

PROGRAMAS E PROMESSAS: SOBRE O (AB-)USO DO JARGÃO COMPUTACIONAL EM TEORIAS COGNITIVAS DA MENTE

W. P. MENDONÇA

Universidade de Konstanz, Alemanha Ocidental

Segundo uma opinião hoje bastante difundida, teorias computacionais da mente do tipo comumente proposto na Psicologia Cognitiva e na Inteligência Artificial distinguem-se de tentativas anteriores pelo rigor dos instrumentos conceituais de que se servem: a introdução, para efeito de construção teórica na Psicologia, do vocabulário retirado das Ciências da Computação permitiria a elaboração de hipóteses precisas sobre os processos cognitivos. Este trabalho defende a tese de que tal opinião é insustentável. O exame do modo de funcionamento real da linguagem com a qual computadores são descritos mostra que a tentativa de descrição e explicação do agir inteligente em termos computacionais não vai além de uma assimilação superficial de níveis de linguagem radicalmente diferentes. Alguns processos fisiológicos envolvidos no exercício das chamadas faculdades da inteligência podem ser efetivamente descritos como processos computacionais. Contudo, a consideração de casos experimentais concretos da Psicologia Cognitiva revela que tais processos são associados às atividades da inteligência através de sua conexão com o agir intencional e sua base criterial. Como estes não podem ser significativamente representados como o resultado da observância de regras esquemáticas ou de procedimentos implementáveis "mecanicamente", as teorias computacionais da Psicologia Cognitiva não devem – ao contrário do que é afirmado por muitos de seus seguidores – ser vistas como uma investigação sobre "a natureza da inteligência". Ao contrário, elas devem ser vistas, no melhor dos casos (quando analisadas apropriadamente), como a descrição sistemática de algumas condições operativas do exercício da inteligência.

According to a widespread view, computational theories of the mind of the type commonly proposed in cognitive psychology and artificial intelligence differ from earlier conceptions in the accuracy of the conceptual instruments which they use: the introduction of the vocabulary of computational sciences in psychology would make possible the elaboration of precise hypotheses about the so-called cognitive processes, being, thus, equivalent to a methodical disciplinaton of our thinking on man as a being capable of thinking and speaking. I argue in this paper that this view is untenable. The examination of the actual way of functioning of the language with which computers are described shows that the attempt to explain intelligent action in terms of computer behaviour does not bring us beyond a superficial assimilation of radically different levels of discourse. To be sure, some neurophysiological processes which are involved in the exercise of the so-called faculties of intelligence can be actually described as computational processes. The consideration of concrete experimental cases of cognitive psychology, however, reveals that these processes are associated with the activities of intelligence through their connection with intentional action and its criterial basis. Since the latter cannot be significantly represented as the result of compliance with schematic rules or procedures which could be "mechanically" implemented, computational theories of cognitive psychology should not be seen – contrary to the opinion of many of their followers – as an investigation of "the nature of intelligence". Rather, appropriately analyzed, it should be seen as, at best, a systematic description of some operating conditions of the exercise of intelligence.

1. Quem conhece a história da Filosofia moderna sabe como tentativas de explicação mecanicista do agir humano conduzem, apesar de seu ponto de partida aparentemente "científico", mas graças à falta inerente de crítica (no sentido kantiano), a doutrinas metafísicas de conteúdo incontrollável. A mais recente entre tais tentativas é a Psicologia Cognitiva contemporânea ou o Cognitivismo, cujos defensores julgam dispor de um novo conceito de mecanismo capaz de resolver os enigmas filosóficos da interação entre corpo e mente (ou alma, ou espírito, ou como se queira). A alguns (a muitos) parecerá inadequado aproximar deste modo o Cognitivismo, isto é, o paradigma da pesquisa experimental e teórica contemporânea no âmbito da Psicologia Geral e da Inteligência Artificial, e as especulações filosóficas da tradição pós-cartesiana. Contudo, as raízes "metafísicas" da nova Psicologia Cognitiva são claramente reconhecidas por um de seus proponentes mais significativos, Jerry Fodor, o qual afirma: "A maior parte da Ciência Cognitiva é Filosofia redescoberta - e, eu penso, justificada" (Fodor 1981, p. 26).

Fodor vê na opinião empirista clássica de que seria possível dar conta da relação entre o pensamento e seu objeto através da concepção das "idéias" como representações mentais ("internas") que se assemelhariam aos objetos do pensamento uma versão primitiva, mas ainda assim antecipadora, em linhas gerais, das teorias cognitivas (Fodor 1978a, p. 60). O elemento novo nessas teorias é a metáfora do computador, isto é, a comparação da seqüência de estados "internos" que supostamente causam o comportamento "externo", concedendo-lhe, ao mesmo tempo, seu caráter inteligente, com o processamento de informações num computador digital. Todavia, Fodor reconhece também aqui a continuidade com a tradição filosófica que remonta a Descartes e Locke: "Na medida em que a teoria representacionista da mente é o conteúdo da metáfora do computador, a metáfora do computador antecede o computador em cerca de trezentos anos" (Fodor 1981, p. 26).

Os problemas desse novo empreendimento "científico" - bem como, aliás, da concepção filosófica que ele procura, implícita ou explicitamente, reabilitar - têm suas raízes no fato de que ele quer operar com explicações causais num âmbito onde elas são inadequadas por princípio. Neste sentido, seus problemas surgem de uma carência de "crítica do conhecimento".

2. Como ponto de partida para a justificação dessa tese oferece-se a consideração dos (na opinião de Fodor) *casos paradigmáticos* da explicação psicológico-cognitiva. Fodor elucida esse paradigma com base em um "silogismo prático": "João crê que choverá se ele lavar seu carro. João quer que chova. Assim, João age com a intenção de lavar o carro." Isto resume, tal é a opinião de Fodor, "a etiologia do 'comportamento de lavar o carro' atribuído a João [John's car-washing behaviour]." Isto quer dizer: "a ação de lavar o carro é um *efeito* da intenção de lavar o carro, e a intenção de lavar o carro é um *efeito* da interação *causal* entre as convicções

[beliefs] de João e seus interesses." (Fodor 1978a, p. 49; grifado por mim).

O uso das expressões "causa" e "efeito" não é concebido aqui como simples "façon de parler". Lidamos aqui, crê Fodor, com a "velha" causalidade (fiscalista). Aliás, não existe para ele uma "outra" causalidade. "Convicções" [beliefs], "vontades" e "intenções" *são*, de acordo com essa opinião, *processos* que estão numa interação causal entre si e causam finalmente o comportamento (nesse caso: o comportamento de lavar o carro). Essa é "a visão realista dos designata dos termos mentais em teorias psicológicas" (Fodor 1975, p. 4). Estados mentais, Fodor afirma além disso, "entram em jogo como agentes causais contribuindo para a produção do efeito comportamental" (Fodor 1981, p. 5).

Certamente, os estados mentais precisam então (se não se deseja admitir agentes causais puramente "espirituais") ser concebidos (interpretados) como *eventos* orgânicos, pois somente processos corporais têm um *efeito* (causal). De fato, Fodor fala de "processos neurológicos envolvidos na etiologia do comportamento" e afirma que eles satisfazem duas descrições diferentes: "uma descrição física, em virtude da qual eles se subordinam a leis causais, e uma descrição psicológica, em virtude da qual eles constituem passos na computação desde o estímulo até a resposta" (Fodor 1975, p. 74). Os processos orgânicos, continua Fodor, "satisfazem descrições como 'armazenar, aceitar, rejeitar, computar, etc, P'" (Fodor 1975, p. 75). Essa é — como Fodor a designa — a descrição "psicológica" dos processos neurológicos.

O problema de Fodor é, portanto, a reinterpretação da explicação cotidiana, "*teleológica*" da explicação do agir humano através do recurso a convicções, interesses ("utilities") e intenções como explicação *causal*. A pergunta fundamental para ele é: como é possível que os chamados estados mentais designados, por exemplo, pelas expressões "convicção" e "intenção" e atribuídos, segundo ele, a um "organismo", não *hipoteticamente* como *disposições comportamentais* (Ryle), mas sim *categoricamente* como *afirmações descritivas*, tenham *conseqüências* comportamentais? E isso quando se insiste, por um lado, que o termo "conseqüência" deve ser interpretado aqui causalmente porque, até onde eu possa perceber, nenhum outro modo de interpretação seria adequado ao caso" (Fodor 1978b, p. 153), e quando se quer, por outro lado, evitar o dualismo cartesiano.

Fodor julga encontrar a solução desse problema na caracterização abstrata das propriedades dos processos neurológicos. Tais processos, envolvidos na produção causal de movimentos corporais, podem ser descritos com categorias neurofisiológicas, mas também sob o ponto de vista de uma exposição estrutural de seu modo de funcionamento. Fodor compreende por isso a caracterização teórica do aparato neurofisiológico como um sistema de processamento de informações. O passo decisivo — e, como se verá mais tarde, funesto — é dado por Fodor quando ele afirma que os termos (teóricos) empregados nessa caracterização funcional designam os

“estados mentais”, as “atitudes proposicionais” do ser no qual ocorrem os processos neurofisiológicos descritos abstratamente. Dito de outro modo: eles seriam a reconstrução ou explicação do que entendemos normalmente por “convicção”, “intenção”, “vontade”, etc.

A teoria procurada por Fodor deve proporcionar um vocabulário tomado das Ciências da Computação com o qual o conteúdo proposicional de tais estados mentais — e isso quer dizer para Fodor: representações dos fatos acreditados (“believed”), intencionados, etc. — possam ser descritos. Para poder requerer causalidade (física) para os processos computacionais interpretados “realisticamente”, Fodor precisa considerar as representações (informações) como “implementadas” no sistema nervoso. “A informação”, diz ele nitidamente, “é dita estar disponível para o organismo quando o evento neural que a codifica é um dos determinantes causais do comportamento do organismo” (Fodor 1975, p. 52). Com vistas a esse objetivo, ele estabelece uma correspondência funcional hipotética entre fórmulas de uma linguagem computacional e estados físicos de um organismo (estados de seu sistema nervoso), de tal modo que a seqüência das fórmulas, como num programa de computador, conduza à descrição funcional do comportamento causado pelos estados físicos correspondentes.

A descrição funcionalista, “psicológica” (no sentido de Fodor), dos processos no sistema nervoso estaria completamente em ordem *se* seus defensores estivessem conscientes do que eles realmente fazem com isso: eles descrevem *eventos neurológicos* num nível de abstração funcional. Isso quer dizer: eles descrevem e “explicam” certos processos e formas de comportamento do corpo humano (sobretudo do cérebro) sob um ponto de vista funcional. E desde que tais descrições e explicações de seqüências regulares sejam confirmadas através de experimentos e observações, elas fazem parte da ciência experimental da Psicologia Comportamental, na qual certas regularidades são explicadas de modo (quase-)causal. Esse empreendimento poderia ser intitulado “Neurofisiologia Comportamental”. E ele estaria conceitualmente em ordem *se* seus proponentes estivessem conscientes do caráter restritivo dos tipos de comportamento que podem ser explicados dessa maneira.

Fodor — e nisso ele é representativo do que se faz em Psicologia Cognitiva — não vê, naturalmente, necessidade alguma de restrição. Ele identifica as descrições funcionais dos estados e processos cerebrais com os *significados* de termos psicológicos ou “mentais”, tais como eles ocorrem, por exemplo, na explicação *coridiana*, teleológica da “ação de lavar o carro atribuída a João”. O problema da exposição adequada da relação entre as ações “inteligentes”, intencionais de um indivíduo e seus elementos motivacionais (vontades, desejos, intenções etc.) sofre, assim, uma “internalização” na medida em que (a) ele é interpretado como (ou mesmo confundido com) a descrição teórica da relação entre os estímulos nervosos de um corpo e suas *reações* macroscópicas e (b) a consideração “externalista” das ações humanas no âmbito de relações interpessoais institucio-

nalizadas não é mais levada em conta (cf. von Wright 1976, p. 416f.).

A tentativa de reduzir *explicações teleológicas*, com as quais estamos, como seres sociais, completamente familiarizados, a *explicações causais* que somente podem ser justificadas através da teorização e experimentação científica implica, portanto, uma reorganização nada marginal na "geografia lógica" (Ryle) de nossa linguagem. A convicção de que lavar o carro leva a chover e a intenção de lavar o carro agora *designam*, segundo o ponto de vista do Cognitivismo, processos no sistema nervoso de João. Em outras palavras: *queremos dizer*, quando dizemos de alguém que ele tem uma intenção, que um certo processo ocorre *nele*, no seu cérebro. Fodor crê que qualquer concepção diferente dessa implica o abandono da idéia de que a intenção é uma *causa* da *ação* de João. "Causalidade entre fenômenos psicológicos" parece *ter de ser* "mediada fisiologicamente", se não quisermos admitir "um ser de alma espiritual" ("ein geisterhaftes Seelenwesen") — eis como Wittgenstein (1980, I § 906; 1967, § 611) caracteriza essa opinião.

3. As "espécies naturais" para propósitos de construção teórica na Psicologia Cognitiva incluem "algumas coisas que o organismo faz, algumas coisas que acontecem no sistema nervoso do organismo e algumas coisas que acontecem no seu meio ambiente" (Fodor 1975, p. 53). Se atentarmos para o fato de que a expressão "coisas que o organismo faz" se refere aqui ao comportamento intencional, "inteligente", o qual é com isso *assimilado* a eventos naturais e tornado objeto de explicações causais-mecanicistas, então podemos reconhecer que o Cognitivismo *mistura as categorias* inapropriadamente. Em particular, a idéia de uma teoria que almeja descrever as manifestações das chamadas "faculdades da inteligência" como se podem descrever as reações de um órgão (corporal) complexo resulta, como vimos, da interpretação indiferenciada de todo "porque" como introdução da especificação de uma causa.

Contudo, a convicção de João de que lavar o carro leva a chover e sua intenção de lavar o carro agora, porque ele gostaria que chovesse, são antes, dito kantianamente, "motivos da razão" ("Vernunftgründe") do agir, ainda que eles nos pareçam, no caso em questão, completamente insensatos. Ao invés de explicar causalmente a ação em questão eles fazem com que esta apareça sob uma certa luz que nos leva a compreender *por que* o agente age de tal modo — por que é "razoável" agir assim. Naturalmente, é possível imaginar ou relembrar circunstâncias em que a questão de saber se esta luz é a "luz da verdade" (Anscombe) se torna uma difícil questão. Mas é um erro *grave* tentar responder essa questão através do recurso a regularidades causais que poderiam ser constatadas no nível fisiológico.

Os proponentes do Cognitivismo cometem esse erro. A interpretação da intenção como um certo estado orgânico (cerebral) tem como consequência a idéia de que nem eu, que em casos normais determino com segurança os motivos do meu agir, nem o psicanalista, que pode descobrir nos

casos menos normais minhas intenções (inconscientes), poderia *saber* por que razão eu agi como agi. Somente o neurologista seria capaz disso. Essa consequência é de fato tirada por Daniel Dennett, ao defender o ponto de vista do Cognitivismo:

“A prática cotidiana de perguntar pelas razões baseia-se na *hipótese* de que nosso raciocínio consciente é uma manifestação confiável do processamento de informação que determina nossas ações e a psicanálise baseia-se na contra-afirmação de que não é assim” (Dennett 1969, p. 160).

Como não tenho acesso direto aos processos neurológicos, essa “hipótese” somente pode ser decidida por um neurologista. Como Dennett corajosamente afirma: “Os neurologistas poderão algum dia fornecer uma confirmação detalhada dessa hipótese” (Dennett 1969, p. 161).

Nesse ponto as “ondas da linguagem” (Wittgenstein) ameaçam crescer desmesuradamente. O significado da palavra “intenção” sofre uma transformação na teoria cognitivista. A dúvida se eu agi intencionalmente e qual a *verdadeira* intenção do meu agir *não* é resolvida através da constatação de uma regularidade causal. Os critérios para a verdade da locução “eu agi assim porque. . .” *não* são os critérios da descrição de um processo no meu cérebro. A causação do movimento corporal, do “aspecto externo de minha ação” (von Wright 1971, p. 86f.), através do evento cerebral que (*supostamente*) acompanha a formação de minha intenção não nos diz *coisa alguma* sobre a intencionalidade da ação a que corresponde, porque a intenção não está *contingentemente* ligada à ação como uma de suas condições causais.

O ponto importante, não observado pelo cognitivista, é que os “motivos da razão” do agir (intencional) não têm a mesma “lógica” das causas (dos eventos naturais). O cognitivista confunde essa distinção e faz surgir a impressão de que um mecanismo oculto da intenção é colocado em movimento por um estímulo nervoso, produzindo assim o agir correspondente. A tarefa da Psicologia Cognitiva consistiria então na descrição exata de tal mecanismo. O acordo incontestado entre cognitivistas de que há algo aqui a ser investigado e descoberto experimentalmente faz lembrar Hume, que chama a atenção para o fato de que os cavaleiros andantes, que saíram pelo mundo à procura de dragões e gigantes, jamais duvidaram da existência de tais monstros.

4. A teoria de Fodor apresenta-se, com razão, como uma variante moderna do Monismo Materialista. Atitudes proposicionais, isto é, o que acreditamos, desejamos, aprendemos e percebemos, formam para Fodor o domínio natural de teorias cognitivistas. Atitudes proposicionais consistiriam essencialmente em representações de seus objetos, implementadas no sistema nervoso, isto é, em representações fisiológicas de seus conteúdos

proposicionais. Fodor considera certos eventos orgânicos como tendo uma "estrutura sentencial" (uma "sentença mental" encarnada no cérebro como parte de uma "linguagem do pensamento"). A atitude global do organismo, caracterizada em termos computacionais, frente a essa 'estrutura sentencial' seria a atitude proposicional que atribuímos ao ser (humano) em cujo organismo os eventos mencionados ocorrem. (Tal atribuição tem normalmente a forma "X crê que p", onde p exprime o conteúdo proposicional). Dessa maneira, atribuições *personais* de atitudes proposicionais são remetidas, para efeito de explicação "científica", a processos *orgânicos*, isto é, "reduzidas materialisticamente". A tarefa é então especificar, para toda atitude proposicional, a representação interna e a relação orgânica global frente a ela, de tal maneira que a existência *dessa* relação frente a *essa* representação seja uma condição *necessária* e *suficiente* para a existência da atitude proposicional correspondente (Fodor 1975, p. 198). Como Dennett afirma: "Acreditar que a neve é branca [...] deve ser o mesmo que ter uma realização ['token'] da tradução em mentalês de 'a neve é branca' instalada de algum modo admirável na cabeça" (Dennett 1977, p. 75).

A postulação de representações internas *não pode*, porém, resolver o problema a que Fodor quer dar uma solução, a saber. como estabelecer a relação entre o pensamento e seu objeto (entre o desejo e a coisa desejada, entre o acreditar e o acreditado)? A pergunta geral "como é que as coisas dão entrada em nosso pensamento?" recebe somente de modo aparente uma solução com a especificação de uma representação interna ("mental") do objeto de referência do pensamento. Pois os mesmos problemas para os quais a representação interna deveria oferecer uma solução surgem com referência a esta. Como um símbolo adicional, a representação interna é, em sua relação projetiva com o mundo, tão enigmática quanto a expressão 'externa' do pensamento (o símbolo lingüístico). Na verdade, somente a práxis da linguagem como "técnica de projeção" coletivamente disponível estabelece a ligação entre pensamento e mundo. A representação mesma, seja ela "externa" ou "interna", só pode desempenhar seu papel contra o pano de fundo do uso da linguagem (pública): o que "é responsável" pelo fato de que alguma coisa no mundo representa outra coisa não é um elemento de um mecanismo, senão de uma práxis.

Como as conexões práticas de nossa vida, com as quais as representações estão entrelaçadas e das quais elas haurem seu significado, nunca aparecem corretamente no campo visual em Fodor, ele precisa, para explicar a relação entre representação e mundo, postular *representações de segundo grau*. O ponto de partida desse movimento teórico é a concepção comentada acima de que "representações internas são postas em correspondência com o que elas representam através de processos computacionais" (Fodor 1975, p. 173). Estes exigem, porém, um sistema de representações internas que estão para aquelas representações, cujas relação com o mundo elas computam, como *representações de representações*:

“Se os sujeitos relamente *calculam* como representações internas devem ser desenvolvidas, então estes cálculos, também, precisam ser definidos sobre representações, i.e., sobre representações de representações. Algumas propriedades da linguagem do pensamento devem, em resumo, ser representadas na linguagem do pensamento, desde que a habilidade de representar representações é, por suposto, uma condição prévia da habilidade de manipular representações racionalmente” (Fodor 1975, p. 112).

Fodor permanece aqui fiel à idéia fundamental do funcionalismo psicológico. Esta afirma ser possível uma análise de nossa vida inteligente tal que esta se revele composta de processos de transformação de informação tão ‘mecânicos’ como as operações primitivas de uma máquina calculadora; “processos tão estúpidos que o apelo a eles em explicações psicológicas não envolve petição de princípio” (Block 1980, p. 171). Esse é o conteúdo da metáfora segundo a qual a mente é um sistema de processamento de informações. Uma explicação da relação entre representação e mundo que remeta ao uso da representação no interior da práxis da linguagem normal não se adapta, evidentemente, a esse quadro. *Toda* referência (implícita ou explícita) a homens (inteligentes) e suas ações no âmbito de uma forma de vida socializada, à qual sempre recorremos para explicar como algo funciona como representação de alguma outra coisa, precisa, quando se quer permanecer no interior do funcionalismo psicológico, *desaparecer* com o tempo, pois senão estaríamos lidando com um caso evidente de petição de princípio.

Os processos que fazem do evento em mim uma representação precisam então ser *completamente analisados em termos de processos maquinais daquele tipo estúpido* aos quais Block se refere. Pois somente sob essa suposição pode-se afirmar que as representações são postas em relação com o mundo através de processos computacionais. Sobretudo, é preciso fazer desaparecer o ser humano (a pessoa), o qual (a qual) usa a representação e para o qual (a qual) a representação é uma representação. Fodor quer resolver esse problema com a hipótese de um sistema hierárquico de representações: quem usa, na sua opinião, as representações, quem as “manipula racionalmente” (expressão de Fodor), não é um agente humano, mas sim outras representações (internas).

5. Com isso abre-se imediatamente a perspectiva de uma construção de representações em níveis sempre crescentes; pois o único meio de evitar um ‘usuário’, para o qual as representações postuladas seriam representações, parece consistir em postular outras representações de nível mais elevado. Fodor não pode prever uma conclusão para essa construção: “Deve haver — talvez precise haver — algum fim para essa hierarquia de decisões racionais. Mas o fim não está à vista” (Fodor 1975, p. 173).

A possibilidade, porém, de que a cadeia de representações postuladas não possa ser completada seria funesta para Fodor. A representação interna

não pode ser concebida de tal modo que ela esteja frente a um "sujeito interno", que a "manipula racionalmente", na mesma relação na qual uma imagem (externa) está frente a uma pessoa. Isso seria uma simples duplicação do problema original em "foro íntimo". A eliminação desse sujeito interno, desse "homunculus", é um problema para toda teoria representacional da mente desde Descartes. E a *importância* da teoria, proposta por Fodor, de um sistema hierárquico de representações — Dennett chama-a de "um dos progressos conceituais mais promissores que já chegaram às mãos do filósofo" — consistiria em fornecer uma *solução promissora* para esse problema, ao imaginar um sistema representacional "sem usuário", isto é, "dando-nos um modelo de veículos de representações que funcionam sem agentes privilegiados para os quais elas seriam objetos de ocupação" (Dennett 1977, p. 73). Se "atrás" de cada nova representação postulada for necessário "imaginar" um usuário fictício (o "agente privilegiado" de Dennett), o qual, como um "homunculus" ou um "demônio", manipularia a representação inteligentemente, faria dela uma representação de algo no mundo e só *aparentemente* poderia ser eliminado pela postulação de representações de nível mais elevado, pois voltaria com a mesma força no próximo nível, *então* Fodor não poderia resolver seu problema e com isso realizar suas promessas funcionalistas. Ao contrário de muitos cognitivistas, Fodor sabe que o valor de sua teoria deve ser medido de acordo com o resgate dessa pretensão. Ele *precisa*, portanto, assumir que existe uma *camada fundamental de representações últimas* que representam um *fim absoluto* no sistema representacional hierárquico e que não trazem mais consigo um problema de "usuário" ou "homunculus".

Que o sistema representacional que se tem em mente aqui possa ser completado dessa maneira é, evidentemente, uma *suposição*. Fodor não tem a menor idéia de como o sistema representacional é estruturado — e menos ainda de suas últimas representações. É *preciso* que elas existam, diz ele continuamente. Mas os argumentos que ele fornece reduzem-se à afirmação de que os "working scientists" da Psicologia Cognitiva se orientam, ainda que implicitamente, por essa suposição. Não se diz, no âmbito dessa Psicologia, que o objeto de investigação são "the workings of the human mind" que a mente é um sistema de processamento de informações e devê ser descrita como tal? Não basta aqui que se faça teorias que mostrem como certos eventos, considerados por nós, *com base no agir de uma pessoa*, como representações do mundo são transformados no cérebro de forma análoga ao processamento de informações num computador. "Se a mente for, afinal de contas, um mecanismo, tais teorias não terão a última palavra em Psicologia" (Fodor 1975, p. 165). Isto quer dizer: a pessoa mesma e seu agir têm de poder ser "decompostos" em "processos estúpidos", ou então a mente humana *não* é um sistema de processamento.

Fodor permanece aqui conseqüente. Teorias cognitivas do reconhecimento de padrões ("pattern recognition"), por exemplo, do tipo "template-

matching” ou “feature-extraction” descrevem, para ele, no melhor dos casos, mecanismos que estão envolvidos *inter alia* na percepção. Elas não nos fornecem uma “teoria geral da percepção” (Fodor 1975, p. 193), porque pressupõem, através do uso de expressões “mentalistas” (intencionais) e ficções antropomórficas (por exemplo, “demônios”), exatamente aquelas habilidades (inteligentes) que elas deveriam explicar. Tais teorias formam para Fodor “uma parte significativa de qualquer explicação do fato de que somos tão bons no que fazemos”. Mas elas precisam ser completadas no futuro por outras teorias ainda não disponíveis. “Porque a longo prazo queremos uma teoria não de um processo racional ou outro, mas sim da racionalidade per se”. E isso requer “a eliminação global de predicados intencionais nas explicações psicológicas” (Fodor 1975, p. 175).

6. A afirmação de que é possível proceder a uma tal eliminação através da construção de um sistema representacional hierárquico sem usuário *parece* tornar-se plausível, para muitos até trivial, através de um argumento freqüentemente utilizado nesse contexto (embora não por Fodor). Não é a estrutura de dados (“data structure”) de um programa de computador — assim diz o argumento — o exemplo de uma estrutura hierárquica na qual a informação é processada sucessivamente por subprogramas ou sub-rotinas que o programa principal “chama” e “usa” por assim dizer diretamente, sem a mediação de um usuário? Subprogramas são eventualmente chamados por subprogramas mais abrangentes e esses por outros ainda mais abrangentes. Mas essa organização hierárquica tem um fim definitivo no programa principal, cujo “sistema executivo” determina a seqüência de operações e a chamada dos subprogramas. O programa principal (ou seu sistema executivo) não é evidentemente chamado por outro programa. Não nos daria então o sistema executivo de um programa de computador um modelo real plausível para o sistema representacional hierárquico imaginado por Fodor? Como diz P.C. Dodwell:

“Tal hierarquia não leva a um regresso infinito porque, como Neisser [. . .] mostra, o sistema executivo é o controlador geral do que ocorre, e não é, em si mesmo, ‘usado’ por um nível superior — ele incorpora a ‘meta’ do programa todo” (Dodwell 1971, p. 371).

Na avaliação desse argumento não se deve esquecer a *banalidade* de que computadores não são dispositivos que “aparecem” na natureza e que programas de computador não surgem por acaso. O homem, como engenheiro e programador, os constrói e usa para calcular. A atividade maquinal é tida como cálculo porque se desenvolve segundo princípios que o programador incorporou no programa. Isso pressupõe que o programador *compreenda* antecipadamente o problema, cuja solução requer a efetuação de um cálculo, e, também, *saiba* através de quais princípios operacionais a solução deve ser alcançada. É verdade que o sistema executivo determina

sem a intervenção humana a seqüência de operações, entre elas as chamadas de subprogramas que executam determinadas operações auxiliares: as expressões antropomórficas "chamar", "comparar", "decidir", usadas na descrição da lógica interna do programa, não aludem, evidentemente, a "homunculi" na máquina, mas caracterizam simplesmente os mecanismos de "chaveamento" que, *sob uma certa interpretação determinada pelo programador e não pela máquina mesma*, efetuam as operações de um cálculo. O sistema executivo — e esse é o ponto decisivo — "incorpora" a meta do programa somente porque foi projetado assim por um programador.

Nesse sentido a "racionalidade" global do programa não é assunto do computador. Ela é *nosso* assunto. Para *nós*, que produzimos, descrevemos e usamos o computador, determinados eventos *valem* como representações do mundo, por exemplo, da distribuição de idades numa população, do saldo de uma conta bancária etc. Outras representações são construídas através de regras iterativas que constituem os subprogramas. Todavia, nem o computador nem o sistema executivo determinam o ponto de vista e o vocabulário que fixam o sentido das representações e o conteúdo das regras. Disso, nós, usuários do computador, temos de cuidar. Exatamente isso justifica a *inutilidade* do programa de computador como um modelo para um sistema representacional hierárquico capaz de dispensar, como supérfluo, o usuário. O (programa de) computador pode fazer seja o que for, mas não *isso*.

7. A argumentação acima não diz coisa alguma contra a possibilidade de usar o modelo do computador para a explicação de *determinados* aspectos do pensar ou da "solução de problemas" nos homens. Alguns *processos concomitantes do pensar*, os quais *acompanham*, portanto, o pensar, podem ser descritos através do uso desse modelo. Com esse objetivo pode-se postular, na investigação do comportamento de certas partes do sistema nervoso *envolvidas* no exercício das "faculdades da inteligência", "variáveis intervenientes" na forma de subprogramas. Todavia, na avaliação crítica das teorias cognitivistas é sempre importante revelar os erros que surgem da circunstância de que os cognitivistas, além da postulação de subprogramas com objetivo de construir um modelo dos processos concomitantes do pensar, procuram descrever o próprio pensamento como um programa de computador.

Aqui os adeptos do modelo do computador ultrapassam os limites do sentido e justificam o alerta crítico *uti non abuti*. Isso pode ser mostrado com base em uma passagem de Dodwell:

"A compreensão de como as sub-rotinas trabalham não explica o princípio de solução de problemas em termos de uma seqüência de passos que dependem das computações das sub-rotinas, do mesmo modo que a compreensão das tabelas de multiplicação não explica a solução da maior parte dos problemas aritméticos. Para isso, deve-se olhar o processo executivo, que incorpora, na máquina, a organização global e a meta do

programa e que é, no ser humano, uma 'orientação por metas' [goal-directedness] menos claramente compreendida" (Dodwell 1971, p. 372).

O problema a que essa passagem remete é o seguinte. Os processos (fisiológicos) acompanhantes do pensar, compreender, perceber hipoteticamente descritos no quadro da Psicologia Cognitiva como seqüências de processos computacionais. Estes computam os resultados ("outputs") de processos computacionais anteriores e tornam os resultados de sua computação disponíveis para computação posterior. Podemos reconhecer aqui a estrutura ordenada dos subprogramas num programa principal. Para a compreensão adequada do programa global num computador *real*, diz Dodwell, não basta a compreensão do modo de funcionamento das sub-rotinas, mas precisamos, além disso, dirigir a atenção ao sistema executivo. E isso basta. No caso do homem também precisamos olhar além dos "sub-programas". Todavia, a "intencionalidade", a "orientação por metas" nos homens é *somente mais complexa*, e por isso "menos claramente compreendida" do que os sistemas executivos de computadores *reais*. Trata-se aqui, no caso dessa *assimilação*, de uma descrição correta?

As operações num computador são passos para a solução de um problema porque estão conectadas entre si de maneira *esquemática*, segundo regras *rígidas*, "estúpidas", que são determinadas pelo tipo de problema a ser resolvido e implementadas no sistema executivo pelo programador. Os processos, por outro lado, investigados e descritos pelo psicólogo cognitivo como processos computacionais são processos do pensar (compreender, perceber) porque ocorrem "simultaneamente" com o pensar (compreender, perceber). Mas nem o pensar (compreender, perceber), nem os critérios segundo os quais julgamos se alguém pensa (compreende, percebe), seguem um sistema rígido de regras que poderia ser aplicado esquematicamente. Ou seja: os processos que podemos descrever parcialmente numa linguagem computacional podem ser identificados como processos do pensar porque ocorrem tipicamente *durante o pensar*. E para decidir se o homem, em cujo organismo tais processos ocorrem, *pensa*, precisamos considerar *suas ações* e o *meio ambiente* (*sobretudo o meio ambiente socio-cultural*) no qual *elas se desenvolvem*. Na maior parte dos casos — e disso os experimentos da Psicologia Cognitiva dão-nos exemplos suficientes — precisamos somente *perguntar-lhe*. Tanto na conversação com os homens, quanto na observação das "finas nuances do comportamento" (Wittgenstein), precisamos evidentemente considerar coisas que são totalmente *dessemelhantes* a um sistema executivo. Regras rígidas que nos digam o que um homem precisa fazer para que digamos dele que ele pensa *não* existem.

Sem que precisemos entrar aqui em mais detalhes, deve ter-se tornado claro que as circunstâncias e os critérios sob os quais usamos o conceito de pensar (compreender, perceber) *não* são *descritíveis* em termos de um sistema de regras; e muito menos podem ser compreendidos como a *observância* real de tais regras. Na perspectiva do "seguir uma regra", a intencio-

nalidade (a "orientação por metas") de um homem é *não compreensível*, e não somente "ainda" não compreendida ou "menos claramente compreendida", como Dodwell sugere. A opinião segundo a qual *tem de* haver um conjunto de regras por trás da aparência desordenada dos fenômenos familiares é um *dogma*. Esse "tem de" é um "tem de" lógico que repousa numa forma de expressão inadequada aos fatos.

8. Consideremos, para ilustrar o que foi dito na seção anterior, um exemplo "clássico" da Psicologia Cognitiva, no qual diferentes sinais acústicos são apresentados "dicoticamente" através de fones de ouvido, isto é, um sinal diferente em cada fone, a um sujeito de teste. Se o exortarmos a dirigir sua atenção a *um* dos canais e lhe perguntarmos então — para sua surpresa — sobre o que foi apresentado no *outro* canal, o sujeito de teste não poderá, tipicamente, relatar mais do que as características determinadas, por assim dizer, *diretamente* pelas propriedades *acústicas* dessa "entrada". Isso inclui, por exemplo, a informação de que o sinal era uma fala (ou um canto), que a voz era masculina (ou feminina) etc. Disso *não* faz parte, normalmente, o *conteúdo* do que foi dito.

Naturalmente, a correção desse experimento, do qual, aliás, há variações mais sutis, não está em questão. Os problemas filosoficamente relevantes emergem, antes, nos comentários que normalmente acompanham sua apresentação e interpretação. Assim, os psicólogos cognitivistas dizem que a informação inicial (a representação da entrada) é processada serialmente em diferentes níveis de abstração até que seja completamente compreendida. A atenção funcionaria como um filtro através do qual a informação precisaria passar para que possa ser reconhecida. Para uma das entradas seriam construídas representações em todos os níveis do processo total de computação: o sujeito de teste compreende e pode reproduzir (em casos típicos) o que foi dito nesse canal. A outra entrada, ao contrário, receberia representações somente nos primeiros níveis ("acústicos"); e tais representações seriam então "contidas" pelo "filtro da atenção", não sendo, portanto, posteriormente processadas. Num caso, a "análise" do que é ouvido estender-se-ia até a camada *semântica* ("profunda") do que foi dito. No outro caso, a "análise" seria interrompida na camada ("superficial") das propriedades *acústicas*. Todavia — e esse é o ponto central —, *quais* são os mecanismos computacionais usados, e *em que medida* eles são postos em ação, *depende* da circunstância de que o *sujeito* de teste *entende* e *segue* as indicações do experimentador, de que ele *dirige*, portanto, sua atenção a um dos canais e *não se importa* com o que ocorre no outro.

Para a atenção mesma não conhecemos um critério esquemático. O experimento descrito mostra, antes, que a atenção necessária tem de ser atribuída e confiada *antes de tudo* (antes de qualquer observação experimental) ao sujeito de teste, e na verdade no sentido mais habitual, não descritível esquematicamente. Isso quer dizer: o sujeito de teste precisa ter *compreendido* o contexto de suas ações e *observado corretamente* as

instruções do experimentador. Parte do que ocorre então pode ser descrita em termos computacionais como construção seletiva de representações de abstração crescente. Mas a *precondição lógica* da existência desse “sistema computacional” é a ação (inteligente) de um sujeito de teste com cuja ajuda e atenção o experimentador precisa contar. Isso parece ser reconhecido pelo próprio Fodor:

“*Seres humanos* têm aparentemente acesso a uma variedade de modos de representação e podem exercer um controle racional sobre os tipos de representações que empregam. Isto é: o modo como os recursos representacionais disponíveis são explorados num caso dado depende do que o agente considera como as exigências da tarefa em questão” (Fodor 1975, p. 194; grifado por mim).

Contudo, o que é atribuído aqui a um agente e que só pode ser descrito, *também por Fodor*, numa linguagem não “objetivista” que se refere a pessoas e seu contexto de ação é apresentado por Fodor, noutra passagem, *de maneira injustificada*, como o efeito de um mecanismo: “São mecanismos atencionais que determinam o modo como as capacidades representacionais disponíveis são exploradas” (Fodor 1975, p. 163). A não-compreensão de que o uso das expressões “mecanismo” e “efeito” é aqui, na melhor das hipóteses, *metafórico* – semelhante ao discurso psicanalista sobre “mecanismo de defesa”, o qual não é mecanismo algum, pois é descrito em termos das ações significativas de uma pessoa num meio social complexo “carregado” de significado (cf. Hamlyn 1953) – conduz Fodor a identificar a intencionalidade de um agente humano com o sistema executivo, *ainda a ser descoberto*, de um computador: “O desenvolvimento eficiente de capacidades computacionais é, ele mesmo, um problema computacional e um problema que seres humanos estão bem equipados para resolver” (Fodor 1975, p. 194).

Nossa vida inteligente é assim vista, em sua totalidade, como um problema computacional. Somente a Psicologia Cognitiva, assistida pela Inteligência Artificial, poder-nos-ia dizer, segundo essa opinião, “que espécie de solução são as pessoas” (Fodor 1975, p. 194), quando tiver decomposto a inteligência humana em processos computacionais elementares do “tipo estúpido”. Então disporemos do “programa da mente”. Se os conceitos de mecanismo e programa, de que se faz uso aqui, não devem ser compreendidos metaforicamente, é preciso assumir, além disso, que tal “programa da mente” poderia ser implementado à vontade. Todavia, porque toda essa especulação tem sua base antes na exploração “realista” de uma metáfora e não em resultados concretos, é preciso que seus proponentes recorram aqui a uma mera promessa:

“Vimos que um modelo da compreensão da sentença é, com efeito, um mecanismo que associa realizações de formas de onda [token wave forms] a mensagens [com significado]. Muito pouco é conhecido sobre o modo

como um tal dispositivo poderia operar, embora eu possa supor que, se começássemos agora e trabalhássemos duramente, estaríamos em condições de construir um em cerca de quinhentos anos" (Fodor 1975, p. 167).

9. Nessa "cruzada otimista" Fodor não é um solitário. O tratamento acrítico das possibilidades de utilização do modelo do computador na Psicologia dá origem, em inúmeros autores, a tais "fantasias da razão em uso especulativo". Que a inteligência humana seja organizada como o sistema executivo de um programa de computador é, porém, uma simples "suposição", uma *crença* insensata que não pode ser justificada através da consideração exata da maneira como computadores são realmente utilizados, como programas são verdadeiramente projetados. O contrário é o caso: o conhecimento do modo de funcionamento do computador mostra que se trata aqui de uma *superstição* psicológica. Pois o cálculo executado pelo computador desenvolve-se de acordo com regras rígidas especificadas pelo programador para a solução de um problema "técnico". Computadores *não aprendem*. Os programas que se modificam a si mesmos fazem isso segundo princípios especificados antecipadamente pelo programador, de tal forma que "atrás" de cada processo que descrevemos como computacional há um agente humano inteligente.

Ao contrário do programa de computador, o agir inteligente não está preso a um *esquema*. É exatamente *isso* que o faz inteligente, "não-mecânico", criativo. Do ponto de vista da lógica do conhecimento, o conceito de esquema (da exposição esquemática dos fenômenos) pressupõe o conceito do agir inteligente. *Agindo* (inteligentemente) descrevemos determinados fenômenos de forma esquemática, construímos aparelhos que operam segundo esquemas projetados por nós. A inteligência que podemos encontrar em programas de computador testemunha somente a inteligência de seus projetistas.

A tentativa de levar avante a idéia de uma análise de nossa vida inteligente em termos de processos quase-mecânicos conduz, portanto, à circularidade e à regressão infinita. O agente humano que se procura eliminar com a postulação de processos computacionais aparece *por razões lógicas*, novamente, "atrás" de cada novo processo. Fodor procura evitar esses "homunculi" ao identificar os processos computacionais descritos em sua teoria com processos *neurológicos* cujo modo de funcionamento é caracterizado abstratamente com expressões antropomórficas. Tais processos não precisam ser "interpretados" por um sistema "interno" adicional para "dirigir" o com-

um lado, e os processos *orgânicos* estudados por Fodor, por outro, é, contrariamente às intenções de Fodor, *consolidada*. Pois tais processos podem ser constatados somente sob a condição de que sejam referidos aos *critérios não esquemáticos do compreender* e com isso ao agir inteligente dos sujeitos de teste — exatamente da mesma forma com o “mecanismo de atenção”, enquanto *veículo fisiológico da capacidade de dirigir intencionalmente a atenção*, precisa ser “ativado” por uma *pessoa “atenta”*, sem o que ele não seria um mecanismo *de atenção*. Nesse ponto Fodor move-se em círculo, como ele mesmo parece reconhecer:

“Esforçamo-nos para fazer experimentos que coloquem em ação os mecanismos brutos, involuntários da cognição; reflexos intelectuais, por assim dizer, com os quais a mente responde quer queira, quer não [willy-nilly] à tarefa. Mas o que encontramos, ao invés disso, são simplesmente as estratégias locais, de propósito especial, que os sujeitos divisam para atender eficientemente às suas instruções. O que o experimento revela principalmente é, portanto, a capacidade do sujeito de imaginar as metas do experimentador e sua disposição geral de fazer o melhor possível para favorecê-las. [. . .] Disseram-nos que a ciência procura explicar as uniformidades que estão subjacentes à desordem superficial dos fenômenos. Assim, é desanimador mergulhar sob os recursos cognitivos complexos e variáveis que os sujeitos humanos trazem para a tarefa de solução de problemas somente para descobrir constantemente novas camadas de recursos complexos e variáveis. (Fodor 1975, p. 194).

10. Do ponto de vista do Cognitivismo, esse resultado é desconcertante. Fodor procurava uma explicação quase-mecânica do comportamento inteligente. Para isso ele postula, por exemplo, um processo do compreender no qual o sinal percebido recebe uma série crescente de representações cada vez mais abstratas — desde suas propriedades meramente acústicas até sua descrição estrutural “profunda”, na qual o sinal é compreendido como “proposição”. Experimentos como o da audição dicótica” descrito acima mostram talvez que alguns eventos no sistema nervoso central funcionam como *condições nomológicas* do exercício da faculdade de compreensão (de uma linguagem). Desde que seja possível, *com base nos critérios habituais*, distinguir graus do compreender, faz sentido considerar as condições neurológicas correlacionadas com esses graus como representações dos diferentes níveis nos quais uma sentença pode ser analisada. Representações implementadas neurologicamente seriam assim *condições necessárias* do compreender, isto é, condições sem as quais não se pode exercer tal faculdade — da mesma forma que uma retina intacta é uma condição necessária da visão. E a especificação do sistema dessas representações poderia ser considerada como a “descrição modelar dos processos que estão subjacentes ao compreender”. O compreender mesmo, enquanto atitude proposicional, não é, porém, *idêntico* aos processos neurológicos que o acompanham, nem *reduzível* a esses. Dito de outro modo: a representação implementada neurologicamente não é o critério com base no qual o “estado” do compreender é consta-

tado. E a razão disso é que o conceito do compreender não se situa no plano dos eventos orgânicos causalmente conectados, mas sim no das ações humanas. Afinal, não é qualquer particularidade no organismo que pode ser considerada como uma representação de "funções" relacionadas com o compreender, mas somente aqueles eventos revelados, em casos típicos, como condições empíricas do "estado" do compreender *que se mostra no agir do sujeito de teste*. Nesse sentido a representação *não* é associada com o que ela representa *através de um cálculo ou uma computação*. Os experimentos da Psicologia Cognitiva, eis o que Fodor constata na citação acima, não permitem contornar esse fato incômodo para as pretensões ambiciosas do Cognitivismo. Ao contrário: tais experimentos obrigam-nos a fazer as distinções que o programa de Fodor queria eliminar, isto é, as distinções entre ações e eventos, entre atribuições pessoais de atitudes proposicionais e constatações de processos orgânicos, entre processamento de representações e sua projeção no mundo através do agir de uma pessoa.

Ao invés de reconhecer nesse ponto que, com a tentativa de reduzir os chamados estados psicológicos a eventos orgânicos "cegos", se persegue somente uma quimera, Fodor, num movimento típico da Psicologia Cognitiva, adia indefinidamente o problema: "A longo prazo — num prazo *muito* longo — queremos chegar a esses reflexos brutos da cognição" (Fodor 1975, p. 195). Contudo, com vagas declarações de intenção Fodor dificilmente conseguirá satisfazer um leitor crítico. De resto, a necessidade constante de suplementar observações concretas com promessas futuristas cujo prazo de resgate é não inferior a quinhentos anos ilustra menos o progresso do que a inadequação conceitual das teorias computacionais da mente.

BIBLIOGRAFIA

- Block, N. (1980). "Introduction: what is Functionalism?" In N. Block (ed.) *Readings in Philosophy of Psychology*, vol. I. London: Methuen.
- Dennett, D.C. (1969). *Content and Consciousness*. London: Routledge and Kegan Paul.
- . (1977). "Critical Notice: *The Language of Thought* by Jerry Fodor". *Mind*, 86, 265-80. [Reimpresso sob o título "A Cure for the Common Code?" in N. Block (ed.) *Readings in Philosophy of Psychology*, vol. 2. London: Methuen, 1981. (As citações foram feitas de acordo com essa reimpressão.)]
- Fodor, J.A. (1975). *The Language of Thought*. New York: Thomas Y. Crowell.
- . (1978a). "Propositional Attitudes". *The Monist*, 61, 501-23. [Citações feitas de acordo com a reimpressão em N. Block (ed.) *Readings in Philosophy of Psychology*, vol. 2. London: Methuen, 1981.]
- . (1978b). "Computation and Reduction". In C.W. Savage (ed.) *Perception and Cognition. Minnesota Studies in the Philosophy of*

- Science*, vol. 9. Minneapolis: University of Minnesota Press. [Citações feitas de acordo com a reimpressão em J.A. Fodor, *Representations*. Sussex: The Harvester Press, 1981.]
- . (1981). "Introduction: Something on the State of the Art". In J. A. Fodor, *Representations*. Sussex: The Harvester Press.
- von Wright, G.H. (1971). *Explanation and Understanding*. Ithaca: Cornell University Press.
- . (1976). "Determinism and the Study of Man". In J. Manninen e R. Tuomela (eds.) *Essays on Explanation and Understanding. Studies in the Formations of Humanities and Social Sciences*. Dordrecht: D. Reidel Publishing Company.
- Wittgenstein, L. (1967). *Zettel* (ed. G. E. M. Anscombe e G. H. von Wright). Oxford: Basil Blackwell.
- . (1980). *Bemerkungen über die Philosophie der Psychologie*. (ed. G. E. M. Anscombe e G. H. von Wright). Oxford: Basil Blackwell.