



La idea de hombre bajo los presupuestos del "humanismo computacional"

Carlos Javier Blanco Martín

En este artículo vamos a tratar, si quiera de forma somera, la *idea de hombre* que, bajo presupuestos computacionalistas, asoma por doquier entre las categorías de la Psicología. Es una cuestión de hecho que el cognitivismo no ha experimentado con animales en la misma medida que el conductismo, y su relevancia en la etología y en la psicología animal es sólo una relevancia secundaria, una "adjetivación" de las técnicas y metodologías previas: el cognitivismo ha de dejar a un lado la programación de ordenadores y hacer uso de los mismos métodos de condicionamiento, las técnicas de observación en espacios naturales o artificiales y los mismos procedimientos de registro conductual. Todas esas técnicas y metodologías forman parte del acervo de una ciencia (conductual) de los animales, ciencia ésta que podrá complementarse, y así lo pide la "filosofía computacionalista", con inferencias sobre la subjetividad de los organismos, reflexiones sobre su cognición, pero el etólogo y el psicólogo de animales no pueden dejar de ser *conductistas* en cuanto a sus prácticas. La llamada "cognición animal" es siempre una reflexión teórica *a posteriori*, una especie de prolongación del debate sobre el alma de los brutos. A diferencia del estudio con humanos, en el campo de la conducta animal es mucho más fácil discernir qué es práctica (adiestramiento, observación, registro) y qué es teoría (preguntas del tipo "¿tienen *vida mental* los animales?"¹)

Lo más sustantivo del cognitivismo, por tanto, se corresponde con lo *específicamente* humano, aunque las "ciencias cognitivas" rebozen en declaraciones

¹ Grandes etólogos, como el mismo Lorenz, han sido convencidos creyentes en la *vida subjetiva* de los animales. En su artículo "¿Tienen vida subjetiva los animales?" (En Lorenz, K. , *Consideraciones sobre las conductas animal y humana*, Planeta-Agostini, Barcelona, 1984; pps. 397-411; publicado por primera vez en el *Jahrbuch der technischen hochschule*, Munich, 1963). La respuesta al título de su artículo es afirmativa. Entre otros argumentos, Lorenz aludía a la célebre conclusión *por analogía*, esto es: si yo tengo experiencia de lo mental, también otros seres *análogos* a mí la tendrán. La adscripción de estados mentales y subjetividad a los animales es un tema muy controvertido, que en este trabajo no podemos tratar de forma sistemática. Desde el punto de vista de un etólogo, surge la cuestión gradualista que, de acuerdo con la teoría de la evolución, exigiría establecer un "corte" en cierto punto de la escala filogenética, si no se quiere caer en atribuciones como las de Loeb, quien defendiera la existencia de vida mental incluso en las bacterias. Konrad Lorenz cita en este artículo a Egon Brunswick, quien apelara a la presencia de ciertos aparatos de cálculo "raciomorfos" que el hombre y los animales deberían poner en marcha en actividades del tipo de la percepción no consciente de los colores o de las formas, centros los cuales estarían subordinados a otros módulos "ánimicos" o superiores y, ciertamente, conscientes. La idea, sin duda, anticipa ciertos elementos de la moderna psicología cognitiva, incluido el estudio de la "cognición animal". Autores más modernos, como Griffin, consideran que la existencia de esa mentación en los animales es una materia para la que hay muchas pruebas a favor (Griffin, *El pensamiento de los animales*, Ariel, Barcelona, 1986).

omnicomprensivas acerca de las máquinas y las especies no-humanas, como conjunto general que forman los llamados *sistemas procesadores de información*. Lo que ha ocurrido, de hecho, no ha sido sino un regreso intensivo a la psicología humana, en contra del conductismo, escuela ésta que se ocupaba preferentemente del entrenamiento de ratas o palomas. En realidad, debido al hecho de que la cultura humana sea notablemente más compleja y polimórfica que la de cualquier otra especie, nos sorprende el "principio de exclusión" que de forma sistemática ha ido operando en el cognitivismo a lo largo de sus trayectorias metodológicas. Exclusión ¿de qué? Pues la exclusión de la *genésis* de todo un conjunto de fuerzas, cuyo poder o incidencia causal sería innegable, y que en la ontología materialista se pueden denominar *fuerzas del inconsciente objetivo*². La determinación de fuerzas o factores *culturoológicos*, especialmente, no es menor que la de otras fuerzas no antropológicas en su origen, impersonales como son las fuerzas mecánicas o gravitacionales, los cambios geológicos, las reacciones químicas etc. De hecho, el *inconsciente objetivo* se gana el epíteto de "objetivo" por lo que tiene de neutralización de las operaciones (de las subjetividades) que, sin embargo, se requieren para que la determinación culturoológica afecte a las conciencias psicológicas individuales. Sin este "principio de exclusión" la Psicología Cognitiva perdería su razón de ser y habría de convertirse en Antropología Cultural, Sociología del Conocimiento, o alguna otra de las ciencias enmarcadas ontológicamente por la idea del *Espíritu Objetivo*.

En principio, la *idea de hombre* que atraviesa el campo psicológico-cognitivo habría de quedar repudiada por nosotros, en cuanto que *idea* filosófica extrínseca a la propia psicología, una idea extragnoseológica. No obstante, la aplicación de un análisis gnoseológico y dialéctico como el presente no puede obviar la totalidad de tipo *fenomenológico* que pide desentrañarse en una crítica filosófica del cognitivismo: la mezcla de ciertas prácticas en cohabitación con ideas de clases y categorías distintas (conexión *metamérica*): conciencia, significado, intencionalidad, y la idea de hombre, que nos corresponde tratar ahora.

1. Implicaciones Ideológicas del nuevo Humanismo Computacionalista

Las objeciones estándar que la Psicología cognitiva ha recibido apuntan, casi todas, al plano metateórico de esta disciplina. El cognitivismo suele ser tildado de: *dualista, mentalista, cartesiano, innatista, formalista, mecanicista, solipsista, teleológico, metafórico*, etc.

Por ejemplo, para muchos autores, el rasgo más inaceptable del cognitivismo es su *dualismo*: aceptar una mente separada del cuerpo. Según Juan Bautista Fuentes y Francisco Robles³:

"Repárese, en efecto, en que la cibernética en general, y por lo tanto las dos versiones [los autores se refieren a la Inteligencia Artificial [I.A.] y a la simulación Cognitiva] del modelo computacional aquí reseñadas, interpretan la conducta inteligente y el pensamiento como una actividad computable y sintáctica separada del cuerpo. De esta manera se obvia la génesis de todos

² El materialismo ontológico no ha dejado de reconocer un *género especial de materialidad* (M2) que connota todos aquellos fenómenos psicológicos o etológicos en los cuales, si se dan relaciones *reflexivas*, estas no son consideradas originarias sino, por el contrario, toda *reflexividad* de la conciencia va a proceder de la objetivación del lenguaje y de otras formaciones sociales de naturaleza supraindividual.

³ Fuentes, J.B. y Robles, F. "Psicocibernética", en Reyes, R. (ed.), *Terminología científico-social*, Anthropos, Barcelona, 1989; p. 89.

los procesos y con ello su mismo significado. A nuestro juicio, el pensamiento se da en el marco de la inteligencia, y esta se nos presenta como una capacidad o función de adaptación al medio que, desde luego, sólo puede explicarse en términos de conducta y aprendizaje. La configuración del aprendizaje se genera a partir de dos factores de singular relevancia: la experiencia y la acción, factores éstos, propios de un legítimo sujeto psicológico e íntimamente ligados a la estructura que sustenta la función del pensamiento, esto es, un cuerpo capaz de tener noticia experiencial de su ambiente y de operar conductualmente en él desde dicha experiencia".

Lejos de analizar el dualismo, en su amplia variedad de versiones, únicamente nos corresponde señalar que el regreso de estas posturas no podría entenderse sin advertir que Noam Chomsky lo hubiera reintroducido. La alternativa al conductismo volvió sus ojos a un lingüista especialmente beligerante con Skinner, el conductista que representaba una posición anti-dualista. Chomsky defendía abiertamente la raigambre cartesiana de su antropología (filosófica). Dualismo el cartesiano que conviene no interpretar de forma errónea. Según un autor que se ha ocupado del problema del dualismo en el filósofo francés, Richardson, para salir al paso del problema de la *heterogeneidad* cuerpo/mente en Descartes, éste había resuelto que no era el sujeto humano como *cosa simple*, una persona que fuera el sujeto de atribución de diversas cualidades, entre las que figurarán las *facultades* del sentir, el imaginar, etc. El objetivo que Descartes se había propuesto explicar era aquel conjunto de fenómenos que se resistían a una explicación física⁴: " (...) *the whole goal of dualism is to introduce some sort of explanation for what is thought to be physically (or, for Descartes, mechanically) inexplicable*". Así pues, para Descartes no habría problemas lógicos desde el momento en que se defendía que la *interacción psicofísica* era una clase completamente distinta a la de la interacción entre cosas físicas⁵:

"The problem of how two such heterogeneous things as mind and body could interact is, as Descartes so clearly saw, no problem at all. The illusion of a problem has two sources: the presumption that psychophysical interaction must be similar in kind to physical interaction, and the conviction that psychophysical dualism must lead to a naïve attempt to attain an exhaustive segregation of properties into those that are attributable to the purely physical and those that are attributable to the purely mental".

No sólo resulta inaceptable la idea de una mente separada del cuerpo, y un conocimiento desligado de la experiencia, como denunciaban Fuentes Ortega y Robles, y muchos otros críticos. Por un lado se dice que el Procesamiento de la Información [P.I.] el alma se ha desprendido del cuerpo dentro de la concepción cognitiva del *sujeto* que, definido como un *sistema procesador de símbolos*, ya no puede ser un espíritu (un alma racional indivisible, una cosa simple). Así, el espiritualismo en Psicología Cognitiva no puede defenderse seriamente sin caer en contradicción, puesto que la propia empresa del Procesamiento de la Información consiste en *destruir* esa unidad del espíritu, dividirla en partes átomas. El componente más ideológico de quienes se autorrepresentan como "cognitivos" ha de consistir, no ya en creerse "científicos de la mente", puesto que su trabajo viene a consistir en *criticar esa mente* (descomponerla, destruirla), sino que reside en el tema del *carácter activo* del sujeto. Que el hombre es un sujeto eminentemente activo ha sido, quizás, el dogma más firmemente defendido por los psicólogos de orientación cognitiva. Y puede

⁴ Richardson, R.C. , "The 'scandal' of cartesian interactionism", *Mind*, (1982), **XCI**, 20-37; p. 36.

⁵ *Ibid*

que no exista un término cuya factura sea más oscura e ideológica, especialmente cuando dicho concepto de *actividad* se opone al de *pasividad*, queriendo reflejar así la nota o marca distintiva entre una psicología cognitiva y un behaviorismo "estímulo-respuesta". La predicación de la *actividad* por parte del sujeto, se ha ido distribuyendo en predicaciones de *actividad* en procesos o facultades especiales de ese sujeto: dicese, entonces, que la "percepción es un proceso activo", "la memoria es activa", etc. Con ello quiere significarse, frecuentemente, que el sujeto *aporta* algo. Sea el sujeto psicológico un sujeto "activo" (metodologías β -operatorias), sea el sujeto, un "procesador de símbolos no interpretados" (metodologías α -operatorias), tenemos, en cualquier caso, una Psicología cuyo aspecto es el de un Jano bifronte, una disciplina con doble plano operatorio, y con un apretado tránsito entre ambos planos, por cierto. La ideología de la *actividad* esencial del hombre, propugnada, entre otros, por un Neisser o un Bransford, choca, por ejemplo, con el neo-conductismo-mecanicista de un Minsky. Por otro lado, Fodor y Simon, vendrían a representar posturas irenistas o reconciliadoras, sin desprenderse en exceso del *desideratum* de una explicación de la conducta en términos mecánico-simbólicos. Ahora bien, la forma de representarse que los psicólogos tienen de lo *activo*, de la actividad (en correspondencia, por ejemplo, con la *creatividad* que Chomsky defendía en su antropología filosófica), no puede coordinarse, sino a través de caminos algo tortuosos, con el mantenimiento de la Psicología Cognitiva en un plano operatorio, esto es, la apelación a un sujeto operatorio en el campo. Así, es frecuente la asimilación que de la idea de "actividad" los autores ejercen con algo semejante a la *volición* (originalmente, en la fisiología de los movimientos aún llamados "voluntarios", esto es, controlados por el córtex), o el *propósito* (como hemos visto en Tolman y en Simon), o con la *intencionalidad* (Searle, Dennett, Fodor). Pero, por encima de todas esas confusiones y asimilaciones, debemos distinguir los dos planos en que la Psicología del P.I., forzosamente, se mueve: por un lado hay que reconocer e incorporar un sujeto, por otra parte hay que destruirlo o resolverlo en componentes no-operatorios. La situación es que, si bien el plano β está presente, la pretensión o *desideratum*, consistente en abrazar o acceder un plano α es más que discutible. El pretendido carácter *subpersonal* (tomando prestada la palabra a D.C. Dennett) de esos procesos psíquicos no-operatorios, ha de ser conectado con el *inconsciente objetivo* al que hacíamos referencia más arriba. La elaboración de modelos "en circuito cerrado"⁶ que simulan o reproducen la actividad β -operatoria del hombre, se hace a costa de la férrea neutralización de unos determinantes que, acaso en un plano causal-ontológico, van a resultar de imprescindible referencia. Estas conexiones entre lo *ontológico* y lo *gnoseológico*, nos parecen de todo punto inexcusables, por lo que toca a una gnoseología *fisiológica* (no meramente *anatómica*). Y es así que podemos detectar las regulaciones que ciertos *principios gnoseológicos* del cognitivismo están ejecutando sobre las configuraciones del campo psicológico. Un *principio de los términos* podrá observarse en la fabricación de una clase de identidad que agrupa a los diversos sistemas procesadores de información (que en alguna parte del texto hemos denominado "principio cibernético"): hombres, máquinas y animales son, todos ellos, y por lo que respecta al procesamiento, máquinas informáticas. Por otra parte, y a lo que a los sujetos humanos les toca, la Psicología Cognitiva como tal no es una Psicología Diferencial ni Patológica, y en este sentido obvia las diferencias intersubjetivas (de personalidad, de inteligencia, de raza) o las variaciones respecto a alguna norma. Aunque ahora ya se dan muchas confluencias (psicopatología, psicología infantil y testología "interpretadas" desde los modelos del P.I.), pero lo que caracteriza a la Psicología

⁶ Pérez Alvarez, M., "Moda, mito e ideología de la psicología cognitiva", *Psicólogos. Papeles del Colegio*, (1985), 4, 45-52.

Cognitiva *qua tale* es el estudio del "hombre adulto", "normal", y con la "base de datos ya formada"⁷.

Para los propósitos de una demarcación entre lo Psicológico-cognitivo y lo culturoológico, cobran especial relevancia los *principios de las relaciones*. Aun cuando hay relaciones sinectivas entre el sujeto del P.I. y el medio social, histórico y cultural en que se desenvuelve, lo cierto es que podemos señalar la presencia de un *principio gnoseológico* que consiste en neutralizar la génesis "extrasomática" o "supraorgánica" de la conciencia, para encapsular la conciencia, reorganizándola sus términos componentes mediante modelos que se pretenden *isomórficos* de la misma, modelos que constan de una suerte de malla relacional, construida artificialmente, para que transcurran o se filtren precisamente las relaciones entre términos que desbordan esa malla. Por tanto, si la "validez ecológica" se llevara a sus últimas consecuencias, abriendo las puertas y ventanas de la casa cognitiva, para que los vientos de la historia y la cultura la penetren, el sistema cognitivo y la misma disciplina psicológica-computacional, se disolverían definitivamente en el seno de esas otras ciencias antropológicas.

Pero sucede que el cognitivismo psicológico, pese a "tener la válvula cerrada" respecto a ese medio cultural, no puede dejar de ser una *antropología filosófica*, una especie de "filosofía del sujeto", por más que la asepsia de los algoritmos, los sistemas de símbolos o de los conexionismos cuasi-neurales pretendan hacer desvanecer ese sujeto (esas *ideas* sobre el sujeto).

Y es así como llegamos al hecho de que las diversas modulaciones dadas ante una *idea de sujeto* no están exentas de implicaciones ideológicas. Desde un análisis gnoseológico en sentido estricto, tales implicaciones no pueden contemplarse sino de forma externa, oblicua. Harían falta datos procedentes de una disciplina sociológica, y no gnoseológica. No obstante, y en consonancia con lo que llevamos dicho respecto a las *prácticas* cognitivistas (que no son muy diferentes de otras prácticas psicológicas), hay en el plano de la *metateoría* (plano filosófico e ideológico) una serie de aspectos que deben contemplarse, siquiera, a vuela pluma.

Ante todo, nos topamos con la pregunta de si está surgiendo algo así como un "nuevo humanismo", no ya cibernético, pero un humanismo sin duda nuevo, si bien deudor de la cibernética: se trata de un humanismo cuyos valores, dicen los propios autores, residen en la libertad y autonomía de un sujeto activo. Ese sujeto activo, racional, responsable, etc., es compatible -no obstante- con la descripción mecánico-simbólica de su cognición *subpersonal*, totalmente ajena a libres arbitrios, responsabilidades morales y cosas así.

Margaret A. Boden es una autora especialmente conocida por un libro, *Artificial intelligence and natural man*⁸, que trata en abundancia estos aspectos relativos a la emergencia de lo que podríamos llamar un nuevo "humanismo computacional" que, a diferencia del defendido por Wiener y los cibernéticos, aboga decididamente por un sujeto mentalistamente considerado. Boden considera imprescindible salir al paso de la contradicción entre "humanismo" y "mecanicismo"⁹:

"(...) aquí sólo anticiparé que la contradicción aparentemente total e ineludible entre los puntos de vista humanista y mecanicista se puede evitar. (...), es posible adherirse sin reservas al acento humanista en la subjetividad en la

⁷ Valle Arroyo, F. "Cognitivismo", pps. 128-131, en Reyes, R (ed.), *Terminología científico-social*.

⁸ Boden, M.A., *Artificial intelligence and natural man*, The Harvester Press, Brighton, 1977. Versión española: *Inteligencia artificial y hombre natural*, Tecnos, Madrid, 1984.

⁹ Boden, M.A., *Inteligencia artificial*, p. 506

psicología sin con ello atentar contra la firme insistencia mecanicista en que los fenómenos psicológicos dependen de procesos causales cerebrales".

La I.A. y la Psicología del P.I. pueden ser, por una parte, mecanicistas, formalistas, reduccionistas, deterministas, etc. , y ello, sin violar o atentar ideas como la libertad de la persona, la dignidad del hombre, la conciencia, la mente, la teleología, etc. Semejantes amalgamas sólo pueden ensayarse bajo una filosofía vulgar y de tinte marcadamente extemporáneo y neoescolástico¹⁰ 10 :

"La psicología debe dar una explicación del significado, o intencionalidad, intrínseco a la vida mental, y debe reconocer las más amplias implicaciones teóricas de semejante explicación. Atribuir significado a un fenómeno es poner al sujeto psicológico como fuente activa de ese fenómeno; correlativamente, es considerar el fenómeno como el objeto psicológico de algún pensamiento o acción de la agente del caso. En una ciencia humana es central e ineludible esa polaridad conceptual entre sujeto y objeto. Cada uno de los conceptos psicológicos más específicos aplicados a los seres humanos --desde 'pasión' a 'percepción', desde 'libertad' hasta 'falibilidad' --se ha de interpretar a la luz de esta distinción psicológica básica".

Así , junto a la elaboración de programas, la Inteligencia Artificial y la Psicología Cognitiva precisan de una Antropología Filosófica, una metafísica humanista que pretende operar en torno a la idea de "hombre", y otras ideas conexas, como "libertad", "pasión". Esa antropología no abandona los dualismos clásicos, el sujeto enterizamente enfrentado al objeto, la intencionalidad sustantivizada *versus* la mecánica-simbólica de nuestros "programas cerebrales", etc. El mentalismo antropológico-filosófico orquestado en torno a una *idea de sujeto*, cohabita con la programación de las máquinas, que es, según expresión de Zenon Pylyshyn, "*un lenguaje técnico con el cual disciplinar la imaginación de cada uno*"¹¹. El nuevo humanismo mecánico considera que la *subjetividad* es su columna vertebral, y defiende la legitimidad científica de la empresa cognitivista, que estudia la *mentalidad* del hombre sirviéndose de las máquinas. Boden, de nuevo¹² 12 :

"Aunque he defendido la práctica de mostrar analogías entre mentes y máquinas, he estado de acuerdo en que las categorías de subjetividad, significado y propósito, tal como se entienden habitualmente, sólo se pueden atribuir a los artefactos en un sentido secundario, derivando su justificación última de la habilidad e intereses de la artífice. Y he reconocido que en la mente de cualquiera que proponga extender estas categorías para incluir sistemas cuya fuente no es la evolución biológica sino la manufactura tecnológica, debe haber problemas ideológicos básicos" .

De acuerdo: Margaret Boden reconoce que profesar este "nuevo humanismo" conlleva compromisos y militancias ideológicas. Podríamos preguntarnos si este "humanismo mecánico", esta antropología de los sujetos *procesadores de información* no es, antes que una filosofía (aunque de factura metafísica) y más que cualquier otra cosa, una *ideología*, en lo que tiene de *conjunto de relaciones erróneas*, ortogramas,

¹⁰ *Op. cit.* p 504.

¹¹ Citado en Boden, M.A. , *op. cit.* p. 511. La referencia original en Pylyshyn, Z. "Script processing in attitude formation and decision making", en Carroll, J.S. y Payne, J.W. (eds.), *Cognition and social behavior*, LEA, Hillsdale, 1976.

¹² Boden, M.A. , *op. cit.* p. 546.

en expresión de Bueno¹³. La antropología mentalista del sujeto *procesador* será, pues, un "paquete" de ortogramas (*falsas conciencias*) a modo de amasijo de materias formalizadas que la conciencia (*falsa*) desprende a resultados de sus operaciones, conciencia que perpetúa sus trayectorias erróneas por una pérdida efectiva de los "mecanismos de corrección".

Dedicar un capítulo a la crítica o destrucción de la *idea de hombre* como núcleo ontológico en torno al cual girase el cognitivismo, desborda los límites de un análisis gnoseológico, pero la crítica filosófica, sin duda, va a requerir como indispensables los resultados que desde la Gnoseología propendan hacia la disolución de ese núcleo aglutinador, en torno al que va a gravitar buena parte de lo que se llamarán "teorías psicológicas", en tanto que *teorizaciones acerca de una idea de sujeto*. Desde el materialismo, no podemos aceptar tan alegremente el bloqueo o segregación de las *conexiones* (o relaciones) que propiamente conciernen al plano del *Espíritu Objetivo*, sin perjuicio de que sustancial u ontológicamente, los términos de aquél sean coincidentes, *sean los mismos* términos, que desde el plano del *Espíritu Subjetivo* van a coordinarse de otra forma. En el cotejo que Gustavo Bueno ofreciera, años ha, entre la *Filosofía del espíritu objetivo* de Hegel y los *Grundrisse* de Marx¹⁴, se señalaba el modo en que Marx había ejercido una *inversión* de un mismo campo ontológico, del que vamos a conservar el nombre de *Espíritu Objetivo*, el mismo campo del que Hegel se ocupara y que hoy se puede vincular directamente con categorías que, bajo una denominación más actual, no son sino las de *Cultura, Sociedad, Mercado*, etc. Categorías que, por tanto, no pueden agotar en su particularidad la *totalización* ontológica que es la construcción del *Espíritu Objetivo*. Bueno señala los antecedentes de esa idea, algunos de los cuales eran el *entendimiento agente y material* de los averroistas, el *paideuma* de Frobenius, lo *superorgánico* de Kroeber, etc. La potencia de dicha construcción ontológica, a la que Marx llegó *vía* Hegel mediante pasos decisivos, no ha de obviarse en un análisis como el presente, en especial cuando se detecta que, bajo el formato de un "humanismo" computacional, o de una antropología filosófica, quiere rehabilitarse el *sujeto* sustancial, portador de predicaciones: libre, moral, responsable, activo, cognoscente, etc., habida cuenta del *principio de las relaciones* que, sistemáticamente, opera en los modelos computacionales de la mente: la segregación de las conexiones o relaciones dadas en el plano del *Espíritu Objetivo*. Desde la tradición del materialismo, el *Espíritu Objetivo* no puede quedar bloqueado, no sólo porque *antecede abstractamente* al *Espíritu Subjetivo*, ni porque también es *posterior*¹⁵.

De esta manera, la psicología cognitiva, cuando estudia o fabrica modelos en los que, inevitablemente, se requieren en un plano ontológico el tipo de *relaciones* que atañen a un *Espíritu Objetivo* (inevitablemente, en especialidades como la psicolingüística, la formación y almacenamiento de conceptos, o la resolución de problemas abstractos) podría reconstruirse como una disciplina *gnosológicamente* diferente de la Culturología o la Historia, por ejemplo, mediante la neutralización de un plano colectivo y objetivado de relaciones, pero también por medio del "filtrado" que esas mismas conexiones objetivas han de conocer a la escala de cada individuo, una escala la psicológica-cognitiva, que ha venido conceptualizándose como *subpersonal* o simbólica (por oposición a la organísmica-integral del conductismo), pero que no es sino la modelización (isomórfica, en el mejor de los casos) de vínculos ya instaurados por la estructura de la cultura y por el desarrollo de la historia, vínculos, muchos de ellos, de los que el sujeto no cobra conciencia (de ahí la importancia de la idea

¹³ Bueno, G. Cuestiones quodlibetales sobre Dios y la religión, Mondadori, Madrid, 1989.

¹⁴ Bueno, G. , "Los 'Grundrisse' de Marx y la 'Filosofía del espíritu objetivo' de Hegel, *Sistema*, (1974), 4, 35-46.

¹⁵ Bueno, G., "Los Grundrisse de Marx...", p. 37.

ontológica de un *inconsciente objetivo*), pero que se presumen formados por cada sujeto empírico en su misma experiencia apotética, que no siempre habrá de ser equivalente a un *autoconocimiento* ("autobservación", que diría Skinner), ni fácilmente traducible en conductas objetivas (movimientos musculares, segregaciones de vísceras), porque es un conocimiento que, si *fenoménico* para el sujeto, es *fisicalista* en la objetividad, objetividad no de las cosas naturales, sino en los productos "objetivados por el trabajo", o las operaciones (agibles o factibles) de los hombres dentro del *eje circular* del espacio antropológico.

Las ciencias histórico-culturológicas han sido caracterizadas como metodologías $\beta 1$ -I situaciones en las que el *objeto se supone ya dado por medio de otras operaciones que lo determinaron (producción)*. Lo que con esas metodologías se reconstruye no es el objeto, sino los cursos de operaciones de quienes lo han construido, operaciones que se suelen definir, no sin ambigüedades, como "supraindividuales" o "extraorgánicas", nociones todas ellas que conciernen a la expresión hegeliana de *extrañamiento*, que se propuso en la génesis constructiva misma de la idea del Espíritu Objetivo. Es así que la Historia y las Ciencias Culturológicas pueden ser entendidas no tanto "ciencias de los hombres", cuanto "ciencias de los objetos producidos por los hombres". Estas metodologías han de considerarse como un momento interno y necesario para acceder a lo que Juan B. Fuentes ha denominado "estructuras morfosintácticas", que habrían de corresponderse con las metodologías $\alpha 2$ -II, y que el citado autor circunscribe no ya sólo a las estructuras lingüísticas, sino a otros *artefactos* o construcciones humanas, por tanto, que *causalmente* (ontológicamente) requieren de la operatoriedad humana, pero en las que *gnoseológicamente* el sujeto (operaciones y fenómenos) ha resultado neutralizado¹⁶ :

"(...) Hemos de contar, en efecto, con una multiplicidad de individuos operatorios, y por tanto con una diversidad de interacciones o interrelaciones operatorias entre ellos, de modo que entre dichas interrelaciones operatorias (o entre sus términos resultantes) broten unas relaciones -ellas mismas no operatorias- que puedan convertirse en la pauta, norma o regla esencial y objetiva de aquellas interrelaciones operatorias. Pero estas relaciones (no operatorias) de (inter)relaciones (operatorias) sólo pueden tener la forma de alguna estructura morfosintáctica como para que, en efecto, no siendo fisicalistas (genéricas), puedan convertirse en la pauta o norma esencial y objetiva, a la vez que específica, de las operaciones y de sus relaciones operatorias".

La situación que estamos describiendo, pensada inicialmente para el Estructuralismo Lingüístico, quiere ampliarse a otras metodologías de las ciencias humanas en las que efectivamente se neutraliza el sujeto operatorio, manteniendo a la vez la *especificidad antropológica* de los objetos del campo. Al respecto del lenguaje, se observa un sistema de oposiciones¹⁷: lenguaje como logos/lenguaje como ουλο α ; generativismo (Chomsky) / estructuralismo (Saussure); metodologías α / metodologías β . Al presentar su crítica de la Gramática Generativa, J. Velarde escribió¹⁸:

"El lenguaje, en cuanto 'sermo', exige como unidades lingüísticas las oraciones (palabras hilvanadas), porque, (...), para los estoicos los significados de las

¹⁶ Fuentes, J.B. , "La Psicología: ¿una anomalía del cierre categorial?", *El Basilisco*, (1992), **11**, 58-71. La cita, en p. 62 y 63.

¹⁷ Fuentes, J.B., "La Psicología...", p. 63.

¹⁸ Velarde, J. , "Metodología de la gramática generativa", *El Basilisco*, (1979), **7**, pps. 4-15.

palabras no se dan aislados, sino que adquieren su valor a través de la acción del sujeto operatorio (el hablante), según que las coloque en un lugar u otro, según opere sintácticamente. El lenguaje es, ante todo, sintaxis, operaciones posibles de los individuos para componer y ensartar, 'capacidad para generar oraciones', en palabras de Chomsky. Y es en la construcción de oraciones en donde el hablante se comporta como sujeto operatorio y no puede ser eliminado del campo lingüístico. La acción del sujeto gnoseológico impide considerar al lenguaje como un conjunto finito y cerrado de elementos -las palabras- y establecer una relación biunívoca entre éstos y los pensamientos o las cosas. Por el contrario, el lenguaje, en la metodología β , es una acción, una facultad que hay que coordinar con la acción del logos que informa al individuo".

Así pues, en aquellas situaciones de las que las *estructuras morfosintácticas* constituyen el "primer analogado", pero que no tienen porqué circunscribirse a fenómenos lingüísticos (situaciones $\alpha 2$ -II), el carácter *universal* y *natural* de dichas estructuras (de raigambre estoica) vendría a oponerse al carácter facultativo, operatorio-intelectual, de las mismas (vía aristotélica). *Mutatis mutandis*, en lo que atañe a la Psicología Cognitiva (similar en tantos puntos a la situaciones examinadas en la Lingüística), podríamos detectar una dicotomía análoga, de larga trayectoria filosófica, y que en el plano categorial habría de llegar hasta nuestros días, constatando, como hemos estado haciendo, el *doble plano operatorio* de las metodologías psicológico-cognitivas.

El empirismo y naturalismo en materia de *cognición*, bascula hacia posiciones muy próximas al conductismo, o a otras tradiciones distantes en todo caso, de la metáfora racionalista-computacional de la mente. Ejemplos podrían detectarse en el área de la Psicología de la Percepción en la línea Gibson-Neisser: el sujeto operatorio es, antes que nada, un sujeto antropológico *genérico*, inmiscuido en un medio ambiente social y cultural. Esta psicología es la que pretende abrir las puertas y ventanas de la Casa de la Mente, sin renunciar al edificio cognitivista, pero precipitándose hacia unidades más amplias, naturales o culturales. Por contra, los llamados *solipsismos metodológicos* (Fodor), o más en general, toda la escuela computacional-simbólica que hemos venido analizando han contribuido a que la doble operatoriedad se desplaza, en este "cartesianismo de nuevo cuño" (que, preciso es decirlo, hunde sus raíces en tradiciones más remotas), aunque no ya hacia la disolución del sujeto cognitivo en unidades (genéricas) fisicalistas, ambientales. La disolución de la especificidad antropológica quiere seguir, por así decirlo, un rodeo en el camino, mediante un pretendido *apriorismo* de leyes lógicas universales, y una genericidad (no antropológica) de los símbolos inscritos en la mente. Con la siguiente cita, relativa a una gnoseología de la Lógica, haremos ver que la estrategia computacional-simbólica, no es sino un camuflaje del *Espíritu Objetivo*, en cuanto que anterior y posterior (abstractamente) a toda forma de conciencia individual. Escribió Bueno¹⁹:

"La escala en la que aparece la racionalidad y la logicidad es, suponemos, la escala de nuestro cuerpo, de nuestras manipulaciones (de nuestras 'operaciones quirúrgicas'). Y aquí pondríamos el privilegio de las 'ciencias formales' (frente a las ciencias reales), su llamado apriorismo, que no haríamos consistir tanto en su vaciedad (en la evacuación de todo contenido, en el 'no

¹⁹ Bueno, G. , "Operaciones autoformantes y heteroformantes. Ensayo de demarcación gnoseológica entre la Lógica formal y la Matemática", I. *El Basilisco*, (1979), 7, pps. 16-39; cita en p. 23.

referirse a la realidad') cuando en su materialidad artificiosa (combinatoria de elementos discretos) en su condición de metros solidarios a nuestro cuerpo manipulador, que no podemos menos de 'llevar siempre con nosotros' cuando nos enfrentamos con el mundo. Traduciendo la fórmula kantiana: es nuestro cuerpo operatorio (no nuestra 'mente', o nuestro 'Ego') aquello que acompaña siempre a todas nuestras 'representaciones' racionales. La Lógica formal no será así tanto el 'reflejo mental' de la Lógica universal, ni la 'trama apriori' del Mundo, cuanto la construcción de un campo cerrado en un espacio de dos dimensiones (las 'leyes' en dirección izquierda / derecha; las 'reglas' en la dirección arriba / abajo) y mantenido dentro de unos márgenes de temperatura precisos. Un universo de símbolos, solidarios a nuestro cuerpo y no siempre coordinables isomórficamente con otras regiones de nuestro mundo, pero entre los cuales broten relaciones necesarias."

2. Los Modelos "Arriba-Abajo": El Mundo Metido en la Cabeza.

Los llamados enfoques "arriba-abajo" (*top down*) van a traerse a colación, pues ilustran sobremanera el conjunto de ideas que esbozamos en el apartado nº 1 de este capítulo: gran parte de los modelos cognitivistas podrían reformularse como una suerte de "repujado"²⁰ que, desde las estructuras o relaciones que pertenecen propiamente a la esfera del Espíritu Objetivo, se van a resolver en pretendidas estructuras subjetivas, por cuyos "huecos" han de filtrarse necesariamente aquellas conexiones artificialmente talladas a escala individual corpórea.

Sin lugar a dudas, los marcos (*frames*) de Marvin Minsky representan una de las estructuras computacionales más conocidas dentro de los llamados modelos "arriba-abajo"²¹. Como ya hemos visto, un *frame* es una estructura que representa el conocimiento en un dominio muy limitado. Esa modalidad de "esquema" produce una descripción de objetos o acciones, comenzando con una estructura invariante, común a todos los casos dentro de un dominio dado, con el añadido de ciertos rasgos variables que se incorporan al "esqueleto". Siguiendo la detallada descripción que B.J. Kuipers²² ofreció de estas estructuras, la lista de propiedades más importantes de los *frames* incluye :

- (1) Descripción: Un marco proporciona una estructura elaborada para crear y mantener una descripción. Se trata de imprimir un *orden*, a partir de una pretendida *experiencia caótica* que es la "vida mental".
- (2) Instanciación: El *frame* produce una descripción del objeto mediante sustitución de los valores predecibles. Esto es, el marco ofrece un número limitado de simples decisiones, de alternativas a elegir.

²⁰ La idea de que los modelos producidos en la Psicología Cognitiva constituyen una suerte de "repujado" de las relaciones objetivas exteriores a la conciencia individual, está tomada de Bueno, G., "Para un análisis gnoseológico de la psicología cognitiva", *Estudios de Psicología*, (1985), **22**, 103-13. En este trabajo emplea literalmente esta expresión , "repujado", haciendo uso de una metáfora artesanal muy sugerente.

²¹ Minsky, M. , "A framework for representing knowledge", en Winston, P. (ed.), *The psychology of computer vision*. McGraw-Hill, N. York, 1975.

²² Kuipers, B.J., "A frame for frames: Representing knowledge for recognition", pps. 151-184, en Bobrow ,D.G. y Collins, A. (eds.), *Representation and understanding. Studies in cognitive science*. Academic Press, Orlando, 1975.

- (3) Valores de omisión (*default values*): Sobre aquellos rasgos descriptivos, acerca de los cuales el marco puede efectuar una predicción suficientemente exacta.
- (4) Variación restringida: Las dimensiones y rasgos de posibles variaciones de cada rasgo son limitadas y de carácter específico.
- (5) Perturbación: Cuando se dan pequeños cambios en una situación dada, se trata de dar cuenta de cómo el sujeto se percata de ellos sin necesidad de recomputar completamente su sistema de *escenas*. Es decir, se pide del marco o escena mental que sea lo suficientemente flexible como para que en él tengan cabida, hasta cierto punto, pequeños cambios inesperados, sin que por ello haya de ocurrir una merma en la "validez" global del esquema para con esa perturbación. Si las transformaciones en la escena tomaran un cariz más radical, se necesitarán recomputaciones más drásticas, al resultar inservible el esquema o esquemas anteriores.
- (6) Justificación. Se refiere al grado de *confianza* de los diferentes rasgos, que constituirán un continuo, por ejemplo, desde las observaciones claras hasta las asignaciones por omisión (lo "implícito").

El *frame* se "activa", para seguir un ejemplo conocido, cuando un sujeto entra en una habitación y ve allí una disposición de muebles, cuadros, ventanas, etc. Si la habitación es "conocida", el individuo que observa esa situación alberga alguna suerte de "esquema", que se cumple de acuerdo con una serie de observaciones que confirman unos valores esperados. Pero, supongamos, que el individuo entra en la habitación y se encuentra, junto a los objetos esperados, un elefante o algún otro animal, persona u objeto desacostumbrado. Desde luego, la observación de este acontecimiento colisiona con el "sentido común" implícito en ese marco o escenario particular, pero no así en un escenario que se supone *típico* o ajustado para ese elefante, por ejemplo, en un zoo, donde el sujeto producirá los esquemas pertinentes en orden a situar allí al elefante y a otros seres exóticos. Este es el tema de las *descripciones dependientes del contexto*²³. Se supone que este es un rasgo *específicamente antropológico* de las descripciones que un sistema cognitivo produce en el proceso de conocer, algo difícil de emular en las máquinas. El hecho de que las personas hiciéramos uso de *conocimiento implícito*, o *tácito*, mundano, etc., frente a la explicitud requerida por las máquinas, ha sido objeto de innumerables polémicas²⁴. Si queremos hacer del hombre una "máquina de inferencias", se preguntaban muchos computacionalistas, ¿cómo podrá ocurrir que el sujeto humano opere tácitamente, llegando a conclusiones que requieren de unos términos que, en principio, no forman parte de las premisas manejadas?

Toda esa constelación de conceptos, todos ellos de naturaleza fenomenológica, cuales son, "mundano", "implícito", "inconsciente", "tácito", han constituido un gran escollo para el "paradigma" representacional-simbólico, al constatarse que no todos los hechos necesitan permanecer explícitamente en un almacén de memoria, para ser operados (transformados) en combinación con otros hechos o términos. Entonces, si la *inferencia* es concebida como el proceso de

²³ Bransford, J.D. y Johnson, M.K., "Contextual prerequisites for understanding: some investigations of comprehension and recall", *Journal of Verbal Learning and Verbal Behavior*, (1972), **61**, 717-26; Bransford, J.D. and Johnson, M.K. , "Considerations of some problems of comprehension", en Chase, W.G. (ed.), *Visual information processing*, Academic Press, Londres, 1973.

²⁴ Polanyi, M. , *Personal knowledge: towards a post-critical philosophy*. Harper, Nueva York, 1964.

transformación de unas configuraciones de términos en otras configuraciones resultantes, de acuerdo con unas reglas fijadas, y los sistemas de procesamiento son entendidos como mecanismos que aplican sus transformaciones sobre el conjunto de términos implicados en el programa, no será lícito apelar una y otra vez a unos sistemas de orden superior hasta llegar a abarcar el universo entero, porque el carácter recursivo del programa (la inclusión de cada subrutina en otro sistema de orden superior) ha de tener un límite. El *modelo* ha de ser más limitado que aquello que ha resultado modelado. Llegamos, con esto, a considerar que *el mundo no "cabe" en los programas*. El recinto de operatoriedad de un algoritmo o programa simbólico es, por definición, restringido.

Bobrow distinguió tres clases de inferencias, para poder dar cuenta del carácter "implícito" de la cognición²⁵:

"I distinguish among three different forms of inference. The term formal inference covers the family of techniques used in predicate calculus representational systems. The term computational inference describes a process in which facts are derived through bounded known computation. The term meta-inference covers techniques by which knowledge about the structure and content of the data base is used to derive further facts consistent with the original set".

Las "técnicas de inferencia formal" consisten en fabricar fórmulas cuantificadas que versan sobre hechos del mundo, y que por medio de *pruebas formales* permiten acceder a cadenas de fórmulas enlazadas por reglas sintácticas. "Dado el suficiente tiempo y espacio", el sistema dispone sobre todos sus términos primitivos. Desde unos "hechos" se accede a otros nuevos por la sola obediencia a unas reglas que gobernarán los procedimientos. Comoquiera que el comportamiento inteligente exige que sean ahorrados los procedimientos exhaustivos propios de los sistemas formales, Bobrow habla de unas "inferencias computacionales", que consisten en el uso de especialistas procedimentales, en la fabricación de unas reglas o mandatos de orden superior que van a permitir el ahorro de datos y pruebas dadas en el más básico nivel simbólico de lenguaje. Por último, nos queda comentar lo que Bobrow llamaba "técnicas meta-inferenciales", útiles para encontrar *hechos* no observables, pero consistentes con el conjunto de datos formalmente descritos en un momento presente. Así, mediante este ejemplo se observa con claridad el modo con el que un sistema procesador puede manejar aquellas informaciones que nosotros suponemos "implícitas", derivadas de un "conocimiento mundano" que, sobre todo por razones económicas, no puede ser manejado por un restringido almacén de hechos sobre los que realizar inferencias. Ese "conocimiento mundano" es suplido por una jerarquía de derivaciones, de las que sólo en la más básica escala de un sistema simbólico podemos hablar, con la debida propiedad, de un sistema lógico-formal. Los ordenadores, en efecto, funcionan con esta jerarquía de niveles, desde un "lenguaje máquina", hasta los niveles "superiores", que son los lenguajes que relacionan los *hechos* que van a describir y manejar la información de manera más inteligible en una *escala antropológica*. Los esquemas, los *frames*, y los modelos mentales son estructuras descriptivas de orden superior aunque están apoyadas, indudablemente, en el más elemental de los códigos de una máquina. Hay una gran diferencia entre una sucesión de "ceros" y "unos" propia de las máquinas de Turing, y los escenarios o marcos (*frames*) capaces de predecir datos no suministrados como input inmediato al programa. Si el programa es capaz de "predecir" que una silla es un objeto que se

²⁵ Bobrow, D.G., "Dimensions of representation", pps. 1-33, en Bobrow, D.G. y Collins, A., *Representation and understanding*, Academic Press, Londres, 1975.

apoya en cuatro patas, la inferencia que a partir del input "silla" realiza la máquina (un rasgo o propiedad de las sillas: "una cosa que tiene cuatro patas") no se debe a ningún tipo de inteligencia o cultura mundana, implícita en el banco del sistema, sino a la extracción de rasgos, extracción o derivación que no es *formal* en el sentido habitual de las pruebas propias de una lógica simbólica, digamos, bivalente (mediante "ceros" y "unos") sino *material* a todas luces, puesto que los términos implicados en la lógica de predicados (de un "lenguaje de orden superior") son *nombres* de los objetos predicados o relacionados mediante otros términos (otros *nombres*), en los que no se ha evacuado ninguna clase de contenido y cuya *referencia* a un objeto material real es, en todo caso, indispensable para un sujeto *antropo-lógico*, no para la máquina misma sino para su usuario. Como escribió Gustavo Bueno²⁶:

"Es erróneo (...) decir que tanto el ordenador como la mano del hombre construyen relaciones ($x=y$), pero que la máquina no las conoce (o advierte) y el hombre sí. En este caso la 'conciencia' se nos daría como un mero epifenómeno. No sería posible entender la diferencia si comenzamos por considerar la relación como algo dado en sí, susceptible de ser conocido: hay que regresar hacia la ontología misma de la relación, a fin de ver cómo es ella la que incluye la conciencia más que reciprocamente. La función del hombre no será 'conocer' una relación que le antecede, sino fundarla. Y ello sólo se entiende si el hombre es el contexto de aquella. Por ello, la diferencia entre el ordenador y el cerebro (o la mente) no la pondríamos en la diferencia que pueda existir entre una supuesta capacidad de advertir la semejanza en un caso y una advertencia de la misma (ya dada) en otro. Nos parece que este planteamiento implica un entendimiento metafísico de la relación de identidad o semejanza (entre A y A, entre 0 y 0) como si se tratase de una relación binaria absoluta, sustancializada : una relación binaria dada tanto por la máquina como por el álgebra, pero que la mente pudiera captar y la máquina no (aunque la máquina la hubiera establecido)".

Desde el materialismo, las relaciones de identidad han de "dialectizarse", puesto que ellas no constituyen nexos sustanciales por sí mismas y la conciencia que el sujeto antropo-lógico (a diferencia de las máquinas) habría de adquirir de tales relaciones sería, así sustantivizadas las relaciones, puramente epifenoménica. Ha de concebirse, en cambio, que la identidad (o semejanza) también es una *negación* de todo un conjunto de propiedades diferenciales en terceros contextos²⁷:

"(...), teniendo en cuenta que la relación no es binaria, advertiremos que la semejanza entre "A" y "A" sólo tiene sentido en un contexto en el que figure el cuerpo humano (dado a su vez entre otros cuerpos), contexto que no cabe atribuir a la máquina en sí sin 'animalizarla'. Si atribuimos pensamiento a una máquina que arroja en su cinta la fórmula $A=A$, es sólo porque no advertimos que esa fórmula está siendo contextualizada por cerebros humanos. Para decirlo rápidamente: si el cerebro se diferencia de la máquina computadora no es porque pueda advertir o captar una identidad que la máquina no puede captar o advertir, sino porque esa identidad es una relación que sólo tiene sentido en el contexto de una actividad cerebral (...), una actividad zoológica antes que mecánica, un 'ejercicio'".

²⁶ Bueno, G. , "Operaciones autoformantes...", p. 31.

²⁷ *Ibid.* p. 31.

Las especulaciones que desde la filosofía de la mente y de la Inteligencia Artificial se producen en relación con la incrustación de *fórmulas simbólicas* ("formas") en los *contextos* ("mundos", "materias"), deberían dar por sentado que las *relaciones de identidad* aprehendidas o desarrolladas en un programa, por reglas o funciones sintácticas, no adquieren su "incrustación semántica" por procedimientos lingüísticos a su vez por que, cualesquiera que sean los niveles de formación y transformación de los términos implicados, estos son *nombres o emblemas* de los objetos reales que en el mundo efectivamente están vinculados o son sujetos de atribución. Así, en la fórmula

AMA (JUAN, MARIA)

Cabe suponer que **AMA** es una función que enlaza de una cierta manera al sujeto **JUAN** con el sujeto **MARIA**, y donde cada uno de los tres *signos mención* (*token*) son miembros de tres respectivas clases *nombradas* por los *signos patrón* (*type*). Pues bien, la fórmula proposicional global, **AMA (JUAN, MARIA)**, es ella misma un *nombre* (o "resumen") de una porción concreta de relaciones, para cuya selección se han excluido otras conexiones que no convienen a la función "**AMAR**", ni tampoco a los sujetos implicados, **JUAN** y **MARIA**. El *contexto* o la "incrustación" que la mera sintaxis de símbolos debe conocer, o expresado de otra forma, los "parámetros" o los "límites de la variabilidad" en las transformaciones de símbolos, no han de considerarse una *contextualización o interpretación* extrínseca, "separada" de las manipulaciones de los símbolos. La máquina no realiza *operaciones* en el sentido con el que, desde el materialismo, atribuimos *operatoriedad* al ser humano (atribución basada, precisamente, en la capacidad de ese *sujeto antropológico* para establecer o construir *relaciones* apotéticas o fenoménicas). Las llamadas "operaciones" de la máquina, el "despliegue" de un programa, no es sino la *relación* misma de los términos (sintaxis), en rigor, un hilvanado o vinculación que se rige por el *determinismo algorítmico*, que es también una *normatividad*, el conjunto de reglas o normas del que se compone un programa en orden a producir un resultado. La secuencia y modo de vinculación entre términos simbólicos es una *normatividad* re-producida por un programador, y luego por un usuario, a partir de la *legalidad* de las conexiones fundadas por los sujetos antropológicos en el mundo cultural y natural²⁸ :

"Si una máquina de Turing puede proceder de suerte que a cada una de sus posiciones sucesivas pueda considerársele determinada por su estado interno q y por un símbolo s impreso en una casilla de la cinta, es porque s no desempeña meramente el papel de un significado (noemático) sino porque actúa en virtud de alguna propiedad o característica física (no necesariamente analógica): en otro caso, la acción del símbolo sería 'mágica' --y sólo cuando los símbolos impresos en la máquina se abstraen de su contenido físico, las computadoras pueden aparecer como 'misteriosos cerebros'. Sobre estas propiedades o características se basa además la posibilidad de que un computador desborde las posibilidades de un cerebro humano. Y, sin embargo, no podría decirse que la máquina 'piensa' : podría decirse que construye figuras de símbolos, los acumula en la cinta, pero sin que los símbolos acumulados, cuando son semejantes entre sí, por ejemplo, sean percibidos como él mismo en virtud de un proceso lógico autoformante que habría de remitir a la naturaleza misma de las sustancias del sistema nervioso".

²⁸ *Ibid.* p. 30.

3. Sentimientos y Emociones en la Antropología Computacionalista.

El problema de la emotividad siempre se ha situado en un punto central en el espacio general de los problemas que se discuten y se ensayan a propósito de representarse el isomorfismo entre los hombres y las máquinas. Introducir los problemas suscitados por la supuesta capacidad emotiva de las máquinas, capacidad a todas luces inverosímil, se exige como un punto inexcusable a tratar, si quiera someramente en un trabajo general como el presente, por cuanto que constituye uno de los *desiderata* de este nuevo humanismo mecánico con el que se pretende bien mecanizar al hombre, bien humanizar las máquinas (sin que haya conflicto en ambas tendencias). Forman parte de la segunda trayectoria las operaciones conducentes a atribuir sentimientos, estados emocionales, ansiedades, etc. a las máquinas, y ello ha preocupado a muchos filósofos de la mente y de la cibernética, desde hace varias décadas. Los programas y teorías ("modelos") que han intentado simular la emotividad humana no han hecho sino perseguir el viejo anhelo cibernético de una "antropología mecánica", aun cuando el mecanicismo cibernético fuera interpretado por muchos espíritus como un verdadero "antihumanismo". Desde los tiempos predecesores de Wiener, la idea de simular o reproducir la persona humana al completo, como compuesto de capacidades intelectuales tanto como sentimentales, ha estado presente en las discusiones y en la literatura, y esos intentos han gozado de mucha popularidad. Huelga decir que nuestro trabajo se ha centrado en aquellas especialidades en las que la Psicología Cognitiva ha ofrecido más elementos de análisis, en áreas como la percepción, el lenguaje y la memoria, esto es, en lo que aún podríamos llamar *entendimiento humano*, pues es en el área de las facultades intelectuales donde la escuela del *procesamiento de información* ha cosechado mayores triunfos. Ahora bien, no han faltado programas en los que una parte de la mente, usualmente llamada "parte caliente" (u horéctica), pretende simularse. Habríamos de enfrentarnos, en este caso, no a una *ciencia psicológica*, en sentido estricto, sino a una antropología filosófica de nuevo cuño que se sirve de las máquinas para reorganizar el campo de ideas y conceptos que gravitan en torno a la *idea de hombre*, y es hacia esta idea a la que se va a *regresar* constantemente, y merece la pena señalar que los filósofos cibernéticos y cognitivos hayan tenido que centrarse en relaciones antropológicas de tipo *angular*, invistiendo a las máquinas de todo un conjunto de notas concernientes a la numinosidad de los ordenadores y robots. Un modo de operar harto frecuente entre estos filósofos (que, más que filosofía escriben "ciencia-ficción" o "futurismo"), consiste en conceder un cúmulo de propiedades --cada una, por sí sola, irrelevante-- para que, una vez llegados ante una cierta sumación de las mismas, la máquina (o el programa) no sea distinguible de cualquier sujeto humano. Es así que la "numinosidad" del programa equivale a un proceso de disfraz antropomórfico, en el que el núcleo (*inteligente*) manipulador de símbolos se reviste de toda una complejidad *protésica*, corpórea, para transformar la primitiva cinta de Turing (o algún equivalente lógico suyo) en un númen análogo al hombre en tantos respectos (no sólo intelectuales, sino emocionales, motrices, etc.). Esto se observa claramente en escritos como los de Paul Ziff en su famoso artículo "El sentir de los robots"²⁹ :

"Una vez sentado que estos robots son, sin duda alguna, máquinas, en todos los demás aspectos serán tan humanos como se quiera. Son de tamaño humano. Una vez vestidos y enmascarados son prácticamente indistinguibles de los seres humanos en casi todos los aspectos: su apariencia, sus movimientos, lo que dicen, etc. Por lo tanto, de no ser por la máscara, cualquier

²⁹ Ziff, P. , "El sentir de los robots", pps. 161-70, en Ross Anderson, A. (ed.), *Controversia sobre mentes y máquinas*, Orbis, Barcelona, 1985. Orig. en *Analysis*, (1959), **Vol. XIX, 3**, pps. 64-68.

persona los tomaría por seres normales. Al no inspirar sospechas de que son robots, no infunden desconfianza. Pero, sin la máscara, todo su entramado metálico queda a la vista. De lo que aquí se trata no es de si podemos diluir la línea entre el hombre y la máquina y atribuir sentimientos a la máquina. La cuestión estriba en si podemos atribuir sentimientos a la máquina para poder diluir esa línea entre el hombre y la máquina".

Este tipo de cuestionamientos, derivado de los que Turing planteara, no dejan de constituir una especie de "antropología predicativa", puesto que se intenta dirimir especulativamente qué predicados son verdaderos de la máquina, y cuáles no, como camino indirecto para acceder a los predicados de los que pueda decirse que son verdaderos (y falsos) del hombre. Así, operar con las predicaciones de la "antropología de las máquinas" no es sino alzar el edificio de la "antropología del hombre" (si se permite esta redundancia). De hecho, muchas de las objeciones que Turing previó, al margen de las de tipo lógico-matemático (semejantes a las llamadas "limitaciones gödelianas" de las computadoras), y otras más extravagantes (por ejemplo, la nº 9, el argumento de la "percepción extrasensorial")³⁰, tienen que ver con una antropología filosófica a la que Turing estaba, precisamente, ridiculizando, en especial, aquella de factura metafísico-predicativa, que venía especialmente reflejada en su quinto argumento³¹:

"5. Argumentos de incapacidades diversas. Estos argumentos responden al esquema: 'Te aseguro que pueden hacerse máquinas que realicen todo lo que has dicho, pero es imposible construir una máquina que haga X', y se citan al respecto diversas X, (...)".

Lo que Turing, con gran sentido del humor, llamaba "argumento de las incapacidades diversas", venía a ser un desmantelamiento de la fuente de críticas "humanísticas" del tipo: "la máquina *no puede* hacer X, y el hombre *sí puede*, en cambio". Independientemente de las posturas adoptadas por cada uno respecto de los particulares X de referencia, cuestión empírica o inductiva en gran medida, el hecho es que la emoción, los sentimientos o, más generalmente, los llamados *qualia* (estados o vivencias cualitativos de la mente) suelen ser atribuidos o despojados de las máquinas según la "teoría" antropológica con la que se está operando. La potencia relativa de cada uno de los *isomorfismos* computacionales se torna muy nebulosa dentro del contexto operacional de una antropología predicativa. La obligada crítica hacia estos modelos consistirá en denunciar la sustantivización de unos nudos de relaciones que se presentan como si no fueran *generadas o producidas* por agentes humanos previos. La evaluación (gnoseológica) ha de presuponer sujetos antropológicos en un sistema artificial isomórfico con respecto al cerebro o mente de los hombres en cuanto que es una manipulación o mediación sobre un material (términos, relaciones), *humanidad* intercalada antes que sustantivizada en las configuraciones tomadas como referencia.

Como quiera que la emotividad de las máquinas es otro caballo de batalla, un auténtico tópicos, en la evaluación de los isomorfismos entre el hombre y la máquina, vamos a examinar con más detalle la propia noción de *isomorfismo* y otras nociones emparentadas (aunque diferentes). Según Bunge³² la *equivalencia* es mucho más fuerte que la *analogía*. Este filósofo distingue entre *analogías formales* y *analogías sustanciales* (que incluyen a las formales), así como las *homologías* (analogías

³⁰ Turing, A., "Maquinaria computadora e inteligencia", en Ross Anderson, A. (ed), pps. 42-43.

³¹ Turing, A. *op. cit.* p. 33.

³² Bunge, M., *Teoría y realidad*, Ariel, Barcelona, 1972; p. 237.

formales y sustanciales a la vez). Se trata ahora de examinar las posibilidades que un modelo artificial tiene como análogo de la mente humana (y viceversa). La "analogía sustancial" queda descartada por los mismos defensores del funcionalismo computacional, quienes dicen buscar las semejanzas hombre-máquina entre ciertas propiedades dentro de un orden más *elevado* con respecto de aquel que se corresponde con la mera composición físico-química o bioquímica del sistema ("protoplasmática", como le gusta decir a H.A. Simon). Así, K. Gunderson, mediante su distinción entre Inteligencia Artificial (I.A) y Simulación Cognitiva (S.C) parece rondar lo que Bunge llamaría "analogía formal laxa" (*algún/algún*) entre las computadoras y el pensamiento, en especial si de lo que estamos hablando es de S.C., es decir, la simulación computacional de procesos cognoscitivos. Una *analogía formal laxa* equivale a una afirmación del siguiente tipo: algunas propiedades formales de la máquina/programa son las mismas que algunas de las propiedades de la mente/cerebro del hombre. Por tanto, se puede decir que --a cierto nivel-- los dos sistemas, computadora y hombre, son análogos por la razón de compartir un conjunto finito de propiedades, si bien con respecto de otras propiedades distintas, para los cuales la computadora y la máquina no se solapan, los dos sistemas han de tomarse como cosas muy dispares. Pero la incorporación de cuantificadores a esta "antropología predicativa" sigue sin limpiar la faz metafísica con la que se presenta.

Otra posibilidad de las que Bunge manejaba en su artículo, era la de la *analogía formal inyectiva (todo/algún)*. Esta sería la opción favorita de quienes defienden la idea de la existencia de subconjuntos o regiones parciales de la Cognición Humana *no simulables* (o *computables*, y aquí empleamos un "o" disyuntivo), por alguna razón muy importante. Roger Penrose ha defendido, entre otros autores, el tipo de argumentos "gödelianos" en contra de la computabilidad universal de los procesos cognoscitivos³³. Es decir, razones *de principio*, razones no derivadas de una pretendida *esencia humana* a la que hubieran de ajustarse los parámetros de normatividad de otros sistemas análogos, las máquinas, sino de la normatividad que aparece precisamente por la confluencia de los diversos algoritmos. Quizás se debe a la persistencia en operar con esa *idea de hombre* como si fuera un núcleo ontológico previo, por lo que Penrose y otros autores (Nagel y Neuman, Lucas) han *regresado* desde las "objeciones gödelianas" hasta una idea *esencial* de hombre aunque esa idea incorpore ahora elementos tomados de la Teoría de la Evolución, la neurología *versus* hardware, etc. Así, según esta antropología, la ausencia de neuronas en una computadora, o la "juventud" filogenética de la máquina en comparación con el hombre constituyen razones *a priori* tan contundentes como las "gödelianas" (procedentes de las matemáticas y de la lógica).

Desde aquí nos vemos obligados a impugnar a los críticos del cognitivismo su incapacidad para salir de un "esencialismo" con respecto de la naturaleza humana, pues por más que sus críticas tengan siempre un fondo de verdad, su potencia ha de verse menguada por no haber sabido apreciar la estructura relacional con la que habría de recuperarse toda idea de hombre. Así, ante la tesis *empírica* (etológica) de

³³ Penrose, R., *La nueva mente del emperador*, Mondadori, Madrid, 1991; orig., *The emperor's new mind*. Oxford Univ. Press, Oxford, 1989. Los aquí llamados "argumentos gödelianos", son razones extraídas desde la Matemática y la Metamatemática, conducentes a rebatir las tesis de la I.A "versión fuerte". Aquí se citan los teoremas de Gödel por antonomasia, si bien hay otras contribuciones de las que no nos ocupamos. Su descripción se encuentra en los primeros cuatro capítulos de la obra de Penrose. Como ya hemos indicado, se debe a J.R. Lucas una célebre utilización de los "argumentos gödelianos" para rebatir la tesis de la inteligencia de las máquinas: "Minds, machines and Gödel", *Philosophy*, (1961), **36**, 112-127. E. Nagel y J.R. Newman en su divulgación, *Gödel's proof* (New York Univ. Press, N. York, 1958), también pretendieron extraer conclusiones acerca del funcionamiento de los cerebros a partir de los teoremas gödelianos.

que las emociones humanas provienen de una suerte de "endulzamiento" de las condiciones de vida a consecuencia de la civilización y la destrucción de los predadores no humanos (el *stress* o la ansiedad serían vestigios de sistemas de alarma de nuestros antecesores)³⁴, cabría oponer la objeción (retorcida o circular, "irrefutable", en el sentido de Popper) de que las "máquinas inteligentes" son, a su vez, un paso más en el "enfriamiento" de tales sistemas, u otra similar.

La distinción debida a Gunderson³⁵ entre *rasgos resistentes al programa* y *rasgos receptivos* sirve para mantener el *desideratum* de la ciencia cognitiva. Esta tesis es consistente con lograr un núcleo de inteligencia susceptible de formalización, deshechando las partes que no se avienen con los simbolismos. Así las cosas, las emociones no tienen cabida en los programas, y a ellas se refiere Gunderson como ejemplo de *rasgos resistentes*, sin duda porque dependen de un hardware humano que posee los sistemas hormonales y nerviosos responsables de cosas como la ira, el dolor, la angustia, los cambios de humor, etc. Gunderson, que está muy lejos de pretender derribar el edificio de la ciencia cognitiva en su conjunto, mezcla así las llamadas "críticas gödelianas" con las limitaciones derivadas del hardware (de la fisiología y la evolución biológica). Mediante la división entre I.A. y S.C. , que tiene la apariencia externa de una demarcación científica (*de distinctioe scientiarum*, que podría coordinarse con la distinción, más efectiva y palpable, entre I.A. y Psicología cognitiva), Keith Gunderson va a proteger los méritos del enfoque computacional que no deberían empañarse por los fracasos, que también existen. El autor quiere deslindar lo que es una "teoría psicológica" (o pretende serlo) frente a aquello que no es ni pretende ser un artefacto inteligente ni un modelo del pensamiento humano. En otras palabras, la I.A. es la disciplina de las máquinas que requieren de inteligencia humana (sin insinuar que las máquinas sean, ellas mismas, inteligentes); por el contrario, la Simulación Cognitiva (S.C.) se encargará de construir programas que son *teorías psicológicas per se* (y esto no tiene nada que ver con la idea de que la máquina puede pensar). En palabras del autor³⁶:

"AI (...) as the discipline dedicated to the design and programming of machines capable of performing tasks that hitherto required human intelligence, and carried withy it no supposition that the machines capable of such tasks were in any way 'themselves' intelligent, nor was it assumed that such machines would be interesting models of human mind or provide us with psychological explanations. CS research, on the other hand, was described as being committed to the view that the machine cum program provides a psychological theory of how human beings think, solve problems, create things and so on. It was never assumed that CS as just described presupposed that a succesful (or useful) simulation would show that machines could think".

Los programas destinados a simular comportamientos emocionales o sensibles, por ejemplo los que K.M. Colby diseñara amparándose en teorías psicoanalíticas³⁷, parecen presentar los mismos problemas filosóficos que cualquier sistema que se encara con la comprensión de frases en lenguaje natural. A fin de

³⁴ Tesis de esta índole son defendidas por algunos etólogos modernos, como Irenhäus Eibl-Eibesfeldt, por ejemplo en su obra *Amor y odio*, Salvat, Barcelona, 1986. Eibl-Eibesfeldt es un continuador casi directo del "Darwin psicólogo", el que escribiera *La expresión de las emociones en los animales y en el hombre* (Alianza, Madrid, 1984; traducción y comentarios de Tomás R. Fernández Rodríguez).

³⁵ Gunderson, K. , *Mentality and machines*, Croom-Helm, Londres, 1985.

³⁶ *Mentality and machines*, p. 168.

³⁷ Boden, M.A., *Inteligencia artificial*, cap. 3.

cuentas, hay que almacenar una red de relaciones conceptuales. Según Boden, que describe estos intentos³⁸ :

"(...),no se puede considerar que un programa competente para manipular cadenas de palabras como 'Yo odio a mi padre' de manera superficialmente racional (aunque neurótica) entiende las correspondientes oraciones comunes. esto hay que destacarlo, porque esta cadena de palabras sugiere a los humanos casi inevitablemente, incluso al programador, una oración común de la misma configuración superficial constituida sobre una rica base semántica en la mente de la persona del caso. Pero el programa de Colby no dispone de esta base semántica. esto no tiene nada que ver con la supuesta incapacidad metafísica de las máquinas para experimentar rencor: se aplican observaciones similares a oraciones como 'Toma el bloque rojo grande' (...). El programa de Colby no entiende ni una oración, ni tampoco una palabra, sino en el sentido más débil posible."

Esta filosofía de la simulación a través de ordenadores, al servicio de "teorías psicológicas", cuando se enfrenta a los problemas relativos a la emoción y los sentimientos, puede adoptar dos posturas: (a) ampararse en una "buena" teoría psicológica y a partir de ella, programar con los medios disponibles, o (b) sea cual sea la teoría psicológica empleada, si no hay una base semántica lo suficientemente rica, reconocer que nunca estaremos ante una simulación aceptable de emociones o sentimientos. Esta segunda posición parece ser la de la escritora M.A. Boden cuando comenta las incursiones del psiquiatra Colby en el campo de los ordenadores. Por lo pronto, podemos decir que, si bien la simulación de emociones y sentimientos presenta problemas filosóficos propios, frente a la clásica simulación de habilidades intelectuales (*problem solving*), comparte con ésta, no obstante, la necesidad de incorporar la semántica de forma adecuada en un sistema que manipula símbolos. Esto es, la computación de las funciones "calientes" sólo puede darse a través de las técnicas estándar de representación de conceptos, con todas las dificultades lingüísticas o filosóficas que eso entraña., con independencia de la "teoría" psiquiátrica o psicológica en la que se haya inspirado el programa.

Lo que está claro es que el nuevo "humanismo mecanicista" exige o promueve la aplicación de unas mismas técnicas (representación del conocimiento) para generar o modelar una suerte de *concepción del hombre* en un sentido integral. No hay que conformarse con que una máquina "comprenda oraciones", o "perciba escenas con sentido", etc. Se trata de que la máquina elabore sus actitudes y sentimientos respecto de lo ajeno y de lo propio, que "vibre" sentimentalmente por sí misma.

Parece que el par *sintaxis/semántica*, que hemos tratado con un mayor detenimiento, se coordina con este otro doblete, *experiencia/expresión*, para construir así una nueva antropología, aunque sea a costa de incluir bajo la *idea de hombre* a unos representantes *sui generis*, las máquinas informáticas. El par *experiencia/expresión* está vinculado con otros muchos, como *cognición/conducta*, *medio interno/medio externo*, *voluntad/representación*, etc. Según escribió A. Hidalgo³⁹: "*Espinosa, (...), es el primero en construir una filosofía racionalista de los sentimientos, el primero en resolver con su antropología determinista la fisura indeterminista entre 'experiencia' y 'expresión' en el marco de su ontología especial del pensamiento*". En efecto, la mente, tomada como algo más que una masa de deseos varios y contradictorios, ya aparece determinada causalmente por factores externos en la antropología espinosista. La división tripartita de las facultades mentales (memoria,

³⁸ Boden, M.A., *op. cit.* p. 88.

³⁹ Hidalgo, A., "Filosofía de los sentimientos", *Luego....Cuadernos de crítica e investigación*, (1988), 11-12, 53-81.

sentimiento, voluntad) quedará destruida por la *relación diamérica* entre *experiencia* y *expresión*, relación que viene desbordada por causas referidas al cuerpo y a lo exterior. Así, en su *Ética demostrada según el orden geométrico*, Espinosa parece acercarse a una noción de *estructura* para explicarse los vínculos entre *deseo*, *conducta* y *conocimiento*. A partir de las investigaciones del profesor Vidal Peña⁴⁰, podemos repetir aquí que la *Ética* espinosista "es un tratado muy poco espiritualista sobre el alma humana, la cual, de alguna manera, no es sino el correlato emotivo, intelectual y pasional de movimientos de las partes del cuerpo... También la *Ética* es un reconocimiento de las pasiones como realidades, y de que el deseo es la esencia del hombre, del que arranca para conocer". Así pues, de acuerdo con estas consideraciones, en Espinosa y, como ahora veremos, también en Leibniz, parecen hallarse esquemas diaméricos de conexión entre *experiencia* y *expresión*, con los cuales salir al paso las "antropologías cognitivo-sentimentales" que no consiguen intercalar una "facultad de conocer" con una "facultad de desear". Lejos de abundar en este *excursus*, necesariamente breve e insuficiente, exponemos también ciertos elementos de la antropología leibniziana, en la que puede advertirse la conexión diamérica entre las *vis representativa* y la *vis apetitiva* ensayada en el siguiente párrafo⁴¹: "la acción del principio interno que produce el cambio o paso de una percepción a otra puede llamarse *apetición*, si bien es cierto que el *apetito* no siempre puede acceder por entero a toda percepción a la que tiende, en cambio, siempre obtiene algo de ella y accede a nuevas percepciones". Basándose en este párrafo, Bueno ha resaltado la partición que Leibniz hizo de la *vis representativa* en partes homogéneas determinadas (**A1**, **A2** ... **An**) cuyo nexo entre determinaciones habría de ser atribuido a la *vis apetitiva*. Pues bien, en lo que se refiere al cognitivismo y a los programas de simulación, estos esquemas diaméricos de conexión entre representaciones y apetitos no están ejercitados. Un programa como el PARRY de K. Colby⁴², pretendía simular a un hombre paranoico que cree de forma patológica que está siendo perseguido por la mafia, programa que trataba de abarcar, en consecuencia, no sólo el *contenido* de un sistema (anormal) de creencias, sino también los estados emocionales asociados a esas creencias, esto es, la tensión, la ira, etc. En este famoso ejemplo de "programa emocional", ¿qué aportación se ha realizado a una teoría de la paranoia o, más en general, de las emociones humanas? Claramente, ninguna. El programador parte de la teoría que más le satisface o le conviene y se pone a "preparar" una máquina para que responda de la forma más ajustada posible a lo que a él, como autor del programa, le parece que debería ser el comportamiento típico de un paranoico. Sin embargo no por diseñar un programa así ha de estar implícita la tesis de que la emocionabilidad humana se reviste dentro de un *medium* simbólico, que su formato es sintáctico. El ordenador, desde luego, debe manejar símbolos y configuraciones de símbolos para arrojar como output palabras y frases coherentes con el input, que también contiene palabras y oraciones en un lenguaje natural. Pero la "legalidad" presupuesta para las emociones (la "teoría psicológica" empleada en la programación) es una legalidad completamente externa a la hipotética vivencia interior de la máquina. Es la misma *exterioridad* de conexiones legales que al programa SAM de Schank y Abelson hubo de imponerse para que la

⁴⁰ Peña, V., "Introducción" a Espinosa, B. de., *Ética demostrada según el orden geométrico*. Editora Nacional, Madrid, 1980; reimp. en Orbis, Barcelona, 1984; la cita en p. 40 de esta última edición.

⁴¹ Tomado de Fuentes, J.B., "La psicología: ¿Una anomalía del cierre categorial?", *El Basilisco*, (1992), 11, 58-71, ver p. 68. También, Leibniz, G., *Monadología*, (edic. trilingüe a cargo de J. Velarde), Pentalfa, Oviedo, 1981.

⁴² La bibliografía sobre diversas versiones de PARRY, de K. Colby (y otros), puede hallarse en Boden, M. A., *Inteligencia artificial*, p. 135, nota nº 5.

máquina "predijera" eventos tales como que tras sentarse uno en la mesa de un restaurante, al poco acude un camarero a preguntar qué se desea para comer. La máquina, el programa, no sabe nada de restaurantes ni de sentimientos. Simplemente, las *temáticas* pueden ser distintas en el simbolismo del programa exactamente del mismo modo que ocurre con los libros: con el mismo alfabeto y reglas de sintaxis del español puede uno escribir *La Regenta* o un tratado de botánica. El sistema simbólico es el mismo, pero las temáticas a las que fueron aplicados pueden ser muy diversas, y así las "dependencias conceptuales" pueden simular, vistas así las cosas, prácticamente cualquier acontecimiento, cualquier pensamiento, porque casi todo se puede representar en palabras y la legalidad que preside esas dependencias conceptuales es exterior y arbitraria. Para quitarle ambigüedad a nuestra denuncia de la *exterioridad* de esos nudos semánticos de "legalidad referencial" (o fenoménica, mundana), traemos a colación todas nuestras consideraciones que hemos estado haciendo en el presente capítulo V en torno a la estructura ontológica de un *Espíritu Objetivo*. Al igual que vimos en nuestra interpretación de los modelos "arriba-abajo", no es el mundo (o la región de conexiones legales que de este mundo se quiera reproducir o modelizar) lo que está introducido en una cabeza. Al entrar en nuestra habitación y ver en ella los objetos esperados, no estamos haciendo sino recomponer (en la esfera de un *Espíritu Subjetivo*) una estructura de copresencias que fluctuará, o bien se reorganizará al completo, en fuerte dependencia o vinculación con lo que de objetivo tiene toda esa urdimbre de *formaciones sociales* (aunque estas *formaciones* no agotan el sentido de un Espíritu Objetivo) que envuelven al individuo psicológico. Estas estructuras presionan al individuo desde fuera de su esfera subjetiva, y la efectividad causal que ellas posean sobre los individuos nada tiene que ver con la conciencia que cada sujeto empírico-psicológico cobre de ellas. Es así que el "esquema de la habitación" pertenece, es cierto, al ámbito de la psicología (en lo que tiene de percepción individual- subjetiva de *una* habitación), pero no es ocioso desconectar esa captación fenoménica de los elementos ordinarios de una habitación (cuyo esqueleto es, en efecto, el *frame*) con la estructura ontológica (o esencial) de las habitaciones, cuyo campo categorial podría designarse, digamos, con el nombre de Urbanística, o alguna otra ciencia que estudia los edificios, sus partes, y sus aglomeraciones. Y esa ciencia urbanística, como disciplina antropológica, desde luego, nos diría cosas muy interesantes a cerca del modo en que se disponen las puertas y demás vanos de un departamento en un edificio, los usos del descanso y la vivienda en una cultura dada, etc. ¿Qué hace la Psicología cognitiva? Sencillamente, y esta es nuestra tesis, el computacionalismo *segrega* la causalidad (ontológica) , la lógica material que preside las relaciones entre objetos, una legalidad que en el cognitivismo se ha de dar como *exterior* al sistema de procesamiento, si bien desde el materialismo una legalidad semejante no ha de considerarse como previa a ningún procesador (puesto que es errónea toda concepción de un sujeto psicológico como un "devorador" o receptor de información), más bien, de lo que ha de dar cuenta una ciencia psicológica es de cómo el sujeto empírico *funda* o construye esas mismas relaciones que no son meramente "exteriores" respecto de un fuero interno, sino más bien envolventes, puesto que exigen las operaciones de otros sujetos, de otras subjetividades.

De la misma manera, lo dicho para los marcos visuales o escenarios será válido para cualquier teoría de los esquemas emocionales. Yerran por completo los filósofos y psicólogos preocupados por el computacionalismo al creer toparse en este terreno con obstáculos insuperables y oscuros "misterios de la mente humana". Nada hay de profundo o enigmático en la emotividad, y esa desvinculación que la Psicología del P.I. exige respecto de las categorías del Espíritu Objetivo aquí no es menos importante que en otros terrenos. La cuestión no se reduce a la constatación de que, en efecto, y por poner un ejemplo, existe una "cultura emocional", con sus variaciones

y sus pautas etnológicas específicas, sino porque en su propio *plano ontológico* la emotividad remite de forma necesaria (sinectiva) a la intersubjetividad y, en el espacio antropológico, a las *relaciones circulares*.

Cuando en la literatura de la filosofía de la mente se menciona a los "estados mentales" como entidades susceptibles de análisis, suelen comprenderse todo un conjunto de especificaciones de ese género vago y hambriente de definición que se supone es el de los "estados mentales". Un estado mental puede consistir en el conocimiento que un sujeto tienen de la palabra "bicicleta", o puede consistir en desear una bicicleta, o también, nótese la diferencia, estar frustrado por no haber conseguido obtener una bicicleta (en un concurso, por ejemplo), o desarrollar una fobia hacia todas las bicicletas en general, quizás tras una desgraciada caída en la niñez. Todos esos son "estados psicológicos" o "mentales", respecto de los cuales toda una tradición analítica ha querido detectar una idiosincrasia lógica, como si ellos mismos, al constituirse como una cierta clase de oraciones o de partes de oraciones, analizables lingüísticamente, fuera legítimo desentrañar en esa clase la peculiaridad lógica que les caracteriza, su forma, y por tanto, saber algo acerca de lo mental, por ejemplo en su coordinación con lo no mental. Por ejemplo, Peirce identificaba la *cualidades* con su presentación *monádica*. Por ende, un dolor de muelas era un fenómeno, un hecho contingente que había que suponer aislado e indivisible, desligado de cualquier otra cosa. Así, bajo este supuesto de aislacionismo de la experiencia cualitativa, la cualidad en Peirce era separada de la factualidad (los hechos "brutales") así como de las Leyes del Pensamiento. Según esto, la estructura monádica del *qualia* "me duelen las muelas" consiste en la adscripción de un predicado **P** (dolor de muelas) a un sujeto **S**, en un tiempo **t**. Esta predicación monádica de lo mental fue puesta en entredicho por Ryle con su doctrina de los *conceptos disposicionales*. Motivaciones, estados de ánimo, sentimientos, etc. fueron concebidos por Ryle como condiciones temporales que aglutinan conjuntos de acontecimientos, o mejor, propensiones o disposiciones que *se actualizan* por las sensaciones corpóreas concretas. La emotividad, al igual que la fragilidad del vidrio, es una cualidad disposicional. El *error categorial* que denunciaba el filósofo británico consistía en interpretar las palabras disposicionales como palabras que denotaban agentes causales subyacentes. Asignar el predicado P ("dolor de muelas") a un sujeto S ("yo", "a mí") no es equivalente a realizar un informe sobre un estado de cosas, una descripción de sucesos que se supone están debidos a ciertas causas, por ejemplo la activación (fisiológica) de ciertas fibras o receptores encargados de la producción y transmisión del dolor. Esto ya lo había puesto de relieve Wittgenstein, que lanzara su argumento en contra de la *identidad*: si yo observo una imagen, y (mediante instrumentos adecuados) observo el estado cerebral durante mis observaciones, estoy viendo dos cosas y no sólo una⁴³. Más explícitamente, le correspondió a Skinner poner en claro la forma en que las predicaciones monádicas en apariencia, cuales son las de tipo mental ("tengo hambre"), encubren en realidad un tejido de relaciones cuya traducción al lenguaje ordinario, no por parecer ridícula, dejaría de ser una de las múltiples reformulaciones conductistas de mi íntima sensación de hambre, por ejemplo: "encubiertamente estoy emitiendo un comportamiento similar al comportamiento implicado en obtener y consumir alimento"⁴⁴. Es decir, el "informe de sensaciones mentalmente experimentadas", no es nada en sí mismo salvo un mero informe, el cual puede obviar o resumir, por cierto, las *condiciones* o las *circunstancias*

⁴³ Este punto ha sido tratado por H. Putnam en su artículo "Mentes y máquinas", versión española en pps. 121-59 de Ross Anderson, A. (ed.), *Controversia sobre mentes y máquinas*. En este trabajo se introduce la noción de unos "estados estructurales", derivados de las condiciones físicas, estados que serían, también, "estados lógicos".

⁴⁴ Skinner, B.F., *Sobre el conductismo*, p. 35.

en las cuales ese comportamiento ha estado determinado o va a ser emitido próximamente. Identificar la *propensión* del hambre, la *cualidad* de sentirse hambriento, con el amasijo de sensaciones fisiológicas concurrentes a esa experiencia mental, es un "error categorial", sin perjuicio de que su exacta correlación, el riguroso emparejamiento entre sucesos corporales y las observaciones ("mentales") que el sujeto emite sobre ellas, interese a la psicofisiología. Pero es un error encontrar en lo *cualitativo* un atributo o rasgo de identificación propio de un reino de lo mental. Lo mental no es inaccesible o incomunicable por su cariz monádico-cualitativo. Es esa sustantivación de las *cualidades mentales* la que hay que romper, y encontrar que en ellas subyace una operatoriedad práctica, controlada por circunstancias externas. Los argumentos que, en relación con "el mundo interno de la emoción y la motivación" Skinner manejó⁴⁵ 46 son muy parecidos a los de Ryle, en tanto que el paso de lo predicativo-cualitativo-mental a lo relacional-conductual, era un paso lingüístico. Lo tradicional había sido convertir nombres en adjetivos y verbos. En una palabra: sustantivizar. Según Skinner⁴⁶ 47:

"La internalización del intelecto se empareja plenamente con la de la vida de la emoción y la motivación. Persiste el hecho de pasar del comportamiento observado a un mundo imaginario interno. Algunas veces es poco más que una práctica lingüística. Tendemos a hacer nombre de adjetivos y verbos, y entonces tenemos que buscar un sitio para las cosas que dicen representar los nombres. Decimos que una soga es fuerte. Y pronto estamos hablando de su fortaleza. A una clase particular de fuerza la llamamos tensión, y luego explicamos que la soga es fuerte porque posee fuerza de tensión. El error es menos obvio, pero más problemático, cuando la materia es más compleja".

Hay una parte de la memoria que los psicólogos han venido denominando *memoria episódica* y que, según Tulving⁴⁷, es el recuerdo de experiencias personales y de sus conexiones temporales, frente a la memoria semántica entendida ésta como sistema para recibir, retener, y transmitir información acerca del significado de palabras, conceptos o clasificaciones de los conceptos. Pues bien, esta parte o faceta episódica de la memoria parece estar mucho más ligada al carácter cualitativo y emocional de muchas experiencias humanas, en las que el almacenamiento de conceptos abstractos no es tan importante. En la Inteligencia Artificial, este concepto de *memoria episódica* no ha gozado de mucho éxito, y sin duda esto se debe a que en los ordenadores se han construido representaciones de tipo conceptual, abstracto: representaciones "impersonales", en una palabra. Abelson⁴⁸ es uno de los militantes de la corriente computacionalista que se ha enfrentado con este dilema de la dualidad humana de las memorias (episódica y semántica), y que más convencido está de la importancia de las memorias episódicas. De hecho, cuando este autor colaboró con Schank⁴⁹ en los *scripts*, estos fueron concebidos como *sistemas de creencias*, más que *sistemas de conceptos* y si bien hemos dicho que las *técnicas* informáticas de

⁴⁵ Sobre el conductismo, cap. X.

⁴⁶ *Op. cit.* p. 149.

⁴⁷ Tulving, E., "Episodic and semantic memory", pps. 401-2, en Tulving, E. y Donaldson, W. (eds.), *Organization of memory*, Academic Press, N. York, 1972. Y también, "Encoding specificity and retrieval processes in episodic memory", *Psychological Review*, (1973), **80**, 352-73.

⁴⁸ Abelson, R.P., "Concepts for representing mundane reality in plans", pps. 273-309, en Bobrow, D.G. y Collins, A., *Representation and understanding*, Academic Press, Londres, 1975; en especial pps. 306-7.

⁴⁹ Schank, R.C. y Abelson, R.P., *Scripts, goals, and understanding*. Erlbaum, Hillsdale, 1977.

representación pueden ser semejantes en uno y otro caso. La teoría que inspira esa programación no parte ya de un "sujeto cognoscente" tomado *in abstracto*, libre de errores, infalible, impersonal, esto es, la máquina ya no se concibe como una suerte de sujeto epistémico universal. La especificidad antropológica con que se quiere barnizar un programa de ordenador es ahora la especificidad que al *script* le da el ser un SPI ideológico, no del todo racional, emocionalmente susceptible, etc. Así, en su simulación de "sistemas ideológicos", Abelson dice que éstos constituyen un tipo de *sistema de conocimiento* guiado no ya por los *datos*, sino por las conjeturas o las *hipótesis*⁵⁰. Así pues, el *script*, como el programa SAM del "guión del restaurante", simula situaciones sociales, de las que sin duda emanan los *recuerdos personales o episódicos*, así como las actitudes o sentimientos hacia ellos. Pero ahora, el *actor* es un sujeto "procesador de creencias", incluso de creencias ideológicas. Se supone que el programa de Abelson simula el modo en que el SPI se sumerge en las situaciones, siendo al tiempo actor y observador ("observador" en un sentido muy próximo al de "observación vicaria") pero está claro que esa "observación" no es sino un input que va a desencadenar un output de una cierta clase, para cuyo desencadenamiento no es necesario apelar a un *medio* de "experiencias personales" o episódicas. La referencia *personal* que el ordenador aparentemente efectúa ha sido *delegada* por sus programadores. Si estoy dialogando con el programa ELIZA, el famoso terapeuta mecánico de J. Weizenbaum, y creo que el programa piensa que yo soy un neurótico, esto no será debido a una *penetración (o Verstehen)* que mi mente ha operado sobre la mente de la máquina. Yo no "leo" la mente de ELIZA, solamente leo sus respuestas escritas en la pantalla y, al igual que sucede en el test de Turing, si nadie me dice que le cuento mis traumas a una máquina, ciertamente podré haber sido engañado por pensar erróneamente que conversaba con una persona, pero ser víctima de un engaño no demuestra en absoluto que la máquina posea algún tipo de inteligencia (a lo sumo, demuestra que yo carezco de ella). Una situación *à la Turing* en la que el programa ELIZA -o alguna versión perfeccionada suya- estuviera *en diálogo* con un sujeto humano no hará sino mostrar el carácter dialógico -o dialéctico- de la situación que se quiere simular, situación que generan dos o más sujetos dialogantes y de la que va a emanar toda emotividad, no como sustancia inmanente al sujeto, sino como conjunto de observaciones (apotéticas) referidas a y concurrentes con la conducta del otro y la propia, ambas entrelazadas en una *gestalt* o situación global.

⁵⁰ Abelson, R.P., "Concepts for representing...", p. 274.