



Rede São Paulo de

Formação Docente

Cursos de Especialização para o quadro do Magistério da SEESP
Ensino Fundamental II e Ensino Médio

São Paulo
2011



UNESP – Universidade Estadual Paulista
Pró-Reitoria de Pós-Graduação
Rua Quirino de Andrade, 215
CEP 01049-010 – São Paulo – SP
Tel.: (11) 5627-0561
www.unesp.br



Governo do Estado de São Paulo
Secretaria de Estado da Educação
Coordenadoria de Estudos e Normas Pedagógicas
Gabinete da Coordenadora
Praça da República, 53
CEP 01045-903 – Centro – São Paulo – SP



SECRETARIA
DA EDUCAÇÃO



Modelos mecânicos da mente

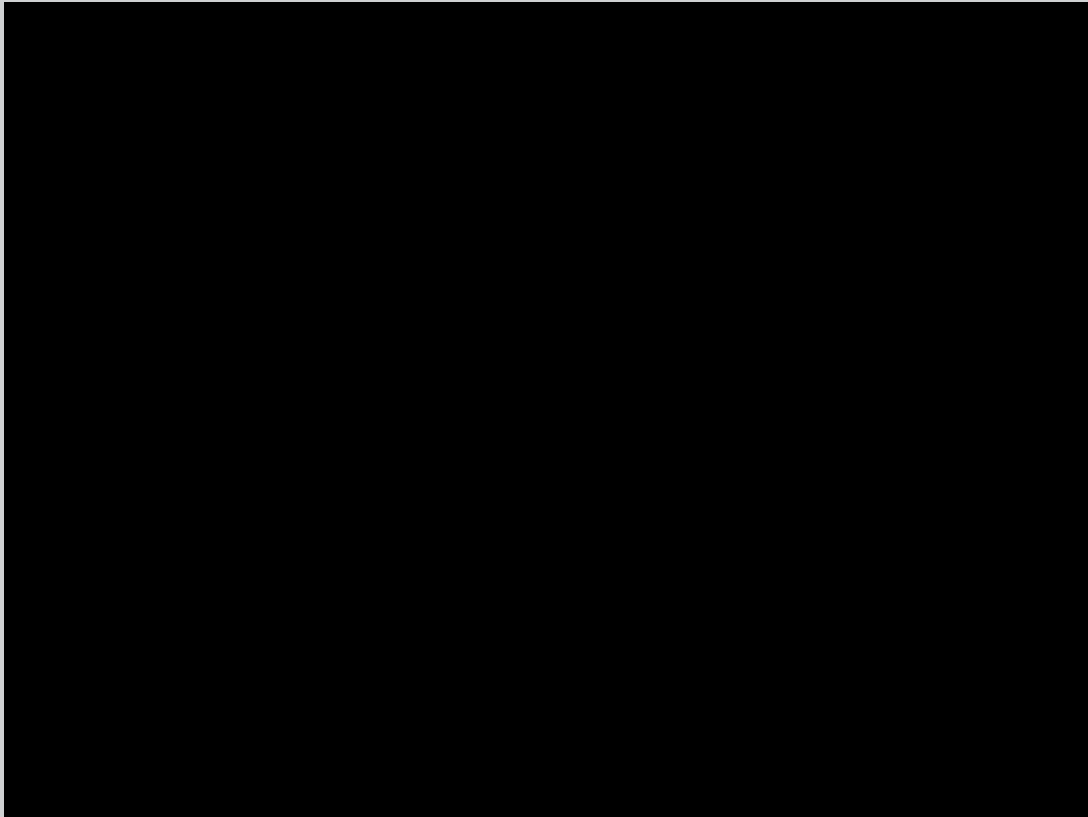


[http://www.tcmoveis.com.br/blog/
wp-content/uploads/2011/03/escher.
jpg](http://www.tcmoveis.com.br/blog/wp-content/uploads/2011/03/escher.jpg)

Sumário

Vídeo da Semana	3
3.1 - A máquina de Turing e a máxima “Pensar é computar”	4
3.2 - Três Vertentes da Ciência Cognitiva	7
Inteligência Artificial	9
(ii) - Redes Neurais Artificiais	10
(iii) - Cognição incorporada e situada	12
Referências	15
Bibliografia	17

Vídeo da Semana



Até agora apresentamos e discutimos problemas decorrentes da dificuldade de se elaborar uma explicação satisfatória da natureza dos estados e processos mentais, bem como da identidade pessoal. Essa dificuldade parece ter suas raízes no fato de que os estudiosos da mente têm que recorrer à própria mente para analisar a mente. Como indicamos no Tema 2, essa circularidade pode comprometer, em certa medida, a objetividade e a universalidade tão almejadas por alguns filósofos e cientistas. Diante dessa dificuldade, o seguinte dilema se apresenta: ou as pretensões de objetividade e universalidade no estudo da mente são abandonadas ou se buscam recursos metodológicos que extrapolem os limites da mente humana para testar suas teorias. Mas quais seriam esses recursos? A proposta dos cientistas cognitivos é, justamente, fornecer um método, conhecido como *Método Sintético de Análise*, que possibilita a criação de modelos mentais artificiais. Conforme veremos na conclusão deste Tema 3, o ideal de objetividade e de universalidade da explicação científica é bastante controverso (na contemporaneidade, esse ideal passa a ser visto sob a ótica dos sistemas complexos que não exclui o pesquisador do seu universo de investigação).

Neste Tema, discutimos o método *Método Sintético* de análise dos estados e processos mentais vigente na Ciência Cognitiva e aceito, não sem resistência, em grande parte da Filosofia da Mente. Trata-se do instrumento fundamental subjacente à modelagem mecânica da mente, já mencionado na disciplina *Teoria do Conhecimento*. A partir de uma breve caracterização desse método, analisamos os seguintes tópicos:

- 3.1 A máquina de Turing e a máxima “Pensar é computar”,
- 3.2 Três Vertentes da Ciência Cognitiva: (i) Inteligência Artificial Simbólica; (ii) Redes Neurais Artificiais e (iii) Cognição situada e incorporada.

Os tópicos acima foram elaborados com o objetivo de fornecer subsídios para um debate crítico acerca do seguinte problema, que direcionará as reflexões do presente Tema: O pensamento humano é estritamente mecânico ou envolve elementos não mecânicos? Como os demais problemas discutidos na disciplina de Filosofia da Mente, este também não possui uma solução consensual entre os seus pesquisadores, deixando em aberto possibilidades de investigação futura apoiadas em testes empíricos.

3.1 A máquina de Turing e a máxima “Pensar é computar”

Conforme indicado na disciplina Teoria do Conhecimento, Alan Turing elaborou a hipótese que pensar é computar através da manipulação de símbolos. Nessa perspectiva, ele argumenta que a natureza do pensamento inteligente pode ser explicada através de modelos mecânicos processadores de informação simbólica que executam a atividade de resolução de problemas. Ao invés de desenvolver um sistema teórico explicativo sobre a natureza do pensamento inteligente, ele propôs um método de elaboração de modelos mecânicos para, efetivamente, realizarem operações que envolvem inteligência.

No artigo *Máquinas e Inteligência* (*Computing Machinery and Intelligence*), originalmente publicado em 1950 na revista *Mind* (então sob a direção de Gilbert Ryle), Turing defende a mencionada hipótese de que *pensar é computar*. Ele argumenta que uma forma de escapar dos labirintos especulativos sobre a natureza da mente (existentes na Filosofia e na Psicologia por séculos) seria a construção, efetiva, de modelos mecânicos do pensamento inteligente. Turing parece incorporar a mencionada máxima de Vico (1744/1976), segundo a qual “saber é fazer”. Para Turing, saberemos efetivamente como funciona a mente ao construir uma máquina pensante.

Turing descreve sua máquina, conhecida como “Máquina de Turing”, capaz de manipular um repertório fixo de símbolos, segundo regras dadas, através de uma fita dividida em células, um leitor móvel de estados, um registro de estados e uma tabela de comportamentos possíveis.

A máquina concebida por Turing não deve ser entendida no sentido comum do termo, como um aparelho elétrico que realiza movimentos físicos semelhantes, por exemplo, aos de uma máquina de costura. Embora ela possa ser implementada em mecanismos físicos, a máquina de Turing é um sistema abstrato universal; trata-se de uma estrutura simbólica, que reúne elementos relacionais que podem (ou não) assumir a forma de máquinas físicas, como é o caso dos computadores.

Turing apresenta os fundamentos de sua máquina universal em um texto de 1936, mas, para os propósitos da nossa análise, vamos apenas discutir o seu conhecido *Teste de Turing*, elaborado no texto de 1950 sob a forma de um “jogo de imitação”. De acordo com esse teste, se um computador puder se fazer passar por um ser humano inteligente num diálogo testemunhado por um juiz, ele teria que ser reconhecido como inteligente. Para compreender essa proposta (à primeira vista bizarra), vamos considerar o seguinte exemplo sugerido por Turing:

Imagine que você, um ser inteligente, esteja trancado em um quarto e que no quarto vizinho exista uma máquina. Um juiz humano, situado em um terceiro cômodo, formulará problemas do tipo: “por favor componha um poema”, “some 2570 a 13.720 e extraia a raiz quadrada dessa soma”, “conte uma piada”, “descreva um entardecer”, “descreva os principais lances do jogo de futebol de ontem” etc. Esses problemas serão escritos e colocados em baixo da porta dos dois quartos, sendo solicitado aos seus ocupantes que os resolvam da melhor forma possível. Dado certo tempo, o juiz recolherá a folha de respostas e caso ele não consiga distinguir quem é o autor das respostas, a máquina terá passado no teste de Turing.

A moral da história é que, dado que você, de fato, é um ser inteligente, que resolveu apropriadamente os problemas e admitindo que a máquina conseguiu resolvê-los de modo igualmente inteligente, então ela teria também que ser considerada inteligente. Como Turing indica, tanto você quanto a máquina poderiam propositalmente tentar enganar o juiz, declarando, por exemplo, que não sabem escrever poemas, que não assistiu o jogo de futebol ou que não consegue fazer cálculos complicados. Sendo admitida a dissimulação, ela não constitui um entrave para a realização do teste de Turing.

Não é difícil imaginar as inúmeras críticas (filosóficas, teológicas, sociológicas, psicológicas, entre outras) sofridas por Turing com sua revolucionária proposta de modelagem da mente através de máquinas. O próprio Turing adianta-se e discute possíveis objeções a sua proposta no texto de 1950. Dentre as objeções mais conhecidas na Filosofia da Mente destacam-se as elaboradas por Dreyfus no livro *O que os computadores não podem fazer*, e por Searle no artigo *Mentes, cérebro e programas (Minds, brains, and programs)* de 1980.

No texto *O que os computadores não podem fazer*, Dreyfus desenvolve críticas veementes ao projeto mecanicista da mente que se iniciara, na década de 1960, denunciando seu aspecto ideológico. Ele ressalta a importância da corporeidade, das sensações e emoções como elementos fundamentais no estudo da mente, as quais os computadores não podem, em princípio, ter, dado que não são seres vivos e não possuem um corpo vivo.

No mesmo viés crítico de Dreyfus, mas em termos mais moderados, no artigo texto *Mentes, cérebros e programas*, Searle propõe o argumento do “Quarto Chinês”, problematizando o teste de Turing. Ele sugere uma situação hipotética em que uma pessoa, que não sabe mandarim, está dentro de um quarto com todas as instruções em seu próprio idioma para manipular símbolos chineses. Essa pessoa recebe cartões com símbolos chineses por uma pequena abertura em seu quarto e, graças às instruções que possui em seu próprio idioma, consegue dar respostas destituídas de significado para elas, que ignora o mandarim, mas que para um juiz externo criam a ilusão de serem respostas significativas e adequadas.

Cabe ressaltar que a pessoa que manipulou os símbolos não sabe mandarim e nem aprendeu o idioma enquanto seguia as instruções em sua própria língua. Neste sentido, ela está apenas seguindo instruções, como uma máquina que opera sem a compreensão do significado do que está fazendo. Este exemplo ilustra, segundo Searle, que o processamento de informações dos modelos mecânicos se limita à manipulação da sintaxe sem entrar no domínio do significado, indispensável ao estudo da mente.

Em resumo, o argumento de Searle procura ressaltar que, embora modelos computacionais possam constituir ferramentas para o estudo da inteligência, eles não constituem um bom instrumento explicativo sobre o funcionamento da mente.

As críticas elaboradas por muitos filósofos, bem como o fato que até o momento nenhuma máquina conseguiu passar no teste de Turing por um tempo satisfatório, colocam em xeque a adequação da proposta de Turing para o estudo da mente. Contudo, o teste de Turing é considerado uma das bases fundamentais da Ciência Cognitiva nas suas três vertentes: (i) Inteligência Artificial Simbólica, (ii) Redes Neurais Artificiais e (iii) Cognição Situada e Incorporada. No tópico 3.2 apresentamos as principais características dessas vertentes.

3.2 Três Vertentes da Ciência Cognitiva

O projeto ambicioso de Turing de explicar a natureza da mente através da elaboração de modelos mecânicos propiciou o surgimento da *Inteligência Artificial*, das *Redes Neurais Artificiais* e da *Robótica Cognitiva*, esta última também conhecida como *Cognição Incorporada e Situada*. Essas áreas de investigação constituem as três vertentes da *Ciência Cognitiva*; elas possuem em comum o **método sintético de análise**, cujos pressupostos centrais são: (i) A hipótese de que a modelagem mecânica de processamento de informação fornece subsídios fundamentais para a compreensão e a explicação da natureza dos estados e processos mentais;

(ii) Estados mentais devem ser analisados em termos *internalistas*: eles são *internos* aos sistemas processadores de informação;

(iii) O teste de Turing é um balizador do sucesso de um modelo – ele indica a compreensão, por parte do programador, das principais etapas do pensamento inteligente responsável pela resolução de problemas.

A partir dos pressupostos acima, modelos mecânicos são elaborados como substitutos de teorias explicativas da atividade mental inteligente. Conforme ressalta Dupuy, a noção de *modelo* assume na Ciência Cognitiva uma nova dimensão explicativa:

[...] se trata de uma idealidade, no mais das vezes formalizada e matematizada, que sintetiza um sistema de relações [...] O modelo é como uma forma abstrata que vem encarnar-se ou realizar-se nos fenômenos (1996, p. 23).

Parte do que se faz nesta área de pesquisa interdisciplinar, que envolve também a Computação e a Robótica, tem interesse estritamente tecnológico e não possui estreita relação com as pesquisas propriamente filosóficas. Contudo, o que interessa para o presente estudo é a sua

parceria com a Filosofia da Mente e a Semiótica, na medida em que ela fornece subsídios para o estudo da natureza informacional dos processos de pensamento.

Além do mencionado artigo de Turing (1950), outro texto escrito por Hilary Putnam em 1960, intitulado *A natureza dos estados mentais* (*The nature of mental states*) deu inicialmente sustentação à hipótese mecanicista da mente. Nesse texto, Putnam adota a perspectiva *funcionalista*, já mencionada, da múltipla realizabilidade.

De acordo com a perspectiva da *múltipla realizabilidade*, diferentes estruturas físicas podem conduzir ao mesmo resultado funcional na realização de uma tarefa. Assim, por exemplo, a operação de adição pode ser realizada através do agrupamento de maçãs, da manipulação de um ábaco de madeira ou de uma calculadora digital. Nessa perspectiva, Putnam (1960) argumenta que estados mentais são estados funcionais que se estabelecem através de relações causais entre estímulos sensoriais (*inputs*), outros estados mentais e comportamentos (*outputs*).

Adotando a hipótese da *múltipla realizabilidade*, os cientistas cognitivos dos anos de 1960 julgavam que para modelar processos mentais seria preciso entender os seus mecanismos funcionais e não o substrato material que os instancia. Essa concepção funcionalista da mente, aliada à proposta mecanicista de Turing, forneceu as ferramentas teóricas necessárias para a modelagem de processos cognitivos, então concebidos como capacidades funcionais de processamento de informação. O funcionalismo computacional constitui os alicerces da área de pesquisa conhecida como *Inteligência Artificial*, que adota o método sintético de análise na modelagem de processos que se supõem presentes no pensamento inteligente.

Uma versão do método analítico de grande interesse para Inteligência Artificial e para as Redes Neurais Artificiais é denominado **método sintético de análise**, cujo pressuposto central é a hipótese de que a modelagem fornece subsídios fundamentais para a compreensão e a explicação da natureza dos estados e processos mentais.

Explicitaremos neste Tema três faces da modelagem mecânica da mente com vistas à tentativa de explicação de sua dinâmica: Inteligência Artificial, Conexionismo e Cognição Situada e Incorporada.

Inteligência Artificial

A Inteligência Artificial (IA) é tradicionalmente dividida em duas vertentes, IA forte e IA fraca, representadas pelos projetos de Simulação Cognitiva e de Reprodução Cognitiva. Ainda que não exista uma linha nítida separando as duas vertentes, podemos dizer que uma distinção básica entre elas reside no suposto grau de realidade psicológica dos modelos elaborados pelos pesquisadores nos seus respectivos domínios de estudo.

Em resumo, a diferença central entre a IA fraca e a IA forte reside no pressuposto de que, de acordo com a IA forte, a modelagem computacional efetivamente *apresenta* um comportamento inteligente: máquinas, devidamente programadas, pensam. Em contraste, os pesquisadores da IA fraca supõem que os modelos mecânicos apenas *simulam* adequadamente tais comportamentos.

Apesar de distintos, os projetos da IA forte e da IA fraca têm em comum a estratégia de investigar as estruturas de símbolos supostamente responsáveis pelo comportamento inteligente. Enquanto a IA fraca se espelha no ser humano para elaborar os seus modelos, os pesquisadores da IA forte pretendem explicar a natureza dos processos mentais responsáveis pelo comportamento inteligente, seja de uma máquina ou de um animal.

Como indicado, o estudo dos processos mentais é feito com o auxílio do método sintético que permite a construção de modelos, na forma de programas, para explicar, através do computador digital, segmentos do comportamento inteligente.

O que deve ser entendido por “comportamento inteligente” é uma questão controversa. Contudo, a tendência geral dos pesquisadores da IA é considerar que uma característica básica desse comportamento é a capacidade de resolver problemas. É nesse sentido que Margaret Boden em *Inteligência artificial e o homem natural (Artificial Intelligence and Natural Man)* afirma:

A inteligência pode ser definida como a habilidade criativa de um sistema para manipular símbolos, ou processos de informação, dadas as exigências da tarefa em questão. (1977, p. 17)

Tendo em vista essa concepção de inteligência, os trabalhos desenvolvidos na IA fraca dizem respeito à simulação do comportamento humano durante a atividade de resolução de

problemas. No caso da IA forte, existe uma preocupação menor por parte dos pesquisadores para com o grau de realidade psicológica humana dos modelos. Retirando o ser humano do centro do universo cognitivo, os defensores da IA forte, como Newel e Simon (1972) e Minsky, (1976, 2006), por exemplo, julgaram ser de pouca relevância a questão de se os processos empregados pela máquina na execução de tarefas que exigem inteligência são semelhantes (ou não) àqueles utilizados pelos seres humanos.

Para realizar a modelagem, os cientistas da IA elaboram um programa particular que possa realizar uma determinada tarefa e consideram, a seguir, quais aspectos desse programa seriam necessários para que um sistema qualquer executasse tal tarefa. Como indica Gonzalez (1984), o pressuposto dessa técnica de modelagem é a de que os macroprocessos manifestos no comportamento inteligente podem ser devidamente explicados reduzindo-os a mecanismos simples, átomos hipotéticos chamados “processos elementares de informação”. Os sistemas que executam tais processos são conhecidos em IA como “sistemas de processamento de informação”.

Apesar dos resultados alcançados pelos modelos mecânicos construídos a partir da década de 1970, durante os anos de 1980 o entusiasmo dos pesquisadores foi progressivamente diminuído diante das várias dificuldades que enfrentaram e que foram apontadas por filósofos como Baker (1976), Dreyfus (1979) Searle (1980) que mostraram a fragilidade de pressupostos da IA na explicação dos processos mentais. Especial crítica sofreu o pressuposto estritamente simbólico da IA que desconsidera o papel do corpo nos processos cognitivos. Para lidar com essa dificuldade, os modelos de redes neurais artificiais entraram em cena.

(ii) Redes Neurais Artificiais

Na tentativa de superar as dificuldades enfrentadas pelos modelos tradicionais da IA, de manipulação de símbolos, cientistas cognitivos (dentre os quais se destacam Hopfield, 1982; Rumelhart e McClelland, 1986; Kohonen, 1989, dentre outros) propuseram a retomada de um tipo de modelagem conhecido como *Redes Neurais Artificiais*. Entre esses primeiros modelos estão aqueles propostos por McCulloch e Pitts em 1943. Tais redes constituem a ferramenta de trabalho da vertente conhecida como *Conexionismo*; elas possuem certa inspiração biológica e possibilitam que leis físicas sejam consideradas no estudo da mente.

As redes neurais artificiais são modelos dotados da capacidade de processamento de informação em paralelo, tendo por base a interação de unidades denominadas neurônio-símile, que lhes permitem aprender com treinamento e melhorar suas *performances*. Um dos objetivos desta nova forma de modelagem é levar em consideração aspectos físicos e não apenas simbólicos como faziam os modelos da IA.

Não entraremos em detalhes técnicos sobre os processos de treinamento de redes neurais, mas julgamos interessante estabelecer um certo paralelo entre eles e a concepção humeana de formação das idéias. Como sabemos, para Hume (1996) as idéias se formam através da repetição de dados sensoriais que se apresentam aos órgãos dos sentidos deixando uma impressão no palco da mente. Tais impressões, inicialmente vívidas, são filtradas dando lugar às impressões menos vívidas e mais abstratas, que constituem as idéias. Uma vez formadas, as idéias são organizadas de acordo com três princípios fundamentais: (a) Associação; (b) Causa e efeito (c) Contigüidade. Esses mesmos princípios podem ser encontrados em muitos dos modelos de redes neurais artificiais, treinados para criar e reconhecer padrões informacionais nos estímulos representativos de diversos objetos que lhes são apresentados.

Entre os modelos conexionistas mais interessantes se destacam as redes neurais de Kohonen (1989) que envolvem processos de auto-organização. Tais redes aprendem, de modo auto-organizado, a criar e a reconhecer padrões informacionais em estímulos representativos de diversos objetos. Uma reflexão filosófica dos processos de auto-organização em redes neurais pode ser encontrado em Gonzalez (1992; 1996). Interessa ressaltar aqui, de forma resumida, as seguintes características centrais dos processos de auto-organização presentes em redes neurais: (1) eles se desenvolvem sem um centro direcionador específico preestabelecido; (2) eles resultam da interação espontânea entre elementos distintos e (3) eles possibilitam a formação de padrões ou sistemas que aprendem com a repetição e com mecanismos de ajuste. Incorporando os recursos (1)-(3), os modelos conexionista auto-organizados são treinados para reconhecer faces, objetos em geral, padrões sonoros, entre muitos outros, independente de uma programação fixa estabelecida no plano simbólico (macroscópico) de análise. Os padrões que eles aprendem a reconhecer emergem da interação entre as unidades neurônio símile, no plano físico, microscópico, de estruturação da rede neural artificial.

Os conexionistas argumentam que os padrões formados nas redes neurais artificiais reúnem elementos indicativos do conteúdo informacional dos objetos reconhecidos, superando dessa forma, as críticas do tipo “Quarto Chinês” que enfraquecia o poder explicativo dos modelos da IA simbólica.

Embora seja discutível o avanço propiciado pelo projeto de pesquisa conexionista, em sua tentativa de superação dos problemas próprios aos modelos da IA tradicional, cabe destacar que modelos de redes neurais são freqüentemente instanciados em máquinas de Turing sem prejuízo para sua *performance*. Por tal razão, os defensores da IA tradicional alegam que, no limite, o processamento em paralelo das redes neurais pode ser reduzido ao processamento simbólico seqüencial próprio da máquina de Turing. Se assim for, a despeito dos esforços dos idealizadores dos modelos de redes neurais de superarem problemas dos modelos simbólicos, alguns deles reaparecem em alguma medida, sobretudo relacionados à redução funcional que propiciam.

Ainda que permaneça em aberto a polêmica sobre a relação do conexionismo com o funcionalismo computacional da IA, o fato é que ambos acabam por subestimar a relevância da ação efetiva dos modelos em sua frágil relação com o meio ambiente, e com o corpo, no estudo dos processos mentais. Como veremos a seguir, uma vez mais, na tentativa de superar essa dificuldade, uma nova vertente da Ciência Cognitiva, denominada *Cognição Situada e Incorporada*, surge na década de 1990 em parceria com a robótica.

(iii) Cognição incorporada e situada

Como indicamos, esta perspectiva teórica começa a se delinear a partir dos anos de 1990 e tem como pressuposto central a hipótese que os processos mentais e o comportamento inteligente estão diretamente relacionados à complexa rede de interações entre os organismos e o meio ambiente. Ao ressaltar o aspecto corporal dos processos mentais, esta abordagem destaca a integração e a interdependência existente entre processos corpóreos em geral (metabólicos, hormonais, etc.) e os processos mentais. Além disso, esta abordagem chama a atenção para as relações que o organismo estabelece com o meio ambiente em que está fisicamente situado e a influência que características ambientais têm no desenvolvimento de processos mentais.

Cognitivistas como Clark (2001, 2008) e Haselager (2004) entendem que é preciso criar modelos robóticos corpóreos e ambientalmente situados que sejam capazes de desempenhar

performances inteligentes análogas às que efetuam os organismos biológicos. A elaboração de tais modelos visa ressaltar a importância dos processos de auto-organização para lidar com os desafios e imprevistos que surgem em ambientes não controlados. Em especial, a Cognição Incorporada e Situada focaliza a relação corpo/ambiente em sua permanente (e de dupla mão) busca por ajustes para a adaptação às novas variáveis que tal interação dinâmica gera. Uma das grandes virtudes desta abordagem é que ela permite levar em consideração uma multiplicidade de fatores hoje reconhecidos relevantes para a compreensão dos processos mentais que até então foram preteridos pelas vertentes da IA e das redes neurais artificiais.

Na contemporaneidade, a Cognição Incorporada e Situada propõe modelos dos processos e estados mentais que possuem uma certa plausibilidade biológica, inclusive levando em conta fatores envolvendo a multiplicação e a complexificação das possibilidades de interação dos modelos mecânicos (agora na forma de robôs) com o ambiente e do ambiente com tais modelos.

Em síntese, o suporte filosófico da Cognição Situada e Incorporada é fornecida pelas teorias de Auto-organização, da Filosofia Ecológica e da Neurofilosofia, as quais investigam fenômenos emergentes da interação dinâmica entre elementos constituintes de sistemas complexos. Contudo, entendemos que apesar dos avanços realizados nesta nova vertente da Ciência Cognitiva, ainda temos aqui os principais pressupostos do método sintético, quais sejam:

- a. A Cognição é uma capacidade individual, interna dos sistemas que processam informação;
- b. Uma forma apropriada de verificar a força explicativa das hipóteses sobre a natureza dos processos e estados mentais reside na elaboração de modelos mecânicos, robôs auto-organizados;
- c. O sucesso da modelagem deve ser avaliado pelo teste de Turing

Não é difícil de se imaginar que a manutenção dos pressupostos do método sintético, e consequentemente, da hipótese mecanicista no estudo dos processos mentais continua gerando grande indignação entre filósofos, teólogos, psicólogos, entre outros. Uma amostra dessa indignação pode ser percebida na seguinte passagem de Kravchenko:

Tenho grande respeito por Alan Turing mesmo que seja por uma simples razão: poucos cientistas foram capazes de enfeitiçar de modo tão profundo as mentes de outros pesquisadores no estudo da cognição, cegando-os em relação a qualquer evidência empírica e/ou argumentos teóricos contra a famigerada hipótese de que pensar é computar (2007).

Como indicamos no início deste Tema, a nossa problemática central, que organizou a presente reflexão, consiste em apresentar evidências sobre a plausibilidade da hipótese mecanicista da natureza do pensamento. Vimos que uma das motivações centrais para a elaboração de modelos mecânicos da mente na Ciência Cognitiva reside na aparente dificuldade gerada pelo fato que os estudiosos da mente têm que recorrer à própria mente para analisar a mente.

Contudo, o argumento da suposta dificuldade de a mente ser seu próprio objeto de estudo, não parece ser uma autêntica dificuldade. Pois, que tipo de problema teríamos aí? Falta de distanciamento do objeto de estudo? Talvez seja interessante diferenciar objetividade e não-envolvimento. Será que o envolvimento em um tema compromete seriamente a objetividade do sujeito? Há inúmeros contraexemplos para desafiar esta crença. Os melhores historiadores costumam ser filhos da própria cultura que estudam e, não raro, testemunhas dos acontecimentos que descrevem. Um avião consegue medir sua própria velocidade com excelente precisão através de inferências, dentre muitos outros exemplos.

Alguns críticos admitem que se o fato de a mente ser concomitantemente instrumento e objeto de estudo constitui um problema sério, não fica claro por que ou como modelos mecânicos da mente ajudam a superar esta dificuldade!? Em Direito, os juízes tendem a rejeitar testemunhas que tenham grau de parentesco com o réu, nem tanto pelo fato de haver envolvimento emocional (a vítima também costuma ter envolvimento emocional com o réu e, no entanto, seu testemunho costuma ser muito valorizado), mas principalmente porque por questões de lealdade ou solidariedade com o réu os parentes podem produzir falsos testemunhos. Porém, se a única testemunha de um crime é um parente de primeiro grau do réu ou da vítima, seu depoimento é levado em conta pelos juízes.

Além das objeções acima, é importante lembrar que nos estudos atuais dos sistemas complexos, métodos recorrentes, nos quais aparentemente há uma circularidade, são amplamente utilizados seja na matemática, na computação, na engenharia etc. (por ex., para calcular o fa-

torial de qualquer número, é muito comum usar algoritmos com recorrência). Computadores são utilizados para fabricar outros computadores ou para fazer um autoteste (para detectar possíveis problemas). Animais fazem autoexames e tomam decisões para superarem algum problema de saúde.

As considerações acima sugerem que o problema em discussão não parece ser de objetividade e universalidade. Ele parece ser de outra natureza: da falta de conhecimento básico, elementar, sobre como seres dotados de mente armazenam e resgatam informações significativas e como elas auxiliam na organização do pensamento e da ação. Somente o conhecimento da região de ativação de um mecanismo, ou de sua interação mecânica com o meio não parece ser suficiente para explicar o funcionamento desse complexo sistema de relações.

Conforme sugere Edgar Morin no livro *Ciência com Consciência*, a nova ciência, seja da mente ou de qualquer outro objeto de investigação, exige que se inclua não apenas o olhar do observador em suas investigações, mas, e principalmente, a Consciência desse olhar que marca a característica central dos sistemas complexos. Entretanto, como entender a natureza da “consciência”? Trata-se, como é consensual na Filosofia da mente, do problema (realmente) difícil, que passaremos a discutir no Tema 4. A nossa sugestão no presente Tema é que os pós-graduandos discutam criticamente as hipóteses aqui apresentadas para que, então, possam compreender o mais difícil dos nossos objetos de investigação: a natureza da consciência.

Referências

- BAKER, L. R. **Saving Belief**: a critique of physicalism, [New Jersey]: Princeton University Press, 1987.
- BODEN, M. **The artificial intelligence and the natural man**. Hassocks: The Harvester Press Ltd., 1977.
 - CLARK, A. **Mindware**: an introduction to the philosophy of cognitive science. Oxford: Oxford University Press, 2001.
 - CLARK, A. **Supersizing the Mind**: Embodiment, Action, and Cognitive Extension. Oxford: Oxford University Press, 2008.
 - DREYFUS, H. **What Computers Can't Do**: The Limits of Artificial Intelligence. Revised edition. New York: Harper and Row, 1979.

- DUPUY, J.P. **Nas origens das ciências cognitivas**. Tradução de Roberto Leal Ferreira Manha. [São Paulo]: Unesp, 1996.
- GONZALEZ, M. E. Q. **Metodologia da descoberta científica e inteligência artificial**. 1984, Dissertação (Mestrado em Filosofia)- Instituto de Filosofia e Ciências Humanas, Programa de Pós-Graduação em Filosofia e Ciência Humanas, Campinas, 1984.
- GONZALEZ, M. E. Q. O nascimento da ciência cognitiva e suas raízes na física do século XIX. In: EVORA, F. R. R. **O século XIX: O nascimento da ciência contemporânea**. Campinas: Unicamp, 1992. (Coleção CLE).
- GONZALEZ, M. E. Q. Ação, causalidade e ruído nas redes neurais auto-organizadas. DEBRUN, M. A. GONZALEZ, M. E. Q.; PESSOA JR., O. (Eds.) **Auto-organização: Estudos Interdisciplinares**. Campinas, [Unicamp], 1996, p. 1-23, v. 18, (Coleção CLE).
- HASELAGER, W.F.G. O mal estar do representacionismo: sete dores de cabeça da ciência cognitiva. In: A. FERREIRA, M.E.Q. GONZALEZ & J.G. COELHO (eds.). **Encontros com as ciências cognitivas**. Campinas: [Unicamp], 2004, p. 105-120. v. 4, (Coleção Estudos Cognitivos).
- KOHONEN, T. **Self-Organization and Associative Memory**. Berlin: Springer- Verlag, 1989.
- KRAVCHENKO, A.V. Whence the autonomy? A response to Harnard and Dror. **Pragmatics & Cognition**, v. 15, n. 3, p. 587-598. Special Issue.
- HUME, D. **Investigação acerca do entendimento humano**. São Paulo: Unesp, 1996.
- MINSKY, M. Automation and Artificial Intelligence. In: **Science, Technology, and the Modern Navy**, Arlington: Office of Naval Research, 1976.
- MINSKY, M. **The Emotion Machine: Common sense Thinking, Artificial Intelligence, and the Future of the Human Mind**. New York: Simon & Schuster, 2006.
- MORIN, E. **Uma ciência com consciência**. Rio de Janeiro: Bertrand Brasil, 2003.
- NEWELL, A.; SIMON, H. A. **Human problem solving**. Englewood Cliffs, NJ: Prentice-Hall, 1972
- PUTNAM, H. The nature of mental states. In: LYCAN, W. **Mind and cognition: an an-**

thology. 2nd. ed. Oxford: Blackwell Publications, 1960-1999.

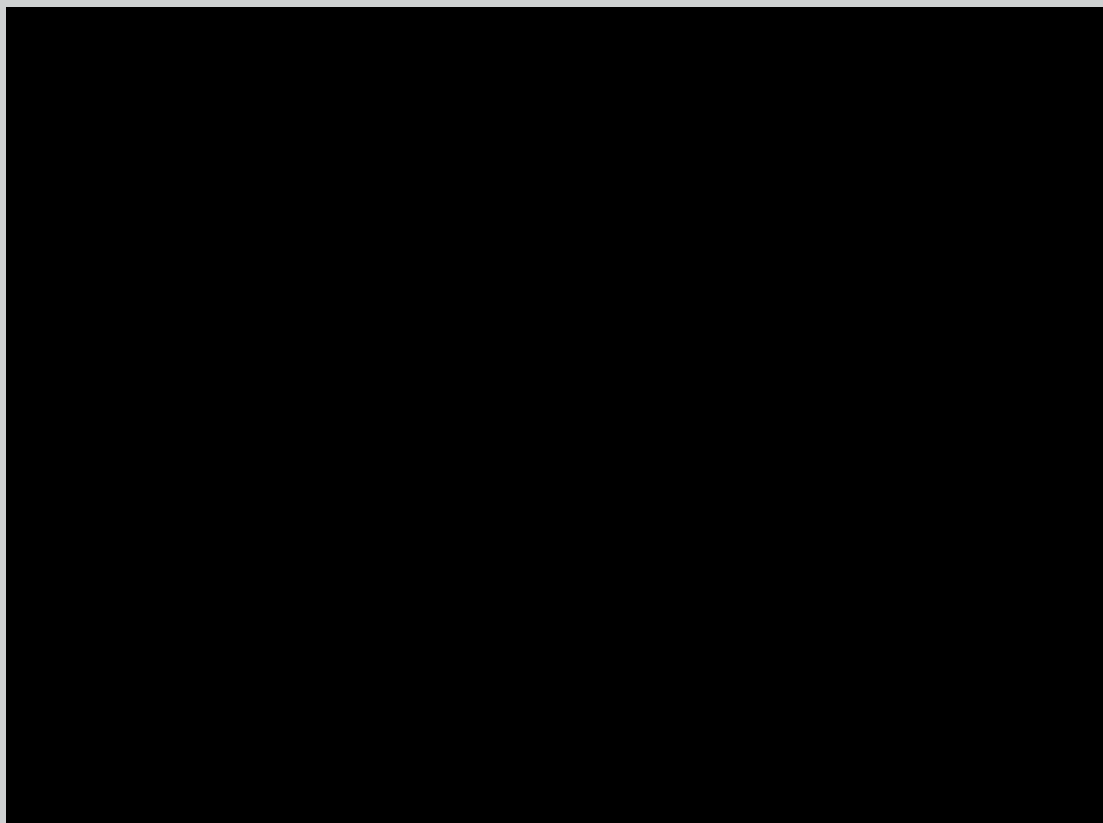
- SEARLE, J. R. Minds, brains, and programs. **Behavioral and Brain Sciences**, v. 3 n. 3, p. 417-457, 1980. Disponível em: <<http://www.bbsonline.org/Preprints/OldArchive/bbs.searle2.html>>. Acessado em: 4 maio 2011.
- TURING, A. 'Computing Machinery and Intelligence'. **Mind**, n. 59, p. 433-460, 1950.

Bibliografia

- DESCARTES, R. **Obra escolhida**. Tradução de J. Guinsburg e Bento Prado Junior, introdução de Giles Gaston-Granger, prefácio e notas de Gerard Lebrun. 3. ed. Rio de Janeiro: Bertrand Brasil, 1994.
- HODGES, A. **Turing**: Um filósofo da natureza. Tradução de Marcos Barbosa de Oliveira. São Paulo: Edunesp, 2001.

Ficha da Disciplina:

A Filosofia da Mente



Maria Eunice Quilici Gonzalez

Mariana Claudia Broens

André Leclerc



Apresentação dos professores-autores:

O conteúdo da disciplina Filosofia da Mente foi elaborado conjuntamente por duas professoras do Departamento de Filosofia da UNESP de Marília: Maria Eunice Quilici Gonzalez ¹ e Mariana Claudia Broens² e pelo professor André Leclerc³ do Departamento de Filosofia da Universidade Federal do Ceará.

Estrutura da Disciplina

Disciplina Filosofia da Mente	Tema 1: Introdução à Filosofia da Mente	1.1 Caracterização da Filosofia da Mente
		1.2 Diferentes abordagens no estudo da mente
	Tema 2: Problemas centrais da Filosofia da Mente	2.1 O problema mente/corpo
		2.2 O problema das outras mentes
		2.3 O problema da identidade pessoal
	Tema 3: Modelos mecânicos da mente	3.1 A máquina de Turing e a máxima “Pensar é computar”
		3.2 Inteligência Artificial
		3.3 Redes Neurais Artificiais
	Tema 4: Intencionalidade e Consciência	4.1 Intencionalidade originária e Intencionalidade derivada
		4.2 Consciência e Subjetividade

1 **Maria Eunice Quilici Gonzalez** é PhD em Cognitive Science, Language And Linguistics pela Universidade de Essex, Inglaterra e professora Livre Docente da UNESP. Tem experiência de pesquisa e de docência em Teoria do Conhecimento, Filosofia Ecológica, História da Filosofia Contemporânea, Ciência Cognitiva e Filosofia da Mente, atuando principalmente nos seguintes temas: informação ecológica, percepção-ação, auto-organização, pragmatismo e Ética da Informação.

2 **Mariana Claudia Broens** é doutora em Filosofia pela Universidade de São Paulo e professora Livre Docente da UNESP. Tem experiência de pesquisa e de docência em Teoria do Conhecimento, História da Filosofia Moderna, História da Filosofia Contemporânea e em Filosofia da Mente, trabalhando os seguintes temas: a abordagem mecanicista da mente, Naturalismo, Auto-Organização e Pragmatismo.

3 **André leclerc** concluiu o doutorado em filosofia na Universidade de Quebec em 1990. Atualmente é Professor Associado da Universidade Federal do Ceará. Atua na área de Filosofia, com ênfase em Filosofia da Linguagem e Filosofia da Mente, trabalhando os seguintes temas: filosofia analítica, conteúdo mental, epistemologia, externalismo, contextualismo, intencionalidade, anti-individualismo, semântica e naturalismo biológico.

Ementa:

A disciplina **Filosofia da Mente** tem por objetivo investigar os principais problemas relacionados à natureza da mente - dos eventos, estados, processos e das funções mentais - bem como sua relação com o corpo e o meio ambiente. Diferentes tentativas de elucidação desses problemas são encontradas nas diversas vertentes da Filosofia da Mente, entre as quais se destacam a abordagem analítica (anglo-saxã) e a continental, além daquelas que possuem influência da filosofia oriental. Dados os limites do presente curso, apresentaremos apenas a abordagem analítica tradicional da Filosofia da Mente. Com esse objetivo, serão tratados os seguintes temas:

Tema 1 - Introdução à Filosofia da Mente: 1.1 Caracterização da Filosofia da Mente; 1.2 Diferentes abordagens no estudo da mente.

Tema 2 - Problemas centrais da Filosofia da Mente: 2.1 O problema mente/corpo, 2.2 O problema das outras mentes, 2.3 O problema da identidade pessoal.

Tema 3 - Modelos mecânicos da mente: 3.1 A máquina de Turing e a máxima “Pensar é computar”, 3.2 Inteligência Artificial, 3.3 Redes Neurais Artificiais, 3.4 Cognição situada e incorporada.

Tema 4 - Intencionalidade e Consciência: 4.1 Intencionalidade originária e Intencionalidade derivada; 4.2 Consciência e subjetividade.

Através da análise dos temas 1 a 4, buscamos familiarizar os pós-graduandos com um campo instigante de investigação filosófica recente em nosso país. Várias das hipóteses aqui apresentadas são bastante controversas e muitas vezes elas se chocam com concepções do senso comum e da tradição filosófica clássica sobre a natureza da mente. Sugerimos que esta disciplina seja estudada com uma atitude que combine a análise rigorosa e a visão crítica das hipóteses aqui apresentadas.

Pró-Reitora de Pós-graduação

Marilza Vieira Cunha Rudge

Equipe Coordenadora

Ana Maria Martins da Costa Santos

Coordenadora Pedagógica

Cláudio José de França e Silva

Rogério Luiz Buccelli

Coordenadores dos Cursos

Arte: Rejane Galvão Coutinho (IA/Unesp)

Filosofia: Lúcio Lourenço Prado (FFC/Marília)

Geografia: Raul Borges Guimarães (FCT/Presidente Prudente)

Antônio Cezar Leal (FCT/Presidente Prudente) - *sub-coordenador*

Inglês: Mariangela Braga Norte (FFC/Marília)

Química: Olga Maria Mascarenhas de Faria Oliveira (IQ Araraquara)

Equipe Técnica - Sistema de Controle Acadêmico

Ari Araldo Xavier de Camargo

Valentim Aparecido Paris

Rosemar Rosa de Carvalho Brena

Secretaria/Administração

Márcio Antônio Teixeira de Carvalho

NEaD – Núcleo de Educação a Distância

(equipe Redefor)

Klaus Schlünzen Junior

Coordenador Geral

Tecnologia e Infraestrutura

Pierre Archag Iskenderian

Coordenador de Grupo

André Luís Rodrigues Ferreira

Guilherme de Andrade Lemeszenski

Marcos Roberto Greiner

Pedro Cássio Bissetti

Rodolfo Mac Kay Martinez Parente

Produção, veiculação e Gestão de material

Elisandra André Maranhe

João Castro Barbosa de Souza

Lia Tiemi Hiratomi

Liliam Lungarezi de Oliveira

Marcos Leonel de Souza

Pamela Gouveia

Rafael Canoletti

Valter Rodrigues da Silva