

# SIMULACIÓN DEL PENSAMIENTO HUMANO \*

*Allen Newell*  
*Herbert A. Simon*

## *Introducción*

ESTE TRABAJO SE INTERESA por la psicología del pensamiento humano. Desarrolla una teoría para explicar cómo algunos humanos intentan solucionar ciertos problemas formales simples. La investigación de donde surgió la teoría <sup>1</sup> está íntimamente relacionada con el campo de procesamiento de la información y la construcción de autómatas inteligentes, de forma que la teoría se expresa en forma de un programa de computador. Los rápidos avances técnicos en el arte de la programación de los computadores digitales para realizar tareas sofisticadas han hecho factible tal teoría.

Se afirma frecuentemente que debe trazarse una cuidadosa línea entre el intento de efectuar con máquinas las mismas tareas que realizan los humanos, y el intento de simular los procesos que usan realmente los humanos para realizar estas tareas. El programa que se discute en este trabajo, GPS (General Problem Solver), confunde en gran medida los dos intentos —para beneficio mutuo—. El GPS se describió anteriormente como un intento para construir un programa de solución de problemas (4, 5), y continúa siendo en nuestra

\* Este artículo está formado con distintos fragmentos de "The Simulation of Human Thought", "GPS, A Program That Simulates Human Thought", y "Computer Simulation of Human Thinking".

<sup>1</sup> Queremos expresar nuestra deuda con J. C. Shaw, que ha sido nuestro colega en la mayor parte de nuestra investigación sobre procesos de información compleja, incluyendo el programa GPS que es el fundamento de este artículo.

presente investigación un magnífico vehículo para explorar el área de la inteligencia artificial. Al mismo tiempo, ciertas variantes del GPS proporcionan una simulación de la conducta humana. Es este último aspecto —el uso del GPS como una teoría de la solución humana de problemas— el que deseamos centrar aquí, atendiendo especialmente a la relación entre la teoría y los datos.

### *El Papel de la Simulación en Psicología*

El camino de la investigación científica en cualquier campo de conocimiento indica una respuesta a dos tendencias opuestas. Por un lado, una poderosa atracción ejercida por los “buenos problemas” —preguntas cuyas respuestas representarían avances fundamentales en la teoría o proporcionarían las bases para importantes aplicaciones. Por otro lado, una fuerte influencia manifestada por las “buenas técnicas”— instrumentos de observación y de análisis que han demostrado que son sutiles y fiables. Los períodos afortunados de una ciencia son aquellos donde estas dos tendencias no paralizan el estudio por su oposición, sino que cooperan para desarrollar la investigación dentro de unos cauces productivos.

Cuando no se satisface sólidamente esta afortunada condición, la ciencia corre el riesgo de una escisión. Algunos investigadores insistirán en trabajar en importantes problemas con métodos que no son suficientemente poderosos y que carecen de rigor; mientras que otros insistirán en atacar problemas que se manipulan fácilmente con instrumentos útiles, al margen de la falta de interés que tengan esos problemas.

Es raro que la tensión que surge del desajuste de fines y medios esté ausente por completo de cualquier ciencia; se podrían mencionar ejemplos de la biología, la meteorología o las matemáticas contemporáneas. Pero ha sido muy evidente en la ciencia de la psicología. En los comienzos de esta centuria, la tesis predominante en psicología fue el asociacionismo. Era una doctrina atomista, que postulaba una teoría de rígidos y escasos elementos, ya fuesen sensaciones

o ideas, que se conectaban o asociaban entre sí sin modificación. Era una doctrina mecanicista, con leyes inflexibles de continuidad en el tiempo y el espacio para explicar la formación de nuevas asociaciones. Estos fueron sus supuestos. La conducta procede mediante la corriente de asociaciones: cada asociación produce sucesores, y adquiere nuevos enlaces con las sensaciones que llegan del medio ambiente.

En la primera década de la centuria, se desarrolló una reacción en contra de esta doctrina a través del trabajo de la escuela de Würzburg. Rechazando la noción de una corriente completamente auto-determinada de asociaciones, introdujo la tarea (*Aufgabe*) como un factor necesario para describir el proceso de pensamiento. La tarea marca una dirección al pensamiento. Una notable innovación de la escuela de Würzburg fue el uso de la introspección sistemática, para arrojar luz sobre el proceso de pensamiento y los contenidos de la conciencia. El resultado fue una combinación de mecanicismo y fenomenismo, que a su vez hizo surgir dos tesis divergentes: el conductismo y el movimiento de la Gestalt.

La reacción conductista insistió en que la introspección era un procedimiento altamente inestable y subjetivo, cuya inutilidad se había demostrado ampliamente en la controversia del pensamiento sin imágenes. El conductismo modificó la tarea de la psicología, en el sentido de explicar la respuesta de los organismos como una función de los estímulos que inciden sobre ellos y la medida objetiva de ambos. Sin embargo, el conductismo aceptó, y aún reforzó, el supuesto mecanicista de que las conexiones entre estímulo y respuesta se formaban y mantenían como funciones simples y determinadas del medio ambiente.

La reacción de la Gestalt realizó un giro opuesto. Rechazó la naturaleza mecanicista de la doctrina asociacionista, pero mantuvo el valor de la observación fenoménica. En muchos aspectos, continuó con la insistencia de la escuela de Würzburg de que el pensamiento es algo más que asociación —el pensamiento tiene una dirección que le da la tarea o la tendencia del sujeto. La psicología de la Gestalt elaboró

esta doctrina de una forma genuina en términos de principios totalistas de organización.

El conductismo y la aceptación de las normas de las ciencias naturales en psicología limitó mucho, durante una generación o más, el rango de los fenómenos de la conducta que el psicólogo, como científico, estaba deseando atacar. A menos que un aspecto de la conducta se pudiese examinar en el laboratorio y se pudiese registrar y medir de una manera completamente objetiva, no era, según el punto de vista predominante, una materia apropiada de estudio.

Hubo en la pasada década una considerable relajación de esta austeridad, aunque no sin recelos y disculpas. Un texto importante de psicología experimental, por ejemplo, al introducir los tópicos de solución de problemas, discernimiento (*insight*), pensamiento y conducta lingüística, señala:

Estos tópicos fueron omitidos frecuentemente de los libros de texto del pasado, quizá por razón de una sutil aura de "mentalismo" que tienen. Históricamente, por supuesto, el pensamiento y el significado fueron los problemas centrales de la psicología. El amplio círculo que los psicólogos norteamericanos han ido dibujando en torno al conductismo parece que los está colocando de nuevo en el núcleo de la ciencia, pero quizá están haciéndolo con técnicas más precisas y con un punto de vista más objetivo que, de otra manera, tal vez no hubiesen alcanzado (7).

No existe necesidad de documentar con detalle este resurgimiento en psicología del interés del tópico central del pensamiento. Esta materia se ha mantenido activa durante el apogeo del conductismo por algunos destacados hombres —recordemos los nombres de Köhler, Tolman, Wertheimer, Bartlett, Duncker y Maier— que no quisieron ceder ante cánones inflexibles de rigor que les impidiese estudiar lo importante y lo significativo. Al mismo tiempo, su trabajo proporcionó fundamentos sobre los que han continuado investigadores más recientes —Luchins-Heidbreder, Harlow, deGroot, Guetzkow y Bruner, por mencionar algunos ejemplos.

Aquellos que miran el pensamiento como el núcleo de la investigación psicológica y que insisten en volver a interesarse por él, no desean regresar a tiempos pasados. Los conductistas y operacionalistas están, por supuesto, en lo cierto al exigir objetividad, claridad y rigor. Pocos psicólogos están satisfechos con la vaguedad del lenguaje de la Gestalt y con el nivel de explicación que permite. Pocos están satisfechos con el lenguaje ecléctico de James, y muchos encuentran excesiva vaguedad y ambigüedad en la hipótesis de "mediación" de sus indecisos descendientes.

La tarea no consiste simplemente en restaurar el pensamiento al centro del escenario psicológico; se debe estudiar el pensamiento con la misma sofisticación metodológica que exigimos para los fenómenos más simples. En consecuencia, la creciente atención hacia los procesos de pensamiento no reafirma solamente su importancia, sino que refleja la creencia en desarrollo de que las técnicas de investigación psicológica han llegado a ser adecuadas, al menos en algún grado, a la materia de estudio. El que la creencia estaba plenamente justificada hace unas décadas, cuando comenzaba el resurgimiento, es un asunto a discutir. La principal tesis de este trabajo es que la creencia está justificada ahora —que han ocurrido los avances tecnológicos que son necesarios para permitir una teoría del pensamiento que pueda formularse y verificarse.

Para comprender los fenómenos complejos, debemos tener instrumentos poderosos de investigación —instrumentos para observar hechos e instrumentos para razonar desde complicadas premisas hasta sus consecuencias. La invención del telescopio y del cálculo jugó un papel crucial en una etapa de la historia de la física, y la invención del ciclotrón y del cuántum produjo los mismos efectos posteriormente en la mecánica. Los nuevos instrumentos de observación hacen visibles los fenómenos anteriormente invisibles, que deben conocerse para que la teoría avance. Los nuevos instrumentos analíticos hacen comprensibles los hechos revelados por el telescopio y el ciclotrón, que de otra manera serían inescrutables. Una ciencia de los fenómenos complejos necesita

máquinas poderosas para observar e instrumentos poderosos para razonar.

Los fenómenos del pensamiento humano son más complejos que los fenómenos que estudia la física. En algunos aspectos, los primeros son más fáciles de observar que los últimos —la conducta verbal humana está presente y se puede oír, ni es submicroscópica ni está tan distante como las estrellas—. (Sin duda, no tenemos instrumentos que sean tan relevantes como los de la física para observar los acontecimientos cerebrales). Pero observable o no, la conducta humana no es fácil de interpretar. Hemos tenido grandes dificultades para construir teorías satisfactorias que la expliquen.

Hasta hace una década, los únicos instrumentos que teníamos para construir teorías sobre la conducta humana eran los instrumentos prestados y adaptados de las ciencias naturales: el operacionalismo y la matemática clásica. Y tan inadecuados son los instrumentos para la tarea, que un psicólogo muy respetado propuso seriamente la doctrina de que debíamos construir una ciencia sin una teoría —seguramente, una doctrina de desesperación.

Con la llegada del computador digital moderno y la aparición del concepto de programa, la situación cambió radicalmente. El computador se inventó como una máquina para hacer aritmética a grandes velocidades. Pero una vez producida, una máquina para hacer aritmética era una máquina que podía manipular símbolos. Era natural preguntar si una máquina tal podría realizar alguno de los procesos más generales de manipulación de símbolos, necesarios para el pensamiento y para la resolución de problemas, al igual que los procesos altamente especializados necesarios para la aritmética. La respuesta, como veremos, es "sí". Existe ahora una sólida evidencia, que revisaremos, de que un computador digital, convenientemente programado, puede desarrollar configuraciones complejas de procesos que son muy paralelos a los procesos observables en los sujetos humanos que están pensando.

Pero la importancia del computador no descansa solamente en su capacidad para manifestar conducta humanoide. Se apoya aún más en el hecho de que podemos especificar con un completo rigor, el sistema de procesos que hacen manifestar al computador esta conducta —podemos escribir un programa que constituya una teoría de la conducta del computador, en el mismo sentido literal que las ecuaciones de la dinámica newtoniana constituyen una teoría de los movimientos del sistema solar. El instrumento analítico absolutamente nuevo del que se dispone para explicar la conducta humana es el programa.

El pensamiento debe explicarse escribiendo un programa para un proceso de pensamiento. Si el programa es complicado —como ocurre normalmente—, de forma que es difícil predecir qué conducta se producirá, codificamos el programa para un computador. Después comparamos la conducta del computador así programado con la conducta de un sujeto humano, que realiza las mismas tareas. De esta forma, el lenguaje de programación proporciona un lenguaje preciso para expresar teorías de los procesos mentales; el computador suministra una máquina poderosa para analizar las consecuencias comportamentales específicas de las teorías, y para comparar estas consecuencias detalladamente —sentencia por sentencia— con la conducta verbal del sujeto humano.

La metodología proporciona una prueba poderosa de la suficiencia de las teorías. Si un programa es vago o incompleto, el computador no funciona —no hace lo que nosotros aseguramos que debe hacer—. A la inversa, si somos capaces de escribir un programa que, realizado en un computador, simula de cerca la conducta humana, podemos afirmar que hemos descubierto un conjunto de mecanismos al menos suficientes para explicar la conducta. No se permite ningún rincón oscuro donde el vitalismo o el misticismo pueda esconderse —ni aun la vaguedad de la hipótesis “medial”.

Estas son grandes pretensiones. Ya es hora de presentar la evidencia para ellas. Sin embargo, sólo podremos hacer

esto después de establecer un poco más cuidadosamente lo que entendemos por una explicación de la conducta y cómo un programa de computador puede constituir una explicación de los procesos de pensamiento humano y de la resolución de problemas.

### *¿Qué es una Explicación?*

Explicar un fenómeno significa mostrar cómo aparece inevitablemente a partir de las acciones o interacciones de mecanismos que se especifican con precisión, que en algún sentido son "más simples" que el fenómeno mismo. Así, una reacción química se explica reduciéndola a las interacciones de átomos que tienen propiedades específicas. Un reflejo espinal se explica reduciéndolo a una secuencia de procesos neurales y sinápticos.

Para los fenómenos complejos puede haber, y normalmente hay, diversos niveles de explicación; no explicamos los fenómenos inmediatamente en términos de los mecanismos más simples, sino que los reducimos a los mecanismos más simples a través de diversas etapas de explicación. Explicamos la digestión reduciéndola a acontecimientos químicos; explicamos las reacciones químicas en términos de procesos atómicos; explicamos los procesos atómicos en términos de las interacciones de partículas subatómicas. Cada aspecto tiene otros aspectos, y el punto de vista científico no acepta ningún nivel de explicación como "último".

Los programas explican la conducta en función de un nivel intermedio de mecanismo, más simple que la conducta misma, pero más complejo que los acontecimientos neurales. Los mecanismos intermedios proporcionan una teoría de la conducta y también un punto de partida para la siguiente etapa de reducción —bien a acontecimientos neurales o aún a otro nivel de mecanismo anterior al neurológico.

Concretamente, el pensamiento humano debe ser explicado en términos de mecanismos simples especificados con precisión, llamados procesos de información elemental. Los procesos de información elemental están organizados en procesos complejos —pensamiento, resolución de problemas,



conducta verbal— mediante programas. Los programas son secuencias largas y ramificadas de procesos elementales. En el desarrollo de la conducta, en cada punto ramificado, se selecciona una continuación determinada y se ejecuta condicionalmente el resultado de una prueba simple (un proceso de información elemental) en función de la identidad o diferencia de un par de símbolos.

En resumen, el estudio y explicación de la conducta humana compleja se realiza de la manera siguiente:

1. La conducta debe explicarse especificando programas que producirán, de hecho, la conducta. Estos programas están compuestos de procesos de información elemental.
2. Los procesos elementales deben explicarse demostrando cómo pueden reducirse a procesos fisiológicos conocidos en el sistema nervioso central y en sus partes dependientes.

Puesto que aquí sólo nos interesa la primera de estas dos tareas de explicación —la reducción de la conducta a procesos de información—, ¿qué garantía existe en contra de la introducción del vitalismo por la puerta falsa? ¿Cómo prevenir que uno de los procesos elementales encubra algún tipo de elan vital? Puesto que no se ha desarrollado la explicación en el segundo nivel, no podemos garantizar directamente que el sistema nervioso humano incluya mecanismos capaces de realizar cada uno de los procesos elementales de información. Pero podemos insistir en que existe algún mecanismo —un mecanismo que puede ser explicado completamente a nivel de la física— capaz de realizar la totalidad de estos procesos. Podemos exigir que los procesos y los programas contruidos, se realicen en un computador digital. Si el computador ejecuta los procesos y, al ejecutarlos, simula el pensamiento humano, entonces ningún misterio vitalista puede estar oculto en los postulados.

No estamos hablando de una burda analogía entre el sistema nervioso y el hardware del computador. El interior de un computador no se parece a un cerebro en mayor me-

dida en que se parece a un proyectil cuando está calculando su trayectoria. Existen varias razones para suponer que los procesos de información simple se realizan mediante mecanismos bastante distintos en el computador y en el cerebro, y desearíamos conocer alguna vez los mecanismos cerebrales como conocemos ahora los mecanismos del computador. Sin embargo, una vez que hemos diseñado mecanismos en un computador para que realice procesos de información elemental, que parecen muy similares a los realizados por el cerebro (aunque mediante mecanismos bastante distintos a los del nivel inmediatamente inferior), podemos construir una explicación del pensamiento en términos de estos procesos de información que es igualmente válida tanto para un computador así programado como para el cerebro.

#### *Programas como Explicaciones*

Hemos descrito un programa como una secuencia de ramificaciones condicionales de procesos elementales de información. Explicar una ruta de conducta mediante un programa es bastante análogo a explicar la ruta de un sistema planetario mediante un sistema de ecuaciones diferenciales. Las ecuaciones diferenciales determinan qué sucederá a continuación (durante el siguiente intervalo "infinitesimal" de tiempo) en función del estado exacto del sistema al comienzo del intervalo. El programa determina qué hará el mecanismo a continuación, en función de su estado exacto en ese momento —este estado depende, a su vez, de la historia previa del sistema y de su medio ambiental actual.

¿Cómo se descubre el programa "correcto" —el que explica la conducta? De la misma forma que se encuentra la teoría correcta para cualquier fenómeno. Una receta es ésta: grabe en cinta magnetofónica a algunos sujetos humanos que están pensando en voz alta mientras solucionan un problema (haciendo observaciones sobre los fenómenos); intente escribir un programa de computador que piense que simulará los protocolos humanos (formule algunas ecuaciones diferenciales); realice el programa en un computador, y determine qué ruta de conducta seguirá cuando se enfrente

a los mismos problemas que los sujetos humanos (integre numéricamente las ecuaciones); compare la simulación con la conducta real (compare las predicciones con los datos); modifique el programa sobre las bases de las discrepancias que se han descubierto (modifique las ecuaciones). Repita el proceso hasta que esté satisfecho con el ajuste.

Algunos investigadores han propuesto independientemente este camino general para la explicación de los procesos mentales superiores, y sus orígenes se pueden remontar al menos a Ach y la Escuela de Würzburg. En época más reciente, quizá los ejemplos más explícitos sean los encontrados en la investigación de deGroot sobre los procesos de pensamiento de los jugadores de ajedrez (2) y en *A Study of Thinking* de Bruner y sus colegas (1). Bruner utiliza el término estrategia, prestado de la teoría matemática de juegos, para lo que nosotros hemos llamado "programa". Lo que añaden el computador digital y las técnicas de programación es la maquinaria que nos da esperanza para seguir este camino, no solamente en principio y en general, sino de hecho y en detalle.

### *Simulación Computacional de un Programa*

La metodología perfilada anteriormente requiere que un computador simule la secuencia de expresiones verbales (u otra conducta simbólica, tal como apretar un botón) de un sujeto humano. Es fácil comprender cómo puede hacerse esto una vez que reconocemos que los computadores no tienen fundamentalmente nada que ver con los números. Sólo es un accidente histórico, el que percibamos a los computadores como mecanismos para manipular símbolos numéricos. De hecho, son artefactos extremadamente generales para manipular símbolos de cualquier tipo; y los procesos elementales necesarios para simular el pensamiento humano podrían realizarse por medio de un computador que no tuviese ninguna capacidad especial para la rapidez aritmética —que no pudiese más que realizar un recuento simple. Los programas que describiremos no hacen un uso esencial de los procesos aritméticos del computador.

¿Qué procesos puede realizar un computador de propósito general? Algunos procesos decisivos son estos:

1. Puede leer un símbolo —transformar un símbolo presentado a un mecanismo de entrada en una representación distinta de ese símbolo en un almacén interno (transformar una configuración de orificios de una tarjeta perforada en una configuración magnética de un almacén central). La relación entre la representación externa e interna es bastante flexible, casi arbitraria, y puede alterarse por medio del programa.
2. Puede mover un símbolo —reproduce en una posición del almacén un símbolo que está introducido en otra posición del almacén, con o sin modificación en la forma de representación.
3. Puede generar un símbolo —crear y almacenar una configuración en uno de su modos de representación interna.
4. Puede comparar dos símbolos, ejecutando un paso de programa si son idénticos y uno distinto si no lo son.
5. Puede asociar dos símbolos, permitiendo acceso a un símbolo (el símbolo asociado) cuando aparece el otro.

Se pueden escribir programas que combinen estos procesos simples en procesos que son ligeramente más complejos. Por ejemplo, un computador puede ser programado para manipular una serie de símbolos como una lista, de forma que pueda realizar operaciones tales como: “Coloca este símbolo al final de esa lista”, o “Encuentra un símbolo en esta lista que sea idéntico con ese símbolo”. Tales estructuras de listas, y los procesos para operar con ellas, tienen mucho parecido con la memoria y la asociación humana.

Finalmente, se pueden construir programas aún más complejos, que permiten a un computador responder a instrucciones semejantes a: “Soluciona los problemas en la siguiente lista”, “Imprime los pasos de la prueba, dando la justificación para cada paso”, “Imprime los procesos utilizados en

cada paso del proceso de resolución de problemas: los métodos, qué es lo que se observa y atiende, qué planes hay contruidos, qué subproblemas se han creado". Cuando se ha alcanzado esta última etapa, la huella que el computador imprime mientras intenta solucionar el problema puede comprobarse, línea por línea, con el registro grabado del protocolo humano de pensamiento en voz alta. Si el flujo de palabras producido por los dos procesos es casi el mismo, entonces el programa de computador que produjo la huella es una explicación del proceso de pensamiento del sujeto humano, en el estricto sentido de la palabra.

### *Verificación de Programas como Teorías*

La frase "casi el mismo" hace referencia a todo el problema de la bondad de ajuste. Desgraciadamente, la teoría estadística existente no ofrece ninguna solución al problema en la situación que hemos descrito, ni puede proponer ninguna respuesta simple. Una respuesta tosca y fácil sería que la evidencia proporcionada mediante cinco minutos de pensamiento en voz alta y la huella correspondiente sea tan amplia, que parezca indudable que se ha conseguido una primera aproximación. Esta es una respuesta subjetiva, y queremos discutir otra, ligeramente más objetiva, aunque más débil:

### *"Prueba de Turing"*

Dos sujetos humanos que solucionen el mismo problema nunca tendrán el mismo programa o producirán el mismo protocolo. Por tanto, cualquier programa aislado sólo puede ser una teoría precisa de la conducta de un sujeto único. Sin embargo, deben existir estrechas semejanzas cualitativas entre los programas y protocolos de clases de sujetos convenientemente seleccionados —caso contrario, no tiene sentido hablar de una teoría de la resolución de problemas humanos. Supongamos que combinamos diez huellas de programas de computador y diez protocolos humanos en una urna. Supongamos que un observador adecuadamente cualificado es inca-

paz de separar, con más éxito que el azar, los protocolos producidos por los programas de computador de los producidos por los humanos. Entonces diremos que los programas que produjeron las huellas de computador superan la prueba de Turing<sup>2</sup> y proporcionan una explicación satisfactoria de los protocolos humanos.

La prueba de Turing puede aplicarse en una forma más fuerte o más débil. La comparación del movimiento seleccionado por un programa de ajedrez con los movimientos seleccionados por jugadores humanos en la misma posición, sería una prueba débil. El programa puede haber seleccionado su movimiento por un proceso bastante distinto al utilizado por los humanos. Porque las circunstancias mismas de la tarea definen qué conductas son apropiadas, y se puede esperar que cualquier mecanismo capaz de conducirse adaptativamente en el medio ambiente manifieste la misma conducta externa. La semejanza de la función no garantiza la semejanza de la estructura o del proceso.

Sin embargo, si los datos se obtienen por la técnica de pensamiento en voz alta o por otros medios que indiquen los procesos utilizados para seleccionar la conducta, normalmente deberá ser posible distinguir caminos distintos que conducen al resultado. Si el programa realiza el mismo análisis que los humanos, observa las mismas características del tablero, domina las mismas artimañas, entonces deduciremos, y adecuadamente, que a cierto nivel de análisis, el programa proporciona una explicación de los procesos humanos. Cuanto más minuciosa y detallada sea la comparación entre el programa y la conducta, mayores serán las

<sup>2</sup> Una prueba de este tipo la propuso por primera vez A. M. Turing, en una discusión sobre si una máquina podía pensar (8). Dados dos canales de comunicación (digamos dos teletipos), uno conectado a un humano, el otro a una máquina, un interrogador humano tendría que identificar qué canal pertenecía a la máquina. Se permitió un interrogatorio ágil, y el problema de la máquina era engañar al interrogador, a pesar de los máximos esfuerzos del humano en el otro canal (suponiéndose que estaría de parte del interrogador) para revelar su identidad.

oportunidades para detectar diferencias entre las conductas predichas y las reales.

Este método de construcción y comprobación de la teoría, se enfrenta al problema de la inducción ni mejor ni peor que otros métodos. Nunca existe, y nunca podrá existir, una garantía de que alguna otra teoría no explicará los datos igualmente bien o mejor. Como en otras ciencias, llegará el momento de enfrentarse a este problema cuando alguien proponga realmente una teoría alternativa que explique los datos igualmente bien y con un detalle comparable. Mientras tanto, la validez de los programas como teorías puede comprobarse de una forma más y más fuerte, presionando el nivel de detalle de equiparación con el nivel de procesos de información elemental.

### *El General Problem Solver*

Para dar materia a estas generalidades, examinaremos el General Problem Solver (4,6). El General Problem Solver fue diseñado para simular la conducta de algunos sujetos humanos específicos al solucionar problemas de lógica simbólica, en una situación de trabajo proyectada por O. K. Moore y Scarvia Anderson (3). Se han registrado en el laboratorio de Carnegie Tech unos 30 protocolos de pensamiento en voz alta. La comparación de éstos con datos obtenidos por Moore en 64 sujetos sin exigirles que pensasen en voz alta, indica que no existe ninguna diferencia sustancial en el proceso bajo las dos condiciones.<sup>3</sup>

El GPS, como llamaremos al programa, se llama "General" porque no está limitado a la tarea para la que fue originalmente diseñado. La simulación manual indica que también puede solucionar los problemas lógicos de Whitehead y Russell, hacer adecuaciones trigonométricas, realizar integración y diferenciación formal y, con una pequeña ampliación del programa, solucionar ecuaciones algebraicas. Como

<sup>3</sup> El Profesor Moore nos proporcionó amablemente los datos completos sobre sus sujetos antes de publicarlos.

veremos, existen razones para esperar que pueda extenderse a un más amplio rango de tareas.

Antes de comparar minuciosamente el GPS con la conducta humana, queremos mostrar cómo soluciona problemas. Por tanto, su programa constituye un sistema de mecanismos, contruidos a partir de procesos de información elemental, que es un sistema suficiente para realizar ciertas tareas que realizan los humanos. Por mucho que se demuestre la necesidad de modificar los detalles del programa para una ajustada simulación humana, en su forma actual, constituye una demostración inequívoca de que un mecanismo puede solucionar problemas por razonamiento funcional.

En términos más simples, el GPS es un programa para razonar sobre medios y fines. Nuestras observaciones hicieron notar que los protocolos de los sujetos de laboratorio contenían muchos enunciados del siguiente tipo: "Tengo anotaciones de todas estas cosas como para saber qué es lo que tengo que hacer con ellas" (Paráfrasis: "Existen aquí algunos medios a mi disposición, ¿para qué fines pueden servir?"). "Estoy pensando en la idea de invertir ahora estas dos cosas ...de esta forma, tendría un grupo similar al del comienzo" (Paráfrasis: "Si uso X como medio, conseguiré Y"). "Ahora estoy buscando la forma para eliminar ese símbolo" ("¿Cuál es el medio de conseguir este fin?"). "Y ahora utilizaré la Regla 1" ("Aplicaré X como medio").

Un examen más minucioso de los protocolos revela que la gran mayoría de los enunciados que existen en ellos, están dentro de este marco general. La simulación de la conducta de estos sujetos requiere un programa que pueda manipular problemas de este tipo de lenguaje funcional. Además, el lenguaje funcional no hace referencia a la materia específica del problema —en este caso, lógica simbólica—. El programa debe estar organizado de forma que sus procedimientos de solución general de problemas estén diferenciados de la aplicación de éstos a una tarea específica. El GPS es un programa de este tipo.

El adjetivo "general" no implica que el GPS pueda razonar sobre todos o la mayoría de los distintos tipos de problemas; o que pueda simular toda o la mayor parte de la



actividad de resolución de problemas humanos. Simplemente significa que el programa no hace ninguna referencia al contenido de la tarea, y por tanto se puede utilizar en tareas distintas para la que fue diseñado.

El GPS opera sobre problemas que pueden formularse en términos de objetos y operadores. Un operador es algo que puede aplicarse a ciertos objetos para producir objetos distintos (como una sierra aplicada a un tronco produce tablas). Los objetos pueden describirse por medio de las características que poseen (las tablas tienen lados planos y paralelos), y por medio de diferencia, entre pares de objetos (una tabla de  $2 \times 4$  pulgadas es más gruesa que una tabla de  $1 \times 4$  pulgadas). Los operadores pueden estar limitados en su aplicación a cierto tipo de objetos (los clavos se utilizan para la madera, no para el acero); y pueden existir operadores aplicados a varios objetos como entradas (inputs), que producen uno o más objetos como salidas (outputs)—la unión de cuatro tablas para producir una estructura—.

Pueden formularse diversos problemas dentro de una tarea, cuyas circunstancias impliquen objetos y operadores: descubrir cómo transformar un objeto en otro; encontrar un objeto con características específicas; modificar un objeto de forma que se le pueda aplicar un operador específico; y así sucesivamente. En el ambiente de la tarea a la que enfrentamos a nuestros sujetos de laboratorio, los objetos eran expresiones lógico-simbólicas (a los sujetos se les dijo que eran mensajes en código). Los operadores eran 12 reglas de lógica para transformar una o dos expresiones de entrada en una expresión de salida. La Figura 1 muestra estas reglas. Por ejemplo, mediante la Regla 1 se puede transformar una expresión de la forma  $(A \cdot B)$  en  $(B \cdot A)$ , invirtiendo así el orden de los símbolos. Los problemas que se plantearon a los sujetos consistían en “recodificar”, aplicando las reglas, una o más expresiones lógicas dadas en una expresión lógica distinta. Un problema, por ejemplo, consistió en transformar  $R \cdot (\neg P \supset Q)$  en  $(Q \vee P) \cdot R$ .

Un enunciado como: “ahora estoy buscando una forma para eliminar esa herradura”, expresa una meta. La meta en este ejemplo consiste en eliminar una diferencia (la “herra-

dura" de la expresión original en contra de la "cuña" en la expresión deseada) entre un objeto y otro. Las metas que los sujetos mencionan en sus protocolos tienen formas muy distintas. Hemos incorporado tres tipos de metas que explican la gran mayoría de enunciados de meta en la actual versión del GPS. Son las siguientes:

*Meta Tipo No. 1.* Descubrir una manera de transformar el objeto  $a$  en el objeto  $b$  (i.e., una secuencia de operadores para conseguir la transformación). Los objetos se forman construyendo expresiones a partir de letras ( $P, Q, R, \dots$ ) y conectivas  $\cdot$  (punto),  $\vee$  (cuña),  $\supset$  (herradura) y  $\neg$  (tilde). Ejemplos son  $P$ ,  $\neg Q$ ,  $P \vee Q$ ,  $\neg(R \supset S) \cdot \neg P$ ;  $\neg\neg P$  es equivalente a  $P$  en todos los casos.

Existen doce reglas para transformar las expresiones (donde  $A$ ,  $B$  y  $C$  pueden ser cualquier expresión o sub-expresión):

- |      |   |  |
|------|---|--|
| R1.  | $A \cdot B \rightarrow B \cdot A$<br>$A \vee B \rightarrow B \vee A$  |  |
| R2.  | $A \supset B \rightarrow \neg B \supset \neg A$   |  |
| R3.  | $A \cdot A \Leftrightarrow A$<br>$A \vee A \Leftrightarrow A$   |  |
| R4.  | $A \cdot (B \cdot C) \Leftrightarrow (A \cdot B) \cdot C$<br>$A \vee (B \vee C) \Leftrightarrow (A \vee B) \vee C$                    |  |
| R5.  | $A \vee B \Leftrightarrow \neg(\neg A \cdot \neg B)$  |  |
| R6.  | $A \supset B \Leftrightarrow \neg A \vee B$   |  |
| R7.  | $A \cdot (B \vee C) \Leftrightarrow (A \cdot B) \vee (A \cdot C)$<br>$A \vee (B \cdot C) \Leftrightarrow (A \vee B) \cdot (A \vee C)$ |  |
| R8.  | $A \cdot B \rightarrow A$<br>$A \cdot B \rightarrow B$  | Sólo se aplica a expresiones principales.                      |
| R9.  | $A \rightarrow A \vee X$  | Sólo se aplica a expresiones principales.                      |
| R10. | $\left. \begin{array}{l} A \\ B \end{array} \right\} \rightarrow A \cdot B$   | $A$ y $B$ son dos expresiones principales.                     |
| R11. | $\left. \begin{array}{l} A \\ A \supset B \end{array} \right\} \rightarrow B$   | $A$ y $A \supset B$ son dos expresiones principales.           |
| R12. | $\left. \begin{array}{l} A \supset B \\ B \supset C \end{array} \right\} \rightarrow A \supset C$                                     | $A \supset B$ y $B \supset C$ son dos expresiones principales. |

Figura 1. Reglas para Transformar Expresiones Lógicas

*Meta Tipo No. 2.* Aplicar el operador  $q$  al objeto  $a$  (o a un objeto obtenido a partir de  $a$  mediante transformaciones).

*Meta Tipo No. 3.* Reducir la diferencia,  $d$ , entre el objeto  $a$  y el objeto  $b$ , modificando  $a$ .

Inicialmente, los problemas planteados a los sujetos establecían metas transformadas; “eliminar la herradura” expresa una meta de reducción; “Y ahora utilizo la Regla 1” establece una meta de aplicación.

Para alcanzar una meta, la consideración de esa meta debe evocar en el que soluciona el problema alguna idea de uno o más medios que pueden ser importantes. El sujeto, por ejemplo, que dice “estoy buscando la forma de eliminar esa herradura”, continúa este enunciado con, “¡Ah!..., aquí está la Regla 6. Aplicaré la Regla 6 a la segunda parte de lo que tenemos aquí”. La aplicación de la Regla 6 (ver Figura 1) surge como un método para eliminar la herradura.

De esta forma el proceso de evocación está representado en GPS por medio de asociaciones entre cada tipo de meta y uno o más métodos para obtener una meta de ese tipo. Estos se muestran en la Figura 2. El método No. 1, asociado con metas de Transformación, está compuesto de: (a) comparar los objetos  $a$  y  $b$  para encontrar una diferencia,  $d$ , entre ellos; (b) establecer la submeta de reducción  $d$  del Tipo No. 3, que si tiene éxito, produce un nuevo objeto transformado,  $c$ ; (c) establecer la submeta del Tipo No. 1 para transformar  $c$  en  $b$ . Si se alcanza esta última meta, se consigue la meta de Transformación original. El emparejamiento del paso (a) se realiza para la primera diferencia más importante (en términos de alguna lista de prioridades).

El método No. 2, para alcanzar una meta de Aplicación, consiste en: (a) determinar si el operador puede aplicarse para establecer una meta del tipo No. 1, para transformar  $a$  en la forma de entrada del operador  $q$  (que llamamos  $C(q)$ ); (b) si tiene éxito, se produce el objeto de salida desde la forma de salida de  $q$  ( $P(q)$ ).

El método No. 3, para alcanzar una meta de Reducción, consiste en: (a) buscar un operador que sea importante para

reducir la diferencia,  $\bar{d}$ ; (b) si se encuentra uno, establecer la meta Tipo No. 2 de aplicación de operador que, si se consigue, produce el objeto modificado.

Para ver cómo GPS va aplicando estas metas-tipo y estos métodos a la resolución de problemas, consideremos un ejemplo concreto. Utilizaremos el problema mencionado anteriormente: "recodificar" la expresión,  $L_1, R \cdot (\neg P \supset Q)$  en la expresión  $L_0, (Q \vee P) \cdot R$ . Explicaremos sobre la marcha las reglas de recodificación en la medida en que sea necesario para comprender el ejemplo.

GPS comienza estableciendo la meta Tipo No. 1 para transformar  $L_1$  en  $L_0$ . Entre los tipos de información utilizables —construidos a partir de procesos elementales—, existen algunas pruebas para las posibles diferencias entre pares de expresiones, por ejemplo:

1. Una prueba para ver si aparecen las mismas o distintas variables (letras) en las dos expresiones;
2. Una prueba para ver si cada variable aparece el mismo o distinto número de veces;
3. Una prueba para ver si una variable o grupo aparece en la misma o distinta posición;
4. Una prueba para ver si un par de conectivas ( $\cdot$ ,  $\vee$ ,  $\supset$ ) son iguales o distintas.

Y así sucesivamente. Las pruebas pueden aplicarse a la totalidad de las expresiones o a las partes correspondientes de las expresiones (por ejemplo,  $(\neg P \supset Q)$  y  $(Q \vee P)$ ).

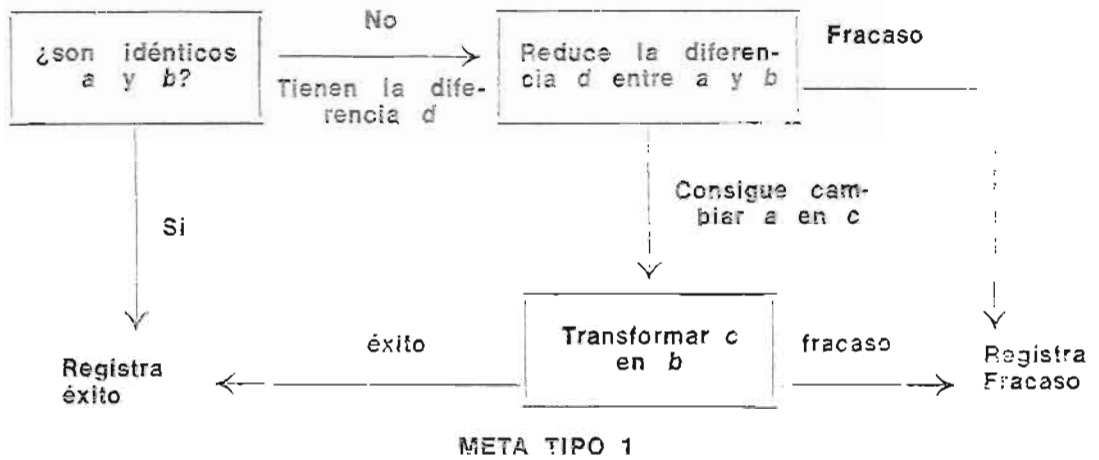
El método No. 1 aplica estas pruebas por orden. Descubre una diferencia en la posición de las  $R$  en  $L_1$  y  $L_0$ , y establece la meta Tipo No. 3 para reducir esta diferencia. Puede utilizar una lista de operadores para cada diferencia, que posiblemente son importantes para eliminar una diferencia de ese tipo. (Estas listas puede construirlas el mismo GPS, examinando el conjunto de operadores utilizables). En este caso, descubre que la Regla 1, que transforma una expresión del tipo  $(A \cdot B)$  o  $(A \vee B)$  en una expresión del tipo  $(B \cdot A)$  o  $(B \vee A)$ , respectivamente, influye en las diferencias de posición. En consecuencia, establece la submeta Tipo No. 2 para

aplicar la Regla 1 a  $L1$ . Esto se puede hacer identificando en la regla  $R$  con  $A$  y  $(\neg P \supset Q)$  con  $B$ , produciendo así  $L2$ ,  $(\neg P \supset Q) \cdot R$  como expresión de salida.

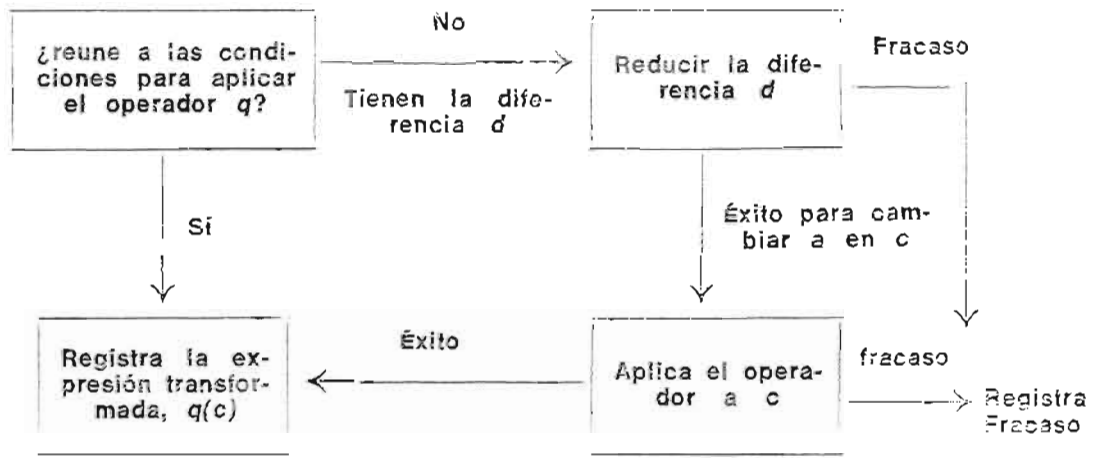
Hecho esto, la meta original establece ahora la nueva submeta para transformar  $L2$  en  $L0$ . Repitiendo el ciclo, se encuentra una diferencia en las conectivas de los lados izquierdos de  $L2$  y  $L0$ , respectivamente; encuentra una regla,  $Rc$ , que cambia las conectivas transformando  $(\neg A \supset B)$  en  $(A \vee B)$ , y por tanto que transforma  $(\neg P \supset Q) \cdot R$  en  $L3$   $(P \vee Q) \cdot R$ .

Una tercera repetición del ciclo básico, descubre la diferencia de posición entre  $(P \vee Q)$  en  $L3$  y  $(Q \vee P)$  en  $L0$ , y aplica  $R1$  para eliminar la diferencia. Finalmente,  $GPS$  descubre que el producto de esta transformación es idéntico a  $L0$ , y declara solucionado el problema. Podemos resumir los pasos de la manera siguiente:

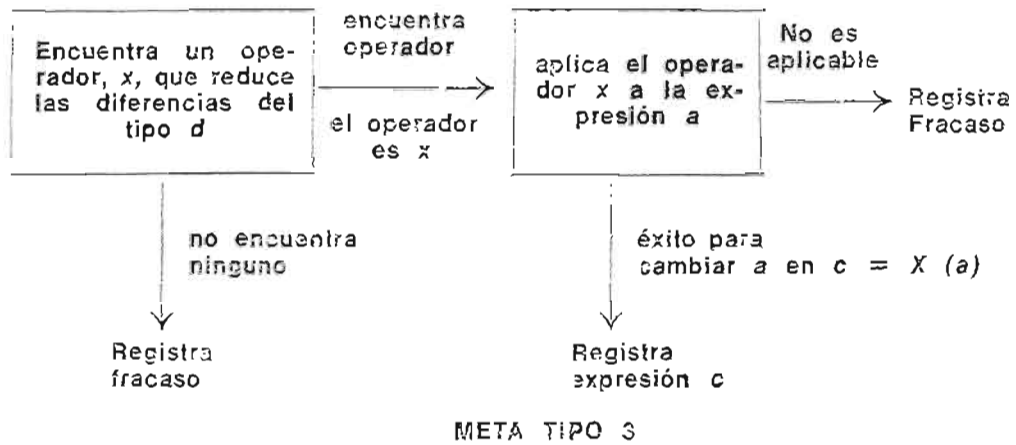
<i>Paso</i>	<i>Expresión</i>	<i>Justificación del paso</i>
$L1$	$R \cdot (\neg P \supset Q)$	Dado
$L2$	$(\neg P \supset Q) \cdot R$	Regla 1
$L3$	$(P \vee Q) \cdot R$	Regla 6 dentro del paréntesis
$L0$	$(Q \vee P) \cdot R$	Regla 1 dentro del paréntesis



META TIPO 1  
Transformar a en b



META TIPO 2  
Aplicar el operador q a la expresión a



META TIPO 3  
Reduce la diferencia d entre a y b

Figura 2. Métodos asociados con las Metas GPS

*Una Comparación con la Conducta Humana*

Una vez concedido que GPS puede solucionar este problema, y otros muchos bastante más difíciles, ¿por qué no suponer que los procesos de GPS se parecen de alguna forma a los procesos que un humano utilizaría para solucionar el mismo problema? Comparemos los procesos del GPS, tal como los hemos descrito, con el contenido del protocolo de un sujeto humano que soluciona el mismo problema. Dejaremos que el lector juzgue si los dos procesos son o no lo suficientemente parecidos. (Ni el problema particular ni el protocolo humano que examinaremos se utilizaron al diseñar el GPS).

<i>Líneas</i>	<i>Simulación</i>	<i>Protocolo</i>
1	$L_0: (Q \vee P) \cdot R$	(Expresión que debe obtenerse) <sup>a</sup>
2	$L_1: R \cdot (\neg P \supset Q)$	(Expresión inicial)
3	Meta 0: Transformar $L_1$ en $L_0$	(Meta establecida por el experimentador)
4	Comparación de diferencias de posición ( $\Delta p$ )	Estoy pensando en la idea de invertir ahora estas dos cosas
5	Meta 1: Reducir $\Delta p$ entre $L_1$ y $L_0$	(¿Pensando en la inversión de qué?)
6	Busca la lista de reglas	Las $R...$
7	Meta 2: Aplica $R_1$ a $L_1$	
8	Comparación: $R_1$ aplicable	
9	Comprueba las funciones de la regla: reduce $\Delta p$	Así tendría un grupo similar al del comienzo
10	Ninguna otra	Pero que parece que es...
11	Decide ejecutar $R_1$ cuando se complete el análisis	Podría dejarlo para el final,

<i>Líneas</i>	<i>Simulación</i>	<i>Protocolo</i>
12	Meta 3: transformar la derecha de $L1$ en la izquierda de $L0$	omitiéndolo, entonces, tendría...
13	La comparación da una diferencia de posición ( $\Delta p$ )	
14	Meta 4: Reducir $\Delta p$ entre la derecha de $L1$ y la izquierda de $L0$	
15	Busca la lista de reglas	
16	Meta 5: aplicar $R1$ a la derecha de $L1$	
17	Comparación: $R1$ fracasa, la derecha de $L1$ tiene $\supset (\Delta c)$	(¿Aplicación de qué regla?)
18.	Meta 6: aplicar $R2$ a la derecha de $L1$	Aplicando... por ejemplo, 2.
19.	Comparación: $R2$ es aplicable	
20	Comprobación de las funciones de la regla: reduce $\Delta p$ ,	
21	pero introduce —no deseado ( $\Delta s$ )	Necesitaría un cambio de signo
22	Meta rechazada	(Intente seguir hablando, si puede)
23.	Meta 7: Aplicar $R3$ a la derecha de $L1$	Bien... miraré la Regla 3
24	Comparación: $R3$ no es aplicable	y no la encuentro muy práctica
25	Meta 8: Aplicar $R4$ a $L1$	Ahora me intereso por la 4



<i>Líneas</i>	<i>Simulación</i>	<i>Protocolo</i>
26	Comparación: R4 no es aplicable <sup>b</sup>	Existen tres partes similares... y... existen puntos como conectivas... parece que funciona con bastante facilidad,
27		
28	Comprobación de las funciones de la regla: no reduce $\Delta p$	pero no cambia el orden.
29	Meta rechazada	
30	Meta 9 a 13: Aplicar desde R5 hasta R9 a L1	
31	Todas las metas fracasan en la comparación	
32	o prueba de la función	
33	Nueva búsqueda de reglas, pero	Necesito cambiar <i>P</i> y <i>Q</i> de forma que,...
34	sin intentar ningún sub-problema	
35	Meta 14: Aplicar R1 a la derecha de L1	
36	Comparación: R1 fracasa, la derecha de L1 tiene $\supset (\Delta c)$	Aquí hay una herradura. Eso no parece
37	Comprobación de las funciones de la regla: reduce $\Delta p$ , nada más	práctico en ningún lugar de aquí
38	Decide ejecutar R1, si es aplicable	
39	Meta 15: Reduce $\Delta c$ entre la derecha de L1 y R1	Ahora busco la forma para eliminar esa herradura
40	Busca la lista de reglas,	
41	la regla con $\supset$ que reduce $\Delta c$	
42	Meta 16: Aplicar R6 a la derecha de L1	Ah... aquí está, la Regla 6

<i>Líneas</i>	<i>Simulación</i>	<i>Protocolo</i>
43	Comparación: $R6$ es aplicable	
44	Comprobación de las funciones de la regla: reduce $\Delta c$	
45	Reduce $\Delta s$	
46	Decide ejecutar $R6$ cuando se complete el análisis	
47	Meta 17: transformar la derecha	
48	de $L1$ después de reducir $\Delta p$ , $\Delta c$ , $\Delta s$ , en la izquierda de $L0$	
49	Comparación: la transformación	
50	de la derecha de $L1$ es idéntica a la izquierda de $L0$	
51	Meta alcanzada, y análisis terminado	
52	Ejecutar $R6$ en la derecha de $L1$ (de la línea 46)	Así, aplicaré la Regla 6 a la segunda parte de lo que
53		tengo aquí
54		(¿Desea hacer eso?)
55		Sí
56	$L2: R \cdot (P \vee Q)$	(OK, aplica a la línea 1 la $R6$ . La línea 2 es $R \cdot (P \vee Q)$ )
57	Ejecuta $R1$ a la derecha de $L2$ (de la línea 38)	Y ahora utilizo la Regla 1.
58		(La Regla 1, ¿en qué parte? Puede usarla sobre la expresión entera o sobre la parte derecha.)

<i>Líneas</i>	<i>Simulación</i>	<i>Protocolo</i>
59		La uso en ambos sitios
60		(Bien, lo hará una después de otra... ¿Cuál sera primero?)
62		Bueno, con $P$ y $Q$
63		
64	$L3: R \cdot (Q \vee P)$	$[R \cdot (Q \vee P)$ . ¿Ahora en la expresión completa?]
65	Ejecuta $R1$ en $L3$ (de la línea 11)	Sí
66	$L4: (Q \vee P) \cdot R$	(En la línea 3. Regla 1... obtiene $(Q \vee P) \cdot R$ )
67	Comparación: $L4$ es idéntica a $L0$	Y... ya está. (Está correcto, OK... no fue demasiado difícil.)

a) Los enunciados entre paréntesis son enunciados del experimentador y comentarios explicativos. Todos los demás enunciados son del sujeto.

b) Pero el sujeto piensa erróneamente que  $R4$  se puede aplicar; por tanto, comprueba sus funciones

Figura 3. Comparación de la huella con el protocolo humano

Reproducimos en la parte derecha de la Figura 3, palabra por palabra, el protocolo humano, omitiendo sólo algunos párrafos introductorios. En la parte izquierda, reproducimos la huella (simulada manualmente) de un programa que, según creemos, se aproxima bastante a los procesos del pensamiento del sujeto. Este programa no es idéntico al descrito por el GPS, puesto que incorpora algunas modificaciones para ajustarlo a los datos empíricos. Sin embargo, los objetos, los operadores, diferencias, tipos de meta y métodos que aparecen en esta huella son básicamente los de GPS.

Si examinamos el protocolo, vemos que el sujeto establece la meta para transformar  $L_1$  en  $L_0$ , observa las dife-

rencias en el orden de los términos de las dos expresiones y piensa en su inversión (líneas 4-9). Al mismo tiempo, el programa establece la misma meta (Meta O), observa la misma diferencia (línea 4), descubre que la Regla 1 reduce esta diferencia (línea 9) y señala la idea de aplicar la Regla 1 cuando el análisis esté completo. (Esta distinción entre acción manifiesta y encubierta es una de las modificaciones introducidas.)

A continuación, el sujeto examina la lista de reglas —existe evidencia explícita de que mira las Reglas 2, 3 y 4— rechazando cada una por no ser aplicable (Regla 3), porque introduciría una nueva diferencia (“Necesitaría un cambio de signo) o porque no realiza la función de cambiar la *P* por la *Q* (Regla 4). Al mismo tiempo, el programa establece la meta de reducir la diferencia en el orden de *P* y *Q* (Meta 4), examinando la lista de reglas y rechazándolas por las mismas razones.

Después, el sujeto observa que la herradura origina dificultades para cambiar la *P* y la *Q* (líneas 36-37) y se propone la meta de eliminar la herradura, descubriendo que la Regla 6 realizará esto (línea 42). Aplica la Regla 6, después la Regla 1 a la sub-expresión de la derecha y luego la Regla 1 a la expresión total, y observa que tiene solucionado el problema. Simultáneamente, el programa emprende una segunda búsqueda de una regla que invierta la *P* y la *Q* (líneas 33-34), pero ahora lo intenta como un subproblema obteniendo otra regla importante que debe aplicar. Considera la Regla 1, establece la submeta (Meta 15) de cambiar la herradura por una cuña, encuentra la Regla 6, comprueba si esto soluciona el problema, después ejecuta las Reglas 6, 1 y 1 (sobre la expresión total), en ese orden.

### *Una Segunda Comparación con la Conducta Humana*

Consideremos ahora un problema algo más complicado, analizando con más detalle la correspondencia entre la huella del GPS y el protocolo del sujeto. Se le pide a un estudiante de ingeniería que transforme la expresión

$$(R \supset \neg P) \cdot (\neg R \supset Q)$$

en la expresión  $\neg(\neg Q \cdot P)$ . El sujeto ha practicado aplicando las reglas, pero sólo ha hecho previamente un problema semejante a éste. Su secuencia de razonamientos sobre el problema como un todo y la sección inicial de su protocolo, se muestran en la figura 4.

Desarrollo de la solución:

1. $(R \supset \neg P) \cdot (\neg R \supset Q)$	$\neg(\neg Q \cdot P)$
2. $(\neg R \vee \neg P) \cdot (R \vee Q)$	Aplicación de la Regla 6 a la izquierda y a la derecha de 1
3. $(\neg R \vee \neg P) \cdot (\neg R \supset Q)$	Aplicación de la Regla 6 a la izquierda de 1
4. $R \supset \neg P$	Aplicación de la Regla 8 a 1
5. $\neg R \vee \neg P$	Aplicación de la Regla 6 a 4
6. $\neg R \supset Q$	Aplicación de la Regla 8 a 1
7. $R \vee Q$	Aplicación de la Regla 6 a 6
8. $(\neg R \vee \neg P) \cdot (R \vee Q)$	Aplicación de la Regla 10 a 5 y a 7
9. $P \supset \neg R$	Aplicación de la Regla 2 a 4
10. $\neg Q \supset R$	Aplicación de la Regla 2 a 6
11. $P \supset Q$	Aplicación de la Regla 12 a 6 y 9
12. $\neg P \vee Q$	Aplicación de la Regla 6 a 11
13. $\neg(P \cdot \neg Q)$	Aplicación de la Regla 5 a 12
14. $\neg(\neg Q \cdot P)$	Aplicación de la Regla 1 a 13 QED.

Protocolo de la primera parte del problema:

Bueno, mirando la parte izquierda de la ecuación, deseamos en primer lugar eliminar uno de los lados utilizando la Regla 8. Parece demasiado complicado trabajar sobre esto. Ahora — no, — no, no puedo hacer eso puesto que eliminaría la  $Q$  o la  $P$  en la expresión total. No quiero empezar por ahí. Ahora busco la manera de eliminar la herradura interna de los paréntesis que aparece a la izquierda y a la derecha de la ecuación. Y no lo veo. Si, si se aplica la Regla 6 a ambos

lados de la ecuación; a partir de aquí, veré si puedo aplicar la Regla 7.

El experimentador escribe: 2.  $(\neg R \vee \neg P) \cdot (R \vee Q)$

Casi puedo aplicar la Regla 7, pero una  $R$  necesita una tilde. Tendré que buscar otra regla. Estoy viendo si puedo cambiar esa  $R$  por una  $R$  con tilde. De hecho, debería haber utilizado la Regla 6 sólo sobre el lado izquierdo de la ecuación. Así, uso la Regla 6, pero sólo en el lado izquierdo.

El experimentador escribe: 3.  $(\neg R \vee \neg P) \cdot (\neg R \supset Q)$

Ahora aplicaré la Regla 7 tal como está expresada. Ambos —perdón, perdón, no puedo hacerlo a causa de la herradura. Así —ahora lo estoy mirando— examino por un segundo las reglas, y miro si puedo cambiar la  $R$  por una  $\neg R$  en la segunda ecuación, pero no veo la forma de hacerlo (*suspira*). Me encuentro algo perdido en este momento.

Figura 4. Sujeto en el Problema D1

La primera porción de la huella de GPS se muestra en la figura 5. El problema inicial de GPS es transformar  $L1$  en  $L0$ . La comparación de  $L1$  con  $L0$  revela que existen  $R$ 's en  $L1$  y que no existen  $R$ 's en  $L0$ . Esta diferencia conduce a la formulación de una meta de reducción, que para facilitar su comprensión recibe el nombre funcional de *Suprimir*. El esfuerzo para alcanzar esta meta conduce a una búsqueda de reglas que desemboca en la Regla 8. Puesto que existen dos formas de la Regla 8, siendo ambas admisibles, GPS elige la primera. (No se indican las variantes de las reglas, pero pueden inferirse fácilmente a partir de la huella). Puesto que se puede aplicar la Regla 8, se produce un nuevo objeto,  $L2$ . Continuando con el método de transformar metas, se genera en el siguiente paso una nueva meta: transformar  $L2$  en  $L0$ . Esto conduce a su vez a otra meta de reducción: reintegrar  $Q$  a  $L2$ . Pero se rechaza esta meta por medio de

una evaluación, puesto que añadir un término es más difícil que suprimir un término. Por tanto, GPS regresa a la Meta 2 y busca otra regla que suprima términos. Esta vez se encuentra la otra forma de la Regla 8 y realiza un desarrollo similar, finalizando con el rechazo definitivo de la Meta 8.

$$L_0 \text{ --- } (\text{--- } Q \cdot P)$$

$$L_1 (R \supset \text{--- } P) \cdot (\text{--- } R \supset Q)$$

- Meta 1. Transformar  $L_1$  en  $L_0$
- Meta 2. Suprimir  $R$  de  $L_1$
- Meta 3. Aplicar  $R_8$  a  $L_1$
- Produce  $L_2 R \supset \text{--- } P$

- Meta 4. Transformar  $L_2$  en  $L_0$
- Meta 5. Añadir  $Q$  a  $L_2$
- Rechazo

- Meta 2
- Meta 6. Aplicar  $R_8$  a  $L_1$
- Produce  $L_3 \text{--- } R \supset Q$

- Meta 7. Transformar  $L_3$  en  $L_0$
- Meta 8. Añadir  $P$  a  $L_3$
- Rechazo

- Meta 2
- Meta 9. Aplicar  $R_7$  a  $L_1$
- Meta 10. Cambiar conectiva a  $\vee$  a la izquierda de  $L_1$
- Meta 11. Aplicar  $R_6$  a la izquierda de  $L_1$
- Produce  $L_4 (\text{--- } R \vee \text{--- } P) \cdot (\text{--- } R \supset Q)$

- Meta 12. Aplicar  $R_7$  a  $L_4$
- Meta 13. Cambiar conectiva a  $\vee$  a la derecha de  $L_4$
- Meta 14. Aplicar  $R_6$  a la derecha de  $L_4$
- Produce  $L_5 (\text{--- } R \vee \text{--- } P) \cdot (R \vee Q)$

- Meta 15. Aplicar  $R_7$  a  $L_5$
- Meta 16. Cambiar el signo de la izquierda a la derecha de  $L_5$
- Meta 17. Aplicar  $R_6$  a la derecha de  $L_5$
- Produce  $L_6 (\text{--- } R \vee \text{--- } P) \cdot (\text{--- } R \supset Q)$

Meta 18. Aplicar  $R7$  a  $L6$

Meta 19. Cambiar conectiva a  $\vee$  a la derecha de  $L6$

Rechazo

Meta 16

Nada más

Meta 13. Nada más

Meta 10. Nada más

*Figura 5.* Huella de GPS en la 1.<sup>a</sup> Parte del Problema D.

Regresando de nuevo a la Meta 2 para encontrar otra regla que elimine términos, GPS obtiene la Regla 7. Se selecciona la variante  $(A \vee B) \cdot (A \vee C) \rightarrow A \vee (B \cdot C)$ , puesto que sólo ésta disminuye términos y tiene un punto como conectiva principal. La Regla 7 no se aplica inmediatamente; primero, GPS descubre que existe una diferencia de conectiva en la expresión izquierda, y después que existe otra en la expresión derecha. En ambos casos, se encuentra y se aplica la Regla 6 para cambiar la conectiva de herradura a cuña, obteniendo sucesivamente  $L4$  y  $L5$ . Pero la nueva expresión revela una diferencia de signo, que conduce de nuevo a la Regla 6, esto es, a la misma regla que antes, pero cumpliendo una función diferente. La Regla 6 produce  $L6$ , que es idéntica a  $L4$  aunque GPS no observa aquí la identidad. Esto conduce en la Meta 19, a detectar de nuevo la diferencia de conectiva; con lo cual la meta es finalmente rechazada, puesto que no representa ningún progreso sobre la Meta 13. Intentos posteriores para encontrar formas alternativas para cambiar signos o conectivas fracasan en la producción de cualquier cosa. Esto finaliza el episodio.

Tenemos ahora una huella muy detallada de lo que hizo GPS. ¿Qué podemos encontrar en el protocolo del sujeto que confirme o refute la afirmación de que este programa es un modelo detallado de las manipulaciones simbólicas que realiza el sujeto? El programa no nos proporciona una salida en lengua inglesa que pueda ponerse en correspondencia biunívoca con las palabras del sujeto. Aún no hemos



dado a GPS una meta para “realizar la tarea y charlar al mismo tiempo”, que sería una reformulación necesaria si quisiéramos intentar una correspondencia con tal detalle. Por otro lado, la huella, repasada por lo que sabemos de cómo se generó, proporciona un registro completo de todo el contenido de la tarea, que fue tomada en cuenta por GPS y el orden en fue atendida. Por tanto, esperamos encontrar cada característica del protocolo que concierne a la tarea, reflejada de una forma esencial en la huella del programa. La inversa no es verdadera, puesto que muchas cosas que conciernen a la tarea seguramente ocurren sin que el sujeto haga comentarios sobre ellas (aun siendo consciente de ellas). Así, nuestra prueba de correspondencia es unilateral, pero exigente. Comencemos con la primera sentencia del protocolo del sujeto:

Bueno, mirando la parte izquierda de la ecuación, deseamos en primer lugar eliminar uno de los lados utilizando la Regla 8.

Vemos un deseo de disminuir  $L_1$  o eliminar algo, y la selección de la Regla 8 como el medio para hacerlo. Esto está en correspondencia directa con la Meta 1, 2 y 3 de la huella. Saltemos ahora a la tercera y cuarta sentencia: “Ahora —no, —no, no puedo hacer eso puesto que eliminaría la  $Q$  o la  $P$  en la expresión total. No quiero empezar por ahí”.

Vemos aquí una expresión directa de la aplicación encubierta de la Regla 8, la subsiguiente comparación de la expresión resultante con  $L_0$ , y el rechazo de este sistema de acción a causa de que elimina una letra que se necesita en la expresión final. Sería difícil encontrar un conjunto de palabras que expresara estas ideas con mayor claridad. A la inversa, si el mecanismo del programa (o alguno esencialmente similar) no actuase, sería difícil explicar por qué el sujeto expresa los comentarios de lo que hace.

Una discrepancia aparece bastante clara. El sujeto manipuló juntas las dos formas de la Regla 8, al menos en la medida que indica su comentario. GPS, por otro lado, con-

sidera por separado cada una de las formas. Posiblemente, el sujeto realizó el programa encubiertamente y registró juntos los dos resultados. Sin embargo, el ajuste sería mejor si GPS hubiese actuado de alguna manera parecida a la siguiente:

Meta 2. Suprime  $R$  de  $L1$

Meta 3. Aplicar  $R8$  a  $L1$

Produce  $L2$   $R \supset \neg P$  o  $\neg R \supset Q$

Meta 4. Transforma  $L2$  en  $L0$

Meta 5. Añade  $Q$  a  $R \supset \neg P$  o añade  $P$  a  $\neg R \supset Q$

Rechazo

Consideraremos más evidencia sobre este punto posteriormente. Regresemos a la segunda sentencia, que saltamos anteriormente:

Parece demasiado complicado trabajar sobre esto.

Ninguna parte del programa está en correspondencia simple con este enunciado, aunque es fácil imaginar alguna explicación posible. Por ejemplo, podría ser simplemente una expresión de la comparación —del hecho de que  $L1$  es una expresión grande que el sujeto no puede absorber en todo su detalle. No existen suficientes datos para determinar qué parte de la huella correspondería a este enunciado, de forma que la sentencia debe soportarse como un elemento no explicado de la conducta del sujeto. Consideraremos ahora las siguientes sentencias del protocolo:

Ahora busco la manera de eliminar la herradura interna de los paréntesis que aparece a la izquierda y a la derecha de la ecuación. Y no lo veo. Sí, si se aplica la Regla 6 a ambos lados de la ecuación; a partir de aquí, veré si puedo aplicar la Regla 7.

Esto está en correspondencia directa con las Metas 9 hasta 14 de la huella. El comentario del final deja claro que la aplicación de la Regla 7 es el principal interés y que se

necesita cambiar las conectivas para conseguir esto. Además, el protocolo muestra claramente que selecciona la Regla 6 como un medio. Las tres reglas seleccionadas proporcionan alguna confirmación de que el sujeto hizo una prueba preliminar de posibilidad —igual que GPS— en el método de meta de reducción. Si no hubo selección en la conectiva principal, ¿por qué no se seleccionó la Regla 5 en lugar de la Regla 6? ¿O por qué no fue seleccionada la forma  $(A \cdot B) \vee (A \cdot C) \rightarrow A \cdot (B \vee C)$  de la Regla 7? Sin embargo, existe una discrepancia entre la huella y el protocolo, porque el sujeto manipula simultáneamente ambas aplicaciones de la Regla 6 (y aparentemente también manipuló simultáneamente las dos diferencias); mientras que GPS las manipula secuencialmente. Esto es parecido a la discrepancia observada anteriormente al manipular la Regla 8. Puesto que tenemos ahora dos ejemplos de procesamiento paralelo, es probable que exista una diferencia real en este resultado. De nuevo, sería mejor que GPS procediese de una manera similar a la siguiente:

Meta 9. Aplica  $R7$  a  $L1$

Meta 10. Cambia conectiva a  $\vee$  en la izquierda y en la derecha de  $L1$

Meta 11. Aplica  $R6$  a la izquierda y a la derecha de  $L1$   
Produce  $L5$   $(\text{---} R \vee \text{---} P) \cdot (R \vee Q)$

Una característica común de ambas discrepancias es que, al formar las expresiones compuestas, no complican los métodos de una manera esencial. Así, en el caso de incluir la Regla 8, los dos resultados se originan desde la misma forma de entrada y sólo necesitan una sola comparación. En el caso de incluir la Regla 7, se hizo una sola búsqueda de una regla y se aplicó simultáneamente la regla a ambas partes, como si sólo se incluyese una sola unidad.

Existen dos aspectos donde el protocolo proporciona información que el programa no es capaz de explicar. Primero, el sujeto manipuló la aplicación de la Regla 8 encubiertamente y ordenó al experimentador hacer las aplicaciones de la Regla 6 sobre la mesa. La versión utilizada aquí del

GPS no hace ninguna distinción entre acciones internas y externas. En este punto fracasa como modelo adecuado. La distinción manifiesta-encubierta tiene consecuencias que corren a lo largo del problema, puesto que las expresiones sobre el encerado tienen características de memoria muy distintas de las expresiones que sólo se generan en la cabeza. Segundo, esta versión del GPS no simula lo suficientemente bien el proceso de búsqueda como para proporcionar una correspondencia con "Y no lo veo. Sí, ..." Esto exige proporcionar un facsímil de la hoja de reglas y distinguir la búsqueda sobre la hoja de las búsquedas en la memoria. Las siguientes sentencias dicen:

Casi puedo aplicar la Regla 7, pero una *R* necesita una tilde. Tendré que buscar otra regla. Estoy viendo si puedo cambiar esa *R* por una *R* con tilde.

De nuevo la huella y el protocolo concuerdan en la diferencia que se ha visto. También coinciden en que esta diferencia no fue atendida anteriormente, aun cuando estaba presente. Alguna delicada estructura de los datos coincide también con la huella. Se toma la *R* de la derecha como la base de la diferencia (*R* por  $-R$ ) en lugar de la *R* de la izquierda, aunque podría tomarse cualquiera de las dos. Esta preferencia surge en el programa (y se supone que también en el sujeto) a partir del hábito lingüístico de trabajar de izquierda a derecha. Este hecho no carece de consecuencias, puesto que determina que el sujeto trabaje sobre el lado izquierdo o sobre el derecho de la expresión; en consecuencia, puede afectar al desarrollo completo de acontecimientos por bastante tiempo. De la misma manera, en el episodio de la Regla 8 el sujeto trabajó aparentemente de izquierda a derecha y de arriba a abajo para llegar a "*Q* o *P*" en lugar de "*P* o *Q*". Esto puede parecer un interés excesivamente detallado por las características, sin embargo, esos detalles soportan el argumento de que lo que está ocurriendo en el interior del sistema humano es bastante análogo a las manipulaciones simbólicas que ocurren en el interior de GPS. El siguiente fragmento del protocolo es:

De hecho, debería haber utilizado la Regla 6 sólo sobre el lado izquierdo de la ecuación. Así, uso la Regla 6, pero sólo en el lado izquierdo.

En esto encontramos una fuerte divergencia de la huella del GPS, aun cuando, cosa curiosa, la huella y el protocolo finalizan en el mismo punto,  $(\neg R \vee \neg P) \cdot (\neg R \supset Q)$ . Tanto el sujeto como el GPS encuentran la Regla 6 como la apropiada para cambiar signos. En este punto, GPS aplica simplemente la regla a la expresión actual; mientras que el sujeto vuelve hacia atrás y corrige la aplicación previa. No existe en el programa nada que corresponda con esto. La explicación más directa es que la aplicación de la Regla 6 en la dirección inversa, la percibe el sujeto como anulando la aplicación previa de la Regla 6. Después de seguir esta línea de razonamiento, toma entonces la alternativa más simple (y aparentemente menos disparatada), que es corregir la acción original.

El segmento final del protocolo dice:

Ahora aplicaré la Regla 7 tal como está expresada. Ambos —perdón, perdón, no puedo hacerlo a causa de la herradura. Así —ahora lo estoy mirando— examino por un segundo las reglas, y miro si puedo cambiar la  $R$  por una  $\neg R$  en la segunda ecuación, pero no veo la forma de hacerlo. (*Suspira*) Me encuentro algo perdido en este momento.

La huella y el protocolo están nuevamente en buen acuerdo. Esta es una de las pocas auto-rectificaciones que hemos encontrado. El protocolo registra la inútil búsqueda de operadores adicionales que afecten las diferencias de signo y de conectiva, siempre con resultado negativo. El comentario final de ligera desesperación puede interpretarse como un reflejo del impacto de varios fracasos sucesivos.

Repasemos los acuerdos y discordancias entre la huella y el protocolo. El programa proporciona una explicación completa de la conducta en la tarea del sujeto, con cinco excepciones de distinta gravedad. Existen dos aspectos donde el GPS no está preparado para simular la conducta del su-

jeto: en la distinción entre el mundo interno y externo, y en una representación adecuada de los lugares donde ocurre la búsqueda de reglas. Ambas son diferencias generalizadas que se pueden remediar. Todavía nos queda por mirar la eficacia con que el GPS puede explicar los datos de estos aspectos de la conducta. El sujeto manipula ciertos conjuntos de items en paralelo, utilizando expresiones compuestas; mientras que GPS manipula todos los items al mismo tiempo. En el ejemplo que examinamos, no aparece ninguna diferencia evidente en el resultado, pero podrían surgir mayores discrepancias bajo otras condiciones. Parece bastante claro cómo podría ampliarse el GPS para incorporar esta característica.

Existen dos casos donde no existe ninguna correspondencia clara en el programa para la conducta del protocolo, orientada claramente hacia la tarea. Uno de éstos, el comentario primitivo sobre "complicación", parece que es un caso fundamentalmente de insuficiencia de información. El programa realiza numerosas comparaciones y evaluaciones que podrían hacer surgir los comentarios del tipo en cuestión. De esta forma, este error no parece demasiado serio. El otro caso, incluyendo el pasaje de "tendría ...", parece más serio. Implica claramente un mecanismo (puede ser un conjunto total de ellos) que no está en el GPS. El añadir el mecanismo necesario para manipular este pasaje podría incrementar significativamente la capacidad total del programa. Por ejemplo, no debe existir ninguna forma razonable de conseguir esto si no es proporcionando a GPS un pequeño recuerdo continuo sobre sus acciones pasadas.

### *¿Cuál es la Efectividad de GPS?*

Para comprobar si este programa proporciona una buena teoría o explicación de la conducta del sujeto humano, podemos plantear dos tipos de cuestiones:

- (a) ¿Cuánto tenemos que modificar GPS para construir un programa que se ajuste a este protocolo del sujeto?

- (b) ¿En qué medida es bueno el ajuste del programa modificado en relación con el protocolo?

En los siguientes párrafos discutiremos algunos enunciados metodológicos que están implicados en estas cuestiones. Naturalmente, estamos considerando una prueba de simulación que es mucho más fuerte que la prueba de Turing, y quizá el lector pueda convencerse a partir de estas pruebas que el GPS y sus variantes superarían la prueba de Turing.

### *Ajuste de un Programa General a la Conducta Específica*

Un programa es la forma que toma una teoría —en este caso, GPS. Pero, como en las ciencias naturales, la teoría se expresa con más propiedad como una clase de programas que debe especificarse y aplicarse a situaciones concretas mediante valores paramétricos, condiciones iniciales y condiciones finales. Por ejemplo, un sujeto puede dar mayor prioridad a una diferencia en variables, otro a una diferencia en conectivas. El GPS se ajusta a estas conductas individuales modificando el programa para representar la diferencia en prioridades —esto es precisamente lo que hace al construir el ejemplo que se muestra en la figura 3.

Como en todo ajuste de la teoría a los datos, debemos mantener nuestros grados de libertad. Si dejamos introducir un cambio paramétrico o un nuevo mecanismo para cada fragmento de conducta que se explica, no llegaríamos a explicar nada. El programa debe ser una descripción cuidadosa de los mecanismos que generan la conducta. Hay que señalar —y es seguro que el lector lo sabe— que la huella es la salida del programa, y no se puede modificar a voluntad. Cualquier cambio en el programa afecta la huella en distintos lugares. Un cambio para reducir una discrepancia —por ejemplo, la aparentemente más exhaustiva búsqueda inicial de la lista de reglas, que mencionamos anteriormente— probablemente introduce nuevas discrepancias. En el caso que recordamos, fuimos incapaces de encontrar algún cambio simple en el programa que eliminase esta discrepancia particular y que permitiese que el resto del ajuste continuase igual.

*Comparación de la Huella con la Conducta*

El computador aún no habla un inglés fluido, y por tanto no podemos comparar la huella con el protocolo del sujeto literalmente, palabra a palabra. La huella dice: "META 1: REDUCE  $\Delta P$  ENTRE L1 Y L0". El sujeto dice: "Estoy pensando en la idea de invertir ahora estas dos cosas". La huella dice: "META 6: APLICA R3 A L1". El sujeto dice: "... después miro la Regla 3". En lugar de tener un computador que hable inglés, desearíamos tener un código (en el sentido psicológico de la palabra) que redujese la conversación humana al "contenido de la resolución del problema". Una vez más, las técnicas carecen actualmente de medios para hacer esto, pero quizá estemos de acuerdo en que una gran cantidad del contenido de las observaciones del sujeto está incluida en las "frases" del computador.

Además, la huella describe los procesos de información con un nivel uniforme de detalle, mientras que el protocolo fluctúa mucho en su claridad —algunas veces proporciona más detalle, pero normalmente ofrece mucho menos. Así, en nuestro ejemplo de búsqueda, es perfectamente posible que el sujeto examine el conjunto entero de reglas, pero no consigue mencionar más que unas pocas. Probablemente es inevitable este tipo de desproporción, al menos en el estado presente de nuestro conocimiento. Lo más que podemos pretender es una huella que evite errores de omisión y contradicción, aunque pueda hablar algunas veces cuando el sujeto permanece silencioso.

No queremos decir que sea arbitrario el nivel de detalle en el protocolo —de hecho, sospechamos que está muy relacionado con los mecanismos y funciones de la conciencia en el problema de solución y aprendizaje. La distinción entre consciente e inconsciente no aparece en los mecanismos actuales del GPS, y por tanto no puede reflejarse de una forma determinada en la huella.

Lo que hemos presentado hasta aquí proporciona una muestra del trabajo que se ha realizado al comparar el GPS, y sus variantes, con la conducta de solución de problemas en sujetos humanos. Hasta la fecha, se ha realizado relativa-



mente poca simulación —es una actividad trabajosa. Nuestro estudio empírico más intenso ha sido una simulación de conducta que cubre un período de 30 minutos. Las siguientes conclusiones, a partir de este estudio intensivo y de otros exámenes menos detallados de otros 20 protocolos, parecen razonablemente ciertas:

1. Medido en términos de tiempo y número de palabras, casi toda la conducta de los sujetos está dentro del marco general de análisis medios-fin. Los tres tipos de meta que hemos descrito, explican al menos las tres cuartas partes de las metas de los sujetos, y los tipos adicionales de metas que se han podido encontrar están estrechamente relacionados con las que hemos mencionado.

2. Los tres métodos que hemos descrito representan la gran mayoría de los métodos aplicados a estos problemas. Un método adicional muy importante —planificación— se ha incorporado al GPS (4), pero los límites de tiempo nos impiden discutirlo aquí. La planificación aparece bajo muy diversas formas en los protocolos, pero en todas ellas tiene como función evitar temporalmente detalles para ver si la línea principal de razonamiento conducirá a una solución.

3. Existen pruebas de que los programas de los sujetos cambian —puesto que aprenden— en el curso de la resolución de problemas. Por ejemplo, al principio tienen que examinar las reglas una a una para encontrar una regla aplicable; más adelante, una vez que crean la meta para reducir una diferencia específica, seleccionan casi instantáneamente una regla que es importante para esa diferencia particular. No aparece una distinción clara entre aprendizaje y resolución de problemas. Ocurre algo de aprendizaje —y se utiliza— durante el ensayo de un problema aislado (hora y media). De hecho, la retícula de metas y submetas del GPS constituye un “aprendizaje sobre el problema”, de forma que en fases sucesivas de la solución, el sujeto se comporta de un modo muy distinto. A la inversa, ocurre algo de aprendizaje cuando el resultado de la actividad específica de solución de problemas se dedica al aprendizaje.

*Conclusión*

En este trabajo hemos descrito un método para el estudio de la solución humana de problemas y otros procesos mentales superiores, dando ejemplos de una aplicación del método e indicando la teoría que surge sobre la solución humana de problemas.

El método consiste en construir una teoría de los procesos centrales en forma de un programa o clase de programas, que demuestre la suficiencia de la teoría para producir conducta de resolución de problemas, realizándolo en un computador, y comprobando la teoría sobre procesos humanos por medio de una comparación entre la huella generada por el programa y el protocolo de un sujeto humano.

La aplicación consistió en construir un programa general de resolución de problemas, capaz de solucionar problemas de lógica y otros campos, demostrando que un computador así programado puede resolver problemas, y comparar sus procesos con los de los sujetos humanos en una situación-problema diseñada por O. K. Moore.

La teoría de la resolución humana de problemas consiste en un programa, construido a base de procesos de información elemental, para que razone en términos de metas y métodos para alcanzar esas metas. Quizá la característica más sorprendente de este programa es que selecciona los caminos que explora determinando, en primer lugar, las funciones que deben realizarse y después descubriendo líneas de acción relevantes para esas funciones. De esta y de otras formas, refleja (e incorpora en mecanismos determinados) el "discernimiento" y la "direccionalidad" que se han observado frecuentemente como características sobresalientes de la resolución humana de problemas. En términos de la nomenclatura actual de la psicología, se podría describir como una teoría "mediacional" que abarca los procesos de la "Gestalt". Su novedad es que es precisa y que, al menos en un área del problema, funciona.

Es fácil señalar dificultades y tareas incompletas. Los métodos sistemáticos para ajustar programas a protocolos

y comprobar la bondad del ajuste no existen. Lo "General" del Problem Solver es demasiado específico aun comparado con los procesos humanos que simula. La construcción y comprobación de programas de aprendizaje apenas acaba de comenzar. Sólo se han escrito los programas más rudimentarios para simular los procesos humanos "simples", y todavía no han sido comprobados. Existe poca información para seleccionar el conjunto correcto de procesos elementales; y aún menos para conectarlos con mecanismos neurales.

Al margen de este inmenso programa de asuntos incompletos, deseamos constatar nuestra convicción de que ya no es necesario hablar sobre la teoría de los procesos mentales superiores en tiempo futuro. Existen ahora instrumentos agudos, suficientes para cortar la dura piel del problema, y estos instrumentos han producido ya una explicación detallada, rigurosa de un área importante de la conducta simbólica humana.

#### REFERENCIAS

1. J. S. BRUNER, J. J. GOODNOW y G. A. AUSTIN, (1956): *A Study of Thinking*, New York: Wiley.
2. DEGROOT, A. D. (1946): *Thought and choice in chess*, Amsterdam: North-Holland.
3. MOORE, O. K. y ANDERSON, S. B. (1954): "Modern logic and tasks for experiments on problem solving behavior", *Journal of Psychology*, 38 (1): 151-160.
4. NEWELL, A., SHAW, J. C. y SIMON, H. A. (1959): "A report on a general problem-solving program for a computer", *Computers and Automation*, 8(7): 10-17. Reimpreso en 1960 bajo el título "Report on a general problem-solving program" en *Information Processing* (International Federation for Information Processing, S. de Picciotto, ed.), Paris: UNESCO, 256-264. 5A.
5. NEWELL, A., SHAW, J. C. y SIMON, H. A. (1960): "A variety of intelligent learning in a General Problem Solver", *Self-organizing systems* (M. C. Yovits y S. Cameron, eds.), New York: Pergamon, 153-189. 7A.
6. NEWELL, A., SHAW, J. C. y SIMON, H. A. (1962): "The Processes of creative thinking", *Contemporary approaches to creative thinking* (H. E. Gruber, G. Terrell y M. Wertheimer, eds.), New York: Atherton, 63-119 6A.

7. OSGOOD, C. E. (1953): *Method and theory in experimental Psychology*, New York: Oxford University Press.
8. Turing, A. M. (1950): "Computing machinery and intelligence", *Mind*, 59 (236): 433-460. Reimpreso en 1956 bajo el título "Can a machine think?" en *The world of mathematics* (J. R. Newman, ed.); New York: Simon and Schuster, Volume 4, 2099-2123. MA.

Versión castellana de JULIO SEOANE \*

---

\* Este artículo apareció originalmente en inglés en la obra *Computer simulation of human behavior*, compilada por J. M. Dutton y W. H. Starbuck y editada por Viley & Sons, que ha cedido a *Teorema* los correspondientes derechos de traducción. Agradecemos al Prof. Herbert A. Simon el amable permiso concedido a *Teorema* para publicar la presente versión castellana.