

COGNIÇÃO, EMOÇÃO E CONSCIÊNCIA: SEGUNDO A ABORDAGEM DE PAUL THAGARD

Leonardo Francisco Costa de Andrade¹

Resumo

O campo da Ciência Cognitiva tem apresentado ao longo das últimas décadas amplo desenvolvimento na compreensão e modelagem dos processos mentais. Fundamentados sobre dois princípios, a noção de representação e a analogia entre processos mentais e processos computacionais, a Ciência Cognitiva contemporânea desenvolveu-se visando à superação das dificuldades impostas a abordagem científica dos estados mentais. A primeira delas, a noção de representação, permitiu a formalização dos conteúdos da atividade mental dentro de uma estrutura teórica. A segunda, a modelagem computacional, proporcionou a construção de evidências científicas por meio de analogias entre dados experimentais e modelos computacionais. Entretanto, mesmo considerando todas estas vantagens possibilitadas pela abordagem algumas dificuldades ainda perduram. A maior delas, evidenciada ao longo dos anos, é a explicação ou consideração do papel desempenhado pelas emoções e pela experiência consciente. O desafio é posto da seguinte maneira: tanto nossas emoções, quanto nossas percepções conscientes aparentam desempenhar um papel fundamental em meio a vários processos cognitivos, entretanto, não são levadas em consideração pela abordagem. Desse modo, o presente trabalho buscará apresentar um panorama geral das pesquisas sob o escopo da CRUM, a fim de expandir e suplementar os termos teóricos computacionais, acomodando o papel desempenhado pelas emoções e experiência consciente.

Palavras-chave: Ciência Cognitiva; Cognição; Emoção; Consciência; Paul Thagard

1. Introdução

O presente artigo tem por objetivo apresentar a relação entre o debate da filosofia da mente e o desenvolvimento da ciência cognitiva. Sendo elaborada uma descrição, segundo a perspectiva do cientista cognitivo Paul Thagard, sobre a centralidade da noção de representação e computação em meio a CRUM (Compreensão Representacional-Computacional da Mente).

¹ Graduado em Psicologia pela Fundação Educacional Araçatuba e aluno de mestrado do Programa de Pós-graduação em Filosofia da Universidade Estadual Paulista – UNESP – *Campus* Marília/SP. if.andrade@unesp.br

Além de evidenciar as principais críticas direcionadas a CRUM, mais especificamente as críticas sobre o papel das emoções da consciência. Por fim, serão apresentadas as pesquisas contemporâneas que visam suplementar e expandir a CRUM, incorporando em seu referencial teórico considerações sobre as emoções e a consciência.

2. Filosofia da Mente e Ciência Cognitiva

O debate sobre a ontologia dos estados mentais, os meios epistêmicos disponíveis e sua relação com o corpo físico são temas recorrentes, na história da Filosofia e da Ciência.

O dualismo substancial, formulado no período moderno, é uma das abordagens mais debatida e criticada no período contemporâneo. O objetivo de Descartes na elaboração de sua teoria sobre a natureza do mental era a descoberta de um princípio indubitável, que o permitisse deduzir os outros princípios subsequentes. O método cartesiano constituísse na suposição de uma entidade com poderes tão vislumbrantes, que permitisse enganar todos e quaisquer dados advindos dos sentidos. Ao aplicar a *dúvida hiperbólica*², esta permitiria verificar a validade das asserções, contestando se sua validade teria como fundamento dados dubitáveis dos sentidos ou deduções racionais indubitáveis.

A conclusão cartesiana era de que não se pode duvidar do próprio pensamento, visto que duvidar é um ato do pensamento. Assim, ao constatar que sua essência era formada por uma substância pensante, sua natureza seria então imaterial, não-mecânica e distinta do corpo.

Apesar de ser uma influente teoria no período moderno, ao longo dos séculos, inúmeras críticas surgiram em relação ao dualismo substancial, possibilitando novas formas de abordar a questão. O foco das críticas à abordagem residiu sobre o problema da causalção mental. Argumentando que se a mente é uma substância imaterial, de que modo essa poderia influir causalmente sobre uma substância material como o corpo na determinação de sua ação.

A forte crítica sobre a relação causal entre mente e corpo no século XX permitiu que o fenômeno mental fosse abordado de uma perspectiva naturalista, possibilitando o levante das mais diversas abordagens fisicalistas sobre a relação mente e corpo.

² A dúvida hiperbólica é um método epistêmico proposto por Descartes com a finalidade de deduzir os princípios da primeira filosofia. Constitui-se como o método de por em dúvida tudo aquilo que se apresenta a razão apenas como verossímil.

Os fatores que permitiram a emergência do paradigma naturalista no estudo da mente foram: a busca pelo distanciamento de uma perspectiva dualista substancial, influenciada por fundamentos teológicos; a adoção de uma postura científica pela Psicologia no final do século XIX; o surgimento dos primeiros computadores baseados sobre o modelo da máquina de Turing, servindo tanto como metáforas quanto modelos epistêmicos e; os desenvolvimentos precusores no campo da neurociência, evidenciando o cérebro como o órgão da mente.

A centralidade de um paradigma naturalista também é o contexto de emergência da disciplina denominada de Ciência Cognitiva, sendo esta um conjunto complexo de interação interdisciplinar. Segundo Wolff (2012, p. 124), a interdisciplinaridade, central em meio a Ciências Cognitivas, permite um agrupamento de disciplinas no estudo da mente em que:

O círculo é fechado: para pensar a mente humana, devemos pensá-la a partir do modelo da máquina, que pode, por sua vez, ser pensada a partir do modelo do cérebro. Feito isso, todas as fronteiras são transpostas – e em ambos os sentidos: na ida, da mente ao cérebro (naturalização da intencionalidade), do homem ao animal (via mente-cérebro), da natureza ao artificial (via o modelo computacional) e, na volta, do artificial à natureza (via o modelo conexionista), depois do animal ao homem (via teorias evolucionistas), do cérebro à mente (via localizações cerebrais), do mental ao social (via teorias epidemiológicas da transmissão das ideias) etc. (WOLFF, 2012, p.124)

Desde então, os estudos no campo da ciência cognitiva passaram a fomentar os debates em suas disciplinas constituintes, sobretudo na filosofia, retomando problemas filosóficos milenares como o problema da relação mente e corpo proposto por Descartes, reformulado contemporaneamente como problema mente e cérebro.

A principal abordagem responsável pela reformulação do problema afirma que (PLACE, U. T. 1956; SMART, J.J.C. 1959), estados mentais são estados cerebrais, seu objetivo era estabelecer relações de identidade entre fenômenos mentais e a estrutura do cérebro, sendo denominado de fisicalismo redutivo ou teoria da identidade. Sua tese central se fundamenta em reduções interteóricas ocorridas em outras esferas ao longo da história da ciência, como por exemplo, as relações de identidade entre o relâmpago e o movimento de cargas elétricas. Porém, a maior dificuldade imposta a redução interteórica é que não há garantias de que as novas e antigas estruturas irão acomodar-se de modo uniforme.

Outra abordagem do problema mente e cérebro é conhecida como fisicalismo não-redutivo ou dualismo de propriedade. Segundo seus adeptos (CHALMERS, D. 1996; JACKSON, F. 1982; NAGEL, T. 1974; SEARLE, J. 1997), os fenômenos mentais, certamente, seriam causados por processos cerebrais, entretanto, existe uma classe específicas de fenômenos mentais como a consciência, a intencionalidade e os qualia que seriam resistentes a uma redução as suas bases físicas neurológicas. Segundo seus autores, uma redução interteórica seria incompatível, pois o cérebro físico apesar de desempenhar papel fundamental na emergência das propriedades mentais, pode não apresentar em sua estrutura neurofisiológica áreas específicas responsáveis pela produção destes fenômenos. Além do que sendo estas propriedades jamais vistas em outros sistemas físicos, pode não ser possível uma redução ou explicação exclusivamente em termos de conceitos físicos habituais.

Uma terceira abordagem fisicalista, que também advoga sobre a incompatibilidade de uma redução interteórica entre estados mentais e cerebrais, é conhecida como fisicalismo eliminacionista. Segundo seus adeptos (CHURCHLAND, P. 1986; FEYERABEND, P. 1975), a dúvida sobre a possibilidade de correspondência em uma redução sistemática entre as descobertas da neurociência e o arcabouço teórico da psicologia popular, não tem como razão à existência de propriedades irreduzíveis, mas ao fato de que a psicologia popular é uma concepção falsa e enganosa sobre as causas do comportamento e da cognição humana. Sendo proposta a eliminação do arcabouço conceitual da psicologia popular, em favor de conceitos neurocientíficos.

O funcionalismo é a abordagem com maior número de adeptos, além de ser a maior responsável pelos desenvolvimentos contemporâneos da ciência cognitiva. A razão de sua ampla defesa e influência são basicamente duas: 1) A emergência da noção de representação, possibilitando a formalização dos conteúdos da atividade mental dentro de uma estrutura teórica. 2) E a modelagem computacional, que proporcionou a construção de evidências científicas por meio de analogias entre dados experimentais e modelos computacionais. Estes dois fatores permitiram a abordagem funcionalista superar o paradigma *Behaviorista* na psicologia, amplamente divulgado na América do Norte, além de se apresentar como uma abordagem que fosse compatível com pressupostos dualistas e fisicalistas.

Segundo os adeptos do funcionalismo (DENNETT, D. 1999; FODOR, J. 1968; PUTNAM, H. 1960; PYLYSHYN, Z. 1984), os estados mentais devem ser caracterizados pelo papel funcional que desempenham na relação entre o *input* ambiental (entrada de dados sensoriais) e o *output* comportamental (efeito comportamental). “Este papel funcional caracteriza-se seja pela interação de um estado mental com outros que estejam presentes no organismo ou sistema” (TEIXEIRA, 1998, p.48).

A abordagem funcionalista serviu ainda como fundamento ao paradigma *cognitivista* da ciência cognitiva, segundo o qual a inteligência seria o resultado de um encadeamento adequado de representações mentais, neste sentido, pensar é o mesmo que realizar computações simbólicas, sendo essa a característica essencial que distingue os seres humanos dos outros animais. Vigorando por toda a década de 70 até entrar em declínio devido à retomada do paradigma *conexionista*. Contudo, no paradigma *conexionista* a noção de representação não é completamente destituída, o foco apenas muda da mente, compreendida como manipulação de símbolos, para o funcionamento cerebral, compreendido como um dispositivo computacional em paralelo. A principal característica do paradigma *conexionista* é que os modelos computacionais passam a apresentar maior similaridade com a realidade biológica do cérebro, realizando computações por meio de processamentos em paralelo.

Na seção seguinte apresentarei, segundo o ponto de vista do cientista cognitivo Paul Thagard, a abordagem denominada de CRUM (*Compreensão Computacional-Representacional da Mente*). Segundo a qual “o pensamento pode melhor ser entendido em termos de estruturas representacionais na mente e procedimentos computacionais que operam naquelas estruturas” (THAGARD, 1998, p.21).

3. Tipos de Representações e os Limites da Crum

Em seu livro, Thagard (1998) apresenta os conceitos de representação e computação. O conceito de representação é composto por lógica, regras, conceitos, imagens, analogias e conexões como os principais meios na elaboração de representações mentais. A computação, compreendida como o tipo de procedimento mental que opera nas representações, é igualmente diversa, dado que “tipos diferentes de representações mentais, como regras e conceitos,

promovem tipos diferentes de procedimentos mentais” (THAGARD, 1998, p.17). Ao longo das décadas de pesquisa, houve muita controvérsia sobre os méritos dos diferentes modos de representação do pensamento humano, mas uma compreensão da mente humana pode se beneficiar mais ao considerar a complementaridade desta diversidade.

Este complexo e integrado agrupamento de representações e computações é denominado pelo autor de CRUM. Segundo Thagard, um ponto importante é que sua abordagem se beneficia de uma analogia de três vias entre mente, cérebro e computadores, possibilitando o surgimento de novas ideias. “As teorias cognitivas por si próprias não são suficientemente precisas para dar origem a estas previsões quantitativas, mas um modelo e um programa podem preencher o espaço entre a teoria e a observação” (THAGARD, 1998, p.24).

Desse modo, o autor estabelece alguns critérios que devem ser avaliados nas considerações sobre as Teorias da Representação Mental. O quadro 1.1 apresenta um breve resumo dos critérios selecionados para avaliar as teorias representacionais da mente.

Tabela 1 – Critérios de avaliação de teorias de representação mental

QUADRO 1.1

Critérios de Avaliação de Teorias de Representação Mental:

- | | |
|---------------------------|-------------------------------|
| 1. Poder de representação | 3. Plausibilidade psicológica |
| 2. Poder computacional | 4. Plausibilidade neurológica |
| a. Resolução de problemas | 5. Aplicação prática |
| I. Planejamento | a. Educação |
| II. Decisão | b. Desenho |
| III. Explicação | c. Sistemas inteligentes |
| b. Aprendizado | |
| c. Linguagem | |

Fonte: Paul Thagard (1998)

O primeiro critério de avaliação é o *poder representacional*, ou seja, a quantidade de informação passível de ser expressa por um tipo específico de representação. O segundo critério é o *poder computacional*, o qual pode ser subdividido em outros três tipos de pensamentos de alto nível: a) *resolução de problemas*, b) *decisão* e c) *linguagem*.

O terceiro critério a ser avaliado é a plausibilidade psicológica. Uma vez que o objetivo da ciência cognitiva é compreender os processos cognitivos humanos, não basta que uma teoria

apresente muito poder representacional e computacional, esta deve ainda estar em coerência com os resultados quantitativos de testes empíricos realizados na área da psicologia.

Semelhante ao critério anterior, o quarto critério de avaliação se refere à plausibilidade neurológica. Assim como a plausibilidade psicológica pressupõe a coerência entre resultados quantitativos de experimentos psicológicos, uma teoria cognitiva deve ser avaliada verificando-se o grau de coerência com resultados de experimentos neurocientíficos.

Apresentados os critérios que guiarão a avaliação das abordagens representacionais, verificaremos alguns dos principais meios utilizados no desenvolvimento de teorias da representação mental. Por questões de objetividade, o desempenho das abordagens representacionais da mente será avaliado segundo os quatro primeiros critérios: 1. Poder representacional; 2. Poder computacional; 3. Plausibilidade psicológica; e 4. Plausibilidade neurológica.

3.1.Lógica

A lógica foi crucial para o desenvolvimento de ideias básicas sobre representação e computação, principalmente no campo da inteligência artificial. Apesar de seu poder representacional substancial é frequentemente argumentado, que como forma de compreensão dos processos cognitivos humanos a lógica apresenta pouca plausibilidade psicológica.

A maior parte de seu poder representacional advém da capacidade de elaboração de inferências dedutivas a partir de fórmulas e conectivos, evidenciando a relação entre as premissas e a conclusão. Entretanto, apesar da utilidade na constituição de raciocínios dedutivos, indutivos e abdutivos, a tradução da linguagem natural para o formalismo lógico é precária ao lidar com premissas que considerem o tempo ou na formulação de questões.

O poder computacional da lógica é avaliado segundo a capacidade de aplicação de regras de inferência a um conjunto de premissas. Para que sejam produzidas computações as regras e premissas devem ser aplicadas por uma pessoa humana ou máquina de modo adequado. Assim, apesar da possibilidade de aplicação lógica na resolução de problemas de planejamento este processo tende a se tornar lento com um número elevado de premissas. Outra questão, é

que o planejamento puramente dedutivo tende a ser *monótono*³, gerando apenas novas conclusões sem rejeitar as anteriores. Outra dificuldade enfrentada pela lógica diz respeito à incapacidade de se aprender com a experiência, sendo necessário a aplicação de um extenso processo dedutivo ao deparar-se com uma situação semelhante.

Na tomada de decisão a aplicação de deduções lógicas não tem qualquer relação com o sucesso da escolha, dado que, muitas vezes, mais de uma alternativa levará ao mesmo objetivo. Para que o processo dedutivo seja eficaz, este deve ser aplicado para verificar a probabilidade de ocorrência das consequências de uma escolha. Contudo, o uso da probabilidade em sistemas computacionais pode ser explosivo de um ponto de vista computacional, uma vez que “o número de probabilidades necessárias pode aumentar exponencialmente na medida em que aumenta o número de proposições ou variáveis do modelo” (THAGARD, 1998, p. 38). Nos problemas de explanação a principal aplicação lógica ocorre na formação do *raciocínio abduativo*⁴, no qual, com base no que já é conhecido, uma hipótese é formada gerando uma explanação sobre um fenômeno.

A aplicação direta da lógica no desenvolvimento de processos educacionais é rara, contudo, as pessoas por vezes aprendem cotidianamente por meio do *raciocínio indutivo*⁵, formando uma generalização com base na observação de um determinado número de casos.

A aplicação da lógica na compreensão da linguagem natural humana é um tema controverso. Existem teóricos que defendem equivalências teóricas entre a linguagem formal e natural, entretanto, há divergências sérias sobre este tema entre linguistas e psicólogos.

Esse tipo de controversa ocorre também em relação à plausibilidade psicológica da representação lógica. Um número reduzido de psicólogos defende, por meio de evidências

³ “Uma função matemática monótona é aquela cujos valores aumentam ou diminuem continuamente sem oscilação; o raciocínio não é monótono porque não incluímos novas crenças continuamente, já que por vezes velhas crenças devem ser abandonadas.” (THAGARD, 1996, p.36).

⁴ O conceito de *Abdução* é empregado pelo filósofo Charles S. Peirce, segundo o qual “todas as ideias da ciência a ela surgem através da Abdução. A Abdução consiste em estudar os fatos e criar uma teoria para explicá-los.” (CP, 5.145).

⁵ O raciocínio indutivo é um tipo de raciocínio lógico muito comum nas ciências, o qual é empregado a fim de gerar generalizações indutivas. Para formular esse tipo de generalização, parte-se de um pequeno grupo de exemplos particulares após ser constatado as principais regularidades, gerando uma generalização que se apresenta como uma hipótese provável.

empíricas, que as pessoas no cotidiano utilizam regras lógicas como o *modus ponens*, o que os leva a concluir que a lógica formal é uma parte importante do raciocínio humano. No entanto, um número maior de experimentos evidencia que a lógica mental pode responder apropriadamente a um número limitado de raciocínios, sendo afirmado que a abordagem lógica, apesar de sua relevância, “não é o único meio possível de entender o pensamento humano” (THAGARD, 1998, p.45). Em relação à plausibilidade neurológica, não há nenhuma evidência consistente de que os neurônios utilizam representações formais ou que o cérebro opera por meio de aplicação de procedimentos lógicos.

3.2. Regras

As regras são estruturas muito semelhantes aos condicionais *se-então* da lógica, contudo, seu poder representacional e computacional é distinto. Além disso, as regras são modelos computacionais com objetivos psicológicos desde o início da ciência cognitiva.

No geral, as regras apresentam um poder representacional inferior em relação à lógica. Ao contrário da representação lógica dotada de muitos símbolos, conectivos e regras de inferência, as regras podem ser divididas em uma parte *se* (condição) e uma parte *então* (ação).

A perda de elegância das regras em relação à lógica, por um lado diminuiu a sua capacidade representacional, mas, por outro, permitiu o desenvolvimento de um poder computacional robusto e com maior plausibilidade psicológica. Interpretar os sistemas baseados em regra como generalizações grosseiras traz muitas vantagens computacionais, que auxiliam na resolução de problemas, na aprendizagem e na linguagem.

Segundo Thagard (1998, p.51) “nos sistemas baseados na lógica, a operação do pensamento é a dedução lógica, mas sob a perspectiva de sistemas baseados em regras, a operação fundamental do pensamento é a *pesquisa*”. Isso significa que em meio a uma situação problema, as regras servem para analisar, os possíveis caminhos que nos permitem sair de um estado inicial para um estado final. Entretanto, em problemas de planejamento com muitas regras potencialmente relevantes ou um campo muito amplo de possibilidades a aplicação de regras pode se tornar difícil. Nesses casos, priorizar as regras que solucionam os pontos-chave do problema é a melhor estratégia. No caso dos problemas de decisão e explanação as regras, em geral, sofrem das mesmas dificuldades enfrentadas pela lógica.

As regras também podem ser aprendidas por generalização indutiva. Contudo, além de serem formadas com base nos exemplos, as regras podem formar-se a partir de outras regras por meio de *chunking*⁶ ou especialização. As regras ainda podem descrever o aumento gradativo da aprendizagem, desde que sejam atribuídos valores numéricos representando a utilidade e plausibilidade associada a ela.

As regras também figuram em uma das principais teorias sobre o desenvolvimento inato da linguagem humana. “De acordo com Chomsky, nossa habilidade de falar e de compreender a linguagem depende de possuímos uma gramática complexa que consiste de regras que conscientemente não sabemos que temos” (THAGARD, 1998, p.56).

O autor ainda afirma que dentre todas as abordagens computacionais-representacionais descritas, as regras são as que possuem mais evidências de sua plausibilidade psicológica. Tais evidências são apresentadas por modelos computacionais baseados em regras que simulam muitos fenômenos psicológicos, como a resolução de problemas criptaritméticos e outros tipos de raciocínio de alto nível. Esses modelos operam buscando num espaço de possíveis inferências, pesquisando a que melhor se adéqua a situação, para eventualmente chegar à conclusão. Outro critério que confirma a plausibilidade psicológica é o fato de que, eles chegam às mesmas conclusões erradas que as pessoas ao realizarem as mesmas operações mentalmente. “Existem muitos exemplos de uma adaptação entre o desempenho de sistemas baseados em regras e o comportamento dos pensadores humanos” (THAGARD, 1998, p.59).

Os sistemas baseados em regras apresentam pouca plausibilidade neurológica, a única evidência é uma analogia grosseira e superficial entre regras e conexões sinápticas dos neurônios.

3.3. Conceitos

Os conceitos são o tipo de representação mental mais discutida ao longo da história da Filosofia e de outras áreas do conhecimento. “O interesse psicológico e computacional na natureza dos conceitos cresceu rapidamente na metade da década de 70, quando os

⁶ O conceito de *chunking* é empregado em alguns modelos computacionais descrevendo o processo de formação de regras por meio da composição de um conjunto de outras regras. Contudo, além dos modelos computacionais há evidências de que o mesmo processo ocorre no pensamento humano, assim como os modelos, com o aumento de performance em uma tarefa devido a experiência o aprendido e a formação de *chunks* tende a diminuir.

pesquisadores introduziram termos como “estrutura”, “esquema” e “manuscrito” para descrever [...] a natureza dos conceitos” (THAGARD, 1998, p.64). Na década de 80, a análise computacional dos conceitos adquiriu um caráter distinto devido ao desenvolvimento de modelos conexionistas locais.

A capacidade representacional dos conceitos sofreu modificações ao longo do tempo. A defesa de que os conceitos devem apresentar definições estritas perdeu força, uma vez que conceitos como justiça, liberdade e consciência, são muito difíceis de serem definidos precisamente. Seu poder representacional foi ampliado ao serem concebidos como representações de entidades em situações típicas. Thagard (1998) propõe um modo estrutural de conceber os conceitos, considerando-os como um conjunto de espaços em branco, podendo ser preenchidos com informações sobre o que se espera que seja tipicamente mantido.

Assim concebidos, os conceitos evidenciam relações hierárquicas em rede, permitindo um modo de organização que raramente são encontrados em sistemas baseados em regras.

A representação conceitual, isoladamente, não tem capacidade de figurar como uma teoria completa da representação mental, mas apresenta características computacionais relevantes quando adicionada aos sistemas baseados em regras, refinando a modelagem do pensamento e ampliando seu poder computacional.

Os conceitos podem aprimorar o poder computacional do sistema pela recuperação de informações na memória por processos como *herança* e *ativação difusa*. Nesses processos, um sistema ao reconhecer um conceito pode recuperar da memória a estrutura hierárquica mantida com outros semelhantes. Outro tipo de recuperação de informação é a evocação de conceitos de modo difuso, nestes casos a estrutura evocada não é necessariamente organizada de modo hierárquico, mas os conceitos permanecem unidos na mesma rede devido às semelhanças de suas características ou situações típicas que representam.

Em problemas de planejamento a aplicação de conceitos é um meio rápido de resolver uma situação com base nas semelhanças. A eficácia desse processo pode ter um potencial desastroso em situações novas ou em semelhanças parcas entre a situação e a sua representação conceitual. Na tomada de decisão a recuperação de informação conceitual pode apresentar vantagens e déficits muito semelhantes ao planejamento. “A aplicação de conceitos é um meio

fácil e rápido de se tomar uma decisão, mas ela nem sempre leva em consideração as preocupações complexas sobre atitudes e objetivos que fazem parte de uma tomada de decisão mais pensada” (THAGARD, 1998, p.69). Na explanação, os conceitos figuram aplicando um esquema que inclui um objetivo a um tipo de padrão que fornece a explicação.

Os conceitos podem ser inatos, formados a partir de exemplos e a partir de outros conceitos. Um exemplo de conceito inato seria o conceito de *objeto*. Evidências empíricas apontam que crianças muito pequenas têm grandes expectativas em relação ao comportamento de muitos objetos. Os conceitos podem ser aprendidos de forma laboriosa por generalização indutiva. A observação de muitos exemplos também modifica a estrutura conceitual na medida em que exemplos adicionais são encontrados. A aprendizagem de conceitos também pode ocorrer por processo combinatório entre conceitos já conhecidos.

Os conceitos podem assumir diversas formas de representação, como os gestos da língua de sinais. Para se apreender uma língua, além de regras, uma ampla estrutura conceitual é crucial. Sendo essa a alegação que sustenta a distinção entre sintaxe e semântica.

Muitas evidências psicológicas atestam, por meio de experimentos diretos, a natureza prototípica dos conceitos. “Os achados sobre os protótipos ajustam-se bem à visão computacional de conceitos como sendo semelhantes à estruturas que listam propriedades típicas” (THAGARD, 1998, p. 74). Contudo, outros estudos sustentam que além dessa relação, estes também têm relação com a capacidade de comparar novos e velhos exemplos, bem como com explanações causais. “Logo, os conceitos podem estar intimamente ligados a regras e exemplos, assim como a padrões típicos” (THAGARD, 1998, p.75).

Poucas evidências relacionam a aplicação de conceitos com a plausibilidade neurológica, a mais relevante diz respeito à semelhança entre a difusão de conceitos em uma rede conceitual e a ativação de neurônios através de impulsos eletroquímicos.

3.4. Analogias

A relevância do pensamento analógico foi reconhecida há muito tempo pelos filósofos, mas a investigação psicológica e a modelagem computacional com analogias são recentes.

Como os conceitos, situações análogas tendem a reunir um pacote de informações relevantes, no entanto, conceitos são mais como generalizações indutivas, ao passo que

situações análogas representam poucos exemplos particulares. A aplicação do pensamento analógico exige que o análogo *alvo* (a nova situação) seja comparado ao análogo *fonte* (a situação antiga). “Cada análogo é uma representação de uma situação, e a analogia é um relacionamento sistemático entre eles” (THAGARD, 1998, p.80). Uma analogia bem sucedida deve apresentar padrões de semelhanças. Entretanto, devem ser diferenciadas semelhanças superficiais de semelhanças relevantes. Semelhanças superficiais, geralmente, são aquelas que pouco auxiliam na resolução do problema, enquanto as relevantes são aquelas que evidenciam uma comparação positiva entre as relações causais do análogo *fonte* e do análogo *alvo*.

O poder computacional das analogias é maior em situações em que há alguma experiência prévia, mas pouco conhecimento geral sobre ela. As analogias são evidentemente poderosas em casos onde não existem regras disponíveis. A aplicação computacional de analogias deve seguir os seguintes passos: 1) identificação do análogo *alvo*; 2) recuperação na memória de um análogo *fonte*; 3) analogia entre análogo *alvo* e *fonte*; e 4) Adaptação da solução da *fonte* para o *alvo*. Em termos computacionais e psicológicos o maior desafio da aplicação de analogias reside na recuperação de um análogo *fonte* compatível, sendo necessária a satisfação de três restrições: similaridade, estrutura e objetivo.

Em situações de planejamento genuinamente novas o análogo *fonte* pode não satisfazer as condições necessárias, tornando mal-sucedida a analogia. Nos processos de tomada de decisão as analogias são frequentemente utilizadas, sobretudo, em decisões judiciais. “O raciocínio legal frequentemente faz referência a casos prévios que servem como precedentes: estes são análogos-fonte que são mapeados para o caso-alvo atual” (THAGARD, 1998, p.85). Nesses casos, as analogias contribuem sugerindo soluções bem-sucedidas e lembrando desastres prévios. As analogias ainda possuem um papel importante na explanação como na analogia entre mente, cérebro e computador da ciência cognitiva.

O pensamento analógico influi no processo de aprendizado de três maneiras: 1) por armazenamento de casos baseados em situações anteriores; 2) aprendizado como resultado direto de uma analogia; e 3) aprendizado por meio de *esquemas analógicos*. O primeiro modo de aprendizado é o mais comum e simples, sendo apenas um prelúdio necessário ao pensamento analógico. O segundo modo se refere a situações em que uma solução é encontrada pela

analogia entre um análogo *fonte* e um análogo *alvo*. O terceiro modo é o resultado de uma generalização, ao comparar vários análogos com soluções semelhantes, é possível a construção de um *esquema analógico* entre eles e aplicá-lo ao análogo *alvo*. As analogias têm um importante papel na produção e compreensão de figuras de linguagem como metáforas. “Tanto a criação de uma metáfora por quem fala quanto a sua compreensão por quem escuta exige a percepção de uma analogia subjacente” (THAGARD, 1998, p.87).

Experimentos psicológicos confirmam o aumento do desempenho na resolução de problema ao serem apresentadas situações com soluções semelhantes. Sendo, ainda confirmado, que o desempenho tende a ser maior quando mais de uma situação com solução semelhantes é apresentada, permitindo a produção de um *esquema analógico*. Experimentos psicológicos no campo da linguagem utilizando metáforas, afirmam que o processo de compreensão literal e metafórico, são processos interativos na compreensão da linguagem. Contudo, não há qualquer evidência sobre as bases neurológicas do pensamento analógico.

3.5. Imagens

A representação mental por imagens, ou melhor, por meio da *imaginação*⁷ foi considerada parte essencial do pensamento humano pelos filósofos do século XVII. Além de figurar como um dos primeiros temas de investigação e de testes empíricos na Psicologia, ao menos, até sua relevância ser mitigada pela corrente behaviorista do século XX. Após a década de 60, com o retorno da psicologia cognitiva, este tema tornou-se novamente um objeto adequado na investigação da cognição humana. A investigação da representação mental por meio da imaginação deve levar em conta além das representações visuais, representações táteis, olfativas, gustativas e até mesmo emocionais.

Apesar do crescente interesse em representações imagéticas na ciência cognitiva, a modelagem computacional ainda apresenta limitações. A razão disso é que a visão robótica é muito rudimentar em comparação com a visão humana. Capturar imagens e vídeos em dispositivos eletrônicos é uma tarefa simples atualmente, entretanto, extrair informação de

⁷ O conceito de imaginação foi empregado a fim de ampliar a noção de representação por imagens para além das representações visuais, sendo igualmente consideradas como representações por imagens as representações gustativas, olfativas, táteis, motoras e até mesmo emocionais.

centenas de milhões de pontos é o verdadeiro desafio enfrentado pelos cientistas da computação contemporâneos, já que a imagem capturada pode ser altamente ambígua.

A representação pictórica dever ser somada a outros tipos de representações para aumentar seu poder representacional. A observação de uma imagem não requer inferências sobre a aparência ou o arranjo espacial dos objetos, dado que são características facilmente percebidas. “As imagens visuais complementam, mas não substituem representações verbais do tipo que vimos nos capítulos anteriores” (THAGARD, 1998, p. 98). Além disso, as representações imagéticas não necessitam que o estímulo esteja presente, mas podem recuperar informações sensoriais da memória reconstituindo o estímulo percebido.

As representações verbais e não-verbais são construtos importantes na resolução de diversas tarefas. Entretanto, o poder computacional das imagens pode ser substancialmente maior em situações que a solução dependa da aparência visual ou da relação espacial entre os objetos. A representação de imagens permite a inspeção, rotação e transformação de objetos e situações representadas mentalmente. A resolução de problemas de planejamento pode ser facilitada pelas representações pictóricas, utilizando-se de representações visuais dos passos a serem cumpridos, semelhante ao processo de resolução de problemas baseado em regras. As transformações visuais também podem auxiliar na resolução de problemas de construção, assim como diagramas servem como auxílio externo na resolução de problemas. Nos problemas de decisão imagens visuais e emocionais tem um papel importante na tomada de decisão. Problemas de explanação também desfrutam dos benefícios da representação imagética, um exemplo, é a hipótese da mudança continental elaborada por Alfred Wegener.

As representações pictóricas são cruciais na aprendizagem. Evidências empíricas apontam que a prática de imaginação mental pode aumentar o desempenho em situações que ainda não ocorreram. As imagens ainda podem auxiliar na generalização e categorização de conceitos, ignorando informações incidentais em favor das propriedades gerais sobre os objetos. Além de figurarem como uma parte importante em meio o aprendizado abduutivo. Estudos com arqueologistas apontam para a importância da abdução visual na explicação das funções de objetos antigos. As imagens também têm um papel relevante na linguagem, principalmente no seu uso metafórico, sendo, por vezes, evocados componentes sensoriais.

A plausibilidade psicológica da tese, de que as imagens são parte importante do pensamento humano, é averiguada por testes empíricos, os quais empregam os processos de rotação e transformação de imagens mentais na resolução de problemas. Contudo, a eficácia deste teste não é razão de consenso sobre a plausibilidade psicológica da representação visual.

Embora a maioria dos pesquisadores de psicologia estejam convencidos através de experimentos como os recém descritos de que os *humanos* utilizam a imagem visual, alguns mais céticos sustentam a ideia de que o mesmo tipo de representação verbal está subjacente a todo pensamento, e que as experiências da imaginação são ilusórias. A rotação, a formação de imagens e outras transformações podem sempre ser imitadas por processos computacionais não imaginísticos em listas de palavras. Entretanto, na última década acumularam-se evidências neurológicas que apoiaram posteriormente a hipótese da imaginação. (THAGARD, 1998, p. 104, adaptado)

As evidências que atestam a plausibilidade neurológica das representações por imagem são basicamente duas. O primeiro grupo de evidências foi alcançada com a aplicação dos mesmos testes psicológicos em pacientes com dano cerebral, verificando que os mesmos déficits percebidos visualmente tendem a manter-se na evocação de imagens mentais. O segundo grupo se refere à medida da atividade cerebral de pessoas operando com imagens mentais, sendo verificado que as mesmas áreas cerebrais ativas na percepção visual também são ativadas quando as imagens mentais são evocadas.

3.6. Conexões

Modelos conexionistas ao enfatizarem conexões entre estruturas simples, são capazes de computações poderosas com um dos graus mais altos de plausibilidade neurológica, dada a inspiração para sua construção pelos neurônios naturais do cérebro humano.

Por meio de computações em paralelo, os modelos conexionistas podem apresentar representações locais que permitem identificar em suas unidades interpretações em termos de conceitos ou proposições. Além de apresentarem representações distribuídas em rede, representando conceitos ou proposições por meio de padrões de ativação em função da interação entre um complexo de neurônios. Em ambos os casos, o processo desempenhado pela rede de neurônios pode ser entendido em termos de *satisfação paralela da repressão*⁸.

⁸ A satisfação paralela da repressão é um processo em que muitas restrições são consideradas simultaneamente, gerando a melhor maneira de resolver o problema segundo os objetivos do sistema. Elaborar um calendário de disciplinas, assim como a captação de imagem por cada um de nossos olhos, são exemplos de satisfação paralela da repressão.

O poder representacional das redes conexionistas é relativamente alto em relação à simplicidade de suas representações. As redes locais especificam interpretações em cada unidades, assim a ativação de uma unidade representa a aplicabilidade de um conceito ou a veracidade de uma proposição, enquanto a inibição pode levar a interpretação oposta. As redes distribuídas são também conhecidas como *redes de alimentação avançada*. Neste tipo de rede nenhuma unidade específica contém a informação relevante que está sendo processada, mas é o padrão de ativação entre as unidades que permite a interpretação. Elas funcionam utilizando-se de assentadores de influxo contendo informações básicas sobre padrões específicos que devem ser reconhecidos, assentadores de fluxo externo que reconhece os padrões gerando um padrão de ativação ou inibição entre as unidades.

As restrições podem ser satisfeitas em paralelo pela passagem repetida da ativação entre todas as unidades, até que depois de um certo número de ciclos de atividade todas as unidades tenham alcançado níveis estáveis de ativação. Este processo é chamado de *relaxamento* [...]. Relaxar a rede significa ajustar a ativação de todas as unidades baseadas nas unidades às quais elas estão ligadas, até que todas as unidades tenham ativações altas ou baixas estáveis. (THAGARD, 1998, p. 116)

Nos problemas de planejamento as redes conexionistas figuram de modo indireto. Isso porque o planejamento é um processo sequencial, no qual a aplicação de regras e analogias funciona como as soluções mais aplicáveis. Entretanto, as redes conexionistas implementam em seu funcionamento regras e analogias. Já os problemas de decisão são um dos maiores sucessos das redes conexionistas. Nas redes conexionistas os problemas de decisão são compreendidos em termos de satisfação paralela da repressão. As repressões positivas podem representar a relação de satisfação entre objetivos e atitudes e repressões negativas representam a incompatibilidade entre duas ações ou dois objetivos, enquanto repressões externas representam os critérios que priorizam um objetivo em detrimento aos outros. De modo semelhante, os problemas de explanação podem ser solucionados aplicando-se a restrição paralela da repressão, a escolha da melhor hipótese pode ser encontrada analisando-se a ativação ou inibição entre as unidades representando as hipóteses e as evidências.

No aprendizado as redes conexionistas aprendem novos padrões de dois modos: acrescentando novas unidades ou modificando os pesos dos elos. No primeiro caso a aprendizagem ocorre acrescentando novas unidades que representem novas atitudes ou

evidências, semelhante aos problemas de decisão e explanação. No entanto, o segundo caso é o mais comum, dado que este método é mais semelhante à estrutura neuronal do cérebro.

A técnica para treinar a rede ajustando os pesos entre os elos é a *retropropagação*. O método consiste em atribuir pesos ao acaso; ativar unidades de influxo com base nos padrões desejados; disseminar a ativação entre as unidades até que as unidades de fluxo externo sejam alcançadas; calcular os erros de ativação; propagar os erros entre as unidades modificando o peso entre os elos; e repetir o procedimento até que os erros sejam minimizados.

Entretanto, há diferenças relevantes em relação a aprendizagem em termos neurológicos. O primeiro deles diz respeito aos padrões de ativação no cérebro, onde nenhum supervisor é necessário para corrigir os erros. O segundo tem relação com o tempo do processo de retropropagação em relação aos processos cerebrais, já que na retropropagação são necessários centenas ou milhões de exemplos para treinar uma única rede.

Na compreensão da linguagem as redes conexionistas figuram muitos experimentos, reconhecidos como problemas de satisfação paralela da repressão. Em um dos experimentos as unidades interconectadas representavam hipóteses sobre as letras e palavras presentes, até alcançar uma estabilidade que permite gerar a melhor interpretação. Outro experimento realizado a fim de que uma rede conexionista determine o sentido exato de uma palavra ambígua, busca verificar as relações estabelecidas entre a palavra e outros elementos textuais, sendo que a melhor interpretação é alcançada em relação ao padrão de ativação das unidades.

Os dois experimentos explicam como os seres humanos compreendem ambiguidades, atestando sua plausibilidade psicológica. Além disso, “os modelos conexionistas [...] não foram apenas utilizados para estimular novos resultados de experimentos psicológicos anteriores, mas também sugeriram novos” (THAGARD, 1998, p.122).

Apesar da superior plausibilidade neurológica das redes conexionistas em relação a seus concorrentes, distinções importantes são verificadas em relação às redes neuronais reais. As redes conexionistas são parcialmente semelhantes em estrutura com referências as redes de neurônios no cérebro, no sentido de que ambas são compostas de elementos simples que excitam e inibem uns aos outros. Entretanto, as redes neuronais do cérebro são compostas de trilhões de conexões, além do que seus neurônios são muito mais complexos. A melhor analogia

para minimizar tais diferenças, seria compreender as unidades artificiais como complexos de neurônios. Conceber essa analogia permite mitigar outra diferença importante, uma vez que “muitas redes locais usam elos simétricos entre unidades, enquanto sinapses conectando neurônios são de uma via. Mas grupos neuronais frequentemente têm caminhos neurais que lhes permitem influenciar uns aos outros” (THAGARD, 1998, p. 123).

3.7. Os limites da CRUM

Apesar do amplo desenvolvimento da ciência cognitiva, tendo a CRUM como um dos principais meios para a compreensão da cognição humana. Questões fundamentais sobre a natureza da vida mental ainda carecem de uma explanação, como o papel das emoções e da consciência. O campo da Filosofia da mente é composto de propostas competitivas, elaboradas por materialistas e dualistas sobre o papel das emoções e da consciência. Desse modo, considerando relevância das alegações sobre o papel das emoções e da consciência, bem como os desenvolvimentos da CRUM, é que o autor propõe *o materialismo integrativo*. “Argumento que a consideração das emoções e da consciência apóiam outra posição – o *materialismo integrativo* – que defende uma integração teórica da computação, da neurobiologia e da experiência consciente” (THAGARD, 1998, p. 142).

A proposta, elaborada no fim da década de 90, de consolidação dessa abordagem tem como fundamento o processo de expansão e suplementação da CRUM. Visando a expansão dos meios de representação apresentados, para então, ser suplementados com considerações sobre a natureza biológica destes fenômenos. “Consequentemente, a experiência humana das emoções parece estar intimamente ligada aos tipos de corpos que temos, e não apenas em nossos processos abstratos de pensamento” (THAGARD, 1998, p.146). Bem como, “a não ser que queiramos abandonar a CRUM [...], necessitamos suplementar a CRUM com considerações biológicas. [...] Para propósitos biológicos pode ser necessário focar aspectos particulares das sensações para entender o papel da consciência” (THAGARD, 1998, p.150)

No entanto, o desafio da emoção e da consciência não são os únicos desafios impostos a consolidação da CRUM, outros pontos deficitários como a natureza das representações, o papel significativo dos ambientes físicos, a pouca ênfase na interação social também figuram

em meio às críticas. Contudo, visando alcançar os objetivos propostos, na seção seguinte verificaremos como pesquisas recentes respondem aos desafios das emoções e da consciência.

4. Suplementando e Expandindo a Crum

Após 24 anos desde a publicação de seu livro analisaremos, ao longo desta seção, as modificações ocorridas na CRUM para suplementar e expandir seus termos, e adequar-se as críticas sobre o papel das emoções e da consciência na explanação do pensamento humano.

4.1. Suplementando a CRUM

Em 1996 Thagard afirmou que a principal estratégia aplicada para suplementar à base teórica da CRUM, seria a consideração do aspecto biológico que compõe o sistema nervoso. Em 2002, Thagard publicou um artigo apresentando de que modo a composição bioquímica do cérebro implica na compreensão das emoções, cognição e da consciência.

Segundo o autor, a maioria dos modelos computacionais, mesmo aqueles baseados em redes neurais artificiais, ignoram o papel das moléculas ao tratarem o processo neuronal apenas como um fenômeno elétrico. “A negligência da neuroquímica nas ciências cognitivas seria apropriada se as propriedades computacionais do cérebro relevantes para explicar o funcionamento mental fossem de fato elétricas em vez de químicas” (THAGARD, 2002, p. 430). No entanto, evidências apontam o contrário, dado que o aspecto bioquímico do cérebro diferencia-se qualitativamente dos modelos conexionistas baseados em disparos elétricos.

Alguns modelos conexionistas atuais, além da taxa de ativação permitem a verificação de padrões de ativação ampliando sua capacidade computacional. Entretanto, mesmo modelos sofisticados permanecem muito aquém da vantagem computacional dos processos bioquímicos. A relevância neurológica dos modelos é que, assim como as células individuais, ambos os processos podem ser compreendidos como computações realizadas por meio de *inputs* e *outputs*, sejam eles químicos ou elétricos.

Podemos pensar em células individuais, sejam neurônios ou não, como computadores que possuem entradas na forma de moléculas que se ligam a proteínas receptoras, saídas na forma de moléculas emitidas pelas células e processos internos realizados por reações químicas envolvendo proteínas. [...] A computação molecular dentro da célula é massivamente paralela, em que muitos receptores podem iniciar simultaneamente muitas reações químicas que ocorrem simultaneamente em bilhões, ou mais, de proteínas na célula. (THAGARD, 2002, p.433)

Os críticos da CRUM afirmariam ser um exagero a descrição do comportamento de proteínas e células segundo uma metáfora com a noção de computação, ignorando a precisão matemática necessária a descrição de Turing. Contudo, tal descrição está muito além de uma simples metáfora. Resultados experimentais e matemáticos comprovam que o processamento molecular é computacional no sentido mais rigoroso. “Assim, a descrição de células e proteínas como realizando cálculos é mais do que metafórica e, portanto, é potencialmente relevante para a compreensão da computação mental” (THAGARD, 2002, p. 433).

Todavia, mesmo que a noção de computação possa ser atribuída a células e proteínas, não significa que outros órgãos possam apresentar propriedades mentais. Essas são características apenas de um único órgão, o cérebro, o qual é formado por bilhões de células específicas com interações entre si que possibilitam a presença de propriedades mentais.

Por fim, são levados em conta três propriedades de uma rede neural seja ela artificial ou natural: o processamento interno dos neurônios; a topografia da rede; e o efeito temporal.

O processamento interno é avaliado segundo a capacidade dos neurônios de processar muitos *inputs* e a complexidade dos *outputs* possíveis. Sob este critério, modelos conexionistas mais sofisticados apresentam uma boa avaliação do seu processamento interno, respondendo de maneiras diferentes aos vários padrões que chegam a eles e produzindo diferentes padrões de resposta. No entanto, o poder computacional de células neuronais é ainda maior. Uma célula pode sinalizar a si mesma por um processo conhecido como *sinalização autócrina*, ou sinalizar uma célula adjacente pelo processo de *sinalização parácrina*, assim como, sinalizar células-alvo muito distantes pela produção de hormônios, um processo conhecido como *sinalização endócrina*, além das sinalizações específicas do cérebro conhecidas como *sinalização sináptica* e *neuroendócrina*.

A topografia da rede refere-se ao padrão de conectividade que permite um neurônio afetar o disparo de outro. Em redes neurais artificiais a topografia da rede é comumente verificada pelo padrão de conectividade estabelecido pela excitação ou inibição dos neurônios. Já as células neuronais podem estabelecer inúmeras vias de conectividade, assim, vias diferentes apresentam funções distintas no sistema. Por exemplo, a integração do movimento pela dopamina e regulação da emoção pela serotonina são processos que quando apresentam

defasagem podem causar doença de Parkinson e depressão. Além disso, o padrão de conectividade do cérebro pode ser modificado por neuromoduladores, que são produzidos por outras células ampliando o poder computacional por meio de *retropropagação*⁹.

O efeito temporal tem relação com a evolução dinâmica da rede ao longo do tempo. As redes artificiais têm uma avaliação de desempenho bem inferior em relação às redes naturais. Nas redes artificiais é muito comum um tipo de ativação síncrona, em que todos os neurônios apresentam atualização de ativação ao mesmo tempo. Já nas redes naturais a ativação tende a ser assíncrona, “os neurônios reais são assíncronos e dependem da história temporal na forma dos padrões de pico que são inseridos neles” (THAGARD, 2002, p. 435). Este é um dos principais fatores na garantia da dinamicidade das redes neurais do cérebro.

4.1.1. O papel das moléculas na cognição, emoção e consciência

Segundo Thagard (2000; 2002; 2008; 2011), as evidências sobre a relevância das moléculas na regulação das emoções é o aspecto central para compreender a influência da composição neurobiológica no processo cognitivo e na emergência da *consciência afetiva*¹⁰.

Ao considerar a neurobiologia da cognição, um grupo muito pequeno de evidências estabelece relação direta entre a presença de neuroreguladores e processo cognitivos de alto nível. Entretanto, há muitas evidências de que os neuroreguladores influenciam as emoções diretamente, bem como de que os processos cognitivos são modulados pelas emoções.

Alguns exemplos que sustentam a primeira afirmativa são a presença de altos níveis de dopamina associada a estados de emocionalidade positiva, como na relação entre altos níveis de oxitocina e a ocorrência de comportamentos sociais como afetividade materna, sentimentos de aceitação e gratificação sexual. Evidências adicionais são encontradas pela administração de medicamentos psicotrópicos regulando estados de humor.

⁹ Termo traduzido do inglês: *Feedback loops*

¹⁰ A consciência afetiva (do inglês *emotional consciousness*) é o fenômeno pelo qual as pessoas experimentam diferentes tipos de emoções, de modo integrado a outras funções cognitivas, como a percepção, julgamento, memória e inferência. Avaliando diferentes graus de intensidade e valência. Assim como, a dinamicidade do processo, provocando mudanças contínuas de humor devido a estímulos diversos.

As evidências sobre a intrínseca relação entre emoções e processos cognitivos, são advindas da psicologia e da neurociência. A intrínseca relação entre neuroquímica, cognição e emoção pode ser observada no aprimoramento da capacidade de resolução de problemas proporcionado pela cafeína, ao bloquear o neurotransmissor inibitório adenosina. Assim como, a ingestão de álcool perturba o funcionamento mental ao inibir os receptores de glutamato e NMDA, importantes no processo de aprendizado. Bem como, há evidências de que pessoas com dano cerebral no córtex ventromedial, responsável pela conexão com áreas que modulam as emoções, apresentam ineficácia na tomada de decisão devido à dificuldade de produzir marcadores somáticos de situações vivenciadas anteriormente.

Mencionei apenas uma pequena parte das evidências que desafiam a divisão psicológica tradicional entre cognição e emoção e a antiga distinção filosófica entre razão e paixão. Mas é suficiente para o propósito em questão, mostrar que a relevância demonstrável da neuroquímica para as emoções é igualmente direcionada a cognição em geral. Se a cognição humana é computação mental, é um tipo de computação determinada pelos aspectos químicos e também elétricos do cérebro. (THAGARD, 2002, p. 442)

Em relação à *consciência afetiva* afirma-se que a maior parte dos sentimentos conscientes como felicidade, tristeza, medo, raiva e nojo são emergentes da atividade cerebral, mas sabe-se pouco sobre seus mecanismos subjacentes. Contudo, “as diversas maneiras pelas quais os neuroquímicos influenciam a emoção sugerem que é improvável que a consciência emocional surja apenas das atividades elétricas do cérebro” (THAGARD, 2002, p. 440).

Em 2008, o autor define *consciência afetiva* como um produto emergente da interação entre áreas coordenadas pela *memória de trabalho*¹¹. Seguindo a definição de consciência de Damásio, como um “sentido interno”, envolvendo vigília, atenção e emoção, propõe-se a subdivisão entre *consciência central* e *consciência ampliada*. “A consciência central requer apenas uma imagem, que é um padrão mental em qualquer uma das modalidades sensoriais, e um objeto como uma pessoa ou outra entidade. [...] A consciência ampliada requer memória

¹¹ “A memória de trabalho envolve o armazenamento de curto prazo de diferentes tipos de informação em diferentes áreas do cérebro e processos executivos de atenção seletiva e gerenciamento de tarefas que envolvem o cíngulo anterior e o córtex pré-frontal dorsolateral” (THAGARD e AUBIE, 2008, p. 816).

que possibilite um *self* autobiográfico” (THAGARD e AUBIE, 2008, p. 813). Propondo-se uma abordagem neurocomputacional da consciência afetiva.

No entanto, os detalhes computacionais fogem ao propósito do artigo, sendo relevante apenas os componentes da abordagem. “Uma teoria neurocomputacional da consciência emocional deve, portanto, ter pelo menos os seguintes componentes: representação, percepção somática, avaliação cognitiva e memória de trabalho” (THAGARD e AUBIE, 2008, p. 814).

O componente da representação exige que uma população de neurônios tenha poder representacional para representar o mundo, estados corporais e suas próprias representações. A representação do mundo ocorre pela captação de estímulos externos, produzindo nos neurônios padrões de disparos que mantém uma dependência estatística causal com os objetos do mundo. De modo similar, os neurônios representam estados corporais por meio de padrões de disparos, mas, neste caso, os estímulos são captados por sensores internos. Para representar suas próprias representações, uma população de neurônios estabelece com outra população de neurônios ciclos de retropropagação com influência causal que ocorre em ambas às direções.

A percepção somática relaciona-se com a regulação emocional pelos neuroreguladores, mas, sobretudo, é influenciada principalmente pela interação entre áreas do córtex e áreas destinadas a regulação das emoções como a amígdala, o hipocampo, núcleo accumbens e outras.

A avaliação cognitiva, parte importante na constituição das emoções é realizada por representações distribuídas, nas quais elementos únicos são representados pela atividade de vários neurônios, e pela participação de neurônios individuais em várias representações.

A memória de trabalho figura como um componente importante na explicação da consciência afetiva, dado seu caráter serial muito distinto dos processos paralelos e assíncronos da atividade cerebral. A memória de trabalho realiza um armazenamento de curto prazo de informações distribuídas em várias áreas do cérebro e também é um importante componente dos processos executivos de atenção seletiva e gerenciamento de tarefas.

Verificaremos agora como estas considerações sobre o aspecto neurobiológico das emoções e da consciência são integrados na constituição de uma teoria neurocomputacional das emoções e da consciência afetiva, possibilitando a expansão da CRUM.

4.2. Expandido a CRUM

A relação entre emoção e cognição nos permite vislumbrar como a CRUM pode ser expandida incluindo as emoções na modelagem computacional. Em 1996, já era suposto que as emoções desempenhariam a seguinte função na resolução de problemas:

As emoções oferecem uma *avaliação* resumida de sua situação de resolução de problemas que fazem duas contribuições importantes para o pensamento subsequente. A avaliação de que certos aspectos de sua situação são extremamente importantes para seus objetivos pode levá-los a *enfocar* tais aspectos, concentrando seus limitados recursos cognitivos no que interessa. Além disso, as emoções fornecem agilidade para a *ação*, garantindo que você será estimulado a lidar com sua situação de resolução de problemas para não ficar perdido em pensamento. (THAGARD, 1996, p. 144)

Assim, uma estratégia possível, seria a implementação de nódulos adicionais representando esse tipo de avaliação de estado geral. Obviamente, isso não permite que o modelo experiencie emoções. Pois, assim como um mapa, um modelo corresponde à realidade, apresentando nível de detalhamento servindo a diferentes propósitos. Mesmo não garantindo emoções, tal implementação permite averiguar o papel das emoções na decisão.

Para o propósito em questão, Thagard (2000) propõe sua *Teoria da coerência emocional*. Segundo sua teoria, os elementos de um processo de tomada de decisão são representações de ações e metas concorrentes. Uma ação ao facilitar um objetivo, estabelece uma relação de coerência e um tipo de restrição positiva. Quando duas ações são incompatíveis na satisfação de um objetivo, elas estabelecem uma relação de incoerência havendo, portanto, uma restrição negativa. Assim, as emoções figuram como *valências* atribuídas a cada uma das metas, priorizando os objetivos mais relevantes.

A teoria em questão tem a vantagem de poder ser replicada computacionalmente por redes artificiais. Em 2005, o autor exemplifica apresentando uma aplicação computacional do modelo. O HOTCO é um modelo computacional de inferência emocional, podendo ser aplicado a vários tipos de pensamento como confiança, raciocínio jurídico e analogias emocionais. Neste modelo, os elementos são representados por unidades de neurônios artificiais e a ativação é interpretada como a medida do grau de aceitabilidade do sistema. Cada unidade é dotada de uma valência, positiva ou negativa, do apelo emocional representado pela unidade. O modelo apesar de ser psicologicamente plausível, não apresenta o mesmo grau de plausibilidade

neuroológica. Contudo, o HOTCO é compatível com modelos mais detalhados como o GAGE, representando a decisão por disparos neuronais organizados em grupos anatómicos, incluindo as principais *áreas de processamento da emoção*¹².

Assim, verificamos como a CRUM pode ser expandida, construindo modelos que considerem o papel das emoções. Mas e a consciência afetiva, seria possível abordá-la da mesma forma? Certamente, um grau de complexidade muito maior é necessário. O modelo deve ser construído considerando todos os componentes descritos anteriormente.

Em 2011, Thagard e Finn propõe abordar a consciência como um tipo de intuição moral. Obviamente, existem outros tipos de experiências conscientes que não apresentam qualquer tipo de julgamento ou intuição moral. A justificativa é dada da seguinte maneira:

Da perspectiva da percepção somática, parece não haver nada de especial nos processos cerebrais envolvidos na intuição moral em comparação com a consciência afetiva em geral. [...] Consequentemente, a diferença crucial deve envolver a avaliação cognitiva, com intuições morais ativando um conjunto mais estreito de objetivos do que a consciência emocional geral. (THAGARD e FINN, 2011, p. 153)

Assim, abordar a consciência como um tipo de intuição moral é mais coerente com as definições sobre consciência elaboradas por filósofos e psicólogos, bem como com o modelo EMOCON proposto em 2008. *O modelo EMOCON*¹³ é um modelo hipotético, conjecturado com base nas partes de outros modelos já em funcionamento como o GAGE, ANDREA, NECO e mais algumas considerações de Damasio (2000) e Morris (2002) sobre as entradas sensoriais. “O modelo EMOCON mostra como combinar percepção somática e avaliação cognitiva em um único sistema que transcende o conflito secular entre as teorias fisiológicas e cognitivas das emoções” (THAGARD e AUBIE, 2008, p. 817).

Ao longo desta seção verificamos como a CRUM pode ser suplementada com compreensões advindas da neurobiologia sobre as emoções e a consciência. Assim como, a expansão de seus referenciais teóricos abarcando tais considerações, e permitindo a consolidação da CRUM como um abordagem mais robusta, flexível e integrada.

¹² O modelo GAGE é composto neurônios artificiais que representam as seguintes áreas responsáveis pelo processamento das emoções: córtex pré-frontal, hipocampo, amígdala e núcleo accumbens.

¹³ A complexidade do modelo extrapola os objetivos deste artigo, podendo ser verificado na íntegra em (Thagard e Aubie, 2008).

5. Considerações Finais

Neste artigo, podemos notar como Paul Thagard apresenta de maneira didática, bem informada e integrada as distinções e similaridades relevantes sobre os diversos tipos de representações, analogias entre sistemas artificiais e naturais e considerações sobre as diversas concepções filosóficas sobre a natureza da mente. Certamente, a CRUM se apresenta como uma abordagem plausível na compreensão da mente ao demonstrar bom desempenho na explicação e compreensão dos fenômenos mentais. Todavia, um aspecto que talvez tenha sido dado pouca ênfase diz respeito à natureza das representações, crítica frequentemente levantada por autores sob um ponto de vista anti representacionista.

Entretanto, considerando-se a ampla gama de explicações dos fenômenos mentais, o bom desempenho de sua avaliação e a integração de perspectivas de diversas áreas do conhecimento, minimizam algumas lacunas explicativas. Além da relevância dos esforços contemporâneos na suplementação e expansão de construtos teóricos e modelos computacionais. Leva-nos a conclusão de que a analogia entre mente, cérebro e máquina da CRUM dificilmente será abandonada por completo, assim como a noção de representação e computação, mas ao longo dos anos, pode ser desenvolvida abarcando considerações mais amplas e precisas sobre o funcionamento mental.

Referências

Paul M. **Matéria e consciência**: Uma introdução contemporânea à Filosofia da mente. São Paulo: Editora UNESP, 2004. 286 p. ISBN 85-7139-519-5.

_____. Paul M. Some Reductive Strategies in Cognitive Neurobiology. *In*: SILVERS, Stuart. **Representation**: Readings in the Philosophy of Mental Representation. [S. l.]: Philosophical Studies Series, 1989. v. 40, cap. 12, p. 279-309. ISBN 978-94-009-2649-3. Disponível em: https://link.springer.com/chapter/10.1007/978-94-009-2649-3_12#citeas. Acesso em: 1 jan. 2020.

DAMASIO, Antonio. **The Feeling of What Happens**: Body and Emotion in the Making of Consciousness. 1. ed. [S. l.]: Harvest Books, 2000. 400 p. ISBN 978-0156010757.

DENNETT, Daniel C. **Brainstorms**: Ensaios Filosóficos sobre a Mente e a Psicologia. São Paulo: Editora UNESP, 1999. 431 p. ISBN 85-7139-645-0.

_____. **Consciousness Explained**. New York, USA: Penguin Books, 1993. 492 p. ISBN 0-14-012867-0.

DESCARTES, René. **Discurso do método**. 3. ed. rev. e aum. São Paulo: Martins Fontes, 2007. 123 p. ISBN 978-85-336-2363-7.

_____. **Meditações Metafísicas**. 1. ed. rev. São Paulo: Martins Fontes, 2000. 155 p. ISBN 85-336-1204-4.

FEYERABEND, P. *Against the method*. London: New Left Books, 1975.

JACKSON, Frank. Epiphenomenal Qualia. **The Philosophical Quarterly**, [s. l.], v. 32, ed. 127, p. 127-136, 1982. DOI 10.2307/2960077. Disponível em: <https://www.jstor.org/stable/2960077>. Acesso em: 15 jan. 2020.

MORRIS, John S. How do you feel?. **TRENDS in Cognitive Sciences**, [s. l.], v. 6, n. 8, p. 317-319, Agosto 2002. Disponível em: https://www.google.com.br/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwjIm7KKn8TtAhWOHbkGHWf3C3QQFjAAegQIAxAC&url=https%3A%2F%2Fthebrain.mcgill.ca%2Fflash%2Fcapsules%2Fpdf_articles%2Fhow_do_you_feel.pdf&usg=AOvVaw1wp6AptXfPAVzPl6pE_e6. Acesso em: 5 out. 2020.

NAGEL, Thomas. What Is It Like to Be a Bat? **The Philosophical Review**, [s. l.], v. 83, ed. 4, p. 435-450, 1974. DOI 10.2307/2183914. Disponível em: <https://www.jstor.org/stable/2183914> Acesso em: 3 jan. 2020.

PLACE, Ullin T. Is consciousness a brain process?. **The Mind-Brain Identity Theory**, London, p. 44-50, 1970. DOI https://doi.org/10.1007/978-1-349-15364-0_2. Disponível em: https://link.springer.com/chapter/10.1007/978-1-349-15364-0_2#citeas Acesso em: 10 jan. 2020.

PYLYSHYN, Zenon Walter. **Computation and Cognition: Toward a Foundation for Cognitive Science**. Mass., The [MIT Press](#), 1984.

PUTNAM, Hilary. **Representation and Reality**. 5. ed. Cambridge, UK: MIT Press, 1996. 136 p. ISBN 0-262-66074-1.

_____. Minds and Machines. In: HOOK, Sidney. **Dimensions of Minds**. New York, USA: New York University Press, cap. 2, p. 20-33, 1960. Disponível em: <https://philarchive.org/archive/PUTMAMv1>. Acesso em: 14 jan. 2020.

SEARLE, John R. **A Redescoberta da Mente**. 1. ed. rev. São Paulo: Martins Fontes, 1997. 379 p. ISBN 85-336-0619-2.

_____. **O mistério da Consciência: E discussões com Daniel C. Dennett e David J. Chalmers**. São Paulo: Paz e Terra, 1998. 239 p. ISBN 85-219-0305-7.

SMART, John J. C. Sensations and Brain Processes. **The Philosophical Review**, [s. l.], v. 68, ed. 2, p. 141-156, 1959. DOI 10.2307/2182164. Disponível em: <https://www.jstor.org/stable/2182164> Acesso em: 11 jan. 2020.

TEIXEIRA, João de Fernandes. **Mentes e Máquinas: Uma introdução à ciência cognitiva**. Porto Alegre: Artes Médicas, 1998. 179 p. ISBN 85-7307-329-2.

THAGARD, Paul. **MENTE: Introdução à Ciência Cognitiva**. Porto Alegre: ArtMed, 1996. 210 p. ISBN 85-7307-460-4.

_____. Coherence as Constraint Satisfaction. **Cognitive Science**. A multidisciplinary Journal. Canadá. v. 22, p. 1-24, 1998.

_____. Cognitive Science. **Stanford Encyclopedia of Philosophy**, Stanford University, p. 1-10, 24 set. 2018. Disponível em: <https://plato.stanford.edu/entries/cognitive-science/> Acesso em: 13 jan. 2020.

_____. **Coherence in thought and action**. [S. l.]: MIT Press, 328 p. v. 13, 2000. ISBN 978-0262201315.

_____; FINN, Tracy. Conscience: What is Moral Intuition?. **Morality and the emotions**, Oxford, ano 2011, p. 150-159. Disponível em: <http://cogsci.uwaterloo.ca/Articles/Pages/Emotion-index.html> Acesso em: 17 set. 2020.

_____; AUBIE, Brandon. Emotional consciousness: A neural model of how cognitive appraisal and somatic perception interact to produce qualitative experience. **Consciousness and Cognition**, Waterloo, ano 2008, n. 17, p. 811-834. Disponível em: <http://cogsci.uwaterloo.ca/Articles/Pages/Emotion-index.html> Acesso em: 18 set. 2020.

_____. The Moral Psychology of Conflicts of Interest: Insights from Affective Neuroscience. **Journal of Applied Philosophy**, Oxford, v. 24, ed. 4, p. 367-380, 12 fev, 2007. Disponível em: <http://cogsci.uwaterloo.ca/Articles/Pages/Emotion-index.html> Acesso em: 14 set. 2020.

_____; KROON, Fred W. Emotional consensus in group decision making. **Mind & Society**, [s. l.], ed. 5, p. 1-20, 2006. DOI 10.1007/s11299-006-0011-5. Disponível em: <http://cogsci.uwaterloo.ca/Articles/Pages/Emotion-index.html>. Acesso em: 14 set. 2020.

_____. How Molecules Matter to Mental Computation. **Philosophy of Science**, Chicago, v. 69, n. 3, p. 497-518, set. 2002. Disponível em: <http://www.jstor.org/stable/10.1086/342452> Acesso em: 16 set. 2020.

WOLFF, Francis. **Nossa Humanidade: De Aristóteles às Neurociências**. São Paulo: Editora UNESP, 2012. 335 p. ISBN 978-85-393-0370-0.

COGNITION, EMOTION AND CONSCIOUSNESS: ACCORDING TO PAUL THAGARD'S APPROACH

Abstract

The field of Cognitive Science has shown, over the last decades, broad development in the understanding and modeling of mental processes. Based on two principles, the notion of representation and the analogy between mental processes and computational processes, contemporary Cognitive Science has developed in order to overcome the difficulties imposed by the scientific approach to mental states. The first, the notion of representation, allowed the formalization of the contents of mental activity within a theoretical framework. The second

computational modeling provided the construction of scientific evidence through analogies between experimental data and computational models. However, even considering all these advantages made possible by the approach, some difficulties still persist. The biggest one, evidenced over the years, is the explanation or consideration of the role played by emotions and conscious experience. The challenge is posed as follows: both our emotions and our conscious perceptions appear to play a fundamental role in various cognitive processes, however, they are not taken into account by the approach. Thus, the present work will seek to present an overview of research under the scope of CRUM, in order to expand and supplement the theoretical computational terms, accommodating the role played by emotions and conscious experience.

Keywords: Cognitive Science; Cognition; Emotion; Consciousness; Paul Thagard