

# O DEBATE SOBRE O REPRESENTACIONALISMO NAS CIÊNCIAS COGNITIVAS

## *THE DEBATE ABOUT MENTAL REPRESENTATION IN COGNITIVE SCIENCE*

*Luis Felipe Oliveira*<sup>1</sup>

**Resumo:** Este artigo investiga o papel que o conceito de representação mental teve no desenvolvimento das ciências cognitivas. Entendemos que estas ciências apresentam três paradigmas centrais que se distinguem, principalmente, pela consideração que fazem sobre a noção de representação mental. O paradigma cognitivista fundamenta-se sobre o entendimento de que a mente opera sobre um conjunto de representações simbólicas armazenadas e manipuladas em processos funcionalmente computacionais. O paradigma conexionista descreve a mente como um sistema de processamento paralelo e distribuído que realiza operações sobre representações subsimbólicas. O paradigma dinâmico da cognição estuda a mente por outras abordagens, inclusive não computacionais, que questionam a necessidade explicativa da noção de representação mental. Um dos modelos mais discutidos na literatura, com relação ao debate entre as visões representacionistas e as não-representacionistas, é o *Watt's Governor*, um mecanismo de autorregulação para caldeiras à vapor. Por um lado, argumenta-se que o *Watt's Governor* é um sistema não-representacional de autocontrole; por outro lado, postula-se que o aparato é um sistema representacional. Haselager et al. (2003) discutem esse embate e concluem que a ubiquidade da representação enfraquece este conceito e expõe a falácia de se considerá-la como condição necessária na explicação de fenômenos mentais. Nesse sentido, discutimos as conclusões desses autores e abordamos a hipótese de que a semiótica pode oferecer contribuições interessantes às ciências cognitivas, em suas considerações sobre os tipos e formas das representações.

**Palavras-chave:** Ciências cognitivas. Representação mental. Cognitivismo. Conexionismo. Dinamicismo.

**Abstract:** This paper investigates the role that the concept of mental representation has had in cognitive science development. We understand that cognitive science has three main paradigms which are distinguished mainly by considerations about the notion of mental representation. The (classical) cognitivist paradigm bears the understanding that mind operates upon a set of symbolic representations stored and manipulated by functionally computational processes. The connectionist paradigm describes mind as a system of parallel and distributed processing which instantiates operations upon sub-symbolic representations. The dynamical paradigm of cognition studies the mind by means of other approaches, including non-representational ones, that put under scrutiny the notion of mental representation as an explanatory necessity. One of the most known and discussed models in literature, considering the debate between representational and non-representational points of view, is *Watt's Governor*, a self-regulating mechanism used in steam machines. On the one hand, it can be argued that a *Watt's Governor* is a non-representational self-controlling system; on the other hand, it can be postulated that such device is a representational system indeed. Haselager et al. (2003) discuss that argument and conclude that the ubiquity of representation undermines this very concept and exposes the fallacy of considering it as a necessary condition in

---

<sup>1</sup> Docente do Curso de Música, Centro de Ciências Humanas e Sociais, Universidade Federal de Mato Grosso do Sul – UFMS. Mestre em Filosofia pelo Programa de Pós-Graduação em Filosofia da UNESP/Marília em 2003. E-mail: luis.felipe@ufms.br

explaining mental phenomena. Thus, we bring that conclusion into question and suggest the hypothesis that semiotics could offer interesting contributions to cognitive science, in its considerations about the sorts and forms of representations.

**Keywords:** Cognitive science. Mental representation. (Classical) cognitivism. Connectionism. Dynamicism.

## **1. Introdução**

Estabelecer um marco preciso para o surgimento de uma nova área científica é, geralmente, uma tarefa controversa e assim também ocorre com as ciências cognitivas. Gardner (1995) atribui esse início ao Simpósio Hixon sobre Mecanismos Cerebrais e Comportamento de 1948; já Dupuy (1996) atribui o surgimento desta área às Conferências Macy, que ocorreram de 1946 a 1953, em um total de dez conferências. Certamente, os maiores nomes do início das ciências cognitivas estavam presentes em ambos os eventos realizados nos Estados Unidos, como John von Neumann, Nobert Wiener, Hebert Simon, Alan Newell, Warren McCulloch, Walter Pitts, Karl Lashley, entre inúmeros outros, o que dificulta uma demarcação mais precisa sobre o surgimento dessa ciência. Podemos dizer que, por volta do final da década de 1940, começaram a surgir eventos e encontros entre uma comunidade científica interdisciplinar que fundaram o que ficou conhecido como cibernética, área que estava interessada em estudar mecanismos responsáveis por comportamentos autocontrolados ou autorregulados, estudos os quais constituem a base para os estudos cognitivos posteriores. Wiener (1961, p. 11) definiram a cibernética como “[...] o campo completo da teoria do controle e da comunicação, seja nas máquinas ou nos animais.”<sup>2</sup> Para tanto, esse grupo interdisciplinar envolvia pesquisadores das áreas da matemática, da lógica, da engenharia, da fisiologia e da neurofisiologia, da psicologia e da antropologia (DUPUY, 1996, p. 9). Em oposição aos pressupostos behavioristas do estudo do comportamento humano, dominantes até então, as emergentes ciências cognitivas buscavam estudar a mente de uma forma empírica mas que, ao mesmo tempo, considerasse estados mentais em sua abordagem, através de uma equivalência formal entre estados lógicos – que podem ser formalizados nas operações de uma máquina computacional – e estados mentais.

---

<sup>2</sup> [...] the entire field of control and communication theory, whether in the machine or in the animal.

Estados mentais são conceitos do senso-comum relacionados a crenças, desejos, percepções, ideias, imaginações, sensações, pensamentos, pelos quais se explica o comportamento inteligente. Estados mentais, nesse entendimento de senso comum, correlacionam-se ao conceito de representação mental das ciências cognitivas, que tem tanto poder causal quanto conteúdos. Tomemos o exemplo de Pitt (2013, grifos do autor): “Crenças e desejos causam os comportamentos que causam porque têm os conteúdos que têm. Por exemplo, o desejo de *tomar-se uma cerveja* e as crenças de que *há cerveja na geladeira* e que *a geladeira está na cozinha* podem explicar o comportamento de levantar-se e ir até a cozinha.” Considerações, portanto, sobre representações mentais, seus poderes causais e seus conteúdos, vinculam-se a inúmeras questões filosóficas, como a questão da intencionalidade do comportamento, as propriedades semânticas das representações mentais, as correlações com atitudes proposicionais, considerações sobre a natureza e as formas das representações mentais, debates sobre internalismo e externalismo entre outras considerações que a Filosofia da Mente traz à cena. No âmbito das ciências cognitivas, nas palavras de Pitt (2013):

Filósofos da mente contemporâneos tipicamente supunham (ou esperavam) que a mente pudesse ser naturalizada — *i.e.*, que todos os fatos mentais têm explicações em termos das ciências naturais. Esse postulado é compartilhado dentro das ciências cognitivas, as quais tentam oferecer considerações dos estados e processos mentais em termos de características do cérebro e do sistema nervoso central. Nesse processo, as várias subdisciplinas das ciências cognitivas (incluindo as psicologias cognitiva e computacional e as neurociências cognitiva e computacional) postulam um número de diferentes tipos de estruturas e processos, muitos dos quais não são diretamente implicados por estados e processos mentais concebidos pelo senso comum. Contudo, permanece um compromisso difundido com a ideia de que estados e processos mentais serão explicados em termos de representações mentais.<sup>3</sup>

Neste trabalho, então, assumimos o argumento de que as ciências cognitivas

---

<sup>3</sup> Contemporary philosophers of mind have typically supposed (or at least hoped) that the mind can be naturalized — *i.e.*, that all mental facts have explanations in the terms of natural science. This assumption is shared within cognitive science, which attempts to provide accounts of mental states and processes in terms (ultimately) of features of the brain and central nervous system. In the course of doing so, the various sub-disciplines of cognitive science (including cognitive and computational psychology and cognitive and computational neuroscience) postulate a number of different kinds of structures and processes, many of which are not directly implicated by mental states and processes as commonsensically conceived. There remains, however, a shared commitment to the idea that mental states and processes are to be explained in terms of mental representations.

tradicionalmente explicam estados mentais em termos de representações mentais. Contudo, além da correspondência entre estados mentais e estados representacionais, fazem-se também necessárias considerações sobre a correspondência entre estados mentais/representacionais/lógicos e estados físicos ou implementacionais.

Se assumirmos que uma das concepções fundacionais da ciência cognitiva é que a mente/cérebro é um sistema computacional, o sustentáculo das representações mentais será estruturas ou estados computacionais. A natureza específica dessas estruturas ou estados depende do tipo de computador que a mente/cérebro hipoteticamente é. Até o momento, a ciência cognitiva focou-se em dois tipos: computadores convencionais (von Neumann, simbólico, ou baseado em regras) e computadores conexionistas (processamento paralelo e distribuído). Se a mente/cérebro é um computador convencional, então os sustentáculos das representações mentais serão estruturas de dados [...]. Se a mente/cérebro é um computador conexionista, então os sustentáculos da representação de estados mentais ocorrentes serão estados de ativação de nódulos ou conjunto de nódulos conexionistas.<sup>4</sup> (ECKARDT, 1999, p. 527-528).

As palavras sintéticas de Eckardt nos remetem ao debate filosófico ou conceitual que percorre a história das ciências cognitivas, sem sugerir ainda que mais recentemente surgem abordagens não-representacionistas que questionam tanto os modelos computacionais apoiados sobre a noção de representação simbólica quanto os modelos que operam sobre representações subsimbólicas. Nesse sentido, entendemos que a história das ciências cognitivas pode ser segmentada em três distintas abordagens, as quais chamaremos de *paradigma cognitivista*, *paradigma conexionista* e *paradigma dinâmico da cognição*. O paradigma cognitivista vincula-se aos modelos computacionais apoiados sobre a noção de representação simbólica; o paradigma conexionista, aos modelos de processamento paralelo e distribuído, com representações subsimbólicas; o paradigma dinâmico da cognição elimina (ou ao menos questiona) a necessidade explicativa da representação mental. Historicamente, esta divisão paradigmática em três abordagens não é rigorosamente

---

<sup>4</sup> If we take one of the foundational assumptions of cognitive science to be that the mind/brain is a computational device the mental representation bearer will be computational structures or states. The specific nature of these structures or states depends on what kind of computer the mind/brain is hypothesized to be. To date, cognitive science has focused on two kinds: conventional (von Neumann, symbolic, or rule-based) computers and connectionist (parallel distributed processing) computers. If the mind/brain is a conventional computer, then the mental representation bearers will be data structure [...]. If mind/brain is a connectionist computer, then the representation bearers of occurrent mental states will be activation states of a connectionist nodes or sets of nodes.

cronológica, ainda que apresente uma certa linearidade temporal, e o surgimento dos paradigmas mais recentes não substituiu ou eliminou os mais antigos. Isto é, muitas das teorias postuladas nas ciências cognitivas assumem, implícita ou explicitamente, a visão computacional da mente. Em áreas ou subáreas de estudos cognitivos que estão mais distantes desse debate filosófico sobre representação mental, muitas vezes se assume o pressuposto da visão computacional *tout court*.<sup>5</sup> Não iremos, neste trabalho, apresentar uma cronologia do debate sobre a representação mental nas ciências cognitivas, nem das teorias envolvidas e postuladas nesse debate, nem iremos considerar historicamente todas as abordagens que vinculam-se a esta ou aquela visão; iremos, diferentemente, apresentar sinteticamente a discussão da representação mental nas ciências cognitivas, examinando os três paradigmas que mencionamos, e, posteriormente, discutir um caso em particular no que se refere às considerações sobre a existência ou não de representações ou de sua necessidade explicativa, em termos cognitivos. Por fim, indicaremos que o debate sobre as representações mentais nas ciências cognitivas pode ser bastante elucidado por considerações advindas da semiótica peirceana.

## **2. A representação mental simbólica**

A célebre frase sem verbos de Jerry Fodor (1975, p. 31) é emblemática com respeito a como ciência cognitiva tradicionalmente vê a mente humana: “sem representações, sem computação. Sem computação, sem modelo.”<sup>6</sup> Trata-se do mote cognitivista, da máxima que inúmeros estudos sobre a mente humana carregam, como dissemos, implícita ou explicitamente. Segundo Pitt (2013, aspas do autor), “a noção de uma “representação mental” é, argumentativamente, em primeiro lugar, um constructo teórico da ciência cognitiva.” E o autor, corroborando a máxima de Fodor, continua, dizendo que a representação mental “é o conceito básico da Teoria Computacional da Mente, de acordo com a qual estados e processos cognitivos são constituídos pela ocorrência, transformação e armazenamento (na mente/cérebro) de estruturas informacionais (representações) de um

---

<sup>5</sup> Confira, neste sentido, a crítica de Oliveira (2014) sobre a ausência de debate sobre os modelos mentais tradicionalmente empregados nos estudos em cognição musical.

<sup>6</sup> No representations, no computations. No computations, no model.

tipo ou de outro.”<sup>7</sup> Como nos coloca Haugeland (1991) uma representação é algo que está no lugar de outra coisa e, portanto, a representa; aquilo que ela representa é, então, seu conteúdo. Esse conceito geral de representação, pode ser adaptado ao conceito mais específico da representação mental: uma representação mental é algo que está na mente/cérebro de um organismo e que representa outra coisa — como objetos do mundo externo, outros seres, outros pensamentos, ideias abstratas etc — e, portanto, a representa; aquilo que ela representa é, então, seu conteúdo.

Dentro do paradigma cognitivista, a representação mental é vista como um símbolo, um objeto que não tem nem relação causal nem formal com aquilo que representa, assim como as palavras da linguagem natural não possuem relações de causalidade ou de semelhança formal com os objetos do mundo aos quais se referem, ou como um V<sup>7</sup> representa um acorde de sétima de dominante. Neste sentido, as representações mentais simbólicas são internas, abstratas, arbitrárias, discretas e infinitas, cada uma delas correspondendo a um único objeto ou propriedade de um objeto. Ao ouvir música, se alguém “é movido por ela, então deve ter passado por um estágio cognitivo que envolve a formação de uma representação interna, simbólica ou abstrata, da música” (SLOBODA, 2008, p. 5). Percebe-se, dessa forma, pela afirmação acima, que dentro desta visão cognitivista, que ouvir música — e ser movido por ela, e entendê-la — é formar representações mentais dos fenômenos sensoriais.<sup>8</sup> Quando alguém fala sobre *o timbre aveludado do violoncelo*, sua mente está operando sobre representações simbólicas dessa qualidade sensorial; ainda mais, para entender que esse timbre sonoro é aveludado, é necessário relacionar as representações mentais do violoncelo, da sensação sonora do timbre do violoncelo, do veludo, da sensação tátil do veludo, e somente então se pode entender a frase *o timbre aveludado do violoncelo*. A compreensão dessa frase é uma operação formal, *i.e.* computacional, sobre um conjunto de representações mentais e seus conteúdos. Existe, nesta assunção, uma equivalência entre relações sintáticas (formais, computacionais) e relações semânticas, que nos permite entender o significado da

---

<sup>7</sup> The notion of a “mental representation” is, arguably, in the first instance a theoretical construct of cognitive science. As such, it is a basic concept of the Computational Theory of Mind, according to which cognitive states and processes are constituted by the occurrence, transformation and storage (in the mind/brain) of information-bearing structures (representations) of one kind or another.

<sup>8</sup> Sloboda (*idem*, *ibidem*) diz que “A natureza desta representação interna, e as coisas que ela permite que o ouvinte faça com a música é a matéria-prima central da psicologia cognitiva da música.”

expressão o timbre aveludado do violoncelo.

Se, em princípio, relações sintáticas podem ser paralelas a relações semânticas, e se, em princípio, pode-se ter um mecanismo cujas operações sobre fórmulas sejam sensíveis às suas sintaxes, então pode ser possível construir uma máquina sintaticamente dirigida, a qual seria o que é necessário para se ter um modelo mecânico da coerência semântica do pensamento; correspondentemente, a ideia que o cérebro é tal máquina é a hipótese fundacional da ciência cognitiva clássica.<sup>9</sup> (FODOR; PYLYSHYN, 1988, p. 30)

No paradigma cognitivista entende-se, portanto, que (ao menos parcialmente) o pensamento compartilha propriedades com a linguagem natural; o pensamento estrutura-se como a linguagem, gerando complexos representacionais potencialmente infinitos a partir de processos combinatoriais sobre um conjunto de finito de estados representacionais primitivos: “a semântica, tanto da linguagem quanto do pensamento é *composicional*: o conteúdo de representações complexas [como o timbre aveludado do violoncelo] é determinado pelos conteúdos de seus constituintes e suas configurações estruturais” (PITT, 2013, Seção 8, ênfase do autor).<sup>10</sup> O paralelismo entre as relações sintáticas e as relações semânticas é o que garante que “a linguagem do pensamento” possa ser investigada empiricamente, através de máquinas computacionais diferentes daquela que é o cérebro humano. Toca-se, neste ponto, em uma problemática bastante discutida na literatura filosoficamente crítica do paradigma cognitivista, a saber, do *isomorfismo*.

Como nos coloca Haselager, Groot e Rappard (2003, p. 8), “um sistema físico é considerado como realizador de um modelo computacional particular quando existe um mapeamento um-para-um entre os estados computacionais e suas estruturas formais, por um lado, e os estados físicos e suas estruturas físico-causais, por outro lado.”<sup>11</sup> Ainda pelo argumento do isomorfismo, assume-se que um sistema físico-computacional seja

---

<sup>9</sup> If, in principle, syntactic relations can be made to parallel semantic relations, and if, in principle, you can have a mechanism whose operations on formulas are sensitive to their syntax, then it may be possible to construct a syntactically driven machine would be just what’s required for a mechanical model of the semantical coherence of thought; correspondingly, the idea that the brain is such a machine is the foundational hypothesis of classical cognitive science.

<sup>10</sup> [...] the semantics of both language and thought is compositional: the content of a complex representation is determined by the contents of its constituents and their structural configurations.

<sup>11</sup> A physical system if thought to realize a particular computational model when there is a one-to-one mapping between the computational states and their formal structure on the one hand and the physical states and their physical-causal structure on the other hand.





revolução conceptual que assinala o advento das ciências cognitivas, e não sua origem,” mesmo porque os primeiros modelos computacionais como aqueles desenvolvidos por McCulloch e Pitts (1943) ou mesmo a máquina de Turing (2004) ainda não tinham essa distinção funcional.

A visão funcionalista da mente associa-se, normalmente, a um entendimento internalista, “pelo qual se pensa que os conteúdos representacionais [e seus poderes causais] são determinados apenas pelas propriedades intrínsecas do indivíduo” (PITT, 2013, Seção 7), ou então se postula, em uma visão menos radical, que se pode distinguir entre conteúdos representacionais em sentido estrito ou lato, sendo que em sentido estrito os conteúdos são governados pela estrutura sintática do sistema e indiferentes e invariáveis ao contexto externo, enquanto que os latos são determinados pelos conteúdos estritos associados a fatores externos. De qualquer forma, o pensamento é um processo (interno) de geração, manipulação (por um conjunto de regras sintaticamente estruturadas) e armazenamento de representações internas cujos conteúdos tem poder causal sobre o comportamento e as ações. Para o cognitivismo clássico o conteúdo e a manipulação das representações mentais não dependem, causalmente, do ambiente externo justamente porque as representações representam as características importantes desse ambiente, como afirmado por Haugeland (1991, p. 62):

Um sistema sofisticado (organismo) desenvolvido (evoluído) para maximizar algum fim (*e.g.*, sobreviver) precisa, em geral, ajustar seu comportamento para características, estruturas, ou configurações específicas do seu meio ambiente em modos que podem não ter sido totalmente pré-arranjados em sua configuração. [...] Mas se as características relevantes não estão sempre presentes (detectáveis), então elas podem, pelo menos em alguns casos, estar representadas; isto é, alguma coisa pode estar no lugar delas, com poder de guiar o comportamento em seus lugares. O que está no lugar de outra coisa, neste modo, é uma representação; aquilo para o qual ela está no lugar é seu conteúdo; e estar no lugar de algo para este conteúdo é representá-lo.<sup>13</sup>

---

<sup>13</sup>A sophisticated system (organism) designed (evolved) to maximize some end (*e.g.*, survival) must in general adjust its behavior to specific features, structures, or configurations of its environment in ways that could not have been fully prearranged in its design ... But if the relevant features are not always present (detectable), then they can, at least in some cases, be represented; that is, something else can stand in for them, with the power to guide behavior in their stead. That which stands in for something else in this way is a representation; that which it stands for is its content; and its standing in for that content is representing it.

### **3. A representação mental subsimbólica**

Pode-se distinguir as abordagens computacionais das ciências cognitivas em dois paradigmas: o paradigma cognitivista — ou cognitivismo clássico — e o paradigma conexionista. Da mesma forma que os modelos computacionais do primeiro paradigma, os modelos conexionistas também têm sua origem localizada durante a cibernética dos anos 40, novamente pela influência dos trabalhos de Alan Turing. Devido a sua natureza puramente lógica, a máquina de Turing possibilitou diferentes caminhos implementacionais, um desenvolvido por von Neumann que levou ao computador serial digital (o modelo cognitivista), e outro caminho implementacional que teve como fundadores o neuropsiquiatra Warren McCulloch e o matemático Walter Pitts, que criaram a primeira modelagem conexionista (GARDNER, 1995, p. 159-160). Dupuy (1996, p. 56) diz que o “número de atores” envolvidos na modelagem cibernética da mente não era apenas dois, o organismo e a máquina, mas eram três os agentes envolvidos: “[...] o organismo em sua estrutura (o cérebro), o organismo em sua função (a mente) e a máquina, com esta última desdobrando-se em máquina lógica (máquina de Turing ou máquina de McCulloch e Pitts) e máquina artificial, material (o computador).”

Em 1943 McCulloch e Pitts publicaram seu referencial artigo *A Logical calculus of the ideas immanent in nervous activity*. É nesse artigo que encontramos a primeira caracterização de uma rede de neurônios como um sistema formal, operando de acordo com procedimentos lógicos. A máquina lógica de McCulloch e Pitts é inspirada na estrutura e no funcionamento cerebrais, incluindo em sua modelagem os tão importantes mecanismos de *feedback*, os quais os cibernéticos consideravam como dirigentes do comportamento. Contudo, apesar da inspiração neurofisiológica desse modelo, sua plausibilidade biológica não é tão pertinente quanto se supunha na época — para os padrões da atual neurociência tal modelagem é considerada como sendo não realista, ou abstrata, por não incluir muitos dos aspectos e propriedades cerebrais de baixo-nível (CHURCHLAND; SEJNOWSKI, 1992, p. 136). Mesmo assim, existe, nas máquinas conexionistas, uma equivalência entre os níveis físico e computacional.

Demonstra ele [McCulloch, no artigo de 1943], em princípio, a existência de uma máquina lógica *equivalente* à de Turing (tendo em vista que tudo

que uma pode fazer, a outra pode fazer, e reciprocamente), que pode ser considerada em sua estrutura e em seu comportamento uma idealização da anatomia e da fisiologia do cérebro. Este resultado, na mente de McCulloch pelo menos, constitui um avanço decisivo, já que não é mais o cérebro em sua função (a mente), mas também o cérebro em sua estrutura, o cérebro material, natural, biológico, que é assimilável a um mecanismo e, mais precisamente, a uma máquina de Turing. É assim que o cibernético julga resolver o velho problema da alma e do corpo, ou, em seus próprios termos, do “*embodiment of mind*.” O cérebro e a mente são um e outro uma máquina, e a mesma máquina. O cérebro e a mente, portanto, são um só. (DUPUY, 1996, p. 59, grifo nosso, aspas do autor)

Essa equivalência implica na existência de uma diferença significativa entre os modelos cognitivistas e os modelos conexionistas, pois se ambos dependem de relações de isomorfismo entre estados físicos e estados computacionais — *i.e.*, para cada estado físico deve corresponder um estado lógico, com as mesmas relações causais —, nas máquinas conexionistas a teoria funcionalista se apresenta diferentemente, pois ela se torna parcialmente dependente do estado físico em questão. Ou seja, não se admite que apenas relações de isomorfismo garantam a correspondência necessária e suficiente entre estados lógico-computacionais e quaisquer outros sistemas físicos que possam implementar satisfatoriamente (isomorficamente) tais estados.<sup>14</sup> Gonzales (1991, p. 96, grifos da autora) distingue, neste sentido, entre o funcionalismo lógico-computacional e o funcionalismo neuro-computacional:

Uma versão informal do funcionalismo nos diz que, apesar de as máquinas e os seres humanos serem fisicamente diferentes, ambos possuem sistemas de processamento de informações, cujas operações mentais/computacionais podem ser entendidas em termos de suas *organizações funcionais*. Atualmente existem dois ramos principais do funcionalismo na Ciência Cognitiva, os quais denominaremos: (i) *funcionalismo lógico-computacional*, (ii) *funcionalismo neuro-computacional*. A diferença central entre (i) e (ii) é que, no primeiro caso, os processos mentais são estudados como se fossem apenas computações abstratas — independentemente de suas realizações físicas e do meio ambiente — que desempenham papéis funcionais no sistema cognitivo. Em contraposição, os funcionalistas neuro-computacionais enfatizam a relevância de considerações acerca dos aspectos *físicos*, *neurofisiológicos* e *ambientais* para o estudo e compreensão dos processos mentais. A diferença de enfoque dos processos mentais existente entre as abordagens (i) e (ii) é particularmente relevante para a nossa presente análise das representações mentais. Como ficará claro a seguir, os funcionalistas

---

<sup>14</sup> Como diz Putnam (1980, p. 134): “*We could be made of Swiss cheese and it wouldn’t matter.*”

lógico-computacionais descrevem as representações mentais como *abstrações* que funcionam como mediadores entre o organismo e o seu meio ambiente. Para os funcionalistas neuro-computacionais, entre os quais esta autora se inclui, as representações mentais são estruturas emergentes da ativação de *unidades neurônio-símile* (*neuron like units*) que se *auto-organizam* em função da informação disponível no meio ambiente.

As unidades de processamento conexionistas, ou neurônios-símile ou neurônios artificiais, podem ser configuradas em um sem número de arquiteturas diferentes. A Fig. 2a apresenta um neurônio artificial típico e a 2b uma arquitetura de rede neural multicamada do tipo *feedforward*, uma rede bastante utilizada para classificação de padrões representados por vetores de entrada. Cada neurônio de uma rede tem ligações com outros neurônios, e os sinais enviados de um para outro podem ser mais fortes ou mais fracos, pois são ponderados. Em um dado momento, na dinâmica temporal de uma rede neural artificial, existe uma configuração de ativação dos neurônios envolvidos, e isto é a representação distribuída de algo. Mas a mesma população neuronal também pode representar, e de fato representa, *n* outras coisas em outros momentos. Estas outras representações são, então, outros padrões de ativação, que são armazenadas na forma de pesos sinápticos, ou configurações dos pesos das conexões envolvidas (que variam de acordo com a arquitetura da rede). Tais pesos possibilitam a recuperação de outros padrões de ativação perante um estímulo específico para determinado padrão. Através desses processos podem ser armazenados inúmeros padrões, de acordo com o treinamento — ou fase de aprendizagem — ao qual se submete uma rede neural.<sup>15</sup>

---

<sup>15</sup> Descrever minuciosamente o funcionamento de uma rede neural artificial está fora do escopo deste trabalho e essa descrição visa apenas oferecer ao leitor não familiarizado com a computação conexionista uma ilustração. Sugere-se, para maiores esclarecimentos sobre redes neurais, a leitura dos textos técnicos de Haykin (2008) e Kung (1993), ou a discussão mais conceitual de Churchland e Sejnowski (1992).

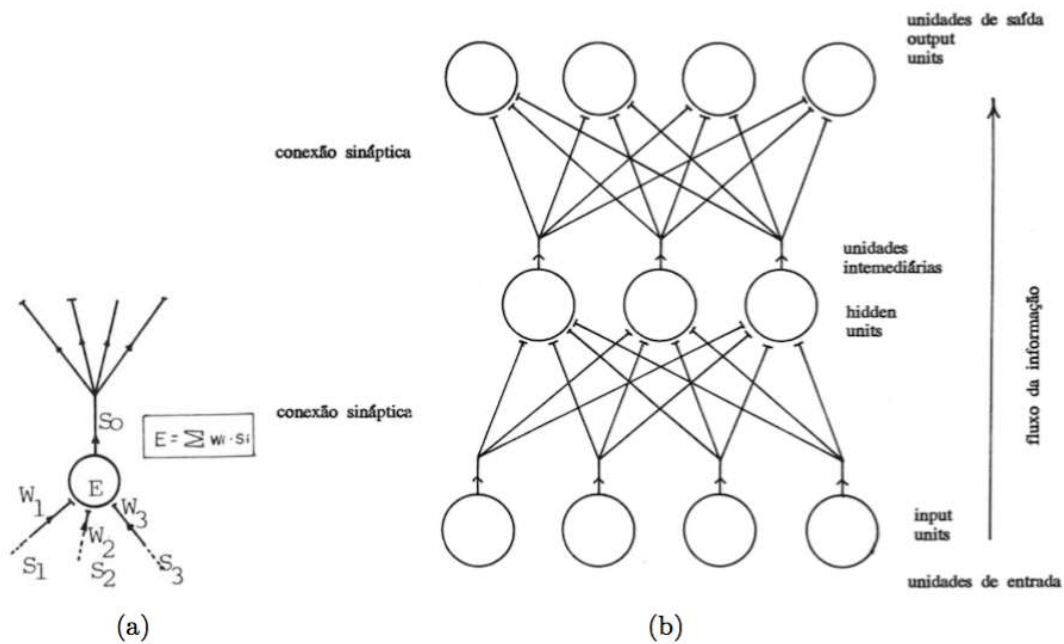


Figura 2: A imagem (a) ilustra uma unidade de processamento de uma rede neural típica, com as conexões de entrada  $S_n$ , os pesos de cada entrada  $W_n$ , a função de ativação  $E$  e as conexões de saída  $S_o$ . A imagem (b) ilustra um exemplo de arquitetura de rede neural artificial multilayer do tipo feedforward. (Imagens retiradas de GONZALEZ, 1991)

Por um lado, arquiteturas conexionistas diversas buscavam e buscam simular (e entender) diferentes organizações de populações neuronais, lembrando-se que a plausibilidade biológica das arquiteturas das redes neurais artificiais é sempre discutível em face dos avanços das neurociências;<sup>16</sup> por outro lado, independentemente da simulação organizacional de populações específicas de neurônios, observou-se que arquiteturas conexionistas distintas possuem propriedades variadas na resolução de problemas lógico-matemáticos. Por exemplo, redes neurais artificiais realizam computações que apresentam a capacidade de aprendizagem (supervisionada ou não-supervisionada) pela exposição a padrões; a capacidade de generalização no reconhecimento de padrões informacionais

<sup>16</sup> Neste sentido, dos modelos conexionistas, é interessante observar que muita plausibilidade biológica pode não ser apropriado para o estudo da mente pois “models that are excessively rich may mask the very principles the models were built to reveal. In the most extreme case, if a model is exactly as realistic as the human brain, then the construction and the analysis may be expensive in computational and human time and hence impractical” (CHURCHLAND; SEJNOWSKI, 1992, p. 137). Acreditamos que eles, os modelos realistas, seriam interessantes na medida em que se quer construir uma mente artificial, mas para explicações científicas sem esta pretensão, esses autores dizem que o importante é “[...] make the model simply enough to reveal what is important, but rich enough to include whatever is relevant to the measurements needed” (idem, ibidem).

estáticos e dinâmicos; a capacidade de classificar e extrair características em conjuntos de dados; capacidade de interpolar e extrapolar funções; robustez ou tolerância a falha e a capacidade de lidar com estímulos ruidosos, entre outras. Encontra-se na literatura diversos trabalhos taxonômicos sobre arquiteturas conexionistas associadas com a resolução de problemas computacionais (HAYKIN, 2008; KUNG, 1993; CHURCHLAND; SEJNOWSKI, 1992, entre outros). Entretanto, as arquiteturas conexionistas as mais variadas compartilham algumas propriedades gerais, às quais podemos vincular, novamente, a discussão das representações mentais, agora sob o novo enfoque que o paradigma conexionista trouxe.

Vimos que, grosso modo, representações mentais podem ser entendidas como objetos internos a uma mente que representam coisas que estão fora dela. Na perspectiva do cognitivismo, entende-se que elas possuem uma natureza estritamente simbólica, ou seja, não há correspondência direta e necessária entre o símbolo e o que ele representa, a não ser a convenção de um código (ou, nas palavras de Fodor, uma ‘linguagem do pensamento’), independentemente de existir ou não um mapeamento um-para-um entre representações e coisas representadas (HASELAGER, 1997, p. 47). De maneira semelhante à reformulação conexionista do funcionalismo, que passou a defender que o nível mental é dependente do, ou ao menos possui estreita relação com o, nível material e dos processos que ali ocorrem, o entendimento das representações mentais também é revisitado no paradigma conexionista. Em termos bastante sintéticos, assume-se que se as arquiteturas computacionais são modelos para as ciências cognitivas — e já vimos, neste trabalho, inúmeras citações que corroboram esta afirmação — o modelo de processamento paralelo e distribuído das redes neurais é mais interessante, entre outras coisas, por ser menos dualista, porque considera que o nível físico é tão importante quanto o nível lógico para a explicação de processos cognitivos (ou, se se quiser assumir o argumento reducionista exposto por Dupuy na citação que há pouco trouxemos, nesta seção, a mente e o cérebro são a mesma máquina).

Dentro do paradigma conexionista, as representações mentais são usualmente denominadas de representação subsimbólica ou distribuída. O entendimento conexionista da representação mental vincula-se ao fato de que desde a década de 50 as neurociências gradativamente deixaram a teoria da codificação local (*local coding theory*) da informação

em função de uma forma de representação distribuída (ou *vector coding* ou multidimensional representation) da informação (CHURCHLAND; SEJNOWSKI, 1992, p. 157-162).

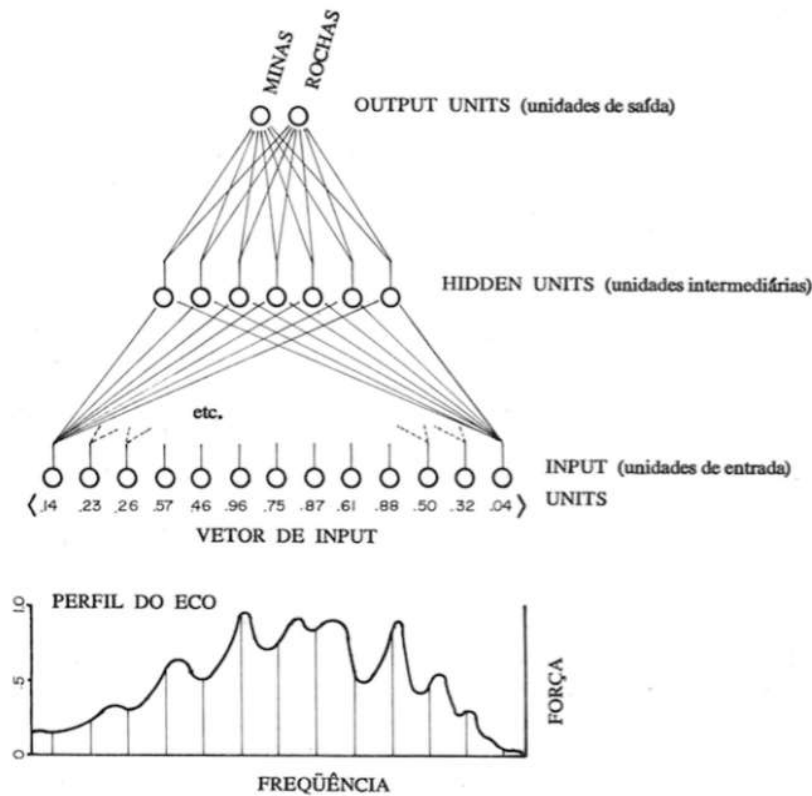


Figura 3: Exemplo de reconhecimento de padrões por processos de aprendizagem supervisionada para duas classes de categorias, a partir da decomposição de um sinal acústico em vetores de entrada. (Imagens retiradas de GONZALEZ, 1991)

Em primeiro lugar, a representação distribuída ocorre sobre um conjunto ou população de neurônios que representa um objeto (enquanto que no *local coding*, cada neurônio individualmente seria responsável por representar uma característica específica do objeto representado). Ou seja, a representação mental não é um objeto único (*token*), um símbolo armazenado, mas encontra-se espalhada sobre um conjunto de unidades integradas, formando um espaço multidimensional. Como ilustração, a Fig. 3 apresenta uma rede neural multicamada do tipo *feedforward* que classifica o tipo de material geológico a partir do sinal acústico captado, o qual é transformado em um vetor de valores que alimenta a camada de entrada da rede. Neste exemplo, os padrões que a rede aprende a classificar não representados em algum dos neurônios da rede, mas é a distribuição (dos pesos das

conexões) que representa ao mesmo tempo cada um dos padrões. Observe na Fig. 3 que todas as unidades são conectadas a todas as demais unidades. Outra característica importante de ser destacada nas arquiteturas computacionais conexionistas é que a população neuronal é sempre a mesma para diferentes representações, fato que van Gelder (1991) chama de sobreposição (no caso do exemplo da Fig. 3 as duas categorias que a rede classifica são representadas nos mesmos neurônios da camada de entrada e intermediária). Em outras palavras, nenhum objeto ou conteúdo ou característica de um objeto é representado por um único neurônio (van GELDER, 1992).<sup>17</sup> Em segundo lugar, outra característica fundamental das representações distribuídas é que elas apresentam redundância, visto que várias unidades representam a mesma coisa. A redundância representacional relaciona-se, ainda, à robustez do sistema quanto a falhas, já que a perda de uma ou algumas unidades neuronais não implica na eliminação completa de informações ou representações (ao contrário, se cada unidade armazenasse exclusivamente um símbolo ou *token*, sua inoperância resultaria na perda desse código, invariavelmente).

#### **4. A visão não-representacional da mente**

As seções anteriores evidenciam, e as inúmeras citações nelas contidas corroboram, o argumento de que as ciências cognitivas vinculam-se de maneira bastante forte aos modelos computacionais da mente, seja pelo viés do cognitivismo clássico — e do modelo do computador serial digital —, seja pelo viés do conexionismo — e o modelo das máquinas de processamento paralelo e distribuído. Vimos que a base para os modelos computacionais da mente dentro dos paradigmas cognitivista e conexionista é a noção de máquina de Turing, diferenciando os modelos (e as teorias sobre eles desenvolvidas) de cada abordagem pelas questões colocadas pela distinção entre as versões lógico-computacional e neuro-computacional do funcionalismo. Pudemos perceber, portanto, que dentro dessas abordagens computacionais, apesar das diferenças encontradas, a noção de representação mental é indispensável. Destarte, podemos estabelecer dois pressupostos

---

<sup>17</sup> Com relação a esta última afirmação, Churchland e Sejnowsky (1992, p. 165) discordam parcialmente: “[...] in some instances, the vector components may correspond to features such as the color, shape, motion, etc. of an object, but in other cases there may be no identifiable microfeature corresponding to any component.



centrais à ciência cognitiva:

- (i) as ciências cognitivas, tradicionalmente, apoiam-se sobre a *visão computacional da mente*,<sup>18</sup>
- (ii) a noção de *representação mental* é, neste contexto computacional, condição necessária para a explicação dos processos cognitivos.

Não obstante, mais recentemente começam a surgir pesquisadores insatisfeitos com estes pressupostos carregados pelos estudos da cognição e tão amplamente aceitos dentro das ciências cognitivas. A crítica à visão computacional da mente se constrói, então, na confluência de diferentes propostas teóricas que, de um modo ou de outro, entendem que modelos cognitivos devem ir além das proposições explicativas baseadas exclusivamente processamento interno de informação (representação), *i.e.*, dos modelos computacionais da mente.

Entretanto, até hoje essa orientação alternativa não tem um nome bem estabelecido, sendo mais um guarda-chuva que cobre um grupo relativamente pequeno de pessoas trabalhando em diversos campos. Propomos como nome o termo enacionismo, para enfatizar a convicção crescente de que a cognição não é representação de um mundo preestabelecido por uma mente preestabelecida, mas, ao contrário, é a atuação de um mundo e de uma mente sobre as bases de uma história da variedade de ações que um ser executa em um mundo. (VARELA; THOMPSON; ROSCH, 1991, p. 9)<sup>19</sup>

As abordagens da cognição adotadas e propostas por este grupo sem nome, que Varela et al. chamam de enacionistas, nós abarcaremos no que chamamos de *paradigma dinâmico da cognição*. Entendemos que, tratando da questão da representação mental, na

---

<sup>18</sup> Associada a esta afirmação existe uma considerável e interessante (e datada) discussão sobre máquinas serem ou não sistemas dotados de estados mentais, *i.e.*, se máquinas computacionais são apenas *modelos da mente* ou se, de fato, elas realizam estados mentais. Neste sentido, Searle (1980), por exemplo, distingue entre o que chama de inteligência artificial forte e fraca. Esta discussão ultrapassa os limites deste trabalho e, aqui, estabelecemos o pressuposto (i), acima, em termos epistemológicos nos estudos da cognição e, portanto, não consideramos e não iremos discutir a tese forte da inteligência artificial.

<sup>19</sup> As yet, however, this alternative orientation does not have a well-established name, for it is more an umbrella that covers a relatively small group of people working in diverse fields. We propose as a name that term enactive to emphasize the growing conviction that cognition is not the representation of a pre-given world by a pre-given mind but is rather the enactment of a world and a mind on the basis of a history of the variety of actions that a being in the world performs.

visão desses pesquisadores — os quais vêm de diferentes áreas, da biologia à psicologia, da inteligência artificial à semiótica, da linguística à música etc — considerações sobre a cognição devem olhar o sistema como algo efetivamente incorporado e atuante no mundo e não como um sistema abstrato de operações lógicas sobre representações simbólicas, destacado de sua materialidade, e também não como uma população altamente conectada de neurônios de um cérebro sem corpo, sem ação no mundo, sem uma história que se correlaciona com essa atuação. Como nos coloca Haselager, Bongers e Rooij (2003, p. 6):

Em primeiro lugar, este argumento destaca que nosso sistema cognitivo está essencialmente fundamentado em nosso cérebro e este está integrado com o nosso corpo. Muito da natureza dos nossos processos cognitivos está determinado pelas possibilidades específicas de ação oferecidas pelo nosso corpo. Isto é, nosso sistema cognitivo não é um sistema de processamento de informação independente, mas um sistema que está essencialmente corporificado. Em segundo lugar, a abordagem dinâmica enfatiza o fato de que cognição e comportamento são resultado da interação entre nosso sistema cognitivo e seu entorno imediato. Desta forma, ela reforça o fato de que cognição e comportamento estão situados em uma rede causal que engloba o cérebro, o corpo e o ambiente. Sobretudo, a abordagem dinâmica confere uma estrutura na qual a cognição e o comportamento são tomados como resultado de interações dinâmicas complexas entre cérebro, corpo e ambiente. O surgimento da abordagem dinâmica coincide com um descontentamento crescente com as abordagens preponderantes da cognição. Em particular, existe uma séria dúvida sobre, e muitas vezes uma forte insatisfação com, a visão tradicional de que a cognição deve ser definida em termos de transformações computacionais sobre representações internas.<sup>20</sup>

Se quisermos ser mais pontuais e mais precisos em apontar razões específicas para esse descontentamento, podemos verificar quais são as ‘dores de cabeça’ que a noção de representação mental infligiu aos pesquisadores vinculados aos paradigmas

---

<sup>20</sup> First, this claim highlights the fact that our cognitive system is essentially grounded in the brain as it is integrated with our body. Much of the nature of our cognitive processes is determined by the specific action possibilities afforded by our body. That is, our cognitive system is not an independent information-processing system, but a system that is essentially embodied. Second, the dynamical approach emphasizes the fact that cognition and behavior are the result of the interaction between our cognitive system and its immediate environment. As such, it stresses the fact that cognition and behavior are embedded in a causal web that encompasses brain, body, and environment. In all, the dynamical approach entails a framework in which cognition and behavior are taken to result from complex dynamical interactions between brain, body, and environment. The advent of the dynamical approach coincides with a growing discontent with the prevailing approaches to cognition. In particular, there is serious doubt about, and sometimes strong dissatisfaction with, the traditional view that cognition should be defined in terms of computational transformations over internal representations.”

computacionais. Haselager (2004) diz que o mal estar do representacionismo<sup>21</sup> decorre de sete ‘dores de cabeça:’ o problema do frame; o problema de estrutura representacional; o problema das raízes do significado; o problema da operacionalização; o problema do observador; o problema da dependência; o problema da ubiquidade da representação.

O *problema do frame* diz respeito à tentativa (fracassada) de formalização do conhecimento de senso comum; *i.e.*, explicar através dos modelos cognitivistas ações que realizamos em nosso dia com bastante eficiência sem a necessidade de muito esforço intelectual, como por exemplo, dirigir por uma cidade conversando com um amigo passageiro. A inteligência artificial teve muito mais sucesso em criar algoritmos que joguem xadrez com o campeão mundial do que em criar uma máquina que consiga servir uma xícara de café e saber quanto açúcar colocar ou um robô que ande de bicicleta. Parece, pois, que alguns domínios são mais propícios para que modelos cognitivistas ou mesmo conexionistas atuem eficientemente e, assim, podem ser considerados como modelos dignos de estudo. Como colocado por Varela, Thompson e Rosch (1991, aspas dos autores):

Essa abordagem [simbólica] da cognição como resolução de problemas funciona, até um certo grau, para tarefas que estão em domínios nos quais é relativamente fácil especificar todos os estados possíveis. Considere, por exemplo, o jogo de xadrez. É relativamente fácil definir os constituintes do “espaço do xadrez:” há posições no tabuleiro, regras para os movimentos, jogadas que são feitas e assim por diante. Os limites desse espaço são claramente definidos; de fato, é quase um mundo cristalino. Não é surpreendente, então, que o jogo de xadrez seja uma arte avançada na computação.<sup>22</sup>

Ao contrário de ações que podemos fazer em domínios tais quais esse mundo cristalino do jogo de xadrez, as ações diárias que realizamos parecem ser melhor explicadas sem a mediação da representação mental e do processamento de informação necessariamente requerido. Gibson (1979, 1966) propôs a teoria da percepção direta, e a

---

<sup>21</sup> Encontra-se, muitas vezes, variações na grafia deste termo, que ora se apresenta como representacionalismo e ora como representacionismo. Tratamo-os como sinônimos.

<sup>22</sup> This [symbolic] approach to cognition as problem solving works to some degree for task domains in which it is relatively easy to specify all possible states. Consider for example the game of chess. It is relatively easy to define the constituents of the “space of chess:” there are positions on the board, rules for movements, turns that are taken, and so on. The limits of this space are clearly defined; in fact, it is an almost crystalline world. It is not surprising, then, that chess playing by computer is an advanced art.

chamou de direta por que explica a relação entre percepção e ação sem a mediação de representação mental, especialmente nessas questões vinculadas ao problema do frame. Gibson explica as ações estabelecendo uma ligação causal não-mediada entre percepção e ação, porque, como colocam Michaels e Carello (1981, p. 47), “para a percepção ser válida (e valiosa), ela deve estar manifesta em ações apropriadas e efetivas em um ambiente. E, por sua vez, porque ações, para serem efetivas e apropriadas, devem estar restringidas por percepções precisas do meio ambiente.”<sup>23</sup>

O *problema da estrutura representacional* diz respeito a forma adequada de representação para explicar os processos cognitivos. Vimos, neste trabalho, que o modelo de representação simbólica, apesar de dominante dentro das teorias cognitivas propostas, limita-se, em seu alcance explicativo a certos domínios; por outro lado, que o advento do conexionismo possibilitou a superação de falhas dos modelos clássicos, embora também apresente limitações em seu alcance explicativo. Encontra-se na literatura uma considerável quantidade de artigos de proponentes de ambos os modelos, com réplicas e tréplicas que visavam estabelecer os limites dos modelos explicativos. Haselager (2004, p. 4) nos esclarece que o problema mais preocupante é que ambos os lados estão certos:

É preocupante que os fundamentos da ciência cognitiva, que se alicerçam nas representações, sejam objetos de acalorados debates sobre sua correta estrutura. Esta questão se torna mais séria ainda quando as críticas a ambas as estruturas representacionais estão basicamente corretas. A questão que surge em decorrência disso é: será a pedra fundamental da ciência cognitiva tão sólida quanto se pensava?

O *problema das raízes do significado* toca um ponto importante e relacionado a algumas das discussões que aparecem nas seções anteriores deste trabalho, especialmente a segunda. Se uma representação é um símbolo, o que estabelece, ou por qual processo se estabelece, a correlação entre uma representação e seu significado? Pesquisadores como Fodor, por exemplo, oferecem teorias como a linguagem do pensamento, que é, basicamente, uma hipótese de que processos mentais são resultado da manipulação sintática de representações simbólicas, cujos significados se estabeleçam sobre “paralelos” entre os

---

<sup>23</sup> For perception to be valuable, it must be manifested in appropriate and effective actions on the environment. And, in turn, for actions to be appropriate and effective they must be constrained by accurate perception of the environment.

domínios sintáticos e semânticos. Ora, o que garante esse paralelo? Na perspectiva do cognitivismo o significado deixa de ser inerente ao sistema, para, na melhor das possibilidades, ser nele colocado por algum agente externo (como o programador). No conexionismo a questão se apresenta diferentemente, de um modo menos linguístico, mas também de maneira problemática, já que as representações são mais inerentes ao sistema, no entanto ainda dependentes de arquiteturas específicas para problemas específicos, e sem conseguir oferecer insight sobre como as representações adquirem significados, inclusive simbolicamente — retornaremos a esta questão mais a frente, com um possível encaminhamento.

O *problema da operacionalização* relaciona-se a abordagens mais recentes da computação, como os autômatos celulares ou os algoritmos genéticos ou a computação evolutiva, nos quais não existe programação prévia de sistemas representacionais, nem em termos simbólicos nem em termos subsimbólicos. Shellard et al. (2010, p. 410) afirmam que, “dentre os métodos de computação adaptativa, a computação evolutiva (CE) é aquele inspirado nas estratégias biológicas de adaptação em populações de indivíduos, como inicialmente descrito por Charles Darwin. CE é normalmente usada para encontrar a melhor solução possível para problemas quando não existem informações suficientes para resolvê-los através de métodos formais (determinísticos).”<sup>24</sup> Essas recentes abordagens computacionais colocam em cheque a tradicional noção de representação das ciências cognitivas porque emprega modelos que lidam com problemas mal formulados, do ponto de vista informacional, e que, portanto, não podem ser representados *a priori*, ou seja, não podem ser operacionalizados através dos métodos tradicionais de computação.

O *problema do observador* relaciona-se a premissas como o mote de Fodor de que sem representação não há modelos de processos cognitivos. Trata-se, quase, de um axioma da ciência cognitiva. Haselager (2004, p. 6) postula que assim como projetamos nossa intencionalidade e desejos sobre objetos e outros seres, como cachorros, gatos, computadores ou robôs, os cientistas cognitivos também o fazem, com respeito a representações mentais: “sempre que observam o comportamento de um sistema,

---

<sup>24</sup> Among computational adaptive methods, Evolutionary Computation (EC) is the one inspired into the biological strategy of adapting populations of individuals, as initially described by Charles Darwin. EC is normally used to find the best possible solution to problems when there is not enough information to solve it through formal (deterministic) methods.

manifestam essa compulsão de atribuir-lhes representações internas.” O autor aponta que um dos fundadores da ciência da cognição, Herbert Simon, já havia nos alertado sobre esse fato, ao afirmar que a complexidade dos nossos processos cognitivos são decorrência da complexidade do ambiente que nos envolve. Nesse sentido, é bastante tentadora a comparação desta afirmação com um dos argumentos associados com teoria da percepção direta de James Gibson, que citamos brevemente alguns parágrafos acima. Gibson entende que a descrição dos processos cognitivos pelos sistemas de processamento de informações internamente representadas associa-se a ideia de que o ambiente é pobre (ou confuso) informacionalmente e, que, assim, um sistema cognitivo precisa elaborar os estímulos sensoriais para poder construir informações que o permitam ligar com seu entorno. Os sistemas perceptivos, no entendimento da teoria da percepção direta, vinculam-se à ideia de que o ambiente é, contrariamente, extremamente rico em termos informacionais e que, portanto, o sistema perceptivo precisa, ao invés de elaborar representações informativas, filtrar as informações ambientais, em um processo de sintonia (self-tuning) entre organismo e ambiente.

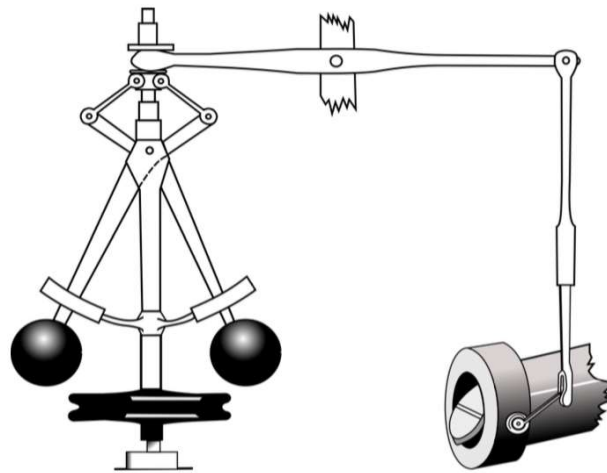
O *problema da dependência* é que os cientistas cognitivos, ao pensar em soluções para problemas — ou quando propõem modelos para a cognição —, não conseguem escapar à tentação de usar a noção de representação mental, não verificando a cada caso se esta noção é realmente necessária. Este vício, esta dependência, decorre da superexposição a esse tipo de solução, criando um hábito tão enraizado que sequer se cogitam outros tipos de explicação. Haselager aponta, como ilustração de tal problema (e de vários dos outros apontados), o controlador centrífugo, ao qual iremos discutir na próxima seção.

Por fim, chegamos ao *problema da ubiquidade da representação*, o qual iremos discutir em maiores detalhes neste artigo, na próxima seção. O problema da ubiquidade da representação é abordado em alguns artigos (HASELAGER, 2004; HASELAGER; BONGERS; ROOIJ, 2003; HASELAGER; GROOT; RAPPARD, 2003), e o sintetizamos aqui. Haselager e seus coautores defendem que se possa ter uma visão ampliada da representação; visão esta que resulta em se enxergar uma ampla variedade de processos como sistemas representacionais, incluindo sistemas mecânicos como o controlador centrífugo. O problema da ubiquidade da representação, então, é que ela perde o seu poder explicativo quando se aplica a qualquer sistema. Certamente, nesse entendimento ampliado

da representação, as noções mais restritas da representação mental simbólica ou subsimbólica não bastam, e o debate sobre a forma da representação mental se transforma o debate sobre a existência ou não da representação mental. O argumento, pois, é que a ubiquidade da representação enfraquece sua condição de fundamento explicativo da ciência cognitiva. Trataremos mais desta questão a seguir.

### **5. O controlador centrífugo e o argumento da ubiquidade da representação mental**

Quando van Gelder (1998), em seu seminal artigo, formula a hipótese dinâmica do estudo da cognição ele assume que sistemas cognitivos podem ser ontologicamente sistemas dinâmicos ou podem ser descritos enquanto tal. Um dos modelos mais populares na discussão dinâmica da cognição, tanto pelo lado daqueles que defendem esta abordagem, quanto pelo lado dos que continuam a estudar processos cognitivos pelos paradigmas computacionais, é o *Watt's governor* ou controlador centrífugo, o qual Tim van Gelder (1995) já havia vinculado à questão da cognição em termos não-computacionais.



*Figura 4: Representação ilustrativa de um Watt's governor, um dispositivo de finais do século XVIII empregado para controlar a pressão de admissão de vapor em máquinas de tecelagem.*

A Fig. 4 ilustra esse dispositivo que foi criado por James Watt para controlar a pressão de admissão de vapor para máquinas de tecelagem. Tal dispositivo pode ser visto como um sistema que controla (ou autocontrola) a pressão de vapor pelo acoplamento

físico com uma máquina movida a vapor. A rotação do motor é transmitida por uma correia ao eixo vertical no lado esquerdo da Fig. 4. Por influência da força centrífuga, em oposição à força da gravidade, a velocidade da rotação transmitida a este eixo faz com que as massas cilíndricas movimentem-se centrífuga ou centripidamente. Articulado às massas, liga-se um eixo de conexão que movimenta a válvula de admissão de vapor do motor. Dessa forma, fecha-se um círculo de causalidade que autorregula a velocidade de rotação do motor em função da manutenção da pressão de admissão de vapor.

Tim van Gelder, assim como todos os dinamicistas que o seguem analisando o controlador centrífugo, entendem que se trata de um sistema que consegue se autorregular sem envolver a noção de representação. O paradigma dinâmico da cognição traz à luz diversos conceitos da teoria dos sistemas dinâmicos aplicados ao domínio da cognição, como a auto-organização, propriedades, estruturas e comportamentos emergentes, acoplamentos informacionais e a crítica recorrente o emprego da representação mental como condição necessária para explicações da cognição e, por extensão, de sistemas que apresentam comportamentos autônomos, mesmo se não o considerarmos, de fato, sistemas cognitivos. Se voltarmos às origens cibernéticas da ciência cognitiva é possível entender que este aparato de controle tem muito em comum com os interesses daqueles pesquisadores (mecanismos de autocontrole com a utilização de *feedback* em relações causais e/ou informacionais). Parece, então, que a ciência cognitiva, em suas próprias origens, optou por enxergar os fenômenos que estudava sempre pelo viés da representação interna; esquecendo-se que, em muitos casos, possam existir outras explicações, às vezes mais simples e diretas — lembremos da questão da dependência. Haselager (2004, p. 6) traduz a solução que van Gelder diz ser aquela que um cientista cognitivo tipicamente apresentaria para solucionar a regulação das máquinas de tecelagem a vapor:

1. Medir a velocidade da roda;
2. Comparar a velocidade atual com a velocidade desejada;
3. Se não há discrepância, voltar ao primeiro passo. Se houver:
  - a. medir a pressão usual do vapor;
  - b. calcular a pressão desejada da pressão do vapor;
  - c. calcular o ajuste necessário na válvula;



4. Fazer o ajuste necessário na válvula reguladora da pressão;
5. Voltar ao primeiro passo.

Esta solução é uma proposta evidentemente representacional. Todos os cálculos são realizados a partir de representações discretas de estados de partes do sistema armazenados internamente em um sistema de processamento computacional. A partir de tais cálculos o sistema de controle efetiva as ações para correção da atuação da válvula, quando necessário. Comparando essas duas perspectivas para a solução do mesmo problema, a dinâmica e a computacional, parece-nos que o modelo de Watt é mais simples, mais eficiente e mais robusto. Assim, corrobora-se o argumento de Haselager de que a representacional decorre mais de um hábito epistemológico do que de uma necessidade efetiva.

Entretanto, existe uma terceira opção. Haselager, Groot e Rappard (2003) discutem o argumento (e o diagrama) postulado por Bechtel (1998) de que o governador centrífugo é isomórfico a um sistema computacional e, portanto, representacional, mesmo sem ser um algoritmo como este que van Gelder hipotetizou como típico do cognitivismo. Na Fig. 5a, pode-se observar as relações causais operantes no *Watt's governor*; na Fig. 5b, paralelamente, observa-se a versão representacional do mesmo diagrama. Bechtel sustenta que existe uma *relação funcional de isomorfismo* entre os dois diagramas e, dessa forma, assume uma explicação representacional do mecanismo de autocontrole.

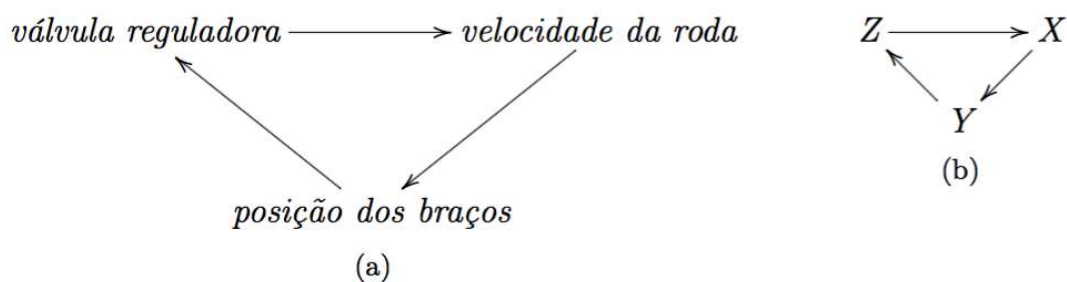


Figura 5: Análise diagramática de Bechtel (1998) do controlador centrífugo (a), e como sistema representacional (b), onde X é o objeto ou evento representado, Y o objeto ou evento representando e Z o sistema que usa Y para coordenar o comportamento com X.

O argumento de Haselager e colaboradores é que posturas como a de Bechtel enfraquecem a capacidade da ciência cognitiva de explicar processos cognitivos, justamente por enxergar tudo como sistema representacional: a tese da ubiquidade da representação (e a ‘dor de cabeça’ dela decorrente). Nos parece que, ao fim e ao cabo, nessa tese da ubiquidade, assume-se uma postura duplamente problemática. Por um lado, critica-se a representação estrito senso dos paradigmas cognitivista e conexionista, pois formas tão restritivas de representação mental não são adequadas a toda a variedade de processos os quais descrevemos como cognitivos. Como visto, representações simbólicas lidam bem com processos linguísticos ou quase-linguísticos, com domínios abstratos e cristalinos, e apresentam dificuldades em sistemas que não requerem este tipo de processamento informacional; representações subsimbólicas lidam bem com fenômenos mais concretos, e conseguem abstrair informação mesmo em ambientes com variações nos padrões de entrada, em mundos não tão cristalinos. Mas, como o *problema do frame* evidencia, para muitas ações corriqueiras as noções estritas de representação parecem mais complicar do que ajudar, especialmente em casos que envolvem a causalidade circular entre percepção e ação. Temos a impressão de que se coloca o debate sobre o representacionalismo como um caso do tipo tudo-ou-nada, como se sistemas cognitivos ou possuem representações simbólicas ou subsimbólicas ou não as possuem.

Por outro lado, critica-se argumentos como o de Bechtel pela sua permissividade (ou promiscuidade) explicativa a partir dessa noção de representação lato senso e ubíqua. Utiliza-se uma versão tão flexível ou tão polivalente do conceito sobre uma gama tão ampla de fenômenos que o emprego dessa própria noção se enfraquece e o conceito basal que deveria servir para explicar os processos cognitivos, inclusive por distinguí-los dos não-cognitivos — e por distinguir a ciência cognitiva do behaviorismo—, perde sua aplicabilidade. Se tudo é representação, e se representação é esse conceito tão vago, a ciência cognitiva perde sua capacidade explicativa. Sobre essa noção lato senso da representação, Haselager, Groot e Rappard (2003) dizem que se o *Watt's governor* é um sistema representacional, uma bicicleta também o é. Será que algum cientista cognitivo diria que uma bicicleta é um sistema cognitivo? Se não é, o que, então, caracteriza um sistema como cognitivo? Seria, de fato, a noção de representação mental?

## **6. Considerações finais**

Neste artigo trouxemos uma célere reconstrução da história da ciência cognitiva — sem pretendermos fazer uma historiografia — na qual destacamos três principais enfoques, sob os quais englobamos a maioria das teorias: o paradigma cognitivista; o paradigma conexionista; o paradigma dinâmico da cognição. Buscamos descrever e elucidar os conceitos de representação mental relacionados a cada uma dessas perspectivas. Tratamos da representação mental simbólica, das suas potencialidades e das suas limitações; trouxemos a versão conexionista da representação subsimbólica e o debate que ela trouxe ao campo dos modelos cognitivos; por fim, apresentamos a crítica à própria noção de representação mental que as diversas abordagens dinamicistas compartilham. Nosso intuito foi o de ilustrar o debate a partir de inúmeras citações que evidenciam as três posições e que destacam argumentos em defesa desta ou daquela posição. Prosseguimos descrevendo as ‘dores de cabeça’ representacionais que Haselager imputa à ciência cognitiva computacional. Dos problemas apresentados, destacamos a questão da ubiquidade representacional e discutimos o sintomático exemplo do controlador centrífugo. Encerramos a seção anterior com a pergunta sobre o que caracteriza um sistema cognitivo e se o conceito de representação seria suficiente, então, para tanto.

Mesmo que não respondendo a tais questões, aqui — e tal resposta extrapolaria os limites deste trabalho —, sugerimos uma hipótese. Como observamos ao fim da seção anterior, muitas vezes o debate acaba tornando-se uma querela do tipo tudo-ou-nada: ou tem-se modelos cognitivos representacionalmente simbólicos, ou tem-se modelos subsimbólicos, ou tem-se modelos não-representacionistas. O incômodo com essa postura muito rigorosa entre um tipo, ou outro, ou nenhum, nos leva a considerações de que a cognição pode ser mais — e arriscaríamos a dizer que, certamente, é mais — do que aquilo que um modelo muito restritivo pode oferecer. O comportamento inteligente engloba inúmeros processos e o debate entre as formas de representacionalismo mais serviu para destacar e distanciar teorias divergentes (que muitas vezes buscavam e buscam explicar fenômenos diferentes) do que integrá-las em abordagens mais amplas e promissoras. Modelos alternativos precisam ser integrados; visões alternativas precisam ser integradas.

Contudo, não estamos defendendo simplesmente que se congregue teorias contrárias ou contraditórias; faz-se necessária, ao contrário, uma abordagem integradora que seja consciente dos alcances e limites explicativos das abordagens existentes e da problemática envolvida com as diferentes formas representacionais. O ponto, então, não é simplesmente juntar-se alhos com bugalhos, mas de saber-se *como* juntar. A integração exige esforços atentos e, talvez se admita, a colaboração de outras áreas — e nesta questão apresentamos uma última hipótese.

O problema colocado à ubiquidade da representação é que, sendo tudo representação, se perde a força explanatória desta noção tão fundamental à ciência cognitiva. Nossa hipótese é que o problema não é ser tudo representação, mas ser tudo a mesma e vaga noção de representação, ou ser tudo descrito pela mesma e vaga noção de representação. Nos parece evidente, então, que para uma superação desse embate, precisamos nos aprofundar sobre este conceito e se a ciência cognitiva não tem ferramentas para tanto, a filosofia peirceana, especialmente a semiótica, pode colaborar substancialmente. O estudo rigoroso dos tipos, das formas, das propriedades das variadas formas de representação, associadas às considerações lógicas do raciocínio e da criatividade, do papel do acaso e da evolução nos processos sógnicos, amplia o entendimento cognitivo sem incorrer aos perigosos holismos apressados e quebradiços. O rigor da semiótica e a amplitude epistemológica da filosofia de C.S. Peirce possibilitam uma visão integradora dos fenômenos cognitivos, sem derrubar as bases que sustentam a ciência cognitiva, de uma forma consciente dos problemas envolvidos no debate sobre o representacionalismo.

## **Referências**

- BECHTEL, W. Representations and cognitive explanations: Assessing the dynamicist's challenge in cognitive science. *Cognitive Science*, v. 22, n. 3, p. 295–318, 1998.
- CHURCHLAND, P. S.; SEJNOWSKI, T. *The computational brain*. Cambridge, MA: MIT Press, 1992.
- DUPUY, J.-P. *Nas origens das ciências cognitivas*. São Paulo: Editora Unesp, 1996.
- ECKARDT, B. V. *Mental representation*. The MIT Encyclopedia of Cognitive Sciences, p. 527–529, 1999.
- FODOR, J. A. *The language of thought*. Cambridge, MA: Harvard University Press, 1975.
- FODOR, J. A.; PYLYSHYN, Z. W. Connectionism and cognitive architecture: A critical analysis. *Cognition*, v. 28, n. 1-2, p. 3–71, 1988.

- GARDNER, H. *A Nova Ciência da Mente: Uma História da Revolução Cognitiva*. São Paulo: EDUSP, 1995.
- GELDER, T. v. *What is the 'd' in 'pdp': a survey of the concept of distribution*. In: RAMSEY, W.; STICH, S. P.; RUMELHART, D. E. (Ed.). *Philosophy and connectionist theory*. Hillsdale: Erlbaum. p. 33–59, 1991.
- \_\_\_\_\_. Defining 'distributed representation'. *Connection science*, v. 4, n. 3-4, p. 175–191, 1992.
- \_\_\_\_\_. What might cognition be, if not computation? *The Journal of Philosophy*, v. 92, n. 7, p. 345–381, 1995.
- \_\_\_\_\_. The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences*, v. 21, n. 5, p. 615–628, 1998.
- GIBSON, J. J. *The Senses Considered as Perceptual Systems*. Boston: Houghton Mifflin Company, 1966.
- \_\_\_\_\_. *Ecological Approach to Visual Perception*. Hillsdale: Lawrence Erlbaum Associates Publishers, 1979.
- GONZALES, M. E. Q. Redes neurais e representação mental: um ensaio sobre harmonia e racionalidade. *Trans/Form/Ação*, v. 14, p. 93–108, 1991.
- HASELAGER, P.; GROOT, A. de; RAPPARD, H. v. Representationalism vs. anti-representationalism: a debate for the sake of appearance. *Philosophical psychology*, v. 16, n. 1, p. 5–24, 2003.
- HASELAGER, W. *Cognitive science and folk psychology: the right frame of mind*. London: Sage, 1997.
- \_\_\_\_\_. *O mal estar do representacionismo: sete dores de cabeça da ciência cognitiva*. In: FERREIRA, A.; GONZALEZ, M.; COELHO, J. (Ed.). *Encontro com as ciências cognitivas*. São Paulo: Coleção Estudos Cognitivos, 2004. v. 4, p. 105–120.
- HASELAGER, W. F.; BONGERS, R. M.; ROOIJ, I. van. *Cognitive science, representations and dynamical systems theory*. In W. Tschacher and J-P. Dauwalder (Eds.) *The dynamical systems approach to cognition: Concepts and empirical paradigms based on self-organization, embodiment, and coordination Dynamics*. Singapore: World Scientific, 2003.
- HAUGELAND, J. *Representational genera*. In: RAMSEY, W.; STICH, S.; RUMELHART, D. (Ed.). *Philosophy and connectionist theory*. Hillsdale, NJ: Erlbaum, 1991.
- HAYKIN, S. *Neural networks: a comprehensive foundation*. New York: Prentice Hall, 2008.
- KUNG, S. Y. *Digital neural networks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- MCCULLOCH, W.; PITTS, W. *A logical calculus of ideas immanent in nervous activity*. *Bulletin of Mathematical Biophysics*, Cambridge, Mass., v. 5, n. 5, p. 115–133, 1943.
- MICHAELS, C. F.; CARELLO, C. *Direct perception*. Englewood Cliffs, NJ: Prentice-Hall, 1981.
- OLIVEIRA, L. F. O estudo da música a partir do paradigma dinâmico da cognição. *Percepta*, v. 2, n. 1, p. 17–36, 2014.
- PITT, D. *Mental representation*. In: ZALTA, E. N. (Ed.). *The Stanford Encyclopedia of Philosophy*, 2013.
- PUTNAM, H. *Philosophy and our mental life*. In: BLOCK, N. (Ed.). *Readings in the Philosophy of Psychology*. Cambridge, MA: Harvard University Press, p. 134–143, 1980.
- SEARLE, J. Mind, brain, and programs. *Behavioral and Brain Sciences*, v. 3, p. 417–457,

1980.

SHELLARD, M., OLIVEIRA, L. F., FORNARI, J. E., and MAZOLLI, J. *Abduction and meaning in evolutionary soundscapes*. In: Model-based reasoning in science and technology. Berlin: Springer, v. 314, n. X, 2010.

SLOBODA, J. A. *A mente musical: a psicologia cognitiva da música*. Curitiba: EDUEL, 2008.

TURING, A. M. *Computing machinery and intelligence*. In: COPELAND, B. J. (Ed.). *The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life, Plus the Secrets of Enigma*. New York: Oxford University Press, 2004.

VARELA, F. J.; THOMPSON, E.; ROSCH, E. *The embodied mind: cognitive science and human experience*. Cambridge, Mass.: The MIT Press, 1991.

WIENER, N. *Cybernetics or Control and Communication in the Animal and the Machine*. 2. ed. Cambridge, MA: MIT Press, 1961.