

LÓGICA EN RAZONAMIENTO ORDINARIO: RAZONAMIENTO PRÁCTICO*

JESÚS MARÍA LARRAZABAL ANTIA
LUIS A. PÉREZ MIRANDA

*Departamento de Lógica y Filosofía de la Ciencia
Universidad del País Vasco/Euskal Herriko Unibertsitatea
Apartado 1249
20080 SAN SEBASTIÁN
ESPAÑA*

O objetivo deste artigo é analizar o papel da lógica como uma verdadeira ferramenta para o raciocínio ordinário. Em primeiro lugar, apresentamos uma visão geral das limitações do enfoque chamado “logicista” que usa a lógica como uma linguagem de representação para teorias ordinárias. Depois, chamamos a atenção para uma nítida distinção entre raciocínio teórico e prático, de tal modo que podemos criticar não somente os raciocinadores idealistas, mas também o raciocínio baseado na racionalidade mínima. Finalmente, esboçamos as linhas principais do que deve ser entendido por raciocínio prático.

The aim of this paper is to discuss the role of logic as a real tool for commonsense reasoning. First, we present a general overview of the limitations of the so-called logicist approach to using logic as a representation language for commonsense theories. Then, we point

*Este artículo es una versión ampliada de la comunicación presentada al ICCS-91. El primer autor ha realizado este trabajo en el marco del Proyecto de Investigación del Gobierno Vasco “PGV9225”. El segundo autor ha disfrutado de una beca del M.E.C. del Gobierno Español en el momento de redacción del presente trabajo. Nuestro agradecimiento a los miembros del Seminario de Lógica de la Universidad del País Vasco (San Sebastián) por las críticas recibidas.

out a sharp distinction between theoretical and practical reasoning, in such a way that we can criticize not only the idealistic reasoners but also reasoning grounded on minimal rationality. Finally, we sketch the main lines of what must be understood by practical reasoning.

INTRODUCCIÓN

Tanto la construcción de teorías como la modelización de los procesos de razonamiento ordinario exigen una contextualización y delimitación previa del propio concepto genérico de razonamiento. Estas consideraciones acerca del razonamiento deben permitirnos establecer no sólo qué partes sustanciales de la lógica pueden resultar útiles para llevar a cabo modelizaciones de esos procesos cognitivos, sino también qué otros recursos técnico-conceptuales de diversa índole son necesarios atendiendo a la diferenciación entre razonamiento teórico y práctico. Desde un punto de vista metodológico analizaremos las diferentes perspectivas existentes en la modelización del razonamiento ordinario, viendo cómo difieren en cuanto a intereses y objetivos prioritarios.

Es precisamente en la filosofía subyacente a la teoría general de la racionalidad donde en realidad se establecen las bases teóricas de lo que entendemos por razonamiento ordinario. La cuestión se centra básicamente en el tipo de procesos que en última instancia nos interesa describir. Por ejemplo, un psicólogo puede estar interesado en la teoría de la competencia y de la actuación respecto al razonamiento centrada tan sólo en humanos, mientras que un teórico de la inteligencia artificial (en adelante IA) puede pretender dar cuenta de ambas parcelas en el caso de agentes artificiales, utilizando, a su vez, todos los recursos computacionales que la informática le ofrece. En este punto surge la polémica en torno a la necesidad de tomar como modelo de los procesos de razonamiento a humanos o a máquinas. Al mismo tiempo debemos

decidir sobre qué tipo de razonamiento ordinario queremos desarrollar nuestra investigación: razonamiento analógico, probabilista, inductivo, abductivo, deductivo, etc. Nuestros intereses epistémicos se centran fundamentalmente en el razonamiento deductivo del sentido común tanto teórico como práctico. Antes de ocuparnos de la importante cuestión de la diferenciación entre razonamiento teórico y práctico, en el siguiente apartado tratamos de señalar algunas de las discusiones y desavenencias de orden epistémico-computacional que ha originado el uso de la lógica en el razonamiento ordinario. En cualquier caso, la motivación subyacente a nuestro trabajo es precisamente la de mostrar que parte de esas diferencias de favor hacia el uso o no de la lógica, en la construcción de modelos computacionales del razonamiento ordinario, tiene su origen más en problemas conceptuales acerca de lo que entendemos por el propio concepto de razonamiento, que en problemas intrínsecos a la lógica como la adecuación o la indecibilidad.

1. SOBRE EL USO DE LA LÓGICA EN EL RAZONAMIENTO ORDINARIO

Desde un punto de vista general no hay consenso en lo que se refiere a la metodología a seguir en la construcción de modelos computacionales del razonamiento. Tradicionalmente han coexistido dos posiciones, en principio, antagónicas: por un lado, aquellos que piensan que la estrategia es diseñar programas de cómputo que razonen sirviéndose de los bien construidos por la lógica matemática, independientemente de si es o no el modo en el que la gente piensa; por otro lado, quienes creen que una aproximación fructífera es intentar hacerse con computadores que no sólo imiten sino también reproduzcan el modo en el que la gente razona, que no es precisamente el modo en el que lo hace la lógica matemática. Si bien estas dos posturas debaten el problema que constituye para la IA la representación y manipulación del conocimiento del sentido

común, por extensión, ambos puntos de vista pueden focalizarse sobre este mismo tipo de conocimiento pero aplicado al razonamiento ordinario.

Los seguidores de Minsky aceptan el argumento de que el mejor modo de proceder es buscando que lo artificial imite a lo natural. En su opinión la lógica no puede enfrentarse con todas las garantías a problemas de la realidad. "Thinking always begins with suggestive but imperfect plans and images; these are progressively replaced by better-but usually still imperfect-ideas" (Minsky 1975, p. 123). En el apéndice de la crítica al enfoque logicista explica por qué los enfoques 'lógicos' no funcionan correctamente. Señala que desde Aristóteles ha habido intentos de representar el razonamiento del sentido común mediante un sistema "logicista" — un sistema que establece una completa separación entre

- (1) "proposiciones" que incorporan información específica,
- y
- (2) "silogismos" o leyes generales de inferencia adecuados.

"The strategy of complete separation of specific knowledge from general rules of inference is much too radical. We need more direct ways for linking fragments of knowledge to advice about how they are to be used". (Minsky *Ibid.*, p. 127). Minsky señala que nadie ha podido confrontar con éxito un sistema tal con un conjunto de proposiciones lo suficientemente extenso de la vida real. En su opinión este tipo de intentos seguirán fallando en función de las características propias del *logicismo* más que por defectos singulares de los formalismos. Esta crítica al logicismo extremo es a nuestro juicio acertada. El problema no es un problema intrínseco a la lógica, sino de su inadecuada aplicación para determinadas tareas.

En micromundos de una complejidad determinada los sistemas puramente logicistas pueden simular pensamiento del sentido

común para propósitos de alto nivel; también son válidos para la axiomatización de pequeños items básicos indispensables en la fundamentación de un plan que guíe nuestra acción posterior. Sin embargo, los condicionantes excesivamente normativos impuestos por el logicismo impiden a los mecanismos formales adecuarse debidamente a problemas de la realidad. Citemos de nuevo a Minsky: "(...) in a simple case one get such systems to 'perform', but as we approach reality, the obstacles become overwhelming. The problem of finding suitable axioms — the problem of 'stating the facts' in terms of always-correct, logical assumptions — is very much harder than is generally believed" (*Ibid.*, p. 124). Por su parte, aquellos que defienden el poder y la precisión de los lenguajes formales son un tanto escépticos respecto a la habilidad para descubrir algo 'nuevo' que sea más útil para construir teorías del razonamiento. Estos últimos mantienen una visión de la IA como "a most abstract inquiry into the possibility of intelligence or knowledge" (Dennett 1978). Del mismo modo, Kolata (1982) señala algunas de las afirmaciones de McCarthy al respecto: "Whether logical reasoning is really the way the brain works is beside the point, (...). This is AI and so we don't care if it's psychologically real". No obstante, la metodología seguida en IA en los últimos años nos muestra que ambas perspectivas no sólo no son contrapuestas, sino que deben ser complementarias para el adecuado tratamiento de las teorías del razonamiento (Pollock 1987). Nuestro objetivo consiste en dar luz a la idea de que puede encontrarse un equilibrio que permita seguir manteniendo una *metodología* de orden lógico en la modelización del razonamiento ordinario, sin desatender a otros criterios más pragmáticos como la adecuación práctica o la realizabilidad.

Israel (1985) señala que las objeciones más importantes históricamente a la aplicabilidad de la lógica en IA han apuntado no a las capacidades expresivas de los lenguajes lógicos, sino a la afirmación imputada de que el razonamiento del sentido común po-

dría ser adecuadamente recogido poniendo en funcionamiento un probador de teoremas presumiblemente correcto sobre la base de un lenguaje lógico. En opinión de Israel la crítica de Minsky a los enfoques lógicos presenta la tendencia equivocada de equiparar pensamiento con pruebas para argumentos, con el fin de abandonar el uso de la lógica formal en la representación. "Minsky, thus, is best taken as objecting to the notion that thinking is proving, and to various of its alleged corollaries" (Israel 1985, p. 433). Hoy en día, sin embargo, es posible construir razonadores automatizados bastante sofisticados con técnicas que difieren en un sentido sustancial de las empleadas por los probadores de teoremas clásicos basados en algún formalismo no monótono. Por ejemplo, el razonador automatizado OSCAR basado en el concepto de *defectibilidad enumerable* (Pollock 1991) es completo para la lógica de primer orden y probablemente adecuado para la concepción basada en argumentos del razonamiento defectible (Pollock 1987).

Las dos opiniones con anterioridad encontradas vuelven a verse enfrentadas en cuanto a la importancia del mantenimiento de la consistencia para nuestros sistemas de representación y manipulación del conocimiento. Los seguidores de McCarthy afirman que la lógica simplemente nos ofrece una formulación precisa de varias nociones de consistencia y — si los enunciados en cuestión son enunciados de primer orden — ofrece procedimientos para mostrar cuándo un conjunto de enunciados es inconsistente. De aquí no puede inferirse que el sentido común deba ser necesariamente consistente, especialmente si tenemos en cuenta toda la información (temporal, espacial, contextual, etc.) que conjuntamente manipulamos para llevar a cabo una acción simple, pongamos por caso. Sin embargo, Minsky no cree que la consistencia sea necesaria o incluso deseable en el desarrollo de los sistemas inteligentes. Para él lo que es importante es saber cómo manejamos paradojas o conflictos, cómo aprendemos de los errores, cómo nos apartamos de las inconsistencias intuitivas. Minsky incide en la necesidad de to-

mar seriamente la posibilidad de la inconsistencia. Para Israel, sin embargo, no hay ninguna razón para que aquellos comprometidos con el uso de la lógica como lenguaje de representación para las teorías del sentido común no debieran tener este punto en cuenta. El uso de un lenguaje lógico estándar no nos prohíbe manejar la inconsistencia más o menos, sea cual sea el modo en que pensemos mejor.

Las reglas de inferencia no tienen nada que ver con la conducta de alguien o de algo. Al igual que en los sistemas estándar, podemos derivar cualquier enunciado arbitrario a partir de una contradicción, en cualquier sistema conocido por el hombre — a partir de cualquier enunciado S podemos derivar, para cualquier n , la conjunción n -aria de S consigo mismo, y también su disyunción n -aria. Pero el que llevemos a cabo o no este tipo de implicaciones parece depender de cada agente racional, de sus recursos (ligados a una arquitectura cognitivo-computacional determinada), intuiciones y propósitos. Respecto al problema de la *relevancia*, que está estrechamente relacionado con lo que ordinariamente entendemos por razonamiento práctico, la lógica clásica no dice nada. Como veremos en el siguiente apartado, la diferenciación entre razonamiento teórico y razonamiento práctico parece ineludible para evitar parte de las confusiones corrientes que se producen debido a una malinterpretación del propio concepto de razonamiento.

La consistencia, como concepto lógico, parece netamente adecuado en tanto nos establece cuándo dos enunciados pertenecientes al mismo conjunto entran en contradicción. Surge un problema cuando trabajamos computacionalmente sobre bases de datos, que en principio nos impide mantener la posibilidad de inconsistencias entre la información. No obstante, la posibilidad de enfrentarnos a situaciones en las que dos opciones contrapuestas deban ser valoradas de cara a tomar una decisión que no puede esperar por más tiempo es un hecho más que corriente en los problemas cotidianos. Algunas respuestas se han dado en relación a

este punto: Pequeno & Buchsbaum (1991) consideran la noción de inconsistencia epistémica. En su opinión, estas contradicciones reflejan no una conducta anormal del estado de cosas sino la incompletud (o vaguedad) de nuestro conocimiento acerca de esta noción. Su idea es presentar una lógica, cálculo y semántica, que haga precisa esta noción y permita razonar sobre estos puntos de inconsistencia, sin trivialidad; también Pollock (1992) aporta algunas soluciones a este problema de la consistencia en el tratamiento de los argumentos autodefectibles.

Para Minsky, el razonamiento lógico es más apropiado para demostrar o confirmar los resultados del pensamiento que para el pensamiento mismo. Sin embargo, no hay duda de que el mayor ímpetu de los lenguajes formales fue precisamente el propósito de obtener especificaciones sistemáticas y formales de la relación de consecuencia lógica. Junto a esto debería quedar claro que los datos manipulados por esta teoría de la consecuencia lógica son juicios de validez para argumentos expresados en los lenguajes naturales. Como señaló Hayes, la axiomatización formal trata de expresar sistemáticamente algún cuerpo de conocimiento en un lenguaje formal para el cual las nociones de consecuencia lógica y derivabilidad pueden ser definidas. "One must keep in mind the goal: that of telling a robot enough of what it needs to know about its world to enable it to do what, say, we want it to do. Hayes isn't engaged in Philosophy of Science; he's engaged in arguing for a particular strategy for designing artificial minds" (Israel 1985, p. 439). Para Israel pueden existir razones para dudar de la adecuación de los formalismos lógicos para el razonamiento ordinario; pero éstas no tienen nada que ver con las afirmaciones acerca de la adecuación de las reglas de prueba deductivamente correctas.

En resumidas cuentas, de la lógica lo que obtenemos son unas determinadas propiedades sintáctico-semánticas más o menos bien estructuradas a través de los teoremas que la metalógica nos ofrece, caracterizando la lógica correspondiente. Otra cuestión es

que los recursos lógicos disponibles resulten necesarios y/o suficientes para determinados propósitos. Si queremos dar cuenta del razonamiento en un sentido global, de modo que incluyamos tanto razonamiento práctico como teórico (diferencia que se establece en el siguiente apartado), inmediatamente nos percataremos de que algunos cambios deben ser introducidos en el cálculo lógico clásico de primer orden.

En un proceso de razonamiento ordinario podemos rechazar algunas de las creencias con las que empezamos, e incluso volverlas a reinstaurar más adelante si así nos conviniera. En el razonamiento debemos tener en cuenta toda la evidencia relevante en la información que disponemos. En efecto, debemos a menudo intentar conseguir más evidencia si la que disponemos es considerada insuficiente, lo que, no obstante, no indica que no nos podamos ver llevados a actuar sobre una base incierta de información si las circunstancias apremiaran. Los juicios acerca de la relevancia y la fuerza de la evidencia son típicamente ellos mismos el producto del razonamiento (Sheridan 1991). Hay que buscar instrumentos lógico-conceptuales que nos ayuden no sólo en la búsqueda de métodos para razonar bajo incertidumbre, sino también métodos para el control del razonamiento, y para situar a los métodos heurísticos bajo bases correctas. Por lo hasta aquí expuesto, no hay argumentos lo suficientemente fuertes como para impedir que el *lenguaje de la lógica* entendido en un sentido amplio pueda servir como instrumento en las tareas de modelización del razonamiento ordinario. En cualquier caso, nuestro comentario y planteamiento inicial exige una argumentación que está basada fundamentalmente sobre ciertas consideraciones de orden epistémico: diferencias entre razonamiento teórico y razonamiento práctico; delimitación del papel del contenido respecto a las actitudes proposicionales; el problema de la relevancia. Se trata de crear un transfondo teórico-conceptual que nos garantice el desarrollo de los modelos computacionales sin

abandonar la metodología y el lenguaje que la lógica nos proporciona.

2. RAZONAMIENTO TEÓRICO Y PRÁCTICO

En la sección anterior han quedado patentes algunas dificultades de orden epistemológico en cuanto a la utilización o no de los formalismos lógicos para la tarea de modelizar el razonamiento ordinario. Si atendemos a la distinción conceptual entre razonamiento teórico y razonamiento práctico, que tratamos de explicar a continuación, vemos que en unos casos una parte considerable de las dificultades desaparece, y en otros toma un cariz diferente.

La diferencia que establecemos entre el razonamiento teórico y el razonamiento práctico está en concordancia con la línea seguida por Pollock (1992). También los trabajos de Doyle (1988, 1988a) y Audi (1989) inciden en la importancia de esta diferenciación, que ya fue recogida en la filosofía clásica principalmente por Aristóteles, Hume y Kant. Ciertamente existen notables diferencias entre estos pensadores clásicos en su modo de entender los procesos de razonamiento ordinario, sobre todo en lo que respecta al razonamiento práctico. No obstante, lo que nos interesa en esta ocasión es recoger la idea primigenia de la necesidad de establecer una diferenciación clarificadora entre ambos tipos de razonamiento.

El razonamiento acerca de qué acciones llevar a cabo es razonamiento práctico. El razonamiento práctico presupone creencias acerca del mundo, y éstas son proporcionadas por el razonamiento teórico. "The former are solved by one's practice, say by taking the right detour; the latter are solved by one's forming, or bringing to bear, the right belief, say by working out a sound proof that an axiom system is consistent" (Audi *Ibid.*, p. 1). La principal función del razonamiento teórico en la racionalidad es proporcionar el trasfondo factual requerido para el razonamiento práctico. Si man-

tenemos la idea de que ambos tipos de razonamiento no pueden ser enteramente separados, podemos establecer la relación siguiente. Primero, el razonamiento práctico debe estar basado sobre conocimiento de la situación ordinaria del agente. Este conocimiento debe ser el resultado del razonamiento teórico. De este modo, cualquier razonamiento práctico debe presuponer razonamiento teórico previo. Esto tiene la consecuencia inmediata de que el razonamiento teórico no puede ser visto como un caso especial de razonamiento práctico. La labor principal del razonamiento teórico es ayudar al agente a resolver problemas de orden práctico. Podría decirse que el razonador teórico sólo entra en acción en respuesta a cuestiones planteadas por el razonamiento práctico. El razonador teórico no razona al azar, sino que está dirigido a intereses, en el sentido de que trata de responder a cuestiones planteadas por el razonador práctico. Tratando de contestar a estas preguntas puede verse llevado a plantearse otras cuestiones que sólo están indirectamente conectadas con el razonamiento práctico, pero en última instancia debe ser el razonamiento práctico el que dirija los intereses del razonador teórico. Dicho de modo más sencillo, tanto la información que se maneja como los *mecanismos inferenciales* (sean éstos de la naturaleza que sean) forman parte del razonamiento teórico. Ahora bien, el control que se ojerce sobre la información disponible desde donde se obtienen las conclusiones, así como el establecimiento del grado de relevancia de esa información para resolver un problema específico, forman parte de las tareas del razonamiento práctico.

Hay que tener además en cuenta que los agentes racionales planifican su conducta tanto a corto como largo plazo, por lo que el estudio del razonamiento ordinario exige a su vez un análisis pormenorizado de la estructura formal de los planes, así como de su significación para el razonamiento práctico. No hay duda de que en el ámbito del razonamiento la lógica debe ser entendida como un instrumento y no como un fin en sí mismo. En última instancia,

son nuestros deseos los que nos llevan a emprender una acción previamente racionalizada a través de un razonamiento práctico, que a su vez emplea mecanismos inferenciales que le son proporcionados por el razonamiento teórico. De nada sirve en estos casos un sistema lógico sobre el que no se imponga ningún tipo de restricciones; así, nuestro potencial inferencial debe adecuarse a las demandas impuestas por las necesidades reales en el momento de realizar una acción. No es lo mismo enfrentarse al problema del razonamiento al margen de cualquier tipo de restricción espacio-temporal que hacerlo en función de unas condiciones específicas de espacio y tiempo. Tampoco podemos olvidar que gran parte de los procesos de razonamiento ordinario que se llevan a cabo tienen como objetivo básico el análisis, deliberación, y posterior elección entre cursos de acciones que usualmente forman parte de planes complejos. Como veremos más adelante la planificación juega un papel fundamental en el estudio del razonamiento práctico. También se deberían considerar las restricciones de arquitectura cognitiva o, en su caso, computacional a las que se ve sometido el agente racional en cuestión. En consecuencia, pensamos que los condicionantes señalados deben ser tenidos en cuenta por cualquier abordaje lógico-formal que pretenda servir como soporte de una teoría general del razonamiento.

Si el razonador teórico trata de responder a preguntas razonando tan sólo a partir de la información ya a su disposición, entonces será incapaz de contestar a muchas de las cuestiones planteadas por el razonamiento práctico (a pesar de que éste sea el modo en el que usualmente funcionan los diferentes mecanismos inferenciales del razonamiento no monótono, esto es, se trabaja de un modo autónomo sobre el razonamiento teórico, sin tomar en consideración su verdadero papel en el marco de una teoría del razonamiento más amplia en la que también tiene cabida el razonamiento práctico). Generalmente, responder a las preguntas prácticamente motivadas requiere investigar el mundo empíricamente. Esto re-

quiere llevar a cabo acciones. Para ello parece imprescindible establecer un bucle que vaya desde el razonador práctico pasando por el razonador teórico y que vuelva nuevamente al razonador práctico. En resumen, los problemas prácticos se refieren a nosotros en cuanto agentes y conciernen a lo que uno debe hacer, mientras que los problemas teóricos se refieren a nosotros como conocedores, y conciernen a cuestiones acerca de lo que es verdadero o falso.

Otra cuestión importante ligada a esta diferencia entre los dos tipos de razonamiento es la discusión surgida en IA a la hora de inclinarse por adoptar razonadores *crédulos o escépticos*. Estos últimos retienen la creencia tal cual cuando hay razones igualmente buenas tanto a favor como en contra. Aquéllos, por su parte, eligen una conclusión al azar. Pollock (1991) señala que ha sido inadecuadamente defendido que la elección entre razonadores crédulos y escépticos es una cuestión más de gusto que de análisis lógico. Sin embargo, opina que los primeros son lisa y llanamente incorrectos. Supongamos que no tenemos mejores razones para pensar que está lloviendo que para creer que no está lloviendo. Si eligiéramos una cualquiera de las dos razones sin mayor convicción, probablemente nadie pensaría que nuestra elección pudiera catalogarse como racional. Pollock (1989) argumenta que esta controversia también surge de una confusión entre el razonamiento teórico (razonamiento acerca de qué creer) y el razonamiento práctico (razonamiento acerca de qué acciones llevar a cabo). En el razonamiento práctico, si no tenemos base suficiente para adoptar una de las creencias entre sí, razonaremos al azar. Ambos tipos de razonamiento pueden tener un tratamiento independiente y autónomo en un alto grado dada su fuerte especificidad. Desde una teoría general del razonamiento, habría que estudiar detalladamente las interrelaciones necesarias entre aquéllos con vistas a crear un modelo cognitivo más general. La comprensión de las interacciones existentes entre ambos tipos de razonamiento nos sirve para situar en el lugar preciso la problemática de la modelización de este pro-

ceso y para ubicar al lenguaje lógico correspondiente (lógicas no-monótonas, lógicas intensionales), dentro del marco de una teoría general del razonamiento que tiene fuertes fundamentos conceptuales en Ciencia Cognitiva.

3. EL CONTENIDO EN EL RAZONAMIENTO PRÁCTICO

Directamente ligada a la diferenciación entre esos dos tipos de razonamiento aparece la cuestión del contenido. Según interpretemos el papel de éste en el curso del razonamiento estaremos ante un tipo u otro de acercamiento desde un punto de vista semántico. Si nos preguntáramos qué es para un estado cognitivo tener una propiedad semántica, hay un número de cosas diferentes sobre las que podríamos optativamente centrar nuestros intereses. ¿Qué es para un estado cognitivo tener una condición de verdad? ¿Qué es para un estado cognitivo ser acerca de algo, o referirse a algo, o ser verdad acerca de algo? ¿Qué es para un estado cognitivo ser un estado intencional? Usualmente en ciencia cognitiva se han visto a estas dos cuestiones como dependiendo de otras dos principales. ¿Qué es para un estado cognitivo tener un contenido? ¿Qué es para un estado tal tener un contenido específico? (Cummins 1991). Sin duda todas estas cuestiones tienen cabida en un análisis de lo que podríamos denominar representación mental del significado. Ahora bien, la pregunta que realmente nos interesa es la siguiente: ¿qué partes del contenido, si alguna, intervienen en los procesos inferenciales relativos al razonamiento práctico?

En muchas ocasiones se dice erróneamente que el razonamiento ordinario, sin mayor delimitación, atiende más a cuestiones prácticas que a consideraciones de consistencia; que el razonamiento basado en las diferentes extensiones de la lógica de primer orden no puede dar cuenta de las cuestiones de contenido ligadas a las actitudes proposicionales tales como creencias, intenciones, etc. En otros términos: que siguiendo las líneas del cognitismo clásico no

es posible representar funcionalmente de forma completa las propiedades semánticas de las expresiones que estamos manipulando en el curso de un razonamiento. ¿Hasta qué punto este problema representacional es relevante en un proceso inferencial? Para algunos psicólogos cognitivos que defienden la teoría de los *modelos mentales* el contenido debe quedar adecuadamente representado en los procesos inferenciales. "People understand the meanings of statements, and so it is odd to suppose that when they reason they throw their understanding away and work with formal rules that are purely syntactic. There is, in fact, a semantic procedure that they can use to reason deductively" (Johnson-Laird 1988, pp. 226-27). En otras palabras, las deducciones no son inmunes al contenido de las premisas. Sin embargo, un análisis más profundo de la cuestión nos pone de relieve que las consideraciones semánticas ligadas al significado en el discurso no intervienen en el proceso mismo de inferencia. El contenido de los pensamientos, representados en enunciados, parece ser independiente del proceso de extraer conclusiones a partir de esos mismos enunciados. A no ser que nos refiramos a su *contenido lógico*, en el sentido defendido por Salmon (1992). En este caso, entenderemos el contenido lógico como el valor semántico primario que un enunciado comparte con sus equivalentes lógicos. Un problema diferente es que tengan incidencia sobre el proceso inferencial las suposiciones o asunciones que cabría formular como parte de un razonamiento suposicional en la construcción de un argumento no lineal, la nueva información recibida a través de la percepción, o las demandas impuestas por el razonamiento práctico (más ligado a lo que los psicólogos denominan *realización*). Además hay que tener en cuenta que el razonamiento ordinario tiene un carácter *defectible*. Las razones tanto teóricas como prácticas de los argumentos que formulamos usualmente no tienen un carácter conclusivo, sino defectible. P puede ser una buena razón para creer Q. Ahora bien, P no tiene por qué derivarse deductivamente a partir de Q. Si bien en la práctica, dada

la información presente, existe la *justificación epistémica* para actuar bajo la creencia de Q. En cualquier caso, los problemas inherentes al funcionalismo computacional en cuanto a la representación del significado no influyen en la caracterización de las teorías del razonamiento defectible y, consiguientemente, tampoco lo hacen en su modelización.

Lo cierto es que parece posible un análisis casi integralmente sintáctico-formal de la cuestión. La presencia, en su caso, tanto de la semántica formal como de la semántica del lenguaje natural, en nuestro modelo, tendría lugar con independencia de los procesos inferenciales. Esto es, podemos 'prescindir' de las interpretaciones con las que cabría recubrir nuestras representaciones sintácticas a la hora de construir un modelo computacional del razonamiento. En el caso de las teorías del razonamiento práctico, el *contenido* queda reducido al papel desarrollado por las actitudes proposicionales. Las teorías computacionales del razonamiento (Cohen and Levesque (1987), Konolige (1991), Pollack (1991)), asumiendo que la *intencionalidad* de la información manipulada viene dada, tratan de recoger el papel estructural que juegan las actitudes proposicionales en la planificación y conducta de los agentes racionales artificiales.

Una opción de estas características exige explicar por qué esta *reducción* relativa al contenido es suficiente para la creación de modelos para razonamiento ordinario. Autores como Quine consideran irrelevantes desde un punto de vista científico a los conceptos relacionados con la psicología popular, o lo que es lo mismo, que las actitudes proposicionales no constituyen en sí mismas ningún objetivo para las teorías científicas. Nosotros, por el contrario, pensamos que las actitudes proposicionales facilitan la comprensión intuitiva de los procesos inferenciales que intervienen en las distintas fases del razonamiento práctico. Esto no quiere decir, sin embargo, que asumamos que el estudio de la representación mental quede reducido al estudio de las actitudes proposicionales;

más bien la tesis aquí supuesta es que éstas forman parte de un conjunto más amplio en el que quedan englobadas las diferentes categorías semánticas que intervienen en los procesos de representación. De hecho, las actitudes proposicionales no configuran sino una parte del contenido. El contenido puede subdividirse a su vez en más categorías (Cummins 1991). En la explicación de los procesos de razonamiento, desde un punto de vista computacional, las actitudes proposicionales ayudan en la construcción de modelos que pueden clarificar el funcionamiento de esos procesos desde la perspectiva del agente racional. Lo que podríamos denominar una explicación de alto nivel.

Cuando hablamos de creencias, intenciones, propósitos, compromisos, etc, nos movemos en el terreno del razonamiento práctico. Nos servimos del razonamiento práctico cuando queremos llevar a adelante nuestros propósitos como parte de un plan que guíe adecuadamente nuestras acciones. Estos son conceptos ligados al contenido intencional que pueden ser de hecho representados en sistemas basados en lógicas intensionales. Las actitudes proposicionales pueden ser comprendidas en función de su forma física y, como a toda expresión sintáctica correcta, cabe recargarla con significados. El tratamiento del problema de los contenidos semánticos queda reducido a la cuestión de las actitudes proposicionales, al menos para las teorías computacionales que se centran en el razonamiento práctico en el sentido indicado. El resto de los contenidos semánticos pasan desapercibidos para todas estas teorías de la representación y del razonamiento. Con ello no se quiere decir que estos contenidos semánticos no tengan un papel importante en la configuración de un modelo cognitivo general, sino más bien que no ayudan en el sentido defendido por algunos psicólogos cognitivos (especialmente Johnson-Laird y sus seguidores desde su teoría de los 'modelos mentales') con vistas a la creación de un modelo computacional del razonamiento.

La manipulación simbólica es llevada a cabo sobre la forma de las expresiones. Que a su vez quepa la posibilidad de cargar a las mismas con significados es algo que en principio no afectaría a la construcción de modelos formales del razonamiento. Una clarificación acerca de la diferencia entre contenido semántico y actitudes proposicionales parece vital para acabar con parte de los malentendidos que se producen cuando pensamos acerca del papel del 'significado' sobre los procesos formales de inferencia en teorías computacionales. Todo parece indicar que son las intuiciones sobre nuestro conocimiento las que guían nuestra labor en la construcción computacional de este tipo de procesos. Las condiciones de verdad de las expresiones lingüísticas nos sirven para entender por qué adoptamos un tipo de creencias a la vez que deseamos otras. Es decir, las distintas 'interpretaciones' de la información colaboran en la configuración formal, pero una vez que ésta es estructurada, el procedimiento podría seguir un camino meramente sintáctico.

4. RACIONALIDAD Y RAZONAMIENTO PRÁCTICO

A partir de este momento podemos ir analizando las distintas situaciones que puede ofrecernos el estudio del razonamiento deductivo ordinario. Por un lado, podríamos hacer un estudio básicamente teórico suponiendo que el agente que realiza las deducciones es de carácter omnisciente, al margen de cualquier restricción espacio-temporal. De hacerlo así, de inmediato nos daremos cuenta que este acercamiento ofrece escasas aplicaciones. La insuficiencia de recursos computacionales hace inviable la posibilidad de modelización de un razonamiento semejante. Igualmente, si no poseemos toda la información necesaria para establecer cuándo un determinado enunciado de nuestro sistema de conocimiento es verdadero o falso, tendremos que implantar técnicas de análisis que nos permitan manipular con un mínimo de garantías y efectividad a

ese razonamiento deductivo incompleto. Es esta clase de razonamiento la que ha dado más juego y ha sido el objeto de interés de numerosas investigaciones en IA: *Asunción del Mundo Cerrado*, *Complección de Predicado*, *Circunscripción*, *Razonamiento por Defecto*, *Razonamiento Defectible*.

A estos tipos de razonamiento, que podríamos considerar como teóricos en función de la diferenciación establecida anteriormente, cabe introducirles determinadas contextualizaciones, actitudes proposicionales, etc, que nos sitúan en la perspectiva del razonamiento práctico. Hay que tener en cuenta que los enfoques de la IA que se enfrentan a problemas de razonamiento con información incompleta, en principio no parece que se planteen la posibilidad de aplicabilidad real de sus teorías en el ámbito más general del razonamiento (aunque quizás ineludiblemente éste sea su último objetivo a largo plazo), sino que más bien tratan de resolver problemas derivados del manejo de ese tipo de información para bases de datos y conocimiento. Sin embargo, nuestro objetivo principal en este apartado pasa por la recopilación de los resultados de los que pudiera derivarse una posible aplicación dentro del esquema de una teoría del razonamiento ordinario para agentes reales, que indirectamente pudieran también ser recogidos por la Inteligencia Artificial. En primer lugar, veremos cómo las teorías excesivamente idealistas no son convenientes desde un enfoque que se interese por la implementación de los modelos del razonamiento. A estas últimas, dependiendo de nuestros objetivos, hay que introducirles ciertas restricciones que, a la larga, nos permitirán ir adecuando nuestros modelos a situaciones más complejas, y que paulatinamente se irán asemejando gradualmente a lo que intuitivamente entendemos por un proceso de razonamiento deductivo ordinario.

4.1 Inadecuación del Razonamiento Ideal

La idealización típica requiere que el agente sea capaz de realizar virtualmente cualesquiera inferencias deductivas a partir de sus creencias, disponiendo a su vez de una capacidad perfecta para elegir acciones apropiadas para su conjunto de creencias. Un ejemplo claro de esta situación lo constituyen los modelos formales de creencia construidos sobre la base de la semántica de los mundos posibles de Hintikka (1962), los cuales exigen que la creencia esté cerrada bajo la relación de consecuencia lógica. Esto significa que si un agente cree un conjunto de fórmulas, entonces debe creer todas las fórmulas que se deriven lógicamente de ese conjunto. Aquí derivación lógica es la relación semántica estándar en la que todos los modelos del conjunto de premisas deben ser asimismo modelos de la conclusión. Como puede verse fácilmente esta noción es independiente de cualquier consideración de orden epistémico o computacional. De este modo, el cierre consecucional es obviamente demasiado fuerte para representar adecuadamente a la creencia; esto es, está usualmente más allá de las capacidades intelectuales de cualquier agente racional natural o artificial determinar si un enunciado particular es lógicamente derivado por su conjunto de premisas.

En resumidas cuentas, no parece posible que cualquier teoría cognitiva que use la condición del cierre deductivo tenga aplicaciones para el caso de agentes tanto artificiales como naturales con unas capacidades limitadas para el razonamiento. Además, el hecho de que esta condición de cierre exija que el agente infiera todas las consecuencias a partir de sus creencias, con independencia de que sean relevantes o no para éste en la resolución de un problema, la convierte igualmente en ineficiente desde un punto de vista práctico. Por lo tanto, no es que los racionales, tal y como veremos de modo más específico en el siguiente punto, no dispongan de una *capacidad inferencial mínima*, sino más bien que, dadas sus limi-

taciones tanto computacionales como cognitivas, o de otro tipo, existe la necesidad de controlar y dirigir esa capacidad inferencial con el fin de que el razonamiento tenga sentido y sea a su vez efectivo. Y es precisamente el tipo de razonamiento que nosotros denominamos práctico quien guía y conduce nuestras capacidades, llamémoslas "lógicas", para razonar abstractamente. El razonamiento teórico es necesario, pero no suficiente para enfrentarse con éxito a problemas de la vida cotidiana. En función de la importancia que le demos a la implementación o no de las teorías del razonamiento, así como a su adecuación con la realidad, trabajaremos con sistemas más o menos omniscientes. En nuestro caso debemos prescindir de sistemas excesivamente ideales dado que nuestro propósito a largo plazo es la modelización de aspectos del razonamiento ordinario que puedan ser implementables.

Si queremos mantener una concepción *lógica* de la racionalidad, que tenga sentido en el marco de los agentes finitos limitados (naturales o artificiales), tendremos necesariamente que adecuarla tanto a las demandas del razonamiento práctico para enfrentarnos a problemas de la vida real, como a las restricciones computacionales en función de los recursos disponibles. Hay que pensar que en las circunstancias en las que no es posible dar una respuesta maximalmente óptima, tal y como se diría en teoría de la decisión, los agentes deben servirse de heurísticos para resolución de sus problemas. La teoría de la racionalidad mínima es uno de los pocos ejemplos en filosofía que pretende dar respuesta a estas cuestiones. Bien es verdad que algunos de sus supuestos no son del todo satisfactorios y deberían revisarse, a nuestro juicio.

4.2 Acerca de la Racionalidad Mínima

La idea del agente mínimamente racional desarrollada por Cherniak (1986), contrapuesta a la del agente epistémico perfecto, puede ser concebida como parte de un programa general para la

construcción de una teoría filosófica del agente racional menos idealizada que las inspiradas en teorías de decisión teórica estándar (Savage (1972), DiFinetti (1975), Jeffrey (1973)). Recuérdese, por ejemplo, que el cálculo probabilista que determina la elección de una acción individual entre distintas posibles, en función del valor de su *utilidad esperada*, es computacionalmente problemático. La teoría de la decisión nos proporciona un método para establecer cuál de entre el conjunto de acciones es maximalmente óptima en el contexto en cuestión. Sin embargo, suelen ser tantas las acciones a analizar y los cálculos probabilistas a realizar que ese método resulta generalmente ineficiente desde un punto de vista práctico. De ahí la necesidad existente de buscar teorías que eviten lo que podríamos denominar la *parálisis práctica* de los procesos de decisión.

El programa para desarrollar teorías menos idealizadas del agente epistémico racional mínimo, tanto para ciencias empíricas como teóricas, se enfrenta con dos problemas extrafilosóficos: estudios psicológicos de razonamiento humano *prima facie* y la teoría de la complejidad computacional. Cherniak (1986, pp. 75-76) se expresa en los siguientes términos: “The psychology of irrationality reveals remarkably persistent and ubiquitous use of formally incorrect heuristics in everyday reasoning. Complexity theory indicates quite stringent practical limitations on even ideal computers with literally the resources of the entire universe ... it is, as a practical matter, Church's Undecidability Theorem applied to formally correct decision procedures all the way down to tautology testing. The “quick and dirty” heuristics may therefore be indispensable to do better than guessing while avoiding computational paralysis”. Por un lado, se toman los resultados en psicología como ilustrativos a la hora de crear modelos de procesos deductivos basados en creencias. Por otro, se adoptan como límites teóricos los resultados de la teoría de la complejidad computa-

cional. Podríamos decir que se trata de una aproximación a caballo entre las propuestas normativas y las propuestas descriptivas.

Desde luego, los modelos de la racionalidad no deberían recubrir procesos que fueran computacionalmente intratables. Por la teoría de la complejidad computacional sabemos que podemos encontrarnos procedimientos formalmente correctos y que, no obstante, sean tan lentos en su ejecución que nos conduzcan a la parálisis computacional. En estas situaciones, como ocurre a menudo en la investigación psicológica o informática, no hay más remedio que recurrir a la utilización de heurísticos que, sin dar como resultado elecciones maximalmente óptimas, eviten esa parálisis manteniendo ciertos criterios normativos de racionalidad. Interesa, por tanto, que cuando menos las tareas sean realizables computacionalmente y mínimamente racionales.

En la teoría defendida por Chermiak se asume la necesidad de un uso mínimo de ciertas partes de la lógica, mientras que al mismo tiempo se debilita su rigidez al tener en cuenta lo limitado de los recursos disponibles por los agentes racionales. Se tienen en consideración los errores y falacias en el curso de los razonamientos, pero manteniendo a su vez ciertas prescripciones mínimas de orden lógico. Nuestro planteamiento, sin embargo, pretende ir más allá en cuanto a la propia idea de racionalidad. No es que se ponga en duda, como decíamos más arriba, que los agentes racionales dispongan de unas capacidades inferenciales mínimas, sino más bien que dadas sus limitaciones cognitivo-computacionales, así como los intereses específicos en la resolución de un problema para la que sólo es relevante una parte de la información disponible, existe la necesidad de controlar esa *capacidad inferencial* con el objetivo de que el razonamiento no carezca de sentido y, por supuesto, sea efectivo. Conjuntamente, si tenemos en mente la modelización del razonamiento de agentes racionales en general (tanto naturales como artificiales), no parece del todo adecuado hacer excesivo hincapié en un estudio de la teoría de la

racionalidad que se reduzca en gran parte a los resultados de los experimentos psicológicos, dado que es posible entender el razonamiento y la racionalidad desde una perspectiva más amplia. La base sobre la que fundamentemos la idea de racionalidad no tiene por qué quedar reducida a los medios que usualmente manejan los agentes humanos en función de su arquitectura cognitiva, sino que es posible servirse de los recursos computacionales que la informática nos ofrece para comprender ese concepto. Pensemos, por ejemplo, en procesos inteligentes de *razonamiento* que para su realización exijan una arquitectura cognitivo-computacional específica diferente a la humana. Además, tal y como hemos visto en una sección anterior, cuando se habla de razonamiento ordinario no hay por qué limitarlo al razonamiento práctico como *implícitamente* hace Cherniak (1986), sino entenderlo como un proceso con dos vertientes (razonamiento teórico y práctico) que se complementan.

5. BASES PARA EL RAZONAMIENTO PRÁCTICO

En esta sección señalamos las líneas generales que a nuestro juicio deberfan seguirse en la construcción de una teoría del razonamiento práctico, de modo que esta última pudiera ser usada para proporcionar las bases sobre las que pudiera sustentarse la modelización de un agente racional artificial.

Independientemente del papel que puedan jugar los contenidos en el tratamiento de los procesos cognitivos generales, mantenemos que los contenidos semánticos en un sentido clásico del término no inciden sobre la modelización del razonamiento práctico. Es decir, podemos construir la estructura del proceso sin tener por ello que considerar las cuestiones unidas al término clásico de *intencionalidad* en el sentido defendido por Brentano o Husserl, o a las cuestiones de *contenido cognitivo o proposicional* (correspondientes a la noción de *Sinn* en Frege o de *Meaning* en Russell). De

hecho, gran parte de los enfoques de la IA establecen un punto de partida similar: Pollock (1989), Levesque (1990), Konolige (1991), Pollack (1991). En un proceso de razonamiento no nos interesa, en principio, la carga de contenido implícito sobre las creencias que esté manipulando el agente. Lo realmente destacable es que ese agente posee determinadas creencias que por deliberación le llevan a interesarse por un propósito concreto que tratará de satisfacer a través de una acción como parte de un plan. Y no el significado de las mismas. En otros términos, lo que modelizamos es la estructura sintáctica que surge del razonamiento derivado a partir de las intenciones de un agente, que probablemente le empujarán a llevar a cabo una acción como parte de un plan para alcanzar sus propósitos. Esto nos lleva a considerar una teoría de la intención y de la planificación en el ámbito del razonamiento práctico.

5.1 Intención, Planificación y Razonamiento Práctico

Un agente racional tiene creencias que reflejan el mundo que le rodea y ciertos deseos acerca de su entorno que en ocasiones desea cambiar. También debe poseer algunos estados internos '*doxásticos*' que deben estar en buena correspondencia con el mundo que le rodea, y algunas disposiciones para formar sus pretensiones y deseos acerca de ese entorno, y además debe tener un mecanismo en el cual sus deseos pueden ponerse en marcha, con el fin de cambiar su entorno inmediato de modo que vaya adquiriendo nuevos estados doxásticos que interactúen con sus disposiciones conativas disolviendo con ello sus primeros deseos. El razonamiento práctico es el que determina las cuestiones a considerar por el razonamiento teórico, con el fin de llevar a cabo determinados propósitos como parte de un plan en el cual un agente está interesado, tomando a su vez en consideración determinadas restricciones en cuanto a recursos a la hora de enfrentarnos con problemas cotidianos concretos (utilización de heurísticos).

Gran parte del trabajo en IA ha sido dirigido principalmente a problemas concretos de procesamiento de datos y a la teoría de la decisión. Pero incluso estos proyectos de IA podría decirse que tienen una conexión íntima con la teoría de la racionalidad. Las conclusiones inferidas y las decisiones tomadas por los sistemas de razonamiento práctico deben ser *razonables*. El problema es dar al menos una explicación parcial de cómo un agente racional responde a una cuestión práctica (Audi (1989)). Si tenemos en cuenta que la racionalidad es un concepto antropomórfico, podemos juzgar la racionalidad de las conclusiones sencillamente usando nuestras intuiciones. Los seres humanos razonan, y juzgan tanto su propio razonamiento como el de sus semejantes en términos de su propia convicción acerca de lo que es o no es un buen razonamiento. Por supuesto, pueden utilizar sus propios juicios intuitivos de racionalidad en el intento de construir caracterizaciones más objetivas de la racionalidad en áreas específicas del razonamiento (Pollock 1991). (Por ejemplo, la caracterización teórico modelista de validez en lógica de primer orden es a menudo vista como una caracterización objetiva de la capacidad inferencial racional en esta área limitada). Pero no podemos perder de vista el hecho de que el último criterio a utilizar para juzgar una estandarización objetiva tal de la racionalidad debe basarse en las intuiciones humanas.

Muchos de los modelos de IA para el reconocimiento de planes han sido construidos sobre la base de técnicas tradicionales para la generación de planes, en las que los planes de un agente son simplemente vistos como *recetas* para alcanzar un propósito, aparte e independientemente del papel que los planes pueden jugar en la vida mental del agente. Em particular, los sistemas para el reconocimiento de planes, como los sistemas de generación de planes clásicos, han ignorado las dificultades que presentan los entornos cambiantes para agentes con limitación de recursos — y han ignorado los modos en que la planificación puede ayudar a los agentes a enfrentarse con estos problemas (Pollack 1991).

Por otra parte, no deberíamos centrarnos en teorías para las cuales el razonamiento práctico consiste tan sólo en sopesar razones del tipo creencia-deseo para cursos de acción que compiten entre sí — este último es un punto de vista que ha primado en filosofía (Davidson (1980), Bratman (1985)), y que ha caracterizado a gran parte del trabajo en la teoría de la decisión (Savage (1972), DiFinetti (1975), Skirms (1980), Jeffrey (1983)) — sino en un tipo de razonamiento práctico en el que se tenga en cuenta el papel de las *intenciones prioritarias* como mediadoras en los procesos de razonamiento. Bratman (1987) propone la teoría de la *planificación de las intenciones*. De acuerdo con esta teoría, las intenciones de mayor prioridad son elementos típicos en planes parciales más amplios para acciones futuras. Y estos planes parciales dan forma a acciones y planes posteriores, ayudándonos a coordinar nuestras actividades a lo largo del tiempo, y por tanto ayudándonos a extender la influencia de la reflexión racional presente en las acciones futuras.

La introducción en nuestro modelo de las intenciones prioritarias nos ayuda a la hora de pensar acerca de las normas de conducta personales. Normalmente, una intención dirigida al futuro está en relación con alguna ocasión particular, una ocasión que es más o menos especificada por la intención particular. Podemos tener la intención de actuar cuando un tipo de ocasión esté presente, o una intención para actuar de un cierto modo sobre una base regular. Esto sugiere que la teoría de la planificación de la intención puede ser extendida de modo que podamos incluir normas de conducta viendo a las mismas como intenciones que son generales en el siguiente sentido: generales con respecto a sus ocasiones de ejecución (Bratman 1991). Paralelamente, un sistema para el reconocimiento de planes necesitaría tener en cuenta la asunción de que los planes de un agente tienden a ser consistentes entre sí. Y esta asunción está justificada por la afirmación de que un modo de recortar el razonamiento en un agente con limitación de recursos es

considerar sólo aquellas opciones que son consistentes con sus planes existentes, una afirmación que surge a partir de la consideración de problemas planteados por la limitación de recursos (Bratman 1987 e 1988).

Una mejor comprensión de las funciones que desempeñan los planes en la vida mental de un agente, y de las restricciones sobre los planes que aparecen a raíz de estas funciones, puede beneficiar no sólo nuestros modelos para la generación de planes, sino también nuestros modelos para el reconocimiento de planes. Pollack (1991) asume las afirmaciones de Bratman (1987) donde se argumenta que es racional para agentes con limitación de recursos formar y adoptar planes, puesto que de este modo se recorta la cantidad de razonamiento práctico necesario y, por tanto, resulta más fácil hacer frente a los respectivos entornos dinámicos.

5.2 Heurísticos como Complemento del Razonamiento Práctico

Pollack (1991) describe la estrategia que denomina '*sobrecarga*' mediante la cual un agente hace uso de sus planes existentes en razonamiento dirigido a fines y pasa por alto el proceso de deliberación que sería en otro caso necesario. Una apreciación de este proceso, al igual que otras estrategias para el razonamiento práctico, puede favorecer tareas tales como el reconocimiento o la generación de planes. También Pollock (1992) incide en la necesidad del uso de estrategias o heurísticos en el razonamiento práctico. A la conjunción de razonamiento práctico y razonamiento teórico Pollock le da el nombre de *intelección*. No obstante, señala que esto trae consigo una construcción estrecha de la racionalidad. La dificultad estriba en que el razonamiento es demasiado lento. En su opinión muchos aspectos del razonamiento deberían ser esencialmente seriales. Los seres humanos ganarían en velocidad si su soporte duro, inherentemente lento, procesara en paralelo. Gran parte del razonamiento no puede ser realizado en paralelo, de modo que

el razonamiento humano incluye muchos proyectos "*no intelectuales*" (procesos que no forman parte de *intelección*) que también tienen que ver con creencias y acciones. Pollock se refiere a estos tipos de procesos intelectuales como '*quick and dirty*'. También señala que estos módulos en los humanos no están exclusivamente vinculados a funciones motoras. La evidencia psicológica sugiere insistentemente (Tversky and Kahneman 1974) que la mayoría de las inferencias tanto inductivas como probabilistas de a diario son llevadas a cabo siguiendo este tipo de módulos.

El razonamiento probabilista adecuado es en muchos casos computacionalmente *irrealizable* y de este modo los humanos parecen echar mano de procesos tales como la *representatividad heurística*, que a menudo resulta incompatible con el cálculo de probabilidades. Esto no quiere decir, sin embargo, que sea irracional hacer uso de tales métodos de aproximación. De hecho, la alternativa del razonamiento explícito es demasiado lenta para muchos propósitos prácticos. Parece que *intelección* puede en principio tratar con cualquier tipo de situación, pero es un proceso lento. En situaciones complicadas puede que no podamos aplicar los módulos, en cuyo caso no tenemos más elección que comprometernos con el razonamiento explícito acerca de la situación. En otros casos, los humanos aceptamos la salida de los módulos a menos que tengamos alguna razón explícita para no hacerlo. Este planteamiento está en concordancia con algunas de las estrategias (o heurísticos) para razonamiento dirigido a intereses propuestas por Bratman et al. (1988), Pollock (1989), Pollack (1991), entre otros.

Por lo visto en esta sección, se pone de manifiesto que los modelos para el razonamiento práctico en IA pueden ser útilmente informados por teorías filosóficas o, en otros términos, que algunas teorías filosóficas pueden ser realizadas computacionalmente. La teoría de la intención de Bratman ha llevado a especificaciones de arquitectura de alto nivel para un agente artificial (Bratman et al. 1988), además de la implementación de un agente prototipo

(Pollack and Ringuette 1990). Igualmente diferentes propuestas ligadas a la teoría de la planificación están siendo introducidas en el razonador automatizado OSCAR (Pollock 1992). Si nuestro propósito es que los agentes racionales actúen inteligentemente, necesitamos razonar acerca de la conducta que intentan llevar a cabo. Por lo tanto, el estudio del uso de planes debería ser un requisito *prima facie* tanto en inteligencia Artificial como en Psicología.

ALGUNAS CONCLUSIONES

En este artículo hemos analizado la importancia de la lógica como instrumento indispensable para el razonamiento del sentido común. Primero, hemos presentado una visión general de las limitaciones del denominado enfoque logicista en el uso de la lógica como lenguaje representacional para teorías del sentido común. Más adelante hemos establecido una clara distinción entre el razonamiento teórico y el razonamiento práctico desde la perspectiva de una teoría general del razonamiento ordinario, lo que nos ha servido para criticar no sólo los que se denominan razonadores ideales sino también algunos aspectos del razonamiento basado sobre la idea de racionalidad mínima. Finalmente, hemos esbozado las líneas generales de la incidencia que tiene para el razonamiento práctico tanto el estudio de los planes como la teoría de la intención. Sin olvidar la necesidad de utilización de técnicas heurísticas para agentes racionales con limitación de recursos que se enfrentan a problemas de razonamiento en entornos dinámicos.

En definitiva hemos visto que es posible la síntesis entre filosofía y teoría del cómputo para algunos casos de razonamiento práctico. No tiene demasiado sentido hacer un estudio del razonamiento ordinario sin haber establecido adecuadamente su fundamento epistémico, tanto si hablamos de agentes naturales como artificiales. ¿Para qué nos interesan teorías sobre el razonamiento que adolecen de la más mínima aplicabilidad, al margen del puro

interés teórico? ¿Qué sentido tienen las teorías psicológicas sobre razonamiento que hablan del mismo como si de un proceso único se tratara, sin establecer las diferencias teóricas oportunas entre las distintas maneras de razonar?

Otra cuestión a destacar es que hay ciertos procesos *racionales* de inferencia que por los motivos indicados no están presentes en los agentes humanos, o que su actuación se ve perjudicada por la condición psicológica específica del agente en cuestión. Psicología e Inteligencia Artificial tienen problemas comunes en el área del razonamiento que debemos solucionar, para lo cual hay que hacer un mayor hincapié si cabe en las teorías de la racionalidad subyacentes con el objetivo de fijar definitivamente los patrones metodológicos y epistemológicos que nos guíen en la definición de los modelos computacionales.

REFERENCIAS

- AUDI, R. (1989), *Practical Reasoning*. (London, Routledge).
- BRATMAN, M.E. *et al* (1988), Plans and resource-bounded practical reasoning. In R. Cummins and J. Pollock (eds.), *Philosophy and AI*. (Cambridge MA, MIT Press).
- . (1985), Davidson's Theory of Intention. In: B. Vermazen y M.B. Hintikka (Eds.), *Essays on Davidson: Actions and Events*. (New York, Oxford University Press).
- . (1987), *Intention, Plans and Practical Reason*. (Cambridge MA, Harvard University Press).
- . (1988), Intentions and Personal Policies. (Report No. CSLI-88-118, Stanford, CA.).

- . (1991), *Practical Reasoning and Acceptance in a Context*. (Report No. CSLI-91-158, Stanford, CA.).
- CUMMINS, R. (1991), *Meaning and Mental Representation*. (Cambridge MA, MIT Press).
- CHERNIAK, C. (1986), *Minimal Rationality*. (Cambridge, MA, MIT Press).
- COHEN, P.R. and LEVESQUE, H.J. (1987), *Persistence, Intention, and Commitment*. (SRI International. Technical Note 415).
- DAVIDSON, D. (1980), *Essays on Actions and Events*. (New York, Oxford University Press).
- DENNETT, D.C. (1978), *Brainstorms*. (Cambridge MA, MIT Press).
- DiFINETTI, B. (1975), *Theory of Probability*. (New York, Wiley).
- DOYLE, J. (1988), *On Universal Theories of Default*. (Carnegie Mellon Computer Science Department, Technical Report CMU-CS-88-111).
- . (1988a), *Artificial Intelligence and Rational Self-Government*. (Carnegie Mellon Computer Science Department, Technical Report CMU-CS-88-124).
- HINTIKKA, J. (1962), *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. (Ithaca NY, Cornell University Press).

- ISRAEL, D. (1985), A Short Companion to the Naive Physics Manifesto. In J. Hobbs & C. Moore (eds.), *Formal Theories of the Commonsense World*. (Norwood NJ, Ablex Publishing Corporation).
- JEFFREY, R. (1983), *The Logic of Decision*. (Chicago, University of Chicago Press).
- JOHNSON-LAIRD, P.N. (1988), *The Computer and the Mind*. (London, Fontana Press).
- KOLATA, G. (1982), How can Computers get Commonsense. *Science* 217, 1237-1238.
- KONOLIGE, K. (1991), Intention, Commitment and Preference. (Menlo Park CA: SRI International (Manuscrito)).
- LEVESQUE, H. (1990), All I Know: A Study in Autoepistemic Logic. *Artificial Intelligence* 42, 263-310.
- MINSKY, M. (1975), A Framework for Representing Knowledge. In J. Haugeland (ed.) (1981), *Mind Design*. (Cambridge MA, MIT Press).
- PEQUENO, T. & BUCHSBAUM, A. (1991), The Logic of Epistemic Inconsistency. In J. Allen, R. Fikes, E. Sandewall (eds.), *Principles of Knowledge Representation and Reasoning. Proceedings of the Second International Conference*. (San Mateo CA, Morgan Kaufmann).
- POLLACK, E.M. & RINGUETTE, M. (1990), Introducing the Tileword: Experimentally Evaluating Agent Architectures.

In *AAAI-90, Proceedings of the Ninth International Conference on Artificial Intelligence*. (Boston, MA.).

POLLACK, E.M. (1991), Overloading Intentions for Efficient Practical Reasoning. (Report No. CSLI-91-149).

POLLOCK, J. (1987), Defeasible Reasoning. *Cognitive Science* 11, 481-518.

———. (1989), *How to Build a Person*. (Cambridge MA, MIT Press).

———. (1991), A Theory of Defeasible Reasoning. *International Journal of Intelligent Systems* 6, 33-54.

———. (1992), *Cognitive Carpentry: a Blueprint for How to Build a Person* (forthcoming).

SAVAGE, J.L. (1972), *The Foundations of Statistics*. (New York, Dover).

SHERIDAN, F.K.J. (1991), A Survey of Techniques for Inference Under Uncertainty. *Artificial Intelligence* 5, 89-119.

SKYRMS, B. (1980), *Causal Necessity*. (New Haven, Yale University Press).

SALMON, N. (1992), On Content. *Mind* 101, 733-751.

TVERSKY A. & KAHNERMAN, D. (1974), Judgement Under Uncertainty: Heuristics and Biases. *Science* 185, 1124-1131.