

# Filosofia da Mente e ciências cognitivas<sup>1</sup>

António Zilhão

De um modo geral, o público interessado em Filosofia sabe que a Filosofia da Mente contemporânea se encontra estreitamente ligada às ciências cognitivas. Porém, nem sempre a natureza desta ligação é efectivamente conhecida. Irei tentar expô-la neste ensaio. Para o fazer, começarei por apresentar os pressupostos básicos sobre os quais ambas estas disciplinas assentam; tentarei, em seguida, tanto mostrar quais são as consequências mais importantes que deles se deixam extrair, como delinear o problema crucial que elas necessitam de resolver; salientarei ainda o facto de que a Filosofia da Mente contemporânea e as Ciências cognitivas se distinguem por serem sensíveis a diferentes aspectos deste problema crucial; finalmente, exporei as dificuldades que a Filosofia da Mente experimenta no tratamento que faz do aspecto do problema ao qual ela é especialmente sensível.

## I

Começemos pela consideração dos pressupostos. A Filosofia da Mente contemporânea e as Ciências cognitivas partilham o mesmo conjunto de pressupostos essenciais. Estes são basicamente dois. O primeiro é uma concepção naturalizada de o que é a mente. Esta concepção deixa-se exprimir por meio de uma definição como a seguinte: uma mente é um dispositivo centralizado de controlo, o qual se encontra realizado nalguma estrutura física, e que comanda o modo como um certo número de objectos físicos complexos se comportam nas suas interacções com o mundo. O segundo é um determinado ponto de vista acerca dos seres dotados de mente. Este ponto de vista deixa-se caracterizar do seguinte modo: os seres dotados de mente podem, e devem, ser adequadamente descritos como sistemas cognitivos. Para que esta caracterização deste segundo pressuposto possa ter um conteúdo substantivo, é necessário complementá-la com uma elucidação da expressão 'sistema cognitivo'. Que é, então, um sistema cognitivo? Basicamente, um sistema cognitivo é um sistema físico cujo comportamento obedece à seguinte lei geral: se o sistema tem um determinado objectivo O, ou desejo D, e se tem armazenadas no seu interior determinadas crenças, ou itens de conhecimento, C1,..., Cn, de acordo com os quais a efectuação de uma acção A conduz à obtenção do objectivo O (ou à realização do desejo D, consoante a terminologia que se preferir), então o sistema efectuará a acção A.

Repare-se que esta lei é uma expressão daquilo a que também se poderia chamar um princípio de racionalidade agencial. Deste modo, a caracterização dos seres dotados

---

<sup>1</sup> Uma primeira versão deste artigo foi publicada em versão electrónica no n.º 6, de Dezembro de 2001, da revista virtual *Intellectu*, a qual se encontra online no site <http://www.intellectu.com/>

de mente como sistemas cognitivos não é senão uma outra forma de os caracterizar como agentes racionais. Todavia, ao contrário do que acontecia com a tradição filosófica clássica, o princípio de racionalidade agencial é agora para ser visto como definindo uma lei natural como qualquer outra, a qual obtém para um conjunto dado de sistemas físicos. Ser uma lei natural como qualquer outra significa basicamente ser uma proposição condicional, o conhecimento da verdade da qual nos permite fazer essencialmente duas coisas. Por um lado, permite-nos prever o comportamento do sistema descrito por ela desde que as condições iniciais do mesmo sejam dadas (neste caso, as condições iniciais seriam as crenças e os desejos do sistema cognitivo particular alvo de consideração num dado momento). Por outro lado, permite-nos explicar *ex post facto* os comportamentos efectivamente exibidos pelo sistema.

Da conjunção da definição de 'mente' apresentada acima com a definição de 'sistema cognitivo' e com a tese de que os seres dotados de mente podem ser adequadamente considerados como sistemas cognitivos, segue-se que a mente humana admite ser definida como o dispositivo de comando, fisicamente realizado no cérebro humano, que controla e dirige o modo como os seres humanos, concebidos como sistemas cognitivos, se comportam nas suas interacções uns com os outros e com o mundo. Este é, portanto, o ponto de partida partilhado pela Filosofia da Mente contemporânea e pelas Ciências cognitivas no estudo que ambas empreendem da mente humana.

Repare-se que estas caracterizações nos colocam de imediato no terreno do mundo físico e das suas leis. Em particular, não se encontra nelas qualquer referência à consciência e aos seus conteúdos, a experiência subjectiva dos quais é considerada pelo ponto de vista tradicional, cartesiano-fenomenológico, como o único ponto de partida possível para qualquer investigação séria acerca da mente. Com efeito, os seres humanos encontram-se aqui caracterizados como sistemas físicos entre outros e o conceito de mente é um conceito que refere o dispositivo de comando, realizado numa estrutura física, que regula o comportamento daqueles organismos ou objectos físicos dotados do género de complexidade comportamental que nos permite considerá-los como sistemas cognitivos. Se este é o caso, então tudo aquilo que for dito acerca dos seres humanos e das suas mentes tem que ser compatível com o mundo físico e enquadrável nas suas leis.

O chamado problema 'mente-corpo', originado pela divisão ontológica do real entre *res extensa* e *res cogitans* proposta pelo cartesianismo, encontra-se assim logo desde o início arredado das preocupações da Filosofia da Mente contemporânea. De facto, não é necessário clarificar como é possível haver interacção causal entre o mundo extenso e o mundo inextenso pela simples razão de que toda a investigação acerca da mente é conduzida *ab initio* no interior do mundo extenso, do qual nunca se pretende sair. Todavia, este problema tradicional deixou um herdeiro. Trata-se do problema 'discurso mental-discurso físico'. Isto é, o problema de determinar que relação obtém entre estes dois tipos de discursos. Em particular, e uma vez que não se pretende sair do mundo extenso, trata-se de clarificar qual é a referência dos predicados mentais e porque processo se podem avaliar com sentido as condições de verdade das frases mentais de carácter declarativo.

Deste modo, a consciência e a subjectividade deixam de se encontrar no lugar fundamental que tinham na Filosofia da Mente tradicional. Isso não significa, porém, que não se empreendam esforços para tentar compreendê-las no interior do novo paradigma. O que isso de facto significa é apenas que o lugar adequado para empreender esses esforços se encontra deslocado. Em vez de problema primeiro, a consciência e a subjectividade são agora vistas como fenómenos associados a sistemas cognitivos extremamente complexos. A sua compreensão nunca poderá por isso preceder a compreensão da cognição menos complexa. Na ordem natural da investigação, a compreensão da consciência e da subjectividade surgem agora no fim e não à cabeça. Mudar o lugar de um objecto numa sequência não é o mesmo que eliminá-lo dessa

sequência. Por isso, não é justo reagir-se a esta postura metodológica acusando, como frequentemente acontece, as Ciências cognitivas e a Filosofia da Mente contemporânea de terem desistido de tratar os Grandes Problemas (com maiúscula) levantados pela Filosofia da Consciência tradicional. Até porque continua a haver um Grande Problema (com maiúscula) que mantém um lugar de destaque, logo à cabeça da nova ordem de prioridades. Vejamos qual é.

Um dos pressupostos destas disciplinas é, como vimos, o de que os seres humanos podem e devem ser vistos como sistemas cognitivos complexos. Ora, a definição *supra* de 'sistema cognitivo' faz apelo ao conceito de 'conhecimento' e a um princípio de racionalidade. O que quer dizer que a conjunção desta definição com o ponto de vista naturalizado associado à definição de mente acima apresentada levanta de imediato o seguinte problema: de que estamos nós a falar, quando falamos em conhecimento e racionalidade nestas circunstâncias, isto é, fora da consciência e dentro do mundo extenso?

Na realidade, se é a posse de conhecimento que faz com que um organismo possa ser adequadamente considerado como um sistema cognitivo e se, como vimos acima, a posse de conhecimento se deixa clarificar em termos da posse de itens de conhecimento ou conhecimentos particulares, aos quais se chama também crenças, e da posse de objectivos, aos quais se chama também objectos de desejo, então a pergunta feita acima pode ser parafraseada do seguinte modo: de que estamos nós a falar quando atribuímos crenças ou desejos a um sistema físico complexo? Por outro lado, como é que um dispositivo físico colocado algures no interior de um sistema cognitivo, por muito complexo que seja, pode adquirir, deter e usar esses itens de conhecimento por forma a controlar e dirigir o comportamento desse organismo? Ou, o que é, na realidade, o mesmo dito por outras palavras: como é que estruturas físicas imersas no mundo da causalidade eficiente podem comportar-se de acordo com a lei geral apresentada acima, a qual é, como foi mencionado, uma lei de racionalidade?

Estas duas perguntas são Grandes Perguntas. E o novo paradigma tem que ser capaz de lhes responder logo de entrada. Mais ainda, ele tem que ser capaz de fazê-lo sem se deixar enredar nos paradoxos solipsistas engendrados pelo ponto de vista cartesiano. Ora, estas duas Grandes Perguntas convergem, por sua vez, para um único Grande Problema. Vejamos com um pouco mais de detalhe qual.

Um sistema cognitivo é um sistema que possui crenças (ou itens de conhecimento) e desejos (ou objectivos). Crenças e desejos são tradicionalmente definidos pela Filosofia da Linguagem como atitudes proposicionais. Por sua vez, a expressão 'o sistema S tem uma atitude proposicional A(p)' é algo que pode ser caracterizado do seguinte modo: o sistema S encontra-se numa determinada relação R com a proposição p. A natureza da relação R é expressa pelo verbo da atitude (nomeadamente, para o nosso caso, crer ou desejar) e o conteúdo do objecto da atitude é dado pelo sentido da proposição que o caracteriza. Ora bem, proposições são representações de estados de coisas, actuais ou fictícios. Isto é, elas são acerca desses estados de coisas. Têm, portanto, a propriedade de ser dotadas daquilo a que Brentano chamou de intencionalidade. Acontece, porém, que, segundo Brentano, esta propriedade nunca poderia ser exemplificada de forma independente por objectos do mundo extenso. Estes limitar-se-iam a ser o que são, não remetendo nem deixando de remeter intrinsecamente para o que quer que seja acerca do que eles poderiam hipoteticamente ser. A intencionalidade encontrar-se-ia, assim, reservada para o domínio do mundo mental inextenso, no qual se encontrariam em actividade poderes desconhecidos no mundo natural, como, por exemplo, o poder de conferir sentido a objectos naturais que não o teriam intrinsecamente.

Assim, se, por um lado, o novo paradigma cognitivo despromoveu a consciência de fenómeno fundador do domínio do mental para algo cuja compreensão se alcançará com a acumulação de conhecimento acerca desse domínio, concebido independentemente dela, por outro lado, o esclarecimento da intencionalidade enquanto fenómeno físico é uma das condições da sua própria possibilidade. Com efeito, o tipo



de relação que um organismo tem com uma proposição, o qual se deixa caracterizar por uma dada atribuição ao mesmo de um género de atitude proposicional, pode ser adequadamente caracterizado por meio do uso que o organismo faz do conteúdo dessa proposição. Mas a especificidade semântica desse conteúdo tem que ser independente desse uso para que a identificação da natureza da relação seja possível. Isso significa então que é necessário que um sistema cognitivo seja capaz de produzir, albergar e usar de forma independente representações de estados do mundo, actuais ou potenciais. Por outras palavras, isso significa que um sistema cognitivo, que, como vimos, é um sistema físico complexo, tenha que ser um sistema intencional.

A clarificação de como é possível haver conhecimento fora da consciência, no mundo extenso, terá então que fazer-se por meio da clarificação de como é possível haver intencionalidade fora da consciência, no mundo extenso. Este é o Grande Problema, daqueles que se escrevem com maiúscula, que tanto a Filosofia da Mente contemporânea como as Ciências cognitivas têm que enfrentar logo desde o início.

## II

Como é então possível haver intencionalidade fora da consciência?

A resposta a esta questão tem que ser dada ao longo de duas dimensões.

A primeira destas dimensões é aquela que incide sobre o esclarecimento da seguinte questão: quais são, ou quais podem ser, os veículos materiais realizadores da intencionalidade num sistema físico. A segunda destas dimensões é aquela que incide sobre o esclarecimento da seguinte questão: qual é o processo por meio do qual esses veículos materiais, quaisquer que eles sejam, adquirem a capacidade de serem acerca daquilo acerca de que são. Aqui verificamos que existe como que uma divisão de trabalho entre as Ciências cognitivas e a Filosofia da Mente. As primeiras têm-se preocupado mais com o esclarecimento da primeira dimensão acima referida, provavelmente por estarem mais associadas a uma tradição de engenharia; a segunda tem-se preocupado mais com o esclarecimento da segunda dimensão, talvez por estar mais associada a uma tradição de esclarecimento conceptual.

Mas comecemos por abordar a primeira dimensão. Quais são os veículos materiais da intencionalidade num sistema cognitivo? A melhor forma de se tentar encontrar uma resposta a esta pergunta é procurando enumerar os constrangimentos a que uma estrutura interna de veículos materiais terá que obedecer para poder representar estados de coisas externos. Uma vez estes constrangimentos conhecidos, então será mais fácil procurar veículos que os satisfaçam. Comecemos então por enumerar estes constrangimentos.

Em primeiro lugar, uma estrutura deste género tem que ter um potencial de diversidade mínimo. Só assim será possível 'passar para dentro' uma quantidade suficiente de diversidade externa de tal forma que a operação de representação possa ser útil.

Em segundo lugar, uma estrutura deste género tem igualmente que ter uma plasticidade mínima. O mundo externo está num constante fluxo e as diferentes unidades que o compõem estão constantemente a sofrer transformações; é por isso necessário que a estrutura que vai servir de veículo à representação possa acompanhar minimamente essas transformações por meio da capacidade de produzir ela própria transformações sobre as suas unidades ou átomos de representação as quais possam acompanhar as transformações externas.

Em terceiro lugar, uma estrutura deste género tem que ter um mínimo de estabilidade. De facto, o processo representativo terá fatalmente que se desenrolar no tempo e é por isso necessário que os veículos utilizados permitam ao sistema reter conteúdos durante o período de tempo necessário para que um processo de transformação se possa completar. Por outro lado, conteúdos gerados num determinado contexto podem ser úteis noutros contextos. Assim, há toda a vantagem em que esses

veículos possam ser retidos sem degradação, ou com o mínimo de degradação possível, durante longos períodos de tempo. Veículos muito instáveis não permitiriam nem uma coisa nem a outra e a putativa informação que eles conteriam não seria fiável para o sistema que os albergasse.

Em quarto lugar, uma estrutura deste género tem que ser minimamente controlável. De facto, na medida em que as capacidades representativas de um sistema cognitivo estejam ao serviço da satisfação das suas necessidades e sirvam, por isso, para controlar o seu comportamento, é necessário que o sistema possa recorrer a elas quando delas necessita e que, portanto, a sua dinâmica seja controlável pelo sistema.

Que espécies de coisas conhecemos nós no mundo que sejam capazes de satisfazer estes constrangimentos? Se pensarmos apenas nas duas primeiras propriedades atrás enumeradas, há uma resposta que parece quase impor-se por ela própria, nomeadamente, a resposta 'os sistemas de símbolos'. Um sistema de símbolos é algo que, por definição, tem uma diversidade mínima de base e que, sobre essa diversidade de base, permite o desenvolvimento de uma diversidade e de uma plasticidade combinatorias praticamente inesgotáveis. Isto não quer dizer que toda a representação materialmente realizada o tenha que ser por intermédio de sistemas simbólicos. A representação analógica é, por exemplo, também possível. Todavia, é um facto que a representação analógica não permite a riqueza combinatoria suscitada pela representação simbólica. Ora, a própria variabilidade e composicionalidade do mundo exterior parecem requerer algo como essa riqueza combinatoria. Deste modo, sofisticação representacional e representação simbólica parecem andar de braço dado. Se este é o caso, então a questão que a seguir se põe é a de tentar descobrir que estruturas físicas com as duas propriedades subsequentemente mencionadas podem servir de veículos materiais a sistemas autónomos de símbolos.

Antes de apresentar uma resposta a esta pergunta gostaria ainda de salientar uma outra vantagem inerente a um sistema de símbolos. É a seguinte. Uma estrutura física que esteja a servir de veículo a representações num dado momento e que, por esse facto, esteja envolvida em processos de combinação de elementos diversos e de transformação dessas combinações efectua o seu trabalho na região localizada do espaço que ocupa. Todavia, se o sistema cognitivo for minimamente complexo, é importante que haja informação previamente adquirida guardada e pronta a ser usada. Essa informação tem, porém que se encontrar guardada numa região diferente do espaço disponível para a execução das tarefas cognitivas, uma vez que há um limite físico à quantidade de informação que se pode encontrar codificada numa mesma região do espaço físico. Assim, é importante que estejam disponíveis os meios locais para, no decorrer de uma dada operação de representação, se ter acesso a estruturas representacionais guardadas numa localização diferente e se poder transportar essas estruturas para o local onde elas estão momentaneamente a ser necessárias. Ora, este é um trabalho que parece especialmente talhado para ser despoletado pela presença de exemplares simbólicos inseridos no contexto da própria actividade representacional.

Regressemos então à última pergunta que formulámos acima. Que estruturas físicas existem que tenham as duas últimas das quatro propriedades acima mencionadas e que possam servir de veículo a actividades representacionais levadas a cabo por sistemas autónomos de símbolos?

Desde os anos 50 que tem sido possível construir estruturas físicas com estas características. Estou a falar, como já deve ter-se tornado óbvio, dos computadores. Pela forma como estão construídas, estas máquinas conseguem fazer com que um conjunto de propriedades próprias dos circuitos eléctricos sejam capazes de realizar fisicamente todas as propriedades que acabámos de passar em revista. Aliás, além destas, elas realizam também uma propriedade fundamental para a representação do conhecimento e que ainda não foi mencionada, a saber, a propriedade da conectividade.

Com efeito, dizer que um sistema cognitivo possui crenças não é o mesmo que dizer que esse sistema possui uma colecção avulsa de representações. Nós estamos



habituaados a considerar que a posse por um agente racional de uma crença implica a posse por esse agente racional de um conjunto de outras crenças. Parece haver uma conectividade essencial entre os elementos de conjuntos identificáveis de crenças a qual é a responsável pelo facto de que aquilo que nós, na realidade, habitualmente fazemos quando atribuímos uma crença a alguém seja atribuir-lhe, mesmo que apenas implicitamente, padrões de crenças e não uma crença individualizada. Assim, qualquer que seja a natureza dessa conectividade é necessário que ela seja realizável pelo veículo material da representação.

De entre os diferentes tipos de conectividade possíveis de serem materialmente implementados numa máquina simbólica, a conectividade lógica tem merecido grande atenção tanto por parte da Filosofia da Mente contemporânea como pela parte das Ciências cognitivas. Isto acontece não só porque ela é efectivamente representável num *medium* computacional, mas também porque ela é um meio mecânico efectivo e facilmente controlável de gerar crenças novas a partir de crenças pré-existentes e de, dessa forma, permitir a construção de vastos sistemas de crenças baseados numa experiência bastante limitada. Ora, algo como esta faculdade parece de facto existir na mente humana. A representação da conectividade existente num qualquer sistema de crenças por meio de uma lógica tem ainda uma outra vantagem. A inferência lógica pode ser descrita como um método de optimização da verdade. Deste modo, se houver algum processo por meio do qual se possa garantir que as crenças primitivas de um sistema cognitivo são verdadeiras, pode ter-se a certeza que as crenças delas derivadas por meio de inferências lógicas preservam essa verdade, o que constitui, em princípio, uma vantagem para o sistema que as alberga.

As ciências da computação mostraram assim que era possível construir máquinas que albergassem representações simbólicas e implementassem materialmente um processamento autónomo e consistente das mesmas. Com isto ficou demonstrado que é possível ter representações e processamento de representações fora da consciência e independentemente dela. Com base nesta constatação, tanto as Ciências cognitivas como a Filosofia da Mente contemporânea desenvolveram a hipótese analógica fundamental na base da qual pretendem compreender a estrutura da mente humana. Esta hipótese é a de que o cérebro humano contém em si uma máquina deste género, realizada num outro *medium*, nomeadamente num *medium* neuronal. Essa máquina seria a mente humana. Dado o que ficou dito atrás, pode considerar-se que a colocação desta hipótese assenta na aplicação ao conhecimento empírico efectivamente disponível de um princípio inferencial habitual no raciocínio científico: o princípio da inferência para a melhor explicação. Este género de inferência é, porém, falível. Só o eventual êxito da linha de investigação baseada na hipótese assim obtida poderá, no futuro, validá-la.

### III

Passemos agora para a tematização da segunda dimensão da resposta à pergunta 'Como pode haver intencionalidade fora da consciência?'. Isto é, partindo do princípio que pelo menos alguns veículos materiais da intencionalidade se encontram identificados, em termos de que é que é possível compreender que cada uma das unidades ou átomos representacionais realizadas nesses veículos seja acerca daquilo acerca do que ela é suposta ser? Repare-se que esta é uma pergunta que pode ter uma resposta trivial. De facto, se o domínio que considerarmos for o das máquinas simbólicas efectivamente construídas por nós até ao momento, o sentido das unidades ou átomos representacionais presentes nas representações que elas albergam é o que é porque nós decidimos que assim seja. Mas o caso que aqui nos interessa considerar não é aquele para o qual existe uma resposta trivial deste género. O caso que nos interessa considerar é o daqueles sistemas cognitivos, como nós próprios, acerca de cujas representações se pode dizer que possuem uma intencionalidade primitiva e não uma

intencionalidade derivada da que lhes é conferida pela intencionalidade primitiva de sistemas mais potentes.

Como dissemos acima, esta é a dimensão de abordagem do problema na qual a Filosofia da Mente contemporânea se especializou. Esta não tem, porém, qualquer êxito conceptual para nos oferecer neste domínio que seja comparável ao êxito tecnológico oferecido pelas ciências cognitivas no primeiro domínio. Mas a discussão prossegue e vale a pena tentar acompanhá-la. Nesta comunicação irei considerar duas tentativas rivais para resolver esta questão, bem assim como as dificuldades que cada uma delas enfrenta.

A primeira é aquela que tenta compreender a intencionalidade em termos da causalidade simples. De acordo com este ponto de vista, uma determinada crença *C* é a crença de que um determinado estado de coisas *E* é o caso pura e simplesmente porque *C* é um efeito de *E* no interior do sistema cognitivo. Esta perspectiva baseia-se numa intuição que parece ser perfeitamente legítima: tanto quanto se sabe, a actividade cognitiva resulta da interacção física entre o sistema cognitivo e o meio externo no qual ele se movimenta; ora, a existência dessa interacção pressupõe que haja contacto físico entre o sistema e aquilo que o envolve. Porque não então tornar simplesmente esse contacto naquilo em termos do que o conteúdo dos itens representacionais se deixaria esclarecer?

Deste ponto de vista então, se um determinado padrão simbólico *S* se forma no interior do aparelho representacional de um sistema cognitivo *K* na sequência de este ter sido causalmente impressionado por um dado aspecto da realidade *R*, então esse padrão simbólico *S* será acerca desse aspecto da realidade *R*. Evidentemente, é preciso ter algum cuidado na formulação deste ponto de vista. Nomeadamente, para determinar que uma dada conexão impacto causal/conteúdo representacional obtém, é necessário que duas outras condições suplementares se encontrem reunidas. A primeira é a de que apenas o aspecto da realidade *R* cause a emergência do padrão *S* em *K*. A segunda é a de que todas as instâncias de *R* têm que dar origem à constituição do padrão *S* sempre que exerçam um impacto causal sobre *K*.

A introdução destas condições suplementares não é gratuita. A primeira tem como objectivo evitar o problema a que, com alguma ironia, se chama o problema da “disjuntivite”. Com efeito, se a acção causal de aspectos diferentes da realidade *P*, *Q*, *R*, etc. sobre *K* tiver o mesmo efeito representacional no interior de *K*, isso terá que significar, pelos próprios dados do problema, que *S* representa a disjunção *P* ou *Q* ou *R* ou etc. A introdução da segunda condição suplementar tem como objectivo evitar que se possa alegar que *S* representa apenas os elementos daquele subconjunto dos fenómenos de tipo *R* que têm a propriedade de terem de facto impressionado o aparelho cognitivo *K* até ao momento presente e não o aspecto da realidade *R* de uma forma geral.

Apesar do seu aspecto *prima facie* razoável, satisfazer de uma forma consistente estas condições suplementares não é tarefa fácil. Com efeito, a nossa experiência com ilusões, *trompe l'oeil*, bonecos, réplicas, museus de cera, etc. e, sobretudo, a nossa experiência com os nossos próprios erros de identificação, mostra-nos claramente que a primeira condição nunca poderá ser satisfeita sem algum género de qualificação.

Por outro lado, passada a primeira impressão de plausibilidade, a satisfação da segunda condição tão-pouco parece ser exequível. Na realidade, é de todo impossível verificar efectivamente se todas as instâncias de *R* dão ou não origem ao padrão *S* sempre que exerçam um impacto causal sobre *K*. Muitas destas instâncias ocorreram, presumivelmente, no passado sem nunca terem tido a oportunidade de afectarem causalmente *K*, muitas ocorrerão previsivelmente no futuro sem tão-pouco terem essa oportunidade, muitas outras ocorrem no presente em localizações que as impedem de ter qualquer impacto causal sobre *K*.

A estratégia normalmente seguida pelos defensores desta primeira perspectiva para tentar resolver estes problemas consiste basicamente na tentativa de estabelecer a existência de uma dependência assimétrica tanto entre representações certas e



representações erradas, isto é, entre a formação de padrões S e padrões não S por resposta a R como entre a formação de padrões não S e padrões S por resposta a não R, como entre interpretações certas e interpretações erradas do conteúdo das representações S efectivamente formadas como efeito de alguns Rs. Esta ideia de dependência assimétrica deixar-se-ia, por sua vez, elucidar em termos de duas ideias. A primeira seria a de que existiria uma correlação regular entre as representações S e os aspectos da realidade R que os padrões S realmente representariam, a qual estaria todavia ausente no caso das representações não S formadas como resposta à exposição a aspectos da realidade R, bem assim como entre as representações não S de aspectos da realidade não R realmente representados por padrões não S, a qual estaria ausente no caso das representações S formadas por resposta a aspectos da realidade não R. A segunda seria a de que existiria uma correlação natural entre as interpretações certas do conteúdo das representações S efectivamente formadas por resposta a alguns Rs que não se encontraria presente nas interpretações erradas do conteúdo dessas mesmas representações.

Esta resposta sofre, porém, de uma dificuldade. Para que a primeira distinção possa ser clarificada em termos do conceito de regularidade, é necessário que esse conceito seja explicitado de um modo independente. Mas não existe um conceito independente de regularidade. Por outro lado, se ele for clarificado, por exemplo, em termos de preponderância estatística, isto é, em termos de frequência dominante num conjunto dado de instâncias, então os contra-exemplos não se fazem esperar. Deste modo, o género de regularidade a que acaba por se fazer apelo é aquele que se encontra associado à existência de condições óptimas, ou pelo menos adequadas, para que o contacto causal tenha o efeito que deveria ter; só que não parece haver qualquer outro processo de caracterizar essas condições que não seja o de as indexar à produção dos efeitos que tornariam a teoria verdadeira. O espectro da circularidade paira assim sobre esta resposta. O mesmo acontece, aliás, no caso da clarificação da segunda distinção em termos de naturalidade. De facto, a naturalidade de uma dada delimitação precisa de R depende ela própria de pressupostos acerca do modo como sistemas representacionais como K tendem a representar aspectos da realidade R.

A segunda das tendências rivais mencionadas acima é a perspectiva teleosemântica. Esta procura resolver os problemas enfrentados pela perspectiva causalista apelando para a teleologia. De acordo com este ponto de vista, o conteúdo de uma representação mental deverá ser dado não pelo aspecto da realidade que despoletou a sua mobilização pelo sistema que a alberga mas antes por aquele aspecto da realidade em resposta ao qual essa representação deveria ter sido mobilizada pelo sistema, independentemente de esse ter sido ou não o caso numa ocasião particular. O termo deveria deve aqui ser entendido no mesmo sentido que preside à sua utilização em contextos biológicos. Isto é, aquilo que os teleosemânticos dizem é que o padrão de símbolos S representa o aspecto da realidade R e não outro porque foi desenvolvido pelo sistema cognitivo para co-variá-lo com o aspecto da realidade R e não com outro.

A afirmação desta co-variação, por sua vez, é entendida como nada tendo de misterioso. Ela é semelhante à afirmação de que o pêlo branco dos ursos polares foi desenvolvido pelos ursos polares para co-habitar com a neve branca que caracteriza o ambiente em que eles normalmente vivem. O pêlo branco é assim um traço que os ursos polares actuais possuem porque a sua presença nos ursos polares passados em associação com as condições externas nas quais eles tiveram que viver facilitou a sobrevivência dos mesmos e, portanto, permitiu a existência actual de ursos polares dotados de pêlo branco. Evidentemente, isso não implica que se se transportar um casal de ursos polares para o Jardim Zoológico de Lisboa a cor do pêlo das suas crias se irá alterar, apesar de essa cor branca ter deixado de ser uma forma de co-habitar com qualquer aspecto do meio ambiente no qual esses ursos polares agora vivem. De modo semelhante, argumentam os defensores da perspectiva teleosemântica, transformações imprevistas na realidade envolvente podem levar a que padrões de símbolos não S



sejam mobilizados pelo sistema K em resposta a instâncias de R e a que padrões de símbolos S sejam mobilizados pelo sistema K em resposta a aspectos do mundo não R sem que isso constitua um problema.

O problema principal que a perspectiva teleosemântica tem que enfrentar é o de que parece ser contra-intuitivo fazer o conteúdo das representações mentais primitivas depender inteiramente de histórias passadas de selecção natural. Este facto costuma ser ilustrado com a metáfora da réplica perfeita. Esta consiste basicamente no seguinte. Se se imaginar que é possível, por meio de um qualquer prodígio tecnológico ou outro, criar um sistema intencional I' que seja uma réplica perfeita, molécula a molécula, de um outro sistema intencional efectivamente existente I, então, das duas uma: ou se teria que dizer, para preservar a teoria, que I' não teria albergadas no seu interior as representações mentais que se encontrariam em I, uma vez que os conteúdos mentais de I' não teriam tido uma história de selecção natural adequada, o que parece ser extremamente contra-intuitivo; ou se teria que dizer que, sendo I' uma cópia física perfeitamente fiel de I, então deveria ter os mesmos conteúdos mentais que I, apesar de nenhuma história de selecção natural ser responsável por eles, o que constituiria um contra-exemplo à validade da teoria.

A estratégia de resposta seguida pelos defensores do ponto de vista teleosemântico para lidar com esta objecção é basicamente a de aceitar a acusação de contra-intuitividade como válida quando se tem o conceito intuitivo de representação como pano de fundo e, simultaneamente, apresentar a sua perspectiva como uma reconstrução teórica do conceito de representação, a validade da qual depende apenas da utilidade teórica do conceito proposto e não da sua capacidade de clarificar a intuição. Pela sua própria natureza, esta estratégia fica refém dos sucessos empíricos futuros a que a reconstrução teórica proposta possa dar origem. Enquanto eles não aparecerem, ela tem que se resignar a ficar indefesa perante as acusações de implausibilidade que lhe são feitas.

Deste modo, o Grande Problema da intencionalidade continua em aberto. Enquanto não for resolvido, de uma forma favorável ou desfavorável às pretensões das Ciências cognitivas, nem a hipótese analógica subjacente ao ponto de vista da Filosofia da Mente contemporânea deixará de ser apenas uma hipótese, entre muitas possíveis, nem o pressuposto de acordo com o qual os seres dotados de mente devem ser apropriadamente considerados como sistemas cognitivos, isto é, como implementações físicas do princípio de racionalidade agencial mencionado no início, terá um conteúdo claro e empiricamente testável. O campo encontra-se assim disponível para o exercício da arte de argumentar por todos aqueles que se deixam sensibilizar por Grandes Problemas como este e que estão dispostos a mergulhar não o olho da mente mas antes as mãos da mente, se assim me posso exprimir, na massa de material conceptual e empírico já acumulada.

#### BIBLIOGRAFIA:

- Aristóteles, "Ethica Nicomachea" in McKeon, R. (ed.) *The Basic Works of Aristotle*. Random House, New York, 1941.
- Block, N. (ed.), *Readings in the Philosophy of Psychology*. Harvard U. P., Cambridge (MA), 1981.
- Brentano, F., *Psychologie vom empirischen Standpunkt*. Duncker & Humboldt, Leipzig, 1874.
- Cummins, R., *Meaning and Mental Representation*. The MIT Press, Cambridge (MA), 1989.
- Dennett, D., *Brainstorms – Philosophical essays on Mind and Psychology*. The Harvester Press, Brighton, 1981.
- Dennett, D., *The Intentional Stance*. The MIT Press, Cambridge (MA), 1988.

Descartes, R., *Méditations sur la Philosophie Première, dans lesquelles sont démontrées l'existence de Dieu et la distinction de l'âme et du corps*. Gallimard, Paris, 1953.

Fodor, J., *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. The MIT Press, Cambridge (MA), 1987.

Gardner, H., *The Mind's New Science: A History of the Cognitive Revolution*. Basic Books, New York, 1987.

Goldman, A., *Epistemology and Cognition*. Harvard U. P., Cambridge (MA), 1986.

Haugeland, J., *Artificial Intelligence: The Very Idea*. The MIT Press, Cambridge (MA), 1985.

Johnson-Laird, P., *The Computer and the Mind: An Introduction to Cognitive Science*. The MIT Press, Cambridge (MA), 1988.

Lycan, W., *Consciousness*. The MIT Press, Cambridge (MA), 1987.

Millikan, R. G., *Language, Thought and Other Biological Categories*. The MIT Press, Cambridge (MA), 1984.

Newell, A., *Unified Theories of Cognition*. Harvard U.P., Cambridge (MA), 1990.

Papineau, D., *Philosophical Naturalism*. Blackwell, Oxford, 1993.

Putnam, H., "Minds and Machines" in Putnam, H., *Mind, Language, and Reality: Philosophical Papers II*. Cambridge U.P., Cambridge, 1975.

Sellars, W., "Empiricism and the Philosophy of Mind" in Sellars, W., *Science, Perception, and Reality*. Routledge, London, 1963.

Simon, H., *Models of Man — Social and Rational*. Wiley & Sons, New York, 1957.

Turing, A., "Computing Machinery and Intelligence" in *Mind*, 59, 1950, pp. 433-460.