un problema o elaborar una estrategia. En el sistema límbico, la amígdala sopesa las señales emocionales, mientras que el hipocampo se encarga de almacenar esas valoraciones para utilizarlas, como experiencias, en el futuro. Los mensajes procedentes del sistema límbico y del cerebro medio llegan al núcleo estriado, que es el que ayuda a tomar una decisión correcta. El córtex orbitofrontal dirige el pensamiento a largo plazo y las pautas de comportamiento.

Para el premio Nobel de Física Murray Gell-Mann (Penrose, 1994) *La consciencia o autoconocimiento es una propiedad, como la inteligencia, que puede acabar evolucionando en los sistemas complejos adaptativos cuando alcanzan ciertos niveles de complejidad [...]. En principio, no me parece posible que nosotros los humanos podamos construir computadores con un grado de conocimiento razonable. Penrose parece atribuir al autoconocimiento alguna cualidad especial que hace improbable que surja de las leyes ordinarias de la ciencia.*

Otro científico que, con su dedicación al problema, dio carta de naturaleza al estudio de la conciencia fue, el ya citado, premio Nobel en 1972, también de Medicina y Fisiología, Gerald M. Edelman. Este premio, compartido con Rodney Porter, lo recibió por su trabajo de postgraduado, en el descubrimiento de la estructura de las inmunoglobulinas, proteínas cruciales en la respuesta inmune del cuerpo humano. Luego siguió trabajando en biología del desarrollo donde descubrió las proteínas de adhesión nuclear importantes para el desarrollo embrionario. Sin embargo, esto no fue más que el preludio, a modo de

entrenamiento, de su gran, difícil y definitivo, proyecto de elaborar una teoría de la mente. Su teoría la expuso en cinco libros: “Neural Darwinism” (Edelman, 1987), “Topobiology” (Edelman, 1988), “The Remembered Present” (Edelman, 1989), “Bright Air, Brilliant Fire” (Edelman, 1992) y el último, al menos por ahora, con Giulio Tononi, “A Universe of Consciousness” (Edelman, 2000). El “quid” de esta teoría es que, así como las tensiones ambientales seleccionan a los miembros más capacitados de una especie, así también las entradas en el cerebro seleccionan grupos de neuronas al reforzar las conexiones entre ellas. Edelman considera que una teoría que diera cuenta de la mente representaría la culminación definitiva de la ciencia, pues entonces ésta podría explicar su propio origen. Las propuestas de Edelman son las siguientes:

- a) Nunca se tendrá una teoría del cerebro realmente satisfactoria mientras no se tenga una teoría realmente satisfactoria de la anatomía neuronal. A esto se le puede hacer la siguiente crítica, dado que la anatomía trata con la estructura corporal, conocer sólo la anatomía no es suficiente. Faltan las funcionalidades neurofisiológicas y el comportamiento.
- b) La mente sólo puede abordarse desde un punto de vista biológico, no recurriendo a la física, la informática u otros enfoques que no tienen en su dominio la estructura del cerebro.
- c) El cerebro no es una máquina “replicable”. Cuando se mira en profundidad; es decir, a todos los niveles y hay un número increíble de ellos, en los cerebros, lo que realmente impresiona es su diversidad. Hasta los gemelos univitelinos muestras grandes diferencias en la organización de sus neuronas. Estas diferencias, lejos de ser ruido insignificante, es algo importantísimo que no se puede soslayar, so pena de cometer un grave error metodológico.

Esta enorme variabilidad y complejidad del cerebro, presumiblemente, está relacionada con el problema de la categorización de los fenómenos y cosas. Wittgenstein, aunque también Kant y otros, fue quien mejor puso de relieve la naturaleza problemática de las categorías cuando señaló que diferentes juegos no tienen nada en común salvo el hecho de ser juegos. Por ejemplo, un niño jugando al tejo, Fisher y Spassky, jugando a ajedrez, alguien intentando resolver el 15-Puzzle o unos marineros ingleses realizando ejercicios navales todos tienen en común el hecho de estar jugando un juego. Para la mayoría de observadores, estos fenómenos parecen tener poco o nada que ver entre sí y, sin embargo,

todos ellos son miembros del conjunto de juegos posibles. Esto define lo que técnicamente se conoce como “conjuntos polimorfos”; es decir, conjuntos no definidos por condiciones ni necesarias ni suficientes. Edelman, da varios ejemplos de estos conjuntos en su libro “Neural Darwinism”. Y es justo la diversidad polimorfa del cerebro la que le permite responder a la diversidad polimorfa de la naturaleza y el mundo fenomenológico. Esta diversidad del cerebro no sólo no es ruido irrelevante, sino que, para Edelman, es *la base misma sobre la que va a realizarse la selección cuando encuentre un conjunto desconocido de correspondencias físicas en el mundo*. Y, descartando la neurona, por su inflexibilidad binaria, propone “grupos” de neuronas vinculadas, o en interacción, como “unidad” de selección. Dichos grupos compiten entre sí en un esfuerzo por crear representaciones eficaces, o “mapas”, de la variedad infinita de estímulos procedentes del mundo. Los grupos que forman mapas con éxito se refuerzan, mientras los demás se debilitan. Estos grupos de neuronas encadenadas resuelven el problema de los conjuntos “polimorfos” mediante el concepto de “reentrada”, que para Edelman es la señalización recursiva en curso entre zonas cartografiadas, de modo que se pueden cartografiar mapas mediante, masivas y recíprocas, conexiones paralelas.

Con todo, Edelman fue muy criticado en su propuesta. Unos, como Gunther Stent lo acusaban de opacidad. Otros, como hizo Crick (Crick, 1989) en la reseña que realizó del libro “Neural Darwinism”, fue mucho más duro y acusó a Edelman de ocultar ideas “presentables”, pero no terriblemente originales, detrás de una *cortina de humo a base de jerga*; y agregó Crick: *“la terminología darwiniana de Edelman tiene menos que ver con cualquier analogía real con la evolución darwiniana que con la grandiosidad retórica”*. Por su parte el filósofo Daniel Dennett en una reseña que hizo al libro de Edelman “Bright Air, Brilliant Fire” (Dennett, 1992), sostenía que éste no había hecho más que presentar versiones bastante burdas de ideas viejas. Y afirmaba que, a pesar de las negativas de Edelman, su modelo era una red neuronal, y la reentrada era simple retroalimentación. Para rematar la cuestión, Dennett terminaba diciendo que: *Edelman también entiende equivocadamente las cuestiones filosóficas que aborda a un nivel, por cierto, bastante elemental*. Finalmente, otros, más críticos si cabe, acusaban a Edelman de tratar deliberadamente de llevarse el mérito de las ideas de otros envolviéndolas con su idiosincrásica terminología. Para Horgan (Horgan, 1998), Edelman es un practicante de la neurociencia irónica, aunque desgraciadamente carece de los datos teóricos requeridos,

quedándose, por decirlo de modo caritativo, en alguien con el cerebro de un empírico y el corazón de un romántico.

En su último libro, Edelman y Tononi (Edelman, 2002), pretenden dar respuesta a las siguientes cuestiones:

1. ¿Cómo surge la conciencia como resultado de procesos neuronales particulares y de las interacciones entre el cerebro, el cuerpo y el mundo?
2. ¿Cómo pueden explicar estos procesos neuronales las propiedades esenciales de la experiencia consciente? Cada uno de los estudios conscientes es unitario e indivisible, pero al mismo tiempo cada persona puede elegir entre un número ingente de estados conscientes distintos.
3. ¿Cómo se pueden entender los distintos estados subjetivos, los llamados “qualia”, en términos neuronales?
4. ¿De qué manera puede ayudar la comprensión que los humanos tienen de la conciencia a enlazar las descripciones estrictamente científicas con el dominio más amplio del conocimiento y la experiencia humanos?

Y establecen la posición de que la conciencia, en toda su plenitud, no surge únicamente del cerebro, sino que las funciones superiores del cerebro precisan interactuar con el mundo y con otros cerebros. Es decir, la conciencia surge como resultado de un tipo particular de proceso cerebral que es, simultáneamente, altamente unificado o integrado y altamente complejo o diferenciado. Las cuestiones a resolver son, pues, entre otras, las siguientes:

¿Por qué razón la simple diferenciación entre luz y oscuridad, cuando la realiza un humano, está asociada con la conciencia, y de hecho la precisa, pero no así, supuestamente, cuando la realiza un aparato? O ¿por qué razón un termistor, sensor físico que puede diferenciar entre el frío y el calor, es incapaz, que se sepa, de generar ninguna cualidad subjetiva o fenoménica, mientras que cuando un ser humano realiza la misma función es consciente del frío, del calor, e incluso del dolor? En términos neuronales, este problema adquiere una cualidad todavía más interesante y paradójica; a saber: ¿Por qué razón la actividad de ciertas neuronas y, o, células gliales, está privada de esa extraordinaria propiedad? Por ejemplo, ¿Cuál es la razón por la que la actividad de las

neuronas de la retina que discriminan entre la luz y la oscuridad no se encuentra directamente relacionada con la experiencia consciente, en tanto que la actividad de ciertas neuronas del sistema visual del cerebro aparentemente sí lo está? O ¿por qué el ser humano es consciente del calor y del frío, pero no es directamente consciente de si la presión sanguínea es alta o baja? Al fin y al cabo, el ser humano dispone de intrincados circuitos neuronales que se ocupan de la presión sanguínea, del mismo modo que se dispone de otros que se ocupan de la regulación de la temperatura general. En términos más generales, ¿Por qué razón el simple lugar que ocupan en el cerebro o la posesión de una característica particular anatómica, fisiológica o bioquímica otorgan a alguna(s) neurona(s) tan alto privilegio que repentinamente infunden en el dueño de ese cerebro el regusto de la experiencia subjetiva, con esas esquivas propiedades que los filósofos denominan “qualia”? Éste es el problema central de la experiencia consciente.

Es necesario en este punto hacer una advertencia metodológica. Es un error esperar que baste con localizar de forma precisa los lugares de actividad del cerebro o entender las propiedades intrínsecas de ciertas neuronas determinadas para explicar por qué la actividad en cuestión contribuye o no a la experiencia consciente. Tal expectativa es un ejemplo perfecto de lo que, como indicó Ryle (Ryle, 1949), es un error categórico, en el sentido específico de atribuir a una cosa propiedades que no puede poseer.

Los perceptos sensoriales; esto es, los constituyentes paradigmáticos de la experiencia consciente, se presentan en varias modalidades; a saber: vista, oído, olfato, gusto, tacto, propiocepción, o percepción del propio cuerpo, quinesia o sensación de la posición corporal, placer y dolor. Además, cada una de estas modalidades comprende muchas submodalidades distintas. Por ejemplo, la experiencia visual incluye: el color, la forma, el movimiento, la profundidad, etcétera. Otros estados conscientes vienen dados por: imágenes, pensamientos, emociones, mismidad, familiaridad, voluntariedad, discurso interior, etcétera.

Al igual que Hilbert y Gödel, su compañero en Princeton, John Archibald Wheeler (Wheeler, 1990) (Wheeler, 1994) creía que se podría comprender y que se comprendería. Gödel creía, falsamente claro, que tal vez entre los papeles de Leibniz estaría la clave de esta pregunta definitiva. No deja de ser, no obstante, irónico que las ideas de Wheeler sugieran que una teoría final será siempre un espejismo, que en cierto sentido la verdad en

vez de aprehenderla objetivamente, uno se la imagina. Según el “it from bit” los seres humanos crean no sólo la verdad, sino también la propia realidad, el “it”, con las preguntas que formulan. Esta visión se aproxima peligrosamente al relativismo o aún a algo peor. Pues como él dijo: *Acepto al cien por cien la idea de que el mundo es un mero producto de la imaginación* (Bernstein, 1985). Y aunque Wheeler es plenamente consciente de que esa opinión es indefendible desde un punto de vista empírico: ¿Dónde estaba la mente cuando empezó a existir el Universo? ¿Qué mantuvo al Universo antes de existir las mentes? Para responder a ello, Wheeler ofrece una paradoja bonita y escalofriante: *En el corazón de toda cosa late una pregunta no una respuesta.*

Desde la perspectiva de la ciencia de la mente; es decir, la psicología, también hubo varios intentos de situar la conciencia dentro de un marco teórico aceptable, pero todos ellos resultaron problemáticos. Así, por ejemplo, la tradición introspeccionista de Kulpe y Titchener (Kulpe, 1909) representa la contrapartida en el campo de la psicología de las posiciones idealistas o fenomenológicas en el ámbito de la filosofía, que pretende describir la conciencia vista por el individuo exclusivamente desde su interior.

Otros, los psicólogos atomistas, postulan que la conciencia estaba formada por partes elementales que podían catalogarse. Por su parte, los conductistas, simplemente intentaban desterrar completamente la conciencia del discurso científico. Sin embargo, los psicólogos cognoscitivos han vuelto a legitimar los conceptos de conciencia y mente, concibiendo la conciencia como un módulo especial o un estadio en el diagrama de flujo que describe la jerarquía del procesamiento de la información. De hecho, los psicólogos cognoscitivos suelen interpretar la conciencia como un cuello de botella de limitada capacidad en el funcionamiento mental causado posiblemente por una limitación no especificada del cerebro. Se han formulado varios modelos de esta guisa de las funciones asociadas a la conciencia, usando técnicas de IA o metáforas tomadas, más en general, de la informática como la de un sistema operativo. Otros psicólogos cognoscitivos, por otra parte, utilizan otra metáfora de la conciencia, en este caso como un escenario unificado, un teatro en el que la información, que procede de diferentes fuentes, es integrada para el control de la conducta (Baars, 1997). Si bien alguna de estas intuiciones pudiera apuntar en la dirección correcta, otras son tan engañosas como atractivas. Lo que, en cualquier caso, es evidente es que esas metáforas no deben suplantar una comprensión científica de la conciencia. En este sentido, los modelos cognoscitivos, hasta ahora, suelen tener poco

o nada que ofrecer con relación al lado experimental o fenomenológico de la experiencia consciente. Desde la perspectiva de estos modelos, sería perfectamente posible prescindir de la conciencia como experiencia fenoménica, y a menudo emocional, siempre que hubiera forma de llevar a cabo sus presuntas funciones, como: el control, la coordinación o la multiplicación. Las interpretaciones cognoscitivas normales no ofrecen ninguna explicación convincente de por qué, en general, la multiplicación realizada por una persona es un proceso consciente dubitativo y lento, mientras que ejecutada por un computador supuestamente no es un proceso consciente. Además, los enfoques estrictamente funcionalistas de la conciencia, centrados en el procesamiento de la información, tienen poco que decir acerca del hecho de que la conciencia precisa de la actividad de sustratos neuronales específicos.

Wilhelm Wundt, el “padre” de la psicología experimental, utilizó el método de la introspección para ahondar en la mente de los seres vivos. Wundt (Wundt, 1922) se basaba en la suposición de que debe existir un conjunto de sensaciones mentales básicas que combinadas de distintas formas, crean el abanico completo del pensamiento y sentimientos. Su objetivo era elaborar el equivalente de la tabla periódica de los elementos químicos, poniendo orden en la serie de “elementos” mentales que subyacen a los estados mentales complejos que conforman la experiencia consciente unificada. En una reacción pendular súbita, después del fracaso de Wundt y seguidores, la psicología se fue al extremo opuesto y abrazó el conductismo (Watson, 1913) (Watson, 1924) (Skinner, 1938), doctrina que sostenía que el comportamiento externamente observable es la única medida verdaderamente científica de lo que ocurre en la mente.

Roger Penrose, premio Wolf de física, y ya citado, es otro científico relevante que incursionó en el problema de la mente. Penrose sugiere que las estructuras llamadas “túbulos”, que se encuentran en las neuronas realizan una modalidad de cálculo denominada “computación cuántica” manejando “q-bits”. Al contrario de los bits, los “q-bits”, se caracterizan por tener en cuenta simultáneamente todas las combinaciones posibles. Es este sentido, se trataría de una forma extrema de proceso paralelo. Penrose era científicamente conocido, como físico-matemático, por sus trabajos en teoría clásica de relatividad general. Con Stephen Hawking estableció muchos de los teoremas fundamentales sobre el comportamiento general de las ecuaciones de Einstein, como los teoremas que permiten demostrar, sin resolver de hecho ecuaciones para un colapso no

esférico, que bajo ciertas condiciones es inevitable la formación de un agujero negro. Teoremas que se aplican a la cuestión estrechamente relacionada de la singularidad inicial del Universo. De hecho, para Lee Smolin, otro físico teórico, Penrose es el físico más importante que ha trabajado en teoría de la relatividad aparte de Einstein. Penrose, sin embargo, se hizo famoso para los no especialistas en relatividad, por dos libros de gran éxito editorial. El primero titulado: “La Nueva Mente del Emperador” (Penrose, 1989), uno de cuyos temas centrales era afirmar que mediante el uso de la consciencia el ser humano está capacitado para ejecutar acciones que están más allá de cualquier tipo de actividad computacional.

Como lo señaló Penrose (Penrose, 1994), los seres humanos, es decir, sus cuerpos y mentes, forman parte de un universo que obedece, con extraordinaria precisión, leyes matemáticas de una extraordinaria sutileza y un amplio alcance. El hecho de que los cuerpos físicos están limitados de forma precisa por dichas leyes es algo aceptado por el moderno punto de vista científico. La cuestión es ¿qué pasa con las mentes? Muchos encontrarán profundamente inquietante la sugerencia de que también las mentes están obligadas a actuar según esas mismas leyes matemáticas. Pero tener que trazar una división clara entre cuerpo y mente estando el primero sujeto a las leyes matemáticas de la física y permitiéndole a la segunda su propio tipo de libertad, no dejaría de ser también perturbador. En efecto, la mente influye con toda seguridad en el modo en que actúa el cuerpo y también debe estar influida por el estado físico del cuerpo. El propio concepto de mente no tendría mucha justificación si la mente no fuera capaz de influir sobre el cuerpo ni de ser influida por él. Además, si la mente fuera meramente un “epifenómeno”, esto es, alguna característica específica pero pasiva, del estado físico del cerebro, que es un subproducto del cuerpo pero no puede reaccionar sobre él, esto podría permitir a la mente tan sólo un papel impotente y frustrado. Pero si la mente fuese capaz de influir en el cuerpo de un modo que dé lugar a que un cuerpo actúe fuera de las limitaciones de las leyes físicas, entonces esto destruiría la precisión de estas leyes científicas puramente físicas. Resulta así difícil de defender la idea, completamente “dualista”, de que la mente y el cuerpo obedecen a tipos de leyes totalmente independientes. Incluso si esas leyes físicas que gobiernan la actuación del cuerpo permiten una libertad dentro de la cual la mente puede afectar constantemente a su comportamiento, la propia naturaleza concreta de esa libertad debe ser un ingrediente importante de esas mismas leyes físicas. Sea lo que

sea, lo que controla o describe la mente, debe ser realmente una parte integral del mismo gran esquema que rige también todos los atributos “materiales” del Universo.

Para Penrose, la conciencia se asienta bien en la formación reticular del “*upper brain stem*” que está activa justamente cuando está consciente y parece ser el responsable de las alertas mentales. O en el hipocampo responsable de la memoria a largo plazo que juega un papel fundamental en la “*conscious awareness*”.

El problema mente-cuerpo tiene dos subproblemas. Uno, ¿cómo un cuerpo material tal como un cerebro evoca la conciencia? Dos, recíprocamente ¿cómo puede una entidad abstracta tal como una mente determinar acciones físicas? Otra cuestión importante es saber si la conciencia confiere una ventaja selectiva a quienes la poseen. Parece que sí. La conciencia es vista como una componente no algorítmica de la actividad mental que permite a los organismos que la poseen dar respuestas no automáticas requiriendo pensamiento, y a veces creatividad y perspicacia. La conciencia es, pues, una propiedad operativa, observable, de ciertos organismos excepto en situaciones patológicas, o administración de drogas, que preservan la conciencia pero inhiben las repuestas motoras que son la manifestaciones externas del comportamiento consciente.

El segundo, titulado: “Las Sombras de la Mente” (Penrose, 1994), puede considerarse una secuela del primero. En él, establece cuatro puntos de vista ¿extremos? acerca de la computación y el pensamiento consciente siguientes:

A) Fisicalismo. Todo pensamiento es computación; en particular, las sensaciones de conocimiento consciente son provocadas simplemente por la ejecución de computaciones apropiadas. Éste es el **enfoque de la IA fuerte**. Este punto de vista es el que se conoce, como interpretación fuerte, dura o científica de la IA o incluso funcionalismo, según el cual la mente puede entenderse exclusivamente en términos de proceso o tratamiento de información. Algunos consideran que éste es el único punto de vista científicamente admisible.

Los integrantes de esta tendencia tratan de responder al problema de la sobredeterminación causal, sin renunciar a la clausura causal del mundo físico, defendiendo que las propiedades causales que aparecen en una explicación mental de

la conducta humana son idénticas a las que aparecen en una explicación fisiológica de la misma. No habría, por lo tanto, dos explicaciones independientes del mismo fenómeno, sino que la primera es la misma que la segunda vista desde un grado de abstracción mayor y descrita mediante el lenguaje que lo refleja. Por ejemplo, la palabra “deseo” utilizada por los mentalistas tiene un grado de abstracción mayor, ya que, como se comentó anteriormente, involucraría excitación sináptica, luego segregación hormonal, acción de los ejecutores locomotores, etc. A su vez, la utilización de un lenguaje de más alto nivel de abstracción parece adecuado para la explicación de los complejos fenómenos mentales haciendo abstracción de los procesos físicos que los sustentan. Con esto se estaría evitando la sobredimensión causal.

En las primeras versiones de las teorías de la identidad, de finales de los años cincuenta, U.T. Place y J.J.C. Smart defendieron una sencilla identidad entre tipos de estados mentales y tipos de estados neurofisiológicos; es decir, entre el dolor y un estado neurofisiológico determinado. Posteriormente, a finales de los setenta, David Armstrong y David Lewis (Lewis, 1966) introdujeron una variación no en la relación de identidad sino en la concepción de los estados mentales. Según éstos, los estados mentales se caracterizan por sus relaciones causales con “entradas” perceptivas y “salidas” conductuales y otros estados mentales intermedios entre ambos.

Esto ha sido asumido por los modelos de redes neuronales multi-capa donde se entiende que en las capas intermedias tienen lugar los procesos de abstracción y decisión. Aunque en este caso no se refiere a problemas puntuales sino a estados mentales lo cual involucra la emoción, la voluntad y, fundamentalmente, la conciencia, los mismos modelos podrán ser plausibles de ser utilizados a un nivel “macro” en la arquitectura neuronal. Esto podría dar lugar a un esquema “composicional” de redes neuronales más que meramente modulares, donde unos subsistemas son, a su vez, elementos de redes cuyas entradas y salidas son de un nivel de abstracción mayor.

Los tipos de estados mentales así concebidos son los que serían idénticos a tipos de estados neurofisiológicos. A estas teorías se las etiqueta como “reduccionismo physicalista”, puesto que reducen lo mental a lo físico o, más precisamente, a lo

neurofisiológico. Este enfoque presenta una grave objeción: la “realizabilidad” múltiple de los estados mentales entre los distintos seres. Ésta se sustenta en una doble base. La intuición de que seres neurofisiológicos distintos entre sí pueden tener los mismos tipos de estados mentales y, por otra parte, el hecho de la plasticidad del cerebro; es decir, la posibilidad de que una parte del cerebro realice funciones que antes no le correspondían es de gran importancia en este asunto. Con ello, se apunta la idea de que la relación entre lo mental y lo neurofisiológico no es de identidad: un tipo mental-un tipo neurofisiológico, sino de realización; esto es, los estados mentales se realizan o implementan neurofisiológicamente de manera múltiple o variable. Técnicamente, en vez de una función biyectiva entre estados mentales y estados neurofisiológicos, lo que se propone es una función sobreyectiva entre estados neurofisiológicos y estados mentales, lo que implica la negación del reduccionismo fisicalista.

Existen varias versiones del punto de vista cartesiano (Sloman, 1992). Alan Mathison Turing (Turing, 1950) fue el primero en formular la hipótesis de que la actividad de la mente humana se puede comparar al software de una computadora. Posteriormente, los creadores de la IA llegaron a la convicción de que el pensamiento humano y su inteligencia eran un programa software que podría aplicarse y hacerlos funcionar sobre cualquier estructura hardware. Sea de silicio, atómico o proteínico. Es decir, consideraban que se podría proyectar una mente para una máquina. Los partidarios de la IA fuerte afirman que la mente es independiente del cerebro: es un programa. Turing afirmaba que no hay mente separada de la materia. Gödel (Gödel, 1990) consideraba evidente que el cerebro físico propiamente dicho se comportaba computacionalmente, pero que la mente es algo que trasciende al cerebro, de modo que la acción de la mente no está limitada a comportarse de acuerdo con las leyes computacionales que él creía debían controlar el comportamiento del cerebro físico. Otros, como Edelman (Edelman, 1992), aunque partidarios de este enfoque, negarían enfáticamente ser partidarios de la IA fuerte, porque dicen adoptar una visión diferente de la IA convencional en lo referente a la interpretación del término computación.

- B) El conocimiento es un aspecto de la acción física del cerebro; y si bien cualquier acción física puede ser simulada computacionalmente, dicha simulación no puede

provocar conocimiento por sí misma. Este enfoque, que se conoce como **IA débil, blanda o tecnológica**, admite que es posible simular la acción del cerebro, pero ésta no tendría efectos mentales. Este punto de vista, lo defiende, entre otros, John R. Searle (Searle, 1985) (Searle, 1990) (Searle, 1998).

- C) La acción física apropiada del cerebro provoca conocimiento, pero ésta nunca puede ser simulada adecuadamente de forma computacional. Este enfoque, que es el de Penrose, establece que **la actividad consciente no puede ni siquiera simularse**.
- D) Mentalismo. El conocimiento no puede explicarse en términos físicos, computacionales o cualesquiera otros términos científicos. Este punto de vista que niega por completo la posición fisicalista y considera la mente como algo completamente inexplicable en términos científicos, es el **punto de vista de la mística, y los misterianos**.

Otros neurocientíficos cuánticos son el premio Nobel de Física de 1973, el galés Brian David Josephson, por sus trabajos sobre el efecto túnel y la supercomputividad que dio origen a un dispositivo, “la unión Josephson” que se usa en supercomputación con superconductores. Josephson, en el simposium organizado en Tucson, en 1994, por el anesthesiólogo Stuart Hameroff, cuya obra sobre microtúbulos inspiró a Penrose sobre los aspectos cuánticos de la consciencia, en la Universidad de Arizona, sugirió que los efectos cuánticos pueden explicar fenómenos psíquicos y hasta místicos. También participaron en dicho encuentro, el matemático André Weil, toda una autoridad en cuestiones psicodélicas, según el cual una teoría completa de la mente debe tener en cuenta la capacidad de los indios sudamericanos que han consumido drogas psicotrópicas para experimentar alucinaciones compartidas. Y Danah Zohar, autor “New Age”, según el cual el pensamiento humano es fruto de *fluctuaciones cuánticas de la energía vacía del universo que es realmente Dios* (Horgan, 1994).

Para, Rita Levi Montalcini (Levi, 2000), la mayoría de los científicos consideran que los hechos de naturaleza físicoquímica que caracterizan las actividades nerviosas no tienen nada que ver con la actividad mental del individuo ni están controlados por ella.

Otros, como Roger Sperry (Sperry, 1964), James y Thorpe piensan que la conciencia es una propiedad que **emerge** de los procesos de los circuitos nerviosos. La conciencia, que los neurofisiólogos y los estudiosos del comportamiento consideran un aspecto “interno” de los procesos cerebrales, un epifenómeno o parafenómeno o aún un producto colateral, cuando no un artefacto de la semántica o un pseudoproblema, según Sperry es una parte integrante de los procesos cerebrales.

El neurobiólogo y premio Nobel Eric Kandel (Kandel, 1998) describe la mente humana como el último gran misterio, y considera que es el gran reto de la biología del tercer milenio y añade: *lo mismo que la cosmología indaga la estructura del Universo, las neurociencias cognitivas indagan la estructura de la mente. El problema que el hombre lleva siglos planteándose, la relación entre cerebro y mente, cuenta hoy para su resolución, no sólo con los conocimientos neurológicos clásicos, sino también con la aportación de la informática capaz de simular las funciones del órgano cerebral, sistemas informáticos con atributos que hasta hace poco se consideraban exclusivos de la mente humana.* Como lo señaló Kandel:

1. La principal cuestión acerca del cerebro es la naturaleza de la conciencia y como varios procesos psicológicos inconscientes se relacionan con el pensamiento consciente.
2. La teoría estructural de la mente de Freud, por cierto algo más que inspirada en Platón y sus dos caballos y el auriga, del Yo, el ello y el superyo, le llevó a considerar ¿cómo los procesos conscientes e inconscientes difieren en su representación en el cerebro?
3. Establecer una definición de trabajo de la conciencia como un estado de percepción consciente, o atención selectiva “writ large”. En esencia, la conciencia en la gente, es una conciencia del Yo, una conciencia de ser consciente. La conciencia se refiere a la capacidad humana no simplemente por experimentar placer o dolores, sino para atender y reflejar dichas experiencias y hacerlo en el contexto de la vida cotidiana y la historia de la vida.
4. Entender la conciencia es, con mucho, la tarea más desafiante que afronta la ciencia. La carrera científica de Francis Crick, el biólogo más creativo e influyente de la segunda mitad del siglo XX es, como ya se ha dicho, prueba de ello.

Los neurocientíficos, que suelen mostrarse extremadamente cautelosos en sus aproximaciones al problema de la conciencia, la mayoría se muestra convenientemente agnóstica acerca de esta cuestión, justificando esta postura por la ignorancia actual sobre este asunto. Y, aunque la mayoría probablemente suscribiría una explicación sistémica, por el momento considera más fructífero dedicarse a la recolección de datos y observaciones y dejar la teoría para el futuro.

2.3. Reduccionismo versus Emergencia

Son estos dos conceptos, conflictivos y polémicos, los que han enfrentado durante décadas, más bien siglos, a científicos y filósofos por igual. La cuestión que se plantea es doble. Por una parte, sería importante saber si estos dos conceptos son contradictorios o, simplemente, complementarios. Por otra, si son contradictorios, ver si es posible establecer una enantiomorfía que, “superándolos”, los englobe. Para dar cuenta de este planteamiento, en primer término, se va a intentar definir ambos conceptos.

2.3.1. Reduccionismo

El primer intento “reduccionista” procede de Tales quien afirmaba que todo estaba hecho con agua. Esto es, había llegado a la idea de que podría ser posible explicar una gran cantidad de cosas complicadas sobre la base de algún principio simple y universal. Posteriormente, esta idea se volvió a dar en Leucipo y Demócrito. Sin embargo, sólo con tales ideas no podían haber progresado mucho, como así fue, por la sencilla razón de que tales ideas eran puramente cualitativas.

Bajo el nombre de “*reduccionismo*” se engloban muchos conceptos y visiones diferentes. Para empezar, el primer “reduccionista” “*sensu lato*” fue, ¿quién sino?, Isaac Newton. Él fue el primero en mostrar la posibilidad de una comprensión de la naturaleza que fuera al mismo tiempo global y cuantitativa. Otros, antes que él de Tales a Descartes, pasando por Ockham y su navaja, propugnaban hacer afirmaciones globales, con el mínimo de principios o leyes sobre la naturaleza, pero ninguno de ellos aceptó el desafío de explicar cuantitativamente las observaciones reales dentro de una teoría física global. Así, en el prefacio a la primera edición de los “*Principia*”, Newton dice (Newton, 1687-1987): “*Me gustaría que pudiéramos derivar el resto de los fenómenos de la naturaleza mediante el*

mismo tipo de razonamiento a partir de principios mecánicos, porque muchas razones me inducen a sospechar que todos ellos pueden depender que ciertas fuerzas". Según Steven Weinberg (Weinberg, 1974), el ejemplo más radical de la pretensión de Newton es el que aparece en el tercer libro de los "*Principia*", donde razona que la Luna estará sesenta veces más lejos del centro de la Tierra que Cambridge. Y, por lo tanto, la aceleración de la Luna hacia la Tierra debe ser menor que la aceleración de una manzana en Cambridge por un factor de sesenta al cuadrado; es decir, 3600. Con este argumento, Newton unifica la mecánica celeste y las observaciones sobre la caída de la fruta de un modo que capta, por primera vez, el enorme potencial del razonamiento matemático no sólo para explicar sistemas idealizados como los planetas que se mueven en sus órbitas, sino finalmente todo.

Este deseo de Newton fue, posteriormente, escenificado por Einstein cuando dijo: "La prueba suprema de un físico es llegar a esas leyes elementales y universales desde las que el cosmos puede ser construido a partir de la pura deducción". Pensamiento que comparten todos los reduccionistas y en especial Steven Weinberg (Weinberg, 1974) cuando dijo: "*Una de las esperanzas más persistentes en el hombre ha sido encontrar unas pocas leyes generales que explicaran porque la naturaleza con toda su aparente complejidad y variedad es del modo que es. En el momento presente lo más cerca que podemos llegar a una visión unificada de la naturaleza es una descripción en términos de partículas elementales y sus interacciones mutuas*".

Ernst Mayr (Mayr, 1985), comentó lo que dijo Weinberg de la siguiente guisa: "*un ejemplo horrible del modo en que piensan los físicos*", y lo tachó de "*reduccionista intransigente*". Mayr es, con Anderson y James Krumhanse, el principal oponente de la tendencia reduccionista. Su libro de 1982 (Mayr, 1982), contiene un bien conocido ataque al reduccionismo. En él habla de dos clases de reduccionismo. El primero, que denomina "reduccionismo teórico", es aquel que propugna, afirma o simplemente desea que, finalmente, todas las ciencias perderán su autonomía y serán todas ellas absorbidas por la física de partículas elementales. Es decir, todas se verán, lisa y llanamente, como ramas de la física de partículas elementales. La segunda clase de reduccionismo la denominó Mayr "reduccionismo explicativo". Aquí la idea es que el progreso en el nivel menor, verbigracia en el nivel de la física de partículas elementales, es necesario para progresar en otras ciencias, como la hidrodinámica, la biología, etcétera.

Hay que distinguir entre gran reduccionismo; es decir, de la visión de que toda la naturaleza es como es, con ciertas condiciones sobre las condiciones iniciales y los accidentes históricos, a causa de leyes universales simples, a las que todas las demás leyes deben ser reducidas, y el reduccionismo menor. Es éste una doctrina que se basa en que las cosas se comportan del modo en que lo hacen a causa de las propiedades de sus constituyentes; por ejemplo, un diamante es duro porque los átomos de carbono de los que se compone se unen unos con otros de manera precisa. Ambos reduccionismos se confunden a menudo porque parte del progreso reductivo en la ciencia se ha dedicado a responder preguntas sobre la composición de las cosas. El gran reduccionismo y el reduccionismo menor se corresponde, más o menos, con lo que Ernst Mayr denominó “reduccionismo teórico” y “explicación reduccionista” (Mayr 1987).

Steven Weinberg (Weinberg, 1974), añade una tercera clase de reduccionismo que llamó “reduccionismo objetivo” y que, en forma resumida, explicó como sigue: Apoyándose en la idea intuitiva de que unas generalizaciones científicas diferentes explican otras, se tiene la sensación de dirección en la ciencia. Es decir, hay flechas de explicación científica, y estas flechas pasan por el espacio de todas las generalizaciones científicas. Tras haber descubierto muchas de esas flechas, es posible mirar el patrón que ha aparecido y percatarse de algo digno de atención y quizás el mayor descubrimiento científico de todos: ¡estas flechas parecen converger en una fuente común! Empiecen en cualquier parte de la ciencia y, como un niño maleducado, no dejan de preguntar “¿por qué?”, al final se desciende siempre hasta el nivel de lo más pequeño.

En la presentación que Weinberg hizo de su artículo “*Reflexiones sobre la física del siglo XX*” (Weinberg, 1995), publicado en 1995 con otros artículos en tres tomos con el título: “*Twenty Century Physics*”, dice que Anderson, otro premio Nobel, presentó en ese mismo trabajo un argumento contra el reduccionismo basada en un fenómeno conocido como **emergencia**, que Anderson identificaba como el “*principio de Dios*” de la física del siglo XX. Y continuaba: “*Cuando ciertos sistemas alcanzan un nivel suficiente de complejidad **emergen** nuevos fenómenos que no existen para los constituyentes elementales de los que se compone el sistema. Por ejemplo, la superconductividad aparece en sólidos compuestos de grandes cantidades de ciertos átomos, pero la superconductividad carece de significado para los átomos individuales. De forma similar, la vida surge de la bioquímica, la bioquímica surge de la física atómica, y la*

*física atómica surge de las propiedades de las partículas elementales, tal y como se describen en el moderno modelo estándar. El fenómeno de **emergencia** libera hasta cierto punto a los fenómenos a gran escala de los detalles de lo que ocurre a escalas pequeñas. Por ejemplo, aunque usted tenga que saber mucho sobre los átomos individuales para saber si una sustancia es un superconductor, los rasgos generales de la superconductividad son universales, no dependientes de las específicas de los átomos que componen el superconductor. Anderson acierta en que el progreso ulterior de la física de las partículas elementales es improbable que sea de mucha ayuda inmediata en la comprensión de los fenómenos que **emergen** como la superconducción, por no mencionar la vida. Y continuaba Weinberg: No creo que nada de esto constituya un argumento convincente en contra de las tendencias reduccionistas de la física de partículas elementales. Después de todo, los fenómenos que **emergen, emergen** finalmente desde la física de partículas elementales, y si usted quiere entender por qué el mundo es como es tiene que comprender por qué las partículas elementales son como son. Esto no puede ser más importante que otros tipos de investigación física pero tiene una “clase” especial de importancia que nos ha llevado algunos de nosotros a dedicarle el trabajo de nuestra vida.*

Nada hay que añadir a esto, salvo señalar que, como a confesión de parte, Weinberg es un reduccionista estricto, hay relevo de pruebas, **la emergencia** es indiscutible.

El gran rompecabezas, sobresale por encima de las leyes de la naturaleza y de las condiciones iniciales. Afecta al papel dual de la vida inteligente como parte del Universo que se busca explicar, y cómo quién lleva a cabo la explicación.

Se puede ilustrar la visión reduccionista del mundo imaginando que todos los principios de la ciencia son puntos en un gran mapa, con flechas que fluyen hacia cada principio desde todos los demás principios mediante los cuales se explica. La lección de la historia es que estas flechas no forman grupos distintos e inconexos, que representan ciencias lógicamente independientes y que no vagan sin objetivo. Más bien están todas conectadas y, si se rastrean en dirección a su origen, todas parecen desgajarse de una fuente común, una ley final de la naturaleza que Dyson llama “conjunto finito de ecuaciones fundamentales”. En efecto, Freeman Dyson (Dyson, 1995), describió el reduccionismo como el esfuerzo para reducir el mundo de los fenómenos físicos a un grupo reducido de

ecuaciones fundamentales. Si se cambia ecuaciones por principios la “*definición*” parece perfecta.

Se dice que un concepto está en un nivel más elevado o más profundo que otro si está gobernado por principios que están más allá de y, o, más cerca de, esa fuente común. De este modo, el reduccionista considera que las teorías generales que gobiernan el aire, el agua y la radiación están a un nivel más profundo que las teorías sobre frentes fríos o tormentas, no en el sentido de que sean más útiles, si no sólo en el sentido de que las últimas se pueden comprender, en principio, como consecuencias matemáticas de las primeras. El programa reduccionista de la física es la búsqueda de la fuente común de todas las explicaciones.

Ahora bien, continúa Weinberg, “*tal como lo he descrito en términos de convergencia de flechas de explicación, el reduccionismo no es un hecho sobre los programas científicos, sino un hecho sobre la naturaleza*”. Para Mayr el “*reduccionismo objetivo*” es lo que él denomina “*reduccionismo teórico*”. Weinberg, sin embargo, prefiere diferenciarlos pues afirman que “*es un orden inherente a la propia naturaleza*”. Y aquí hay que criticar a Weinberg, pues más que un orden inherente a la propia naturaleza, es, como él afirma en su artículo “*La propia naturaleza*”, “*cómo se ve la propia naturaleza*”; es decir, no es una cuestión ontológica, sino epistemológica.

Hay que distinguir el reduccionismo como programa para la investigación científica y el reduccionismo como visión de la naturaleza. Por ejemplo, la visión reduccionista pone énfasis en que el tiempo meteorológico se comporta como lo hace a causa de los principios generales de la aerodinámica, el flujo de radiación, etc., y por accidentes históricos como el tamaño y la órbita de la Tierra. Pero para predecir el tiempo que hará mañana puede que sea más útil pensar en frentes fríos y tormentas, etcétera. El reduccionismo puede ser una buena guía para un programa de pronóstico del tiempo o puede no serlo, pero proporciona la intuición necesaria de que no hay leyes autónomas del tiempo que sean independientes de los principios de la física. Tanto si tenerlos en cuenta ayuda al meteorólogo o no, los frentes fríos son como son a causa de las propiedades del aire, el vapor de agua, etcétera; los que, a su vez, son como son a causa de los principios de la química y la física. No se conocen las leyes finales de la naturaleza, pero se sabe que no se expresa en términos de frentes fríos o tormentas.

Sobre reduccionismo también se posicionó Murray Gell-Mann. Según él (Gell-Mann, 1996), cuando se habla de este asunto, la gente se lanza toda suerte de epítetos. Para Gell-Mann, la única posición sensata al respecto es que las leyes de la física son fundamentales en el sentido de que las demás leyes se construyen sobre ellas. Pero esto no significa que las otras leyes se puedan derivar de las leyes físicas. Y esto es así porque hay que añadir los rasgos especiales que proceden de la historia y que subyacen a las demás ciencias. Y para ejemplificar y fundamentar su aserto, señala que la física y la química sí se derivan de las leyes fundamentales, pero, incluso en este caso, en sus diversas ramificaciones la formulación de las cuestiones apropiadas incluye una buena cantidad de información adicional sobre condiciones particulares que no se dan en cualquier parte del Universo. Por ejemplo, en el centro del Sol no hay física del estado sólido. En el Universo primordial, cuando la materia era principalmente una sopa de quarks, ni siquiera había física nuclear. Así pues, en cierto sentido, también estos temas implican algo más que leyes fundamentales. El resto de las ciencias depende en gran medida de accidentes particulares en la historia del Universo: accidentes astronómicos, geológicos, biológicos y de la historia de la humanidad. Por ejemplo, para obtener los detalles de la biología terrestre hay que añadir a las leyes fundamentales una gran cantidad de información suplementaria. El hecho de que la física de partículas sea fundamental no significa que la biología pueda reducirse a la física de partículas, ni siquiera en principio, a menos que se adjunte esa información adicional. Es más, en la práctica es esencial estudiar la biología en su propio nivel, y lo mismo pasa con la psicología, la sociología, la historia y todo lo demás, porque cada nivel tiene leyes propias. Aunque en principio esas leyes puedan derivarse de las del nivel inferior, más una cantidad de información adicional, la estrategia razonable es construir escaleras tanto ascendentes; explicaciones mecanicistas, como descendentes; leyes empíricas. Todas estas ideas pertenecen a lo que Gell-Mann denomina “la doctrina de la **emergencia**”.

En suma, existe una forma fuerte de reduccionismo que viene expresada por la proposición << x “**no es sino**” y >>. Un ejemplo de este tipo de reduccionismo es el intento del atomismo lógico, de Russell (Russell, 1924) y Wittgenstein (Wittgenstein, 1957), para edificar una epistemología basada en proposiciones menos cuestionables; es decir, aquellas basadas en datos de los sentidos, tales como: “Me parece oír un do mayor en este momento”. “Esto aparenta ser una superficie plana y dura”. De este modo, la proposición “hay una tumbona en el jardín” no sería sino una lista de proposiciones

preestablecida finitamente sobre datos sensoriales. En esta forma de reduccionismo, un concepto de alto nivel queda reducido a un lenguaje de menor nivel si la verdad de un conjunto definido de enunciados en este último lenguaje es necesaria y suficiente para garantizar la verdad de la proposición de alto nivel. En este caso, la proposición de alto nivel es simplemente una abreviatura de la lista de proposiciones necesarias y suficientes en el lenguaje reductor de bajo nivel. El problema es que este intento fracasa para las proposiciones sobre objetos físicos. Aunque sea relativamente fácil hallar suficientes proposiciones basadas en datos sensoriales, resulta imposible especificar de forma finita un conjunto necesario y suficiente cuya verdad sea intercambiable con las de la proposición de orden superior.

Ludwing Wittgenstein, en su “Tractatus” (Wittgenstein, 1957), había llevado al límite la línea seguida por el atomismo lógico de Russell (Russell, 1924) que buscaba unos cimientos epistemológicos firmes para todo el conocimiento, en forma de “proposiciones atómicas” sobre datos provenientes de los sentidos. La idea, acariciada por Hume, Berkeley, Kant y los empiristas británicos, era la siguiente: Uno puede equivocarse al decir que hay una tumbona en el jardín, pero difícilmente errará al enunciar fragmentos de sus propias percepciones; verbigracia, percibo la nota si bemol. Dado que estas proposiciones se limitan a reportar datos sensoriales, son menos susceptibles de error que las relativas al mundo exterior. En suma, el atomismo lógico intentaba construir proposiciones sobre el mundo exterior a partir de combinaciones lógicas de proposiciones atómicas relativas a datos sensoriales. De hecho, como lo señaló Stuart A. Kauffman (Kauffman, 2003), a quien se sigue en este apartado, el “Tractatus” buscaba llevar tal empresa a término... y casi lo logró. Fue el propio Wittgenstein quien, veinte años después, dinamitó la propuesta, dándole el tiro de gracia con sus “Investigaciones Filosóficas” (Wittgenstein, 1988). En ellas, destruyó el concepto de un nivel de descripción privilegiado y abrió el camino a la comprensión de que, a cualquier nivel, los conceptos forman típicamente círculos codefinicionales.

El asunto, bien mirado, resulta ser aún peor, pues hasta en el caso de una tumbona real, la idea de que se puedan construir proposiciones sobre objetos reales, tales como tumbonas, combinando proposiciones sensoriales es errónea. Y el error es trascendente. El programa para reemplazar proposiciones sobre objetos físicos mediante conjuntos de proposiciones sobre datos sensoriales requiere, para Kauffman, “equivalencia lógica”. Esto significa que

dondequiera que aparezca una proposición sobre la tumbona a un alto nivel de descripción, debe ser posible especificar una lista de proposiciones sobre datos sensoriales cuya certeza conjunta sea necesaria y suficiente para que la proposición acerca de la tumbona sea verdadera. El problema estriba en que resulta imposible preestablecer de modo finito conjuntos de proposiciones sobre datos sensoriales, que sean lógicamente equivalentes a proposiciones sobre tumbonas. Para lograrlo habría que ser capaces de preestablecer de modo finito algo sobre un conjunto de proposiciones relativas a proposiciones sobre datos sensoriales cuya verdad fuese necesaria y suficiente para validar una proposición sobre una tumbona ubicada en el jardín. En síntesis, el problema radica en que no existe un conjunto de condiciones finitamente predefinibles sobre proposiciones acerca de datos sensoriales cuya certeza sea lógicamente equivalente a cualquier proposición sobre una tumbona real en un jardín.

Wittgenstein (Wittgenstein, 1988) inventó el concepto de “juego del lenguaje”; es decir, un grupo de conceptos que permite contemplar el mundo de algún modo novedoso. Para ello proponía varios ejemplos. Uno, referente a la actuación de un jurado. Otro referente al Reino Unido entrando en guerra con Alemania. La clave del argumento de Wittgenstein se halla en que no es posible, en general, reducir proposiciones de alto nivel a un conjunto finito y preestablecido de proposiciones necesarias y suficientes de más bajo nivel. En lugar de ello los conceptos a alto nivel están codefinidos. Esto no siempre es un inconveniente, sino todo lo contrario. Muchos conceptos, sobre todo científicos, útiles surgen en forma de conceptos codefinidos. Por ejemplo, tal y como decía Poincaré, la fórmula de Newton que relaciona fuerza con masa y aceleración, es bueno que sea circular, pues ese círculo codefinido es el centro de una red de conceptos más amplia, aplicable al mundo real, tal red abarca, articula, discrimina y categoriza ese mundo. También Wallace (Wallace, 1858) y Darwin (Darwin, 1970) dieron un salto hacia un nuevo círculo definicional con sus conceptos de selección natural, variación heredable y adaptación, creando otro círculo virtuoso.

Asimismo, existe una dificultad sistemática al tratar de reducir un juego de lenguaje a otro. Del mismo modo, las descripciones de actos humanos y, por extensión, las de agentes autónomos, parecen implicar un juego de lenguaje distinto del de las descripciones en términos de nuevos eventos físicos. Entre otras características críticas del juego del lenguaje de “*acciones y hechos*” está el que, comparada a una hipotética

descripción física “*completa*”, la descripción de acciones y hechos destaca los aspectos relevantes con respecto a los objetivos de un agente autónomo. Y, como en otros casos, se encontrarán muchas condiciones suficientes, pero será difícil que a la vez se encuentren condiciones que sean necesarias.

El diccionario, al ser autocontenido, da una pista de por qué las ideas reduccionistas pueden estar equivocadas. Ser autocontenido significa que todas y cada una de las palabras del diccionario están definidas en términos de otras palabras del diccionario. Esto, que permite juegos tan entretenidos como el “*Wish*”, tiene efectos desastrosos para el reduccionismo. Los conceptos se hallan definidos en redes, conectadas de algún modo al mundo real mediante definiciones “*ostensivas*”; esto es, definiciones dadas por medio de ejemplos. Se clasifica el mundo de multitud de maneras, los juegos de lenguajes de Wittgenstein, que parecen no ser deducibles unos a otros en el sentido estricto de las condiciones necesarias y suficientes. Y lo peor es que incluso al nivel de la teoría física más elemental también parece ocurrir lo mismo. Por ejemplo, la termodinámica clásica es una teoría perfectamente definida en sí misma. Ludwig Boltzman, Willard Gibbs y otros, se esforzaron en crear esa mecánica estadística basada en las leyes de Newton operando sobre un conjunto de partículas idealizadas en un espacio de fases 6-dimensional, tres para las coordenadas espaciales y tres para el momento. Se consideró un éxito el que los conceptos termodinámicos clásicos de temperatura, presión y entropía fueran reducidos a propiedades estadísticas de conjuntos ideales de partículas de gas: La temperatura se convirtió en la energía cinética media de las partículas. La presión, en el momento transferido a las paredes del recipiente. La entropía en una medida del número de microestados por macroestado. La cuestión es saber si, o no, la mecánica estadística constituye un conjunto de proposiciones necesarias y suficientes respecto a la termodinámica clásica. La respuesta parece ser negativa. Aunque la mecánica estadística, basada en las fuerzas newtonianas, produce un conjunto de condiciones suficientes, dicha mecánica no alcanza, a la vez, la condición de necesaria. De hecho es posible construir otras mecánicas estadísticas consistentes, basadas en partículas que no respondieran a las leyes newtonianas. Así pues, en el propio campo donde se supone que la reducción ha tenido éxito, parece no existir un conjunto finitamente predefinido de condiciones necesarias y suficientes de bajo nivel relativo al conjunto de mecánicas estadísticas posibles que sea suficiente como reducción de la termodinámica clásica.

Hay una forma más débil de reduccionismo que consiste en la caracterización de un objeto, concepto o fenómeno de alto nivel en términos de un conjunto de condiciones suficientes, aunque no simultáneamente necesarias, de menor nivel. En este sentido, la mecánica estadística seguramente “abarca” la mecánica clásica. Muchos argumentan que esta forma más débil de reduccionismo basta. Muy bien, pero ¿basta para qué? No es fácil ser claro al respecto. Resulta tentador decir: “La temperatura no es sino energía cinética media de los átomos del sistema”. Desde luego con ello se simplifica el mobiliario ¿ontológico? del Universo. Ahora bien, este paso ¿ontológico? habitual en el reduccionismo, no siempre está justificado. Tómense estos dos ejemplos de **emergencia** dados por el especialista en física de la materia condensada, el ya mencionado, Philip Anderson premio Nobel de Física de 1977. El primero, tiene como protagonista al oro, metal amarillo y maleable familiar a todos. Pues bien, en ninguna parte de la descripción que hace la mecánica cuántica del átomo de oro se encuentran esas propiedades macroscópicas. Además, no hay forma de llegar deductivamente a dichas propiedades macroscópicas selectivas a partir de la mecánica cuántica que subyace en dichos átomos. En lugar de ello, se observan tales propiedades, se encuentran pautas bien definidas de ellas y entonces se trata de enlazarlas a condiciones suficientes en la descripción que la mecánica cuántica hace de la materia.

El segundo ejemplo, concierne a la ruptura de simetrías. El caso típico sería el de un mástil que se mantiene vertical sobre un plano “paralelo” a la superficie terrestre. El mástil es inestable respecto a la gravedad y cae al poco tiempo. Podría caer en cualquier dirección, pues el sistema es totalmente simétrico respecto al plano antes de que caiga el mástil. En un momento dado, éste se precipita y queda apuntando en una dirección concreta. Así pues, la simetría inicial del sistema se rompe, pero no es posible deducir de ello la forma en que quedará rota.

Philip Anderson, en un ensayo publicado en Science en 1972 (Anderson, 1972), reflexiona sobre los límites del reduccionismo científico, ante la pretensión de los físicos de partículas de que su investigación era la más fundamental e importante de las investigaciones científicas, todas las demás “cosas” que había en la ciencia eran simples “detalles”, peor aún, simple ingeniería. Anderson reconocía los éxitos extraordinarios del reduccionismo en todos los campos científicos hasta el punto de que era aceptado sin rechistar por la inmensa mayoría de los científicos en activo. Su tesis es que los

mecanismos de nuestras mentes y cuerpos, y de toda la materia animada o inanimada de la que se tiene cualquier conocimiento detallado, se supone que están controlados por la misma serie de leyes fundamentales. Pero, según Anderson, el conocimiento de las leyes básicas que rigen para la física no basta para comprender plenamente todos los fenómenos. Si los físicos de partículas no pueden vaticinar el comportamiento del agua, qué decir de la conducta de los humanos. La realidad tiene una estructura jerárquica, cada uno de cuyos niveles, como lo mostró Hartman, es independiente hasta cierto grado de los niveles superiores e inferiores. Por ello, en cada fase se necesitan conceptos, leyes y generalizaciones distintas y completamente nuevas y se exige un grado de inspiración y creatividad tan grande como en la fase anterior. La psicología no es biología aplicada, ni tampoco la biología es química aplicada. Si hay un ámbito de la naturaleza que haya resultado ser más que la suma de las partes, ése no es otro que la naturaleza humana.

Reflexionando en términos de causalidad, la versión “no es sino” del reduccionismo ha tendido a ver la causalidad actuando sólo hacia arriba, desde el comportamiento de los átomos hasta sus consecuencias causales en el comportamiento de entidades complejas como los mamíferos. Sin embargo, la causalidad hacia abajo también es real y no es mística en absoluto. Hay muchos más neumáticos viejos que viejas ruedas de carro tiradas en los arcones de las carreteras. El automóvil desplazó a la calesa y al carro de bueyes.

Como señaló Hartmann (Hartmann, 1964), los principios del ser y las leyes físicas que rigen lo inorgánico, prevalecen también sin limitaciones en los estratos superiores. Dichos estratos son: inorgánico → orgánico → anímico → intelectual. Éstas fueron sus palabras: *Así pues, la naturaleza orgánica predomina sobre la inorgánica. No fluctúa libremente de por sí sino que presupone las circunstancias y legitimidades de lo material; se asienta en ellas aun cuando esto no basta ni mucho menos para integrar lo viviente. Asimismo el ser anímico y la conciencia están condicionados por el organismo sustentador, y obligados a surgir con él exclusivamente en el mundo [...].*

Lorenz (Lorenz, 1974), afirma que hay especulación metafísica cuando, se transgreden las fronteras, tanto hacia arriba como hacen los reduccionistas, como hacia abajo como hizo Leibniz con el monadismo. Por lo que respecta al reduccionismo, una cosa más: Idealmente, toda ciencia natural comienza con una descripción de los fenómenos sometidos a estudio. De ahí, se pasa a una ordenación de los fenómenos descritos; a

continuación, se llega a la abstracción de los factores legítimos imperantes en ellos; finalmente, mediante la experimentación, se verifican las leyes abstractas obtenidas. Ahora bien, como la física fundó su desarrollo en las dos últimas etapas, los reduccionistas, quieren eliminar de la investigación de la conciencia las dos primeras, lo que es un grave error metodológico. Además, y como lo demuestra el caso de Hellen Keller, que aquí es pertinente (Keller, 1905) por dos razones, las dos primeras fases son, en el estudio de la conciencia, imprescindibles. En primer lugar, porque la conciencia es personal e intransferible; esto es, en cierta medida, irreproducible en los demás. Sin embargo, como es sabido, el fundamento y razón de ser del método experimental se centra sólo en fenómenos reproducibles e ignora por lo tanto, como lo señaló René Dubois (Dubois, 1986), las manifestaciones imprevisibles... De hecho, el experimentador hace lo posible por disimular e incluso eliminar este tipo de manifestaciones... Por otra parte, el caso Keller es un ejemplo fabuloso de la descripción, por Sullivan, de la conciencia de Keller.

Como réplica a los reduccionistas, para quiénes la única causa de organización, yace en las partes, Aristóteles con su causa formal, Hegel con el surgimiento del espíritu en la naturaleza y Bergson (Bergson, 1970) con su simple e irrepetible autoorganización creativa, aseguran que el todo es lo predominante.

2.3.2. Emergencia

Otra posición es el de la **emergencia**. En este caso, se afirma que si bien la conciencia tiene su origen en eventos físicos del cerebro, no se reduce a éstos, sino que **emerge** de ellos de modo análogo a cómo las propiedades del agua **emergen** de la combinación química de dos átomos de hidrógeno y uno de oxígeno, pero no pueden reducirse a la suma de las propiedades individuales del hidrógeno y del oxígeno. Lo mismo sucede con la sal, elemento nutricional formado por dos compuestos: un gas, el cloro, brillante y venenoso, y un sólido, el sodio, metálico. En suma, estos casos de **emergencia** de propiedades son algo normal en sistemas complejos holísticos donde el todo es más que la suma de las partes. Hay otras variaciones a la misma tesis pero, en general, atribuyen a la conciencia un valor marginal, al menos desde un punto de vista de la explicación. En cualquier caso, todas insisten en que no existe una sustancia de la “conciencia” distinta de la sustancia del “cerebro”.

Como lo señaló Daniel Hillis (Hillis, 1988), a veces un sistema con muchos componentes simples exhibirá un comportamiento del todo que parece más organizado que el comportamiento de las partes individuales. El caso más simple son los copos de nieve. Las formas constituidas por el hielo son consecuencia de reglas locales de interacción que gobiernan las moléculas de agua, aunque la conexión entre las formas y las reglas está lejos de ser obvia. Después de todo, son las mismas reglas de interacción las que causan que el agua repentinamente se convierta en vapor en su punto de ebullición y provoque burbujas para formar una corriente. Las reglas que gobiernan las fuerzas entre las moléculas de agua parecen mucho más simples que las de los cristales o burbujas; sin embargo; todos esos fenómenos complejos son consecuencia de esas reglas. Tales fenómenos, se denominan “comportamientos **emergentes**” del sistema.

Sería muy conveniente que la conciencia fuese un comportamiento **emergente** de neuronas conectadas aleatoriamente, en el mismo sentido que en los casos anteriores del agua. Entonces puede ser posible construir una máquina pensante simplemente por “colgar” juntas una cantidad suficientemente grande de neuronas artificiales en una red. La noción de **emergencia** sugeriría que tal red, una vez alcanzada alguna masa crítica, debería comenzar a pensar espontáneamente. Todo esto, claro está, en el supuesto de que la glía no jugara ningún papel al respecto.

Ésta es una idea muy atractiva y seductora, porque permitiría la posibilidad de construir “conciencia” e inteligencia sin entenderla previamente. Entender la inteligencia es difícil y probablemente lo será durante algún tiempo al menos. De ahí que la posibilidad de que la inteligencia pueda **emerger** espontáneamente a partir de las interacciones de una gran colección de partes simples tiene un atractivo considerable para los investigadores y diseñadores de máquinas inteligentes. Desgraciadamente, a partir de esa idea no surge un enfoque práctico para su construcción. El concepto de **emergencia** en sí mismo ni ofrece una guía de cómo construir tal sistema, ni perspicacia acerca de por qué debería funcionar.

Irónicamente, la aparente inescrutabilidad de la idea de conciencia e inteligencia como un comportamiento **emergente** es la causa de su continua “popularidad”. La **emergencia** ofrece una manera para creer en la causalidad física mientras, simultáneamente, mantiene la imposibilidad de una explicación reduccionista del pensamiento. Para aquellos que

temen las explicaciones mecanicistas respecto a la mente humana, la ignorancia de cómo las interacciones locales producen comportamiento **emergente**, ofrece un *reasoning fog* (niebla en el razonamiento) en la cual “ocultar el alma”.

Modelos de comportamiento **emergente** son las redes de neuronas y demás modelos conexionistas y los autómatas celulares, así como los modelos evolutivos. Los antireduccionistas interpretan estos procesos como un “schism” entre los racionalistas simbólicos, a los que se oponen, y los gestaltistas a quienes apoyan. ¿De que lado se está? A los informáticos sólo les interesan los aspectos prácticos de la cuestión. ¿Cómo se debería construir una inteligencia **emergente**? ¿Qué información debería ser necesaria conocer para tener éxito? ¿Cómo puede esa información ser determinada por experimentos? El sistema **emergente** más fácilmente imaginable debería ser una implementación del pensamiento simbólico antes que una refutación del mismo, como ocurre, “mutatis mutandis”, con el origen de la inteligencia humana. Después de todo no hace tanto tiempo en que ésta se produjo. ¿Cuáles serían los requerimientos de almacenamiento de información de tal sistema de pensamiento **emergente**?

La posibilidad de alcanzar efectos globales basados en perturbaciones locales, se da tanto en el cerebro, concretamente en el modelo del sistema tálamo-cortical (Edelman, 2002), como en ciertos sistemas físicos. Generalmente, en condiciones alejadas del equilibrio, pueden darse regímenes dinámicos en los que la “longitud de correlación” entre los elementos constituyentes aumenta de forma drástica, lo que significa, precisamente, que cualquier perturbación local tiene efectos globales fuertes y rápidos (Nicolis, 1989). La posición metafísica de Edelman es el realismo limitado y la epistemológica, la epistemología de base biológica, cuya idea clave es la noción de que los conceptos no son, en primera instancia, sentenciales; es decir, no son proposiciones de un lenguaje, sino que son construcciones que el cerebro desarrolla al levantar mapas de sus propias respuestas al lenguaje.

Con el nombre de “**emergencia**” o “fulguración” se describe el hecho de que el todo presenta propiedades que no tienen individualmente las partes. Por ejemplo, los aminoácidos de una bacteria carecen de la propiedad de “autorreproducirse”; no obstante, el conjunto, con otras sustancias, posee esa propiedad. Por eso, cuando se observa un fenómeno parcialmente, como algunas variables escapan a la observación, el sistema

representado por el resto puede desarrollar propiedades notables y hasta “milagrosas”. Ejemplo, los conjuros logran milagros, en apariencia, simplemente porque no todas las variables significativas son observables. Es casi seguro que alguna de las propiedades milagrosas del cerebro, por ejemplo, tener premoniciones, inteligencia, etc., sean milagrosas en este sentido.

El hecho es que desde la teoría cuántica se hizo a los sucesos, no a las partículas; o sea, la “materia”, los elementos de la física, de modo que la “materia” no forma parte de la materia última del mundo, y es meramente una forma conveniente de coleccionar acontecimientos o eventos en haces. Es decir, la física ha venido haciendo a la materia menos material. Del mismo modo, la psicología ha estado haciendo a la mente menos mental. Así, desde ambos extremos, la física y la psicología se han ido acercando y haciendo más factible la doctrina del “movimiento mental” sugerida por la crítica de la conciencia de William James. Para Russell tanto la mente como la materia son formas convenientes de grupos de eventos.

Para David J. Chalmers (Chalmers, 1996), la conciencia de las sensaciones pudiera ser una propiedad de todos los sistemas funcionales, incluidos los termostatos, mientras que la conciencia de la identidad personal y la libertad de actuación quedan restringidas a los sistemas sofisticados dotados de lenguajes. En cambio, Searle sostiene que la conciencia es una propiedad natural de los sistemas biológicos complejos, y afirma que la verdadera laguna explicativa se encuentra en el área de la volición y los actos libres. Dennett, a su vez, ve en el supuesto conflicto entre determinismo y libre albedrío otra más de las falsas dicotomías que surgen siempre que uno se olvida de llevar hasta el final las implicaciones del diseño evolutivo.

2.4. Teoría del caos y conciencia

Walter Freeman (Freeman, 1995), el neurobiólogo de la Universidad de California en Berkeley y el filósofo Andy Clark, de la Universidad de Washington en St. Louis, sugieren, **basándose en la teoría del caos y la emergencia**, que las propiedades de la conciencia surgen a partir de fluctuaciones minúsculas en los elementos que la conforman y que, en consecuencia, es imposible predecirlas a partir de esos elementos. Instalados en la teoría del caos, Freeman y seguidores sostienen que todos estos son seudoproblemas

que surgen como consecuencia de la negativa a aceptar la aserción de Hume según la cual el concepto de “causa” es una construcción de la psicología humana. Freeman, uno de los primeros investigadores de las oscilaciones neuronales de 40 hercios como signo de consciencia, afirma que dichas oscilaciones tal vez desempeñen algún papel en la consciencia pero, era muy poco probable que fueran la clave para la solución del problema de la misma. Por tanto, resulta ser una vía muerta, un callejón sin salida. Él aboga por un modelo más complejo basado en la teoría del caos. Según un experimento llevado a cabo en el “Instituto de Ciencia No Lineal” de San Diego, California, las neuronas es el primer ejemplo de lo que se conoce como “computación caótica”. En efecto, parece ser que cuando el sistema nervioso y más en concreto el cerebro recibe un estímulo cada neurona se comporta de un modo concreto, singular e impredecible. No es hasta una fracción de segundo después cuando la corriente caótica se amortigua y surge un patrón estable que determina lo que se denomina “pensamiento”.

Los sistemas caóticos parecen fortuitos, pero en realidad encierran un orden oculto, que se puede describir con objetos matemáticos llamados atractores. Estos sistemas poseen una especial sensibilidad a las condiciones iniciales. Esto se conoce como el efecto mariposa (May, 2004), frase que procede de una conferencia que dio en Washington capital Edward Lorenz en 1972 y que se titulaba: “¿El batir de las alas de una mariposa en Brasil desencadena un tornado en Texas?” Esta noción no era nueva, de hecho, apareció en un impactante y breve relato de Ray Bradbury de 1952 (Bradbury, 1952), titulado: “El Ruido de un Trueno”. En el relato, la mente de una mariposa prehistórica altera el resultado de unas elecciones presidenciales. Y la popularidad de la frase, se debe probablemente a la conocida obra de James Gleick, “Chaos” (Gleick, 1987), cuyo primer capítulo se titula precisamente: “El Efecto Mariposa”. Trazando los patrones de disparo de grandes grupos de neuronas éstos, demostró Freeman, forman patrones caóticos. Esta conducta podría explicar la capacidad del cerebro para reaccionar frente a datos complejos preceptuales con asombrosa rapidez. Ahora bien, humildemente, Freeman reconoció que, en el mejor de los casos, su teoría era una pieza más del rompecabezas.

Más aún, el efecto mariposa tiene un antecedente en el relato de ciencia-ficción de W. Moore (Moore, 1955) titulado “Aniversario fatal”. En donde, lo que vale para la naturaleza también vale para la historia en donde basta un pequeño detalle para mudar su curso. Al menos esto fue lo que aprendió, en cabeza propia, el protagonista del relato que,

expedido por una máquina del tiempo del siglo XX al campo de batalla de Gettysburg, en julio de 1863, un acto aparentemente insignificante, invierte la suerte de la batalla y así borra “su” mundo, en el que una confederación sudista independiente domina los EE.UU., para dejar su lugar al mundo actual donde la potencia hegemónica está representada, como es sabido, por los EE.UU. “postlincolnianos”.

El rasgo clave que distingue entre el comportamiento simple y el caótico es lo que científicamente se denomina “alinealidad”. El comportamiento de un sistema lineal es una simple superposición de los comportamientos de sus elementos y causas. En otros términos, el todo no es más que la suma de las partes. En cambio, en un sistema no lineal las causas que actúan en combinación pueden producir efectos nuevos y completamente inesperados. De hecho, la relación entre la entrada y la salida es bastante menos directa.

Oliver Sacks acertó plenamente cuando refiriéndose a los procesos cerebrales dijo (Sacks, 1986): *Los procesos mentales, que constituyen nuestro ser y nuestra vida, no son abstractos y mecánicos, también son personales, y como tales implican no sólo la clasificación y la ordenación en categorías, sino también una actividad continua de juicio y sentimiento.*

En apoyo de esta opinión sobre que los procesos mentales afectan a toda la persona, viene el descubrimiento de que los humanos tienen un segundo “cerebro” en el tubo digestivo. En efecto, dicho tubo está tapizado por más de cien millones de células nerviosas, una cantidad muy similar a la que alberga la medula espinal. Hasta ahora se pensaba que la única misión de las neuronas intestinales consistía en regular las funciones intestinales, digestivas y los movimientos peristálticos, pero ahora acaba de comprobarse que también tiene otras funciones. Entre ellas, producir sustancias que gobiernan los estados anímicos humanos, así como otros compuestos químicos que mitigan el dolor. De hecho, en este “cerebro” entérico se han descubierto hasta treinta clases distintas de neurotransmisores a los que gobiernan las actividades intelectuales de la cavidad craneal. Por cierto, las “mariposas” en el estómago cuando se está en circunstancias emocionales intensas no es sino el resultado del funcionamiento, a “toda máquina”, de dicho “cerebro”.

Stephen Gould, y en esto coincidía con Gunther Stent, creía que el cerebro sencillamente no era capaz de resolver ciertas cuestiones. Gould, que se consideraba un materialista

anticuado, creía que la mente **emerge** de las complejidades de la organización neuronal, cuestión ésta que no se entiende bien del todo.

El problema de la mente se puede enfocar y abordar de distintas maneras. En primer lugar, está su aspecto histórico: ¿Cómo y por qué se volvió tan listo el “homo sapiens”? Alfred Wallace (Wallace, 1858). y Charles Darwin (Darwin, 1970) ofrecieron una respuesta general: la selección natural. Charles Darwin, en el “Origen del hombre” (Darwin, 1871), afirma que la diferencia entre la mente humana y la de los demás seres era “de grado y no de naturaleza”. Este juicio, respetado por los entendidos, recibió respaldo siglos después cuando se supo que el ser humano comparte el 98% de genes con chimpancé. Ahora bien, la pregunta va de suyo: Si esa herencia genética compartida pudiera explicar el origen evolutivo de la mente humana ¿por qué ningún chimpancé escribe artículos, demuestra teoremas, cocina, usa herramientas, etc.? En realidad cada vez hay más pruebas de que, contrariamente a la tesis de Darwin sobre la continuidad entre los demás y la especie humana, media un profundo y amplio hiato entre el intelecto y la “mente” animal.

Para poder superar ese hiato, lo primero es establecer claramente cuales son las diferencias, o mejor las características diferenciales que constituyen, según Marc Hauser (Hauser, 2009), la “humanidad”; o sea, lo característico de lo humano. Estos rasgos característicos que distinguen a la mente humana de las mentes animales son:

A) “Computación generativa”. Es decir, la capacidad para crear una variedad prácticamente ilimitada de “expresiones” ya sean organizaciones de palabras, secuencias de notas musicales, conjuntos de acciones o cadenas de símbolos matemáticos. La “computación generativa” se refleja en el uso de herramientas. Otras criaturas se valen de útiles, pero las fabrican de un solo material y para una finalidad exclusiva. En cambio, el hombre combina sin dificultad diversos materiales para crear sus útiles que a menudo emplea para varios fines. Por ejemplo, el orangután se sirve de una hoja para cubrirse; los humanos utilizan, un simple lápiz, hecho de varios materiales para más de un fin. Uno de los útiles elementales del ser humano, el lápiz del número 2, el que se usa para cumplimentar un test, ilustra la excepcional libertad de la mente humana en comparación con el limitado alcance de la cognición animal. El lápiz se sostiene por la madera, se escribe con la mina y se borra lo escrito

con la goma del otro extremo, sujeta por un anillo de metal. Cuatro materiales, cada uno con una función, unidos en un objeto. Aunque este útil fue concebido para escribir, puede servir para sujetar el pelo en un moño, señalar una página o para espantar un insecto molesto. En cambio, los útiles de los animales, como los palitos que usan los chimpancés para extraer “termitas de los hormigueros”, están compuestos de un solo material, fueron “diseñados” para una función única y nunca se usan para otras. En suma, no poseen las propiedades combinatorias del lápiz. **Y justo estas propiedades combinatorias son las que pueden dar lugar a funcionalidades y comportamientos emergentes.**

La computación generativa comprende los dos tipos de operaciones siguientes.

- 1) La recursión o recurrencia, que consiste en la repetición de una regla para crear nuevas expresiones. Una frase breve puede incrustarse en otra, repetidamente, y crear descripciones más extensas y ricas del pensamiento humano como, verbigracia, la sencilla pero poética expresión de Gertrude Stein: “Una rosa, es una rosa, es una rosa”. Se tiene en acción la recurrencia en otro sencillo elemento que se encuentra en el equipo de multitud de campistas: el vaso telescópico retráctil. Para realizar este dispositivo, el fabricante necesita programar sólo una regla sencilla: añadir al último segmento un aro de diámetro algo mayor y repetir el procedimiento hasta que se obtenga el tamaño del vaso deseado. Los humanos se valen de operaciones recursivas como ésta en prácticamente todos los aspectos de la vida mental: desde el lenguaje, la música o las matemáticas hasta la generación de una gama ilimitada de movimientos en las piernas, manos o boca. En cambio, los únicos atisbos de recurrencia que se aprecian en los animales proceden de la observación de su sistema motor en acción. Es un hecho obvio, que todas las criaturas poseen maquinaria motora recursiva, integrada en un equipamiento operativo estándar o para andar, sitúan una mano o pata delante de otra, y así una y otra vez y otra vez... Para alimentarse pueden asir un objeto y llevárselo repetidamente a la boca, hasta que su estómago les comunica que ya están saciados. En las mentes animales, ese sistema recursivo se encuentra encerrado bajo llave en las regiones motoras del cerebro, siendo inaccesible para otras regiones cerebrales. Su existencia lleva a pensar que una etapa crítica en la adquisición de la peculiar forma de pensamiento humano no fue que la

recurrencia evolucionase hasta convertirse en una nueva forma de computación, sino la liberación de la recurrencia de su “prisión” motora, dejándola acceder a otros dominios del pensamiento, las formas en que la recurrencia se liberó de esa función restringida está vinculada con uno de los otros ingrediente humanos: la promiscuidad de interfases. Dicha promiscuidad se da en la interactuación de los saberes lingüísticos junto con los cómputos que requieren, con otros dominios de conocimiento que resaltan poderosamente las excepcionales facultades humanas para establecer conexiones promiscuas entre sistemas de conocimiento.

- 2) La combinatoria. Por su parte la operación combinatoria consiste en conjuntar elementos discretos para engendrar nuevas ideas, tal vez expresables con palabras nuevas “parasol”, o formas musicales, entre otras posibilidades.
- B) “Combinación promiscua de ideas”. Por ella, el ser humano conecta sin esfuerzo pensamientos de campos de conocimiento muy dispares. Ello permite unir al ser humano lo que sabe, por ejemplo, sobre arte, sexo, espacio, causalidad y amistad. Una mezcla tal engendra leyes nuevas, nuevas relaciones sociales y nuevas técnicas. Éste es el caso, verbigracia, cuando se concluye que está prohibido (dominio moral) empujar a otra persona (dominio de actos motores) intencionadamente (dominio de la psicología) ante un tren en marcha (dominio de los objetos) para salvar la vida (dominio moral) a otras cinco (dominio numérico).
- C) Uso de símbolos mentales. El ser humano tiene la capacidad de convertir espontáneamente experiencia sensorial (real o imaginaria) en un símbolo, que se puede conservar para uno mismo manifestar a otros a través del lenguaje, el arte, la música o la codificación informática. En suma, capacidad de codificar experiencias sensoriales reales o imaginadas que forman la base de un sistema de comunicación rico o complejo. Tales símbolos pueden ser para uno mismo o manifestados a los demás mediante palabras o imágenes.
- D) Pensamiento abstracto. Los seres humanos son los únicos en la Tierra, y el Universo conocido por ellos, con pensamiento abstracto. En efecto, a diferencia de los animales, anclados en buena medida, en experiencias sensoriales o perceptivas, muchos de los pensamientos de los humanos no tienen una conexión clara y desde luego inmediata con sucesos de este tipo. Sólo los humanos pueden reflexionar sobre

la apariencia de los unicornios o los alienígenas, sobre espacios de n dimensiones, sobre sustantivos o verbos, sobre los números, el infinito o Dios. Este pensamiento abstracto permite, pues, trascender lo audible, “olible”, tangible, visible o “saborable” y remontarse a lo no inmediatamente accesible.

Sin embargo, no siempre esto fue así, los antropólogos piensan que el cambio relevante y profundo se produjo hace unos 800.000 años, en la era paleolítica y se aceleró hace unos 50.000. Todo ello basándose en restos arqueológicos. Sin embargo, dichos restos guardan ¿guardarán? silencio ¿eterno? sobre los orígenes y las presiones selectivas que desembocaron en los cuatro ingredientes anteriores que integran la “humanidad”, la singularidad del ser humano.

Las preciosas pinturas rupestres de Altamira y Lascaux, hacen ver que los antepasados humanos comprendían la doble naturaleza de sus pinturas: son objetos que, a su vez, se refieren a objetos y a sucesos. Pero no revelan si sus autores y admiradores expresaban sus preferencias estéticas sobre esas obras de arte mediante símbolos organizados en clases gramaticales, sustantivos, adjetivos, verbos, etc., o si imaginaron que esas ideas podían ser comunicadas igualmente bien mediante sonidos o signos dependiendo de sus sistemas sensoriales. Análogamente, tampoco los instrumentos más antiguos que se han hallado, flautas de hueso y marfil de hace unos 35.000 años, cuentan historia alguna sobre su uso, ni dicen si en ellos se tocaron unas pocas notas repetida, monótona y machaconamente, o si el artífice, también era artista y concibió, como Wagner, la inclusión recursiva de unos temas en otros.

Lo que si se puede afirmar, no obstante, es que con total seguridad, todos los humanos, desde los cazadores-recolectores de la sabana africana hasta los financieros de Wall-Street, poseen al nacer los cuatro ingredientes de la “humanidad”. Sin embargo, las formas de agregarlos y combinarlos en las “recetas” para la creación cultural sí varían considerablemente de un grupo a otro. Así, las culturas humanas, difieren en idioma, composición musical, normas morales o artefactos. Desde un punto de vista de una cultura, las prácticas ajenas resultan, a menudo, extravagantes, a veces repulsivas, con frecuencia incomprensibles o absurdas y, en ocasiones, inmorales. Ningún otro animal despliega, tan amplia gama de estilos de vida.

Lo que sí se sabe, en cambio, es que el tamaño del cerebro no es criterio infalible para determinar si un animal posee capacidades de “humanidad” o no. Por ejemplo, el cerebro de una ballena asesina pesa 5.620 gramos; mientras que el de un humano pesa 1.350 gramos. Y en el otro extremo, el cerebro de la musaraña etrusca sólo pesa 0,1 gramos. En suma, el tamaño del cerebro no explica, por sí solo, la singularidad de la mente humana, pero sí afecta a ciertos aspectos estructurales del cerebro y, por consiguiente, a ciertos aspectos del pensamiento.

Se ha comprobado que la mayoría de los tipos de células nerviosas cerebrales, junto con sus mensajeros químicos, son los mismos en todas las especies de vertebrados incluidos el ser humano. Además, la organización general de las estructuras de la corteza cerebral es, en buena medida, la misma en monos, simios y humanos. Dicho de otro modo, los humanos comparten un buen número de rasgos cerebrales con otras especies. Las diferencias residen en el tamaño relativo de determinadas regiones de la corteza y en la forma en que éstas se interconectan. Esas diferencias dan vida a pensamientos sin parangón en el resto del reino animal.

Sobre la consciencia se ha escrito mucho y poco atinado. Como lo señaló Stuart Kauffman (Kauffman, 2003): *Dado lo poco sensato y lo mucho inútil que se ha vertido sobre la consciencia, una hipótesis más no empeorará las cosas.* Y, a continuación, plantea su hipótesis como sigue: *La consciencia está asociada a una “toma de decisiones” de alta resolución, conducente a los comportamientos alternativos en los agentes autónomos moleculares, que abarca los reinos cuántico y clásico y equivalente a la persistente propagación de bucles interconectados y percolados de coherencia cuántica, los cuales simultáneamente y sistemáticamente pierden coherencia, transformándose en clásicos. El paso al comportamiento clásico representa la “mente” actuando sobre “materia”. La consciencia es la experiencia interna que el agente tiene de esa red percolada de coherencia cuántica que sostenidamente sufre decoherencia hacia el comportamiento clásico.* Cierto, retórico y oscuro. Sin embargo, para él es innegable que la consciencia es un epifenómeno del comportamiento del cerebro. Y que este epifenómeno puede entenderse.

2.5. Mente y Computador

Los partidarios de la IA fuerte afirmaban que era sólo cuestión de tiempo el que las computadoras superen el Test de Turing (TT). Y, en efecto, el sistema Siri, asistente de voz del iPhone 4S, que permite la navegación en Internet, diseñado por Apple “pasó” el TT, al superar con éxito, la batería de preguntas que le planteó “*New York Times*”. La propuesta que se hace en esta Tesis es que cualidades mentales tales como la conciencia **emerjan** naturalmente como una consecuencia del comportamiento computacional complejo y que los computadores poseen entendimiento de un dominio de problemas simplemente en virtud de ser capaces de ejecutar conjuntos de tareas representativas de dicho dominio.

Los procesos que pueden resolver problemas con respuestas definidas pero no computables son deterministas pero son más potentes que las computadoras. Ya que los futuros estados de un sistema no computable no pueden predecirse a partir del pasado por mera computación, la no computabilidad puede confundirse fácilmente con el no determinismo. En este sentido, los procesos mentales son inherentemente más potentes que los computacionales.

La reflexión acerca de un dominio de problemas que trascienda los dominios de la computación y el razonamiento formal dentro del dominio parece ser una cuestión capital en lo concerniente a la creatividad. Para Penrose, la visión platónica de que las verdades matemáticas son tan reales, posiblemente más reales que, los objetos del mundo físico, es algo incontrovertido. Desde esta perspectiva y para lo que aquí interesa, el conjunto de Mandelbrot es una réplica aproximada, aunque no exacta, de un conjunto como un todo. La propiedad de que cada parte de un conjunto de Mandelbrot sea un microcosmos del conjunto completo, sugiere que la realidad física puede tener una estructura similar, con cada átomo conteniendo un universo de átomos que a su vez contiene universos. Tal infinita escalabilidad, de acuerdo con la visión científica predominante actualmente de que la realidad física puede emular la realidad matemática, es muy sugerente en si misma y por las circunstancias que de ella derivan. En primer término, la propiedad de que cada parte, no importa su tamaño, contenga una réplica estructural del conjunto completo es similar al código genético que replica el organismo adulto en cada célula del embrión.

En segundo lugar, los universos escalables, también conocidos como universos de Mandelbrot no tienen “partículas elementales” en cualquier sentido absoluto. Pero limitaciones sobre observaciones en un nivel de la escala acerca de fenómenos en otros niveles de las escalas pueden causar universos a un nivel para comportarse como partículas fundamentales y exhibir efectos cuánticos cuando se contemplan desde un nivel más alto. Las actuales teorías físicas si son compatibles con esta visión. El factor de escala entre sucesivos universos podría estar relacionado con el número de partículas en el universo $\approx 10^{80}$, o mediante los efectos cuánticos, a la constante de Planck: 10^{-35} .

La pretensión de Penrose, en su primer libro (Penrose, 1989), era refutar la pretensión de los defensores de la IA fuerte en el sentido de que los computadores podrían “replicar” todos los atributos de los humanos incluida la conciencia. Para ello, basaba su argumentación en el famoso teorema de incompletud de Gödel (Gödel, 1931) que afirma que “todo sistema de axiomas consistente y suficiente para generar la aritmética de Peano (Peano, 1889) genera afirmaciones que no pueden probarse o refutarse con tales axiomas”, de ahí que sea incompleto; sin embargo, un ser humano podría ver que son verdaderas o falsas. La razón es que los humanos no están limitados a realizar sólo inferencias deductivas. De hecho un buen modelo de la mente tendría que asegurar, entre otras cosas, la posibilidad de hacer inferencias no deductivas, y éstas proporcionarían una manera de obviar el resultado de Gödel.

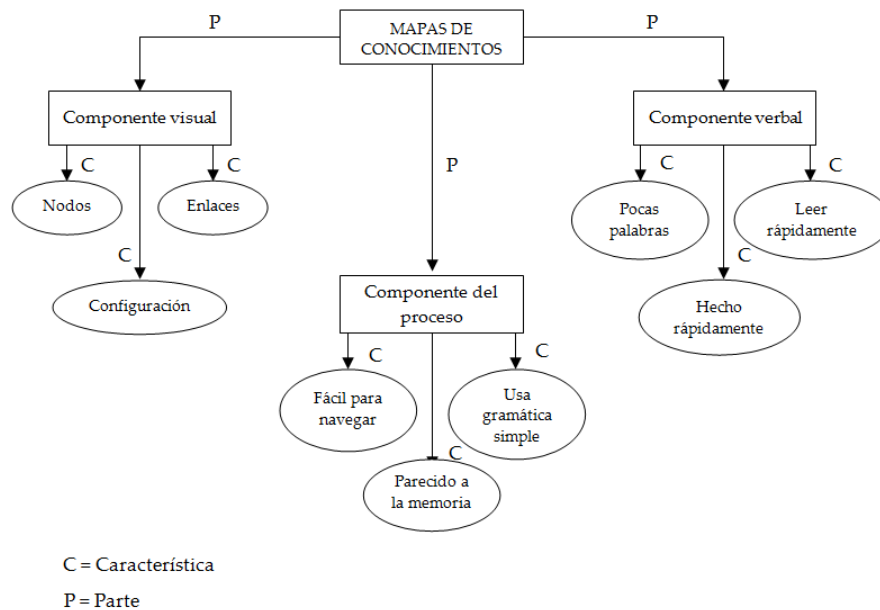


Figura 2.1. Mapa de conocimiento autorreferente.

La esencia de la fórmula gödeliana es que es “autorreferente”. Aunque no es fácil dar ejemplos fuera de las matemáticas de entes autorreferentes, porque no abundan, existen en todos los dominios de la vida. Por ejemplo, definir un mapa de conocimiento en términos de un mapa de conocimiento como se muestra en la figura 2.1, es un caso de definición autorreferente. Otro, este procedente de la literatura, es la definición de soneto de Lope de Vega, que aparece en su obra “La niña de plata”, es otro ejemplo de autorreferencia. He aquí el famoso “Soneto de Violante”:

Un soneto me manda hacer Violante,
que en mi vida me he visto en tal aprieto;
catorce versos dicen que es soneto:
burla burlando van los tres delante.
Yo pensé que no hallara consonante
y estoy a la mitad de otro cuarteto;
mas si me veo en el primer terceto
no hay cosa en los cuartetos que me espante.
Por el primer terceto voy entrando
y parece que entré con pie derecho,
pues fin con este verso le voy dando.
Ya estoy en el segundo, y aun sospecho
que voy los trece versos acabando;
contad si son catorce, y está hecho.

Las paradojas de la conciencia surgen porque un ser consciente puede darse cuenta de sí mismo al igual que de otras cosas, pero no puede construirse como si fuera divisible en partes; es decir, un ser consciente, además de saber que sabe, puede considerarse a sí mismo y a su acción.

Una neurocientífica, por ella denominada de sistemas, Patricia Goldman-Rakic se dedica a investigar la corteza donde cree que está la sede de la facultad de razonar y otras decisiones. Ella espera demostrar que la psicología, la psiquiatría y otras “macrodisciplinas” de la mente pueden explicarse mediante modelos más reduccionistas centrados en procesos neuronales, genéticos y moleculares. Un aspecto importante de su trabajo es la memoria operativa o de trabajo, que permite mantener el hilo de una conversación, leer un libro, jugar una partida de cartas, realizar de memoria cálculos

aritméticos elementales, etc. Muchos neurocientíficos creen que una mejor comprensión de esta memoria ayudará a resolver enigmas como, verbigracia, el libre albedrío, la conciencia, el problema del ensamblaje, entre otros. Es decir, servirá para colmar la laguna explicativa. Sin embargo, a pesar de sus trabajos y, sobre todo, manifestaciones, dicha laguna permanece, de momento, infranqueable. Eso al menos es lo que opina Timothy Beardsley (Beardsley, 1997). Para Chris Bremser (Horgan, 2001): *Cualquiera que tenga una mínima noción de programación de computadores reconocerá que esta laguna explicativa es exactamente igual a la diferencia entre el código máquina, unos y ceros, y los verdaderos programas. Cualquiera puede conectar un dispositivo de mayor alcance a la memoria de un computador, pero si no tiene, o no infiere, los constructores programáticos operativos de nivel superior, no se conseguirá gran cosa.* Douglas Hofstadter (Hofstadter, 1980), trató este tema más extensamente en su conocida obra “Gödel, Escher, Bach”.

Como lo citó Alan Ross Anderson (Anderson, 1964), desde el año 1950 hasta 1964, se han publicado más de mil artículos acerca de si las máquinas pueden o no “pensar”. Naturalmente, si se amplía el periodo hasta la actualidad, el número de artículos ha crecido al menos en medio orden de magnitud. Sin embargo, la respuesta final cada vez parece más incierta.

Hay un argumento curioso presentado por J. R. Lucas (Lucas, 1961) según el cual el argumento o tesis mecanicista implica que el modelo mecánico de la mente funciona según “principios mecánicos”; es decir, que se puede comprender la operación en función de las operaciones de sus partes, y la operación de cada parte estará determinada por su estudio inicial y la construcción de la máquina o, a lo sumo, será una selección al azar entre un número determinado de operaciones determinadas. Si el mecanicista construye una máquina que sea tan complicada que ese principio no tenga validez para ella ya no es una máquina... cualquiera que haya sido la manera con que ha sido construida. Dijérase más bien que él creó una mente... pues no sería el total de las partes. Y, más adelante, continúa, enfáticamente, diciendo lo siguiente: *No existe ningún límite arbitrario a la investigación científica; pero ninguna investigación científica jamás agotará la infinita variedad de la mente humana.*

Ernest Nagel y James R. Newman (Nagel, 1958), afirman paladinamente que [...] el teorema de Gödel indica que la estructura y la potencia de la mente humana son mucho más complejas y sutiles que cualquier máquina no viviente hasta ahora imaginada. Dicha afirmación, empero, fue rotundamente rebatida por Michael Scriven (Scriven, 1960) señalando que el teorema de Gödel (Gödel, 1931) no es un obstáculo mayor para un computador que para las personas... De hecho, de la misma manera que un ser humano puede reconocer la verdad de una fórmula indemostrable comparando lo que dice con lo que sabe en el caso, lo mismo puede hacer un computador.

Hilary Putnam (Putnam, 1960), señala en opinión del doctorando, muy acertadamente, que los diferentes problemas y enigmas que constituyen el tradicional problema de la mente y el cuerpo son de carácter totalmente lingüístico y lógico: los pocos “hechos empíricos” que hay en esta esfera apoyan tanto al uno como al otro de los puntos de vista. Putnam, en su artículo, trata de mostrar que todos los puntos controvertidos surgen en conexión con cualquier sistema de computación capaz de responder preguntas acerca de su propia estructura y que, por tanto, no tienen nada que hacer con el carácter único, si es que es único, de la experiencia subjetiva humana. Y después de señalar que preguntas tales como ¿cómo sé que tengo un dolor? son “desviantes”, en el sentido de que se desvían de una regularidad semántica en el lenguaje natural apropiado, y sin embargo, ¿cómo sabe alguien que otra persona tiene un dolor? no lo es. Y demostró, usando máquinas de Turing, MT, que la diferencia de estado entre las dos preguntas se refleja en el caso de máquinas. A continuación, plantea, en relación con el problema mente-cuerpo, la cuestión de saber si o no es permisible identificar los acontecimientos mentales y los acontecimientos físicos. Para ello, arguye que es posible construir una analogía lógica, sin ningún abuso especioso de la analogía. Para obtenerla, identifica una teoría científica con un cálculo parcial interpretado en el sentido de Carnap (Carnap, 1953). A continuación, imagina una MT capaz de generar teorías, las prueba, suponiendo que es posible “mecanizar” la lógica inductiva en cierta medida, y acepta teorías que satisfacen ciertos criterios; por ejemplo, el éxito en la predicción. Si la MT puede escudriñarse a sí misma mientras está operando, puede formular teorías acerca de su propia estructura y someterlas a prueba. En este sentido, todas las consideraciones a favor y en contra de la mente y el cuerpo podrían parangonarse a consideraciones a favor y en contra de la aseveración; verbigracia, de que el estado A es en realidad idéntico al hecho de que un elemento físico de la máquina E hace algo concreto y definido. Es decir, la condición

necesaria y suficiente de estar en el estado A, si E, está encendido. Esto, evidentemente, es una proposición “sintética” para la máquina.

Posteriormente, Putnam establece una distinción entre estados mentales y estados lógicos usando de nuevo la MT. En efecto, cuando se describe un computador por medio de una tabla de estados y entradas, los estados son evidentemente “estados lógicos”. De hecho, la descripción lógica de la computadora no incluye ninguna especificación de “carácter físico” de toda la máquina; esto es, si está hecha, con chips, cartón, seres humanos o cualquier otra cosa. En suma, la computadora, cual MT, es algo abstracto que puede ser realizada físicamente por un número casi infinito de diferentes maneras. En cuanto la máquina se realiza físicamente, la máquina tiene un número casi infinito de estados adicionales, “estados estructurales”, a los lógicos; por ejemplo, por qué un elemento no funciona, etc. Los computadores actuales y otras máquinas, por ejemplo, los automóviles, tienen elementos de control de sí mismos, y, en este caso, están en disposición de describir algunos de los estados estructurales y así son similares a los humanos que pueden **percatarse** de algunos de los defectos de funcionamiento de su cuerpo y con diferentes grados de “fidedignidad”. Del mismo modo que hay dos descripciones posibles del comportamiento de un computador; a saber, el proyecto estructural del mismo y la tabla de estados y entradas-salidas, también hay dos posibles descripciones de la psicología humana: La “conductista” que incluye las teorías que emplean “construcciones hipotéticas” incluyendo las anatomofisiológicas, y tiene por objeto proporcionar una descripción “fiscalista” completa del comportamiento humano, en términos relacionados con la física y la química y la “mentalista”. En este caso, se busca dar una descripción más abstracta de los procesos mentales humanos, en función de “estados mentales”, sin especificar la realización física, en caso de que haya alguna, y de “impresiones”, los símbolos sobre la cinta en una MT. Esta descripción formaba parte del programa de psicología clásica. Como es sabido la psicología clásica fracasó en el empeño, unos creen que por cuestiones metodológicas, Putnam cree que por razones empíricas: los “estados mentales” y las “impresiones” no constituyen un sistema causal cerrado en la medida en que lo constituyen las MT.

La analogía entre los estados lógicos de un computador y los estados mentales de un ser humano, por una parte, y los estados estructurales del computador, no solo el actual, y los estados físicos de un ser humano por otro, es, heurísticamente hablando, muy

prometedora, hasta el punto de que, una “exploración detallada” de la misma podría, cuando menos, facilitar una mayor clasificación de la noción de estado mental. Para llevar a cabo esta exploración detallada es necesario tener en cuenta los siguientes rasgos que parecen distinguir los estados lógicos y mentales respectivamente de los estados estructurales y físicos:

- a) La organización funcional: solución de problemas, pensamiento, etc., del ser humano o del computador, que puede ser definida en función de las secuencias de los estados mentales o lógicos, respectivamente, y de las “verbalizaciones” que los acompañan, sin referencia al carácter de la “realización física” de esos estados.
- b) Los estados parecen estar íntimamente relacionados con la verbalización.
- c) En el caso del pensamiento racional o de una computación, el “programa” que determina cuáles estados suceden a cuáles otros, está abierto a una crítica racional.

Finalmente, está el problema de la “identidad” mente-cuerpo o, para ser más precisos, la “**identificación**” de los estados mentales con los estados físicos correspondientes o estados lógicos con los estructurales correspondientes. Wittgenstein empleó, al respecto, el siguiente argumento contra la identificación: Si observo una imagen posterior y, simultáneamente, observo el estado de mi cerebro, con la ayuda de un instrumento adecuado, observo dos cosas, no una sola. Russell, por su parte, llevó el asunto de la identidad al extremo absurdo de mantener que lo único que jamás se ve son partes de los propios cerebros. Análogamente un Russell mecánico podría “argüir” que lo único que jamás uno observa son los propios chips. Ninguno de los dos Russell está en lo cierto: el ser humano observa acontecimientos en el mundo externo, y el proceso de la “observación” implica acontecimientos en el cerebro. De hecho, una buena hipótesis de trabajo y, por tanto de partida, podría ser: *Uno se encuentra en el estado mental Ψ si, y sólo si, uno se encuentra en el estado cerebral S* . Ahora bien, en primer lugar, sería sólo una mera correlación; esto es, una simple “generalización empírica”. Sin embargo, lo verdaderamente interesante sería el caso que surgiría si se tuviera un sistema elaborado y teóricamente desarrollado de “correlaciones” como la señalada. A partir de ahí, se podría a empezar a postular cosas tales como: *Es imposible, en principio estar en el estado mental Ψ sin estar en el estado cerebral S* . Y podría perfectamente ocurrir que la “imposibilidad en principio” se elevara a lo que se llama imposibilidad conceptual

(Hanson, 1958); es decir, no se podría concebir que alguien estuviese en el estado mental Ψ sin estar en el estado cerebral S.

Y termina Putnam concluyendo que ya no se puede creer que el problema del cuerpo y la mente sea un auténtico problema teórico, o que su “solución” arrojaría la menor luz sobre el mundo. Y continúa señalando la importancia del hecho de que dos problemas sobre distinto asunto sean idénticos en todos sus aspectos lógicos y metodológicos, puesto que las conclusiones que pudieran obtenerse en uno serían también válidas en el otro. Para rematar señala que si el problema de la mente y del cuerpo se identifica con cualquier problema que no sea meramente conceptual, entonces o bien: (a) ningún argumento empleado alguna vez por un filósofo arroja la menor luz sobre él; o (b) algún argumento filosófico a favor del mecanismo es correcto, o (c) algún argumento dualista demuestra que tanto los humanos como los computadores tienen alma.

Ahora bien, el computacionalista misteriano más inverosímil es Marvin Minsky, uno de los fundadores de la Inteligencia Artificial con: John McCarthy, Claude Shannon, Herb Simon, Oliver Selfridge, Allen Newell, Nathaniel Rochester, Ray Salomonoff, Trenchard More, y Arthur Samuel (Pazos, 1987). Según Minsky, el cerebro no es nada más que una máquina muy complicada cuyas propiedades se pueden replicar por los computadores. Minsky, un archirreduccionista, desprecia a los que duden de que los computadores pueden ser conscientes. Para él la consciencia era una cuestión trivial, es meramente un tipo de memoria a corto plazo. En este sentido, lenguajes recursivos como el LISP (“List Processing”), que permiten seguir la traza de sus distintos pasos de proceso, son sumamente conscientes. En su libro “The Society of Mind” (Minsky, 1985) sostenía que el cerebro contiene muchas estructuras diferentes y muy especializadas que evolucionan a fin de solucionar diferentes problemas. Es decir, la mente tiene muchos métodos distintos para abordar un único problema, por sencillo que sea. Para él, los humanos tienen muchas capas de redes de máquinas que están aprendiendo, cada una de las cuales ha evolucionado para corregir gazapos o adaptar los otros instrumentos a los problemas del pensamiento. Es, así, poco probable que el cerebro se pueda reducir a una serie particular de principios o axiomas, o leyes como pensaba Boole (Boole, 1854) pues el ser humano se enfrenta al mundo real y no a uno matemático definido por axiomas.

Daniel C. Dennett se interesa por la conciencia e, igual que Minsky, la concibe como pensamiento abstracto de alto nivel. Es conocido como el principal defensor del modelo informático de la mente. Una idea interesantísima de Dennett son sus “bombas de intuición” (Dennett, 1996) deliciosos cuasi “experimentos mentales”, sin ninguna argumentación sólida detrás de ellas. Ejemplos de dichas bombas son: La cueva de Platón, El demonio de Descartes, la visión de Hobbes del estado de la naturaleza, el contrato social de Rousseau, el imperativo categórico de Kant, y la “habitación china” de Searle, que fue para quien las creó. Estas bombas no son argumentos, sino relatos. En vez de llegar a una conclusión, si se usan correctamente, bombean una intuición. Así, para él, la idea de que la conciencia es una máquina virtual es una “bomba de intuición”.

A continuación, señaló algo evidente y es que a medida que cambia la tecnología cambian los modelos que se usan para analizar el cerebro. Así, los mecanismos de relojería, las poleas, etc., eran entusiásticamente empleados en los días de Descartes. En la era industrial, las máquinas de vapor, las dinamos y las centralitas telefónicas fueron la analogía adecuada. Hoy son los computadores. Sin embargo, quizás, la metáfora tecnológica más penetrante que se ha usado para explicar lo que pasa en el cerebro, es la escritura; es decir, la idea de pensar en los procesos cerebrales en términos de señales, de transferencia de mensajes escritos.

2.6. Características del estado consciente

Searle y Nagel adscriben dos características al estado consciente; a saber:

- a) “Unidad”. La naturaleza unitaria de la consciencia se refiere al hecho de que las experiencias humanas se presentan a los seres humanos como un todo unificado. Todas las diferentes modalidades sensoriales se funden en una experiencia consciente única y coherente. Así cuando uno acerca un espliego, o lavanda campestre, y se huele su exquisita fragancia y simultáneamente se ve su bello color lila y uno percibe este espliego en el campo sobre un fondo verde en la orilla de un arroyo. La percepción no es sólo total durante el momento de la experiencia, si no que sigue siendo completa semanas después, cuando se recuerda. Con independencia de que participan diferentes órganos como la vista y el olfato y que cada uno usa su propia vía de acceso al cerebro, convergen en el cerebro de tal modo que se unifica la

percepción. Esta naturaleza unitaria de la consciencia plantea un problema difícil, pero quizás no insuperable. Esta naturaleza unitaria puede analizarse descomponiéndola en partes. En un paciente quirúrgico cuyo cerebro está separado entre los dos hemisferios cerebrales, hay dos mentes conscientes, cada uno con su propia percepción unificada. Esto constituye el así denominado problema fácil de la consciencia.

La unidad de la consciencia es una variante del problema del enlace (binding) identificado, en primer lugar, en el estudio de la percepción visual. Una parte íntima de experiencia, el placer subjetivo, verbigracia al ver y oler las lavandas es como la visión y el olor del espliego en los jardines, están conectados juntos y unificados con la visión del campo de lavanda y todas las demás imágenes componentes de la percepción individual. Cada una de estas componentes de la experiencia subjetiva de un individuo, está mediada por diferentes regiones del cerebro dentro de los sistemas visual, olfativo y emocional del receptor. La unidad de la experiencia consciente implica que el proceso de conexión debe, de alguna manera, enlazar e integrar todas estas áreas separadas en el cerebro.

Como un primer paso para resolver el problema fácil de la consciencia, se necesita preguntar si la unidad de la consciencia, una unidad de pensamiento a alcanzar por los sistemas neuronales que media la atención selectiva, es localizada en uno o justo unos pocos sitios, que permitirían manipularlos biológicamente. La respuesta no es clara. Así Gerald Edelman, opina que es probable que esté ampliamente distribuida a través del córtex y el tálamo. Como resultado, Edelman afirma que es inverosímil que se sea capaz de encontrar la consciencia a través de un simple conjunto de correlatos neuronales. Por su parte, Crick y Koch creen que la unidad de la consciencia tendrá correlatos neuronales directos por qué verosímilmente implican un conjunto específico de neuronas con “firmas” neuroanatómicas y moleculares específicas. Los correlatos neuronales, argumentaron, probablemente sólo requiere de un pequeño conjunto de neuronas actuando como un receptor, el foco de atención. Entonces, para ellos, la tarea inicial es localizar, dentro del cerebro, el pequeño conjunto de neuronas cuya actividad se correlaciona mejor con la unidad de la experiencia consciente y luego determinar los circuitos neuronales a los cuales pertenece. ¿Cómo encontrar esta pequeña población de células nerviosas que podrían

mediar la unidad de la conciencia? ¿Qué criterios deben reunir? Crick y Koch (Crick, 2004) (Kandel, 2006) proponen, como se verá más adelante, el “claustrum”.

- b) Subjetividad. La segunda característica de la conciencia, “awareness”, es la que plantea el desafío científico más formidable. Cada una de las experiencias personales en un mundo de sensaciones privadas y únicas que es un mucho más real a cada uno que las experiencias de los demás. Uno experimenta sus propias ideas, humores y sensaciones directamente, mientras que sólo se pueden apreciar las experiencias de otras personas indirectamente por observar o escuchar acerca de ellas. Por consiguiente uno puede preguntarse, si la respuesta al jazmín que se huele o al azul que se ve; esto es, el significado que tiene para uno, es el mismo que para otros. La cuestión aquí no es un asunto de percepción “per se”. No es si cada uno ve una tonalidad similar del mismo azul. Esto es relativamente fácil de establecer registrando el patrón en las células nerviosas únicas del sistema visual de cada individuo. El cerebro reconstruye la percepción de un objeto, pero el objeto percibido; verbigracia, el color azul, parece corresponder a las propiedades físicas de la longitud de onda de la luz reflejada. Realmente, la cuestión es el significado que ese azul tiene para cada individuo. Lo que no se entiende es cómo la actividad eléctrica en las neuronas da lugar al significado que cada uno adscribe a ese color. El hecho de que la experiencia consciente es única para cada persona plantea la cuestión de si es posible determinar objetivamente cualesquiera de las características de la conciencia que son comunes a todos. Si finalmente los sentidos producen experiencias que son completa y personalmente subjetivas, no se puede, el argumento va de suyo, alcanzar una definición general de conciencia basada en la experiencia personal.

Searle y Nagel ilustran la dificultad de explicar la naturaleza subjetiva de la conciencia en términos físicos como sigue: Supóngase que se tiene éxito en registrar la actividad eléctrica de las neuronas en una región conocida como importante para la conciencia mientras la persona sometida a estudio realiza algunas tareas que requieran atención consciente. Por ejemplo, supóngase que se identifican las células que se disparan cuando uno ve y es consciente del color lila del espliego. Entonces se ha dado un primer paso en el estudio de la conciencia; a saber, se ha encontrado lo que Crick y Koch han denominado el “correlato neuronal” de la conciencia para esa percepción. Naturalmente,

esto constituiría un gran avance en la solución del problema porque indica con toda precisión un material concomitante con la percepción consciente. De ahí se pasaría a realizar experimentos para determinar si éstos correlatos también se funden en un todo coherente, el campo verde al lado del arroyo y el lugar, etc. Pero para Searle y Nagel éste es el problema fácil de la consciencia. El problema difícil, arduo o duro, que de las tres formas se denomina, es el segundo misterio el de la experiencia subjetiva. Es decir, ¿cómo responde uno a la imagen lila de la lavanda con un sentimiento que es distintivo para cada uno? Hasta ahora, no se sabe cómo el disparo de neuronas específicas conduce a la componente subjetiva de la percepción consciente, incluso en el caso más simple tanto de hecho, de acuerdo con que se carece de una teoría adecuada de cómo un fenómeno objetivo, tal como señales eléctricas en el cerebro, pueden causar una experiencia subjetiva tal como la tristeza. Y dado que la ciencia, tal y como se practica actualmente, es reduccionista, la visión analítica de eventos complicados mientras la consciencia es irreductiblemente subjetiva, tal teoría cae más allá del alcance humano por ahora.

De acuerdo con Nagel la ciencia no puede abordar la cuestión de la consciencia sin un cambio metodológico significativo, un cambio que debería capacitar a los científicos para identificar y analizar los elementos de la experiencia subjetiva. Estos elementos, probablemente serán componentes básicos de funciones cerebrales, similares a cómo los átomos y las moléculas son componentes básicos de la materia, pero que existen en una forma que aún uno no se imagina. Las reducciones efectuadas rutinariamente en ciencia no son problemáticas, piensa Nagel. La ciencia biológica puede explicar fácilmente como las propiedades de un tipo de materia particular surgen de las propiedades objetivas de las moléculas de que están hechas. De lo que carece la ciencia es de las reglas para explicar cómo las propiedades subjetivas, la consciencia, surge de las propiedades de objetos, células nerviosas conectadas.

Nagel argumenta que la carencia completa de entendimiento en los elementos de la experiencia subjetiva no debería hacer desistir de descubrir los correlatos neuronales de la consciencia y de las reglas que relacionan los fenómenos conscientes a procesos celulares en el cerebro. De hecho, es sólo por acumulación de tal información que se estará en posición de pensar acerca de la reducción de algo subjetivo a algo físico y objetivo. Pero para llegar a una teoría que soporte dicha reducción, primero se tiene que descubrir los

elementos de la conciencia subjetiva. Este descubrimiento, dice Nagel, será de enorme magnitud y tendría extraordinarias implicaciones, requiriendo una revolución en biología y más probablemente una transformación completa del pensamiento científico.

El propósito de los científicos trabajando en la consciencia es mucho más modesto que lo anterior. Pasan de la cuestión peliaguda y siguen realizando avances notorios en entender la percepción y la memoria sin tener en cuenta la experiencia individual. Por ejemplo, se han producido avances en entender las bases neuronales de la percepción del color azul sin afrontar la cuestión de cómo cada persona responde al mismo azul. El problema duro.

Lo que no se entiende es el problema duro de la consciencia; esto es, el misterio de cómo la actividad neuronal da origen a la experiencia subjetiva. Crick y Koch han argumentado que una vez se resuelva el problema “fácil” de la conciencia, su **unidad**, se será capaz de manipular aquellos sistemas neuronales experimentalmente para resolver el problema difícil. ¿Cómo se podría encontrar la pequeña población de células nerviosas que podrían mediar en la unidad de la conciencia? ¿Qué características tendrían que cumplir? En el último artículo de Crick y Koch, antes de la muerte de Crick, el 28 de julio de 2004, se centraron en el “*claustrum*” una pequeña lámina de tejido cerebral localizado bajo el córtex, como el lugar que medía la unidad de la experiencia. Poco se conoce acerca del “*claustrum*” salvo que está conectado a y cambia información con casi todas las regiones motoras y sensoriales del córtex, así como, con la amígdala, que juega un importante papel en las emociones. Crick y Koch comparan al “*claustrum*” con un director de orquesta. Verdaderamente, las conexiones neuroanatómicas del “*claustrum*” reúnen las características de un director; puede enlazar, juntar y coordinar las distintas regiones necesarias para la unidad de la conciencia.

El psicólogo cognitivo Vilaganur Ramachandran, un amigo y colega de Crick, describió la dedicación de éste al “*claustrum*” durante las últimas semanas de vida de Crick, como sigue: “Tres semanas antes de su muerte, lo visité en su casa en la Jolla. Tenía 88 años, un cáncer terminal y dolor, y estaba tratado con quimioterapia; sin embargo, había estado trabajando ininterrumpidamente en su último proyecto. Su gran escritorio, que ocupaba la mitad de la habitación, estaba cubierto de artículos, correspondencia, sobres, números recientes de “*Nature*”, un portátil (a pesar de su renuencia a las computadoras) y de libros recientes sobre neuroanatomía. Durante las dos horas que estuve allí, no hizo ninguna

mención de su enfermedad, sólo unas ideas al vuelo acerca de la base neuronal de la conciencia. Él estaba especialmente interesado en una estructura diminuta denominada “*claustrum*” que había sido largamente ignorada por las principales líneas eruditas. Cuando ya me iba dijo: “*Rama, yo pienso que el secreto de la conciencia cae en el “claustrum” ¿verdad? ¿Por qué sino debería esta estructura diminuta estar conectada a tantas áreas del cerebro?*” Y me hizo un disimulado guiño conspiratorio. Fue la última vez que lo vi”. Con independencia de lo poco que se sabe del “*claustrum*”, continuaba Crick, él buscaba iniciar un instituto que se centrara en su función. En particular, buscaba determinar si el “*claustrum*” se conmuta cuando una percepción inconsciente, subliminal de un estímulo dado por los órganos sensoriales de una persona se tornan una percepción consciente.

Un ejemplo de tal conmutación, que tenía intrigados a Crick y Koch, es el de la rivalidad binocular. En este caso dos imágenes, verbigracia barras horizontales y verticales se le presentan a una persona simultáneamente de tal manera que cada ojo sólo ve un conjunto de barras. La persona puede combinar ambas imágenes e informar que ve una parrilla, pero más comúnmente la persona verá primero una imagen, luego la otra, alternando las barras horizontales y verticales una y otra vez de un lado para otro espontáneamente.

Usando MRI, Eric Sumer y sus colegas en el College University de Londres, han identificado las áreas parietal y frontal del córtex como las regiones del cerebro que llegan a estar activas cuando una persona en atención consciente pasa de una imagen a la otra. Estas dos regiones juegan un papel especial en centrar la atención consciente sobre objetos en el espacio. A su vez, las regiones prefrontal y parietal posterior del córtex, parecen transmitir la decisión respecto a la imagen que está siendo mejorada para el sistema visual, que entonces, trae la imagen a la conciencia. Realmente, la gente con daños en el córtex prefrontal tiene problemas para conmutar de una imagen a la otra en situaciones de rivalidad binocular. Crick y Koch pueden argumentar que las áreas parietal y frontal del córtex están reclutadas por el “*claustrum*”, que conmuta la atención de un ojo al otro y unifica la imagen presentada a la conciencia intencional por cada ojo.

La interacción de la conciencia con los mecanismos cerebrales sería recíproca, pero le correspondería a la conciencia, en la cumbre de la escala jerárquica representada por partículas subatómicas, átomos, moléculas, células y circuitos cerebrales, desempeñar la

función organizadora y dirigir la actividad del conjunto de entidades subcelulares, celulares y supracelulares que forman el cerebro. **Esta teoría de la conciencia como propiedad emergente incluye el fenómeno en el ámbito de las ciencias cognoscitivas y plantea la hipótesis de que la conciencia interactúa con los circuitos cerebrales ejerciendo una influencia causal activa sobre ellos.**

Crick y Koch, por su parte, creen que la conciencia es una propiedad específica del cerebro humano y que no es posible tener una comprensión de la misma o de otros fenómenos mentales si se trata el cerebro como una caja negra cuya estructura interna se desconoce. En otros términos, Crick y Koch niegan cualquier posibilidad al conductismo para explicar la conciencia.

En resumen, como lo señaló Chalmers (Chalmers, 1996) los investigadores usan la palabra “consciencia” o “conciencia” de muchas formas diferentes. Por ello, cuando menos es conveniente diferenciar los problemas que caen bajo ese epígrafe. Para ello conviene distinguir entre las dos clases de problemas siguientes:

2.6.1. Problemas fáciles

A pesar de su adjetivo, estos problemas de fáciles no tienen nada y no son en absoluto triviales, pues encierran en sí mismos la complejidad propia de los problemas biológicos y psicológicos. Entre ellos figuran los siguientes:

- a) ¿Cómo discierne el sujeto entre un estímulo sensorial y otro y cómo reacciona ante ellos según lo pertinente en cada caso?
- b) ¿Cómo integra el cerebro la información que le llega de fuentes dispares y se sirve de ella para controlar el comportamiento?
- c) ¿Por qué pueden los individuos verbalizar sus estados internos?

Aunque estas preguntas están asociadas a la conciencia, todas se refieren a los mecanismos objetivos del sistema cognoscitivo. En consecuencia, se tiene la certeza y desde luego la esperanza de que la investigación en neurología y psicología cognitiva acaben por darle una respuesta satisfactoria a dichas cuestiones. De ahí el adjetivo de fáciles. El rasgo descriptivo que comparten estos problemas es que todos se refieren al

cómo se realiza una función cognoscitiva o del comportamiento. En definitiva, todos preguntan por la manera cómo el cerebro efectúa alguna tarea, en cuanto la neurofisiología especifica un mecanismo neurohormonal apropiado que muestra cómo se efectúan las funciones, los problemas fáciles se resuelven.

2.6.2. El Problema Arduo o Duro

Con estos adjetivos se califica, por el contrario, la cuestión de cómo los procesos físicos del cerebro dan lugar a la conciencia. En palabras de Crick y Koch: un explicación completa de la manera en que las experiencias subjetivas surgen de los procesos cerebrales. Es decir, se refiere al aspecto interior del pensamiento y la percepción, a la sensación que las cosas producen en el sujeto, como son verbigracia: el sufrimiento de una gran pena, el gozo de la felicidad o la cualidad meditativa de un momento en que una persona se “pierde” ensimismada en sus pensamientos. Estos fenómenos son los que plantean el verdadero misterio de la mente. Este problema va más allá de los problemas que se refieren a cómo se desempeñan las funciones. Aun cuando cada función del comportamiento o cognoscitiva relacionada con la conciencia se explica quedaría al fin un misterio: ¿Por qué la ejecución de esas funciones va acompañada de experiencia consciente?

Para ilustrar la diferencia entre unos y el otro, considérese el siguiente experimento mental concebido por el filósofo australiano Frank Jackson. Supóngase, que María, una neuróloga del siglo XXIII, es la más destacada experta del mundo en los procesos cerebrales a los que se debe la visión del color. Pero, lamentablemente, María ha pasado toda su vida en una habitación blanca y negra y nunca ha visto ningún color. Ella sabe todo lo que hay que saber acerca de los procesos físicos del cerebro: su biología, su estructura y su función. Con ese conocimiento sabe todo lo que hay que saber sobre los problemas fáciles: cómo distingue el cerebro entre estímulos, integra la información y produce comunicaciones verbales. Por su conocimiento de la visión del color, sabe cual es la correspondencia entre los nombres de los colores y las longitudes de onda del espectro de la luz visible. Pero aún así, hay algo, crucial, acerca de la visión del color que a María se le escapa: Cuál es la experiencia de un color, el rojo, por ejemplo. Esto quiere decir que se dan hechos relativos a la experiencia consciente que no se deducen de los hechos físicos concernientes al funcionamiento del cerebro. En realidad nadie tiene la

más remota idea de por qué la experiencia consciente acompaña a estos procesos físicos. Ya que la experiencia de un color es una representación interna del hecho objetivo (cuerpos que reflejan luz visible con determinada longitud de onda mientras absorben el resto de las ondas) y no el hecho mismo. Esto se ve claramente en el hecho que no todos los seres perciben igual las longitudes de onda, Los perros ven la realidad en diferentes tonos de gris ciertas experiencias parecen indicar que las abejas sólo perciben los tonos amarillos, mientras que los felinos perciben el infrarrojo que está por debajo del rango de la luz visible para el ser humano. ¿Por qué cuando el cerebro procesa la luz de una longitud de onda tiene la experiencia del verde? ¿Por qué siquiera se tiene una experiencia? ¿No podría un autómatas inconsciente haber efectuado las mismas tareas igual de bien?

Chalmers, a continuación, argumenta acerca de la dificultad que, para explicar el problema arduo, presentan las teorías físicas. Para él, estas teorías son más adecuadas para la tarea de explicar por qué ciertos sistemas presentan cierta estructura física y cómo realizar ciertas funciones. La mayoría de los problemas científicos tienen esta forma. Por ejemplo, para explicar la vida se ha de descubrir la reproducción, adaptación y metabolismo de un sistema físico. Pero el de la conciencia es un tipo de problema completamente diferente, pues va más allá de la explicación de la estructura y la función. Por supuesto, la neurofisiología algún día descubrirá la naturaleza del correlato neuronal de la conciencia, los procesos cerebrales asociados más directamente con la experiencia consciente. Incluso, se podría establecer la correspondencia entre procesos específicos del cerebro y componentes de la experiencia relacionados con ellos. Ahora bien, mientras no se sepa porque esos procesos dan lugar a la experiencia consciente, no se habrá cruzado lo que el filósofo Joseph Levine llama la laguna explicativa entre los procesos físicos y la conciencia. La teoría propuesta cubre ese hecho.

Para Crick y Koch (1996) la mejor forma de abordar el problema de la conciencia es concentrarse en la búsqueda y eventual hallazgo de correlatos neuronales de la conciencia; es decir, de los procesos del cerebro responsables directos de la conciencia. Para ellos, el problema arduo se puede descomponer o mejor dicho plantear, mediante las siguientes cuestiones:

a) ¿Por qué siquiera se experimenta algo?

- b) ¿Qué lleva a una experiencia consciente concreta, como la “azulidad” del azul?
- c) ¿Por qué algunos aspectos de la experiencia subjetiva no pueden ser comunicados a otras personas? En otras palabras, ¿por qué la experiencia no es transferible?

Para responder a la última pregunta, usan lo que denominan representación neuronal explícita. Para explicar lo que significa “explícito” en ese contexto, ponen el siguiente ejemplo: en respuesta a la imagen del rostro, las células ganglionares se disparan por toda la retina, en muy buena medida como los píxeles de una pantalla de televisión, para generar una representación implícita de la cara. Simultáneamente, pueden responder a muchísimos otros rasgos de la imagen, como las sombras, las líneas, la iluminación desigual, etc. Por el contrario, ciertas neuronas situadas en un lugar elevado de la jerarquía de la corteza visual, responden principalmente al rostro o incluso al rostro visto desde un ángulo determinado. Estas neuronas le sirven al cerebro para representar al rostro de una forma explícita. Su pérdida, debida a un accidente cerebro-vascular o a otro tipo de lesión, lleva a la *prosopagnosia*; es decir, a la incapacidad de un individuo de reconocer conscientemente rostros familiares, incluso el suyo propio, aunque todavía pueda reconocer los rostros como tales. Análogamente, los daños que sufran otras partes de la corteza visual pueden hacer que uno pierda la capacidad de experimentar el color, pero sin dejar de ver matices del blanco y del negro aún cuando no haya ningún defecto en los receptores del color en el ojo.

Algo similar ocurre con el lenguaje. Se piensa mucho más de lo que se habla. De acuerdo al filósofo Carlos Vaz Ferreira, “no todo el psiqueo puede reducirse a psiqueo formulable”. Considerando el psiqueo y aún el pensamiento en sus primeras fases como representación implícita de conceptos, luego el lenguaje sería una representación explícita. Análogamente, a los casos anteriores, existen casos en los cuales, debido a una lesión cerebral, una persona pierde la capacidad de traducir los conceptos en palabras (esto se denomina afasia); sin embargo, puede comprobarse que las demás capacidades de razonamiento de la persona permanecen intactas. Más aún, debido a la plasticidad neuronal, en muchos casos se observan niveles sorprendentes de recuperación, lo que muestra la capacidad del cerebro de asignar otras neuronas a cumplir las funciones faltantes.

En cada área, la información visual se re-codifica, normalmente de forma semi-jerárquica. Las células ganglionares retinales reaccionan a una mancha de luz. Las neuronas de la corteza visual primaria son las indicadas para responder a las líneas o bordes. Las neuronas de un nivel superior podrían preferir un contorno móvil. Más arriba aún están las que reaccionan ante los rostros y otros objetos familiares. En la cima residen las que se proyectan hacia las estructuras premotrices y motrices del cerebro, donde excitan las neuronas (*llamadas neuronas de proyección*) que ponen en marcha determinadas acciones tales como hablar o esquivar un coche.

Chalmers, Crick y Koch creen que los aspectos subjetivos de una experiencia tienen que estar estrechamente relacionados con el disparo de las neuronas correspondientes a esos aspectos (los correlatos neuronales). Para Crick y Koch la razón por la que María no sabe que es ver un color es que nunca ha tenido una representación neuronal explícita de uno en su cerebro, sólo la ha tenido de las palabras e ideas asociadas a los colores. Es evidente que la experiencia de un determinado color no es un comportamiento que resulte del aprendizaje (aún María con su absoluta inexperiencia en colores, si algún día se le mostrara uno tendría la experiencia habitual de ese color desde la mismísima primera vez que lo observe). Sin embargo, el aprendizaje modifica y hace más compleja la experiencia del color. Los seres humanos suelen asociar los colores con determinadas cosas, verbigracia, el rojo se considera un color “agresivo” posiblemente porque se asocia con la sangre, el negro se asocia con lo negativo, porque en la oscuridad la inseguridad aumenta, etc. Las asociaciones inciden enormemente en la experiencia y esto es algo que los pintores explotan habitualmente (claroscuros, colores que sugieren estados de ánimo del autor, limitación de ciertos lienzos o aún de la entera obra del autor a determinados colores, etc.). Algo similar ocurre con la experiencia del olfato. Es común oír a muchas personas hablar de aromas dulces como la canela o “amargos” como el vinagre. Sin embargo, los aromas no pueden tener esas cualidades que son propias del sentido del gusto, sino que son el resultado de las asociaciones que hacen más compleja, con el aprendizaje, la experiencia olfativa.

Para describir una experiencia visual subjetiva, la información debe transmitirse al área del cerebro que se encarga de la reacción motriz donde queda disponible para la verbalización u otras acciones. Esta transmisión supone siempre una re-codificación de la información de modo que la información explícita expresada por las neuronas motrices

guarda relación con la información implícita, pero sin ser idéntica a ella, expresada por el disparo de las neuronas asociadas la experiencia del color, en algún nivel de la jerarquía visual. No es posible, pues, comunicar con palabras e ideas la naturaleza exacta de una experiencia subjetiva. Sin embargo, es posible comunicar una diferencia entre experiencias subjetivas tal como, verbigracia, distinguir el rojo del naranja, y ello es posible porque una diferencia en un área cortical visual de alto nivel seguirá estando asociada a una diferencia en las zonas motrices. En consecuencia, nunca se podrá explicar a otras personas la naturaleza de cualquier experiencia consciente, sino sólo su relación con otras experiencias.

Este supuesto se basa en la idea de que la conciencia humana no es más que un conjunto de eventos físicos. Pero ésta también puede, y el doctorando añade debe, cuestionarse. De hecho, la existencia de determinados eventos físicos es lo que hace posible la conducta humana, verbigracia, la acción de ir al frutero, pero lo que la constituye es una serie de relaciones, por ejemplo racionales, entreveradas con otras acciones o estados mentales distintos de las relaciones entre eventos físicos. Esta línea argumentativa, tiene importantes consecuencias para el propio planteamiento del problema mente-cuerpo. Por una parte, parece directamente relevante para el problema de la causalidad mental, tal como éste se plantea generalmente, pues en la medida en que las explicaciones neurofisiológicas y las mentales no compiten entre sí, no habría un problema de sobredeterminación causal. Por otro lado, también parece ir más allá, pues al aceptarse relaciones *sui generis* en un ámbito distinto del de los eventos físicos, se abriría la puerta a una explicación del lugar de lo mental en el mundo al margen de la manera habitual de entender las relaciones causales en cuestión. Se rechazará así uno de los supuestos principales de la filosofía de la mente del último tercio del siglo XX: explicar el lugar de lo mental en el mundo consiste en explicar la causalidad mental.

2.7. Resultados obtenidos a partir del estado de la cuestión

2.7.1. Tipología de los investigadores de la conciencia

Como resumen del estado de la cuestión, a continuación se presenta, por una parte, una tipología de los distintos enfoques para abordar, o no, el problema de la conciencia y, por

otra, lo que en opinión del doctorando, constituye las cuestiones fundamentales en lo que concierne al problema de la conciencia.

1. **Misterianos o Mistéricos.** Unos, como Nagel, Jerry Fodor, Colin McGinn o David Chalmers, aseguran que el problema de la conciencia es completamente irresoluble: *Nadie tienen la más mínima idea de cómo algo material puede ser consciente.* Otros, como Eccles, Roger Penrose o el filósofo Bill Seager, recurren a los fundamentos de la mecánica cuántica para adherirse a teorías, similares al antiguo idealismo, que sostienen que únicamente lo mental es real, o el panpsiquismo, según el cual todo el Universo posee cualidades mentales.

El hecho de que no se sea consciente de los procesos neuronales conducentes a una decisión no quiere decir, según Crick, que los procesos no se produzcan. Algunos filósofos han cuestionado que haya algún modelo puramente materialista que pueda explicar realmente la conciencia. El filósofo materialista Owen Flanagan bautizó a estos filósofos como los nuevos “mysterians”, en analogía con el grupo de rock de los sesenta “Question Mark and the Mysterians”, intérpretes de la conocida canción “96 Tears”. Es decir, el término acuñado por Flanagan (Flanagan, 1991) sirvió para catalogar a aquellos que sugerían que la conciencia podría no llegar a explicarse del todo en términos científicos convencionales; ni, para el caso, en ningún otro término. A éstos les denominaba “misterianos”. Por cierto, el término no era original de esta banda, sino de la película japonesa “Los Misterianos” acerca de una invasión de alienígenas.

Para defender su postura, los “misterianos” tomaron prestada una idea de Noam Chomsky, pionero del enfoque genético y modular de la mente. Era ésta (Chomsky, 1988) la distinción que haría el lingüista entre “problemas”, que parecen solubles, al menos en principio, mediante métodos científicos convencionales y “misterios”, que parecen insolubles incluso en principio. Todo ello debido a la capacidad cognoscitiva de cada uno. Chomsky consideraba misterios a la conciencia, el libre albedrío y otros aspectos relacionados con la mente.

Fue el filósofo Thomas Nagel quien ofreció la, tal vez, más clara formulación del punto de vista “misteriano” (¿empresa oximorónica?) en su famoso ensayo “What is

it Like to Be a Bat?” (Nagel, 1979). Nagel asumió que esa experiencia subjetiva es un atributo fundamental de los humanos y de muchos animales de nivel elevado, como, verbigracia, los murciélagos. Y dice Nagel: *Pero, varíe como varíe la forma, el hecho de que un organismo tenga experiencia de alguna manera consciente significa, básicamente, que hay algo consistente en ser tal organismo.* Nagel sostenía que por mucho que se aprenda sobre la fisiología de los murciélagos, no se podrá nunca saber realmente a qué se parece ser uno de ellos, porque la ciencia no puede penetrar en el reino de la experiencia subjetiva. Nagel, a pesar de todo, es un misteriano débil pues defiende la posibilidad de que la filosofía y, o, la ciencia puedan un día revelar una manera natural de colmar el abismo que existe entre las teorías materialistas y la experiencia subjetiva.

Sin embargo, no sucede lo mismo con Colin McGinn, un misteriano fuerte, que cree que la mayoría de las cuestiones filosóficas referentes a la mente son insolubles por estar situadas más allá de las capacidades cognoscitivas humanas. Estas fueron sus palabras: *La mente puede sencillamente no ser lo suficientemente grande para comprender la mente* (Zohar, 1990). McGinn (McGinn, 1989) niega rotundamente la capacidad de resolver el misterio y dice: *de algún modo, tenemos la sensación de que el agua del cerebro físico se convirtió en el vino de la conciencia pero no sabemos nada en absoluto de la naturaleza de esta conversión. Las transmisiones neuronales simplemente parecen un tipo impropio de misterio con los que traer la conciencia al mundo.* Uno de los problemas que, para él, está más allá de esos límites es el problema mente-cuerpo. McGinn define su punto de vista como superior a la que él llama postura “**eliminacionista**”, según la cual el problema mente-cuerpo no es en absoluto ningún problema. Es perfectamente posible, según McGinn, que los científicos consigan formular una teoría de la mente capaz de vaticinar el resultado de los experimentos con gran precisión y de aportar una gran riqueza de beneficios médicos. Pero una teoría eficaz no es necesariamente una teoría comprensible. *No hay ninguna razón de peso por la que parte de la mente humana no pueda desarrollar un formalismo con esas notables propiedades vaticinadoras, pero no podemos dar sentido a ese formalismo en la parte de nuestra mente que entiende las cosas. Así, en el caso de la conciencia puede ser que se llegue a una teoría que sea análoga, en ese aspecto, a la teoría cuántica, una teoría que sea realmente una*

buena teoría de la conciencia; pero entonces no seríamos capaces de interpretarla ni de comprenderla. (McGinn, 1991).

También un materialista y reduccionista como Christof Koch sostuvo también una postura misteriana tanto en la conferencia de Tucson como en un artículo que publicó en *Nature* (Koch, 1996). Allí y entonces reconoció que la teoría neuronal de la conciencia tal vez no lograra resolver viejos enigmas filosóficos como el problema mente-cuerpo y la cuestión del libre albedrío, y ello porque esos arcanos superaban sencillamente el radio de acción de la ciencia. Para avalar su postura, echó mano de la cita final de la monografía oracular de Ludwig Wittgenstein (Wittgenstein, 1957): *De lo que no se puede hablar mejor callar*, y de otra cita de otro gigante del siglo XX, Harry el Sucio, el policía duro protagonizado por Clint Eastwood: *Un científico debe conocer sus limitaciones, amigo*.

Daniel G. Dennett acusa a los misterianos de vitalistas. Dennett concibe la conciencia, igual que Minsky, como pensamiento abstracto de alto nivel. A Dennett se le conoce como el principal defensor del modelo informático de la mente. Esta postura le ha llevado a chocar frontalmente con filósofos como John Searle que mantienen que los aspectos más importantes de la conciencia, la intencionalidad y la cualidad objetiva, nunca podrán programarse. Dennett ejemplifica perfectamente lo que McGinn denomina postura eliminacionista pues sostiene (Dennett, 1995) que la conciencia y la sensación de poseer un yo unificado, era una ilusión fruto de la interacción de muchos “subprogramas” distintos que funcionan ininterrumpidamente en el hardware del cerebro. Y, sin embargo, un poco como McGinn, cuestiona la comprensión de la ciencia actual al decir: *En la ciencia moderna “late una curiosa paradoja”. Una de las cosas que hacen que la ciencia progrese tan rápidamente en estos tiempos es cierta tendencia que aleja a la ciencia de la comprensión humana. Cuando se pasa de tratar de modelar las cosas con ecuaciones elegantes a realizar grandes simulaciones informáticas... podremos terminar con un modelo que modele exquisitamente la naturaleza, los fenómenos en los que estamos interesados, pero que no comprendamos. Es decir que no comprendamos de la manera como comprendíamos los modelos antiguamente.*

2. **Evolucionistas o empíricos.** Como Steven Pinker, quienes sostienen que la conciencia puede explicarse en términos funcionales, neurológicos y cognoscitivos. Para los empíricos, los místicos caen en el dualismo cartesiano. Ellos propugnan que la conciencia no es más que una actividad muy evolucionada del sistema nervioso y hormonal y más en concreto del cerebro. Contradictoriamente, entre los misterianos se encuentra Steven Pinker quien, al final de su libro “How the Mind Works” (Pinker, 1997), concluye diciendo que la conciencia, el libre albedrío, el ego y otros enigmas relacionados con la mente, son probablemente irresolubles.
3. **Campos electromagnéticos.** Esta hipótesis, desarrollada por Johnjoe McFadden y publicada en Journal of Consciousness, afirma que la actividad sincrónica de las neuronas durante la vigilia genera una serie de campos electromagnéticos que serían algo así como el soporte de la conciencia. Además de contener todo tipo de información neuronal, estos campos electromagnéticos llegarían a influir en cualquiera de las acciones humanas, incluido el libre albedrío, al provocar la excitación de algunas neuronas y la inhibición de otras.
4. **Funcionalismo. La teoría computacional de la mente.** Desarrollado y defendido por Crick y Koch, y Dennett. Éste afirma que la conciencia se genera exactamente en la parte posterior del córtex cerebral, naciendo de reacciones bioquímicas del cerebro. De modo que todo el comportamiento mental es íntegramente explicado por la interacción de las células cerebrales. De sus experimentos, con pacientes con lesiones cerebrales, concluye que distintas partes del cerebro pueden fundirse unas con otras para crear un sentido de la conciencia.

La concepción sobreyectiva de las relaciones mente-cerebro se denominan “funcionalismo”. Éste mantiene el análisis causal de los estados mentales, sustituyendo la relación de identidad por la de realización, introduciendo una importante distinción entre dos niveles: el funcional y el material. El primero es el propiamente mental, los estados mentales son funciones caracterizadas por una red de relaciones causales entre “entradas, salidas” u otros estados mentales. Es decir, el nivel funcional guarda algún tipo de relación con el nivel material; a saber, la realización o implementación pero son niveles distintos. De este modo, el funcionalismo defiende que, de hecho, los estados mentales se realizan en el sistema

neurofisiológico humano, pero que bien podrían implementarse en sistemas, materiales o no, distintos.

Esta concepción funcionalista tuvo su exponente máximo en el funcionalismo computacional de Hilary Putnam de acuerdo con el siguiente dicho: “La mente es al cerebro lo que el software es al hardware de un computador”, Es decir, la mente no es más que un conjunto de programas que se implementa físicamente en el cerebro humano y teóricamente, en cualquier otro sistema con la complejidad suficiente. Los procesos mentales son un conjunto de procesos algorítmicos de transmisión de información codificada simbólicamente. La objeción a este enfoque es que, según la conocida tesis de Jerry Fodor, la información se computa sintácticamente, es “el lenguaje de los pensamientos”, vino del conocidísimo argumento de la habitación china de Searle. Es decir, la mente no sólo es sintaxis, sino que también es semántica y pragmática.

Retomando la cuestión principal de la causalidad mental ¿podría el funcionalismo explicar la causalidad mental sin negar la clausura causal del mundo físico y sin caer en la sobredeterminación causal? La dificultad radica en la distinción crucial entre los niveles: funcional y neurofisiológico. En efecto, por una parte, la clausura causal del mundo físico haría causalmente innecesarias las propiedades del nivel funcional, el nivel propiamente mental según el funcionalista. Por otra parte, la insistencia en la eficacia causal de las propiedades del nivel funcional, no sólo las del nivel neurofisiológico, implicaría un nuevo problema de sobredeterminación causal. Los defensores del fisicalismo reduccionista han propuesto evitar este indeseable dilema mediante la identidad entre causas mentales, caracterizadas, eso sí, por su rol causal, y causas neurofisiológicas. Simultáneamente, evitarían la base que sustenta la objeción de la realizabilidad múltiple relativizando dichas identidades a especies o algún otro tipo de estructura. Así lo proponen David Lewis y Jaegwon Kim (Kim, 1981) Los detractores de esta propuesta aducen que, si con eso se consiguiera explicar la causalidad mental, sería a costa de renunciar a lo propiamente mental, pues las causas mentales son idénticas a; esto es, reducibles a, causas neurofisiológicas. Por ello, desde mediados de los 80, se abrió camino la alternativa consistente en defender algún tipo de dependencia, distinto de la identidad, entre el nivel mental y el neurofisiológico. Esta relación de dependencia sin identidad se

anticipó en el tiempo a la noción de “sobreveniencia”. Lo mental no es idéntico sino que sobreviene a lo neurofisiológico. El núcleo de esta relación de sobreveniencia psicofísica se halla en que, si bien las propiedades mentales son distintas de las neurofisiológicas, aquellas dependen de éstas porque no puede haber diferencias en cuanto a las propiedades mentales sin diferencias en cuanto a las propiedades neurofisiológicas o, dicho de otro modo, la indiscernibilidad en cuanto a las propiedades neurofisiológicas implica indiscernibilidad en cuanto a las propiedades mentales.

En la relación de “sobreveniencia” hay dos aspectos fundamentales: Uno, es compatible con la distinción funcionalista entre niveles. Dos, busca dar cuenta de las intuiciones materialistas que se hallan en la base de los temas de la identidad, mediante la noción de indiscernibilidad. Una vez que se fijan las propiedades neurofisiológicas relevantes, quedan fijadas *ipso facto* las mentales, en cualquier situación posible. Según los defensores de la sobreveniencia, para ser materialista no hace falta defender la identidad. Entre esos dos tipos de propuestas, basta con la indiscernibilidad.

Los antirreduccionistas afirman que el marco del reduccionismo fisicalista es inaceptable, porque implica que toda explicación causal, no sólo las mentales de la conducta humana, es en realidad una explicación física. Es decir, no hay explicaciones genuinas causales en química, biología o ciencias sociales, y habrían de recurrir a sus propias causas. Pero aceptada la clausura causal del mundo físico, eso sería caer en la sobredeterminación causal. Para obviar esto, algunos como Jerry Fodor han tratado de articular un modelo de explicación causal que no implique renunciar al genuino poder explicativo de las distintas ciencias de la física.

5. **Emergentistas.** Para Kim la noción de sobreveniencia no es la solución antirreduccionista al problema mente-cuerpo porque es parte del problema. Hoy, desde los noventa, está en baja una respuesta al problema que niega uno de los supuestos comunes a reduccionistas y antirreduccionistas, **los llamados “emergentistas” entre los que se incluye John Searle. Estos defienden que las propiedades mentales emergentes a partir de las neurofisiológicas tienen poderes causales propios. Con esto se negaría la clausura causal del mundo**

físico: en la explicación causal de algunos fenómenos físicos aparecería algunas propiedades distintas de las físicas. A los emergentistas no les importa, pagar ese precio por obtener una respuesta a la causalidad entre tipos distintos de propiedades, pero sería tan mágica como la postulación cartesiana de relaciones causales entre distintos tipos de sustancias.

Hay, además de éstas, otras formas de replantearse el problema mente-cuerpo, de modo que su solución pasaría con confrontarse con uno de los dos extremos siguientes: o el materialismo fisicalista que reduce los poderes causales de lo mental, o el dualismo que postula más que explica la causalidad mental sobre el mundo: Peter Strawson (Strawson, 1959) y John McDowell (McDowell, 1985) han replanteado el problema como sigue: la existencia de causas mentales de la conducta humana es problemática porque ésta se puede explicar siempre en virtud de causas neurofisiológicas. Sin embargo, ¿por qué ha de asumirse que las explicaciones mentales y las neurofisiológicas buscan dar cuenta del mismo fenómeno?

6. **Fenomenología.** Varela y otros que estudian la conciencia desde la perspectiva del sujeto. Según esta escuela, sólo la unión de la neurociencia y la psicología conductista resuelven el problema.

	Científicos	Informáticos	Humanistas-Lingüistas	Filósofos	Psicólogos Cognoscitivos
Misterianos	Levi Montalcini, Huxley, Sherrington, Planck, Gardner, Kandel		Chomsky, Schalz, Fernández	Platón, Hamilton, Descartes, Leibnitz, Thomson, Stent, Locke, McGinn, Chalmers, Vaccero, Reid, Fodor, Spinoza, Nagel	Pinker
Funcionalistas Conceptualistas	Taylor, Wheeler, Bemser	Turing, Minsky, Kurzweil		James, Dennett, Lucas, Putnam, Churchland, Chalmers	Fodor

Cuánticos	Penrose, Zohar, Pribram, Hameroff, Bohr, Bridgman, Delbrück, Eccles, Beck, Seager, Josephson			Hume	Titchener
Reduccionistas	Crick, Koch, Rguez- Delgado, Gould, Stent, Schrödinger, Goldman- Rakic, Penfield Sperry			Searle, Russell, Wittgenstein. Demócrito, Bacon, Flanagan, Ryle	
Holistas Jerárquicos	Bertalanffy, Lorenz, Edelman, Sacks, Gell-Mann, Kauffman, Varela			Anaxágoras, Hartmann	
Caóticos	Freeman, Clarck				

Tabla 2.1. Cuadro Sinóptico de los “Concienciólogos”, según profesión y enfoque.

Si la conciencia evolucionó con el tiempo, entonces la capacidad de engrenrarla debe haber estado presente en el Universo desde el principio. Esto lleva a una especie de pansiquismo, del que participa el propio Chalmers; es decir, la creencia de que la propia materia física, hasta el nivel incluido de las partículas subatómicas, posee la conciencia en potencia. Para evitar esta conclusión, los neurofilósofos Paul y Patricia Churchland (Churchland, 1990) niegan la realidad de los estados mentales. No se hace ciencia si lo que se está haciendo en lo concerniente a la conciencia no acaba casando con lo que sucede en el mundo real de manera demostrable; es decir, algo que sea verificable. Todo lo referente a esta tipología aparece resumido en la tabla 2.1.

3. EL PROBLEMA ABORDADO Y LA SOLUCIÓN

Hasta ahora el problema de la mente se abordaba en un marco conceptual que consideraba la ontología del Universo formada sólo por energía y materia. Esto dio lugar al dualismo entre cuerpo y espíritu, mente y cerebro, etc. Sin embargo, al completar la ontología con la información se supera esta dicotomía y la mente puede abordarse de forma natural científicamente. Naturalmente, este marco no es suficiente, por eso también se van a considerar los constructos que permiten considerar, científica y tecnológicamente, la información y su tratamiento: los holones e informones. Finalmente, en dicho marco, y con tales constructos, y ésta es la parte original de esta tesis, se verá como una de las características más importantes de la conjunción materia-energía-información; esto es, la emergencia de propiedades, es algo que, con relativa facilidad se consigue. Como además la tesis que aquí se propugna es que las propiedades mentales son emergentes, ergo el problema de la conciencia puede y debe abordarse dentro de este marco y con el enfoque emergentista.

3.1. El marco conceptual: La ontología del universo. El Modelo EMI

3.1.1. La información el tercer constructo del Universo

Que la información, junto con la materia y la energía, es uno de los tres constructos que componen la ontología del Universo, es algo de lo que ya muy pocos científicos dudan. Sin embargo, queda mucho trabajo por hacer para establecer, igual que desde Einstein ya existen entre materia y energía, las interrelaciones entre información y energía y entre información y materia. Aquí ahora se muestra una de ellas: la de la energía y la información. Ésta es una de las aportaciones de esta tesis.

Como lo muestra reiteradamente la historia de la ciencia, la percepción del mundo es producto de la experiencia histórica del ser humano. No fue hasta que se tuvo una experiencia relevante con las máquinas de medir el tiempo, los relojes, cuando se desarrolló el concepto de tiempo. Hasta ese momento, el tiempo tendía a verse de una manera más bien difusa y como algo cíclico más que lineal. A partir de la aparición de los relojes, con el reloj de péndulo de Christian Huygens, que permitía definir el tiempo en términos de unidades pequeñas, iguales y repetitivas, fue cuando el tiempo

exhibió propiedades de homogeneidad y continuidad y con un sentido de que pasaba minuto a minuto o segundo a segundo. De modo similar, no fue hasta que se tuvo una experiencia relevante y significativa, con la creación de dispositivos que generaban energía, cuando ésta comenzó a ser abstraída del concepto de materia. Hasta ese momento, ambos conceptos, materia y energía, eran indistintos. Un objeto podía estar caliente o frío del mismo modo que podía ser duro o blando. Ambas eran propiedades específicas de la materia. Fue sólo después de que aparecieron las máquinas *energéticas* lo que obligó a clarificar el concepto y la definición de energía.

Actualmente, se está en una encrucijada histórica similar. Hasta los años 40 del siglo XX, prácticamente no se tenía ninguna experiencia acerca de máquinas o artefactos de información. Fue con la aparición de las computadoras electrónicas y su uso generalizado y ubicuo, cuando se hizo una necesidad sentida el establecer y definir, con toda la precisión posible, el concepto de información y crear y desarrollar un marco teórico de la misma. Como lo señaló Gordon Scarrott (Scarrott, 1986) es necesaria una *ciencia de la información* que debería investigar las propiedades naturales de la información tales como: función, estructura, comportamiento dinámico, características, estadísticas, etc.

Si se hiciera una encuesta a nivel mundial de pregunta única: ¿de qué se compone el Universo? La inmensa mayoría respondería, eso sí de distintas formas más o menos equivalentes: de materia. Un número relativamente reducido de respuestas, sería, y esta vez con más precisión, de materia y energía. Y, finalmente, un número ínfimo de encuestados, tal vez, próximos al 1/10000, daría la respuesta, actualmente más acertada, siguiente: de materia, energía e información. Dentro de los que dan esta respuesta estarían científicos e ingenieros. La razón de esto, es que ambos colectivos son conscientes de que, verbigracia, los robots de una fábrica están contruidos de algún material, *materia*, trabajan mediante algún tipo de motor, *energía*, y sólo harán cosas útiles si tiene abundantes instrucciones, *información*, que le diga cómo actuar adecuadamente. Por su parte, los biólogos conocen a la perfección que un ribosoma de una célula se construye con aminoácidos, *materia*, y se alimenta con la *energía* generada por la conversión del ATP (adenosintrifosfato) en ADP (adenosindifosfato), pero sería incapaz de sintetizar proteínas sin la *información* suministrada por el ADN (ácido desoxirribonucleico) del núcleo celular.

El resultado de todo ello es que hasta hace no mucho todo el mundo estaba de acuerdo en que el mundo físico, Universo o la *fábrica del mundo*, que de estas tres formas se le denomina, se componía de *materia y energía*, que, de acuerdo con la famosa ecuación de Einstein (Einstein, 1905) (Einstein, 1907): $E = mc^2$, eran intercambiables. Hoy a estos dos componentes hay que añadirle un tercer elemento, la información. La información no es materia ni energía, pero resulta ser una propiedad de la realidad tan esencial como la materia y la energía. Un siglo de investigaciones ha puesto de manifiesto, fehacientemente, que la información desempeña una función esencial en los sistemas y procesos físicos.

Hoy una línea de pensamiento iniciada por John Archibald Wheeler, de la Universidad de Princeton, considera que el mundo físico está, fundamentalmente, hecho de información; la energía y la materia serían accesorias. En efecto, Wheeler, como consecuencia de las *paradojas* de la física cuántica, propuso un cambio paradigmático absolutamente revolucionario. Según él, la lección más profunda que ofrece la mecánica cuántica, es que, de alguna manera, los fenómenos físicos resultan definidos por las preguntas que se les plantean. *En cierto sentido*, afirmó, *es un Universo participativo*. Y añadía que puede que la base de la realidad no sea el cuanto, que, pese a su carácter elusivo, sigue siendo un fenómeno físico, sino el *bit*; es decir, la respuesta a una pregunta de sí o no. Siguiendo esta orientación, se está intentando reformular la física cuántica en términos de la teoría de la información. Ya se ha encontrado que el principio de indeterminación de Heisenberg, la dualidad onda-partícula y la no localidad, pueden formularse de manera más potente en su contexto. Mientras tanto, otros teóricos de otros campos están *conjurando* experimentos mentales que desvelen la clave del enigma de una vez por todas.

John Archibald Wheeler (Wheeler, 1994), una de las máximas autoridades en relatividad general, fue uno de los primeros físicos de categoría reconocida en afirmar que la realidad podía no ser completamente física; que, en cierto sentido, el cosmos podía ser un fenómeno participativo que exigía el acto de la observación y, por ende, la conciencia propiamente dicha. En la década de los sesenta, Wheeler contribuyó a popularizar este notorio principio antrópico. Esencialmente, dicho principio sostiene que el Universo debe ser como es porque si fuera de otro modo, el ser humano no estaría en él para observarlo. Wheeler (Wheeler, 1990) también llamó la atención acerca de ciertas relaciones

interesantísimas entre la física y la teoría de la comunicación de Claude Shannon (Shannon, 1963) que le llevaron a establecer que así como el edificio de la física se levanta sobre una entidad elemental e indivisible: el cuanto, definida por el acto de la observación, lo mismo ocurre también con la teoría de la comunicación, cuyo cuanto es el bit, una observación en forma de mensaje que ofrece una de dos opciones; a saber, cara o cruz, si o no, cerrado o abierto, verdadero o falso, cero o uno. Wheeler se percató aún más de la importancia de la información tras realizar un experimento mental que denominó de la “elección retardada” y que ponía de manifiesto la extrañeza del mundo cuántico. Dicho experimento es una variante del clásico experimento cuántico de las dos rendijas, que demuestra la naturaleza esquizofrénica de los fenómenos cuánticos. Cuando los electrones se dirigen hacia una barrera que contiene dos rendijas, actúan como ondas; es decir, pasan por ambas rendijas e inmediatamente forman un patrón de interferencia, producido por el solapamiento de las ondas al golpear a un detector en el lado alejado de la barrera. Ahora bien, si se cierra una rendija, cada vez, los electrones pasan a través de la rendija abierta como simples partículas y, en consecuencia, desaparece el patrón de interferencia. En el experimento de la elección retardada, el experimentador tiene la elección de dejar ambas rendijas abiertas o cerrar una de ellas “una vez que los electrones han franqueado ya la barrera”, con los mismos resultados. Sorprendentemente, los electrones parecen saber de antemano cual será la elección del experimentador para observarlos. Llevado a cabo el experimento en los primeros noventa, confirmó la predicción de Wheeler. Éste explicó el enigma con una nueva analogía, al parangonar el trabajo del experimentador con el de alguien que juega a hacer veinte preguntas en plan sorpresivo.

En abstracto, en el **problema de las veinte preguntas**, se trata de que alguien piense en un número ente 1 y $10^6 \approx 2^{20}$ y otro, mediante veinte preguntas, debe descubrir cuál es el número. La solución consiste en preguntar por sucesivas dicotomías. Así ¿está el número entre 1 y 50.000? ¿Está entre.....?. Es decir, reduciendo cada vez los números posibles a la mitad, y así sucesivamente. Finalmente, el número se obtiene en menos de $\log_2 10^6$ preguntas. Una variante más interesante consiste en que se pudiese mentir, una o dos veces; entonces ¿cuántas preguntas se necesitarían para obtener la respuesta correcta?. Claramente hacen falta más de n preguntas para averiguar uno de los 2^n objetos porque no se sabe cuando se dijo la mentira. En general, este problema no está resuelto, pero, lamentablemente, así es como sucede en la realidad. Este caso es análogo al siguiente:

Supóngase un juego, da igual un solitario que uno competitivo en el que se puede hacer trampa una o dos veces, sin que el juego se destruya. De hecho seguiría siendo un juego definido matemáticamente de manera completamente precisa, aunque distinto. Simplemente se habría enriquecido un poco, se habría hecho más general. Sin embargo, si uno toma un sistema de axiomas y se permite la adición de una o dos proposiciones falsas, el resultado carece automáticamente de sentido, ya que una vez que se tiene una proposición falsa se puede deducir lo que se quiera. ¿Donde está la diferencia? Probablemente resida en el hecho de que en el juego sólo se permite un cierto tipo de movimientos, mientras que en matemáticas, una vez que se introduce una afirmación incorrecta uno puede inmediatamente obtener la proposición, $0 = 1$. Entonces tiene que haber una manera de generalizar el “juego” de las matemáticas de modo que se puedan cometer unos pocos errores y en lugar de obtener un sin sentido absoluto, simplemente se pueda obtener un sistema más amplio. Wheeler, en el contexto de la mecánica cuántica aderezada con la teoría de la información, dio la siguiente variante del juego: En esta variante del viejo juego de las veinte preguntas, una persona sale de una habitación, mientras el resto del grupo, o eso cree al menos el excluido, selecciona a una persona, lugar o cosa. El excluido regresa entonces a la habitación y trata de adivinar qué tienen en mente los demás haciendo una serie de preguntas a las que sólo se puede contestar con un sí o con un no. Sin que lo sepa el adivinador, el grupo ha decidido hacerle trampa pues la primera persona preguntada pensará en un objeto sólo “después de que” el interrogador haya hecho la pregunta. Los demás harán lo mismo, dando una respuesta que sea compatible no sólo con la pregunta inmediata, sino también con todas las anteriores. Como explicó Wheeler, *la palabra no estaba en la habitación cuando entró el interrogador aunque creía que estaba*. Del mismo modo, el electrón antes de que el físico decida cómo observarlo no es ni onda ni partícula. En cierto sentido es irreal; existe en un limbo indeterminado. “Hasta que no se hace la pregunta no se consigue nada”. La situación no puede declararse hasta que no se hace la pregunta. Pero el hacer una pregunta concreta impide y excluye hacer otras. Wheeler ha condensado estas ideas en una frase: *it from bit*. Cada “it”; esto es, cada partícula, cada campo de fuerza, inclusive el continuo espacio-tiempo propiamente dicho, extrae por entero su función, su significado y su mismísima existencia, aunque de manera indirecta en algunos contextos, “from” respuestas obtenidas con aparatos a preguntas de tipo si o no, elecciones binarias o bits. Wheeler decía que *a medida que aumenta la isla de nuestro conocimiento aumenta*

también el litoral de nuestra ignorancia. Sin embargo, también estaba convencido de que los humanos encontrarían algún día “La Respuesta”.

En este trabajo se va a presentar un modelo del mundo basado en sus tres constructos: Energía, Masa e Información. Basándose en los trabajos de Tom Stonier (Stonier, 1990) y Yingxu Wang (Wang, 2008) y de Pazos y colegas. Se van a presentar, además, y es lo importante de este trabajo, ejemplos, dos cualitativos y dos cuantitativos, de las relaciones entre energía e información.

3.1.2. El Modelo EMI

Tom Stonier (Stonier, 1990) señaló cualitativamente relaciones entre energía, materia e información. El calor constituye pura energía interactuando con la materia. La aplicación de calor, por si misma, no aporta ninguna entrada de información, sino todo lo contrario. Un incremento de calor provoca que moléculas u otras partículas se muevan más aleatoriamente. Ello causa que un cristal se disuelva, el hielo se licue, luego el líquido se vaporice. En cada etapa, el sistema está menos organizado. La aplicación de calor a un sistema, produce una aleatorización de sus componentes, y ocasiona desorden en el Universo. La recíproca, retirar calor de un sistema, tal como condensar un gas, o congelar un líquido, produce un incremento de la organización. Tales procesos de enfriamiento se asocian, consiguientemente, con un incremento de información. Por lo tanto, se puede establecer lo siguiente: *El calor es el producto de la “energía” interactuando con la materia, y la estructura representa el producto de la “información” interactuando con la materia.* La aplicación de energía, expresada ella misma como calor, causa que las partículas, moléculas, *plasmones*, etc., vibren y se muevan aleatoriamente. Por el contrario, la aplicación de información causa que las partículas se unan en patrones determinados y movimientos ordenados. En este sentido, puede considerarse al calor como la antítesis de la organización.

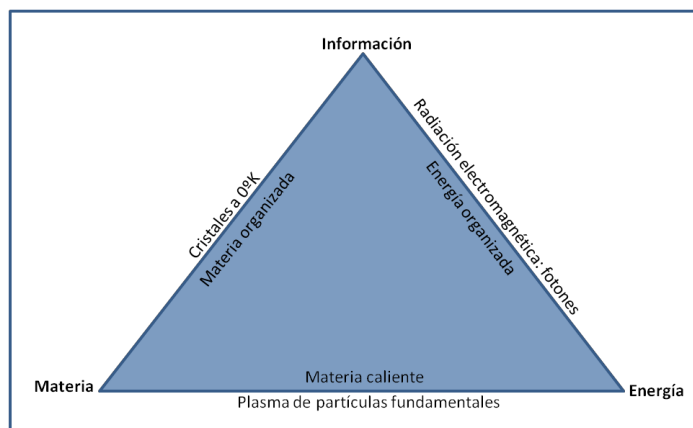


Figura 3.1. Relaciones entre materia, energía e información, según Stonier.

Ahora bien, si el calor es la antítesis de la organización, y, por implicación, la energía la antítesis de la información, eso no excluye la posibilidad de que la energía y la información puedan interactuar para proporcionar una combinación que pueda verse como *información energizada* o, alternativamente, *energía estructurada*. La *energía* y la *información* no deben verse como los opuestos de un sistema bipolar, antes bien deben considerarse de acuerdo con el modelo EMI y, tal y como se ve en la figura 3.1, como dos vértices de un triángulo, con la materia como tercer vértice. Dicha figura representa un modelo conceptual que debería definir los límites del mundo físico; siguiendo los tres lados de un triángulo, los extremos, en términos de los fenómenos siguientes:

- a) Una combinación de pura energía y materia careciendo de información, debería comprender el plasma de las partículas fundamentales.
- b) Una combinación de materia e información pura, careciendo de energía, debería ejemplarizarse por un cristal a 0° K.
- c) Una combinación de información y energía, careciendo de materia, consistiría en partículas sin masa; verbigracia, los fotones, viajando a través del espacio desprovisto de materia.

Por otra parte, las medidas de espacio y tiempo establecen información acerca de la organización y distribución de la materia y la energía. La información contenida en un sistema físico es directamente proporcional al espacio que ocupa, en tanto que el tiempo está inversamente relacionado con la información. La idea de que un incremento de tiempo entre dos eventos implique que el sistema tiene menos información, no debería

confundirse con el contenido de información estructural de un sistema que perdura a lo largo del tiempo. Obviamente, un sistema que sobrevive a los estragos del tiempo contiene más información que uno que se desintegra.

Hay otras dos razones que hacen que se considere que la información es más que algo que se manipula dentro de la cabeza de los seres humanos. La primera, es la transmisión y manejo de la información en forma física: teléfono, telégrafo, ondas de radio, etc. La segunda, fue la demostración inequívoca de que el ADN (ácido desoxirribonucleico) y el ARN (ácido ribonucleico) vehiculan información genética. Es decir, cómo una sustancia física, el ADN, vehicula información. Y esto desde hace más de mil millones de años, en tanto que el cerebro humano, sólo lleva manejando información desde hace cinco millones de años.

La cuestión que ahora se plantea es la siguiente ¿puede existir información fuera del cerebro humano? La respuesta, teniendo en cuenta el ADN, es, obviamente, sí. Sin embargo, la pregunta anterior se refiere más específicamente a la información creada o acumulada por los seres humanos. Y de nuevo la respuesta es afirmativa. Las bibliotecas, museos, etc., almacenan información fuera del cerebro. La información humana; esto es, la información producida por los seres humanos, puede existir como las formas de materia o energía, cuya realidad física es independiente de los seres humanos.

Pero, la información es, como la electricidad, *etérea*. Es por eso, por lo que no parece razonable el comparar una información, tal como un pensamiento escrito en un papel, con un producto material cualquiera. Sin embargo, tanto la electricidad como la información tienen realidad física. En efecto, del mismo modo que existen diferentes formas de energía: mecánica, química, eléctrica, calorífica, luminosa, sonora, nuclear, etc., también existen diferentes formas de información, de las cuales una es la humana. Esta información, a su vez, puede ella misma almacenarse y comunicarse en una amplia variedad de formas y representarse de muy diferentes maneras.

Es importante señalar que los medios de propagar información implican habitualmente *impulsos de ondas*: luz, sonido, radio, o de electrones o, aún, impulsos que afectan a la materia y a su organización. El hecho de que la información pueda dividirse en *paquetes* pequeños y discretos se utiliza en telecomunicación para conmutar varios

paquetes y así permitir a varios usuarios que usen la misma facilidad simultáneamente. La idea de que la información es una entidad independiente es algo que se considera natural desde los trabajos de Nyquist, Hartley (Hartley, 1928), Shannon (Shannon, 1948), etc. Así por ejemplo, Bell (Bell, 1968) afirmó: *Información... es una cantidad medible que es independiente del medio físico que la vehicula*. Sin embargo, esto no significa que tenga una realidad *física*, hasta el punto de que Bell compara la información a un término más abstracto, el de *patrón*; pero sí implica una existencia de sí misma. Es decir, la información se trata como una entidad abstracta, pero esta idea no se sigue hasta su conclusión lógica: la información *existe*.

En efecto, la *información existe*. No necesita que se *perciba* para existir. No necesita que se *entienda* para existir. No requiere inteligencia para interpretarla. No tiene que tener *significado* para existir. Existe. Sin esta premisa es imposible entender el universo físico y desarrollar una teoría de la información. Y sin esta teoría, no sólo se vuelve imposible, por ejemplo, convertir el desarrollo del software en una verdadera ingeniería con base científica, sino que también es imposible entender el comportamiento de los sistemas de información avanzados y complejos sean éstos biológicos, sociales, ..., incluso algunos fenómenos de la física cuántica, etc. Más aun, sin la información difícilmente puede entenderse el Universo. Y difícilmente podrá abordarse, con cierta garantía de éxito, el problema cerebro-mente.

En resumen, la información, en informática cognitiva, se considera como artefactos abstractos genéricos que pueden modelarse, procesarse y almacenarse por cerebros humanos. La(s) teoría(s) de la informática cognitiva proporciona(n) una nueva percepción acerca de la información y la informática (ciencia de la información), en los aspectos siguientes:

- a) La información es el tercer constructo esencial para modelizar el mundo.
- b) Cualquier producto y, o, proceso de las actividades mentales humanas derivan en información.
- c) Energía, materia e información pueden transformarse entre sí.
- d) El software obedece a las leyes de la nueva informática y de la informática cognitiva.

Este punto de vista invita a considerar cuestiones fundamentales como: ¿Cuándo se

detendrá el aumento de capacidad de almacenamiento de información en dispositivos de memoria? ¿Cuál es la capacidad de información última de un dispositivo que pese, verbigracia, menos de un gramo y ocupe un cm^3 , ya que ése es, aproximadamente, el tamaño de un chip? ¿Cuánta información se necesita para describir el Universo? ¿En qué momento se produjo la primera información? O, lo que es más interesante para lo que aquí concierne, ¿qué relaciones existen entre la información y la energía?

Según Yingxu Wang (Wang, 2008), tal y como se muestra en la figura 2, el modelo Energía- Materia- Información, de ahora en adelante EMI, establece que el *mundo natural* (MN), que forma el contexto de la inteligencia humana y la ciencia y tecnología, es dual. Un aspecto de él es el *mundo físico* (MF), y el otro es el *mundo abstracto* (MA), donde la materia (M) y la energía (E) se usan para modelizar el primero, y la información (I) al último; es decir:

$$MN \cong MF \parallel MA = f(E, M) \parallel g(I) = h(E, M, I),$$

Donde \parallel denota una relación paralela y f , g y h son funciones que determinan, respectivamente, un cierto MF, MA, o MN.

De acuerdo con el modelo EMI, la información juega un papel vital en conectar el mundo físico con el mundo abstracto. Los modelos del mundo natural han sido bien estudiados en física, química y otras ciencias naturales. Sin embargo, la modelización del mundo abstracto aún es una cuestión a ser explorada. En esta exploración toman parte ciencias como: informática, ciencias cognitivas y del cerebro, inteligencia artificial, teoría de la información, etc. De hecho, y en especial las relaciones entre E-M-I y sus transformaciones, se consideran como una de las cuestiones fundamentales actuales en ciencia e ingeniería.

El *mundo natural* MN (E, M, I), particularmente la parte del *mundo abstracto* MA (I), se considera de diferentes maneras por los individuos, debido a las distintas percepciones y contextos mentales entre ellos. Esto indica que aunque el *mundo físico* MF (E, M) es el mismo para todo el mundo, el *mundo natural* MN (E, M, I) es único para los distintos individuos, porque el *mundo abstracto*, como una parte de él, es subjetivo dependiendo de la información que cada individuo obtiene y percibe.

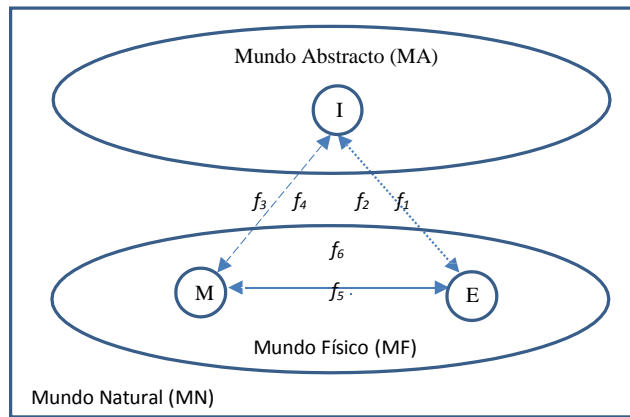


Figura 3.2. El Modelo EMI de la visión general del mundo.

Como muestra el modelo dado en la figura 3.2, la información es el tercer constructo para modelizar el *mundo natural* complementando la materia y la energía; es decir, para modelizar el *mundo abstracto* y sus interacciones con el *mundo físico*. De acuerdo con la teoría cognoscitiva de la informática, la información es cualquier propiedad o atributo de entidades en el *mundo natural* que puede abstraerse, representarse digitalmente y procesarse mentalmente. La ciencia de la información o informática estudia la naturaleza de la información, su procesamiento, y, junto con la física, vías de transformación entre información, materia y energía.

La transformabilidad entre E , M e I puede ilustrarse mediante la figura 3.2, donde todas las funciones genéricas obedecen a las ecuaciones siguientes:

- 1) $I = f_1(E)$, $E = f_2(I) \stackrel{?}{=} f_1^{-1}(I)$;
- 2) $I = f_3(M)$, $M = f_4(I) \stackrel{?}{=} f_3^{-1}(I)$;
- 3) $E = f_5(M)$, $M = f_6(E) = f_5^{-1}(E)$;

en donde la interrogación sobre el signo igual denota que, hasta ahora, la existencia de función directa o inversa es hipotética.

En cierta medida, la teoría de la información contemporánea, para Wang, es la ciencia que trata de encontrar posibles soluciones $f_1, f_2, f_3 \vee f_4$. Como puede verse en el siguiente punto, las interrogaciones anteriores, al menos en parte; es decir, las referentes a la ecuación (1), ya han sido despejadas, tal y como se prueba en el punto 3.2.4.

3.1.3. Equivalencia entre Materia-Energía-Información

Examínese ahora la interconversión entre energía e información, usando un fenómeno físico bien conocido: el choque de dos bolas de billar, tal y como se muestra en la figura 3.3. Supóngase que dos bolas de billar, un blanca y una negra, ruedan sobre una mesa de billar a igual velocidad. La negra se desplaza en dirección noreste y la blanca en dirección sureste, cuando chocan ambas revierten su dirección. La cuestión que, entonces, se plantea es la de saber si ellas cambiaron entre sí energía o información. Realmente ellas cambiaron poca energía pues la velocidad a la que continúan tras el choque es prácticamente similar. Sin embargo, lo que se alteró sustancialmente fue su dirección y esto lleva a preguntarse lo siguiente: ¿es la conservación del momento un reflejo del hecho de que los dos cuerpos meramente cambiaron información?

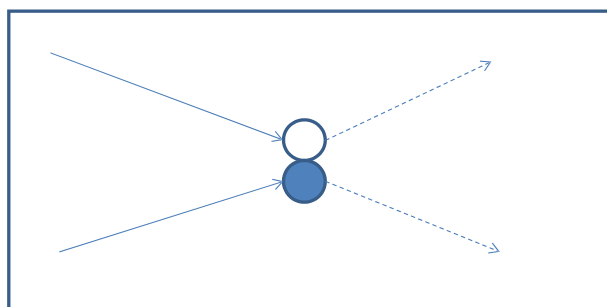


Figura 3.3. Contenido de información de los cuerpos en movimiento.

Comiéncese por explorar la interconvertibilidad entre energía e información. Para ello, pártase de la ecuación del destino de la humanidad, obtenida por Einstein, como consecuencia de su teoría de la relatividad restringida, siguiente:

$$E = m \cdot c^2 / \pm \sqrt{1 - v^2/c^2}. \quad [1]$$

Si una partícula carece de masa; o sea $m_0 = 0$, entonces la ecuación anterior implica, al menos, que la partícula no está desplazándose a la velocidad de la luz, c , y su energía debe ser cero. Es decir, si $m_0 = 0$:

$$E = \frac{0 \cdot c^2}{\pm \sqrt{1 - v^2/c^2}} = 0. \quad [2]$$

Sin embargo, si la partícula se desplaza a la velocidad de la luz; esto es, $v = c$,

entonces, $v^2 / c^2 = 1$, y la ecuación anterior [2], se hace indeterminada; o sea:

$$E = \frac{0 \cdot c^2}{\pm\sqrt{1-1}} = 0/0 : \text{Indeterminado.} \quad [3]$$

Esto significa que E puede tener un valor, pero no puede determinarse mediante la ecuación [1]. Un argumento similar puede hacerse para el *momento relativista*, p, dado por la ecuación:

$$p = m_0 v / \pm\sqrt{1 - v^2/c^2}. \quad [4]$$

Considérese ahora una partícula sin masa desplazándose a una velocidad distinta de c . No debería poseer ni energía, ni momento. No obstante, al menos en teoría, una tal partícula podría existir. Como un fotón, no poseería masa en reposo; sin embargo, y al contrario de un fotón, no debería desplazarse con velocidad c y, por lo tanto, no poseer momento. Con todo, podría poseer velocidad; consiguientemente, podría representar una unidad en movimiento constando de pura información. Tal partícula hipotética tendría ciertas propiedades.

El momento lineal de un fotón, dado por la ecuación [4], también puede expresarse como sigue:

$$p = h \nu / c. \quad [5]$$

Siendo h , la constante de Planck y ν la frecuencia. Por otra parte, y como es archiconocido, la relación entre la frecuencia ν y la longitud de onda λ viene dada por la ecuación:

$$\nu = c / \lambda. \quad [6]$$

Y sustituyendo [6] en [5], se obtiene:

$$p = h / \lambda \Rightarrow \lambda = h / p. \quad [7]$$

Sustituyendo ahora [4] en [7], da:

$$\lambda = h / [m_0 v / \pm\sqrt{1 - v^2/c^2}] = h (\pm\sqrt{1 - v^2/c^2}) / m_0 v. \quad [8]$$

El mismo tipo de argumento puede hacerse con respecto a las ecuaciones [2] [3] y [4]; es decir, con partículas cuya masa en reposo es cero, siempre que $v = 0$, y la ecuación se hace indeterminada y λ puede tener un valor. En los demás casos, λ se hace infinito, lo que conduce a los dos *postulados*, propuestos por Stonier (Stonier, 1990) interesantes, siguientes:

1. Un informón es un fotón cuya longitud de onda ha sido ampliada hasta el infinito.
2. Un fotón es un informón viajando a la velocidad de la luz.

Por lo tanto, un informón puede considerarse como un fotón que parece haber parado de oscilar: A velocidades distintas de c , su longitud de onda aparece infinita y, en consecuencia, su frecuencia es cero. Una vez un informón se acelera hasta la velocidad de la luz, cruza un umbral que le permite ser percibido como teniendo energía. Cuando esto ocurre, la energía E llega a ser una función de su frecuencia ν ; es decir:

$$E = h \nu.$$

Recíprocamente, a velocidades distintas de c , la partícula ni exhibe energía ni momento; sin embargo, podría retener al menos dos propiedades de la información; a saber: su velocidad y su dirección. En otros términos, a *velocidades distintas de la luz, un cuanto de energía se convierte en un cuanto de información; o sea un informón.*

Lo anterior sugiere la posibilidad de que el Universo esté relleno con informones. Sin embargo, las técnicas actuales no están diseñadas para determinar la presencia de tales partículas porque al desplazarse a velocidades distintas de c , fracasan en interactuar con la materia; es decir, no poseen momento. Un informón poseyendo información que lo hace equivalente a la luz verde, no debería ser *visible* como luz verde hasta que alcance una velocidad de c . Del mismo modo que el ojo humano no percibe la radiación electromagnética perteneciente a cierta frecuencia y por consiguiente fracasa en detectar la luz infrarroja hasta que se le dota de los aparatos adecuados, no se perciben partículas sin masa a velocidades inferiores a c .

3.1.4. Ejemplos de la Relación Energía-Información

Desde que, en 1905, Einstein escribió su famosa fórmula de equivalencia entre materia y energía, conocida con el nombre de *ecuación del destino de la humanidad*: $E = m c^2$, se sabe que la energía y la materia son convertibles entre sí (Einstein, 1905). En efecto, el paso de la materia a energía, se probó de forma dramática con las bombas atómicas. El paso inverso, la conversión de energía en materia, se ha observado experimentalmente en los aceleradores de partículas. La producción de pares de partículas electrón y positrón se ha conseguido mediante rayos de alta energía cuando golpean el núcleo de un átomo. Por ejemplo, bajo ciertas circunstancias, un fotón de alta energía puede convertirse en un electrón y un positrón. Pues bien, como se verá a continuación, la energía también puede convertirse en información y viceversa. Esto ocurre bajo circunstancias en las cuales el sistema exhibe una disminución de la entropía o un incremento de la energía potencial. Como sucede en la *inestabilidad de Bernard* y en algo tan cotidiano como levantar un objeto del suelo. Y lo que es aún más importante, la información puede convertirse en energía.

A) La *inestabilidad de Bérnard*, representa un ejemplo chocante de lo que se acaba de decir. Dicha inestabilidad aparece bajo ciertas circunstancias cuando la superficie más baja de una capa horizontal de un líquido se calienta. En estas circunstancias, un gradiente de temperatura vertical erige un flujo de calor permanente que va de abajo hacia arriba. Si el gradiente no es demasiado grande, el calor es vehiculado sólo por conducción. Sin embargo, si el calor aplicado crece lentamente se alcanza un valor umbral en el cual la convención se convierte en importante para la transferencia de calor de abajo hacia arriba. En la *inestabilidad de Bérnard*, este proceso puede implicar un movimiento de moléculas altamente organizado. Como lo señalaron Prigogine y Stengers (Prigogine, 1985): *La inestabilidad de Bérnard es un fenómeno espectacular: el movimiento convectivo que se genera consiste en realidad en una organización espacial compleja del sistema; millones de moléculas se mueven coherentemente, formando células hexagonales de convención de un tamaño característico*. En lugar de producir un nuevo incremento en la desorganización con la posterior aplicación de calor, la *inestabilidad de Bérnard* crea estructuras que ayudan a transferir el calor a lo largo de las capas del líquido. Prigogine y Stengers también hacen la siguiente observación importante: *el calor*

transferido fue considerado un tipo de “despilfarro” en termodinámica clásica. En la célula de Bénard se convierte en una fuente de orden. Éste es, pues, un caso claro donde la aplicación de energía, se convierte en un aumento de la organización del sistema y por tanto de la información. Esta forma particular de organización se mantiene sólo mientras hay un flujo de calor de suficiente magnitud cruzando la capa horizontal del líquido. Una vez se retira la energía, la estructura colapsa, la información desaparece.

- B) La energía potencial. Cada vez que se coge un objeto del suelo y se le coloca sobre una mesa, se gasta energía en alterar la organización del Universo. Para realizar esa operación, uno necesita realizar un trabajo, pero también se crea una situación termodinámicamente menos probable; por lo tanto, se ha *incrementado* el contenido de información del Universo. Tradicionalmente, siempre ha sido un misterio el saber que sucedía con la energía gastada cuando se levanta un objeto. Para explicarlo, los físicos se vieron forzados a inventar un dispositivo *contable*: la energía *potencial*. Dicha energía ha sido siempre una anomalía. El objeto sobre la mesa está desprovisto de energía. Él no se moverá hasta que se le aplique una fuerza. Pero hay otra explicación más sencilla: la energía gastada llega a convertirse en información. La pérdida de energía iguala el incremento de información del sistema. Para enfatizar esto: el *trabajo* es un fenómeno transitorio; por otra parte, el *producto del trabajo* implica un cambio en el estatus de información del sistema, un cambio que permanece fijo hasta que es sujeto a otra fuerza o a posterior trabajo. En línea con esta conceptualización, se puede definir la *energía potencial* como: *un término que describe un estado en el cual el gasto de energía resulta en un aumento en el contenido de información del sistema.*

Una vez considerados los dos ejemplos cualitativos, se van a considerar los dos ejemplos cuantitativos siguientes:

En sistemas microscópicos, cantidades termodinámicas tales como trabajo, calor y energía interna no permanecen constantes sino que fluctúan. De hecho, se han observado violaciones aleatorias de la segunda ley; sin embargo, está en equilibrio térmico: $\langle \Delta E - W \rangle \leq 0$, donde ΔE es la diferencia de energía libre entre estados, W el trabajo ejecutado sobre el sistema, y $\langle \rangle$ el conjunto medio. Sin embargo, el

control por retroalimentación capacita para manipular selectivamente sólo aquellas fluctuaciones que violan la segunda ley tal como saltos hacia arriba, usando la información acerca del sistema. Este experimento mental muestra que empleando control por retroalimentación, puede usarse la información como una fuente de energía libre. De hecho, Szilard (Szilard, 1929) desarrolló un modelo que convierte un bit de información acerca del sistema a $K_B T \ln 2$ en energía libre o trabajo. En otros términos, la segunda ley de la termodinámica se generaliza como sigue: $\langle \Delta E - W \rangle \geq K_B T I$, donde I es la información mutua contenida obtenida por las medidas efectuadas.

Hasta ahora, la idea de una rectificación térmica simple por control por retroalimentación ha encontrado aplicaciones tales como la reducción del ruido térmico y la rectificación de una corriente atómica a baja temperatura. Por otra parte, el demonio de Maxwell, tipo Szilard, capacita y permite evaluar tanto la entrada, contenido de información utilizado, como la salida, energía obtenida, del control por retroalimentación y relacionarlos operacionalmente. Por consiguiente, proporciona un “test-ground” ideal de la conversión información-energía y juega el papel crucial en los fundamentos de la termodinámica. No obstante, su realización experimental ha sido elusiva, hasta ahora. A continuación, se consideran estos dos casos:

- C) El ruido térmico. Se denomina capacidad de canal, y se representa por C , la cantidad de información máxima que se puede transmitir sobre un canal de comunicación. Matemáticamente, viene dada por: $C = \max_{\{p(A_i) \mid i = 1, 2, 3, \dots, n\}} I(E, S)$, siendo: $I(E, S)$ la información mutua, E representa la entrada y S la salida; es decir, $I(E, S) = H(E) - H(E/S) = H(S) - H(S/E)$, que representa la cantidad media de información recibida sobre la entrada, una vez conocida la salida del canal. El valor viene dado en bits/símbolos de entrada.

Los ingenieros de telecomunicación saben muy bien el papel que juega la potencia de la señal sobre la calidad de la información y añaden que es intuitivamente evidente. Sin embargo, el papel de la amplitud o ancho de banda, insisten, lo es menos. Ahora bien, como es sabido, por la fórmula de Hartley, Tuller-Shannon (Rifá, 1991), para canales *tiempo discretos* y *amplitud continua* (TDAC) gaussianos,

la capacidad de canal viene dada por:

$$C = W \log (1 + \gamma) \text{ bits/segundo.}$$

En efecto, imagínese un ruido térmico; o sea, causado por la temperatura. En este caso, como es de sobra conocido, la potencia del ruido es proporcional a la temperatura y al ancho de banda de las frecuencias del canal W . Sea $\gamma = \beta^2 / \sigma^2$, la relación entre la limitación de la potencia media asociada al canal y la varianza de la señal asociada al ruido; esto es, la relación señal/ruido. El valor σ^2 , cuando la fuente de ruido es sólo térmica, viene dado por: $\sigma^2 = kTW$, siendo T la temperatura absoluta y k una constante. Por su parte $\beta^2 = M/n$, es la energía disipada por cada símbolo emitido; siendo M la limitación de la energía media a la entrada del canal y $n = [2WT]$ la longitud de las secuencias de entrada. Entonces se tiene que:

$$C = W \cdot \log \left(1 + \frac{\beta^2}{kTW} \right) \text{ bits/segundo.}$$

Y si se supone W muy grande, resulta:

$$\lim_{W \rightarrow \infty} C = \frac{\beta^2}{kT} \log e = 1,44 \frac{\beta^2}{kT}.$$

Naturalmente, el valor de C depende del sistema de medidas usado; cuando se expresa en grados Kelvin y β^2 en julios, resulta ser que:

$$1 \text{ bit/segundo} = \frac{\beta^2}{T} \cdot 10^{23}.$$

Esto expresa una limitación fundamental, la cual es que para conseguir un bit de información a la temperatura absoluta T se necesita una energía de $T \cdot 10^{23}$ julios. Por ejemplo, a la temperatura ambiente de 20 grados centígrados se tiene que: $1 \text{ bit} = 3 \cdot 10^{21}$ julios. Como puede verse, existe una relación precisa y claramente definida entre información y energía.

D) *El demonio de Maxwell.* En 1929, Leo Szilard (Szilard, 1929) inventó un protocolo de investigación con el cual una hipotética inteligencia, denominada, en

honor a su creador, *demonio de Maxwell* (Maxwell, 1871), bombeaba calor de un entorno isoterma y lo traducía o convertía en trabajo. Después de una intensa y extensa controversia que duró unos ochenta años, finalmente se llegó a la conclusión de que el rol demonio no contravenía y menos violaba la omnipresente segunda ley de la termodinámica, la de la entropía creciente. Pero eso implica que, en principio, se puede convertir información en energía libre. Sin embargo, la demostración experimental de esta conversión información-a-energía, se mostraba muy elusiva, hasta que Toyabe y colegas (Toyabe, 2010), probaron, en 2010, que una manipulación *nonequilibrium feedback* de una partícula browniana basada en la información acerca de su ubicación, consigue una conversión información-energía, tipo Szilard. Es decir, Shoichi Toyabe y colegas realizaron un experimento en el que desarrollan y muestran un nuevo método para evaluar el contenido de información y las cantidades termodinámicas de sistemas de retroalimentación y demuestran, por primera vez, la conversión información-energía tipo Szilard, usando una partícula coloidal sobre spiral- stair-like potential; es decir, un potencial similar a una escalera de caracol.

En el mundo real, sistemas biológicos que utilizan membranas están formados con tales demonios trabajando. Por ejemplo, el alga verde *Valonia* consiste en una esfera hueca con líquido dentro. Se sabe que dicho líquido contiene concentraciones de potasio mil veces mayor que el agua del mar donde habita. Similarmente el riñón humano está depurando moléculas de la sangre, excretando las potencialmente dañinas, incluyendo el exceso de agua. Un riñón requiere energía a fin de ejecutar su tarea. Como tal, es uno de los muchos ejemplos de máquinas biológicas que convierten energía en información. Es decir, los demonios biológicos efectúan el trabajo de *demonio de Maxwell*, seleccionando moléculas y disminuyendo la entropía, pero lo hacen sólo como resultado de una entrada de energía.

Considérese, ahora, tal y como se muestra en la figura 3.4(a), una partícula microscópica sobre una potencial, tipo escalera de caracol. Se establece la altura de cada escalón proporcional a la energía térmica $K_B T$, donde K_B es la constante de Boltzmann y T la temperatura. Sujeta a fluctuaciones térmicas, la partícula salta entre escalones, estocásticamente. Aunque la partícula a veces salta a un escalón superior, los saltos hacia abajo, de acuerdo con el gradiente, son más frecuentes

que los saltos hacia arriba. De este modo, en término medio, la partícula desciende por la escalera a menos que sea empujada hacia arriba externamente, tal y como se muestra en la figura 3.4(a). Considérese, ahora, el siguiente control con retroalimentación. Mídase la posición de la partícula en intervalos regulares, y si se observa un salto hacia arriba, se coloca una *pared* tras la partícula para prevenir subsiguientes saltos hacia abajo, tal y como se muestra en la figura 3.4(b). Si se repite este procedimiento, es de esperar que la partícula acabe por subir todos los escalones. Nótese que, en el caso ideal, la energía para colocar la *pared* puede ser despreciable; esto implica que la partícula puede obtener energía libre sin ninguna inyección de energía directa. En tal caso, ¿qué impulsa a la partícula a subir las escaleras? Esta aparente contradicción a la segunda ley de termodinámica, representada por el *demonio de Maxwell*, inspiró a muchos físicos para generalizar los principios de la termodinámica. Ahora se entiende que la partícula está dirigida por la *información* ganada por la medida de la ubicación de la partícula.

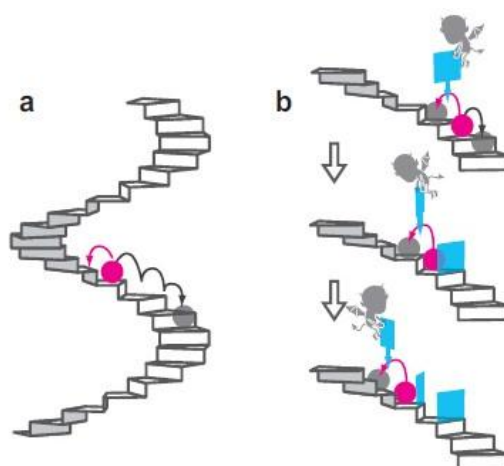


Figura 3.4. Ejemplo de partícula microscópica en Espiral.

Toyabe y colegas, demuestran que la energía libre se obtiene por un control por retroalimentación usando información acerca del sistema; es decir, convirtiendo la información en energía libre, como la primera realización de un demonio de Maxwell tipo Szilard. La energía libre o trabajo obtenido se compensa por el coste de la energía del demonio para manipular la información; no se viola la segunda ley de la termodinámica cuando se considera el sistema total que incluye tanto la partícula como el demonio. Éste consiste en dispositivos macroscópicos en la práctica, tales como computadores en el sistema; y el dispositivo microscópico que

gana energía a expensas del consumo de energía de un dispositivo macroscópico tal y como se muestra en la figura 3.5. En otros términos, usar información como el *medio* de transferir energía. Esta conversión información-energía puede utilizarse para transportar energía a nanomáquinas incluso aunque no sea posible manejarlas directamente. En dicho esquema, se representa la conversión información-energía con un demonio macroscópico y un sistema microscópico como un ejemplo de motor de Szilard. El motor de Szilard puede alcanzar el ratio del 100% de conversión de la información a la energía; 1 bit ($\ln 2$ nat en logaritmos naturales) de información se convierte a energía libre o trabajo de $K_B \ln 2$ en el sistema microscópico a expensas de un consumo de energía de $K_B \ln 2$ en el demonio microscópico.

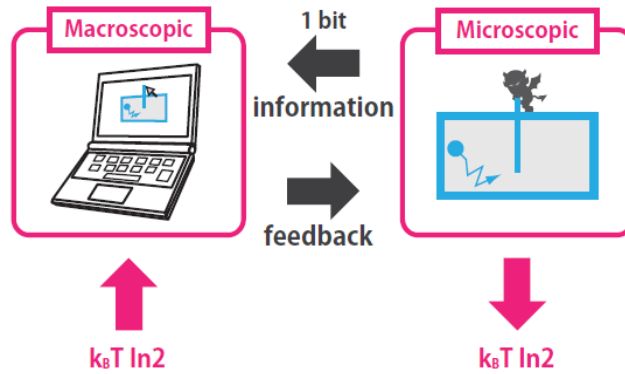


Figura 3.5. Conversión Información-Energía.

Como puede verse, las interrelaciones entre información y energía no son sólo cuestión de experimentos mentales, como sucede con el *demonio de Maxwell* de Szilard, sino que es algo materialmente realizable, como lo muestran los ejemplos acabados de dar.

Quedan dos cuestiones importantes abiertas. Una, igual que la materia está formada de fermiones y la energía de partículas denominadas cuantos, se trata de buscar el componente, informón de la información. Dos, establecer, las interrelaciones entre materia e información, con lo que se cerraría el triángulo formado por los constructos elementales de la ontología del Universo. De esta segunda cuestión abierta, según Wang, Pazos y colegas están llevando a cabo un trabajo, cuyos resultados esperan poder presentar en un futuro próximo.

3.2. Los Constructos: Holones e Informones

El marco conceptual para abordar desde la ciencia cognitiva, y la IA, el problema de conciencia, consta de sólo dos elementos fundamentales o constructos conceptuales siguientes:

3.2.1. Informón

3.2.1.1. Definición

Cualquier teoría acerca de la informática en general y del diseño y desarrollo software más en concreto, tiene que considerar; esto es, establecer y definir, el elemento fundamental de la misma: la información y su componente básico el “informón”. Etimológicamente, información procede del termino latino “informare” cuyo significado de “inform” y “are”, es “dar forma” y del sufijo griego “on” entidad, ser, partículas que fusionadas dan “ser informacional”, o “ser con forma”. Es decir, informón es una entidad informacional. Analícese ahora el término información. Si se busca en los diccionarios, uno se encuentra que, según el DRAE, Información es la acción y efecto de informar e informar es dar noticia de algo (RAE, 2001). Por su parte, en el Webster’s New World Dictionary (Guralnik, 1986), se encuentran, entre otros, los siguientes significados: noticias, conocimiento adquirido de cualquier manera, hechos, datos, etc. En efecto, y eso es lo que es trascendental aquí, la información puede presentarse en forma de datos, noticias y conocimientos. Este enfoque, coincide, tal y como se muestra en la tabla IV.1, con multitud de autores (Lage, 2004), salvo que alguno de ellos, reducen toda la información a lo que, más adecuadamente, debe denominarse noticias.

Conceptualmente, para que se produzca información debe existir algún grado de incertidumbre. Como lo señaló Bateson (Bateson, 1979): *Información es cualquier diferencia que produce una diferencia*. Es decir, si algo caracteriza a la información es que es o produce sorpresa (Schank, 1996). Lo que hace que algo sea digno de interés y consideración, es aquello que aparece organizado alrededor del concepto de fracaso de expectativas; las previsiones cobran interés no cuando se cumplen, sino justamente cuando fallan. Por ello, información es el valor de la sorpresa, medida como el inverso de la probabilidad esperada de un evento o fenómeno. Es decir, cuando se habla de cualquier

fenómeno, no es posible aprender algo válido de él, si está totalmente determinado. Teniendo en cuenta la definición de Bateson y ampliándola, se puede definir información del modo siguiente: “Una diferencia, causada por un proceso subyacente, casi siempre dirigido por propuestas y, o, intereses, capaz de transformar una estructura cognitiva vehiculada por una colección de símbolos que tiene el potencial de alterar el estado cognitivo de un agente cognoscitivo que le dota de significado”.

En consecuencia, sólo puede haber información si existe un elemento de variabilidad. Ésta debe presentarse, como acaba de decirse, en un número finito de estados distintos, pues, en otro caso, no sería posible considerarlos todos. Entonces, ¿cuál sería el significado de los no considerados? Lo único que se puede decir de ese fenómeno finito es en que estado está actualmente. Todo ello permite establecer la siguiente definición formal de información: *Sea un fenómeno variable que se presenta en un número finito de estados, se dice que hay información sobre dicho fenómeno cuando se determina previamente el estado actual del fenómeno sometido a consideración.*

3.2.1.2. Niveles y elementos de información

Los distintos elementos que intervienen en el manejo y transmisión de la información (Lage, 2004), que establecen una categorización en niveles de la misma, son los siguientes: Soporte, señales, signos, datos, noticias, conocimientos y sabiduría.

La información va desde la infraestructura que la vehicula, formada por su soporte, señales y signos, a la superestructura que la usa para la acción, constituida por conocimientos y sabiduría, pasando por la estructura, conteniendo los datos y las noticias (que es lo que habitualmente se denomina información). Es decir, los tres primeros elementos, conforman la infraestructura de la información, los dos del medio, la estructura y los dos últimos la superestructura. Estos cuatro últimos conforman lo que genéricamente se denomina información. El paso de unos niveles a otros se puede dar esquemáticamente como sigue:

- a) Soporte, medio que transmite señales o soporta fenómenos (en griego lo que aparenta ser), verbigracia: Aire, Agua, ADN, canales, humo, nervios, etc., son distintos soportes + variaciones en dicho soporte =.

- b) Señales + Código =
- c) Signos + Patrón =
- d) Datos = Hechos, imágenes, sonidos, valores de variables, texto, código, ... + Interpretación + Significado + Estructura + Relevancia + Propósito =
- e) Noticias: Datos interpretados, resumidos, organizados, estructurados, filtrados, formateados + Acción + Aplicación =
- f) Conocimientos: Ideas, normas, procedimientos, casos, reglas, modelos, intuiciones que guían las decisiones y actuaciones + Selección + Experiencia + Principios + Restricciones + Aprendizaje + Intuición =
- g) Sabiduría: Juicios éticos y estéticos y axiológicos, preferencias, gustos, etc.

3.2.1.3. Aspectos de la información

Clásicamente, y teniendo en cuenta la naturaleza triádica de los signos, hay tres aspectos a considerar en el estudio formal de la información (Morris, 1938); a saber:

- a) Sintáctico, que trata sobre cómo están estructurados los símbolos.
- b) Semántico, que concierne a lo que denotan o significan.
- c) Pragmático, que afecta a cómo se interpretan o usan.

Aspectos que, de modo similar, también considera Arsac (Arsac, 1970) y hoy, con ligeras variaciones, acepta todo el mundo (Paradela, 2003). Sin embargo, siendo importante la información en si misma, no lo es menos la componente de comunicación que implica. Intuitivamente, el significado de una “comunicación” es una combinación de semántica y pragmática. Esta última, incluye, inherentemente, consideraciones tales como el entorno y los estados mentales de los que se comunican, que son externos al propio lenguaje. Las máximas de Grice de la comunicación son, en psicología, la piedra maestra del proceso de discusión, y son las siguientes (Grice, 1975):

- a) “Cantidad”: Hacer la contribución, a la conversación, tan informativa como sea necesario, pero no más, para reducir la confusión.
- b) “Calidad”: Intentar hacer sólo contribuciones verdaderas; es decir, no decir lo que se cree que es falso o de lo que se carece de la adecuada evidencia para establecerlo.

- c) Relación: Emplear sólo aquello que sea relevante para el asunto entre manos.
- d) Modo, manera o forma: Evitar la oscuridad y la ambigüedad; ser breve y ordenado. Es un anhelo generalizado en casi todos los ordenes de la vida, el economizar palabras, naturalmente sin perder nada del significado de lo que se quiere decir. Baltasar Gracian (Gracian, 1963) decía: *“Lo bueno si breve, dos veces bueno; y aún lo malo, sí poco, no tan malo: más obran quintas esencias que fárragos”*. Las formulas, tablas, etcétera, ayudan a este propósito.

Las formas más concisas, elegantes y potentes de transmisión de ideas y conceptos son: las fórmulas y ecuaciones matemáticas y la poesía; ambas comparten la siguiente cualidad característica: La poesía es, como las fórmulas y las ecuaciones, la forma del lenguaje más concisa y cargada de significado. La diferencia está en que, en las fórmulas y ecuaciones, el idioma es universal, cosa que no sucede en la poesía.

No existe consenso respecto a la distinción entre datos, noticias y conocimientos. Una posible razón para esta falta de consenso estriba en el hecho de que se mezclan distintas perspectivas en las discusiones acerca de conceptos que, como sucede en el caso de “información”, resultan ser “polimorfos”, o sea, cuando no es posible definirlo, como se hace con las definiciones clásicas, mediante un conjunto de atributos, características, y, o, propiedades, y, o, condiciones necesarias y suficientes universalmente validas. Por ejemplo: “coche”, no es lo mismo para un ingeniero mecánico que para un ecologista o para un planificador de tráfico; el concepto “coche” admite, distintas definiciones dependiendo del contexto. Todo lo contrario sucede con los conceptos matemáticos formales establecidos que tienen definición única.

En general, no hay forma de distinguir datos, noticias y conocimientos a partir de una base “representacional” de los mismos; es decir, vistos aisladamente como elementos y estructuras en un papel o una máquina o cualquier otro sistema, pues usan los mismos signos y señales. Por eso, cualquier distinción basada en el tamaño o complejidad de la representación está, con probabilidad rayana en la certeza, condenada al fracaso. Una alternativa estriba en ir más allá de la representación e identificar cómo y con qué propósito se usan estas estructuras; esto es, que papeles juegan en el proceso, generalmente de toma de decisión, en el que participan. En consecuencia, para poder categorizarlos, ya que no es posible clasificarlos, hay que tener en cuenta, además de las

estructuras por ellos mismos representadas, su “interpretación” dentro de los distintos contextos en que se aplican y por quién son interpretados y aplicados. Esto conduce directamente al problema del “marco de referencia” de los aspectos de interpretación y los “agentes”, hombres, y, o, máquinas, etc., ejecutando la interpretación que están interrelacionados. Con el término “agente” se quiere decir un sistema, natural y, o, artificial, con capacidad de inferencia y razonamiento sobre estructuras abstractas y de ejecutar acciones en base a esas capacidades. La cuestión crucial es dilucidar a que “agente” debe asignársele un(os) elemento(s) concreto(s) en forma de dato(s), y, o, noticia(s) y, o, conocimiento(s), o si éstos pueden considerarse como objetivos y, o, elementos independientes de un intérprete particular. El concepto de “holón”, que se considerará más adelante, permitirá superar esas dificultades.

De todo lo anterior se deduce que, aunque pueda tomar distintas apariencias y usos, la información es una y, por eso, uno debe ser abstractamente su elemento básico que, a falta de mejor nombre, el autor y con quien lleva a cabo su investigación, deciden denominarle “informón”. Hay que aclarar que dicho término lo usó, por primera y única vez, Uttley (Uttley, 1970) al aplicar algunas ideas de la teoría matemática de la comunicación de Shannon (Shannon, 1948), al perceptrón de Rosenblatt formado por unidades de separación lineal (“neuronas”) lo que constituía una fuerte limitación (Rosenblatt, 1958).

Es decir, toda información, con independencia del soporte que se utilice, de las señales que use y de los signos a que dé lugar y emplee, es, ni más ni menos, un conjunto de informones. Fíjese que no se dice que puede representarse, es mucho más que eso, está constituida por informones. Por otra parte, los tres aspectos a considerar en el estudio formal de la información (Morris, 1938): sintaxis o estructura de los signos, semántica o significado de los mismos y, pragmática, o interpretación y, o, uso de dichos signos, están contemplados en el informón. El informón es tanto un dato a nivel sintáctico, como un registro de un MIS a nivel semántico, como una ontología a nivel pragmático. Por ejemplo, las bases de datos están constituidas por informones tipo datos. Los sistemas de información proporcionan informones tipo noticias estructuradas. Y las bases de conocimiento se estructuran con informones en forma de ontologías.

Así, se define el informón como el elemento básico de la información que tiene sentido para un holón y le permite tomar decisiones y ejecutar acciones adecuadas. El informón puede tomar la forma de un dato, de una noticia y de un conocimiento. La unión de informones como datos y los holones correspondientes, de informones como noticias y los correspondientes holones con los que interactúan, forman los así, aunque inadecuadamente, denominados sistemas de información. Finalmente, la unión de informones, en forma de ontologías, como conceptualizaciones consensuadas de conocimientos, y los correspondientes holones, en este caso en forma de motores de inferencia, que las manejan y se interrelacionan forman los Sistemas Basados en Conocimientos, en general, y los Sistemas Expertos en particular.

3.2.2. El “Holón”

3.2.2.1. La parábola de los relojeros. Definición de holón

Como ya se ha dicho, no existe información sin algo que la procese. Dicho lo anterior, el elemento básico del proceso de la información que aquí se considera es el que, para usar el término empleado por Koestler (Koestler, 1967) basándose en “la parábola de los relojeros” de Simon (Simon, 1962), se va a denominar holón.

Herbert, A. Simon (Simon, 1962), planteó la parábola de dos relojeros, en el sentido de diseñadores y, sobre todo, constructores de relojes. En dicha parábola, Simon concluía que los sistemas evolucionan mucho más rápida y seguramente de un nivel de complejidad determinado a otros niveles de complejidad significativamente superior, si se consiguen formas intermedias estables que actúen como pasos intermedios hasta “alcanzar” el objetivo. Simon ilustraba el proceso según el cual era posible “crecer”, minimizando riesgos, hasta llegar a sistemas complejos. La idea es que un sistema complejo resultaría sumamente frágil si no constara de subsistemas interrelacionados y, en la medida conveniente, autónomos. La parábola, sucintamente, es la siguiente: Había una vez dos relojeros, Tempus y Hora, que trabajaban, naturalmente en Suiza. Hora montaba sus relojes como, por entenderse, se monta un castillo de naipes. Bastaba, por tanto, que una pieza, verbigracia un muelle, se saliera de su sitio, o, por exagerar, le “atacara” un golpe de tos al relojero mientras colocaba una tuerca, para que todo se le viniera abajo y se viese forzado a retomar, cual Sísifo (Seeman, 1958), la tarea desde el

principio. Tempus, con más astucia y visión de futuro, seguía otro procedimiento más eficaz. Ensamblaba unidades independientes y complementarias, hasta que oía el “tic tac” señal que el trabajo estaba felizmente terminado. De este modo, se precavía de que el golpe de tos o la pieza rebelde, paradigma de la maldad de las cosas inanimadas, echaran a perder su trabajo. Pues, en el peor de los casos, se estropeaba sólo una unidad o módulo del sistema total; es decir, se malograba una unidad que formaba parte o estaba comprendida en otra unidad de orden superior, lo que era un incordio, pero no una tragedia. Tempus, consiguió fabricar muchos más relojes que Hora y al final se quedó con, prácticamente, toda la clientela. Retomando esta hipótesis contrastada, Arthur Koestler (Koestler, 1967), siguiendo la máxima holista aristotélica de que “el todo es más que la suma de las partes” (Aristóteles, 1998), y teniendo en cuenta la parábola de los dos relojeros, acuñó el neologismo “holón” que, etimológicamente, viene de “holos”; es decir, totalidad, e incorpora el sufijo “on” que, como ya se ha dicho, significa “parte” o “partícula”. El concepto de holón está inspirado en las estructuras recursivas autosimilares presentes en los sistemas organizativos biológicos. De hecho, un holón lo definió Koestler como una parte identificable de un sistema que tiene entidad única, compuesta de partes subordinadas y que, a su vez, es parte de un todo mayor.

La idea de jerarquía y de sus constituyentes partes-todo, u holones ha de remontarse a los filósofos presocráticos, padres del atomismo, Leucipo y Demócrito. Éstos desarrollaron el concepto abstracto de átomo y lo usaron para desarrollar una filosofía que pretendía, explicar todos los eventos observados. Aristóteles, por su parte, al establecer en su metafísica que el todo era más que la suma de las partes, fue el iniciador del holismo. Además, usó la jerarquía como metodología para acumular y conectar conocimiento biológico. Este concepto de jerarquía fue, quizás, la forma dominante de ver la conexión entre los órdenes natural, humano y sobrenatural de los seres durante la edad media. En el siglo XVII Leibinz propuso la monada como unidad irreducible para explicar no sólo el mundo material, sino también el mundo espiritual.

Al comienzo del siglo XX hubo un agitado interés en el holismo y la jerarquía que debe su génesis al impacto de la teoría de la evolución de Wallace y Darwin. Posteriormente, en 1926 Jan Smuts en su obra “*Evolution and Holism*”, vio las profundas conexiones entre los mundos natural y social y su concepto de holismo influenció claramente las ideas de Wilber. Éste, en 1980, cita a Smuts al comienzo de su principal trabajo “*The*

Atman Project” donde usa el concepto de jerarquía: *En todas partes se ve en la naturaleza nada más que totalidades*. Mientras todos estos distintos entramados de ideas incluyen la consideración de redes jerárquicas y niveles y órdenes de desarrollo, no fue hasta el trabajo de Arthur Koestler donde se propuso una teoría completa de holarquía y holones.

Koestler, primero, vio la necesidad de algún modelo que pudiera unir e integrar la visión del mundo reduccionista y mecanicista de las psicologías “científica” y conductista, con la visión del mundo holista y humanista de las psicologías freudiana, rogeriana y de la gestalt. En segundo lugar, reconoció la importancia y relevancia de los procesos evolutivos en las ciencias sociales y buscaba proporcionar algún sistema teórico que pudiera aplicar conceptualizaciones evolutivas a ambas realidades. En tercer término, buscaba desarrollar un modelo de sistemas sociales humanos que fueran igualmente naturales para analizar el micronivel de la individualidad y el macronivel de la colectividad. Él buscaba proponer algún modelo básico de explicación que fuera relevante a lo largo y ancho de todo el rango de la actividad humana y su entorno. Koestler reconoció que este constructo “holón” tenía, de hecho, unos muy venerables y antiguos ancestros en la filosofía occidental. Varios e importantes filósofos, incluyendo Leibniz y Hegel habían dirigido su atención a la importancia de cosas tales como jerarquía y niveles de desarrollo. Koestler se veía a si mismo en una línea de pensadores tales como esos que buscaban presentar juntos diferentes búsquedas y escuelas de esfuerzos científicos en lugar de perseguir la continua especialización en el conocimiento científico que había caracterizado las modernas escuelas científicas. La teoría holónica era el intento de Koestler de conseguir una filosofía de la ciencia integradora y esperaba que esta teoría o algo similar formasen la base de cualquier futura visión del mundo futuro.

El término holón es, como ya se ha dicho, una combinación del vocablo griego “holos” que significa todo o totalidad y el sufijo “on” el cual como en las palabras, protón, neutrón, etc., que indica, además de ser, partícula o parte. El holón, entonces, es una parte-todo. Es un punto nodal en una jerarquía que describe las relaciones entre entidades que son totalidades autocompletas y entidades que pueden verse como partes, dependientes de otras. Así, dependiendo de cómo se centre el foco moviéndose arriba, abajo, y, o, a lo largo y ancho de los nodos de una estructura “jerárquica” así la percepción de lo que es un todo y lo que es una parte cambiará.

Koestler, con la parábola de los relojeros de Simon quiso mostrar dos cosas. Una, que los sistemas complejos evolucionarán desde los sistemas más sencillos mucho más rápidamente si hay formas estables intermedias; es decir, si están jerárquicamente organizadas. Dos, y más importante, él buscaba mostrar que los sistemas complejos resultantes serán siempre jerárquicos y que la jerarquía es el resultado natural y ubicuo del desarrollo de formas estructurales. Después de establecer la importancia universal de la jerarquía en el desarrollo de sistemas complejos, Koestler continuó proponiendo que estas jerarquías podrían analizarse en términos de los nodos o formas intermedias estables mediante las cuales se define su estructura. Fue justo a estas formas intermedias a las que Koestler les confirió la nueva etiqueta de “holón”.

Koestler vio la teoría holónica como una filosofía de la ciencia amplia que mostraba una salida al interminable debate acerca de los méritos del reduccionismo y del holismo. Koestler señaló que en cualquier y todo orden existente, de los sistemas físicos a los sociales pasando por los químicos y los biológicos, enteramente autosoportados, no existían entidades no interactuantes. Y lo que es más importante, que las entidades pueden verse situadas en una relación holárquica entre ellas. Él llamaba a los sistemas de tales entidades Open Hierarchical Systems (OHS) y a éstos, subsecuentemente, se les denominó holarquías. Un holón, como Koestler entendía el término, es una parte identificable de un sistema que tiene una identidad única; sin embargo, está hecho de partes subordinadas y, a su vez, es parte de un todo mayor. Los holones de Koestler no fueron pensados como entidades y objetos, sino como una forma sistemática de relacionar estructuras teóricas. En otras palabras, los holones eran puntos de referencia arbitrarios para interpretar la realidad. Los holones de Koestler son postulados y “determinados” sólo fuera de las reglas relacionales y estrategias que ayudan a dar sentido a la realidad.

Dado que los holones se definen por la estructura de una jerarquía cada holón identificado puede verse el mismo como una serie de subjerarquías anidadas de la misma manera que las muñecas rusas, las matrioscas, son una serie inclusiva de muñecas dentro de otras muñecas. Los holones son, entonces, simultáneamente partes y todos porque ellos son siempre partes de jerarquías más amplias y siempre contienen subjerarquías. Los holones son simultáneamente todos autocontenidos de sus partes subordinadas, y partes dependientes cuando se ven en la dirección inversa. Por consiguiente, los holones pueden verse como puntos de referencia en series jerárquicas u holarquías.

Koestler ha propuesto un conjunto bastante detallado de principios holónicos y mostró que el constructo holón tiene una aplicación muy amplia. A su vez Wilber ha colocado el constructo holón firmemente en el centro de su marco integrador comprensivo para conectar conocimiento. Wilber ha ampliado la teoría holónica en un nuevo enfoque, para entender las relaciones de muchos dominios de conocimiento diferente. Las concordancias entre los principios holónicos de Koestler y los 20 postulados de Wilber son claras pero no completas. En efecto, hay varios aspectos de la teoría de Koestler que, hasta ahora, no han sido explorados por nadie, incluido Wilber. Éstos incluyen los conceptos de:

- a) Cambio holístico. Sistemas de entrada-salida que buscan la manera de cómo se disparan los holones y cómo éstos escudriñan y filtran las entradas.
- b) “Arborización”, “reticulación” y “canales de regulación”. Tanto tener aptitudes como maneras de ver cómo los holones pueden relacionarse entre sí.
- c) La “salud” holónica y cómo cambian los holones y los principios de Koestler sobre equilibrio holónico, desorden y regeneración, que ofrecen fértil terreno para posterior estudio.

Por su parte, Wilber a estos principios añade los siguientes:

- a) No exclusión que significa que se pueden aceptar las declaraciones válidas de verdad; es decir, las pretensiones de verdad que pasan los test de validez para sus propios paradigmas en sus propios campos, sean estos hermenéuticos, científicos, etc., en tal grado que realizan afirmaciones acerca de la existencia de sus propios fenómenos descubiertos y realizados, pero no cuando hacen afirmaciones acerca de la existencia de fenómenos realizados por otros paradigmas. Este principio se refiere a la aceptación de conocimiento parcial, pero válido, que ha sido recogido por disciplinas centrándose en aspectos particulares de los holones. Mucho de ese conocimiento ha sido el resultado de paradigmas (matrices disciplinares/metodologías) reduccionistas.
- b) Plegado/Desplegado que se define como sigue: La no exclusión con frecuencia desvela una idea que estaba velada. En cualquier corriente evolutiva, sucesivas olas trascienden e incluyen a sus predecesoras, y así cada ola es adecuada y cada ola sucesiva es más adecuada. En breve, cada ola es holista y cada ola sucesiva es más holista. Este principio “unfolding/enfolding” se refiere a la aceptación de la

naturaleza holista y evolutiva del conocimiento y sus métodos. Este principio se relaciona con la idea de que todas las bases y métodos de conocimiento están conectados y pueden ilustrar a los demás.

- c) “Enactment” “Promulgación”. Poniendo todos estos modos de inquisición juntos, como un “edicto”, “enactment”, y descubrimiento de cognición “turquesa”, resulta en un pluralismo metodológico integral. Los fenómenos son “enacted”, paridos y desvelados por prácticas, entonces uno se percata de que lo que parece ser “fenómenos o experiencias conflictivas” son simplemente experiencias diferentes, y completamente compatibles, paridas por diferentes prácticas. De este modo, “enactment” se refiere a la capacidad de situar y proporcionar un nuevo contexto integrador para que todos los enfoques parciales sean reduccionistas y holistas. Son precisamente estas tres capacidades las que se hacen disponibles cuando se ve el holón como una unidad de análisis para una teoría integral.

3.2.2.2. Características de los holones

El término “holón” se emplea, entonces, para denominar entidades que exhiben, simultáneamente, un comportamiento autónomo; es decir, se comportan como un todo, y no autosuficiente; esto es, se comportan como una parte de un todo mayor. Verbigracia, las moléculas que son, simultáneamente: todo, cuando integran a los átomos que las forman; y parte, cuando, por ejemplo, se integra en una célula como el ADN, que constituye el nivel inmediatamente superior en tamaño y funcionalidad, que no en importancia, pues todos los niveles son decisivos e imprescindibles, sin átomos no hay moléculas, etc. Cuando el holón se siente todo, afirma su autonomía, “identidad”, etc., con lo que adquiere, al menos en los niveles más elaborados, seguridad “sicológica”. Una de las principales características de los holones es su autorreplicación lo que proporciona la funcionalidad de recursión y le puede proporcionar la capacidad de autoorganización y autopoiesis. Esto significa que se precisan capacidades de cooperación. En consecuencia, se puede “definir” un “holón” como un elemento con entidad propia que presenta un comportamiento autónomo, autoorganizativo, recursivo y cooperativo. Autonomía significa que cada holón debe ser capaz de crear, controlar y monitorizar la ejecución de sus propios planes y, o, estrategias de actuación y realizar acciones preventivas y, o, correctivas adecuadas; esto es, racionales frente a sus propias disfunciones. En 1981, el premio Nobel de Medicina Roger Sperry hizo uso del concepto de holón en su área de

conocimiento. Posteriormente, Suda a partir de 1989, aplicó el concepto de holón a los sistemas industriales, con lo que el término se hizo ubicuo. En general, el holón es la unidad básica estable y coherente en sistemas complejos tanto naturales (biológicos y sociales) como artificiales (informáticos o industriales). Un holón contiene siempre una parte de procesamiento de información y, opcionalmente, otra de procedimiento físico. Con el término holón, el doctorando concretamente denomina, siguiendo a Alonso y otros (Alonso F., 2004), las entidades de proceso que exhiben simultáneamente:

- a) Un comportamiento *autónomo*; esto es, se comportan como un todo, por lo que requieren capacidades de cooperación. Entendiendo por autonomía la capacidad de una entidad para generar su propio comportamiento; o sea, sus propios planes y, o, estrategias de actuación y comportamiento, en definitiva su “teleología” y controlar su ejecución, así como su propio estado.
- b) No *autosuficiente*; es decir, se comportan como parte de un todo mayor. Cuando el holón se siente “todo”, afirma su autonomía, “identidad”, etc., con lo que adquiere, como ya se ha dicho, al menos en los niveles más elaborados, seguridad “sicológica”.
- c) Capacidad de cooperación. Por cooperación, se quiere decir, en este contexto, un proceso por el cual un conjunto de dichas entidades desarrolla planes, comúnmente aceptados, que se ejecutan en forma distribuida. Cooperación significa que los diferentes holones deben ser capaces de negociar y ejecutar planes mutuamente aceptables, “joint intentions” y llevar a cabo acciones preventivas y, o, correctivas frente a disfunciones globales. Dicho de otro modo, un holón es el elemento básico de proceso de información, en cualquiera de sus niveles, con entidad propia capaz de presentar un comportamiento autónomo y cooperativo.
- d) Carácter abierto. Un sistema holónico debe permitir la inclusión de nuevas unidades, eliminación de las existentes y modificación de las capacidades funcionales y comportamentales de las existentes, con intervención exterior mínima. Los holones y, o, sus funcionalidades pueden venir de distintas “mentes” (heterogeneidad).
- e) “Autorreplicación”. Una de las principales características de los holones es su autorreplicación lo que proporciona la funcionalidad de recursión y la capacidad de autoorganización y autopoiesis o autoproducción específica del nivel celular.
- f) Estabilidad. Todo holón, de acuerdo con el contexto y circunstancias, posee cuatro impulsos, a saber:

- 1) Actividad o impulso a continuar siendo totalidad.
- 2) “Comunión” o impulso para seguir siendo una parte.
- 3) Trascendencia o impulso de superación y ascendencia.
- 4) Disolución, o impulso a descender y descomponerse.

Estos impulsos son los que le dan estabilidad al holón y le permiten desarrollar sus capacidades. Como lo señaló Octavio Paz, los seres humanos han descubierto, al finalizar el siglo XX, que son parte de un inmenso sistema o “conjunto de sistemas” que va de las plantas y los animales a las células, las moléculas, los átomos y las estrellas. Es decir, un eslabón más en la “cadena del ser” como llamaban los antiguos filósofos al Universo. Una de las principales características de los holones es su granularidad múltiple, que se manifiesta a través de la replicación de estructuras autosimilares. Esto permite establecer una descomposición, jerárquica y, o, heterárquica, en forma fractal que se puede denominar “holarquía”. Una holarquía puede definirse, pues, como un sistema de holones autorregulados que, según las circunstancias, cooperan y, o, compiten, y, o, colaboran teleológicamente; esto es, para alcanzar una meta u objetivo global. En suma, una holarquía define las reglas básicas y específicas de lo que genéricamente se puede denominar la colaboración de los holones que la forman y, en consecuencia, limita su autonomía y, o, condiciona sus procesos de toma de decisión. Dicho en otros términos, el carácter colaborativo de los holones se manifiesta tanto en los procesos teleológicos de planificación y asignación de ellos mismos y los recursos que necesitan para la consecución de las metas globales del sistema en que se integran, como para la superación de sus carencias, en recursos y, o, capacidades en la ejecución de sus propias metas. Su fortaleza y oportunidad procede del hecho de ser unidades independientes; es decir, módulos, en el tratamiento de problemas cooperativos en entornos heterogéneos. Su eficacia proviene de su característica de ser, a su vez, elementos subordinados de otros holones de nivel superior que los contienen y, en cierta medida, controlan. Las formas intermedias proveen la funcionalidad propia de la totalidad mayor.

Obviamente, los holones tienen que tomar decisiones racionales y, como se aplican a una nueva disciplina científico-tecnológica, dichas decisiones racionales deben ser similares a las científicas. Estas decisiones pueden modelizarse con los pares “holón(es)”, “informón(es)”, en los cuales los holones representan dicho mundo. Se supone,

naturalmente, que los holones pueden elegir entre distintas alternativas u opciones. Asimismo, se supone que el mundo es como es y que esto se refleja en su representación mediante los informones, y que lo que le ocurre al holón depende, por una parte, de como sea el mundo en un momento determinado y, por otra, de las opciones epistémicas y gnoseológicas que elija el holón. Se acepta, también, que el holón no conoce el estado actual del mundo pero al menos es capaz de distinguir entre dos estados posibles del mismo. Las valoraciones de los holones no atañen directamente a las opciones o a los estados del mundo, sino a los pares “opción-estado”; es decir, lo que se valora es haber elegido una opción particular en un estado definido del mundo.

3.2.2.3. Niveles holónicos

Al igual que los informones pueden tomar la forma de dato, noticia o conocimiento, los holones que los manejan presentan estructuras, desde las más simples a las más complejas, dependiendo de la complejidad funcional asignada a los mismos. Más en concreto, la construcción del software se presenta mediante un modelo computacional holónico formado por diferentes niveles denominados:

1. “Instrucción”: Lo constituyen los holones primarios que son entidades propias y cooperativas que tratan únicamente *datos* y producen nuevos datos o *noticias simples*. Estos holones están especializados y realizan operaciones *primitivas* básicas. Por ejemplo, si a un holón primario *operador lógico* le llegan dos datos que son las edades de dos individuos, el holón puede sacar como noticia que un individuo es mayor, igual o menor que otro.
2. “Componente”: Un holón componente surge al estructurarse holárquicamente (jerárquicamente/heterárquicamente) los holones *primarios* del nivel “instrucción” proporcionando un nivel superior al de “instrucción”. *Un holón componente* tiene una funcionalidad superior a la suma de sus holones “instrucción” y tiene capacidad de producir *noticias más elaboradas* y, o, *conocimientos*. Si los informones que trata el holón componente corresponden a estructuras de datos, el componente proporcionará noticias típicas de un sistema de información. En cambio, si los holones tratan estructuras de datos que corresponden a conceptualizaciones de conocimiento, el resultado será informones de conocimiento que proporcionan información propia de

los sistemas basados en el conocimiento. Por ejemplo, un holón componente puede representar un programa o un subprograma.

3. “Entidad”: El holón “entidad” se forma mediante relaciones holárquicas de holones componentes. Un holón *entidad* presenta creencias, motivaciones e intenciones y cambia su comportamiento basándose en experiencias previas. Incorpora la propiedad de *proactividad*; es decir, actúa por su propia iniciativa para alcanzar metas que el mismo es *capaz de generar*. Los holones “entidad” tratan con informones de noticias y conocimientos, dependiendo de la estructura subyacente de los datos. Por ejemplo, un holón entidad puede representar y actuar como un agente software. Los sistemas transaccionales son informones tipo datos y holones clase programas. Los sistemas de información son informones tipo noticias estructuradas y holones tipo programas y fórmulas. Finalmente, los sistemas basados en conocimientos son, ni más ni menos, que informones en forma de ontologías y holones en forma de agentes.

4. “Organización”: Se denomina organización holónica a una holarquía de entidades colaborativas. Un holón *organización* designa una agrupación formal y estable de holones entidad y ofrece una interfaz bien definida de comunicación con el exterior, que constituye en si misma una unidad independiente e identificable (con entidad propia). Las actividades de cada organización vienen determinadas por procesos de cooperación con otras entidades u organizaciones. Una *organización* presenta una meta común y los holones *entidad*, que tienen sus propias metas, muestran un comportamiento cooperativo dirigido por la meta común de esa “organización”. En la “organización”, los holones “entidad” son libres para participar en la misma, abandonarla y actuar de forma autónoma como “entidad”, o formar nuevos holones “organización” con otras entidades. Una organización holónica precisa en su funcionamiento de los siguientes elementos:
 - a) Estructuras de información y ontologías compartidas por sus holones “entidad”.
 - b) Interfaces de cooperación basadas en protocolos de interacción.
 - c) Infraestructura para el paso de mensajes entre las entidades.

- d) Mecanismos de toma de decisiones que ayuden a los holones “entidad” en sus actividades de planificación de tareas, negociación, intercambio de información, etc.
- e) Técnicas y reglas de descomposición y asignación de tareas entre los miembros de la “organización”.
- f) Facilidades para la monitorización del estado de una tarea distribuida, y la planificación y control de todas las acciones que conforman la tarea.

Por ejemplo, un sistema multiagente (MAS) es un tipo de holón “organización”. Sin embargo, no hay que confundir agentes con holones y, en consecuencia, sistemas de agentes con holones. Como se verá, más adelante, los holones son más que agentes.

3.2.2.4. Dominios de cooperación y colaboración holónica

Además de estos niveles, es importante considerar dentro de una estructura holónica el dominio de cooperación y colaboración holónico. Se define un tal dominio como un espacio lógico en el que los holones “entidad” u “organización”, cada uno con sus propias metas dentro del dominio, operan y se comunican entre sí, proporcionando con ello el contexto en el que estos holones pueden localizarse, contactar e interactuar unos con otros; por ejemplo, un dominio en una plataforma de Internet. Dentro de un dominio, ésta puede ser:

- a) Simple. En este caso, cualquiera se compromete a cooperar con cualquier otro siguiendo un conjunto de protocolos de interacción establecidos. Aquí se consideran aceptables respuestas no cooperativas; verbigracia: “rechazo”, “no entendimiento”, etc. Todos los holones deben presentar facilidades de cooperación simple.
- b) Complejo. Esta clase de cooperación está encaminada a alcanzar una meta compartida por varios holones; por ejemplo, el acuerdo de un plan común para ejecutar las tareas asociadas a un problema distribuido. Una organización holónica se diferencia de un dominio holónico en que la organización está constituida por holones “entidad”, tiene una meta común a todos los holones y éstos participan voluntariamente en la organización para alcanzar las metas de la misma. Mientras que el dominio es un espacio de cooperación y colaboración para que los holones

“entidad” y “organización”, que participan voluntariamente en el mismo, puedan alcanzar cada uno de ellos su propia meta.

En un *dominio*, los holones se conciben como integrantes en una sociedad, cuyas actividades globales deben ser contempladas. Estas actividades deben ocurrir de acuerdo con un conjunto de convenciones y leyes *sociales* tanto de carácter general como de carácter específico. Estas leyes regulan las relaciones sociales que es necesario establecer para llevar a cabo dichas actividades. La estructura social del dominio viene determinada por los roles, las relaciones sociales identificadas entre éstos y las convenciones y leyes sociales que rigen estas relaciones. El *dominio* define entonces el comportamiento global esperado mediante las metas sociales que deben contemplarse, los roles necesarios para alcanzar dichas metas, las relaciones existentes entre estos roles y un conjunto de convenciones y leyes sociales de carácter global que regirán dichas relaciones. Una forma plausible de caracterizar el dominio es mediante la identificación de las relaciones sociales potenciales existentes entre sus actores y, más concretamente, entre los diferentes roles identificados.

El holón asume la responsabilidad de llevar a cabo sus tareas individuales. Ahora bien, al estar en un dominio de cooperación asume también unas responsabilidades sociales. La necesidad de garantizar la realización de estas tareas sociales conduce a la necesidad de identificar tanto las relaciones sociales necesarias para llevar a cabo dichas tareas, como las leyes sociales que regulen dichas relaciones. Estas leyes, impuestas por el dominio, deben ser respetadas por el holón y forzadas por el dominio con el fin de permitir el funcionamiento correcto del mismo, de acuerdo con el comportamiento global esperado. Son, por tanto, las leyes que regulan las relaciones sociales las que permiten obtener un comportamiento socialmente responsable a los holones *entidad y organización*.

Plantearse la obtención de un comportamiento socialmente responsable durante el diseño del holón no sería adecuado en un entorno dinámico y heterogéneo como es un *dominio*, en el que, por ejemplo, los patrones de interacción se establecen dinámicamente. Los distintos modelos de coordinación y negociación y la propia estructura social, constituyen las abstracciones conceptuales para establecer las leyes que deberán regir tanto la actuación de los holones como sus interrelaciones, con el fin de alcanzar un

comportamiento socialmente responsable. Los medios de coordinación y negociación que soportan dichos modelos constituyen el mecanismo utilizado por el dominio para este fin.

La “configuración” que se crea en un dominio de cooperación y colaboración es dinámica. Mientras que el holón *organización* forma una configuración estable, los holones de un dominio son más inestables, participan en el dominio de cooperación para alcanzar sus propias metas, respetando las leyes sociales del dominio y nada más alcanzarlas desaparecen del mismo. Un holón entra a formar parte de un dominio de cooperación y colaboración interactuando con un holón interfaz del dominio de modo que pueda participar en el mismo.

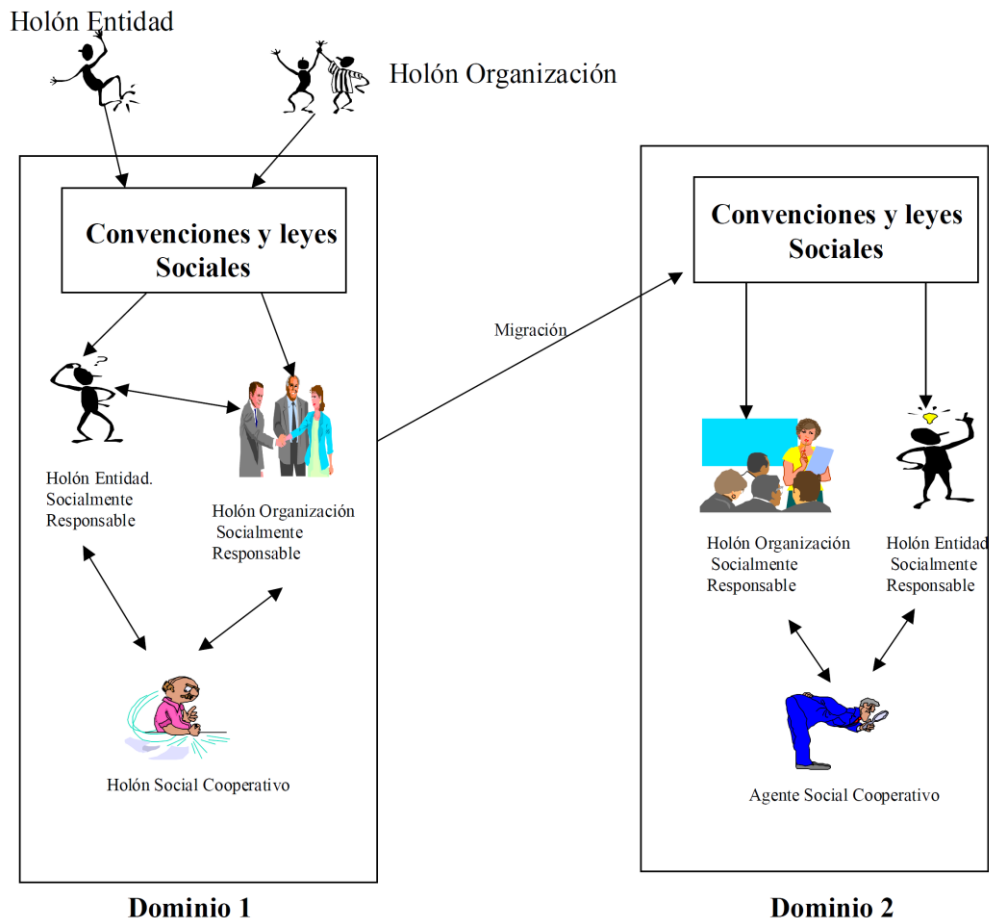


Figura 3.6. Dominios de Colaboración y Cooperación.

Los holones se incorporan en el *dominio* para alcanzar sus propias metas. El *dominio* acepta o rechaza su entrada dependiendo de que el rol con el que desean participar sea apropiado en ese *dominio* y acepten las convenciones y leyes sociales. En caso

afirmativo, se incorporan como holones *socialmente responsables*, acatando las normas del dominio. Normalmente, el dominio incorpora una serie de holones sociales cooperativos con la misión de atender a las necesidades de los holones que participan en el dominio y de hacerles cumplir las leyes sociales, tal y como se muestra en la figura 3.6.

La interacción entre holones ocurre siempre a través de sus respectivos dominios de cooperación. Un dominio de cooperación comprende los siguientes elementos:

- a) Estructuras de información y ontologías compartidas por sus holones miembros.
- b) Interfaces de cooperación basadas en protocolos de interacción.
- c) Infraestructura para el paso de mensajes entre las organizaciones y el dominio de cooperación. Si el dominio de cooperación se localiza en el mismo dispositivo físico, el paso de mensajes puede conseguirse por memoria compartida. En otro caso, los mensajes deben representarse, codificarse y enviarse a través de una red de comunicaciones. Este aspecto se aborda en el modelo de comunicación.
- d) Mecanismos de toma de decisiones que ayuden a los holones en sus actividades de planificación de tareas, negociación, intercambio de información, etc.
- e) Técnicas y reglas de descomposición y asignación de tareas entre los miembros del dominio, así como facilidades para la planificación y el control de tareas dentro de un holón (síntesis).
- f) Facilidades para la monitorización del estado de una tarea distribuida, y la planificación y control de todas las acciones que conforman la tarea.

Un sistema de información holónico, es, pues, una holarquía que integra el conjunto de actividades de información desde la identificación de necesidades, la producción, la selección y adquisición, el procesamiento analítico-sintético y la recuperación y diseminación de información para responder de forma dinámica a las necesidades de los usuarios en forma de organización virtual. Un tal sistema se sustenta en una red de elementos autónomos y con capacidad de cooperar. La organización la determina el propósito y el compartir normas, políticas y reglas. Dentro de una red holónica, cada configuración del proceso puede identificarse como una empresa, institución, etc., virtual. Una componente software es cualquier pieza de código escrita con anterioridad que define un elemento de conceptualización y puede ser invocada para proveer una

funcionalidad que la componente encapsula. Las componentes, típicamente, se empaquetan de acuerdo con normas efectivas de modo que puedan ser invocadas desde lenguajes, entornos y ambientes múltiples. En este contexto, un “webservice”, puede entenderse como un holón.

3.2.2.5. Holones versus Agentes

Debido al confusionismo existente al respecto, se van a dar, tal y como se muestra en la tabla 3.1, las diferencias y similitudes entre holones y agentes. Lo más destacable de la misma son las dos situaciones siguientes: Con respecto a la autoorganización, para unos, como es el caso de Barbat, los holones y los agentes son distintos; sin embargo, para otros, como es el caso de Passer, So y Durfle, y Tambe, son iguales.

PROPIEDAD	HOLON	AGENTE
1. Autonomía	Si	Si
2. Reactividad	Si	Si
3. Proactividad	Si	Si
4. Habilidad Social	Si. La interfaz humana es específica de cada holón.	Si. La interfaz humana se implementa generalmente por uno o varios agentes especializados.
5. Cooperación	Si. Los holones nunca rechazan de manera deliberada la cooperación con otro holón.	Si. El agente puede competir y cooperar.
6. Reorganización	Si. Holarquías. Las holarquías aquí se pueden implementar utilizando varios enfoques para federaciones en esquemas tales como: facilitadores, o mediadores	Si. Jerarquías, organización horizontal, heterarquías, etc. Las holarquías en este caso se pueden implementar usando SMA; o sea, sistemas multiagente.
7. Racionalidad	Si	Si
8. Aprendizaje	Si	Si
9. Benevolencia	Si	Si
10. Movilidad	Los holones raramente necesitan movilidad para la ejecución de sus tareas. Si así fuera, la tendrían.	Si
11. Recursión y Asociación	Si. Ésta es la gran diferencia entre un holón y un agente. Mientras el holón tiene capacidad de recursión o asociación entre varios de ellos que posean un objetivo común, los agentes no pueden agruparse, ni son recursivos.	No existe ninguna arquitectura recursiva como tal, pero algunas técnicas son utilizadas para definir federaciones que simularán los diferentes niveles recursivos.
12. Procesamiento de la Información y Físico	Si. La separación es explícita, aunque la parte de procesamiento físico es opcional.	No existe una separación explícita
13. Actitudes Mentales	Si. Los holones no necesitan razonar acerca de sus propias actitudes mentales o aquellas de otras unidades de control.	Si

Tabla 3.1. Diferencias y similitudes entre holones y agentes.

Pero la diferencia sustancial y de verdad discriminatoria es que mientras que los holones tienen capacidad de recursión e integración entre varios de ellos que tengan un objetivo común, los agentes ni tienen recursión ni capacidad de integrarse. Otra característica fundamental de los holones, en tanto que operan con un conjunto restringido de opciones y estados del mundo que les viene dado por la situación y el problema sobre el que actúan, y no les es posible construir una estructura de preferencias completa y coherente, es la de ser elementos satisfacedores (no maximizadores) en el sentido de Simon (Simon, 1982); es decir, buscan una solución suficientemente buena en vista a las restricciones existentes y el coste relativo de hacer algo mejor. En este sentido se comportan como los seres humanos.

Evidentemente, dichas valoraciones, de acuerdo con las distintas situaciones y problemas, deben adoptar distintas formas que van desde la más sencilla ordenación lineal hasta la racionalidad acotada o satisfactoria de Simon (Simon, 1945), pasando por los holones bayesianos entre otros. En efecto, el agente racional clásico es un “maximizador”, en tanto que en la propuesta de Simon, donde introduce la noción de “satisfacción” (Simon, 1982) retomando la de Tarski (Tarski, 1956), es la de un “agente” que funciona en condiciones de “racionalidad acotada”. Esto es debido, básicamente, a que:

- a) Los holones, van a operar con un conjunto suficientemente restringido de opciones y estados del mundo, que le vienen dados por la situación y el problema sobre el que actúan.
- b) Raras veces es posible construir una estructura de preferencias completa y coherente, ni tampoco calcular la utilidad esperada para cada opción. Si éste fuera el caso, la cuestión se resolvería mediante métodos bayesianos sin problema. En efecto Thomas Bayes (Bayes, 1763) desarrolló, en su trabajo de 1763, un procedimiento para resolver este tipo de cuestiones. Curiosamente, dicho trabajo permaneció prácticamente ignorado, hasta que Pierre Simon de Laplace lo discutió en uno de sus trabajos.
- c) Es capaz de distinguir entre opciones satisfactorias y aquellas que no lo son.

Es decir, de acuerdo con Simon, los holones son satisfacedores (satisfacer), no maximizadores y en este sentido se comportan como los seres humanos. Según Simon (Simon, 1981), hay que ir hacia una solución satisfactoria; esto es, una solución

suficientemente buena en vista a las restricciones existentes y el coste relativo de hacer algo mejor, antes que una solución satisfaciendo. Naturalmente, los pares holones-informones y las estructuras que sobre ellos se construyen, se enfrentaran, inexorablemente, tanto a situaciones conflictivas, como a otras concurrentes y cooperativas, en las que los esquemas de decisión que usen se fundamentarán tanto en la teoría de von Neuman y Morgenstern (von Neuman, 1953) como en la teoría de juegos cooperativos de Nash (Nash, 1953) y cualquier otra teoría o regla de decisión que permita a los holones, de modo efectivo, tomar la decisión más adecuada, en el momento más oportuno y de la forma más eficiente posible.

De hecho, si no fuera por el carácter continuo que representa, el término que mejor le cuadraba a lo que Koestler denominó holón, es el de “homeomeron”, en el sentido de “homeomería” de Anaxágoras (Empédocles, 1996), y, tal vez, al carácter holárquico mejor le convendría el término de “jerarquía desarraigada”, en contraposición al concepto de jerarquía “dionisíaca”, por el “oxímoron” que representa el hecho de que, en una holarquía, convivan simultáneamente dos planteamientos opuestos como son el de jerarquía y la heterarquía. El filósofo neoplatónico Dionisio Areopagita, introdujo el concepto de jerarquía en un contexto teológico. El término significa, literal y etimológicamente, dominio a través de lo sagrado. Dionisio llegó a la conclusión de que el cielo estaba organizado jerárquicamente; aseguraba que había nueve niveles. Dios era la cumbre, los arcángeles formaban el consejo de dirección y Jesucristo ocupaba un puesto a la diestra de Dios. A continuación, estaban los ángeles, tronos, dominaciones, potestades, querubines, serafines, etc. Pero, para él, también el infierno era jerárquico y la estructura era similar. Lo que no determinó Dionisio, fue si el cielo significaba cooperación y el infierno confrontación o ¿que? Con los holones, se contemplan ambas alternativas. Por eso, con los holones, se puede representar tanto una simple sentencia de control hasta un sistema de agentes, pasando por un subprograma, “demonios”, etcétera.

3.2.3. Postulados de la Teoría

Definidos los elementos constitutivos de la teoría propuesta, se pasa a presentar los postulados que la soportan. Éstos son los siguientes:

- P1. Complementariedad. Todo sistema de información, biológico o artificial, se compone de sólo dos elementos complementarios: Holones e Informones. Formalmente:

$$SI = \{H, I\}$$

- P2. Cómputo. En todo sistema de información se puede establecer la “operación” de cómputo, representada por \circ , que consiste en transformar holones e informones de entrada en otro(s) de salida. Es decir, los holones e informones se interrelacionan para obtener nuevos informones y holones; de acuerdo con el esquema siguiente:

$$(\mathcal{P}(H) \setminus \{\emptyset\} \circ ((\mathcal{P}(H) \setminus \{\emptyset\}))) \rightarrow \mathcal{P}(H' \cup I' \setminus \{\emptyset\}),$$

Donde $\mathcal{P}(S)$ es el conjunto de las partes de S , que a su vez es el conjunto de todos los subconjuntos; esto es, $\mathcal{P}(H) = \{X: X \subseteq S\}$

- P3. Satisfactibilidad. Todo sistema de información se rige por criterios de satisfactibilidad y no de optimización o maximización.

3.2.4. Experimentación

Cualquier teoría que se proponga se quedaría en mera especulación si no fuera capaz de predecir y, o, explicar ciertos hechos en su dominio de aplicación. Justamente, estas predicciones y, o, explicaciones, constituyen los “experimentum crucis” que permiten, ya que no corroborar formalmente, al menos falsar la teoría. En este sentido, la teoría aquí propuesta predice lo que a continuación se explicita en los dominios siguientes:

- A. Genética. En el dominio del ADN, la teoría afirma que habida cuenta de que varios informones distintos, en este caso en forma de codones, dan el mismo aminoácido, y el código genético actúa como holón, entonces dependiendo del codón del que proceda el aminoácido éste se comportará de una manera o de otra. Como consecuencia de ello, muchas enfermedades genéticas de etiología única tienen “prima facie” como causa el codón del que proceden. Para confirmar esto se puede realizar el siguiente experimento. Cójense varios individuos con la enfermedad de etiología única X , obsérvese si todos ellos tienen en la proteína Y el aminoácido Z proveniente del codón o triplete T . Y véanse, como contraprueba, otros individuos

que no presentan dicha enfermedad y compruébese que tienen codones distintos. Si se confirma el supuesto, es obvio que queda establecida, como supone la teoría, la relación íntima, hoy aún no aceptada, ni siquiera considerada, de tripletes y enfermedades.

- B. Conciencia. Como viene de verse a lo largo de esta tesis, el premio Nobel por el descubrimiento de la estructura del ADN Francis Harry Compton Crick en 1990 y un joven colaborador suyo Christof Koch (Crick, 1990) proclamaron en Seminarios sobre Neurociencias que había llegado la hora de estudiar el tal vez más elusivo e ineludible de los fenómenos: la consciencia, convirtiéndola así en un tema legítimo para la investigación empírica. Ambos rechazaban la creencia de muchos de sus colegas en el sentido de que la consciencia no se podía definir, y menos aún estudiar. Bien al contrario, para Crick y Koch, siguiendo a William James, la consciencia o conciencia, y todas sus formas, dirigidas hacia objetos, bien del mundo sensible o altamente abstractos o internos, parecen necesitar el mismo mecanismo subyacente; es decir, un mecanismo que combina la atención con la memoria a corto plazo. Sostenían, además, que no se podía esperar alcanzar una verdadera comprensión de la consciencia, ni por otra parte, de cualquier otro fenómeno mental, tratando al cerebro como una caja negra. Para Crick el componente crucial de su definición de consciencia es la palabra atención y hace hincapié en que ésta implica algo más que un simple procesamiento de información. En sus palabras (Crick, 1988): *“Somos conscientes de que tomamos una decisión, pero no somos conscientes de lo que nos hace tomar dicha decisión”*. Esto ya lo había sugerido Sigmund Freud (Freud, 1956) cuando reflejando su inquietud sobre el inconsciente que ya se había desarrollado en el siglo XIX, planteó la hipótesis de que la mayoría de los actos humanos estaban determinados por impulsos inconscientes. Hoy la neurociencia sabe que eso es cierto, como se vio en los experimentos de Decker y colegas y Libet y colaboradores.

Para el premio Nobel de Física Murray Gell-Mann (Penrose, 1994) *La consciencia o autoconocimiento es una propiedad, como la inteligencia, que puede acabar evolucionando en los sistemas complejos adaptativos cuando alcanzan ciertos niveles de complejidad [...] . En principio, no me parece posible que nosotros los humanos podamos construir computadores con un grado de conocimiento razonable.*

Penrose parece atribuir al autoconocimiento alguna cualidad especial que hace improbable que surja de las leyes ordinarias de la ciencia.

Sin embargo, es innegable que la consciencia es un fenómeno emergente del comportamiento del cerebro, y que puede entenderse. La teoría aquí propuesta afirma que dicho fenómeno es el resultado de observar el proceso de cómputo de holones e informes por holones que los observan y categorizar a otros holones de más bajo nivel que trabajan sobre ellos. Véase el siguiente ejemplo dado por la tabla 3.2. En ella, se pueden ver los cuatro casos siguientes:

1. Se es consciente de lo que se sabe. Un grupo de holones, trabaja sobre informes de conocimiento y los holones que los manejan y da cuenta de ese trabajo. Es justamente la consciencia.
2. Se es consciente de lo que no se sabe. Holones de un cierto nivel quieren trabajar sobre informes de conocimiento y se percatan de que no existen.
3. No se es consciente de lo que se sabe. En este caso, no se han usado holones de nivel suficiente para estar en el caso 1.
4. No se es consciente de lo que no se sabe. No se han usado holones para convertir la situación en el caso 2.

Todo ello implica que en el cerebro debe haber algún “mecanismo” formado por neuronas, y, o, glía y sus respectivas actividades, que se activen siempre y cuando el cerebro reciba una actividad que requiera consciencia. Dicho mecanismo debe, anatómica y físicamente, tener muchas entradas y pocas salidas; es decir, debe ser más receptor que emisor. El candidato propuesto por Crick y Koch, es el “claustrum”. Es éste una pequeña lámina de tejido cerebral localizado bajo el córtex. Poco se conoce acerca de él, salvo que está conectado y cambia información con, casi todas las regiones sensoriales y motoras del córtex, así como la amígdala que juega un papel trascendental en las emociones. Además, Koch y Crick, comparan el “claustrum” con un director de orquesta. Es decir, las conexiones neuroanatómicas del “claustrum” reúnen las características de un verdadero director, puesto que pueden enlazar juntas y coordinar las distintas regiones sensoriales y cognitivas necesarias para la unidad de la conciencia. En suma, a día de hoy el “claustrum” es el candidato con mayor probabilidad para ser el soporte material de la consciencia, de acuerdo con

la teoría aquí propuesta. Actualmente, se está llevando a cabo con el Profesor Ortiz de la UCM, un trabajo para verificar esta hipótesis.

		CONSCIENCIA	
		Consciente	Inconsciente
C O N O C I M I E N T O	Conoce	Se es consciente de lo que se sabe. Por ejemplo, Al aprobar el carnet de conducir, se “sabe” que se sabe conducir	No se es consciente de lo que uno sabe; verbigracia, cuando uno es un conductor versado, no es consciente de que sabe conducir pues lo hace automáticamente
	Ignora	Se es consciente de lo que se ignora. Es el caso que ocurre cuando uno quiere tener el carnet de conducir	Uno es inconsciente respecto a su ignorancia. Este estadio es habitual en la infancia, lo que no es preocupante. Lo grave es cuando uno no evoluciona y se queda en este estadio.

Tabla 3.2. Consciencia Versus Conocimiento

C. Computación. Toda la computación puede explicarse con holones e informones de una manera trivial. Por ejemplo, los sistemas transaccionales son informones tipo datos (Bases de Datos) y holones clase programas. Los sistemas de información son informones tipo noticias estructuradas como sistemas de bases de datos y holones tipo programas y formulas. Los sistemas basados en conocimientos son, ni más ni menos, que informones en forma de ontologías y holones en forma de agentes. Finalmente, las Máquinas de Turing son un informón, en forma de signos 1 y 0 sobre la cinta, y un holón en forma de la tabla de actuación. Curiosamente, una Máquina de Turing específica, como, verbigracia, la de la suma, descrita en la cinta como parte de la Máquina de Turing Universal (MTU), es un informón, que trata como tal la MTU, y no un holón, como pudiera parecer “prima facie”.

3.3. Emergencia de propiedades

3.3.1. Emergencia o *fulguratio*

Como ya se ha visto en el capítulo 2, con este nombre, se describe el hecho de que, en un sistema, el todo presenta propiedades, y, o, funcionalidades que no tienen individualmente las partes. Por ejemplo, los aminoácidos de una bacteria carecen de la propiedad de *autoreproducirse*. Sin embargo, el conjunto, con otras sustancias, posee esa propiedad. Ésa es la razón de que cuando se observa un fenómeno parcialmente, al

escapar algunas variables a dicha observación, el sistema representado por el resto puede desarrollar propiedades notables, incluso *milagrosas*.

Se entiende, pues, por propiedades y, o, funcionalidades emergentes un conjunto de cualidades que no son inherentes a cada una de sus partes, ni pueden explicarse mediante interacciones lineales de propiedades elementales de los componentes. Por ejemplo, la hipótesis conexionista sostiene que la inteligencia es una *propiedad emergente* resultado del comportamiento de ciertos sistemas sencillos con características constitutivas.

Konrad Lorenz (Lorenz, 1974) el padre de la etología y premio Nobel de Medicina, hablando sobre la formación de nuevas propiedades en un sistema, teniendo en cuenta la teoría estratigráfica de Nicolai Hartman (Hartman, 1964), enfatizó tres hechos que revisten importancia en la investigación comparativa de la constitución y filogénesis de los mecanismos cognoscitivos. Estos hechos son los siguientes:

1. *Independencia de sistemas elementales*. Los sistemas sencillos y los más primitivos, son perfectamente capaces de funcionar con plena independencia.
2. *Emergencia funcional*. A menudo, aunque no siempre, surge una función nueva y compleja mediante la integración de varias funciones existentes y más simples que estaban capacitadas para funcionar por sí solas e independientemente de la integración posterior y que, lejos de desaparecer o sacrificar su transcendencia, siguen funcionando como componentes indispensables de la nueva unidad.
3. *Novedad*. Es absolutamente ocioso el buscar en los sistemas secundarios aislados de funcionamiento independiente, esas propiedades del sistema que sólo aparecen sobre los planos de una integración superior.

Ejemplos de emergencia de propiedades abundan en la naturaleza. Tal vez, el más conspicuo, sencillo y vulgar es el de la sal de mesa o sal común, un elemento nutricional importante para los seres humanos, compuesta de un gas brillante y venenoso, el cloro, y el sodio, un metal corrosivo. Otro caso es el del agua cuyas propiedades emergen de la composición química de sus elementos constitutivos, dos gases: el hidrógeno y el oxígeno. En este caso, el oxígeno es un elemento que favorece la combustión, en tanto que el agua se usa para apagar incendios.

En fin, que la capacidad de emergencia, *fulguratio* en latín, de propiedades, puede verse como un resultado del holismo (Aristóteles, 1998), que presentan los sistemas. En efecto, sistemas compuestos a veces sencillos, la mayoría no excesivamente complejos y, en general, todos los complejos, manifiestan propiedades que ni explícita ni implícitamente estaban presentes en los componentes o partes del sistema, haciendo bueno la famosa y aparentemente mística sentencia aristotélica, definitoria del holismo de que *el todo es más que la suma de las partes*.

Lo que de verdad distingue entre el comportamiento *normal* y el complejo caótico es lo que, científicamente se denomina *alinealidad*. En un sistema lineal existe una relación simple, y casi siempre explícita y evidente, entre causa y efecto, entre la entrada y la salida. Si la causa C produce el efecto E, y la causa C' produce el efecto E', C y C' juntas producen E y E' juntas. Ninguno de ambos vínculos causales interfiere en el *camino* del otro. Es decir, el comportamiento en un sistema lineal es una simple superposición de los comportamientos y sus causas. En otros términos, el todo no es más que la suma de las partes. En cambio, en un sistema *alineal*, o sea, no lineal, el todo es más, a veces mucho más, que la suma de las partes. Esto es, las causas que actúan en combinación, pueden producir, y efectivamente producen, efectos nuevos y completamente inesperados. De hecho, la relación entre la entrada y la salida, es bastante menos directa. Un ejemplo palmario de alinealidad es la retroalimentación positiva que desemboca en la producción de aludes de nieve y otros efectos perniciosos, la mayoría de las veces y, en menos pero deseables casos, en lo que se denomina *círculos virtuosos*. Es este desfase en contra *versus* a favor, por lo que se suele identificar la retroalimentación positiva con el mal y la destrucción. La realidad es que la retroalimentación positiva es *simplemente ciega*. El hecho de que se destaquen sus efectos negativos es, en el peor de los casos, una cuestión meramente estadística.

Para Daniel Hillis, (Hillis, 1995) hay una aproximación evolutiva capaz de producir la complejidad suficiente como para generar inteligencia, y esta evolución puede conseguirse utilizando los computadores. Para ello hay que disponer las cosas para que, de manera espontánea, evolucionen programas inteligentes dentro del computador.

Se comienza introduciendo secuencias aleatorias de instrucciones y se deja que estos programas compitan e interactúen entre sí, se apareen y produzcan nuevas generaciones

de programas. Si se colocan en un mundo donde su supervivencia dependa de la solución de un problema, entonces las generaciones sucesivas irán resolviendo el problema cada vez mejor, y al cabo de unos cientos de miles de generaciones resolverán el problema muy bien.

Se trata pues, de encontrar la forma de explotar determinados principios generales de organización para producir algo que va más allá de los humanos. Como ya se sabe, la ingeniería deja de funcionar bien cuando se vuelve muy complicada. Hoy haciendo uso de computadores, que se valen de un proceso muy diferente, para conseguir cosas de alta complejidad. Este proceso se alimenta a sí mismo; esto es, es autocatalítico y da la sensación, cada vez más, que los humanos caminan aceleradamente hacia una singularidad en la corriente a la creación de máquinas pensantes, el santo grial de la IA dura.

En resumen, todo parece encajar, si no se confunde ontología, lo que *es*, con epistemología, lo que *se conoce*. En este sentido, no existe contradicción, si no complementariedad, si los sistemas se consideran en los tres aspectos que inexorablemente hay que tener en cuenta cuando se pretenden entender los sistemas, a saber, el estructural, el funcional y el comportamental: la estructura, o partes constituyentes; la función, o relaciones entre dichas partes; y el comportamiento, o funcionamiento real y teleológico de los anteriores.

El caos enseña que reglas simples pueden dar lugar a comportamiento complejo. Este descubrimiento tiene un aspecto muy positivo: significa que sistemas que hasta ahora se han considerado demasiado complicados de entender, en realidad pueden estar gobernados por reglas simples. Su contrapartida, sin embargo, es que puede que el mero hecho de ser capaces de escribir las reglas para un sistema no ofrezca mucha comprensión del mismo.

Existe, también, una especie de teoría recíproca del caos que tiene, asimismo, profundas implicaciones para la comprensión de los sistemas basados en reglas. Esta compañera del caos es la teoría de la complejidad o *pléctica*, desarrollada en primera instancia por el premio Nobel de física Murray Gell-Mann (Gell-Mann, 1996). La pléctica se centra en la tendencia que presentan ciertos sistemas a mostrar comportamiento simple, simple en

cierto nivel de descripción, pero no en el de los componentes del sistema. El núcleo filosófico de la pléctica es el concepto de *emergencia*, por el que un sistema puede trascender a sus componentes, de modo que, de nuevo, se da el aristotélico *el todo es mayor que la suma de las partes*. Por ejemplo, los teóricos de la complejidad ven un derrumbamiento del mercado de valores como una respuesta emergente de un sistema monetario complejo a la acción de un número enorme de inversores individuales. Ningún inversor por sí solo puede hacer que se derrumbe el mercado. Pese a ello, cuando las interacciones entre inversores resultan encaminarse a lo largo de una trayectoria dinámica no lineal concreta, sus respuestas colectivas se refuerzan mutuamente y el colapso es el resultado inevitable.

No es coincidencia que la complejidad y el caos estén relacionados. Ambos son parte inherente de la teoría de la dinámica no lineal. En cuanto se admita la flexibilidad de la no linealidad en los modelos de la naturaleza, se encontrarán estos dos fenómenos. Y si no se admite dicha flexibilidad...., bien, la ignorancia puede dar la felicidad, pero comprada ésta a un precio que no merece la pena pagar.

Tanto el caos como la pléctica desafían una hipótesis profundamente arraigada relativa a causas y efectos que Ian Steward (Steward, 1997) denominó *conservación de la complejidad*. Según esta idea, reglas simples siempre implican comportamiento simple. Si se acepta esto, entonces se sigue que todo comportamiento complicado debe ser resultado de reglas complicadas. Esta hipótesis, habitualmente tácita, es la razón, verbigracia, de que se piense en la complejidad de los seres vivos como un enigma: ¿de dónde *procede* la complejidad? Hasta ayer mismo, nadie apenas se había atrevido a sugerir que no tiene por qué proceder de parte alguna.

Un problema con la conservación de la complejidad es que si la simplicidad se hereda directamente desde las reglas al comportamiento, entonces es difícil reconciliar un universo complejo con la simplicidad de sus reglas. La respuesta habitual es que la complejidad del Universo surge de la interacción de un gran número de componentes simples; es decir, es complejo a la manera en que son complejos una guía telefónica o un diccionario. Pero, recientemente, la idea de que la complejidad se conserva se ha enfrentado a una serie de desafíos matemáticos. Uno, es el descubrimiento del caos, en el que la complejidad surge de la interacción *no lineal* de un número pequeño de

componentes. Otro es la *pléctica* que pone el acento en lo contrario: que interacciones altamente complejas, las cuales a menudo tienen lugar en sistemas compuestos de muchos elementos idénticos, conspiran para crear pautas a gran escala pero simples: fenómenos emergentes.

Muchos de los grandes misterios de la ciencia parecen ser, cuando menos, fenómenos emergentes. Mente, forma biológica, estructural, social... Resulta tentador y realmente atractivo llegar a la conclusión de que el caos y la complejidad deben tener todas las respuestas a estos misterios. Pero realmente sólo es el inicio y primer paso de un largo, difícil y accidentado camino.

¿Qué es un sistema complejo? No hay una definición matemática consensuada, pero la idea general es que no debería ser posible describir el comportamiento del sistema de forma *inversa*, incluso si tiene elementos de organización definidos. Los sistemas complejos no son ni ordenados ni aleatorios, aunque combinan elementos de ambos tipos de comportamiento de una manera muy escurridiza aunque sorprendente. Hay muchos tipos de sistemas complejos. En efecto, la complejidad puede ser puramente:

- A) Espacial. El sistema exhibe pautas complicadas, pero éstas no varían con el tiempo. Un ejemplo es la molécula de ADN, con su doble hélice y su intrincada secuencia de *codones* o ternas de nucleótidos, aproximadamente 10^9 en los seres humanos, que determinan las operaciones químicas necesarias para producir la vida.
- B) Temporal. En este caso, la estructura espacial es sencilla en cualquier instante, pero cambia de una forma complicada a lo largo del tiempo. Un ejemplo, es el precio de mercado de algún bien, verbigracia el oro, que en cualquier instante es un simple número, pero que cambia *erráticamente* en intervalos de tiempo muy cortos. A la larga siempre parece crecer.
- C) Espacio-Temporal. Cuando, como sucede con el cerebro humano, el sistema es complejo tanto en el espacio como en el tiempo. En el espacio, con la gía y sus más de 10^9 neuronas todas ellas conectadas de un modo individualmente organizado, aunque complicado, y sus pautas de señales eléctricas y químicas en continuo cambio.

Además, los sistemas complejos pueden ser también *adaptativos*; es decir, capaces de responder a influencias externas, o incluso a los resultados de su propio comportamiento y *aprender* de ellos: cambian la respuesta. Ejemplos son los ecosistemas y el propio cerebro humano.

En este sentido, un sistema complejo no es sólo *complicado*. Las moléculas de cadena larga en un caramelo de café son tan *complicadas* como las de ADN: contienen un número similar de átomos y se necesitaría el mismo espacio para hacer la lista detallada de su estructura. Pero no son igual de *complejas*, pues carecen de los elementos organizados del ADN. Las moléculas del caramelo son básicamente tan sólo conjuntos aleatorios de átomos de C, O y H, sujetas a unas pocas reglas generales acerca de cuantos de cada tipo hay en una región dada. La diferencia entre *complejidad* y complicación es como la que hay entre el texto de Hamlet y una tabla de números aleatorios.

Un área donde la *complejidad* está arrojando luz es la teoría de la evolución. Los biólogos llevan mucho tiempo intrigados por la capacidad de los sistemas vivos para hacerse cada vez más organizados, en aparente violación de la Segunda ley de Termodinámica, que establece que cualquier sistema cerrado en equilibrio térmico aumentará su entropía; esto es, se hará cada vez más desordenado. Parte de la respuesta al enigma es que los sistemas vivos no son cerrados ni están en equilibrio; toman energía de fuentes exteriores y están continuamente en movimiento. Pero también los océanos toman energía y se mueven de modo que estos únicos rasgos no explican el comportamiento extrañamente teleológico; o sea, finalista, de las criaturas vivientes.

3.3.2. Clasificación de fenómenos emergentes: *Simplejidad* y *Complicidad*

Para afrontar la complejidad, se necesita una teoría razonable o efectiva de la *emergencia*. La respuesta no está en abordar frontalmente la complejidad tal como se intentó en el proyecto del genoma humano. Eso nuevamente retrasa el colapso inexorable que se producirá cuando sea abrumador el exceso de información, en forma sobre todo de datos y algunas noticias, imposibles de organizar como conocimientos. Se precisa una teoría más holista en la que pautas de alto nivel pueden entenderse sin remitirlas a reglas de bajo nivel. Steward propone un nombre para ello: *informática*, pero inmediata y humildemente añade que no tiene un ello para unir al nombre.

Las matemáticas actuales, no obstante, muestran que una tal teoría es posible, en principio. La teoría moderna de los sistemas dinámicos es a la vez reduccionista y holista. Es reduccionista cuando utiliza ecuaciones diferenciales para reducir el flujo de fluidos a las interacciones de un número enorme de cubos minúsculos. Pero tiene también otras armas en su armería: principios cualitativos a gran escala, como continuidad, conectividad y simetría. Dispone tanto de la aritmética de bajo nivel de las ecuaciones como de la geometría de alto nivel de los atractores. Por eso, en esta área de la matemática ambos enfoques se encuentran en el centro, y se fusionan en algo más poderoso que cualquiera de los dos por separado. Sin embargo, su poder es limitado. No existe todavía un enfoque comparable a la ciencia en general.

La teoría de la complejidad es un paso, en dicha dirección, pero adolece de una visión demasiado limitada de la naturaleza del problema. Cohen y Steward (Cohen, 1991), clasificaron los fenómenos emergentes en dos tipos cualitativamente deferentes: *simplejidad* y *complicidad*. La *simplejidad* es emergencia del tipo carretera *hormiga de Langton*: una pauta a gran escala que ocurre dentro de un sistema basado en reglas pero cuya deducción detallada, a partir de las reglas, es enormemente larga y poco informativa o quizá desconocida. Otro ejemplo de la ciencia real es la creencia de los físicos en que la estructura de las redes cristalinas es una consecuencia de la MC. Hay todo tipo de evidencia indirecta de que esto es cierto, pero no hay ninguna demostración conocida. Precisamente ahora, la teoría de la complejidad trata principalmente de *simplejidades*.

La *complicidad* es mucho más escurridiza pero también mucho más importante. Ocurre cuando interaccionan dos, o más, sistemas basados en reglas. En tales circunstancias no es insólito encontrar que nuevas regularidades de alto nivel emergen de las interacciones. A menudo estas interacciones son desconocidas, puesto que no son parte de las reglas de ninguno de los subsistemas. Un ejemplo de *complicidad* en el mundo real es la evolución del succionado de sangre. Esto ocurrió cuando las reglas para la anatomía de los mamíferos primitivos (sangre) interaccionaron con las reglas para un ancestro del mosquito, que había desarrollado un órgano para succionar líquidos, probablemente agua. Por una coincidencia creativa, resultó que la trompa succionadora de agua era capaz de penetrar en la piel humana. Esta colisión de dos espacios de desarrollo hizo que coevolucionaran de una nueva manera, no inherente a ninguno de ambos espacios de desarrollo por separado. El resultado de esta coevolución cómplice es un insecto que está

adaptado para succionar sangre *humana*. Una consecuencia es la malaria una *simplejidad* más predecible erigida en la cima de la *complicidad* inicial. Existen reglas simples porque la simplicidad *emerge* de interacciones complejas en niveles de descripción más bajos. El Universo es una pluralidad de reglas que se solapan. Y en los huecos que quedan entre las reglas está el *País de las Hormigas*, en el que simplicidad y complejidad no sólo se conservan, sino que se transmutan una en otra.

Un caso especial de emergencia se da en lo que Lyn Margulis consideraba la variación hereditaria más significativa (Margulis, 1995), la procedente de las fusiones, de lo que los rusos, especialmente Konstantin S. Mereschkovsky, denominaban *simbiogénesis* y el norteamericano Ivan Emanuel Wallin llamaba *simbioticionismo*. Wallin aplicaba el término a la incorporación de sistemas genéticos microbianos en los progenitores de células animales o vegetales. El nuevo sistema genético, una fusión entre microbio y célula animal o microbio y célula vegetal, es ciertamente distinto de la célula ancestral. En analogía con la tecnología informática, en vez de partir de cero creando módulos totalmente nuevos, la idea de la simbiosis equivale al acoplamiento de módulos preexistentes. Las fusiones dan como resultado seres nuevos y más complejos. Y añadía Margulis que dudaba que se formaran nuevas especies sólo a partir de mutaciones aleatorias.

Como es sabido, la simbiosis es una asociación física entre organismos con beneficios mutuos; es decir, la convivencia a conveniencia de diferentes especies en el mismo tiempo y lugar. El trabajo de Margulis sobre la simbiosis partía de los sistemas genéticos citoplasmáticos. A todos se les enseña que los genes están en el núcleo y que éste es el control central de la célula. Sin embargo, ella pronto se percató de que existen otros sistemas genéticos con pautas hereditarias diferentes. Resultándole sumamente intrigantes los revoltosos genes fuera del núcleo. El ejemplo más famoso es un gen citoplasmático, bautizado como *killer*, presente en la protista *Paramecium Aurelia*, cuya herencia seguía ciertas reglas. Tras veinte años de investigación, resultó que el gen *killer* era un virus dentro de una bacteria simbiótica. Casi todos los genes extranucleares derivan de bacterias u otra clase de microbios. Esto llevó a Margulis a la convicción de que eran entidades cohabitantes, seres vivos. Pequeñas células vivas que residen dentro de células mayores. Esto condujo al estudio de la simbiosis moderna.

Lo importante en este caso, es la *simbiogénesis*; esto es, simbiosis a largo plazo que conduce a nuevas formas de vida y sobre si dicha simbiogénesis ha ocurrido y sigue ocurriendo. En suma, lo relevante es la simbiogénesis como fuente principal del cambio evolutivo. O sea, la simbiogénesis es el resultado de una convivencia a largo plazo, especialmente entre microbios, y es la principal fuente de innovación evolutiva en todos los linajes de organismos superiores no bacterianos.

Margulis (Margulis, 1966), trató la cuestión del origen de todas las células no bacterianas. El origen de las células bacterianas es el origen de la vida misma. Tras diez años de investigación escribió un libro, en el que se detalla el papel de la simbiosis en la evolución celular, que conduce directamente al origen de la división celular mitótica y de la sexualidad meiótica (Margulis, 1970 y 1993). La cuestión principal es cómo bacterias diferentes forman consorcios que, bajo presiones ecológicas, se asocian y experimentan cambios metabólicos y genéticos tales que de la estrecha integración de sus comunidades surge la individualidad en un nivel de organización más complejo. Lo que se considera es el origen de las células nucleadas (protoctistas, animales, hongos, plantas verdes) a partir de las bacterias.

Durante miles de millones de años, las únicas formas de vida en la Tierra fueron las células bacterianas carentes de núcleo, células procariotas o procarióticas. Todas se parecen y desde la atalaya humana parecen aburridas. Nada más lejos de la realidad, las bacterias son la fuente de la reproducción, la fotosíntesis, el movimiento, en fin de todas las propiedades interesantes de la vida excepto, quizás, el habla. Todavía existen en gran número y variedad y gobiernan la Tierra. En cierto momento, apareció en escena un nuevo tipo de célula más complejo: la célula eucariótica, la que forma los cuerpos animales y plantas. Estas células contienen ciertos orgánulos, entre ellos un núcleo diferenciado y son los elementos de construcción de todas las formas de vida familiares. ¿Cómo se produjo esta revolución? ¿Cómo apareció la célula eucariótica? Probablemente se originaron a partir del ataque de un predador, cuando alguna forma bacteriana oscilante invadió a otra por supuesto, en busca de alimento. Algunas de estas invasiones evolucionaron, de modo que asociaciones antes feroces y, o, deletéreas se hicieron benignas. Cuando estas bacterias invasoras se establecieron en el interior de sus indolentes huéspedes, la unión de las fuerzas de ambos creó un nuevo conjunto que era, en efecto, mucho más que la suma de sus partes y evolucionaron nadadores más rápidos

capaces de transportar enormes cantidades de genes. Algunos de éstos recién llegados se mostraron singularmente competentes en la lucha evolutiva. A medida que la célula moderna evolucionaba se adicionaron nuevas asociaciones bacterianas.

Una de las evidencias a favor de la *simbiogénesis* son las mitocondrias, orgánulos presentes en la mayoría de las células eucariotas y que tienen su propio ADN. Además, del ADN nuclear, que compone el genoma humano, los humanos tienen un ADN mitocondrial. Las mitocondrias humanas constituyen un linaje completamente distinto, pues se heredan exclusivamente de las madres. No hay ADN mitocondrial paterno. Así en todo hongo, animal o planta, y en la mayoría de protoctistas, existen al menos dos genealogías paralelas. Esto, en sí mismo, es un indicio de que en algún momento estos orgánulos fueron microorganismos de vida libre.

David Luck y John Hall, han encontrado, en los 90 del siglo XX, algo que ni siquiera estaban buscando; esto es, serendípicamente, un ADN peculiar exterior al núcleo, a los cloroplastos y a las mitocondrias. Este ADN, predicho, más o menos, un cuarto de siglo antes por Margulis, puede interpretarse como un vestigio de antiguas bacterias invasoras, cuya asociación inicialmente agresiva presagiaba la integración. Este sistema genético no nuclear, que Hall y Luck, revelaron en las algas verdes podría representar lo que en otro tiempo fueron bacterias. Para Margulis, existe un isomorfismo entre el crecimiento, reproducción o comunicación de estas bacterias móviles formadoras de alianzas y el pensamiento, alegrías, sensibilidades y estímulos humanos. Si Margulis está en lo cierto, estas bacterias simbiotes organizan una gran cantidad de conocimientos. Esta profusión de seres parciales representa la base física del *anima*: alma, vida, locomoción, una defensa del materialismo en el sentido más craso. De hecho, cuando se juntan dos clases diferentes de proteínas móviles, se mueren. Se desplazan por sus propios medios. Los bioquímicos y biólogos celulares pueden mostrar el número común denominador del movimiento de la locomoción. *Anima*. Alma. Una interpretación plausible de éstas proteínas móviles es que son reliquias de las bacterias nadadoras incorporadas por los seres que se convirtieron en ancestros de los humanos.

El sistema de movimiento mínimo es tan caracterizable, física y químicamente, que existe un consenso total en cuanto a la composición de las *proteínas de la motilidad*. Bioquímicos y biólogos celulares están de acuerdo en todos los detalles. Ahora sólo falta

comprender cabalmente el papel de la simbiogénesis en el origen evolutivo de esta estructura. Esta comprensión seguramente arrojará nueva luz sobre la base física del pensamiento. De hecho, el pensamiento y el comportamiento humanos resultan mucho menos misteriosos cuando se comprueba que la capacidad de elección y la sensibilidad están ya exquisitamente desarrolladas en las células microbianas que se convirtieron en los ancestros humanos.

Una de las cuestiones básicas relativas al origen de la vida es entender que un organismo o sistema vivo consiste, a grandes rasgos, en una relación simbiótica entre un metabolismo y un replicador. El metabolismo, compuesto por proteínas y demás, extrae energía del entorno, y el replicador contiene la impronta del organismo, con la información necesaria para crecer, hacer reparaciones y reproducirse. Ambos se necesitan: el replicador contiene la información necesaria para producir las proteínas, el ADN y las otras moléculas que constituyen el metabolismo y proporcionan la fuerza motriz del organismo, y el metabolismo suministra la energía y materias primas necesarias para la construcción y función del replicador. La cuestión es ¿quién fue primero, el metabolismo o el replicador? ¿o es que no pueden existir el uno sin el otro, de manera que tuvieron que evolucionar juntos? El químico Harold Urey y el biólogo Stanley Miller, allá por los años 50 del siglo XX, demostraron que los sillares de las proteínas, los aminoácidos, se forman espontáneamente a partir de *tierra, fuego y agua*. Sin embargo, la síntesis de moléculas más complicadas requería replicadores, de manera que el origen de los metabolismos estaba mucho menos claro (Farmer, 1996).

Rick Bagley y J. Doyne Farmer demostraron que los metabolismos podían emerger espontáneamente a partir de moléculas básicas y evolucionan sin la presencia de un replicador. Esto es, los replicadores serían ellos mismos y la información estaría almacenada simplemente en la llamada *sopa primordial*. Partiendo de componentes simples, verbigracia aminoácidos, se trataba de obtener proteínas complejas; es decir, cadenas largas y de alta diversidad de aminoácidos. El principio básico de una red autocatalítica es que, aunque ninguno de sus componentes puede producirse por sí solo, cada uno es producido por al menos una reacción entre algunos de los otros ingredientes del caldo. Se trata de un sistema simbiótico, en el que todo coopera para que el metabolismo funcione, el todo es más que la suma de sus partes. Si la replicación normal es como la monogamia, la reproducción autocatalítica es como una orgía. Todos estos

procesos eran simulados en computador. Luego, en unión con Walter Fontana, demostraron que el sistema podía evolucionar: nuevas *proteínas* emergían espontáneamente, compitiendo con las presentes y alterando el metabolismo.

3.3.3. Sistemas Constitutivos

Los sistemas pueden ser vistos tanto desde una perspectiva científica como desde una tecnológica. Desde el punto de vista científico, cualquier disciplina, y los sistemas no son una excepción, tiene por objeto validar las ideas, es decir, demostrar que las ideas son correctas y viables, aunque sólo sea en un sistema *ad hoc*. Desde el punto de vista tecnológico, aparte de que las ideas que sean viables, los productos también tienen que funcionar y, lo que es más importante, de forma rutinaria.

El verdadero problema se encuentra dentro de los sistemas como técnica; es decir, en la posibilidad de construir artefactos que pueden emular las capacidades funcionales del cerebro humano, lo que especialmente suele ser clasificado como el pensamiento.

A continuación, se mostrará cómo un sistema no excesivamente complejo es capaz de poseer propiedades que no estaban, ni explícita ni implícitamente, presentes en las partes del mismo, haciendo buena la famosa frase, y aparentemente la oración mística que define el holismo: *el todo es mayor que la suma de sus partes*.

En realidad, como es señalado por Bertalanffy, elementos complejos pueden, como se muestra en figura 3.7, clasificarse según (a) su número, (b) su tipo y (c) las relaciones entre elementos. En los casos (a) y (b) el conjunto se puede entender como una suma de elementos tomados por separado. En el caso (c) tanto los elementos como sus interrelaciones deben conocerse. Las características del primer tipo se denominan *sumatorias* y del segundo, *constitutivas*. Otro punto es que las características *sumatorias* sean las mismas dentro y fuera del conjunto, en donde si se suman las características como un todo, los elementos se comportan como si estuvieran separados. Las características constituyentes dependen de las relaciones específicas dentro del conjunto. Para entender estas características, se necesita conocer tanto las partes como sus relaciones. Las características físicas del primer tipo son, por ejemplo, el peso y el calor, consideradas como una suma de movimientos de las moléculas. Un ejemplo de

características del segundo tipo son las características de isómeros, cuando la composición general de los compuestos es lo mismo pero los átomos están dispuestos de manera diferente en la molécula.

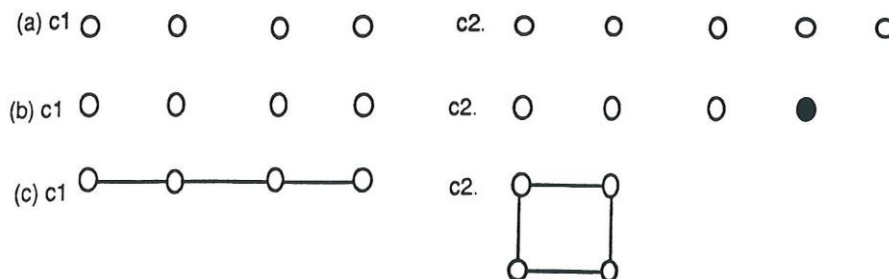


Figura 3.7. Clasificación Compleja.

Características constitutivas no pueden ser explicadas por las características de partes separadas. Así, el conjunto cuenta con características nuevas o emergentes, en comparación con sus elementos. Sin embargo, si la suma de las partes contenidas en un sistema y la relación entre ellos es conocida, el comportamiento del sistema se puede deducir de sus partes. En otras palabras, aunque la composición gradual de una suma es posible, un sistema, como una suma de partes interrelacionadas, tiene que ser concebido como instantáneamente integrado.

Como se ha indicado, el holismo, en este caso en forma de *fulguratio*, es posible en computadores, a lo cual se debería añadir, al menos en teoría es importante pero no suficiente. En ciencia en general y particularmente en los sistemas, la teoría debe guiar la experimentación y esto, a su vez, debe probar o refutar la teoría mediante la presentación de pruebas. En este sentido, los ejemplos anteriores dejaban fuera de duda el propósito del autor y también la viabilidad de holismo en los computadores. Sin embargo, el hecho de que las ideas funcionan no necesariamente significa que el problema ha sido resuelto. El holismo es necesario aplicarlo en la tecnología, de lo contrario sería *ad hocismo*. De hecho, si las cosas se quedaran así, se estaría, en contra del problema general de los viajeros del comercio, en el reducido a un vendedor particular ambulante con un particular itinerario, lo cual es fácil de resolver.

Por supuesto, la importancia de la emergencia o *fulguración* alcanza su esplendor cuando las funcionalidades emergentes son relevantes y significativas, lo que, como se explicó

anteriormente, es sólo el caso con el último ejemplo propuesto. Pero esto es una pregunta de grado o, mejor dicho, una cuestión cuantitativa. Por lo tanto, una vez que la viabilidad de esta propiedad ha sido científicamente demostrada, lo único que queda es implementar su “rutina”, es decir, tecnológicamente y al gusto del consumidor. Así las cosas, si se lograra una fulguración tecnológica, el problema que se presenta no es ni tecnológico ni científico, sino ético y sociológico: la capacidad de los computadores para actuar por sí mismos. En otras palabras, aquí se tienen las facetas del mito del Golem: por un lado, si una criatura puede superar a su creador y, por otra parte, si una máquina puede actuar independientemente de su creador. Pero, de nuevo, ésta es otra historia.

Los filósofos teístas y los místicos medievales acuñaron el término *fulguratio*, rayo, para enunciar el acto de una nueva creación. Sin duda, según comentó Konrad Lorenz, para expresar así la influencia directa ejercida desde arriba, desde Dios. Por una casualidad etimológica, ese término acierta a representar la emergencia de la *nada*.

Hay un hecho suficientemente conocido en ciencia: hay que validar las ideas; es decir, demostrar que las ideas funcionan. Así pues, aquí, se dará a continuación un ejemplo, que demuestra que la emergencia de propiedades *fulguratio* dentro del holismo, es viable. Dicho ejemplo se puede observar en la figura 3.8, cuando dos sistemas independientes entre sí, se enlazan, surgen de improviso unas propiedades inéditas del sistema; es decir, unas propiedades que no existían antes y de las cuales tampoco se había entrevisto el menor indicio. De acuerdo con Hassenstein (Hassenstein, 1966) este ejemplo ilustra el concepto de “propiedad sistémica” y, en particular, la propiedad de emergencia o *fulguración*.

En efecto, sean los dos circuitos (a) y (b) de dicha figura. Como se ve, el primero está compuesto por un conductor que enlaza los dos polos de una batería de fuerza electromotriz E_0 , por ejemplo, de potencial V_0 al que se le incorpora una resistencia óhmica R y un condensador C . En el segundo, se cambia el condensador, por una bobina de inducción. Al medir la tensión V existente entre los bornes al cerrar el interruptor I , se obtienen los respectivos diagramas que aparecen a la derecha de la figura. En (a) el condensador se carga paulatinamente venciendo la resistencia R hasta alcanzar la tensión V_0 . En (b), la corriente eléctrica, reprimida al principio por la autoinducción de la bobina, se incrementa durante el tiempo necesario para alcanzar la intensidad determinada por la

ley de Ohm; entonces la tensión de V es teóricamente nula, porque la resistencia total del circuito se concentra en R .

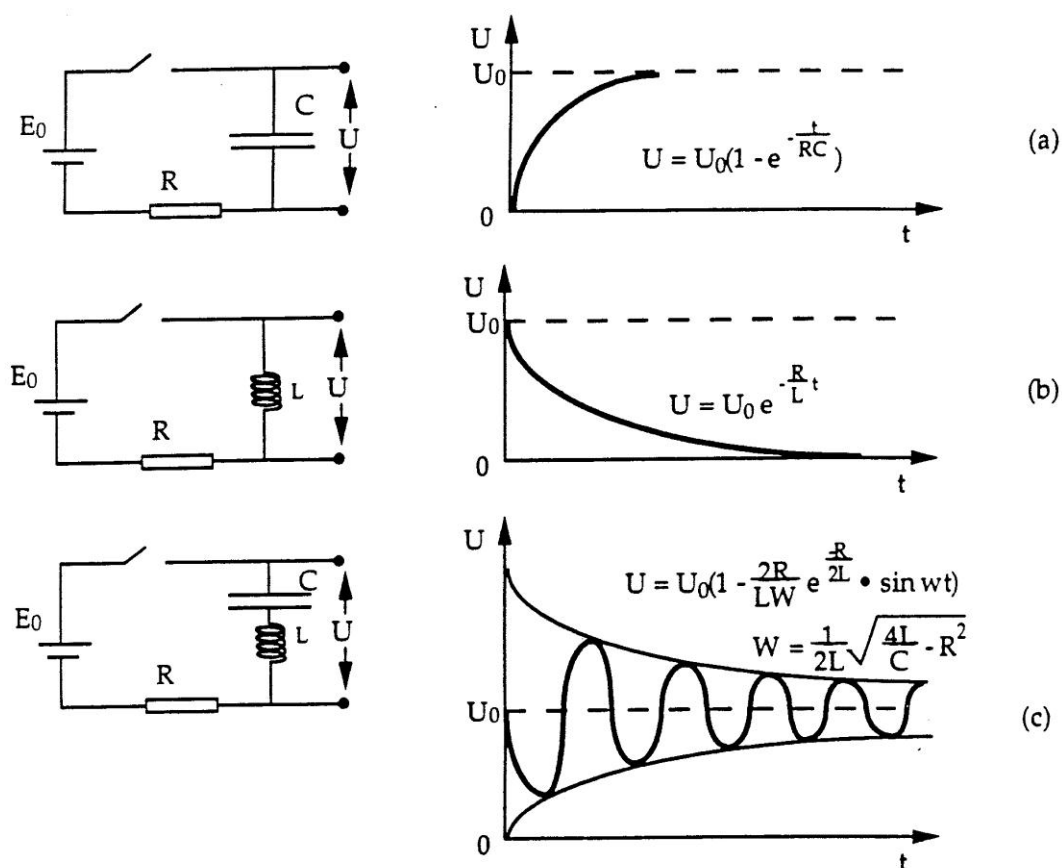


Figura 3.8. Ejemplo de *Fulguratio* en electricidad.

Si, como se muestra en (c), se unen los dos circuitos anteriores, creando un nuevo circuito que contiene, además de la batería, la autoinducción y la resistencia, entonces se obtiene una oscilación de V , amortiguada en el tiempo. Es evidente que el comportamiento de este nuevo circuito no resulta de la superposición de (a) y (b) sumados, a pesar de que el sistema (c) pueda pensarse como combinación de ambos. Finalmente, señalar que dicho esquema es válido, verbigracia, para los valores siguientes:

$$C = 0,7 \times 10^{-9} F; L = 2 \times 10^{-3} Hy; R = 10^3 \Omega; \lambda = 1,2 \times 10^{-6} s.$$

Este último valor, que es el mismo para las tres curvas, define también el eje temporal.

En suma como lo señaló Konrad Lorenz (Lorenz, 1974) una cadena lineal de elementos puede, en ciertos casos, generar un sistema cuyas propiedades difieren, no gradualmente, sino radicalmente de las de sus predecesores dando lugar a emergencia de propiedades o a una “*fulguratio*” en el sentido literal del término.

Con este ejemplo material de emergencia de propiedades en sistemas, se pone en evidencia, en este texto reiterada, la máxima aristotélica, paradigmática del holismo, siguiente: “El todo es más que la suma de las partes”.

3.3.4. Emergencia de Propiedades: Casos de estudio Computacionales.

Para demostrar que las ideas funcionan, se darán varios ejemplos de programación, que demuestran que el holismo es viable.

Caso 1

Supóngase que se tiene, como se muestra en la figura 3.9, una estructura de árbol de datos binario, almacenado como una lista de tres ítems (el padre, hijo de la izquierda, hijo de la derecha)

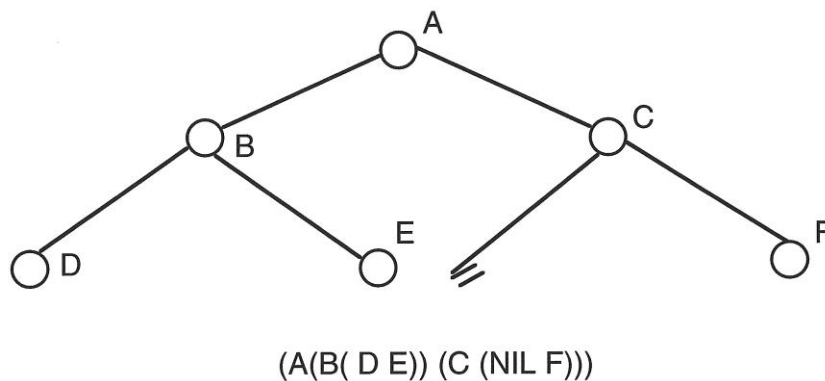


Figura 3.9. Árbol Binario.

Considérense dos funciones, preorden y postorden, las respectivas misiones de introducir un nuevo elemento en el árbol como hijo-izquierdo del nodo terminal, ubicado en la parte izquierda, e introducir un nuevo elemento en el árbol como hijo derecho del nodo terminal, situado en el lado derecho. Su uso puede variar considerablemente dependiendo del tipo de datos almacenados por estos árboles. Si, por ejemplo, esta estructura fuera a

utilizarse para almacenar una lista ordenada de elementos, puesto-preorden sería utilizado para introducir el menor y poner postorden para introducir el elemento principal en la lista. También serían útiles para el manejo del árbol como una pila (estructura FIFO) o cola (LIFO estructura) A continuación, se presenta, el código LISP de estas dos funciones.

```
(DEFUN -poner en orden previo (elementos del árbol)
  (cond ((null árbol) list de elementos del árbol)
        ((átomo árbol) (list de elementos del árbol nil))
        (T (list (car árbol)
                  (put-preorden elemento (cadr árbol))
                  (caddr árbol))))))
```

```
DEFUN put-postorden (árbol de elementos)
  (cond ((null árbol) list de elementos del árbol)
        ((átomo árbol) (list de elementos del árbol nil))
        (T (list (car árbol)
                  (cadr árbol)
                  (put-postorden elemento (caddr árbol))))))
```

La integración de estas dos funciones, respetando su estructura y sin repetir líneas de código similares, daría la siguiente función como resultado. Esto ha sido llamado puesto-preorden, como uno de sus componentes, de modo que se sea capaz de respetar el mecanismo recursivo.

```
(DEFUN - poner-en preorden (elementos del árbol)
  (cond ((null árbol) (list de elementos del árbol))
        ((átomo del árbol) list de elementos del árbol nil))
        (T (list (car árbol)
                  (cadr árbol)
                  (put-postorder elemento (caddr árbol))
                  (put-preorder elemento (cadr árbol))
                  (caddr árbol))))))
```

Si esta función se aplica al árbol en la figura 3.9, para introducir el elemento G, además de aumentar el factor de ramificación del árbol, ahora ya no es binario, e introduciendo G en varios puntos generaría una subrama inesperada en el árbol.

El árbol resultante se presenta en la figura 3.10, donde la subrama 1 debería ser el resultado de la función inicial, poner-preorden. Y la subrama 2, el resultado esperado de la función inicial puesto-postorden, cuya subrama 3 es el fruto de la integración y, por lo tanto, un resultado emergente holístico.

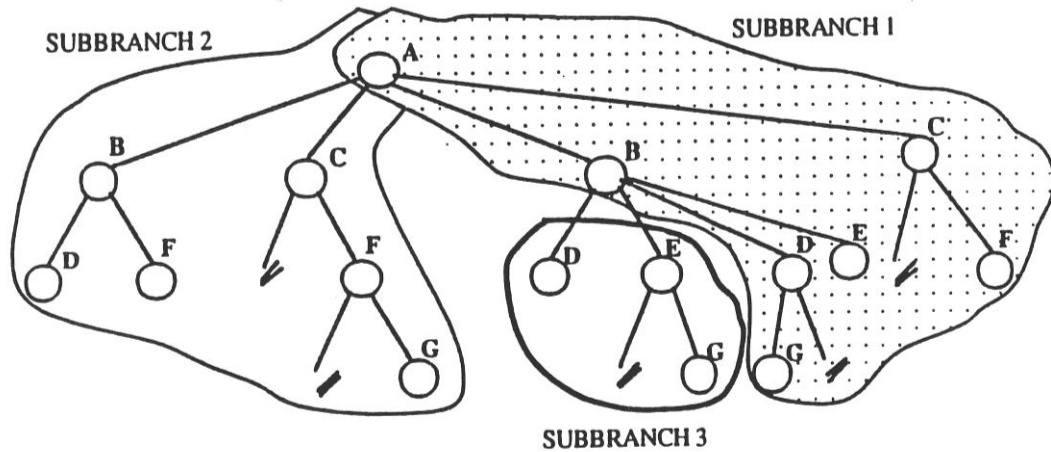


Figura 3.10. El árbol resultante de la integración de las funciones puesto-preorden y puesto-postorden.

Es claro que las dos funciones originales no dan ninguna indicación de que la subrama 3 podría aparecer. Por lo tanto, es obvio que la totalidad, que es la integración de las dos funciones principales, es mayor que la suma de sus partes. Sin embargo, realmente pudiera parecer que nada nuevo surgió, dado que las funciones tratan con árboles binarios y el cambio de un árbol binario a uno más ramificado no implica un cambio cualitativo importante. Ni la aparición de un nuevo subárbol es, al menos, cualitativamente hablando, una emergencia en sentido estricto, es más bien un cambio cualitativo o, a lo sumo, topológico. Por ello, a continuación, se dará un ejemplo implicando tales cambios cualitativos.

Caso 2

Supóngase ahora que se tienen dos funciones aritméticas, basadas en bucles, implementadas iterativamente. Se continuará con el lenguaje LISP, a pesar de que no es el más adaptado a la programación iterativa. Estas funciones iniciales serán la multiplicación y suma de dos números, definidos como sigue.

```
(DEFUN multiplication (n m)
  (setq a 0)
  (dotimes (x n)
    (dotimes (y m)
      (setq a (+ a 1))))))
(DEFUN add (n m)
  (setq y m)
  (dotimes (z n)
    (setq y (+ y 1))))
```

La integración de estas dos funciones, respetando sus estructuras, sin repetir líneas similares de código y colocando la función de bucle suma dentro del cuerpo de los bucles de la función multiplicación, dará como resultado el siguiente código.

```
(DEFUN unión (n m)
  (setq a 0)
  (setq y m)
  (dotimes (x n)
    (dotimes (y m)
      (dotimes (z n)
        (setq a (+ a 1)
              (setq y (+ y 1))))))
  (format nil "The values obtained are ~d and ~d" a y))
```

Como muestra la función unión, el índice del segundo bucle de multiplicación y el valor acumulado de la adición, reciben el mismo nombre. Esto es precisamente donde se alcanza la estructura integrada, la cual está ligada con la integración funcional obtenida por secuenciación sin respetar las dos funciones originales, lo que conduce a lo buscado después de la emergencia.

Como resultado de la integración de las funciones $F(n, m) = n * m$ y $G(n, m) = n + m$. Se obtiene una tercera función H la cual se equipara con la siguiente fórmula.

$$H = n^2 * \left[E \left(\frac{m-1}{n+1} \right) + 1 \right]$$

Donde E es la parte entera de la función. El resultado devuelto por la función *unión* en la variable a es una verdadera sorpresa, teniendo en cuenta las fórmulas iniciales, que de ninguna manera llevan a sospechar este resultado final. Sin embargo, nadie esperaría que una función de suma integrada con una función de multiplicación diera esta función. Esto demuestra lo viable que es el holismo en los sistemas complejos.

Caso 3

Finalmente, el caso más interesante, es el siguiente. La máquina elige las tres funciones elementales siguientes:

$$x_t = k \quad (1)$$

$$x_t = x_{t-1} \quad (2)$$

$$x_t = (x_{t-1})^2 \quad (3)$$

La primera, es una función muy sencilla que expresa el estado de un sistema estático, invariable en el tiempo. La segunda, expresa el estado actual como el cuadrado del estado inmediatamente anterior. Finalmente, la tercera es una sucesión recurrente que expresa el estado x_t de un determinado sistema en un momento concreto en el tiempo t , en función del instante anterior $t-1$.

Al analizar las dos últimas funciones, se ve que tienen comportamientos muy sencillos. La función (2) es invariante en el tiempo y, por lo tanto, representa, en todo momento, un estado igual al estado inicial x_0 , que se supone comprendido en el intervalo $(0,1)$. La función (3) representa un estado como el cuadrado del estado anterior, por lo tanto, con la suposición de que el estado inicial pertenece al intervalo $(0, 1)$, se tiene que el estado que representa esta sucesión tiende hacia el valor límite $I = 0$. Si se realiza una composición de (2) y (3) a través de la operación diferencia, se encuentra la sucesión:

$$x_t = x_{t-1} - (x_{t-1})^2 = x_{t-1} (1 - x_{t-1}) \quad (4)$$

Esta sucesión (4) tiene un comportamiento uniforme e independiente del valor inicial X_0 , siempre que éste se encuentre comprendido dentro del intervalo abierto $(0, 1)$. En tal caso, el valor de x_t tiende rápidamente al valor límite 0, según se incrementa el valor de t . Esto se puede comprobar empíricamente sin más que calcular unos pocos valores de x_t .

La demostración es la siguiente: Se sabe que una de las propiedades de las sucesiones es: que toda sucesión monótona creciente que está acotada, converge hacia un punto que coincide con el supremo de sus cotas inferiores (5). Se puede demostrar fácilmente que la sucesión (4) es monótona decreciente una vez que, reformulando la expresión (4), se obtiene la expresión equivalente siguiente:

$$x_t = x_t / (1 - x_{t-1}) \quad (5)$$

donde de forma evidente se obtiene que:

$$x_t / (1 - x_{t-1}) > x_t$$

puesto que x_{t-1} pertenece al intervalo (0, 1), y por lo tanto

$$x_{t-1} > x_t$$

Por otra parte, resulta también fácil de demostrar que 0 es el supremo de las cotas inferiores, puesto que siempre se puede encontrar un x_t , tan cercano a 0 como se quiera, pero nunca menor que 0. Demostradas las dos condiciones expresadas por (5), se puede dar por comprobado que, efectivamente, el límite al que tiende x_t es 0.

En cuanto a la sucesión expresada por (1) $x_t = k$, se trata de una sucesión que adopta un valor constante para todo valor de t y, por lo tanto, tiende a ese valor constante k .

Pues bien, la máquina combinó dichas funciones en una única mediante el producto, obteniendo la nueva función x_t definida por la siguiente expresión:

$$x_t = k x_{t-1} (1 - x_{t-1})^t \quad (6)$$

Si se intenta definir el comportamiento de (6) tomando como base el comportamiento individual de sus componentes (1) y (4), se podría pensar que si dos funciones tienden a 0 y a una constante, su producto tenderá al producto de los límites, es decir, 0. Sin embargo, el comportamiento de (6) no se parece en nada a lo que sería de esperar, produciéndose, por lo tanto, la emergencia de funcionalidades no contempladas en ninguno de los componentes individuales, que enlaza con uno de los resultados más sorprendentes obtenido en el pasado inmediato en la teoría del Caos.

En suma, que el comportamiento descrito para la sucesión x_t , en (6) por complicado y rico, es absolutamente impredecible a partir de las sucesiones iniciales (1), (2) y (3) y de las operaciones diferencia y producto, utilizada para la composición. Es, por lo tanto, un caso claro de holismo en el sentido tratado en este trabajo.

En efecto, uno de los descubrimientos más sorprendentes de los últimos años es que algunos sistemas deterministas muy simples, con pocos elementos y reglas de actuación sencillas y fijas que no encierran en sí mismas ningún elemento de azar, pueden generar un comportamiento tan complejo y difícil de comprender y predecir que podría calificarse de aleatorio. A la aleatoriedad así generada se la denomina caos, comportamiento caótico, turbulento o estocástico. Considérese el sistema obtenido en (6):

$$x_t = k x_{t-1} (1 - x_{t-1})^t \quad (6)$$

donde k es un parámetro que vale más de 1 y menos de 4. Puede comprobarse que si el valor de x_{t-1} está comprendido entre 0 y 1, el nuevo estado x_t también lo estará, por lo que si x_0 se encuentra en el intervalo abierto (0,1) el estado del sistema siempre será un valor en ese intervalo. Sin embargo, el comportamiento de dicha función que, a primera vista, parece carente de todo interés es, realmente, bastante interesante y muy sorprendente, puesto que, a pesar de la simplicidad de la ecuación que define el estado del sistema, ocurre que pequeñas variaciones en el valor del parámetro k , dan lugar a cambios espectaculares en su comportamiento. Si k es superior a 1 pero inferior a 3, las sucesiones de x_t son muy regulares, tendiendo al valor $(k-1)/k$, independientemente del valor de x_0 . Dado que este valor parece atraer a todas las sucesiones que pueden generarse con distintos valores iniciales de x_0 , recibe el nombre de atractor. Cuando k supera el valor $k_1 = 3$, sucede algo curioso, pues el valor de x_t , aún siendo regular, oscila entre dos sucesiones que tienden a dos valores fijos x_1 y x_2 ; por ello, se dice que el atractor ha duplicado su dimensión y el estado del sistema oscila entre cuatro valores; si se incrementa aun más el valor de k , aparecen nuevos valores k_3, k_4, \dots, k_n para los que el atractor prosigue duplicando su dimensión por lo que los valores de x_t n tienden a un atractor de dimensión, 2^n . Hasta aquí, el comportamiento del sistema va complicándose, pero sigue siendo regular y definido por un movimiento oscilatorio entre 2^n estados. Sin embargo, si k supera el valor crítico $k_c = 3,569945\dots$, el valor de x_t irá recorriendo de forma totalmente caótica y errante todo el intervalo (0,1), sin que siga ninguna pauta reconocible; el atractor se vuelve caótico, causando que la sucesión de estados del sistema

sea impredecible y merezca el calificativo de aleatoria. En el comportamiento de este sistema pueden observarse dos zonas bien diferenciadas en las que la capacidad de predicción varía notablemente. Para cualquiera de los valores de k inferiores a k_c , el sistema es determinista y un pequeño error en la medida del estado actual del sistema no impide hacer predicciones razonables sobre el estado futuro del mismo. Sin embargo, si el parámetro k supera el valor crítico, el sistema se vuelve caótico y cualquier error infinitesimal en la medición del estado actual x_t impedirá obtener, incluso, una predicción aproximada.

La comprensión de este fenómeno está resultando crucial para ayudar a entender la capacidad de predicción de la ciencia. Hay fenómenos, tales como los movimientos orbitales de los planetas, que pueden predecirse con gran exactitud a pesar de los errores que, inevitablemente, se producen en la medición del estado en un instante dado. Sin embargo, hay otros fenómenos tales como los movimientos turbulentos de los fluidos, en los que no parecía existir una clara relación entre causa y efecto, por lo que se pensaba que intervenían en ellos procesos aleatorios. Además, aunque la capacidad para conocer en detalle la estructura compleja de los sistemas se ha incrementado notablemente en los últimos años, la de integrar ese conocimiento se ha visto frenada por la falta de un marco conceptual apropiado; así, aun siendo posible obtener un mapa detallado y completo del sistema nervioso de ciertos organismos simples, no puede deducirse de él su comportamiento.

El descubrimiento del caos supone una revolución para el desarrollo de modelos científicos. Por una parte, establece nuevas limitaciones sobre la capacidad de modelar sistemas y realizar predicciones a largo plazo. Igualmente, la esperanza de que, gracias a un mejor conocimiento de las fuerzas y partículas elementales, la física pueda predecir mejor el comportamiento de la materia, puede resultar infundada; la interacción entre los componentes en una escala reducida puede, como lo señala el principio de Heisenberg (Heisenberg, 1971), llevar a un comportamiento global muy complejo en otra escala mayor que, en general, no puede deducirse del conocimiento de los componentes elementales. Pero, por otra parte, el determinismo que subyace en todo proceso caótico muestra que muchos fenómenos complejos, que parecían gobernados por mecanismos aleatorios y que, hasta ahora, resultaban difíciles de estudiar, son más predecibles de lo que se había pensado. La teoría del caos puede ayudar a establecer ese marco conceptual

que permita la comprensión de fenómenos a escalas macroscópicas. Se ha demostrado en los últimos años que un número creciente de sistemas que exhiben un comportamiento complejo pueden modelizarse mediante un simple atractor caótico. Entre ellos, puede citarse la transición del flujo laminar al turbulento de un fluido, la oscilación de los niveles de las concentraciones en reacciones químicas inestables, los latidos de las células del corazón y muchos osciladores eléctricos y mecánicos.

La teoría del caos constituye, pues, una analogía para comprender cómo el comportamiento complejo, tal como la inteligencia, puede emerger en virtud de simples interacciones no lineales entre unos cuantos componentes sencillos. La inteligencia puede explicarse, entonces, como el resultado de la existencia de “atractores caóticos” en una red de “neuronas”. Estos “atractores” constituyen estados de equilibrio coherente con la interpretación de datos y resultados hacia los que tiende el sistema y de los que difícilmente puede escapar. Killen indica que puede encontrarse evidencia de la existencia de estados de equilibrio difíciles de evitar en muchos aspectos del comportamiento humano; por ejemplo, es muy improbable que una persona aprenda un idioma extranjero sin que tenga algún acento: los fonemas del idioma materno resultan ser estados estables, o “atractores”, a los que los labios, lengua y paladar no pueden escapar. Algunos “atractores” existen desde el nacimiento; por ejemplo, las acciones reflejas. Otros, como los relacionados con el lenguaje, se adquieren durante las etapas tempranas del crecimiento y se estabilizan rápidamente, siendo difíciles de variar. Algunos, tales como los hábitos, se adquieren a lo largo de la vida y son más moldeables.

4. RESULTADOS, CONCLUSIONES Y FUTURAS LÍNEAS DE TRABAJO

4.1. Resultados

Nagel afirmaba que la ciencia no podía abordar el problema de la conciencia, sin un cambio metodológico significativo, un cambio que capacitaría a los científicos para identificar y analizar los elementos de la experiencia subjetiva. Estos elementos serían los componentes básicos de las funciones cerebrales. Por su parte, Spinoza y Sherrington afirmaban que la energía no puede producir el menor cambio en el mundo de las sensaciones y pensamientos; esto es, en la mente. Pues bien, el doctorando va más allá y propone, por una parte, un cambio de modelo para entender la mente. El modelo EMI que toma de Stonier, Wang y Pazos y colegas, que tiene como principal novedad la incorporación de la información al binomio materia y energía tradicional. Y éste es un aporte innovador. Además, importa, asimismo de Pazos y colegas, un enfoque metodológico nuevo, basado en dos constructos: holones e informones, que permiten abordar con naturalidad el problema mente-conciencia. Y éste es otro aporte innovador. Finalmente, y ésta es la aportación original de esta tesis, demuestra, mediante varios casos, cómo, dentro de ese marco y ese enfoque metodológico, la emergencia de propiedades funcionales es algo asequible, incluso dentro de las limitaciones computacionales actuales.

Estos resultados dieron lugar a la producción, hasta el momento, de dos artículos:

- *EMI: ontoepistemologic model to solve the mind-brain problem.*
- *Property Emergencly: a relevant issue to solve the mind-brain problem.*

4.2. Conclusiones

En primer lugar, el problema mente-cerebro es abordable mediante métodos científicos. En segundo lugar, al menos el problema de la unidad de la conciencia es soluble mediante métodos neurocientíficos y computacionales. Finalmente, la emergencia funcional de propiedades, algo necesario para entender la mente, es algo ciertamente asequible mediante métodos computacionales.

4.3. Futuras líneas de investigación

1. Como se vio en el estado de la cuestión, el problema mente-cerebro, desde cualquier punto de vista que se mire, resulta ser un auténtico galimatías, que por no tener ni siquiera tiene una clasificación del mismo. Por eso, es urgente el conseguir una taxonomía o, cuando menos, una clasificación formal de las teorías y, o, planteamientos del mismo. Ésta será la inmediata tarea a la que se va a dedicar el doctorando.
2. Crear un sistema basado en el conocimiento que integre computacionalmente todos los casos de emergencia de funcionalidades obtenidos hasta el momento de una forma individual y algo *ad hoc*.
3. Búsqueda de una(s) partícula(s) que soporte(n) el concepto de la tercera dimensión del modelo EMI: el informón o infón.
4. Demostración de la equivalencia entre Materia e Información, pues es un problema abierto que hasta ahora no ha sido demostrado.

5. BIBLIOGRAFÍA

- [1] Aleo V.: "The Future of Knowledge: Increasing Prosperity through Value Networks". Verna Allee. Butterworth-Heinemann Title. Holland, October 2002.
- [2] Alighieri, D.: "La Divina Comedia". (2 Vol.) S.A. de Promoción y Ediciones. Madrid. 1992.
- [3] Alonso F., Lopez G., Pazos J., Rodriguez-Paton A., Silva A., Soriano F. J.: "Fundamental Elements of a Software Design and Construction Theory: Informons and Holons". Proceedings of the International Symposium of Santa Caterina on Challenges in the Internet and Interdisciplinary Research (SSCCII), Italia. 2004. PP: 21-35.
- [4] Anderson, A. R. (Ed.): "Mind and Machines". Prentice Hall, Inc. Englewood Cliffs. N. J. 1974.
- [5] Anderson, P. W.: "More is Different". Science, N° 177. 1972. PP: 393-396.
- [6] Aristóteles: "Metafísica". Editorial Gredos S.A. Madrid, 1998.
- [7] Atkins, P.: "Galileo's Finger: The Ten Great Ideas of Science". Oxford University Press, 2004
- [8] Baars, B. J. A., Banks, W. P. & Newman, J. B.: "Essential Sources in the Scientific Study of Consciousness". M.I.T. Press. New York, N. Y. 1997.
- [9] Baars, B. J. A.: "Inside the Theater of Consciousness: The Workspace of the Mind". Oxford University Press. New York, N. Y. 1997.
- [10] Bacon, F.: New Atlantis. In, Great Books of the Western World. Chicago, Ill. 1952. PP: 199-214.
- [11] Bateson, G.: "Mind and Nature: A Necessary Unit". Bantam Books. New York, N. Y. 1980.
- [12] Bateson, G.: "Steps to a Ecology of Mind". Ballantine Books. New York. 1972.
- [13] Bateson, G.: "Steps to on Ecology of Mind". Ballantine Books. New York, N. Y. 1972.

- [14] Beardsley, T.: "The Machinery of Thought". Scientific American. August. 1997. PP: 78-83.
- [15] Bell, D.A.: "Information Theory and its Engineering Application". Sir Isaac Pittman & Sons. London. 1968.
- [16] Bergson, H.: "L'Evolution Créatrice (1907)". En, A. Robinet (ed.), "Oeuvres", Presses Universitaires de France (PUF). París, 1970.
- [17] Bernstein, J.: "Physicist John Wheeler: Retarded Learner". Princeton Alumni Weekly. 9, October, 1985.
- [18] Bertalanffy, L.M. von: "An Essay on The Relativity of Categories". Philosophy of Science, 225. 1955. PP:243-263.
- [19] Bertalanffy, L.M. von: "Robots, Hombres y Mentes". Ediciones Guadarrama, S.A. Madrid, España. 1974.
- [20] Bertalanffy, L.M. von: "The Mind-Body Problem: A New View". Psychosomatic Medicine, 24. 1964. PP: 29-45.
- [21] Block, N. J., Flanagan, O. J. and Güzeldere, G.: "The Nature of Consciousness: Philosophical Debates". MIT Press. Cambridge, Mass. 1997.
- [22] Boole, G.: "An Investigations of the Laws of Thought on Which Are Founded The Mathematical Theories of Logic and Probabilities". Dover Publications, Inc. New York, N. Y. 1854.
- [23] Borel, P.: "Discours Nouveau Prouvant la Pluralité des Mondes". Genève. 1657.
- [24] Bridgman, P. W.: "Remarks on Niels Bohr's Talk. The Microscopic and the Observer". In Le Lionnais, F. (Ed.), "La Methode dans les Sciences Modernes", Editions Science et Industrie. Quovadis Daedalus 87. Paris, 1958. PP:117-122.
- [25] Brillouin, L.: "La Science et la Théorie de l'Information". Ed. Masson, Paris 1959.
- [26] Brillouin, L.: "Scientific Uncertainty and Information". Academic Press. New York, N. Y. 1964.
- [27] Buffon: "Discours sur La Nature des Animaux". Dans, Histoire Naturelle. Vol. 4. Imprimerie Royale. Paris, France. 1753.

- [28] Campbell, D. T.: "Evolutionary Epistemology". In, Schilpp, P. A. (Ed.): "The Philosophy de Karl R. Popper". Open Court Publisher, La Salle, IL. 1974. PP: 412-463.
- [29] Capra, F.: "The Web of Life: A New Scientific Understanding of Living Systems". Anchor Books. New York, 1996.
- [30] Carnap, R.: "The Interpretation of Physics". In, Feigl, H. & Brodbeck, M. (Eds.): "Readings in the Philosophy of Science". Appleton-Century-Crofts, Inc. New York, N.Y. 1953.
- [31] Carroll, L.: "Through the Looking-Glass", Great Britain 1872. In "The Complete Illustrated Works of Lewis Carroll", Chancellor Press, London, UK. 1982. PP: 121-231.
- [32] Chalmers, D. J.: "The Conscious Mind". Oxford University Press. New York, N. Y. 1996.
- [33] Chalmers, D. J.: "The Puzzle of Conscious Experience". Scientific American, 273, December, 1995. PP: 80-86.
- [34] Chomsky, N.: "Language and Problems of Knowledge". MIT Press. Cambridge. Mass. 1988.
- [35] Churchland, P. M. and Churchland, P.S.: "Could a Machine Think?". Scientific American. January, 1990. PP: 32-37.
- [36] Churchland, P. M.: "Matter and Consciousness". Bradford Book. The MIT Press. Cambridge, Mass. 1984.
- [37] Churchland, P. M.: "The Engine of Reason, the Sent of the Soul: A Philosophical Journey into the Brain". The MIT Press. Cambridge, Mass. 1995.
- [38] Cohen, J. & Steward, I.: "Chaos Contingency and Convergence". Nonlinear Science Today. 1:2. 1991. PP: 9-13.
- [39] Cole, D. J.: "Thought and Thought Experiments". Philosophical Studies 45 (May). 1984. PP: 431-44.
- [40] Copérnico, N.: "Sobre las Revoluciones". Ediciones Altaya, S.A. Barcelona, España. 1997.

- [41] Craik, K.: "The Nature of Explanation". Cambridge University Press. Cambridge, U.K. 1943.
- [42] Crick F., "La Ciencia e l'Anima. Un'Ipotesi Sulla Coscienza", Rizzoli, Milano, 1994.
- [43] Crick, F. H. C. & Koch, C.: "Are we Aware of Neural Activity in Primary Visual Cortex?". *Nature*, 375, 1995. PP: 121-123.
- [44] Crick, F. H. C.: "Neural Edelmanism. Trends in Neurosciences". Vol. 12. Nº 7, 1989. PP: 240-248.
- [45] Crick, F. H. C.: "The Astonishing Hypothesis". Charles Scribner's Sons. New York, N.Y. 1994.
- [46] Crick, F. H. C.: "Toward a Neurobiological Theory of Consciousness". *Seminary in The Neurosciences*. Vol. 2, 1990. PP: 263-275.
- [47] Crick, F. H. C.: "What Mad Pursuit". Basic Books. New York, N.Y. 1988.
- [48] Cusa, N.: "Un ignorante discurre acerca de la mente". Círculo de Estudios Cusanos de Buenos Aires, Biblos, 2005.
- [49] Darwin, C.: "Of the Origen of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life". London, U.K. 1859. En español "El Origen de las Especies". EDAF, Ediciones – Distribuciones, S.A. Madrid, España. 1970.
- [50] Deecke, L.; Grözing, B.; Kornhuber H.H.: "Voluntary Finger Movement in Man: Cerebral Potentials and theory". *Biol. Cybernet.* 23, 1976. PP: 99-119.
- [51] Delbrück, M.: "Mente y Materia. Un Ensayo de Epistemología Evolutiva". Alianza Editorial, S.A. Madrid, España. 1989.
- [52] Dennett, D. C.: "Bombas de Intuición". En, Brockman, J.: "La Tercera Cultura". Tusquets Editores, S.A., Barcelona, España. 1996. PP: 170-185.
- [53] Dennett, D.C.: "Consciousness Explained". Little Brown. Boston, Mass. 1991.
- [54] Dennett, D.C.: "Review of Bright Air, Brilliant Fire". *New Scientist*. 13, June, 1992. PP: 48.

- [55] Descartes, R.: “Meditationes De Prima Philosophia”. In, “Quibus Dei Existentia & Animae Humanae Á Corpore Distinctio, Demonstrantur”. Apud Danielem Elsevirium. Amstelodomi. 1642.
- [56] Descartes, R.: “Oeuvres et Lettres”. Bibliothèque de la Pléiade. Paris, France. 1947.
- [57] Dubois, R.: “Un Dios Interior”. Salvat Editores, S.A. Barcelona, España. 1986.
- [58] Dyson, F.: “The Scientist as Rebel”. The New York Review of Books. May 25. New York 1995. PP: 31-33.
- [59] Eccles, J. & Popper, K. R.: “The Self and its Brain”. Springer-Verlag. Berlin, Germany. 1977.
- [60] Eccles, J. and Beck, F.: “Quantum Aspects of Brain Activity and the Role of Consciousness”. Proceedings of the National Academy of Science, Vol. 89, December, 1992. PP: 11.357-11.361.
- [61] Eccles, J.: “A Unitary Hypothesis of Mind-Brain Interaction in the Cerebral Cortex”. Proceedings of the Royal Society of London. Series B. Biological Sciences, 240, 1990. PP: 433-451.
- [62] Eccles, J.: “How the Self Controls Its Brain”. Springer-Verlag. Berlin, Germany. 1994.
- [63] Edelman, G. M. and Tononi, G.: “El Universo de la Conciencia”. Editorial Crítica, S. L. Barcelona, España. 2002.
- [64] Edelman, G. M. Neural Darwinism: “The Theory of Neuronal Group Selection”. Basic Books. New York, 1987.
- [65] Edelman, G. M.: “Bright Air, Brilliant Fire”. Basic Books. New York, N.Y. 1992.
- [66] Edelman, G. M.: “The Remembered Present”. Basic Books. New York, N.Y. 1989.
- [67] Edelman, G. M.: “Topobiology”. Basic Books. New York, N.Y. 1988.
- [68] Egger, M. V.: “La Parole Intérieure”. Paris, France. 1881.
- [69] Einstein, A.: “Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig?”. Annalen der Physik, 18. 1905. PP: 639-641.

- [70] Einstein, A.: “Über das Relativitätsprinzip und die aus demselben gezogene” Folgerungen. Jahrbuch der Radioaktivität und Elektromik 4. 1907. PP: 411-462.
- [71] Empédocles de Agrigento, Anaxágoras de Clazomenas: “Los Filósofos Presocráticos”. Vol. III. Editorial Planeta De Agostini. Madrid, España. 1996.
- [72] Enriques, F.: “Scienza e Razionalismo”. Zanichelli. Bolonia, Italia. 1992.
- [73] Flanagan, O.: “Consciousness Reconsidered”. The MIT Press. Cambridge, Mass. 1992.
- [74] Flanagan, O.: “The Science of the Mind”. MIT Press. Cambridge, Mass. 1991.
- [75] Fodor, J.: “The Modularity of Mind”. MIT Press. Cambridge, Mass. 1983.
- [76] Fodor, J.: “The Trouble with Psychological Darwinism”. London Review of Books. 22 January, 1998. PP: 11-13.
- [77] Freeman, A.: “¿Qué Son los Pensamientos?”. En, H. Swain (Ed.): “Las Grandes Preguntas de la Ciencia”. Crítica, S.L. Barcelona, España. 2002.
- [78] Freeman, W.: “Societies of Brains: A Study in the Neurosciences of Love and Hate”. Lawrence Erlbaum Associates, Hillsdale, 1995.
- [79] Freud, S.: “Conferencias de Introducción al Psicoanálisis”. En, “Obras completas”. Biblioteca Nueva. Madrid, España. 1948-1968.
- [80] Freud, S.: “Introduction à la Psychoanalyse”. Payot. Paris, France. 1975.
- [81] Freud, S.: “Proyecto de una Psicología para Neurólogos”. En, “Obras Completas”. Editorial Nueva. Madrid, España. 1922.
- [82] Gamow, G.: “Biografía de la Física” Alianza Editorial. Madrid, 2001.
- [83] Gardner, M.: “Computers Near the Threshold?”. Journal of Consciousness Studies 3. Nº 1. 1996. PP: 89-94.
- [84] Gaston, duc de Lévis; “Maximes et réflexions sur différents sujets de morale et de politique. éd. Renouard, 1812, vol. 1, p. 224-
- [85] Gell-Mann, M.: “Plectica”. En J. Brockman (Ed.), “La tercera cultura. Más allá de la Revolución Científica”. Tusquets Editores S.A., Barcelona, 1996. PP: 298-313.
- [86] Gleicks, J.: “La Creación de una Ciencia”. Seix Barral. Barcelona, España. 1987.

- [87] Gödel, K.: "Kurt Gödel, Collected Works". Vol. II. S Feferman et alii (Eds.). Oxford University Press. New York, N.Y. 1990.
- [88] Gödel, K.: "Über Formal Unentscheidbare Sätze der Principia Mathematica". Vol. LIX, und Verwandter Systeme. In Monatshefte für Mathematik und Physik, 38. 1931. PP: 173-198.
- [89] Goethe, J. W.: "Fausto. En, Obras completas". Aguilar, S.A. Madrid, España. 1963.
- [90] Greenfield, S.: "The Private Life of the Brain: Emotion, Consciousness and the Secret of the Self". John Wiley & Sons. New York, N.Y. 2001.
- [91] Hameroff S. R., Kaszniak, A. W. & Scott A. C.(Eds.): "Toward a Science of Consciousness: The First Tucson Discussions and Debates". MIT. Press. Cambridge, Mass. 1996.
- [92] Hanson, N. R.: "Patterns of Discovery". Cambridge University Press. London, U.K. 1958.
- [93] Hartley, R.V.L.: "Transmission of Information". Bell Systems Tech Journal. 7. 1928. PP: 535.
- [94] Hartmann, N.: "Der Aufbau der Realen Welt". De Gruyter. Berlin, Deutschland. 1964.
- [95] Hebb, D.: "Mente e Pensiero". Il Mulino. Bologna, Italia. 1980.
- [96] Heisenberg, W.: "Physics and Beyond: Encounters and Conversations". Harper and Row. New York. 1971.
- [97] Heisenberg W.: "Physics and Beyond". Harper New York, 1971. En español "Diálogos sobre la física atómica". In, Heisenberg, Werner, Bohr and Schrödinger: "Física Cuántica", Círculo de Lectores. Barcelona, España. 1996. PP: 343-368.
- [98] Hilbert, D.: "Mathematische Problem". Nachrichten von der Königl. Gesellschaft der Wiss. Zu Göttingen. 1900. PP: 253-297.
- [99] Hobbes, T.: "Leviatán o La Materia, Forma y Poder de un Estado Eclesiástico y Civil". Alianza Editorial. Madrid 2001.

- [100] Hofstadter, D.: “Gödel, Escher, Bach”. Vintage Books. New York, N. Y. 1980.
- [101] Homero: “Iliada”. Editorial Gredos, S.A. Madrid, 1996.
- [102] Horgan, J.: “Can Science Explain Consciousness?”. Scientific American. July, 1994. PP: 88-94.
- [103] Horgan, J.: “El Fin de la Ciencia”. Ediciones Paidós Ibérica, S. A. Barcelona, España 1998.
- [104] Horgan, J.: “La Mente por Descubrir”. Ediciones Paidós Ibérica, S.A. Barcelona, España. 2001.
- [105] Hume, D.: “Investigación sobre el Entendimiento Humano”. Losada. Buenos Aires, Argentina. 1939.
- [106] Humphrey, N. A.: “History of the Mind”. Harper Perennial. New York, N.Y. 1993.
- [107] Huxley, J.: “Evolution: A Modern Syntesis”. George Allen & Unwin. London, U.K. 1942.
- [108] Huxley, J.: “New Bottles for New Wine: Essays”. New York: Harper. 1957.
- [109] Huxley, T.H.: “Methods and Results: Essays”. D. Appleton. New York, N.Y. 1901.
- [110] Huygens, C.: “Oeuvres Completes”. 22 Vol. Martimus Nijhoff. La Haya. 1850-1888.
- [111] James W.: “Ensayos sobre el empirismo radical”. Ensayo II. Nueva York, 1912.
- [112] James, W.: “The Principles of Psychology”. Henry Holt. New York, N. Y. 1890.
- [113] Johnson, Steven. Sistemas Emergentes: O qué tienen en común hormigas, neuronas, ciudades y software. Ediciones Turner/Fondo de Cultura Económica. Madrid, 2004.
- [114] Jung, C.: “Der Geist der Psychologie” (Consideraciones Teóricas Sobre la Naturaleza de lo Psíquico). Eranos Jahrbuch, 1946.
- [115] Kandel, E. R.: “I Mistero della Mente”. B. Levy é E. Servan-Schreiber (Eds.). Le Science, CD-Room. 1998.

- [116] Kandel, E. R.: "In Search of Memory: The Emergence of a New Science of Mind". W. W. Norton & Co. New York, 2006.
- [117] Kauffman, S.A. : "Investigaciones". Tusquets Editores, S.A. Barcelona, España. 2003.
- [118] Keller, H.: " Die Geschichte Meines Lebens". Lutz. Stuttgart, Deutschland. 1905.
- [119] Kierkegaard, S.: "Diario". Editorial Morcelliana, 12 vols. Brescia, 1980.
- [120] Killen, M. "Culture, self, and development: Are cultural templates useful or stereotyp cal?". *Developmental Review*, 17. 1997. 239-249.
- [121] Koch, C.: Observación hecha en "Toward Scientific Basic for Consciousness". Tucson, Arizona, April 1994. PP: 12-17.
- [122] Koch, C. "Hard-Headed Dualism". *Nature*, Volume 381, Issue 6578. 1996. PP: 123-124.
- [123] Koch, C.: "La Consciencia. Una aproximación neurobiológica". Editorial Ariel. Barcelona. 2005
- [124] Kottler, M. J.: "Charles Darwin and Alfred Russel Wallace: Two Decades of Debate over Natural Selection". In David Kohn (eds.): "The Darwinian Heritage". Princeton, 1985. PP: 367-432.
- [125] Kulpe, O. and Titchener, E. B.: "Outlines of Psychology, Based upon the Results of Experimental Investigation". MacMillan. New York, N.Y. 1909.
- [126] Kurzweil, R.: "The Age of the Spiritual Machines: When Computers Exceed Human Intelligence". Penguin Books. 1999.
- [127] Laercio, D.: "Vidas, Opiniones y Sentencias de los Filósofos más Ilustres" (2 Vol.). Teorema. Barcelona. 1985.
- [128] LeDoux, J.: "El Cerebro Emocional". Editorial Planeta, S.A. Barcelona, España. 1999.
- [129] Leibniz, W. G.: "Monadología". Pentalfa. Oviedo, España. 1981.
- [130] Leibniz, W.G.: "Nuevo Sistema de la Naturaleza". Aguilar. Madrid, España. 1926.
- [131] Leucipo, Demócrito: "Los Filósofos Presocráticos". Vol IV.: Editorial Planeta De Agostini. Madrid, España. 1996.

- [132] Levi Montalcini, R.: "La Galaxia Mente". Editorial Crítica, S.L. Barcelona, España. 2000.
- [133] Levine, J.: "Materialism and Qualia: The Explanatory Gap". *Pacific Philosophical Quaterly*. 64. 1983. PP: 354-361.
- [134] Libet B., Wright EW Jr., Feinstein B. and Pearl D.K.: "Subjective Referral of the Timing for a Conscious Sensory Experience: A Functional Role for the Somatosensory Specific Projection System in Man". *Brain*. 102(1). Mar, 1979. PP:193-224.
- [135] Libet, B.: "Cerebral Processes that Distinguish conscious Experience from Unconscious Mental Functions". In Eccles, J.C. and Creutzfeldt, O.D. (Eds.) "The Principles of Design and Operation of the Brain". Springer Verlag, Berlin, 1990. PP: 185-205.
- [136] Libet, B.: "The Neural Time-Factor in Perception, Volition and Free Will". *Revue de Metaphysique et de Morale*. 2, 1992. PP: 255-272.
- [137] Llull, R.: "Ars generalis ultima or Ars magna" (The Ultimate General Art). 1305.
- [138] Locke, J.: "Ensayo sobre el Entendimiento Humano". (2 Vol.) Editora Nacional. Madrid. España. 1980.
- [139] Lorenz, E.: "The Essence of Chaos". University of the Washington Press. 1993.
- [140] Lorenz, K.: "La Otra Cara del Espejo". Plaza & Janes, S.A. Editores. Esplugas de Llobregart, España. 1974.
- [141] Lorenz, K.; "On Aggression". University Paperback. London, U.K. 1967.
- [142] Lucas, J. R.: "Minds, Machines and Gödel". *Philosophy*, Vol. XXXVI. 1961. PP: 112-127.
- [143] Lyons, W.: "Matters of the Mind". Edinburgh University Press. Edinburgh, Scotland, U.K. 2001.
- [144] MacKay, A. L.: "Diccionario de citas científicas: La cosecha de una mirada serena". Ediciones de la Torre. 1992. PP: 201. Maxwell's Collected Works II. 1869. PP: 760.
- [145] MacKay, D. M.: "Cerebral Organization and the Conscious Control of Action". In J. C. Eccles (Ed.): "Brain and Conscious Experience". New York, 1966.

- [146] MacKay, D. M.: "The Science of Communication – A Bridge between Disciplines". Conference in the College of The University North Staffordshire. Keele, 1961.
- [147] Margulis, L.: "Gaia es una Pícaro Tenaz". En, Brockman, J.: "La Tercera Cultura. Más allá de la Revolución Científica". Simon & Schuster. 1995.
- [148] Margulis, L.: "Symbiosis in Cell Evolution". W. H. Freeman. San Francisco, CA. 1993.
- [149] Margulis, L.: "The Origin of Eukaryotic Cells". Yale University Press. 1970.
- [150] Margulis, L.: "The Origin of Mitosing [Eukariotic] Cells". The Journal of Theoretical Biology. 1966.
- [151] Maxwell, J.C.: "Theory of Heat". Longmans, Green and Co. London. 1871.
- [152] Maxwell, J.C.: Revisión de Stewart, B. y Tait, P. G.: "Paradoxical Philosophy". Nature, nº 19. Also Maxwell's Collected Works II, 1869. PP: 760.
- [153] May, R.: "El Mejor Tiempo Posible Para Vivir. El Mapa Logístico". En, Farmelo, G. (Ed.): "Fórmulas Elegantes: Grandes Ecuaciones de la Ciencia Moderna". Colección Metatemas Tusquets Editores, S.A. Barcelona, España. 2004.
- [154] Mayr E.: "How Biology Differs from The Physical Sciences". In D. J. Depew and B H Weber, (eds.): "Evolution at a Crossroads: The New Biology and the New Philosophy of Science". The MIT Press, Cambridge MA, 1985. PP: 43-63.
- [155] Mayr E.: "The Growth of Biological Thought". Harvard University Press. Cambridge, 1982.
- [156] Mayr E.: "The Limits of Reductionism". Nature, nº 331, 1987. PP:475.
- [157] McGinn, C.: "Can We Solve the Mind-Body Problem". Mind, 98. 1989. PP: 349-366.
- [158] McGinn, C.: "The Problem of Consciousness". Blackwell Publishers. Cambridge, Mass. 1991.
- [159] Minsky, M.: "The Society of Mind". Simon and Schuster. Nueva York, 1987.
- [160] Moore, W.: "Aniversario fatal". Mondadori, Milán, 1955.

- [161] Müller, F., Kunesh, E., Binkofski, F. and Freund, H. J.: “Residual Sensorimotor Functions in a Patient After Right-Sided Hemispherectomy”. *Neuropsychologia*, 29. 1991. PP: 124-145.
- [162] Nagel, E. & Newman, J. R.: “Gödel’s Proof”. New York University Press. New York, N.Y. 1958.
- [163] Nagel, T.: “What Is It Like to Be a Bat”. In, T. Nagel: “Mortal Questions”. Cambridge University Press. New York, N.Y. 1979.
- [164] Newton, I.: “Philosophiæ Naturalis Principia Mathematica“, en “Principios matemáticos de la filosofía natural“. Edición de Eloy Rada. Alianza Editorial, Madrid, 1987.
- [165] Nicolai Copernici Torienensis: “De Revolutionibus Orbium Coelestium”. Libri VI. Johann Petreius. Notimbergae. 1543.
- [166] Nicoli, G. and Prigogine, I.: “Exploring Complexity: An Introduction”. W. H. Freeman. San Francisco, Ca. 1989.
- [167] Norretranders, T.: “The User Ilusion: Cutting Consciousness Down To Size”. Viking. New York, N.Y. 1998.
- [168] Pauli, W.: “Über den Einflub der Geschwindigkeitsabhängigkeit der Epeketronenmase auf den Zeemaneffekt”. *Zeüitschrifl für Physik*, 31. 1925a. PP: 373-385.
- [169] Pauli, W.: “Über den Zusammenhang des Abschlusses der Odectronengruppen im Aform mit der Komplexstruktur der Spektren”. *Zeitschriif für Phys. K*, 31. 1925b. PP: 765-783.
- [170] Pazos, J.: “La Nostrificación y Einstein”. To Be Published. 2012.
- [171] Peano, G.: “Arithmetices Principia, Nova Methodo Exposita”. Bocca Turin, Italy. 1889.
- [172] Penfield, W.: “The Mystery of the Mind: A Critical Study of Consciousness and the Human Brain”. Princeton University Press. Princeton, N.J. 1975.
- [173] Penrose, R.: “The Emperor’s New Mind”. Oxford University Press. New York, 1989.
- [174] Penrose, R.: “Shadows of the Mind”. Oxford University Press. New York, 1994.

- [175] Pinker, S.: "How the Mind Works". Norton. New York, N. Y. 1997.
- [176] Planck, M.: "Problemas Aparentes de la Ciencia". En, M. Planck "Autobiografía científica y últimos escritos". Nivola Libros y Ediciones S.L. Madrid, 2000. PP: 59-60.
- [177] Platón: "Fedón". En, "Platón, Menón. Crátilo. Fedón". Editorial Planeta de Agostini, S.A. Madrid, España. 1997
- [178] Plutarco: "Las Vidas Paralelas". (5 Vol.) Imprenta Nacional. Madrid, España. 1821-1822-1830.
- [179] Pribram, K.: "Brain and Perception". Lawrence Erlbaum Associates. Hillsdale, New York, N. Y. 1991.
- [180] Prigogine, I. & Stengers, I.: "Order Out of Chaos: Man's New Dialogue with Nature". Flamingo Fontana Paperbacks. London. 1985.
- [181] Protágoras: "Discursos Demoledores". En, Sofistas: Testimonios y Fragmentos. Editorial Planeta De Agostini, S.A. Madrid, España. 1997.
- [182] Putnam, H.: "Mind and Machines". In, Hook, S. (Ed.): "Dimensions of Mind: A Symposium". New York University Press. New York, N.Y. 1960.
- [183] Rheingold, H. "La Realidad Virtual". Baskerville. Bolonia, Italia. 1993.
- [184] Rifá, J. and Huguet, U.: "Comunicación Digital. Teoría Matemática de la Información. Codificación Algebraica. Criptología". Masson, S.A. Barcelona, España. 1991.
- [185] Rodríguez Delgado, J. M.: "Brain and Mind". Ciba Foundation Series, Nº 69. October, 1979.
- [186] Rodríguez Delgado, J.M.: "Physical Control of the Mind". Toward a Psychocivilized Society. Harper & Row. New York, N. Y. 1969.
- [187] Rodríguez Delgado, J.M.: "Triunism: A Transmaterial Brain-Mind Theory". Ciba Foundation Series, X. 1979. PP: 385.
- [188] Rose, S. (Ed.): "From Brains to Consciousness?. Essays on the New Sciences of the Mind". Penguin Books. Harmondsworth. 1999.

- [189] Rosenblum, B., Kuttner, F.: "The Only Objective Evidence for Consciousness". *Journal of Mind and Behavior* 27 (1). 2006. PP:43-56.
- [190] Russell, B. A. W.: "An Outline of Philosophy". George Allen & Unwin. Ltd. London, U.K. 1927.
- [191] Russell, B. A. W.: "Logical Atomism". In, J. H. Muirhead (Ed.): "Contemporary British Philosophy". Vol. 1. 1924.
- [192] Russell, B. A. W.: "The Analysis of Mind". George Allen & Unwin. Ltd. London, U.K. 1921.
- [193] Russell, B.: "Religion and Science". George Allen & Unwin, Ltd. London, U.K. 1935.
- [194] Ryle, G.: "The Concept of Mind". Hutchinson. London, U.K. 1949.
- [195] Sacks, O.: "L'Uomo che Scambio sua Moglie per un Capello". Adelphi. Milano, Italia 1986.
- [196] Scarrott, G.: "The Need for a "Science" of Information". *Journal Information Technology*, 1(2). 1986. PP: 33-38.
- [197] Schopenhauer, A.: "De la Cuádruple Raíz del Principio de Razón Suficiente". Editorial Gredos, S.A. Madrid, España. 1981.
- [198] Schopenhauer, A.: "The World as Will and Representation". Dover. New York, N.Y. 1966.
- [199] Schrödinger, E.: "Mind and Matter". Cambridge University Press. Cambridge, Mass. 1958.
- [200] Scriven, M.: "The Compleat Robot: A Prolegomena to Androidology". In, Hook, S. (Ed.): "Dimensions of Mind: A Symposium". New York University Press. New York, N.Y. 1960. PP: 118-142.
- [201] Searle, J. R.: "How to Study Consciousness Scientifically". In, Cornwell, J. (Ed.): "Consciousness and Human Identify". Oxford University Press. Oxford, U.K. 1998. PP: 21-37.
- [202] Searle, J. R.: "Minds, Brains and Science". Harvard University Press. Cambridge, Mass. Searle, J. R.: *La Mente è un Programma? Le Scienze*. Marzo, 1990. PP: 259.

- [203] Searle, J. R.: "The Rediscovery of the Mind". MIT Press. Cambridge, Mass. 1992.
- [204] Searle, J.: "El misterio de la Conciencia". Ediciones Paidós Ibérica. 2000.
- [205] Shannon, C. E. & Weaver, W.: "The Mathematical Theory of Communication". University of Illinois Press. Urbana, Ill. 1963.
- [206] Shannon, C.E.: "A Mathematical Theory of Communication". The Bell System Technical Journal. 623-656. July, October, 1948. PP:379-423.
- [207] Shear, J.: "Explaining Consciousness: The Hard Problem". MIT Press. Cambridge, Mass. 1997.
- [208] Sherrington, C. S.: "Man and His Nature". Cambridge University Press. Cambridge, Mass. 1951.
- [209] Sherrington, C. S.: "The Integrative Action of the Nervous System". Constable. London, U.K. 1906.
- [210] Simon, H.: "I Misteri della Mente". Levi, B. e Servan-Schreiber J.J. (Eds.). Le Scienze, CD-rom. September. 1998.
- [211] Skinner, B.F.: "The Behavior of Organisms". Appleton-Century. New York, 1938.
- [212] Sloman, A.: "The Emperor's Real Mind: Review of Roger Penrose The Emperor's New Mind". Artificial Intelligence, 56. 1992. PP: 355-396.
- [213] Sperry, R. W.: "Problem Outstanding in the Evolution of Brain Function". James Arthur Lecture on the Evolution of The Human Brain. The American Museum of Natural History. New York, N.Y. 1964.
- [214] Spinoza: "Ethica Ordine Geometrico Demonstrata". Dover Publications, 1951.
- [215] Stent, G. S.: "Introducción y Panorámica General". In, Delbrück, M.: "Mente y Materia. Un Ensayo de Epistemología Evolutiva". Alianza Editorial, S.A. Madrid, España. 1989.
- [216] Stent, G.: "The Coming Golden Age". Natural History Pres. Garden City, N.Y. 1969.
- [217] Steward, I.: "Does God Play Dice? The New Mathematics of Chaos". Penguin Books Harmondsworth. 1997.
- [218] Stonier, T.: "Information and the Internal Structure of the Universe". Springer-

- Verlag. London. 1990.
- [219] Sutherland, K.: “¿Qué es la Conciencia?”. En H. Swann (Ed.): “Las Grandes Preguntas de la Ciencia”. Crítica, S. L. Barcelona, España. 2003.
- [220] Tales, Anaximandro, Anaxímenes, Pitágoras, Alcmeón y Jenófanes: “Los Filósofos Presocráticos”. Editorial Planeta De Agostiini, S.A. Madrid, España. 1995.
- [221] Thorpe, W. H.: “Ciencia Hombre y Moral”. Editorial Labor, S.A. Barcelona, España. 1967.
- [222] Thorpe, W. H.: “L’Etologia e La Coscienza”. In V. Somenzi (Ed.): “La Física della Mente”. Bollati Boringhieri. Torino, Italia 1969.
- [223] Titchener, E. B.: “An Outline of Psychology. MacMillan”. New York, N.Y. 1901.
- [224] Tononi, G., Sporns, O. and Edelman, G. M.: “A Measure for Brain Complexity: Relating Functional Segregation and Integration in the Nervous System”. Proceedings of the National Academy of Sciences of the United States of America, 91. 1994. PP: 5033-5037.
- [225] Toyabe, S; Sagawa, T., Ueda, M.; Muneyuki, E. and Sano, M.: “Information Heat Engine: Converting Information to Energy by Feedback Control Internet”. Sept., 2010.
- [226] Turing, A. M.: “Computer Machinery and Intelligence”. Mind, Vol. LIX, N° 236. 1950. PP: 433-460.
- [227] Varela, F.: “El Yo Emergente”. En, Brockman, J.: La Tercera Cultura. Tusquets Editores, S.A. Barcelona, España. 1996. PP: 196-208.
- [228] Vinci, L. da: “Il Codice Atlantico”. 1483-1518.
- [229] Von Neumann, J.: “The Work of the Mind” 1947. Collected Works (6 Vol.) Vol. 1. Macmillan, New York, 1961-1963. PP: 1-9.
- [230] Wallace, A. R.: “On the Tendency of Varieties to Depart Indefinitely from the Original Type”. Journal of the Linnean Society, Vol. 3. 1-July, 1858.
- [231] Wang, Y.: “Software Engineering Foundations: a Software Science Perspective”. Taylor & Francis Group, LLC. Boca Ratón, Fl. 2008.

- [232] Warner, R. and Szubka, T.: "The Mind-Body Problem: A Guide to the Current Debate". Blackwell. Cambridge, Mass. 1994.
- [233] Watson, J. B.: "Behaviorism". University of Chicago Press. Chicago, Ill. 1924.
- [234] Watson, J. B.: "Psychology as the Behaviourist Views it". *Psychological Review*, XX. 1993. PP: 158-177.
- [235] Weinberg, S.: "La Propia Naturaleza". En "Plantar cara. La ciencia y sus adversarios culturales". Ediciones Paidós Ibérica S.A. Barcelona, 2003. PP: 67-78.
- [236] Weinberg, S.: "Unified Theories of Elementary-Particle Interactions". *Scientific American* 231. 1974. PP: 50-55.
- [237] Wheeler, J. A.: "At Home in the Universe". American Institute of Physics Press. Woodbury, N. Y. 1994.
- [238] Wheeler, J. A.: "Information, Physica, Quantum: The Search for Links". In, W. H. Zurek (Ed.): "Complexity Entropy and the Physics of Information". Addison Wesley. Reading, Mass. 1990.
- [239] Whyte, R. W.: "The Unconscious Before Freud". Basic Books, Inc. New York, N. Y. 1960.
- [240] Wigner, E.P.: "Philosophical Reflections and Synthesis". Springer Verlag. Berlin, Germany. 1977.
- [241] Wittgenstein, L.: "Investigaciones Filosóficas". Crítica, Grijalbo. Barcelona, España. 1988.
- [242] Wittgenstein, L.: "Tractatus Logico-Philosophicus". Revista Occidente, S.A. Madrid, España. 1957.
- [243] Wundt, W.: "Vorlesungen Ubre die Mensche und Tíresele". Verlag von Leopald Voss. Leipzig. Deutschland. 1992.
- [244] Zilard, L. Z.: "Über die Entropieverminderung in Einem Thermodynamischen System bei Eingriffen Intelligenter Wessen". *Zeitschrif für Physik*, 53. 1929. PP: 840-856.
- [245] Zohar, D.: "The Quantum Self. William Morrow". New York, N. Y. 1990.