

LAS HEURÍSTICAS SIMPLES SE ENCUENTRAN CON LA MODULARIDAD MASIVA¹

PETER CARRUTHERS

Dep. Filosofía, Univ. de Maryland

Resumen

Este artículo investiga la coherencia entre la propuesta de una organización modular masiva de la mente (propugnada por algunos defensores del programa de investigación de la psicología evolucionista) y el enfoque de las heurísticas simples. Se discute una serie de potenciales conflictos entre los dos programas, pero finalmente son desestimados. De todos modos, el programa de las heurísticas simples sí termina socavando uno de los muchos argumentos propuestos para apoyar la modularidad masiva, al menos en el modo en que esta última es comprendida por los filósofos. Así que un resultado de la argumentación será forzarnos a reexaminar la mejor manera de caracterizar la noción de *modularidad* en la ciencia cognitiva, para que la tesis de la modularidad masiva no deba ser abandonada completamente. Lo que está en juego en esta discusión es si existe una noción de “módulo” que esté bien motivada, de modo que tengamos una buena razón para pensar que la mente humana deba ser masivamente modular en su organización. Argumentaré (sobre el final) que tal noción existe.

PALABRAS CLAVE: Heurísticas; Modularidad masiva; Arquitectura cognitiva; Psicología evolucionista

Abstract

This chapter investigates the extent to which claims of massive modular organization of the mind (espoused by some members of the evolutionary psychology research program) are consistent with the main elements of the simple heuristics research program. A number of potential sources of conflict between the two programs are investigated and defused. However, the simple heuristics program turns out to undermine one of the main arguments offered in support of massive modularity, at least as the latter is generally understood by philosophers. So one result of the argument will be to force us to re-examine the way in which the notion of *modularity* in cognitive science should best be characterized, if the thesis of massive modularity isn't to be abandoned altogether. What is at stake in this discussion, is whether there is a well-motivated notion of “module” such that we have good reason to think that the human mind must be massively modular in its organization. I shall be arguing (in the end) that there is.

KEY WORDS: Heuristics; Massive modularity; Cognitive architecture; Evolutionary Psychology

¹ “Simple heuristics meet massive modularity”, incluido originalmente en P. Carruthers, S. Laurence y S. Stich (eds.) *The innate mind 3: Foundations and the Future*, Oxford University Press, 2007. Traducción de Tomás Balmaceda. Agradecemos a Oxford University Press la autorización para esta traducción.

Introducción

En las últimas décadas, se han desarrollado dos programas de investigación diferentes en ciencia cognitiva, cada uno de los cuales es inequívocamente evolucionista o adaptacionista en su orientación, y nativista con respecto a la arquitectura y gran parte del contenido de la mente humana. Uno es el programa de la psicología evolucionista, con su defensa de la modularidad masiva de lo mental (Gallistel, 1990, 2000; Tooby y Cosmides, 1992; Sperber, 1996; Pinker, 1997). El otro es el movimiento de las heurísticas simples, con su idea de una “caja de herramientas adaptativas”, de procedimientos cognitivos simples (Gigerenzer *et al.*, 1999). Adicionalmente, cada uno de estos programas se ve obligado a explicar la variabilidad cultural en términos de la aplicación concreta de tales módulos / heurísticas en las diversas condiciones locales.

Mi pregunta es ésta: ¿se trata de programas de investigación *en pugna* o se complementan el uno con el otro? Los defensores de estos programas no mencionan con frecuencia la existencia del otro. Aun así, ambos están trabajando para construir explicaciones que sean plausibles, no sólo en términos evolutivos, sino también en relación con los datos de la psicología comparada. Y parece asimismo que ambos están trabajando para explicar (o tratar de explicar) cómo la cognición puede ser realizada en procesos que sean tratables computacionalmente. Sin embargo, existen algunas razones para pensar que estos programas ofrecen explicaciones de la cognición humana que son incoherentes entre sí, que se cuestionan, tal como veremos.

Voy a comenzar con una breve elaboración y elucidación de las tesis gemelas de que la cognición es masivamente modular y que está construida a partir de una caja de herramientas adaptativas de heurísticas simples. Me dirigiré/volveré a continuación hacia la pregunta sobre su coherencia, sosteniendo (hacia el final) que los dos programas de investigación deben ser considerados como compañeros naturales y colaboradores mutuos, en vez de como competidores. Pero para que éste sea el caso, la tesis de la modularidad masiva de la mente debe pasar por una (bien motivada) transformación.

2. Modularidad masiva

Los modularistas sostienen que cuando se considera la mente desde una perspectiva evolutiva, se obtienen predicciones específicas sobre su arquitectura, siendo la más importante la modularidad masiva de su organización. Esta conclusión puede alcanzarse a través de distintas vías

de razonamiento (pero que se apoyan mutuamente). Esbozaré dos de ellas aquí (para una discusión más extensa, vd. Carruthers, 2005.)

2.1 El argumento de la biología

El primer argumento se remonta a Simon (1962) y concierne a los sistemas funcionales complejos en general, y a la biología en particular. De acuerdo con esta línea de pensamiento, deberíamos esperar que tales sistemas se construyan a partir de subsistemas disociables, de tal modo que el ensamblaje completo pueda armarse gradualmente, agregando un subsistema sobre otro subsistema de manera que la funcionalidad del todo pueda ser amortiguada en la medida de lo posible, en caso de daños a las partes.

Simon (1962) usa su famosa analogía de los dos relojeros para ilustrar el punto. Un relojero arma un reloj por vez, agregando microcomponente sobre microcomponente, uno cada vez. Esto hace que se vuelva fácil y se olvide del orden apropiado de las partes y, en el caso de una interrupción, quizás deba recomenzar desde cero. El segundo relojero primero construye un conjunto de subcomponentes a partir de partes de microcomponentes, y luego los combina en subcomponentes más grandes, hasta que eventualmente el reloj queda completo. Esto facilita la organización y secuencia de todo el proceso y lo vuelve menos vulnerable a las interrupciones.

De manera consistente con esta visión, existe una gran cantidad de evidencia desde distintos campos de la biología que muestra cómo los sistemas funcionales complejos están contruidos a partir del ensamblaje de subcomponentes. Cada uno de estos subcomponentes está contruido con otros de nivel inferior, cumpliendo un rol distintivo en el funcionamiento del todo, de tal manera que si uno de ellos resulta dañado o se pierde, la funcionalidad de las partes remanentes resulta al menos parcialmente intacta. Esto es cierto para las operaciones de las células, el ensamblaje entre células, de órganos completos y de unidades multiorgánicas como una colonia de abejas (Seeley, 1995). Y, por extensión, deberíamos esperar que sea cierto también en el caso de la cognición.

La predicción en esta línea de razonamiento, entonces, es que la cognición estará estructurada partiendo de sistemas disociables, cada uno de los cuales tiene una función distintiva asignada, o un conjunto de funciones (deberíamos esperar que muchos sistemas cognitivos tengan un conjunto de funciones, más que una sola función, ya que la multifuncionalidad abunda en el mundo biológico. Una vez que un componente ha sido seleccionado, puede ser cooptado y moldeado para realizar otras funciones). Esto nos da una noción de “módulo” cognitivo que se acerca mucho

al sentido cotidiano en que uno habla de un minicomponente musical como “modular”, dado que se puede comprar un reproductor de cassettes, y puede funcionar de manera independiente al reproductor de discos compactos, etcétera. Dicho rápidamente, un módulo es simplemente un *componente* disociable.

De manera consistente con la predicción hecha más arriba, existe una gran cantidad de evidencia de tipo neuropsicológico de que algo así como una modularidad masiva (en el sentido cotidiano de “módulo”) es de hecho verdadero de la mente humana. Las personas pueden tener su sistema de lenguaje dañado, manteniendo a la vez mucho de la cognición restante intacta (afasia); las personas pueden carecer de la habilidad de razonar acerca de estados mentales, mientras siguen siendo capaces de muchas otras cosas (autismo); las personas pueden perder su capacidad de reconocer caras humanas; alguien puede perder su capacidad de razonar acerca de los engaños en el intercambio social, mientras sigue manteniendo capacidades paralelas de razonar acerca de riesgos y peligros, etc. (Sachs, 1985; Shallice, 1988; Tager-Flusberg, 1999; Stone *et al.*, 2002; Varley, 2002).

2.2 El argumento de la tratabilidad computacional

La segunda línea de razonamiento en apoyo de la modularidad masiva parte del supuesto de que los procesos mentales deben ser realizados computacionalmente en el cerebro.² Y el argumento, derivado de Fodor (1983, 2000), es que los procesos computacionales necesitan ser encapsulados para ser tratables.

El argumento puede formularse así.³ Si un sistema de procesamiento puede ver cualquier ítem arbitrario de información en el curso de su procesamiento, entonces los algoritmos que el sistema “corre” deberán ser también arbitrariamente complejos, hasta el punto de resultar computacionalmente intratables. Porque tales algoritmos deberán especificar, respecto a cada ítem de información al que puedan tener acceso, qué paso

² Esta suposición es común en todas las ciencias cognitivas clásicas. Ha sido desafiada recientemente por los partidarios del conexionismo distribuido. Pero más allá del éxito o no de las redes conexionistas en el reconocimiento de patrones, existen razones de principio para pensar que tales redes son incapaces de la clase de aprendizaje “de golpe” y de la actualización de variables de los que los seres humanos y otros animales son manifiestamente capaces. Vd. Gallistel, 2000; Marcus, 2001.

³ Con esto no quiero dar a entender que apruebo este argumento exactamente como está expuesto. Algunas de sus suposiciones serán desarmadas y cuestionadas, y el argumento será reconstruido en el curso de nuestra discusión.

deberán tomar en el momento siguiente, presumiblemente diferente para cada ítem de información, si el sistema debe ser sensible al contexto. A la inversa, para que los algoritmos de un sistema resulten computacionalmente tratables, deberán existir restricciones a la cantidad de información a la que tendrán acceso.

De manera consistente con el argumento de Fodor, lo que más de un cuarto de siglo de investigación en Inteligencia Artificial parece haber enseñado es que los procesos computacionales necesitan ser divididos en un conjunto de subsistemas modulares para ser tratables (Bryson, 2000; McDermott, 2001). Nótese que esta línea del argumento no comienza con un compromiso con un tipo particular de arquitectura computacional (como la arquitectura asumida por Brooks, 1986), diciendo “Hey, este sistema es modular, por lo tanto la cognición es modular”. Más aún, el argumento es una metainducción de corrientes recientes en Inteligencia Artificial. La afirmación es algo así como esto: en la última docena de años, virtualmente todos en Inteligencia Artificial han convergido en una arquitectura modular de un tipo u otro. Esto ha sido forzado por la experiencia de intentar diseñar sistemas que verdaderamente funcionen. Entonces esto nos da buenas razones para pensar que cualquier agente real debe tener una organización interna modular.

De todos modos, bien podría ser, tal como veremos, que la noción de “módulo” con la que se trabaja en Inteligencia Artificial no sea exactamente la misma que la de Fodor. Pero la premisa en la metainducción anterior puede ser cuestionada de modo más directo. Porque existen en oferta también arquitecturas computacionales de carácter amodular, como la arquitectura SOAR de Newell (1990). Ahora bien, puede afirmarse que existe cierta confusión terminológica a este respecto. Aunque todos pueden coincidir en que un sistema es un módulo sólo si es específico de dominio (y, por lo tanto, que una arquitectura es amodular si los sistemas internos *no son* específicos de dominio), diferentes tradiciones de investigación significan cosas diferentes por “dominio”. Entonces, cuando alguien que viene de una tradición *dice* que su arquitectura es amodular, podría ser en realidad el caso de que sea modular en el sentido explicado en la otra tradición. Déjenme que elabore esto.

Los psicólogos del desarrollo y los cognitivos tienden a pensar en los dominios en términos de clases de contenido, o clases de problemas. Cuando dicen que el desarrollo es un proceso específico de dominio, lo que quieren decir es que proceden en pasos diferentes y siguiendo trayectorias diferentes en las diversas áreas de competencia cognitiva que los seres humanos adultos demostramos (psicología del sentido común, física de sentido común, biología de sentido común y demás). Llámese a esto “domi-

nio de contenido”. Los psicólogos evolucionistas y/o los teóricos de la modularidad masiva, en contraste, piensan en los dominios como caracterizados por una *función*. En este último sentido, el dominio de un módulo *se supone* que tiene que ver con la arquitectura total del sistema cognitivo. Llámese a esto “dominio de tareas”. La confusión surge de modo natural, y puede pasar fácilmente desapercibida, debido a los *muchos* módulos de tareas específicas postulados por la psicología evolucionista que son *también* en naturaleza específicos de contenido (el módulo de la psicología de sentido común se centra en estados mentales; el módulo de la física de sentido común es acerca de movimientos físicos de objetos inanimados; el módulo detector de tramposos es acerca de costes y beneficios del intercambio y así). Pero no hay nada en la noción de un módulo *per se* que requiera esto, desde la perspectiva de la psicología evolucionista.

Cuando alguien proveniente de la tradición de la psicología cognitiva afirma “He construido un sistema que es amodular en su arquitectura”, lo que probablemente esté queriendo significar es “He construido un sistema que no opera sobre ningún tipo específico de contenido”. Y es cierto que, por ejemplo, el SOAR de Newell –que fue diseñado para tomar decisiones– puede adquirir la habilidad de tomar decisiones acerca de distintos tipos de asuntos. Pero todavía puede ser un sistema modular desde la perspectiva del psicólogo evolucionista (en tanto podría consistir en sistemas aislables con funciones específicas cuyas operaciones internas estén encapsuladas). En realidad, hay que mirar y ver. Y cuando se mire a SOAR, *parecerá* que está organizado modularmente (a pesar de publicitarse a sí mismo como amodular). Porque diferentes objetivos y subobjetivos vienen con “marcos” de información relevante adjunta. Cuando debe descifrar qué hacer para conseguir un objetivo, el programa sólo puede ver lo que está en el marco relevante. Así que sus operaciones resultan encapsuladas.⁴

Debe notarse de todos modos que la información contenida en un “marco” dado puede cambiar con el tiempo. Esto requiere que nosotros distingamos entre concepciones de la encapsulación *débilmente modulares* y *fuertemente modulares*. En el sentido fuerte, decir que un sistema dado está encapsulado de todo salvo de la información en su base de datos propietaria es decir que no puede tener acceso a ninguna otra información en ningún momento durante su existencia, no importa lo que suceda. Éste es un sentido en que los “marcos” de SOAR no están encapsulados, ya que pueden alterarse con el tiempo y de hecho lo hacen. En el sentido más

⁴ Qué sentido de “encapsulado” está involucrado aquí se explicará en las próximas secciones, especialmente la sección 8.

débil, de todos modos, podemos decir que un sistema está encapsulado si únicamente tiene acceso a cualquier información que esté contenida en su base de datos propietaria en *ese* momento.

Parece que no existe una buena razón para insistir en una interpretación fuerte de la modularidad. Porque la concepción más débil nos da todo lo que *necesitamos* de la modularidad, que es que las operaciones del sistema sean computacionalmente tratables. Piénsese en la facultad del lenguaje, por ejemplo, que es considerada por Fodor (1983) como uno de los módulos arquetípicos. La base de datos procesadora de este sistema seguramente no está congelada todo el tiempo. Siempre se pueden aprender nuevos verbos irregulares, por ejemplo, y seguramente serán considerados como pertenecientes a la base de datos de procesamiento.

Uniendo estas dos líneas de razonamiento, entonces (desde la biología y desde la tratabilidad computacional), lo que conseguimos es la predicción de que la mente humana debe consistir en una cantidad de sistemas funcionales y multifuncionales, que son en algún grado disociables entre sí y cuyos procesos computacionales internos están encapsulados en relación con la mayor parte de la información guardada en algún otro sitio de la mente en ese momento. Ésta es la tesis de la modularidad masiva de la mente, tal como generalmente se la concibe.⁵

3. Heurísticas simples

Cuando la psicología evolutiva comienza con la reflexión sobre los sistemas biológicos en general y propone un programa de investigación para descubrir los elementos de una mente modular, el foco inicial del movimiento de las heurísticas simples es más limitado. Comienza con la consideración de la racionalidad o irracionalidad del razonamiento humano.

Los psicólogos han afirmado desde 1970 que los seres humanos son extremadamente *malos* en muchas clases de razonamientos. Por ejemplo, numerosos estudios que involucran la tarea de selección del condicional de Wason sugieren que las personas no son buenas en la tarea de discernir en qué circunstancias una simple afirmación condicional puede ser verdadera o falsa (Wason y Evans, 1975; Evans y Over, 1996). Y los seres

⁵ Existen, claro, muchas objeciones a las tesis de la modularidad masiva también. La mayoría tiene que ver con el aparente holismo de los procesos centrales cognitivos en los seres humanos para la inferencia y la formación de creencias (Fodor, 1983, 2000) y con la distintiva flexibilidad y creatividad de la mente humana. Estos hechos hacen difícil ver cómo la mente puede consistir sólo (o en gran medida) de módulos. Éste no es el lugar para identificar y responder estas objeciones. Ver Carruthers, 2002a, 2002b, 2002c, 2003, 2004.

humanos cometen falacias con frecuencia, especialmente cuando razonan acerca de la probabilidad, como la falacia de la conjunción, o la de negligir la línea base, etc. (Kahneman y Tversky, 1982). Pero es evidente que, para nosotros, movernos más allá de estas afirmaciones sobre hechos para hacer juicios sobre la irracionalidad humana bien podría requerir que hagamos algunas suposiciones sobre la *naturaleza* de la racionalidad.

De hecho, la pregunta “¿Qué es la racionalidad?” es la misma que la pregunta “¿Cómo *deberíamos* razonar?”. Tanto filósofos como psicólogos han asumido tradicionalmente que deberíamos razonar de manera válida donde sea posible, y, en general, de modos *confiables* (por ejemplo, en dominios de inferencia no demostrativa, como en la ciencia). Pero de hecho, la verdad no puede ser nuestro único objetivo. Necesitamos también *suficientes* verdades en un marco de tiempo lo suficientemente corto como para permitir la toma de decisiones y la acción. Más aún, el razonamiento y la toma de decisiones deben realizarse en procesos computacionalmente *tratables*, si es que deben ser computacionalmente realizables.

Por ejemplo, tradicionalmente se ha asumido que cualquier nuevo candidato a ser una creencia debería ser revisado en su consistencia con las creencias ya existentes, antes de ser aceptado. Pero de hecho, la revisión de la consistencia es un proceso demostradamente intratable, si se plantea de un modo exhaustivo. Considérese cómo se podría revisar la consistencia de un conjunto de creencias en una tabla de verdad. Aun si cada línea pudiera ser revisada en el tiempo en que le toma a un fotón de luz viajar el diámetro de un protón, entonces incluso después de 20 mil millones de años la tabla de verdad para un conjunto de sólo 138 creencias (2.138 líneas) seguiría sin ser revisada completamente (Cherniak, 1986).

Existe un campo completo de la ciencia computacional dedicado al estudio de tales problemas, llamado “teoría de la complejidad”. Pero es importante darse cuenta de que la intratabilidad computacional para los objetivos de la ciencia *cognitiva* puede incluir problemas que no deberían ser caracterizados como intratables por los científicos de la computación. Esto es así porque nuestro objetivo es explicar los procesos que suceden en tiempo real (en segundos o minutos, en vez de en milenios) y porque necesitamos operar con supuestos acerca de la velocidad del procesamiento de nuestros cerebros (*mucho más* lentos que las computadoras modernas, aunque con muchos más procesos en paralelo), y con supuestos sobre nuestra capacidad de memoria. En efecto, esto significa que la idea de una *intratabilidad* computacional, para los propósitos de la ciencia cognitiva, no admite una definición formal. Pero eso es tal como debe ser, ya que estamos lidiando aquí con la ciencia empírica.

Esta línea de pensamiento conduce a la idea de la *racionalidad naturalizada*. Necesitamos procesos de razonamiento que sean lo *suficientemente* confiables, pero también lo *suficientemente rápidos y frugales* (en términos de los recursos temporales y computacionales requeridos), dadas las demandas de la vida de un ser humano normal. Y por supuesto lo que cuente como “rápido” y “frugal” no es algo que pueda ser especificado *a priori* por los filósofos. En cambio, estas cosas dependerán de las propiedades de los sistemas computacionales empleados por los cerebros de los mamíferos en general, y por el cerebro humano en particular, y de las demandas de las tareas con que nuestros ancestros se enfrentaron con regularidad.

Éste es el trasfondo en el que se ha desarrollado el programa de las heurísticas simples. El objetivo es encontrar procedimientos computacionales que sean rápidos y frugales, pero que sean lo *suficientemente* confiables en un medio dado y que valga la pena tener. Por ejemplo, una heurística explorada por Gigerenzer *et al.* (1999) es el *reconocimiento*—si se pregunta cuál de dos ciudades de Alemania es la más grande, la heurística dice que hay que seleccionar la que uno ya conoce, si es que uno conoce sólo una. Esta heurística demuestra ser notablemente exitosa, aun cuando se oponga a otras opciones más elaboradas, como la regla de Bayes, la regresión múltiple, etc.; y demuestra ser exitosa a través de un rango amplio de tipos de decisión (incluyendo la selección de compañías que seguramente tengan un buen desempeño en la Bolsa de Valores).

Nótese que existe un punto importante de similaridad entre el movimiento de las heurísticas simples y la evolución del programa de la psicología evolucionista. Es que ambas ponen un énfasis similar en la tratabilidad computacional de sus mecanismos cognitivos. Pero cada una parece seguir una estrategia diferente en la búsqueda de procesos computacionalmente tratables. Una postula un conjunto de heurísticas simples; la otra, un conjunto de módulos encapsulados. Parecen ser acercamientos distintos—quizás incompatibles—al mismo problema. Hablaremos de estos temas en las secciones que siguen.

4. ¿Un par inconsistente?

¿Son, entonces, inconsistentes los programas de investigación de la modularidad masiva y de las heurísticas simples? A primera vista, por lo menos, parecería que cada uno puede incorporar elementos del otro sin inconsistencia, y quizás obteniendo un beneficio mutuo. De este modo, un modularista masivo puede creer que algunos de los procesos que ocurren internamente dentro de un módulo son de carácter heurístico. Por ejemplo, en vez de buscar exhaustivamente a través de toda la información

en su base de datos propietaria, un módulo podría adoptar una heurística satisficente [satisficing], que deje de buscar cuando ha encontrado un ítem de información que es lo *suficientemente bueno* como para ser usado en esta tarea. Asimismo, un modularista podría aceptar que las heurísticas simples desempeñan un papel al orquestrar las *interacciones entre* los módulos y su influencia en la conducta. De modo similar, los partidarios de las heurísticas simples seguramente aceptarían que al menos algunos de los procesos que importan en la creencia o que conducen a una decisión son de carácter modular.

Además, los teóricos de la modularidad masiva enfatizan que parte del sentido de una arquitectura modular es que los diferentes módulos pueden estar contruidos de tal modo que disponen de información específica relativa a sus dominios y/o pueden desplegar algoritmos que se adaptan a un conjunto específico de demandas de las tareas asignadas. Una idea similar parece encontrarse dentro del marco de trabajo de la heurística simple, a través de la noción de *racionalidad ecológica*. La idea es que, en conexión con cualquier heurística dada, habrá un rango de diferentes entornos y tipos de entornos dentro de los cuales la heurística operará con un grado significativo de fiabilidad. Y podemos pensar acerca de la heurística como siendo seleccionada (por la evolución, por aprendizaje individual, o a través del éxito en una cultura particular) para operar en esos entornos y de ese modo incorporar (en un sentido) información sobre ellos.

Sin embargo, tales aspectos pueden resultar insuficientes para diluir la impresión de conflicto entre los dos programas de investigación. Porque podría parecer que están ofreciendo modelos en competencia sobre la arquitectura innata total de la mente. Mucho de lo que sucede dentro del programa de las heurísticas simples es un intento de modelar varios aspectos de la *toma de decisiones*; y muchas personas asumen que los sistemas de toma de decisión no deben ser modulares (ciertamente no pueden ser específicos de dominio en el sentido de específicos de contenido de “dominio”, de todos modos). Además, muchas de las heurísticas discutidas por aquellos que trabajan dentro del programa de heurísticas simples parecen aplicarse a dominios ampliamente diversos y evolutivamente distintos; y algunos de ellos también podrían ser aprendidos. En contraste, generalmente se concibe la hipótesis de una modularidad masiva desde el supuesto de que la mente consiste en un conjunto de sistemas modulares, dedicados a dominios de especial importancia evolutiva.

De todos modos, el razonamiento práctico puede concebirse de hecho como modular, en el sentido relevante de “módulo”. Recuérdese que la modularidad es acerca de la encapsulación, y no necesariamente acerca

de la especificidad del dominio (en el sentido de contenido). Un módulo de razonamiento práctico sería un sistema que puede tomar cualquier creencia o deseo como input, pero cuya forma de procesarlo es encapsulada. Según está esbozado en Carruthers (2002a) en relación a un módulo de decisión en animales, un sistema de este tipo podría estar configurado para que tome el deseo P más fuerte en un momento dado como el input inicial, y luego pregunte a los otros sistemas modulares de generación de creencias, y/o inicie una búsqueda en la memoria de creencias de la forma $Q \supset P$. Cuando un condicional de esta forma se recibe como input, se comprueba Q en una base de datos de un esquema de acción para ver si es algo factible; si lo es, sigue adelante y lo hace; si no, comienza una nueva búsqueda de creencias de la forma $R \supset Q$. Y así continúa. Quizás tenga también una regla simple para detenerse: si hay que inspeccionar más de n condicionales o si pasa más de t tiempo sin éxito, hay que detenerse y seguir hacia el próximo deseo fuerte. Este tipo de proceso daría lugar a un sistema encapsulado; no a uno específico de dominio.

Uno podría fácilmente imaginar, entonces, que las operaciones de un sistema tal podrían estar suplementadas por un conjunto de heurísticas, tales como: si uno quiere algo, primero debe acercársele. O, en otro ejemplo, mucha literatura sobre navegación sugiere que los niños y otros animales operan con un conjunto anidado de heurísticas cuando están desorientados (Shusterman y Spelke, 2005). La secuencia parece ser algo como esto: si uno no sabe dónde está, buscar un hito direccional (por ejemplo, una marca de referencia en la distancia o la posición del sol); si no hay tal hito, buscar en las propiedades geométricas del entorno; si la información geométrica no es de ayuda, buscar una marca próxima sobre el terreno. Asimismo, es factible que el sistema de razonamiento práctico pueda emplear una variedad de heurísticas sucesivas, con una condición de final de búsqueda (por ejemplo, cuando se decide con quién casarse, Gigerenzer *et al.*, 1999). Y así. Nada de esto parece inconsistente con la afirmación de que el sistema de razonamiento práctico es modular.

Y con respecto al hecho de que las heurísticas parecen aplicarse a través de lo que son dominios evolutivos distintos, recuérdese la metáfora de la “caja de herramientas adaptativas”, que parece ser central en el modo en que se piensa acerca de la mente en el programa de las heurísticas simples. Una manera de elaborar esto sería proponer la existencia de procedimientos heurísticos que puedan instanciarse de forma múltiple dentro de un rango de sistemas modulares distintos (entonces, una heurística dada es un *tipo* de regla de procesamiento, que puede ser instanciada muchas veces dentro de diferentes sistemas en el cerebro, en vez de un sistema de procesamiento por derecho propio). En efecto,

no hay una razón para pensar que cada módulo tenga que desplegar un conjunto *único* de algoritmos para procesar su dominio. Podría haber un rango de tipos de algoritmos/estrategias heurísticas que podrían ser adoptadas una y otra vez por los diferentes sistemas modulares.⁶ Tampoco hay ninguna razón para pensar que los sistemas modulares deban usar algoritmos “de maximización” o exhaustivos. Por el contrario, las presiones de velocidad y tiempo deberían conducir a la evolución de reglas de toma de decisiones “rápidas y sucias”, tal como conducen a tales reglas en la cognición.

Igualmente (y además), la modularidad masiva ciertamente no es inconsistente con el aprendizaje. De hecho, muchos módulos se entienden mejor como módulos aprendidos. Y algunos módulos probablemente sean construidos por otros mecanismos de aprendizaje en vez de ser innatos (los de *lectura* se me vienen a la mente como ejemplo). Además, mientras algunos de estos mecanismos pueden parecer enfocados a un solo dominio de contenido, otros podrían involucrar la interacción de una variedad de diferentes módulos y mecanismos de aprendizaje modular (y por lo tanto, aparentando generalidad de dominio). Y luego bien podría ser que exista un juego de principios heurísticos, de operaciones, que pueda ser seleccionado entre los de la “caja de herramientas” de implementación de uno de estos módulos, si es apropiado. El proceso de aprendizaje podría involucrar en parte la selección de la herramienta apropiada de la caja.

5. ¿Se ve afectado el argumento de la tratabilidad computacional?

Hemos visto que la hipótesis de la modularidad masiva parece ser totalmente consistente con las afirmaciones hechas por el programa de las heurísticas simples. Parece, no obstante, que el éxito del segundo programa debe socavar uno de los argumentos principales *en apoyo* a la modularidad masiva —específicamente, el argumento de la tratabilidad computacional. Porque parece que los mecanismos computacionales de base

⁶ Marcus (2004) explica cómo la evolución con frecuencia opera empalmando y *copiando*, siguiendo una adaptación. Primero, los genes que resultan en una microestructura dada (un banco particular de neuronas, por ejemplo, con un conjunto dado de propiedades de procesamiento) son copiados, dando lugar a dos o más instancias de tales microestructuras. Luego, en segundo lugar, algunas de las copias pueden ser adaptadas a nuevas tareas. *Algunas veces*, esto involucrará modificar el algoritmo de procesamiento que es implementado. Pero frecuentemente sólo involucrará la provisión de nuevas conexiones de inputs y/o outputs para el nuevo sistema.

heurística ofrecen un camino para que la computación sea tratable sin necesidad de una encapsulación informacional. Si es así, entonces los procesos cognitivos pueden ser computacionalmente tratables (porque son de base heurística) *sin* que se estructuren a través de líneas modulares, lo que hace colapsar el argumento de que “la cognición debe ser modular para que pueda realizarse en una forma computacionalmente tratable”.

Para poder evaluar esta objeción, necesitaremos realizar un examen más detallado de la noción de encapsulamiento. Pero propongo comenzar este examen dando inicialmente un paso hacia atrás: hacia las consideraciones sobre la tratabilidad computacional que supuestamente dan lugar a las demandas de que la cognición debe ser construida a partir de los sistemas encapsulados.

Podemos distinguir dos modos diferentes en los que un sistema puede no ser computacionalmente tratable. Uno es que sus algoritmos requieran consultar *demasiada información* para llegar a una solución en tiempo real. Por ejemplo, considérese un mecanismo muy sencillo de comprobación de consistencia. Dado un candidato a convertirse en una nueva creencia, avanza con esfuerzo a través del conjunto total de las creencias del sujeto, buscando una contradicción explícita. Aunque el algoritmo ejecutado aquí pueda parecer extremadamente simple (esencialmente se parece a cualquier par de creencias de la forma P, P), un intento de tomar cada creencia como input (ya sea secuencialmente o en paralelo) seguramente no sería factible para una mente/cerebro como el nuestro. Llámese a la demanda correspondiente en los sistemas computacionalmente tratables *frugalidad de información*. Podemos decir que la cognición necesita ser realizada en sistemas de información frugal para ser tratable.

El otro modo en el cual un sistema puede no ser computacionalmente tratable es si los algoritmos que corren son *muy complejos* para ser ejecutables de modo factible en tiempo real. Considérese, por ejemplo, un comprobador de consistencia que opera usando la siguiente heurística cruda, que sólo requiere considerar conjuntos relativamente pequeños de creencias. Para cualquier candidato a ser una nueva creencia, selecciona de manera aleatoria un conjunto pequeño de los cientos de creencias existentes y genera una tabla de verdad para el conjunto de proposiciones seleccionadas, comprobando cada línea para ver si existe un set completo de “F”. Es fácil ver que la cantidad de tiempo y memoria de trabajo que este sistema necesitaría para completar su tarea se elevaría exponencialmente con el tamaño de su conjunto de inputs. Como hemos notado antes, aun cuando comprobara cada línea de la tabla de verdad en el tiempo que le toma a la luz viajar el diámetro de un protón, comprobar la consistencia de tan sólo 138 creencias le llevaría al sistema más tiempo que el que

ha transcurrido desde el nacimiento del universo –¡sin contar el tiempo necesario para generar la tabla de verdad en primer lugar! Llámese a la demanda correspondiente en los sistemas computacionalmente tratables *frugalidad de procesamiento*. Podemos decir que la cognición necesita ser realizada en sistemas de procesamiento frugal si debe ser tratable.⁷

El argumento de la tratabilidad computacional, entonces, nos conduce a pensar que la cognición consiste en sistemas que son *tanto* frugales en cuanto a la información *como* frugales en cuanto al procesamiento. Pues bien, una manera de hacer un sistema frugal en información sería negarle el acceso a cualquier tipo de información almacenada. Esto nos daría un arquetipo de un módulo de input, tal como lo exploró Fodor (1983). Sería un sistema que puede recibir y procesar información transducida sensorialmente, pero que no puede acceder a ninguna de las informaciones guardadas en otro lugar de la mente. Pero, por supuesto, esto no puede ser un modelo general de cómo debe ser un módulo, si es que estamos buscando extender la noción de modularidad a los sistemas centrales que operan sobre creencias (y deseos). Una vez que pasamos de los módulos de *input* a considerar los supuestos módulos *centrales*, ya no podemos pensar más en el encapsulamiento.

En este punto, el camino natural para avanzar supone distinguir entre el input de un módulo y *la base de datos de procesamiento* del módulo (Sperber, 2002; Carruthers, 2003). Un módulo central no específico de contenido podría ser un sistema que tome cualquier información almacenada como input, pero de modo encapsulado con respecto a su procesamiento; tanto sin acceso alguno a la información almacenada al ejecutar su algoritmo (en cuyo caso el sistema está encapsulado completamente), como con acceso restringido a una base de datos que es relevante para la ejecución de esos algoritmos (en cuyo caso, el sistema es encapsulado en un grado inversamente proporcional al tamaño de la base de datos).

⁷ Debe hacerse notar que las nociones de *demasiada* información y de un procesamiento que es *muy complejo*, como se desarrollan aquí, se mantienen en gran medida indeterminadas. La reflexión de sentido común sobre las circunstancias de la vida humana puede darnos algunos sentidos de los tipos de escalas de tiempo dentro de las cuales los sistemas cognitivos deben cumplir con sus tareas, por supuesto. Pero falta la mayoría de los otros parámetros que necesitarían ser tenidos en cuenta. No sabemos mucho acerca de la velocidad de procesamiento de los sistemas cerebrales, cuando se consideran desde un nivel cognitivo de descripción, en comparación al neurológico. Tampoco sabemos mucho acerca de la capacidad de memoria de varios sistemas que pueden estar involucrados. Por ello, cualquier juicio que hagamos acerca de si un sistema dado es o no lo suficientemente frugal deberá ser tentativo, y a expensas de futuros descubrimientos en ciencia cognitiva.

Con esta sugerencia tosca sobre la mesa, la cuestión se resume en si la tratabilidad computacional (y de ahí la *frugalidad*) requiere la encapsulación informacional. En tal caso, para el sistema computacional y sub-sistema: (a) necesitamos poder identificar su input, y distinguirlo de su base de datos de procesamiento (si la hay); y (b) su base de datos de procesamiento debe ser un conjunto pequeño del total de información disponible. Por el contrario, si el programa de las heurísticas simples sugiere un modo de tener frugalidad en ausencia de (a) o (b), entonces el argumento de la tratabilidad computacional para la modularidad masiva resultaría socavado.

6. Heurísticas y bases de datos de procesamiento

Para poder ver si el programa de **heurísticas simples** socava el argumento a favor de la encapsulación del procesamiento, necesitamos examinar si las aplicaciones particulares de ese programa de investigación –tales como las heurísticas de reconocimiento, de “tomar el mejor”, o “tomar el más reciente” y demás– pueden apoyar una división apropiada entre el *input* del sistema y su *base de datos de procesamiento*.

La heurística de reconocimiento opera más o menos de la siguiente manera. Cuando se tiene que decidir cuál de dos elementos consigue la máxima puntuación en una misma dimensión (por ejemplo, cuál de dos ciudades alemanas es la más grande), si uno solamente reconoce uno de ellos, entonces debe seleccionar ése (si los dos son reconocidos, entonces se tendrá que utilizar alguna otra heurística). Para nuestros propósitos, el punto importante que hay que resaltar es que la heurística de reconocimiento opera de modo completamente encapsulado. No hay otra información en la mente que influya o pueda influir en el resultado, excepto quizás la información que está de algún modo involucrada en el proceso de reconocimiento mismo. Una vez que el sistema ha recibido la tarea de juzgar el tamaño, sólo debe tener en cuenta los objetos presentados para determinar cuál de ellos evoca reconocimiento.⁸ No se necesita consultar otra información (ni puede serlo, de hecho, o al menos no de modo interno a la operación heurística de reconocimiento misma) y el proceso inferencial involucrado es muy simple. Por tanto, parecería que las instanciaciones de la heurística de reconocimiento merecen ser consideradas como *módulos* en el sentido tradicional.

⁸ De ahí que las bases de datos de procesamiento para el sistema puedan consistir en un conjunto de conceptos, junto con cualquier información requerida para el reconocimiento del objeto. Esto probablemente sea un pequeño subconjunto del total de información contenida dentro de una mente.

Ahora considérese la heurística de “tomar el mejor”. A diferencia del reconocimiento, esta heurística requiere que el sistema busque y consulte más información con respecto al ítem en cuestión. Pero no busca en *toda* la información disponible. Específicamente, busca el ítem de información que concierne a los dos elementos-diana que han sido encontrados con mayor frecuencia en el pasado para discriminar entre elementos de ese tipo a través de la dimensión requerida. Gigerenzer *et al.* (1999) han mostrado que la heurística puede tener un desempeño casi tan bueno como un conjunto de algoritmos de procesamiento más sofisticados, pero que puede hacerlo siendo mucho más frugal en la información que se usa y en las demandas que plantea a los recursos computacionales del sistema.

En este caso, no es fácil ver cómo puede trazarse la distinción entre *input* y *base de datos de procesamiento*. Uno puede intentarlo, diciendo que el subconjunto relevante del total de información que puede consultar durante el procesamiento un sistema que instancia la heurística “tomar el mejor” consiste en sus creencias acerca de la validez relativa de un aspecto junto con sus otras creencias acerca de los aspectos en cuestión. Cuando se le plantea una pregunta acerca del tamaño relativo de dos ciudades alemanas, por ejemplo, debe mirar primero sus creencias acerca de qué propiedades de las ciudades han correlacionado mejor con el tamaño en el pasado. Si tener un equipo de fútbol en la primera división es el mejor predictor, entonces consultará a la base de datos más amplia: ¿tiene alguna de estas ciudades un equipo de fútbol de primera división? Si recibe de vuelta la información de una de ellas y lo tiene, entonces selecciona ésa como la de mayor tamaño. Si ninguna o las dos lo tiene, avanza al siguiente mejor predictor de tamaño en la base de datos de procesamiento. Y así sucesivamente.

Nótese, de todos modos, que cuáles sean las creencias que tal sistema *puede* consultar en el curso de su procesamiento es una función de qué creencias *tiene* en realidad. Si se cree que tener una tasa alta de crímenes es el mejor predictor del tamaño de una ciudad, entonces *ésa* es la información que hay que buscar. Y en principio cualquier creencia *podría* tener un impacto en el procesamiento. Entonces, parece que nuestra mayor esperanza de encontrar un lugar para la noción de “encapsulamiento” aquí sería adoptar una idea de nuestra anterior discusión sobre la arquitectura de planes SOAR. Podríamos considerar las creencias específicas que un sistema que instancia “tomar el mejor” adquiere como la selección de una base de datos de procesamiento funcionalmente individuada a partir de un conjunto más amplio de información almacenada en relación a cada dimensión de comparación, de tal modo que el sistema sólo *puede* considerar ese conjunto más reducido para responder a una cues-

ción dada. Pero tiene que admitirse que esto parece bastante forzado y poco natural.

Ahora considérense los procesos heurísticos que descansan sobre fenómenos tales como la *saliencia* de un trozo de información en un contexto o la *accesibilidad* de esa información dada la historia reciente de su activación (esto último está muy relacionado con la heurística que Gigerenzer *et al.*, 1999 llaman “tomar el último”). Considérese la comprensión del lenguaje, por ejemplo, en el tipo de modelo proporcionado por Sperber y Wilson (1996), en donde la accesibilidad de creencias desempeña un papel central. En su explicación, uno de los factores en la interpretación es la saliencia en el entorno actual, y otro es cercanía [*regency*] relativa (por ejemplo, si un ítem de información ha sido activado o no anteriormente en la conversación).

¿El proceso de comprensión podría entenderse como encapsulado, si no obstante en principio *cualquier* creencia *podría* ser considerada saliente en el entorno actual, o *podría haber sido* activada previamente? Si fuera así, tendríamos que pensar que el proceso de comprensión, mientras se despliega en el curso de un conjunto de intercambios lingüísticos, crea una suerte de módulo de comprensión local “sobre la marcha”, cuyas condiciones de encapsulamiento se modifican continuamente mientras la conversación continúa. ¿Pero en qué se convierte la idea de que hay algún subconjunto del total de información disponible que el sistema de comprensión puede tener en cuenta, si es que *cualquier* ítem de información *podría haber sido* saliente?

Podría responderse que, de todos modos, estamos lidiando aquí con un sistema encapsulado *efímero*, creado a partir de los recursos de un sistema de comprensión duradero, a partir de hechos sobre el entorno reciente. Dada la historia previa de la conversación, algunos de los ítems de información son mucho más accesibles que otros. Entonces, un proceso de búsqueda que opera sobre principios de accesibilidad sólo puede tener en cuenta esa información, y otra información en la mente *no puede* influir en el proceso de comprensión. Ciertamente, si los hechos anteriores sobre la conversación hubieran sido diferentes, entonces otra información habría tenido una influencia sobre la comprensión de la oración en cuestión. Pero esto no altera el hecho de que, dada la historia previa de la conversación, esa información ahora *no puede* tener una influencia.

Aunque hay un sentido en el que esta réplica funciona, su victoria es pírrica. Porque la noción de modularidad resultante es altamente problemática. La ciencia cognitiva, tal como cualquier otra ciencia, tiene la misión, *inter alia*, de descubrir y estudiar las propiedades de un conjunto de *clases naturales* dentro de su dominio. Y una clase natural, para ser

un objeto de estudio que valga la pena, debe tener un cierto tipo de *estabilidad*, o una recurrencia regular. Por el contrario, el estado de un sistema de comprensión que ha experimentado una historia conversacional específica, y que por tanto tiene una distribución particular de grados de accesibilidad entre sus representaciones, es algo que puede existir sólo una vez en la historia del universo. Esta combinación particular de principios de procesamiento y accesibilidad (que da lugar a la “base de datos de procesamiento” de un módulo “sobre la marcha”) podría no ocurrir nunca más.

Si la ciencia cognitiva debe conseguir el tipo de generalidad que una espera de una ciencia, entonces necesita cincelar sus clases en junturas *recurrentes*. Esto requiere que pensemos el sistema de comprensión como un *solo* sistema en el tiempo, operando parcialmente sobre principios de accesibilidad que ayuden a hacer sus operaciones frugales en cuanto a la información. Del mismo modo, incluso en el caso de SOAR (para volver a un ejemplo que apareció mucho antes en nuestra discusión, si bien cosas similares pueden decirse acerca de “tomar el mejor”, discutido recientemente), probablemente debemos pensar que éste es *el mismo* sistema que se emplea en la búsqueda de una variedad de diferentes objetivos, en los que la frugalidad de información está garantizada al organizar su base de datos en “marcos” enlazados a cada tipo de objetivo. No debemos pensarlo como un conjunto completo de sistemas encapsulados que se superponen (uno por cada sistema de procesamiento con su “marco” emparejado) que comparten un mismo conjunto de algoritmos.

7. Información de input versus base de datos de procesamiento

Algunos ejemplos de procesamiento seleccionados del programa de las heurísticas simples parecen poner en aprietos la noción de un sistema *encapsulado*, donde esta noción se explica en base a la distinción intuitiva entre *input* y *base de datos de procesamiento*. Merece la pena, no obstante, preguntar a su vez directamente cómo se podría explicar esa distinción. Los argumentos anteriores intentan *usar* la distinción entre input y base de datos de procesamiento, sin decir en qué consiste esta distinción. Pero ¿cómo debe trazarse?⁹

⁹ Nótese que la distinción es sólo problemática con respecto a los módulos centrales, cuyos inputs pueden incluir creencias u otros estados proposicionales almacenados. Cuando el módulo en cuestión es un módulo input, la distinción es clara: el input al sistema es la información que le llega de sus transductores sensoriales y cualquier otra información que se utilice en el procesamiento puede considerarse como parte de la base de datos de procesamiento.

Una manera de hacerlo sería pensar el input de un sistema como *cualquier cosa que lo active*. Pero ésta no parece una propuesta muy plausible. Claro que querríamos permitir que, una vez que fue “activado” por algo (por ejemplo, por un deseo que ganó a otros en la competencia por controlar los recursos del sistema de razonamiento práctico), un sistema pudiera reclamar información de otros sistemas, sin por ello considerar tal información como perteneciente a la base de datos de procesamiento del sistema en cuestión. La Inteligencia Artificial, tal como se la practica en la actualidad, está llena de redes de sistemas distintos que pueden pedirse entre sí información, una vez que han sido “activados”. Pero si fuéramos a adoptar la propuesta anterior, entonces deberíamos decir que había un sentido en el que, en realidad, formaban *un* gran sistema, ya que la información producida por cada uno formaría parte de la base de datos de procesamiento de cada uno de los otros.

Otro camino que podríamos tomar sería decir que la base de datos de procesamiento de un sistema, para que cuente como tal, debería ser una base de datos *dedicada*, cuyos contenidos no estén disponibles a otros sistemas. Esto encaja bastante bien con el modo en que se piensa acerca del módulo del lenguaje. Uno podría considerar la base de datos de procesamiento para la facultad del lenguaje como un conjunto de ítems de información adquiridos, específicos del lenguaje –sobre reglas gramaticales y fonológicas, por ejemplo–, que no están disponibles para ningún otro sistema. No parece bien motivado, de todos modos, insistir en que los sistemas de memoria y de procesamiento deban alinearse uno-a-uno del modo en que se deriva de esta sugerencia. Porque la modularidad es una tesis acerca del *procesamiento*, no una tesis sobre el almacenamiento. Es muy poco claro por qué no podría haber sistemas de almacenamiento de información multipropósito a los que cierto número de sistemas de procesamiento pudieran acceder durante su actividad. Ni está claro por qué se vería comprometido por ello el estatus modular de estos sistemas de procesamiento.

Otra alternativa es concebir la base de datos de procesamiento de un sistema como el conjunto de información que éste *debe* consultar para poder ejecutar sus algoritmos. El *input* al sistema (si el sistema no es de contenido específico) podría en principio venir de cualquier lado. Pero una vez que el sistema se activa, requerirá comenzar a consultar algunas partes de su base de datos de procesamiento para poder manejar su input (el sistema no necesita intentar consultar *toda* la información en su base de datos de procesamiento, por supuesto. Éste es uno de los lugares donde ayuda imaginar diversas heurísticas y reglas de búsqueda operando dentro de un módulo dado).

Esta propuesta parece encajar perfectamente con los modos en que algunos tienden a pensar acerca del módulo del lenguaje o el módulo de la teoría de la mente, por ejemplo. Cuando se recibe un input lingüístico, el módulo del lenguaje *debe* comenzar a consultar su base de datos de reglas gramaticales y fonológicas, su lexicón y así sucesivamente. En el curso del procesamiento, *podría* también solicitar información a otros sistemas sobre los intereses o conocimientos probables del hablante; y las respuestas a estas consultas podrían entenderse como inputs futuros del sistema. Y lo mismo cuando la descripción de un ítem de conducta es recibida por el sistema de teoría de la mente: *debe* comenzar consultando su base de datos de información adquirida (por ejemplo, con respecto al significado convencional de un apretón de manos en la cultura que lo rodea, o acerca de los estados mentales atribuidos previamente a un actor). Y en el curso de su procesamiento, también, *podría* solicitar información a otros sistemas acerca de las consecuencias físicas probables de la conducta observada del sujeto, por ejemplo.

De todos modos, existen buenas razones por las cuales uno no puede explicar la encapsulación en términos de la información que el sistema *debe* consultar. Considérese el módulo de razonamiento práctico esbozado más arriba. Una vez que ha sido activado por un deseo, sus algoritmos *lo obligan* a buscar información bajo una cierta forma condicional. Pero estas creencias condicionales pueden en principio provenir de cualquier lado y ser acerca de cualquier cosa. Entonces, si decimos que la base de datos de procesamiento para un sistema es el conjunto de creencias que *se ve obligado a* consultar, entonces casi todas las creencias podrían pertenecer a esta base de datos, y el razonamiento práctico resultaría ser, después de todo, radicalmente des-encapsulado y a-modular.¹⁰

Una opción final es recurrir a la distinción entre llevar a cabo una búsqueda de información por uno mismo y enviar una petición de información a otro. Podríamos decir que la base de datos de procesamiento para un módulo es la información que (el módulo) consulta, en lugar de la que recibe de otros sistemas. Pero concentrémonos ahora en ese aspecto del

¹⁰ ¿Caigo en petición de principio al asumir que el razonamiento práctico es modular? No. Porque lo que está en cuestión es si existe una noción de “encapsulamiento” modular que pueda ser puesta al servicio de la tesis de la modularidad masiva; y esta noción de encapsulación debe seguramente sostener que el razonamiento práctico es modular. Además, el sistema de razonamiento práctico, tal como se ha esbozado antes, sí parecía ser computacionalmente tratable. Si se supone que la tratabilidad requiere encapsulamiento, entonces necesitamos de nuevo una noción de encapsulamiento que encaje con las operaciones de tal sistema.

requisito del razonamiento práctico para las creencias condicionales que involucren una búsqueda entre la información condicional *almacenada* (en oposición a peticiones a otros sistemas para ver si pueden generar tal información en esas circunstancias). ¿Por qué no podría el sistema de razonamiento práctico hacerlo por sí mismo? Si la memoria es accesible en base al contenido o por la forma sintáctica, se podría buscar bajo la descripción “condicional con *P* como consecuente”. ¿Existe alguna razón para pensar que llevar a cabo una búsqueda tal convertiría al razonamiento práctico en intratable (o más intratable que si fuera el resultado de la petición a otro sistema)?

Parece entonces que la distinción entre *input* y *base de datos de procesamiento* no puede hacer el trabajo requerido en el contexto de una tesis de modularidad masiva de la mente, por lo menos no si queremos dar cabida a sistemas modulares con inputs sin restricciones, que puedan solicitar información a otros sistemas, que puedan llevar a cabo búsqueda de información ellos mismos durante su procesamiento, etc. En este punto, la noción de *encapsulamiento*, y con ella la de *modularidad*, parecen encontrarse en un verdadero aprieto en el contexto de la tesis de la modularidad masiva de la mente.

8. ¿Dónde queda la modularidad?

En consecuencia, el programa de las heurísticas simples supone una dificultad considerable para la afirmación de que la cognición debe ser construida a partir de sistemas *encapsulados*, si por un sistema encapsulado se entiende un sistema capaz de recibir cualquier cosa como input, pero que sólo puede acceder a una base de datos de información limitada en el curso de su procesamiento. Pero ¿es así como *debemos* entender la noción de encapsulación? ¿Existen otras alternativas abiertas para nosotros?

Puesto del modo más neutral posible, podemos afirmar que la idea de un sistema encapsulado es la idea de un sistema cuyas operaciones *no pueden ser* afectadas por la *mayoría* o por *toda* la información almacenada en otra parte de la mente. Pero existe un margen de ambigüedad aquí. Podemos hacer que el operador modal tome un alcance restringido con respecto al cuantificador o podemos hacer que tome un alcance amplio. En su forma restringida, un sistema encapsulado sería así: con respecto a la mayoría de la información almacenada en la mente, el sistema en cuestión *no podría ser* afectado por *esa* información en el curso de su procesamiento. Llámese a esto “encapsulación de alcance restringido”. En su sentido amplio, por otro lado, un sistema encapsulado debería ser así: el sistema es tal que *no puede ser* afectado por la mayor parte de la infor-

mación mantenida en la mente en el curso de su procesamiento. Llámese a esto “encapsulación de alcance amplio”.

La encapsulación de alcance restringido es la que se da por supuesta en la literatura filosófica sobre modularidad, siguiendo a Fodor (1983).¹¹ La mayoría de nosotros nos hacemos una idea de la modularidad en términos de un cuerpo de información determinado, tal que esa información no puede penetrar en el módulo. Y esta manera de ver el asunto se ve reforzada si explicamos el encapsulamiento en términos de la distinción entre el input a un sistema y su base de datos de procesamiento, porque de este modo se supone que existe un cuerpo de información determinado (la información en la base de datos de procesamiento) que *puede* afectar las operaciones del módulo, implicando que todas las otras informaciones *no pueden afectarlo* (excepto si son tomadas como inputs).

De todos modos, puede ser verdad que las operaciones de un módulo no puedan ser afectadas por la mayoría de la información en una mente, sin que pueda establecerse una subdivisión entre la información que puede afectar el sistema y la información que no puede hacerlo. Ya que podría ser el caso que los algoritmos del sistema estuvieran configurados de modo tal que sólo un monto limitado de información pueda ser consultado antes de que la tarea se complete o aborte. Póngase de este modo: un módulo puede ser un sistema que sólo *debe* considerar un subconjunto pequeño de la información disponible. No es esencial si lo hace vía encapsulamiento, según se entiende tradicionalmente (la variedad de alcance restringido), o vía heurísticas de búsqueda frugal y reglas de fin de búsqueda (encapsulamiento de alcance amplio). Lo importante es que sea frugal en cuanto a información y frugal en cuanto a procesamiento.

Finalmente, entonces, el siguiente argumento está equivocado: si los procesos cognitivos deben ser realizados de modo tratable, entonces la mente debe estar construida a partir de sistemas de procesamiento que estén encapsulados en el sentido del alcance restringido. Por tanto, el argumento en favor de la modularidad masiva, tal como es tradicionalmente concebido (por los filósofos), también es incorrecto. Pero todavía contamos con el argumento de que la tratabilidad computacional requiere un

¹¹ De todos modos el uso del término “módulo” en la literatura de Inteligencia Artificial es probablemente bien diferente (comunicación personal de Joanna Bryson, Jack Copeland, John Horty, Aaron Sloman). Debe ser cercana a la combinación del sentido cotidiano de “módulo” significando un componente de procesamiento individuado funcionalmente junto con un requerimiento de lo que aquí yo llamo “encapsulamiento de alcance amplio”. De ser así, entonces el argumento de la modularidad masiva de las tendencias recientes en Inteligencia Artificial esbozado en la sección 2 puede todavía sostenerse dado este sentido de “modularidad”.

encapsulamiento en sentido amplio; el argumento que afirma que si los procesos cognitivos deben ser computacionalmente tratables, la mente debe estar construida a partir de sistemas cuyas operaciones son tanto frugales en información como frugales en procesamiento, y eso significa que estos sistemas sólo deben acceder a un subconjunto pequeño de la información total disponible mientras ejecuta sus tareas.

¿Significa esto que la tesis de la modularidad mental masiva tiene insuficiente apoyo y debe ser rechazada? Bueno, por supuesto que eso depende de qué sigamos entendiendo por “módulo”. Todavía tenemos en juego el argumento de la biología acerca de que deberíamos esperar que la cognición esté construida a partir de sistemas separables, de tareas específicas (esto es, en el sentido cotidiano de “módulo”). Y todavía tenemos un argumento proveniente de la tratabilidad computacional acerca de que los sistemas necesitan ser tanto frugales en información como frugales en procesamiento. Esto requiere que estos sistemas deban ser encapsulados de alcance amplio (necesitan ser sistemas que no puedan acceder más que a un pequeño subconjunto del total de información disponible antes de completar sus tareas). Y depende de nosotros decir que éste es el modo en que la tesis de la modularidad masiva debe ser comprendida.

Más aún, todavía tenemos en juego un argumento metainductivo de las corrientes recientes en Inteligencia Artificial. Los investigadores que asumieron el propósito de construir sistemas inteligentes han convergido mayormente en arquitecturas en las cuales el procesamiento dentro del sistema total está dividido entre un amplio conjunto de sistemas de procesamiento de tareas específicas, que pueden pedirse información entre sí y acceder a bases de datos compartidas. Pero muchos de estos sistemas desplegarán algoritmos de procesamiento que no son compartidos por los demás. Y muchos de ellos no sabrán o se preocuparán sobre lo que están haciendo los otros. El hecho de tal convergencia es, entonces, una buena evidencia de que es así como estará también organizada la mente.¹²

Claro que el término “módulo” ha sido usado de *muchas* diferentes maneras dentro de la literatura en ciencias cognitivas, desde Fodor

¹² De hecho, la convergencia es en realidad más amplia, abrazando la ciencia de la computación en general. Aunque el lenguaje de la modularidad no es tan usado por los científicos computacionales, puede argumentarse que se piensa en el mismo concepto bajo el nombre de “programas orientados a objetos”. Muchos lenguajes de programación como C++ y Java requieren ahora sistemas de procesamiento total que tratan algunas de sus partes como “objetos” que pueden ser consultados o informados pero donde el procesamiento que tiene lugar en esos objetos no es accesible para otros. Y la arquitectura resultante es considerada casi como inevitable siempre que se supera un cierto umbral en el grado de complejidad global del sistema.

(1983) en adelante. Una reacción a este lío de diferentes usos podría ser pedir que el término sea dejado de lado y que las personas deban describir con otras palabras qué es en lo que están pensando. Pero resulta muy útil esto de tener un único término para expresar lo que uno quiere expresar. Y dado que uno es explícito acerca de cómo un término está siendo usado, no debería resultar ninguna confusión. Propongo, entonces, que por “módulo” digamos algo parecido a “un sistema de procesamiento distintivo de tareas específicas cuyas operaciones son tanto frugales en información como en procesamiento (y de ahí que es entonces encapsulado en sentido amplio)”.¹³ Y la tesis de la modularidad masiva entonces se convierte en la afirmación de que la cognición debe estar construida a partir de tales sistemas. Comprendida así, la tesis de la modularidad masiva de la mente está bien apoyada y es consistente completamente con las ideas nucleares del programa de investigación de las heurísticas simples.

Lo que en realidad importa finalmente no es cómo *se llama* al sistema en cuestión, sino más bien qué podemos afirmar que conocemos acerca de la arquitectura de la mente humana. Hay un rango de diferentes argumentos (no todos los cuales pueden ser sondeados en este capítulo; para otros ejemplos, Carruthers, 2005), junto con un conjunto de programas de investigación en ciencia cognitiva e Inteligencia Artificial, todos los cuales sugieren que la mente está, de hecho, compuesta por una multiplicidad de sistemas de procesamiento distintos. Estos sistemas hablarán unos con otros y se consultarán, pero lo harán operando de modo altamente independiente entre sí. Y su procesamiento será *frugal* (tanto usando algoritmos construidos especialmente para las demandas de tareas o usando heurísticas o atajos de un tipo u otro, o ambos). Personalmente, estoy inclinado a expresar este resultado diciendo que la mente es masivamente modular en su organización, pero lo que importa aquí es el resultado en sí, no la manera en la cual es descrita.

Menciones

Estoy agradecido a Stephen Stich por una conversación que provocó el grupo principal de ideas en este capítulo, y a Richard Samuels por un intercambio temprano que también fue productivo. Gracias a Mike Anderson, Clark Barrett, Stephen Laurence, Rui Mata, Stephen Stich y Peter Todd por los comentarios a un borrador.

¹³ Para una consideración de un conjunto más amplio de argumentos en apoyo de la modularidad masiva, y la noción resultante de “módulo” que no es la definitivamente esbozada aquí (incorporando la idea de *inaccesibilidad*), vd. Carruthers (2005).

Bibliografía

- Brooks, R. (1986), "A robust layered control system for a mobile robot", *IEEE Journal of Robotics y Automation*, RA-2, pp. 14-23.
- Bryson, J. (2000), "Cross-paradigm analysis of autonomous agent architecture", *Journal of Experimental and Theoretical Artificial Intelligence*, 12, pp. 165-190.
- Carruthers, P. (2002a), "The cognitive functions of language", *Behavioral and Brain Sciences*, 25:6, pp. 657-674.
- (2002b), "Modularity, language, and the flexibility of thought", *Behavioral and Brain Sciences*, 25:6, pp. 705-719.
- (2002c), "Human creativity: its evolution, its cognitive basis, and its connections with childhood pretence", *British Journal for the Philosophy of Science*, 53, pp. 1-25.
- (2003), "On Fodor's Problem", *Mind and Language*, 18, pp. 502-523.
- (2004), "Practical reasoning in a modular mind", *Mind and Language*, 19, pp. 259-278.
- (2005), "The case for massively modular models of mind" en R. Stainton (ed.), *Contemporary Debates in Cognitive Science*, Blackwell.
- Cheng, K. (1986), "A purely geometric module in the rat's spatial representation", *Cognition*, 23, pp. 149-178.
- Cherniak, C. (1986), *Minimal Rationality*, MIT Press.
- Evans, J. y Over, D. (1996), *Rationality and Reasoning*, Psychology Press.
- Fodor, J. (1983), *The Modularity of Mind*, MIT Press.
- (2000), *The Mind doesn't Work that Way*, MIT Press.
- Gallistel, R. (1990), *The Organization of Learning*, MIT Press.
- (2000), "The replacement of general-purpose learning models with adaptively specialized learning modules", en M. Gazzaniga (ed.), *The New Cognitive Neurosciences* (segunda edición), MIT Press.
- Gigerenzer, G., Todd, P. y the ABC Research Group (1999), *Simple Heuristics that Make us Smart*, Oxford University Press.
- Hermer, L. y Spelke, E. (1996), "Modularity and development: the case of spatial reorientation", *Cognition*, 61, pp. 195-232.
- Hermer-Vasquez, L., Spelke, E. y Katsnelson, A. (1999), "Sources of flexibility in human cognition: dual-task studies of space and language", *Cognitive Psychology*, 39, pp. 3-36.
- Kahneman, D., Slovic, P. y Tversky, A. (eds.) (1982), *Judgment under Uncertainty: heuristics and biases*, Cambridge University Press.
- Marcus, G. (2001), *The Algebraic Mind*, MIT Press.
- (2004), *The Birth of the Mind: how a tiny number of genes creates the complexities of human thought*, Basic Books.

- McDermott, D. (2001), *Mind and Mechanism*, MIT Press.
- Newell, A. (1990), *Unified Theories of Cognition*, Harvard University Press.
- Pinker, S. (1997), *How the Mind Works*, Penguin Press.
- Sachs, O. (1985), *The Man who Mistook his Wife for a Hat*, Picador.
- Shallice, T. (1988), *From Neuropsychology to Mental Structure*, Cambridge University Press.
- Simon, H. (1962), "The architecture of complexity", *Proceedings of the American Philosophical Society*, 106, pp. 467-482.
- Seeley, T. (1995), *The Wisdom of the Hive: the social physiology of honey bee colonies*, Harvard University Press.
- Shusterman, A. y Spelke, E. (2005), "Investigations in the development of spatial reasoning" en P. Carruthers, S. Laurence y S. Stich (eds.), *The Innate Mind: structure and contents*, Oxford University Press.
- Sperber, D. (1996), *Explaining Culture: a naturalistic approach*, Blackwell.
- (2002), "In defense of massive modularity" en I. Dupoux (ed.), *Language, Brain and Cognitive Development*, MIT Press.
- Sperber, D. y Wilson, D. (1996), *Relevance*. Blackwell (segunda edición).
- Stone, V., Cosmides, L., Tooby, J., Kroll, N. y Wright, R. (2002), "Selective impairment of reasoning about social exchange in a patient with bilateral limbic system damage" en *Proceedings of the National Academy of Science*, 99, pp. 11531-11536.
- Tager-Flusberg, H. (ed.) (1999), *Neurodevelopmental Disorders*, MIT Press.
- Tooby, J. y Cosmides, L. (1992), "The psychological foundations of culture" en J. Barkow, L. Cosmides y J. Tooby (eds.), *The Adapted Mind*, Oxford University Press.
- Varley, R. (2002), "Science without grammar: scientific reasoning in severe agrammatic aphasia" en P. Carruthers, S. Stich, y M. Siegal (eds.), *The Cognitive Basis of Science*, Cambridge University Press.
- Wason, P. y Evans, J. (1975), "Dual processes in reasoning?", *Cognition*, 3, pp. 141-154.