

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO

PUC-SP

Luciana Akemi Nakabayashi

A contribuição da Inteligência Artificial (IA)

na Filosofia da Mente

MESTRADO EM TECNOLOGIAS DA INTELIGÊNCIA E DESIGN DIGITAL

SÃO PAULO

2009

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO

PUC-SP

Luciana Akemi Nakabayashi

A contribuição da Inteligência Artificial (IA)

na Filosofia da Mente

MESTRADO EM TECNOLOGIAS DA INTELIGÊNCIA E DESIGN DIGITAL

Dissertação apresentada à Banca Examinadora como exigência parcial para obtenção do título de MESTRE em Tecnologias da Inteligência e Design Digital pela Pontifícia Universidade Católica de São Paulo, sob orientação do Prof ° Doutor Luís Carlos Petry.

SÃO PAULO

2009

Banca Examinadora

Yoshiaki Nakabayashi

In memoriam

Agradecimentos

Ao orientador Professor Doutor Luís Carlos Petry, por todas as orientações e pelo apoio com que me presenteou ao longo do curso.

Ao Professor Doutor João de Fernandes Teixeira, notório intelectual, que com presteza e generosidade foi além das obras escritas e me levou a renovar leituras e encontrar o fio deste estudo e, mais, fez manter viva a minha vontade de compreender temática tão instigante quanto a da Inteligência Artificial (IA) no âmbito da Filosofia da Mente.

À Professora Lúcia Rodrigues Máximo, minha gratidão pela amizade, carinho, compreensão e dedicação. Por entender a dificuldade de colocar no papel tantas e novas informações, acompanhou com paciente estímulo o semear de palavras, compartilhou a relação entre os conectivos e, sobretudo, soube deixar viva a minha expectativa de colheita; a possível satisfação de eu vir a cumprir o percurso a que me propus.

Aos Professores Oswaldo Giorgi e Laura Mazzei, do Centro Paula Souza (CEETEPS), pelo acolhimento dos relatórios semestrais das atividades realizadas no curso de Mestrado, oportunidade em que a confiança foi renovada e demonstrada a cada semestre em que foi preciso promover o afastamento das funções docentes por mim exercidas junto à prestigiosa instituição que representam.

À CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior pelo apoio financeiro.

A todos que comigo compartilharam suas experiências e conhecimento, amigos compreensivos e preocupados em valorizar o meu interesse em pesquisar e escrever dissertação sobre “A contribuição da Inteligência Artificial (IA) na filosofia da Mente”.

Sumário

Resumo.....	8
Abstract.....	9
Introdução.....	10
Capítulo I Fundamentos conceituais para uma abordagem da inteligência no ciberespaço.....	13
1.1 O que é Inteligência Artificial (IA).....	19
1.2 O desenvolvimento da Inteligência Artificial (IA) como Disciplina	23
1.3 Breve conceito de Ciência Cognitiva	26
1.4 Paradigmas clássicos das abordagens conexionista e cognitivista.....	29
1.5. Críticas Filosóficas e argumentação de uma Inteligência Artificial (IA) forte.....	33
Capítulo II Abordagens joaninas da Filosofia da Mente rumo à Inteligência Artificial (IA) no ciberespaço	35
2.1 A natureza do Mental.....	36
2.2 O problema mente - cérebro.....	37
2.3 Maquinismo e Consciência	38
2.4 Breve histórico da obra de John Searle.....	45
2.5 A experiência do quarto chinês.....	48
2.6 Searle e a crítica ao computacionalismo	53

2.7 Searle e a crítica ao Cognitivismo_.....	56
Capítulo III Inteligência Artificial (IA) ou entendimento automático?_.....	59
3.1. O que é um computador?_	60
3.2 Homens versus máquinas_.....	62
3.3 A máquina permitindo ao homem ocupar-se de si	64
3.4 A invenção da cibernética_.....	66
3.5 Modelo antropológico e resíduo cultural_.....	68
3.6 O corpo pós-humano_.....	72
3.7 Os limites do ciberespaço.....	75
3.8 Inteligência Artificial (IA) na ficção científica.....	79
Capítulo IV Inteligência Artificial (IA): interagindo com robôs de conversação brasileiros.....	82
4.1 Vôo de reconhecimento.....	83
4.2 Cybelle, a primeira robô de experiência brasileira na web	86
4.3. Informação, caixa preta e interação_.....	93
Considerações finais.....	100
Referências Bibliográficas.....	104

Resumo

A presente pesquisa tem como tema central investigar o conceito de Inteligência e, especificamente, de Inteligência Artificial (IA), derivado das Ciências da Computação e sua repercussão nas pesquisas da Filosofia da Mente. Procura conceituar e entender tal perspectiva à luz da Tecnociência e da Cibernética, tendo em vista os conceitos que buscam compreender a mente humana pela imitação de seu comportamento, aplicando os conceitos à investigação e simulação de diálogos: os chamados *chatterbots*.

A pesquisa parte de um estudo metodológico e bibliográfico diversificado dos conceitos de Inteligência Artificial (IA) e das concepções de Filosofia da Mente, enfocando a proeminência do tema em seus aspectos interdisciplinares. Discute as perspectivas consolidadas na comunidade acerca do tema, especialmente as abordagens de Searle (1984, 1992, 1997), Gardner (1994, 1995) e, no panorama brasileiro, de Teixeira (1990, 1994, 1995, 1998, 2000, 2004, 2008), incidindo nas relações entre os conceitos de intencionalidade, cérebro, experimento mental e cognitivismo, bem como algumas tendências críticas e de contestação de seus limites formais.

Valendo-se de pesquisas anteriores sobre o tema e de metodologia que mescla elementos computacionais e de filosofia da mente, analisa o experimento conceitual do quarto chinês de Searle (em seus aspectos de sintaxe e semântica). A partir deste experimento e seus resultados, apresenta a perspectiva da investigação da Inteligência Artificial (IA), como pertencente à natureza da cognição, apoiada na teoria da cognição, nomeadamente de fenômenos como categorização e identificação de objetos, resolução de problemas, decisão e consciência.

A pesquisa culmina na análise do conceito de Inteligência Artificial (IA), na Filosofia da Mente, proposto como um elemento noemático, constituindo-se como condição para a produção do conhecimento científico, aplicado este nas áreas de Tecnologia da Inteligência e Design Digital. Aplica os conceitos e resultados alcançados à definição e extensão do conceito de hipertexto, tomando-o como um mecanismo computacional capaz de estruturar diálogos hierarquizados, indexados, ao modo da formulação de Robot de Primo (2001) e Roth Coelho (2001).

Por fim, os pressupostos investigados são apresentados em resultados observados nos chamados *chatterbots* na área de Inteligência Artificial (IA), destacando suas características e sua importância no atual contexto de sua utilização computacional no ciberespaço.

Palavras chave: Inteligência, Inteligência Artificial (IA), Filosofia da Mente, Conhecimento, Cognição, Chatterbots.

Abstract

The current research has as central theme investigate the concept of intelligence, and specifically of Artificial Intelligence (AI), derived from the Computer Science and its effect on research of Philosophy of Mind. It tries to consider and to understand such perspective to the light of Technoscience and Cybernetics, in view of the concepts that look for to understanding the human mind by the imitation of its behavior, applying the concepts to the investigation and simulation of dialogues: called Chatterbots.

The research starts from a methodological and bibliographic study diverse from the concepts of Artificial Intelligence (AI) and the concepts of Philosophy of Mind, focusing on the prominence of the issue in its interdisciplinary aspects. Discusses the prospects consolidated in the community about the issue, especially the approaches of Searle (1984, 1992, 1997), Gardner (1994, 1995) and, in the Brazilian landscape, from Teixeira (1990, 1994, 1995, 1998, 2000, 2004, 2008), focusing on relations between the concepts of intent, brain, mental and cognitive experiment, as well as some trends of criticism and defense of its formal limits.

Drawing up of previous searches on the subject and methodology that merges computing elements and philosophy of mind, analyzes the conceptual experiment of the Chinese Room of Searle (in aspects of syntax and semantics). From this experiment and its results, it presents the prospect of Artificial Intelligence (AI) investigation as belonging to the nature of cognition, supported by the theory of cognition, particularly for phenomena such as categorization and identification of objects, problem solving, decision and conscience.

The search culminates in the analysis of the concept of Artificial Intelligence (AI) in Philosophy of Mind, proposed as a noematic element, being constituted as prerequisite for the production of scientific knowledge, implemented in the areas of Intelligence Technology and Digital Design.

Applies the concepts and achievements to the definition and extension of the concept of hypertext, taking it as a computational mechanism capable of structuring hierarchical dialogues, indexed, so as to the formulation of Robot Primo (2001) and Roth Coelho (2001).

Finally, the assumptions investigated are shown in results observed in the so-called Chatterbots in the area of Artificial Intelligence (AI), highlighting its characteristics and its importance in the current context of its computer utilization in cyberspace.

Key words: Intelligence, Artificial Intelligence (AI), Philosophy of Mind, Knowledge, Cognition, Chatterbots.

Introdução

Esta dissertação tem como tema central a investigação do conceito de Inteligência e, mais especificamente, de Inteligência Artificial (IA), com o propósito de resumir e compreender o estágio em que se encontram tais estudos e a possibilidade de se construir máquinas inteligentes.

Para tal empreendimento, valemo-nos de pesquisas anteriores sobre o tema, utilizando uma metodologia que combina elementos computacionais e de filosofia da mente. Pesquisamos os estudos realizados pelo cientista da Filosofia da Mente, João de Fernandes Teixeira, que nos explica a questão dos fenômenos mentais, e a importância das abordagens teóricas que contribuíram para compreendermos vários conceitos relevantes da Inteligência Artificial (IA) na Filosofia da Mente.

Dentre os diversos estudos, destacamos os princípios históricos da Inteligência Artificial (IA), o teste de Turing, a abordagem dos funcionalistas que defendiam a ideia de que a descrição das funções mentais tem o mesmo estatuto da descrição de um software, ou seja, de que estados mentais não são físicos, mas, sim, funcionais.

Desta maneira, também, tornou-se essencial entendermos os conceitos da abordagem cognitivista, cujo objetivo é estudar o processo de aquisição do conhecimento que procura entender e imitar os processos mentais dos seres humanos, no contexto da Inteligência Artificial (IA).

Fomos buscar também os estudos dos conexionistas que enfatizam o modelo de funcionamento do cérebro, dos neurônios e das conexões neurais.

Enveredamos pelos estudos sobre o desenvolvimento da inteligência de Piaget e de Gardner, de maneira a conceituar as suas mais diversas interpretações, que envolvem a aquisição do conhecimento e o processamento de informações.

Sendo assim, partimos do princípio de que a Inteligência Artificial (IA) pode ser interpretada como um tipo de inteligência produzida pelo homem, para dotar as máquinas de algum tipo de habilidade que simule a inteligência do próprio homem.

Abordamos os estudos de Searle sobre a Filosofia da Linguagem, Filosofia da Mente e Filosofia da Ciência. Searle foi um crítico da Inteligência Artificial (IA) forte, ciência essa que acreditava que poderiam se criar mentes artificiais. Searle, por sua vez, através de seu experimento da metáfora do quarto chinês, inferiu que sintaxe não é garantia da existência da semântica e refutou os teóricos da Inteligência Artificial (IA) forte e do funcionalismo.

A presente investigação também se preocupou em entender sobre robôs de conversação, os chamados *chatterbots*, mais especificamente da primeira robô da web brasileira, Cybelle, que demonstrou a possibilidade de interação com a máquina.

Além dela, procuramos conhecer e compreender também outros robôs de interação da web brasileiras, o robô Ed e a robô Sete Zoom cujo funcionamento é semelhante ao da Cybelle.

Porém, a implementação desses *chatterbots* levou-nos a diversos questionamentos, os quais permanecem sem respostas: - “Pode um robô pensar e interagir como um ser humano?” “A interação homem-máquina, através da linguagem natural, se iguala a um encontro interpessoal humano?” “O cérebro e ou a mente humana podem ser reproduzidos artificialmente?”

Todo esse percurso está aqui apresentado em quatro capítulos.

No primeiro capítulo, Fundamentos conceituais para uma abordagem da inteligência no ciberespaço, definimos os vários conceitos de inteligência e Inteligência

Artificial (IA), usando para isso ferramentas da web, mais especificamente da enciclopédia digital Wikipédia, e estudos de autores como Gardner e Piaget.

No segundo capítulo, Abordagens joaninas da Filosofia da Mente rumo à Inteligência Artificial (IA) no ciberespaço, apresentamos o percurso histórico da Inteligência Artificial (IA), que surgiu a partir de pesquisadores como Marvin Minsky, Hebert Simon e McCarthy, que acreditavam ser possível imitar o comportamento da inteligência humana, por meio de máquinas. Apresentamos ainda, neste capítulo, os estudos dos filósofos da mente, João de Fernandes Teixeira e Searle.

O terceiro capítulo, Inteligência Artificial (IA) ou entretenimento automático, fundamenta o princípio histórico da cibernética e sua trajetória, até chegar ao fascinante mundo do ciberespaço,

O quarto capítulo retoma alguns conceitos, discutidos em capítulos anteriores, para observar como uma máquina interage com um ser humano, através de um mecanismo de conversação, o *chatterbot*, ou seja, um robô virtual que simula conversação.

Capítulo I

1. Fundamentos conceituais para uma abordagem da inteligência no ciberespaço

Intitulamos o presente capítulo como “Fundamentos conceituais para uma abordagem da inteligência no ciberespaço”. Nele, é apresentado o conceito de inteligência a partir de sua forma dicionarizada, buscando investigar o mesmo conceito em autores, cuja abordagem direta faz do conceito de inteligência o centro de suas preocupações.

Optamos por eleger o modo wiki, como nosso centro norteador, por uma razão fundamental, que é o fato de nosso objeto situar-se dentro do universo digital do ciberespaço, concebido como uma hipermídia aberta e seguindo as sugestões, da Wikipédia.

“A inteligência significa faculdade que tem o espírito de pensar, conceber, compreender; capacidade de resolução de novos problemas e de adaptação a novas situações; discernimento; juízo, raciocínio; talento; compreensões fáceis, nítidas, perfeitas, profundas, de qualquer coisa; apreensão, percepção; pessoa de grande capacidade intelectual.”¹

Portanto, do ponto de vista metodológico, optamos por recorrer ao conceito na sua forma “wikificada”, a saber, a sua abordagem na Wikipédia, a enciclopédia, digital no ciberespaço. Não que com isso venhamos a desconhecer ou nos desinteressar pelas suas versões diversas encontradas em livros, autores e pesquisadores individuais e mesmo em dicionários técnicos especializados, na forma de livro papel.

O conceitual de inteligência apresentado pela Wikipédia pode ser definido quando se estimula a reflexão, em que a inteligência e a imagem são conceitos e aspectos da natureza humana.

¹ Disponível em: http://pt.wikipedia.org/wiki/inteligencia_artificial acesso em 05 Abril 2008.

Embora não tenhamos total conhecimento do conceito de inteligência, sabemos que a Psicologia é uma ciência que estuda o comportamento humano e outras questões como aprendizagem e inteligência, dentre outros fenômenos.

A Inteligência, de certa forma, envolve a criatividade. Alguns vinculam a palavra criatividade a artistas, músicos, arquitetos e publicitários. Na verdade, criatividade exige manipulação de idéias para criar algo novo e utiliza o mesmo esforço da organização mental.

A criatividade é essencial para resolver problemas de qualquer natureza, seja num poema ou numa prova de física. Nesse sentido, um dos principais requisitos para a manifestação inteligente é a faculdade de fazer associações, que, por sua vez, requer criatividade.

A capacidade para fazer justaposições com criatividade é um estilo de pensamento que pode ser considerado como a mais válida definição de pessoa inteligente, a facilidade de construir metáforas e ver relações não evidentes para os outros.

O pensamento inteligente envolve todos os meios de fazer conexões na vida cotidiana, de forma rápida e natural, tais como metáforas, analogias, formulação de hipóteses, imaginação de cenários, combinações, figuras de linguagem, simbolismo, visualização mental, trocadilhos e outras maneiras de conectar uma coisa à outra, analisando os significados e integrando o pensamento criativo e crítico.

Embora todos sejam dotados mentalmente para raciocinar com inferências e fazer conexões entre significados, metáforas, analogias e abstrações, a pessoa mais inteligente destaca-se pela sua velocidade, versatilidade e abrangência nos assuntos.

Porém, o estudo da inteligência envolve necessariamente a criatividade, que pode ser classificada de duas formas.

- **criatividade individual:** é a maneira natural do indivíduo conhecer suas idéias.

-**criatividade coletiva:** é a forma criativa de conhecer e organizar uma equipe ou grupo.

Entretanto, ainda não existe uma definição consensual sobre o que é inteligência, nem a certeza de como a adquirimos ou como a aperfeiçoamos.

A Inteligência Artificial (IA) fica de certa forma comprometida, visto que ainda não foram desvendados todos os mistérios do cérebro humano, muito menos de modelar o funcionamento da mente no computador. A idéia de modelar o funcionamento da mente está ligada à tentativa de simular a atividade dos neurônios. Para alguns cientistas, simulando a atividade dos neurônios, seria possível simular redes de neurônios e, com isso, simular a produção do pensamento.

De acordo com (Lima, 1999),

“Apesar de não haver uma definição aceita de forma unânime sobre a inteligência, aceita-se que o ser humano possui um tipo de inteligência característica, que se traduz pela capacidade de compreender e transformar o mundo à sua volta.”

Contudo, antes de modelarmos o funcionamento da mente no computador, precisamos entender como o cérebro humano funciona e como a inteligência é adquirida. Existe, sim, ainda uma questão a ser entendida. O que é inteligência? Embora tal questão não esteja resolvida, apresentaremos os conceitos de inteligência, de acordo com vários autores.

Sternberg, 1996, diz que a inteligência é a capacidade de solucionar problemas abstratos; um indivíduo é inteligente, na medida em que é capaz de pensar em termos

abstratos. Pode-se conceituar a inteligência como a capacidade que o ser humano possui de criar e ou modificar coisas; a capacidade de extrair a essência de um determinado conteúdo e, ainda, refletir e discutir sobre o mesmo.

Para Gardner (1995, p.21), “Uma inteligência implica a capacidade de resolver problemas ou elaborar produtos que são importantes num determinado ambiente ou comunidade cultural.” Ele identifica e conceitua diversos tipos de inteligência, que enumeraremos mais adiante.

Até pouco tempo atrás, a inteligência era medida somente a partir de uma série de testes psicométricos, como o teste de QI, sendo que este teste privilegiava o raciocínio matemático, e os indivíduos que possuíssem boa capacidade de memorização e raciocínio rápido seriam os mais qualificados.

Todavia, com relação a esse teste, Gardner (1994 a, p.13) argumenta que “[...] os testes apresentam poder de previsão para o sucesso acadêmico, mas relativamente pouco poder preditivo fora do contexto escolar.”

Sendo assim, pode-se dizer que a inteligência, é muito mais ampla do que os testes conseguem medir, até mesmo porque a maioria dos testes mede apenas o conhecimento lógico-matemático dos indivíduos, não conseguindo avaliar, por exemplo, sua capacidade de criação (Gardner, 1995).

Os testes de QI, não enfatizam nenhuma visão do processo da resolução de um problema, pedem apenas para se encontrar a resposta correta. Gardner, 1994a, coloca-nos ainda como limitação desse teste: o fato de ser composto de, na maioria das vezes, tarefas desconexas, distantes da vida cotidiana, e raramente avaliam a habilidade de assimilar novas informações.

Os estudos de Piaget surgiram como um contraponto à moda estabelecida pelos testes de inteligência. Para Piaget, o importante não são as respostas, mas sim as linhas de raciocínio desencadeadas para chegar até elas, inclusive as linhas que levam a uma conclusão errônea.

Para Piaget, o conhecimento não está no sujeito nem no objeto, mas ele se constrói na interação do sujeito com o objeto. Na medida em que o sujeito interage com os objetivos é que ele produz a capacidade de conhecer e produzir o próprio conhecimento.

Piaget afirma que existe um mecanismo básico de aquisição do conhecimento. Esse mecanismo consiste num processo de equilíbrio com dois componentes inter-relacionados: assimilação e acomodação. A assimilação refere-se à incorporação de novas informações aos esquemas já existentes e acomodação refere-se à modificação desses esquemas. A relação entre assimilação e acomodação é altamente interativa.

Porém, para Gardner (1994 a), atualmente a psicologia do processamento de informação está mais em voga entre os estudiosos da mente. O processamento de informações visa a descrever as etapas mentais envolvidas para a resolução de um problema (ou até mesmo quando não se encontra a solução ou encontra-se uma solução errada).

Assim, de acordo com o pesquisador:

“De fato, uma meta final da psicologia de processamento de informação é descrever, tão exaustiva e escrupulosamente quanto possível, as etapas em que o desempenho de um indivíduo possa ser simulado em um computador”. Gardner (1994 a, p.17).

Mas, de certa forma, existe um conceito errôneo de inteligência que é aquele que associa quantidade de conhecimento a inteligência elevada. Sendo assim, faz-se a suposição de que alguém que possui muito conhecimento automaticamente pode ser

considerado inteligente. Porém, o que torna alguém inteligente não é a quantidade de conhecimento que possui, mas, sim, a habilidade de adquirir novos conhecimentos, o que se pode chamar de aprender a aprender constantemente.

Nesse sentido, cabe aqui uma pergunta: É possível desvendar o funcionamento do cérebro, de forma a reproduzi-lo numa máquina para que ela possa criar coisas novas, assim como os seres humanos? Isso nos leva novamente à questão do processo criativo, pois como já vimos a inteligência envolve muito mais do que o pensamento lógico.

Segundo Damásio (1996, p.222),

“O cientista criador tem muito em comum com o artista poeta. O pensamento lógico e a capacidade são atributos necessários a um cientista, mas estão longe de ser suficientes para o trabalho criativo. Aqueles palpites na ciência que conduziram a grandes avanços tecnológicos não foram logicamente derivados do conhecimento pré-existente: os processos criativos em que se baseia o progresso da ciência atuam no nível do subconsciente”.

Damásio coloca inclusive que o juízo crítico, por exemplo, é uma perspectiva muito próxima da alma e não se pode supor que a alma tem como base o funcionamento cerebral do ser humano.

Porém, a grande questão que permanece em relação à inteligência é a possibilidade de existir um mecanismo que possa aumentá-la ou aperfeiçoá-la. Todavia, Gardner (1994 a), coloca que as abordagens de Piaget e as de processamento de informações são falhas no que tange à Biologia, ou seja, todas focalizam um determinado tipo de resolução de problemas lógicos e ou lingüísticos.

Gardner (1994 a, 1995) defende a abordagem dos sistemas simbólicos, abrangendo sistemas de símbolos mais amplos do que apenas os sistemas lógico-lingüísticos e numéricos, envolvendo também símbolos musicais, corporais, espaciais e até mesmo pessoais. Esta é a perspectiva da Teoria das Inteligências Múltiplas.

1.1 O que é Inteligência Artificial (IA)

Etimologicamente a palavra inteligência vem do latim *inter* (entre) e *legere* (escolher). Inteligência significa, então, aquilo que nos permite escolher entre uma coisa e outra. Na prática, inteligência é a habilidade de escolher, dentre várias possibilidades, aquela que permitirá realizar de forma eficiente uma determinada tarefa. A palavra artificial vem do latim *artificiale* e significa algo não natural, isto é, algo produzido pelo homem.

Portanto, Inteligência Artificial (IA) é um tipo de inteligência, produzida pelo homem, para dotar as máquinas de algum tipo de habilidade que simule a inteligência do próprio homem.

Segundo McCarthy (1956), a Inteligência Artificial (IA), pode ser definida como a ciência que estuda a emulação do comportamento da inteligência humana, por meio de máquinas.

No entanto, a definição de Inteligência Artificial (IA) mais abrangente e largamente aceita é a de Minsky, 1968. “Inteligência Artificial (IA) é a ciência de fazer máquinas fazerem coisas que requereriam inteligência, caso fossem feitas pelo homem”.

Para Durkin, 1994, a Inteligência Artificial (IA) é o campo de estudo da ciência que pesquisa formas de fazer o raciocínio do computador simular o raciocínio humano. Enquanto que Weber, 1998, conceitua a Inteligência Artificial (IA) como o ramo da Ciência da Computação, dedicado ao estudo das técnicas computacionais que representam alguns aspectos da cognição humana. Assim, pode-se perceber que a Inteligência Artificial (IA) esta baseada em duas idéias:

- O estudo do processo do pensamento humano, para entender o que é inteligência; e

A representação de tais processos via máquina (computador);

Segundo Teixeira (1990, p.10), “A mente humana funciona como um computador, e por isso o estudo de programas computacionais é a chave para se compreender alguma coisa acerca de nossas atividades mentais.”

A Inteligência Artificial (IA) fundamenta-se na idéia de que é possível modelar o funcionamento da mente humana através do computador. Lévy, 1998, diz que os fundadores da Inteligência Artificial (IA) foram Herbert Simon, John McCarthy e Marvin Minsky, que acreditavam firmemente que a inteligência é um mecanismo.

O cérebro, nessa visão, é uma máquina, e os neurônios são processadores de informação. Diante deste enfoque, de que o cérebro é uma máquina, seria perfeitamente possível modelá-lo dentro do computador.

Ora, o mundo em que vivemos o “mundo real”, é muito maior e mais complexo do que o micro mundo digital do computador, construído por nós mesmos. Entretanto, a Inteligência Artificial (IA) levanta várias questões, do tipo: Como ocorre o pensar? Como o homem extrai conhecimentos do mundo? Como a memória, os sentidos e a linguagem ajudam no desenvolvimento da inteligência? Como surgem as idéias? Como a mente processa informações e tira conclusões, decidindo por uma coisa, em vez de outra? Lima, 1999, comenta que essas questões são fundamentais para que se possa simular o raciocínio humano e para que a simulação do funcionamento da mente possa ser implementada no computador.

Um dos principais pontos de discussão, atualmente, é como desenvolver a capacidade de criar coisas novas. Há alguma técnica de Inteligência Artificial (IA), já existente, que possibilite ao computador criar?

Partindo da citação de Lévy (1998, p.110), verifica-se que isso ainda é possível.

“A Inteligência Artificial depara-se, na simulação do raciocínio científico, com o mesmo tipo de problema teórico, que em suas outras aplicações. Para programas serem autônomos, deveriam ter capacidade para aprender e aplicar uma hierarquia indefinida de conhecimento (metaconhecimento), capacidade essa à qual ainda não chegou nenhum programa. Por outro lado, as simulações automatizadas de raciocínio científico funcionam no marco de uma problemática fixa. Não se conhece nenhuma que saiba formular novas perguntas, embora isso pareça ser a principal vocação do pesquisadores.”

Marvin Minsky e Seymour Papert lançaram as bases para o surgimento do paradigma simbólico na Inteligência Artificial (IA); esse paradigma aborda a simulação da inteligência; não através da construção de hardware específico, mas no desenvolvimento de programas computacionais que operam sobre dados ou representações. Teixeira (1998).

Marvin Minsky, a partir de um de seus projetos, desenvolvidos no MIT, juntamente com Seymour Papert, desenvolveu a construção de uma máquina que pudesse ver o suficiente para ser capaz de usar mãos mecânicas a fim de resolver problemas do mundo real. Ele chega, então, à conclusão de que nenhum método sozinho é capaz de trabalhar bem. Minsky (1985, p.324). Para um robô diferenciar a forma de um objeto, por exemplo, não lhe basta utilizar a visão, precisa também de outros tipos de conhecimento.

Esse paradigma envolve a mudança de definição do que é inteligência. Em vez de definir a inteligência como a capacidade para solucionar problemas, passa-se a perceber a inteligência como resultante da representação mental, que não é nada mais do que uma atividade simbólica. O paradigma simbólico é que dá ênfase aos processos cognitivos.

A Inteligência Artificial (IA), também pode ser abordada do ponto de vista conexionista. Na abordagem conexionista, a ênfase é o modelo de funcionamento do cérebro, dos neurônios e das conexões neurais e está envolvida por uma série de

subáreas: Matemática, Biologia, Filosofia, Informática, Engenharia, Linguística e Psicologia.

De acordo com Teixeira, (1990, p.13),

“Mas foi realmente sobre a filosofia que o impacto da IA foi maior: criar uma máquina pensante significa desafiar uma velha tradição que coloca o homem e sua capacidade racional como algo único e original no universo. Mais do que isto, criar uma máquina pensante significa dizer que o pensamento pode ser recriado artificialmente sem que para isto precisemos de algo como uma “alma” ou outra marca divina.”

Entretanto, existem questões que preocupam os filósofos até hoje, como o problema mente e corpo (estados mentais), a questão do materialismo e do dualismo, cujas definições e comentários estaremos abordando no próximo capítulo, dentre outras questões.

Conforme Teixeira, J. de F. (1990, p.15),

“Se o pensamento humano pode ou não ser mecanizado, como pretendem os teóricos da IA, é uma questão que ainda permanece em aberto. Tudo dependerá ainda de realizações futuras e de algum tipo de consenso a que os filósofos ainda hesitam em chegar. Para se ter uma noção mais precisa do que a IA propõe como programa de pesquisa, é preciso saber um pouco de sua história, de suas realizações até agora, e saber, em linhas gerais, como funciona um computador, o que para muitos é ainda um mistério.”

Portanto, permanecemos com questionamentos acerca da possibilidade da construção de robôs pensantes e interação com seres humanos, que ainda nenhum estudioso pôde responder em definitivo.

1.2 O desenvolvimento da Inteligência Artificial (IA) como Disciplina

Embora exista um grande consenso entre os pesquisadores do campo da Inteligência Artificial (IA), de que esta ciência foi fundada em 1956, na famosa conferência de Dartmouth, o intento de reproduzir a inteligência data de muitos antes, desde a Antiguidade, havendo registros históricos disso na Literatura, na Engenharia e na Filosofia.

Na metade do século XX, existia um contexto propício para o florescimento da Inteligência Artificial (IA) como ciência. As Ciências Humanas e as Ciências da Computação forneciam ferramentas interessantes para esse surgimento.

A Inteligência Artificial (IA) nasceu como uma ciência interdisciplinar. A Filosofia, a Psicologia, a Antropologia e a Biologia, admitiam conceber o ser humano como um sistema complexo, por sua vez, a Matemática, a Teoria da Computação e a Cibernética mostravam como uma estrutura era capaz de processar informações. No encontro dessas concepções é que surgiu a Inteligência Artificial (IA).

Entretanto, a Inteligência Artificial (IA) logo se tornaria independente das disciplinas que lhe deram origem. Seus pesquisadores e estudiosos desenvolveram concepções próprias da inteligência, que foi se afirmando em torno de seus paradigmas clássicos: Baseando-se neste princípio clássico da Inteligência Artificial (IA) e conceituando sobre a abordagem Cognitivista e a abordagem Conexionista, as quais iremos fazer referências posteriormente.

Nos anos seguintes, as dificuldades, e a falta de resultados convincentes esfriaram aquele otimismo inicial dos pesquisadores da Inteligência Artificial (IA), pois reproduzir no computador os comportamentos inteligentes do ser humano era uma tarefa muitíssimo mais árdua do que se imaginava.

As pesquisas tornaram-se mais específicas e menos ambiciosas. Esse fechamento teve seu ápice na década de 1970, a utilidade prática foi adotada como um critério de verdade e levou os pesquisadores a uma busca por resultados concretos, tomando lugar dos sonhos das décadas de 1950 e 1960.

Como consequência, a Inteligência Artificial (IA) concebeu a maioria de seus cientistas contemporâneos ausentes das discussões epistemológicas e dos que duvidavam de tudo quanto aos sonhos colocados no passado. O lado positivo de tal postura foi que os conhecimentos adquiridos passaram a ser utilizados em larga escala, porém, desastrosamente, a Inteligência Artificial (IA) viu-se desvinculada de um projeto de compreensão da inteligência.

Esse processo permitiu o crescimento da Inteligência Artificial (IA) enquanto campo acadêmico, porém, afastou os pesquisadores das discussões relacionadas ao estudo dos processos cognitivos e inteligentes, e as próprias discussões sobre os fundamentos teóricos da disciplina.

Esse processo também provocou um distanciamento entre o conceito de inteligência utilizado na Inteligência Artificial (IA) e o conceito de inteligência humana que se desenvolveu nas Ciências Humanas, a partir das novas descobertas da Psicologia e das Neurociências.

Além disso, os paradigmas clássicos começaram a dar sinais de saturação. O desenvolvimento da Inteligência Artificial (IA) em cima das mesmas e velhas técnicas foi gradativamente mostrando limites, como em qualquer ciência.

Dessa maneira, a Inteligência Artificial (IA) precisou novamente abrir-se, e buscar inspirações para superar seus modelos clássicos e encontrar modelos alternativos. Nas décadas de 1980 e 1990 surgiram propostas de novos paradigmas, de

novas técnicas (muitas delas nascidas de hibridação das técnicas clássicas), e de uma busca por novas concepções teóricas, provenientes de um renovado contato com outras disciplinas.

No entanto, Helder Coelho, 1999, nos diz que:

“O colapso dos sonhos utópicos dos anos 60 deu oportunidade aos pragmáticos, os quais aproveitaram para, no fim da década de 70, afastar imediatamente a disciplina do trabalho sobre os fundamentos. (...) Mas, na medida em que se alargou o espectro de complexidade dos novos problemas reais a resolver, tornou-se imperioso regressar de novo à zona das áreas nucleares da Inteligência Artificial (IA).”

Porém, a inteligência de certa forma possui um caráter sistêmico e passível de estruturação. Por consequência disso, é válido almejar a construção artificial de sistemas inteligentes por meio de computação.

Logo, a emergência da inteligência em sistemas artificiais talvez seja possível com uma arquitetura integradora de diversos mecanismos, os quais, por si só, ainda necessitam de definição e esclarecimento.

1.3 Breve conceito de Ciência Cognitiva

A Ciência Cognitiva é formada por diferentes disciplinas que têm como objetivo estudar o processo de aquisição de conhecimento. A Ciência Cognitiva, por sua vez, é também o campo de estudos que procura entender e imitar os processos mentais dos seres humanos, no contexto da Inteligência Artificial (IA).

Essa ciência tenta modelar a aquisição de informações feita pelo homem na busca do entendimento do mundo, que é subjetivo e depende de vários fatores, quais sejam:

- Experiências conscientes e inconscientes.
- Estímulos sensoriais externos.
- Contexto sócio-cultural.

Entretanto, a Ciência Cognitiva Contemporânea - o representacionismo e a Inteligência Artificial (IA) desenvolvida pelo MIT, nos anos 1970, herdou os pressupostos da teoria clássica da representação.

Este tipo de Ciência Cognitiva que, em grande parte, prevalece até hoje, desenvolveu uma visão da cognição e do chamado “modelo computacional da mente” em que ambos são definidos como computações de representação simbólica, muito próximas da noção de idéia cartesiana e, por isso mesmo, pressupõem o “fantasma na máquina”, a mesma pressuposição que fazia com que Descartes reconhecesse as limitações dos autômatos que os impediriam de terem uma vida mental semelhante a nossa, mas em que a tecnologia pudesse avançar. (Teixeira, 2004).

Essas são as principais dificuldades conceituais encontradas nas tentativas de formulação de modelos cognitivistas da consciência: uma máquina poderia fazer tudo o

que um ser humano faz, mas nunca poderia adquirir consciência do que está fazendo (Turing, 1950 apud Teixeira, 1995).

No entanto, na visão do funcionamento mental, defendida por Minsky (1985 apud Teixeira, 1995), a idéia de consciência aparecia pelo fato de que o sistema possuiria um módulo superior ou controlador dos outros, que coordenaria as atividades do sistema como um todo, dando-lhe unidade.

Esse módulo superior, para efetuar controle, teria acesso (externo) aos outros módulos e conteria em si o modelo do sistema - e é este último aspecto que o equipararia a uma idéia de consciência.

Essa concepção da consciência como módulo controlador enfrenta todas as dificuldades geradas pelo regresso infinito, afinal, o que monitora o módulo controlador? Será que podemos encontrar uma semelhança de propriedades entre o módulo controlador e aquilo que habitualmente chamamos de consciência? Este módulo superior não poderia ser conhecido, ultrapassando o escopo da investigação realizada pela Ciência Cognitiva (Teixeira, 1995).

Porém, a idéia de conhecimento, como representação dos sistemas especialistas da década de 1970, parece estar na raiz da dificuldade tecnológica aparente envolvida na construção destes sistemas: explosão combinatória, rigidez de estrutura e assim por diante (Teixeira, 2004).

Pelo menos, é o que se pode observar que, enquanto a filosofia neste último século tentou derrubar a noção de representação e evitar o mentalismo nas suas concepções sobre o conhecimento, a Ciência Cognitiva nas últimas décadas tentou se consolidar fundamentando-se na noção de representação (Teixeira, 2004).

Porém, para os filósofos, esta noção sempre constituiu um dos aspectos mais problemáticos a serem enfrentados pelas teorias da cognição.

A Ciência Cognitiva procurou encontrar suas definições de conhecimento, representação, inferência, etc com a filosofia analítica, porém, a própria filosofia analítica, nos últimos anos, veio colocando em questão suas noções tradicionais de representação e conhecimento.

Portanto, para incorporar os resultados da reflexão filosófica contemporânea, a Ciência Cognitiva teria de situar-se para além da noção tradicional de representação, deixando de imaginar a representação como fundamento da cognição e passando a tê-la como um fenômeno que ocorre no mundo.

Nossa representação do mundo é parte do mundo e não sua condição de possibilidade – isto parece ser a grande reversão operada pela filosofia contemporânea na sua crítica às teorias que partem de uma noção tradicional de representação (Teixeira, 2004).

A criação da Ciência Cognitiva apresenta-se, assim, como um estímulo à filosofia, permitindo definir questões científico-cognitivas fundamentais, de uma forma coerente.

1.4 Paradigmas clássicos de abordagens conexionista e cognitivista

Os modelos teóricos acerca do conhecimento e da inteligência foram organizados no decorrer do século XX, dentro e ou a partir de paradigmas. De acordo com Thomas B. Kuhn, um paradigma consiste em: “[...] considero paradigmas as realizações científicas universalmente reconhecidas que durante algum tempo fornecem problemas e soluções modelares para uma comunidade de praticantes de uma ciência.” (Kuhn, 1975:13) ²

Para o presente escopo de nossa dissertação, trataremos das abordagens conexionista e cognitivista, embora sem nos aprofundarmos, organizadas ao modo de paradigmas clássicos na questão da Inteligência Artificial (IA).

Abordagem Conexionista

Denominada também de biológica ou ascendente, enfatiza o modelo de funcionamento do cérebro, dos neurônios e das conexões neurais. Os pioneiros dessa corrente foram McCulloch, Pitts, Hebb, Rosenblatt e Windrow.

Os primeiros modelos de redes neurais artificiais surgiram em 1943, quando foi possível sua representação e formalização matemática. No entanto, a corrente conexionista sofreu grande impacto quando os cientistas Marvin Minsky e Seymour Papert publicaram, em 1969, o livro “Perceptrons”, no qual criticavam e sustentaram que os modelos das redes neurais não tinham sustentação matemática suficiente para que se lhes pudesse atribuir alguma confiabilidade.

Apesar das pesquisas nesta área terem continuado, foi somente na década de 1980 que o físico e biólogo do instituto de Tecnologia da Califórnia, John Hopfield conseguiu recuperar a credibilidade da utilização das redes neurais.

²Disponível em: http://eis.bris.ac.uk/~plajb/research/papers/Kuhn_for_DLB.pdf acesso 13 Fevereiro 2009.

O conhecimento de um sistema conexionista é armazenado de forma distribuída no sentido de que a informação não é estruturada em blocos centralizados, mas sim difundida por toda a rede. As redes neurais não trabalham com símbolos como aqueles definidos nos sistemas simbólicos.

Ao invés disso, trabalham com conceitos que são representações categóricas adquiridas pela experiência formadas através da adaptação a características invariáveis dos estímulos de entrada.

Um aspecto fundamental dessas representações é a ausência de uma explicação semântica. Isso significa que, na prática, seres humanos não podem construir nem interpretar representações de redes neurais de uma forma precisa e significativa, como ocorre com as representações simbólicas.

Abordagem cognitivista

Também denominada de descendente ou simbolista, dá ênfase aos processos cognitivos, ou seja, à forma como o ser humano raciocina. Objetiva encontrar uma explicação para comportamentos inteligentes, baseada em aspectos psicológicos e processos algorítmicos. Os pioneiros dessa corrente são McCarthy, Marvin Minsky, Newell e Simon.

Um sistema simbólico é como um processador de símbolos, onde tanto a entrada como a saída é composta por símbolos. A abordagem simbólica da Inteligência Artificial (IA) concentra-se em desenvolver agentes inteligentes baseados em tais sistemas.

Os sistemas simbólicos representam o conhecimento explicitamente através de símbolos, estruturas de símbolos e regras simbólicas. Em um agente simbólico, essas estruturas são projetadas desde o início para atender especificamente ao problema

tratado. Esse é o nível de meta-conhecimento, o conhecimento sobre o conhecimento a ser representado.

Devido a essa característica, os problemas abordados por modelos simbólicos precisam estar bem definidos antes de serem resolvidos.

Uma outra característica dos sistemas simbólicos é o uso de representação localizada do conhecimento. Isso significa que o conhecimento necessário para o agente fica disponível em blocos simbólicos que são acessados diretamente, de maneira centralizada. Essa é, na verdade, a essência dos símbolos: representar conceitos do mundo de forma íntegra e bem definida.

É preciso ressaltar ainda que as representações simbólicas são semanticamente interpretáveis. Embora os símbolos, por si só, não possuam significado próprio, como visto anteriormente, aos sistemas por eles compostos pode ser atribuída uma interpretação.

A Inteligência Artificial (IA) simbólica possui um rico conjunto de técnicas bem definidas de representação do conhecimento. Podemos, por exemplo, citar: lógica formal, frames, redes semânticas, scripts, e uso de regras explícitas.

O importante a enfatizar é que todas elas são intercambiáveis, em princípio, ou seja, uma representação simbólica pode fazer tudo o que qualquer outra representação simbólica faz. Isso ocorre porque, no fim das contas, toda representação computacional resume-se a um denominador comum – bits, na memória.

As primeiras modelagens da Inteligência Artificial (IA) surgiram na década de 50 e tiveram como base de produção os algoritmos e a lógica dos predicados. A formalização da lógica facilitou o processo de formalização e representação dos conhecimentos a serem utilizados pelos programas de computador. Inicialmente, esses

conhecimentos se restringiram a esquemas de raciocínios para jogos, aplicações matemáticas e simuladoras.

Essas modelagens basearam-se na hipótese do sistema de símbolos físicos, segundo a qual, um conjunto de estruturas simbólicas e um conjunto de regras de manipulação dessas estruturas são os meios necessários e suficientes para se criar inteligência.

O sistema cognitivo atua sobre os símbolos, estes, por sua vez, atuam sob uma perspectiva sintática, e não semântica.

A explicação cognitivista é puramente computacional, associada com o rápido aparecimento dos computadores, na década de cinquenta. A natureza do conhecimento humano está afastada da consciência, está afastado do “eu“ do indivíduo, desta forma é considerado um fenômeno, para os cognitivistas.

1.5 Críticas filosóficas e a argumentação de uma Inteligência Artificial (IA) forte

Os críticos da Inteligência Artificial (IA) forte, dos quais podemos destacar John Searle, Hubert Dreyfus, dentro outros, inseriram no debate questões de ordem filosófica e epistemológica, questionando qualquer possibilidade efetiva da Inteligência Artificial (IA) forte. Seriam falsos, assim, os próprios pressupostos da construção de uma inteligência ou consciência semelhante à humana em uma máquina.

Searle é bastante conhecido por seu argumento sobre o quarto chinês (ou sala chinesa), o qual estaremos tratando mais detalhadamente, no próximo capítulo.

Searle inverte a questão colocada por Minsky a respeito do Teste de Turing. Seu argumento diz que ainda que uma máquina possa parecer falar chinês, por meio de recursos de exame comparativo com mostras e tabelas de referência, binárias, isso não implica que tal máquina fale e entenda efetivamente a língua.

Ou seja, demonstrar que uma máquina possa passar no teste de Turing, não necessariamente implica que ela seja um ser consciente, tal como entendido no sentido humano.

Porém, Dreyfus, 1997, em seu livro “O que os computadores ainda não conseguem fazer”, faz uma crítica ao raciocínio artificial e argumenta que a consciência não pode ser adquirida por sistemas baseados em regras ou lógica; tampouco por sistemas que não façam parte de um corpo físico.

No entanto, Dreyfus, deixa em aberto a possibilidade de um sistema robótico alcançar a Inteligência Artificial (IA), baseado em Redes Neurais, ou em mecanismos semelhantes.

O primeiro neuro computador a obter sucesso (Mark I Perceptron) surgiu em 1957 e 1958, criado por Frank Rosenblatt, Charles Wightman e outros.

Entretanto, ultimamente, os revezes na interpretação da Inteligência Artificial (IA) vêm contribuindo para a relativização de todo o seu valor anteriormente determinado.

Definir uma concepção de inteligência humana é uma tarefa árdua e a possibilidade de se criar uma base plástica, artificial, em que possa ser gerada inteligência, deverá resultar numa inteligência necessariamente diferente da humana, na medida em que se compõe de elementos totalmente diversos dos encontrados em seres humanos.

A inteligência Artificial (IA), tal como a entendemos, pode ser vista como o fruto do cruzamento de uma base biológica com um complexo simbólico e cultural, talvez impossível de ser reproduzido artificialmente.

Retomando alguns pontos vistos anteriormente, pode-se dizer que tanto Dreyfus assim como Searle, desacreditam da possibilidade de uma Inteligência Artificial (IA) forte; quer dizer, desacreditam da construção de uma inteligência ou consciência semelhante à humana em uma máquina.

No que se refere aos cognitivistas, eles apresentam uma abordagem que representa o conhecimento através de símbolos, estruturas de símbolos e regras simbólicas. Para eles, o sistema cognitivo atua sobre os símbolos, porém estes atuam sob uma perspectiva sintática e não semântica, não possibilitando, portanto, a construção de uma consciência. .

Capítulo II

2. Abordagens joaninas da Filosofia da Mente rumo a Inteligência Artificial (IA) no ciberespaço

O presente capítulo aborda e resume alguns questionamentos históricos da Inteligência Artificial (IA) na Filosofia da Mente. Para isso, estaremos utilizando uma metodologia baseada em tecnologias da inteligência, com representação cognitiva no ciberespaço, investigando a Inteligência Artificial (IA) como pertencente à natureza da cognição e de fenômenos como categorização, fenômenos mentais, consciência e cérebro.

Pesquisamos também a exploração do uso da tecnologia cognitiva no ciberespaço, como facilitadora do conhecimento e reelaboração de estratégias com fins táticos de pesquisa.

Discutimos ainda mais demoradamente as questões da Inteligência Artificial (IA) na Filosofia da Mente, perguntando-nos: “Pode uma máquina pensar?”

Diante dessa pergunta e outras que surgiram ao longo do trabalho, sentimo-nos motivados e incentivados a buscar respostas, juntos aos filósofos da mente, mais especificamente, João de Fernandes Teixeira e Searle.

2.1 A natureza do mental

Definir o conceito de Filosofia da Mente é tarefa nada fácil. Desde tempos remotos, os filósofos estudam os fenômenos da mente e como representá-los. Ainda hoje, nem mesmo os cientistas advindos de outras áreas conseguiram nos dar uma resposta precisa do que vem a ser a mente e os fenômenos mentais. O objetivo deste capítulo é somente conceituar algumas dessas questões; não pretendemos nos aprofundar em outras ramificações que a filosofia se desafia a investigar.

Embora a Psicologia, a Neurofisiologia, a Metafísica, e a própria Filosofia, tenham feito pesquisas que contribuíram enormemente na tentativa de desvendar os fenômenos mentais, pode-se dizer que eles são característicos da natureza do próprio homem, são únicos. Por mais que a ciência tenha evoluído, permanecem ocultos na mente do ser humano.

A partir de 1940, a questão dos fenômenos mentais ganhou grande destaque, por parte dos filósofos, devido ao surgimento das grandes evoluções tecnológicas, o computador digital, que passou a ser comparado ao cérebro humano, o que gerou grande polêmica entre os filósofos, fazendo com que se perguntassem “pode uma máquina pensar?” Ora, diante tantos questionamentos, pensamos que os estudiosos mais avalizados para dar respostas são os Filósofos da Mente.

2.2 O problema mente - cérebro

O filósofo cartesiano Descartes, foi pioneiro do pensamento que mente e corpo são duas coisas diferentes. Descartes ativou a polêmica com o monismo, também chamado de materialismo, que sustenta a teoria de que os fenômenos mentais são idênticos, pois mente e cérebro são a mesma coisa, ou seja, os fenômenos mentais são iguais aos fenômenos físicos.

Porém, Descartes é defensor da teoria dualista, também chamada de espiritualismo, que sustenta que mente e cérebro não são a mesma coisa. Para justificar sua teoria, ele criou um verdadeiro guia de anatomia humana, como até hoje é considerado, porém suas justificativas não são muito claras. Nessa mesma época, surgiram outros filósofos que refletiram e estudaram com profundidade a natureza da alma (ou da mente).

Mas, somente a partir do século XX é que a Filosofia da Mente surgiu, fazendo com que retornassem as questões dos fenômenos mentais e das teorias metafísicas, principalmente com a relação mente-cérebro, que voltou a adotar o monismo materialista, devido às pesquisas dos neurofisiologistas.

Um dos mais notáveis filósofos da época, final da primeira metade do século XX, Gilbert Ryle, dizia que o problema mente corpo não deveria sequer ser considerado um problema; assim, vieram ainda outros filósofos sustentar suas idéias. Entretanto, não há como afirmar qual das teorias deve-se seguir: monismo materialista ou dualismo espiritualista. Talvez aqui tenhamos mais um questionamento para ser respondido pela Filosofia da Mente.

2.3 Maquinismo e Consciência

Alan Turing, matemático que ainda hoje é considerado uns dos grandes estudiosos como referência para as pesquisas nas áreas da Ciência da Computação e da Inteligência Artificial (IA), em resposta à famosa questão “pode uma máquina pensar?” argumenta que isso só poderá ser possível quando tivermos uma definição exata do que vem a ser “pensamento” e “máquina”.

Afinal, em que se baseia o teste de Turing?

Baseia-se no “jogo da imitação”. A questão é saber se o comportamento humano é igual ao comportamento da máquina.

Nessa experiência, uma pessoa, isolada, faz uma série de perguntas para serem respondidas por um computador e por uma outra pessoa. O indivíduo que faz as perguntas deverá descobrir qual dos interlocutores é a máquina e qual é o ser humano.

A idéia geral de Turing é que uma máquina se igualaria a um ser humano caso seu comportamento fosse indistinguível deste. Esse tipo de jogo tem sido um desafio para os pesquisadores da Inteligência Artificial (IA).

De acordo com Teixeira, J. de F. (2008, pp.58-59),

“[...] até hoje nada melhor do que esse teste apareceu para fornecer critérios a seres artificiais. Se uma máquina puder conversar, ela deverá ter consciência do que fala. Conversas são indícios de mente e consciência.”

Desta forma, sendo impossível distinguir o comportamento da máquina do comportamento de um ser humano, seria correto atribuir à máquina pensamentos e estados mentais?

Se dissermos que seres humanos pensam, e se constatarmos que a máquina imita perfeitamente o comportamento humano, deduziremos que esse é um critério que nos permite atribuir pensamentos à máquina.

A Inteligência Artificial (IA) teve como objetivo construir máquinas pensantes e programas computacionais que pudessem imitar as atividades humanas mentais. A partir da década de 1960, pode-se dizer que as duas grandes realizações que a Inteligência Artificial (IA) exibiu foram os cálculos complexos e os jogos de xadrez, dentre outros, o que trouxe um grande interesse aos filósofos, criando estes uma teoria chamada de Funcionalismo, acrescentando agora também o estudo dos sistemas artificiais.

Os funcionalistas defendiam a idéia de que a descrição das funções mentais tem o mesmo estatuto da descrição de um software, ou seja que os estados mentais não são físicos, mas, sim, funcionais, com certos comportamentos sensoriais de entrada (input) e outros estados mentais que são a saída (output).

Ora, segundo os funcionalistas, o pensamento não é somente de uso exclusivo dos seres humanos. Desde que uma máquina dotada de sistema (programas de computadores) possa executar as mesmas tarefas que um ser humano, talvez seja possível atribuir-lhe pensamentos.

Conforme explica Teixeira. J. de F. (1998, p.45), na perspectiva do funcionalismo,

“[...] se computadores são um tipo especial de arranjo material, ou seja, uma combinação de elementos materiais de silício ou de qualquer outro elemento da natureza, e se eles puderem realizar tudo o que uma mente humana realiza, não haveria nenhuma razão para supor que mente e matéria são diferentes. Poderíamos igualar mentes e máquinas, cérebro e mentes.”

De acordo com Teixeira, J. de F. (1998, p.49), e ainda na perspectiva do funcionalismo, cabe dizer que:

“Um marciano pode ter um sistema nervoso completamente diferente do meu, mas se o sistema nervoso desse marciano puder executar as mesmas funções que o meu, o marciano terá uma vida mental igual à minha.”

Por sua vez, o funcionalismo também adota determinadas regras e algumas estratégias. Sendo assim, faz suas pressuposições básicas, conforme citadas abaixo:

- a) A realidade dos estados mentais, de certa forma, é constituída de regras e estratégias;
- b) estados mentais não podem reduzir-se a estados físicos, não importando de que matéria seja feitos e, sim, às suas funções, regras e estratégias;
- c) estados mentais geralmente são designados a um papel de input e output, sendo necessário utilizar de estratégia e raciocínio lógico, que somente um ser humano possui.

Todavia, Teixeira, J. de F. (2000, p. 125) diz:

“[...] a partir dessa pressuposição, o funcionalismo pode rejeitar o dualismo substancial e propor uma teoria neutra, nem materialista, nem dualista. Na realidade, o funcionalismo é compatível com o dualismo e com o materialismo não reducionista.”

De certa forma, o funcionalismo busca uma compatibilidade com o materialismo não reducionista. Ora, o quem vem a ser o materialismo não reducionista na concepção da psicologia? Pode-se dizer que identificar a mente com entidades ou propriedades físicas é uma forma direta de materialismo, assim como alegar que a psicologia é redutível à biologia e que neurofisiologia é redutível à física.

Porém, se for mostrado que todos os eventos mentais ou psicológicos são redutíveis à neurofisiologia, e, por sua vez, que a neurofisiologia é redutível (talvez através da química) à física, então será mostrado que a mente não é nada além daquilo que é físico ou corpóreo.

Sendo assim, a redução é realizada em dois passos:

1-Redução da linguagem de uma área de estudos à linguagem de outra área de estudos.

2-Alegação de que a área de estudos reduzida é idêntica à área de estudos redutora.

Ademais, de acordo com a pressuposição do funcionalismo, anteriormente explicada, dessa forma, ele passa a ser de grande interesse da Filosofia da Mente. Contudo, todas essas informações dão-nos uma chave para abirmos novos caminhos para a tecnologia da Inteligência Artificial (IA).

De acordo com Teixeira, J. de F. (2000, p.125)

“A possibilidade de variar o tipo de material, com o qual podemos construir mentes tem dado ao funcionalismo um lugar de destaque na filosofia da mente nas últimas décadas. Essa possibilidade tem servido de fundamento para as realizações tecnológicas oriundas de disciplinas novas como é o caso da Inteligência Artificial. A noção de uma Inteligência Artificial como realizações de tarefas inteligentes – ou seja, a possibilidade de replicação mecânica de segmentos da atividade mental humana - por dispositivos que não têm a mesma arquitetura nem a mesma composição biológica e físico-química do cérebro foi a grande motivação para o aparecimento das teorias funcionalistas.”

Ainda dentro desse contexto, existe uma analogia feita com os estados mentais, em que a mente pode ser considerada um software do cérebro. Entretanto, um software estipula um conjunto de regras lógicas, sendo que a atribuição dos estados mentais é feita por meio de observações de comportamentos (behaviorismo), nos quais se acredita que exista certa ligação lógica entre comportamentos e estados mentais.

Sendo assim, poderíamos supor que um dia talvez exista a possibilidade, de revelar um tipo de software ou sistema operacional existente na mente humana, possibilidade esta investigada pelos os pesquisadores da Inteligência Artificial (IA) desde 1970, por acreditarem que comportamentos humanos poderiam ser descritos em um programa de computador.

Ora, o teste de Turing coloca uma situação que nos permite pensar nos filmes de ficção científica em que se convive com robôs quase humanos, digamos, “replicantes”, embora tenham um “pensamento” e “consciência” que funciona somente com uma determinada programação (linguagem de computador), limitada e projetada pelo homem.

Portanto, é fácil observar “o pensamento” de um robô, pois ele apenas responde a perguntas e respostas previamente elaboradas. Sendo assim, coloca-se em questão a viabilidade ou não da construção do robô pensante, cujo estudo faz parte das investigações da Filosofia da Mente.

Segundo Teixeira, J. de F. (1994, pp.39-40)

“Mas se não sabemos o que é pensamento nem tampouco o que é consciência, isto não significa que tentar abordar tais problemas mediante a investigação da Inteligência Artificial seja uma tarefa fadada ao fracasso. É muito difícil profetizar negativamente sobre aquilo que a ciência e a tecnologia poderão vir a descobrir algum dia.”

Sabemos que o jogo de imitação de Turing tinha por objetivo mostrar como máquinas inteligentes poderiam reproduzir com perfeição o comportamento humano, a ponto de se tornarem indiscerníveis, uma do outro. Ou seja, à percepção de um observador humano, não seria possível dizer qual a ação teria sido determinada por um algoritmo ou se fora fruto do livre arbítrio de mente humana.

John Searle, filósofo da linguagem, considera o aspecto intersubjetivo do jogo de imitação como insuficiente para definir se o programa embutido na máquina compreende ou não aquilo que responde ao interrogador.

Em 1980, Searle escreveu o artigo “Mente, Cérebro e Programas”, no qual se posiciona criticamente acerca das pesquisas que visavam esclarecer os processos mentais, como causas do comportamento das partes do cérebro, ou seja, a teoria da

Inteligência Artificial (IA) fraca, que considera o computador como ele é, ou seja, uma máquina.

Para Searle, a Inteligência Artificial (IA) fraca permite a harmonia entre as abordagens biológicas e o que se sabe sobre o mundo, enquanto a Inteligência Artificial (IA) forte, ao tentar criar mentes artificiais, estaria condenada ao fracasso, pois nenhum programa de computador seria suficiente para fornecer um sistema inteligível.

Porém, a Filosofia da Mente é tão minuciosa em suas investigações que, de certa maneira, tenta nos mostrar vários caminhos para refletir sobre o que é mente e os fenômenos mentais. Junto a ela, tentaremos buscar alguma resposta para nossos questionamentos. Com o surgimento de novas disciplinas, como a Inteligência Artificial (IA) e a Ciência Cognitiva, os filósofos se reanimaram e voltaram a debater os problemas da natureza da representação mental.

Segundo Teixeira, J. de F. (1998, p.44),

“[...] inteligência resulta da representação mental, e esta nada mais é do que atividade simbólica. O que nos distingue de outros animais menos inteligentes é nossa capacidade de produzir e manipular símbolos. Este é o real caráter distintivo da inteligência humana: a produção e manipulação de símbolos que dão origem às atividades cognitivas superiores, como a Matemática e a linguagem.”

Os pesquisadores da Inteligência Artificial (IA) dedicam-se a vários estudos, dentre eles, a simulação de comportamento inteligente. A Inteligência Artificial (IA) pode ser considerada como um método laboratorial muito complexo, por meio do qual são testadas hipóteses através dos sistemas causais internos, responsáveis pela cognição humana (conhecimento).

Conforme Teixeira, J. de F. (1998, p.134),

“[...] a simulação do comportamento inteligente deve ter como ponto de partida os comportamentos simples, mundanos, que não requerem a existência prévia de representações. Isto constitui uma guinada radical em relação ao estatuto de representação, que passa a ser visto como um fenômeno tardio na ordem vital. A

cognição não se inicia com a representação e sim com a interação do organismo com o seu ambiente onde dois fatores são fundamentais: a percepção e a locomoção.”

Depreendemos dessa fala que, necessariamente, o homem faz a leitura do mundo antes de criar as suas representações.

2.4 Breve histórico da obra de John Searle

O filósofo norte americano John Searle é dono de uma das obras filosóficas mais importantes de nosso tempo. Nascido em 1932 e ligado à tradição analítica da filosofia contemporânea, começou sua produção filosófica abordando temas de Filosofia da Linguagem que, progressivamente, se ampliaram para abranger tópicos da Filosofia da Mente e da Filosofia da Ciência.

Suas principais obras foram a teoria da intencionalidade, a nova teoria da consciência, teses da construção da realidade social, críticas ao computacionalismo, dentre outras. Dentre todas essas abordagens, faremos alguns comentários sobre a questão da teoria da intencionalidade e a crítica ao computacionalismo.

Entretanto, vamos primeiro definir o que vem a ser o computacionalismo, contra o qual, Searle se bate e do qual o cognitivismo deve ser distinguido. Pode-se chamar de computacionalismo a crença de que os processos mentais consistem essencialmente em manipulação simbólica lógica, ou seja, na computação.

Para Ulrich Neisser, 1967, geralmente ocorre uma confusão entre o computacionalismo e cognitivismo. O cognitivismo é oriundo da Psicologia Cognitiva contemporânea que, por sua vez, possibilitou o advento do computador, devido à simulação dos processos cognitivos em máquinas.

Desde o surgimento da teoria matemática da “máquina de Turing”, até o desenvolvimento dos computadores de processamento paralelo, chamados de “metáfora computacional”, a dádiva mais importante para o computacionalismo, foi obtida a clareza conceitual do que veio a ser o hardware e o software e permitindo, desta forma, a teorização sobre a mente em relação ao cérebro. Quanto à questão do computacionalismo, estaremos discutindo-o detalhadamente mais adiante.

Porém, a idéia introduzida por Hilary Putnam, 1961, era que o cérebro poderia ser comparado com o hardware e a mente como o software. Esta idéia repercutiu muito durante o século XX e este, por sua vez, sedimentou um programa de pesquisa, que logo seria conhecido pelo termo de Inteligência Artificial (IA). Seu objetivo era projetar um computador e um programa que operasse de forma idêntica a um ser humano, o que nos levaria a compreender nossos processos cognitivos, e marcaria, desta forma, o nascimento da tese da Inteligência Artificial (IA) forte.

Embora já tenhamos explicado anteriormente o “Teste de Turing”, vale ressaltar que durante muito tempo foi visto como uma espécie de crença religiosa na Inteligência Artificial (IA) forte. Acreditava-se que ele estava plenamente apto a dizer o contrário a qualquer um que duvidasse que uma “máquina pode pensar”.

Entretanto, Searle defendia a teoria da intencionalidade, que pode ser entendida como a propriedade da consciência, segundo a qual todo o estado consciente refere-se a algo diverso da própria mente. Ora, consciência é sempre consciência de alguma coisa, esse é o sentido de se dizer que a consciência é intencional. Por exemplo: quem ama, ama alguém; quem representa, representa alguma coisa; e assim por diante.

Nessa teoria da intencionalidade, Searle elabora conceitos que talvez hoje sejam fundamentais para a Filosofia da Mente. Todavia, a Filosofia da Mente veio trilhando um caminho natural que, partindo de um problema específico da Filosofia da Linguagem e passando pela teoria dos atos de fala, chegou à teoria geral da intencionalidade.

Esta teoria geral da intencionalidade começou a ser exposta numa resposta às pretensões da Inteligência Artificial (IA) forte. O argumento hipotético, criado por Searle em 1980, “metáfora do quarto chinês”, o experimento mental mais célebre de

nosso tempo, tem por objetivo refutar os teóricos da Inteligência Artificial (IA) forte e do funcionalismo. Searle baseia-se no fato de que a sintaxe (gramática) não é garantia da existência da semântica (sentido).

2.5 A experiência do quarto chinês

A experiência do quarto chinês é apresentada e trabalhada por Searle com o intuito de refutar os teóricos da Inteligência Artificial (IA) forte, que acreditavam ser possível criar seres artificiais. Ele a sistematiza em seu livro “Mente, Cérebro e Ciência” do qual extraímos o texto a seguir:

“Bem, imaginemos que alguém está fechado num quarto e que neste quarto há vários cestos cheios de símbolos chineses. Imaginemos que alguém, como eu, não compreende uma palavra de chinês, mas que lhe é fornecido um livro de regras em inglês para manipular os símbolos chineses. As regras especificam as manipulações dos símbolos de um modo puramente formal em termos da sua sintaxe e não da sua semântica. Assim a regra poderá dizer: “Tire do cesto número um símbolo esticado e ponha-o junto de um símbolo encolhido do cesto número dois.” Suponhamos agora que alguns outros símbolos chineses são introduzidos no quarto e que esse alguém recebe mais regras para passar símbolos chineses para o exterior do quarto. Suponhamos que, sem ele saber, os símbolos introduzidos no quarto se chamam “perguntas” feitas pelas pessoas que se encontram fora do quarto e que os símbolos mandados para fora do quarto se chamam “respostas às perguntas”. Suponhamos, além disso, que os programadores são tão bons a escrever programas e que alguém é igualmente tão bom em manipular os símbolos que muito depressa as suas respostas são indistinguíveis das de um falante chinês nativo. Lá esta ele fechado no quarto manipulando os símbolos chineses e passando cá para fora símbolos chineses em resposta aos símbolos chineses que são introduzidos. Com base nesta situação tal como escrevi de nenhum modo se pode aprender chinês pela simples manipulação desses símbolos formais. Ora, o cerne da história, é apenas este: em virtude da realização de um programa formal de computador do ponto de vista de um observador externo, esse alguém comporta-se exatamente como se entendesse chinês, mas de qualquer modo não compreende uma só palavra de chinês.” Searle (1984, p.40),

O argumento do quarto chinês é um assunto que tem gerado muita polêmica. Uma delas remete à questão: podem os computadores pensar? Nesse sentido, o objetivo da experiência de Searle consiste, em síntese, em atribuir intencionalidade a um computador e ver se, a partir da execução de um programa, ele poderia compreender os conteúdos que os seus símbolos representam para um observador exterior.

Enquanto Turing tenta derivar o “interior” da máquina a partir do seu comportamento exterior, Searle imagina-se, ele próprio, dentro do computador. A idéia é que, como Searle, dentro da caixa, é uma máquina de Turing e não compreende chinês, apesar de se comportar como se compreendesse então a Unidade Central de

Processamento (CPU), que também é uma máquina de Turing, também não precisa compreender chinês, ainda que se comporte como se o compreendesse. Ou seja, o computador não seria capaz de apreender a natureza significativa própria dos estados mentais.

Segundo Teixeira, J. de F. (1998, p.77),

“estados conscientes desempenham o papel de um “observador externo” que toma “decisões” diante de processos não-computáveis. Ora, como conceber um análogo a estes estados não-computáveis sem romper com uma hipótese materialista? É preciso encontrar na natureza algo semelhante, algo que possa servir de fundamento para uma abordagem científica da consciência.”

Sendo assim, pode-se pensar que a efetivação do programa apropriado do computador para a compreensão do chinês não é suficiente para lhe dar uma compreensão dessa língua, então, também não basta para dar a qualquer outro computador digital tal compreensão.

Em outras palavras, se puséssemos um computador digital qualquer, que apenas soubesse inglês, no lugar de Searle, esse computador não precisaria saber chinês para funcionar corretamente. Portanto, seria indiferente se estamos na presença de um computador biológico ou feito de outro material, intencional ou não, consciente ou não.

Desta forma, seria impossível, fosse qual fosse o sistema computacional, perceber os conteúdos significativos dos símbolos chineses, se apenas tivéssemos os símbolos. As regras de manipulação dos símbolos não são suficientes para nos dar o seu conteúdo.

Sendo assim, como os computadores de silício só operam símbolos segundo regras, não se pode atribuir-lhes qualquer conteúdo significativo. O computador simplesmente sintático não compreende sequer o significado da sua própria linguagem. Limita-se a manipular símbolos que não interpreta.

Esses questionamentos intrigaram e intrigam a Filosofia, a Psicologia e a Inteligência Artificial (IA), uma vez que esta realça a semelhança entre o funcionamento do cérebro humano e o funcionamento dos computadores digitais.

De acordo com Searle (1984 pp.35-36),

“Segundo a versão mais extrema desta concepção, o cérebro é justamente um computador digital e a mente é um programa de computador. Poder-se-ia resumir esta concepção - dou-lhe o nome de Inteligência Artificial (IA) forte – dizendo que a mente está para o cérebro tal como o programa está para o hardware do computador.”

Embora, tenhamos tipos de computadores que servem para apoiar os programas que constituem a inteligência humana, qualquer sistema físico que tivesse um programa correto de entrada e saída, teria uma mente. Os que defendem essa concepção acreditam que:

“[...] é apenas uma questão de tempo, até que os cientistas de computadores e os que trabalham na Inteligência Artificial (IA) projetem o hardware apropriado e os programas que serão o equivalente dos cérebros e das mentes humanas. Serão esses os cérebros e mentes artificiais que de todos os modos constituem o equivalente dos cérebros e mentes humanas.” Searle (1984, p.36)

Todavia, Searle propõe-se a refutar tal concepção. A natureza da refutação de Searle tem a ver com a definição exata do que vem a ser um computador digital, independentemente de qualquer estágio da tecnologia. O computador processa as informações em códigos binários, numa seqüência de zeros e uns. Esses símbolos não têm significado algum, ou seja, não têm conteúdo semântico. Os computadores, por definição, só têm sintaxe.

De acordo com Searle (1984, p.39),

“[...] a mente tem mais do que uma sintaxe, possui também uma semântica. A razão por que nenhum programa de computador pode alguma vez ser uma mente é simplesmente porque um programa de computador é apenas sintático, e as mentes são mais do que sintáticas. As mentes são semânticas, no sentido de que possuem mais do que uma estrutura formal, têm um conteúdo.”

Portanto, Searle entra em confronto direto com a pretensão da teoria computacional, que acredita ser possível reproduzir uma Inteligência Artificial (IA), produzida por qualquer artefato físico desde que bem montado. Para os funcionalistas, o computador feito com outro material, totalmente diferente, poderia executar o programa cognitivo e reproduzir a configuração exata da mente; isto seria uma questão de tempo.

Nessa ótica, poderia ser possível as máquinas pensarem, uma vez que os símbolos formais possibilitariam conduzir os processos lógicos realizados pelo cérebro para um mecanismo inorgânico.

Porém, ambas as motivações podem ser refutadas pela afirmação de que a mente é um fenômeno biológico natural. Enfim, Searle propõe as seguintes teses que refutariam a teoria da Inteligência Artificial (IA) forte: A primeira, assumindo que os cérebros causariam mentes; segunda, que a sintaxe não é suficiente para gerar a semântica; terceira, que os computadores só são capazes de ser definidos por uma estrutura formal sintática; quarta, que a mente possui conteúdos mentais semânticos. A seguir, as conclusões de Searle (1984 pp.48-51):

“Conclusão I – Nenhum programa de computador é, por si mesmo, suficiente para dar uma mente a um sistema. Os programas, em suma, não são mentes e por si mesmos não chegam para ter mentes.”

“Conclusão II – A maneira como a função cerebral causa mentes não pode ser apenas em virtude da ativação de um programa de computador.”

“Conclusão III – Tudo ou mais que causou mentes deveria ter poderes causais, pelo menos, equivalentes aos do cérebro.”

“Conclusão IV – Para qualquer artefato que pudéssemos construir, o qual tivesse estados mentais equivalentes aos estados mentais humanos, a realização de um programa de computador não seria por si só suficiente. Antes, o artefato deveria ter poderes equivalentes aos poderes do cérebro humano.”

Portanto, nenhum programa pode facultar uma mente ao computador e nem sequer pode ser entendido como tal. Não é o programa que leva as funções cerebrais a causarem a mente.

De acordo com Teixeira, J. de F. (1998, p.116),

“Será nosso cérebro capaz de produzir uma noção de complexidade que nos permita descrevê-lo? Este problema se desdobra imediatamente na dificuldade envolvida em representar a multiplicidade das conexões que devem estar presentes no cérebro. Esta multiplicidade pode ser tão complexa e intrincada que, mesmo que nela encontremos algum tipo de padrão e geração de um modelo de cérebro, mesmo com o auxílio de computadores, pode facilmente levar-nos a um problema do tipo NP (significa tempo não determinístico polinomial), ou seja, não poderíamos, num tempo razoável, produzir sequer um “retrato” aproximado de nosso próprio cérebro. E, neste caso, como poderíamos estabelecer todas as possíveis conexões entre seus neurônios – conexões que seriam responsáveis pelo aparecimento de formas mais complexas de vida mental. Em outras palavras, como simular aquilo que não podemos sequer representar?”

De fato, se não somos capazes de compreender totalmente e representar a multiplicidade de conexões presentes em nosso cérebro, como reproduzi-lo em uma máquina?

Ainda conforme Teixeira, J. de F. (1998, p.116),

“Uma descrição completa do cérebro será sempre mais complexa do que o próprio cérebro que a produz. Ora, como pode o cérebro produzir algo mais complexo do que ele mesmo? E como o próprio cérebro poderia compreender e reconhecer como sendo verdadeiro algo mais complexo do que ele mesmo? O problema do reconhecimento de tal teoria ou descrição pode levar a um impasse de difícil solução: não seria possível assegurar que tal descrição, uma vez atingida, é a correta. Ora, se o cérebro não pode produzir algo mais complexo do que ele mesmo, a possibilidade de replicá-lo através de sistemas artificiais fica afastada. Pelo menos a possibilidade de construir uma réplica do cérebro em laboratório.”

Tudo o que causa a mente deve ter, essencialmente, o mesmo poder causal dos cérebros humanos, o programa por si só não é suficiente. Searle explica a estrutura das ações humanas, ele esclarece como e em que sentido o comportamento é semelhante aos estados mentais internos. Entenda-se por comportamento, o comportamento humano voluntário e intencional, demonstrando como a ação humana se harmoniza com a explicação do problema Mente Corpo e rejeita a Inteligência Artificial (IA).

2.6 Searle e a crítica ao computacionalismo

Partimos do princípio, da realidade invencível e fundamental da consciência para poder explicar os fenômenos mentais de Searle, 1992, que de certa maneira ele aprofunda em sua crítica ao computacionalismo, com dois argumentos tão simples e devastadores quanto o mais antigo e conhecido argumento do quarto chinês.

O ataque de Searle ao computacionalismo se divide em três frentes, uma para cada uma das três alegações centrais dessa abordagem da mente. A primeira alegação afirma que “o cérebro é um computador digital”, a segunda, que “a mente é um programa computacional”, e a terceira, que “as operações do cérebro podem ser simuladas em um computador digital”. O argumento de Searle contra a segunda tese central do computacionalismo evidencia que a dimensão sintática não é suficiente para explicar o que faz a mente. Veremos agora como ele ataca a primeira e a terceira premissas dessa abordagem da mente.

Para refutar a terceira das teses centrais do computacionalismo, ele o faz de forma desconcertante e simples. Sua resposta é: sim, as operações do cérebro podem ser simuladas em um computador digital. As operações do cérebro, as das moléculas de um composto químico, as das condições meteorológicas, do crescimento de uma planta e de tudo o que obedecer a padrões em todo o universo. Mas, assim como, ao simular o comportamento de um furacão, nós não produzimos um furacão nem todas as suas propriedades, ao simular o comportamento de um cérebro, nós não produzimos a consciência e suas propriedades emergentes.

Dessa forma, sobra a primeira alegação, a central, a fundamental tese do computacionalismo, de que “o cérebro é um computador digital”, a única que poderia trazer ainda algum ânimo à Inteligência Artificial (IA). Mas, Searle (1992) nos oferece a

mais desconcertante de suas teses. Analogamente à questão três, ele responde, sim. Um cérebro é um computador digital.

Conforme Searle (1992, p.208),

“Seguindo a definição de computação padrão oferecido em livros-texto, é difícil ver como evitar os seguintes resultados:

1- Para qualquer objeto, existe alguma descrição de tal objeto tal que sob esta descrição o objeto é um computador digital.

2- Para qualquer programa e para qualquer objeto suficientemente complexo, existe uma descrição deste objeto sob a qual este está implementando tal programa. Então a parede atrás de minhas costas, por exemplo, está neste exato momento implementando o programa Word, porque existe nela algum padrão de movimentos moleculares que é isomórfico com a estrutura formal do Word. Mas se a parede esta rodando o Word, então esta é uma parede grande o bastante para estar implementando qualquer programa incluindo qualquer programa implementado no cérebro.”

O que Searle diz é que um programa pode ser executado por qualquer coisa organizada para reagir digitalmente, pois a sintaxe não é algo físico como gravidade ou massa e se encontra somente nos olhos do observador.

Ainda, conforme Searle (1992, p. 209),

“A múltipla realizabilidade de processos computacionalmente equivalentes em meios físicos diferentes não é somente um sinal de que os processos são abstratos, mas também de que eles não são intrínsecos ao sistema. Eles dependem de uma interpretação de fora.”

Dessa forma, com essas explicações, Searle demonstra que as crenças básicas do computacionalismo são falsas, como possivelmente carecem de um sentido preciso.

Para ele, a consciência é real e o programa computacional está nos olhos de quem vê (só existe para a consciência). Sendo assim, a intencionalidade não pode ser explicada nem identificada com computação.

Entretanto, retomando a questão da “intencionalidade” do computador, o que significa dizer então que um computador processa informação? Nenhum computador processa completamente informação, diz Searle (1992), e assim nenhum cérebro

também o faz. Computadores nos ajudam a transmitir informações e a processá-las, eles não fazem a parte mais difícil do serviço. Um agente externo (programador) codifica alguma informação, cujo significado já esteve previamente acordado e estabelecido, de uma forma que possa ser processado pelo hardware.

Então, o computador, através de uma série de estágios elétricos, transforma novamente os sinais elétricos em caracteres, nos quais a informação está codificada, para que um agente externo possa interpretá-lo, tanto sintática quanto semanticamente, uma vez que o hardware não tem nenhuma sintática ou semântica intrínseca. Searle (1992, p.223).

Recentemente, Christian Kaernbach (2006), nos explica em seu artigo “No virtual Mind in the Chinese Room”, que Searle escolheu o *chatterbot*, ou seja, programa de computador que tenta simular um ser humano na conversação, como uma parodia da visão do teste de Turing (1950). No entanto, com o *chatterbot*, o computador poderia conversar. Ainda assim, o que para nós seria, um simples bate-papo, para a máquina, tratar-se-ia de um verdadeiro obstáculo a ser superado.

Ora, o que nos leva a pensar que o argumento do quarto chinês continua sendo um assunto muito polêmico, que ainda se encontra em discussão com a questão da sintaxe e semântica. Entretanto, para Christian Kaernbach (2006), “a situação do quarto chinês”, é passível de realização.”

2.7 Searle e a crítica ao Cognitivismo

O cognitivismo diz que pensar é processar informação, mas o processamento de informação é justamente manipulação de símbolos. Os computadores fazem manipulação de símbolos. Assim, a melhor maneira de estudar o pensamento é estudar os programas computacionais de manipulação de símbolos, quer existam em computadores ou em cérebros.

De acordo com os cognitivistas é importante e fundamental caracterizar o cérebro ao nível de seu funcionamento como sistema de processamento de informação e não ao nível das células nervosas nem ao nível dos estados mentais conscientes.

Segundo Searle, uma das razões que leva alguns estudiosos a acreditar que o cognitivismo é verdadeiro provém de experiências que mostram que diferentes tarefas intelectuais exigem diferentes intervalos de tempo para serem executadas. Isto ocorre tanto com as pessoas quanto com os computadores, o que leva os cognitivistas a inferir que o sistema humano trabalha com os mesmos princípios que um computador.

A segunda razão tem a ver com a lingüística, e diz que as regras formais da gramática de uma língua são semelhantes às regras formais que um computador segue, portanto, existe algum sentido unitário em que o cérebro e o computador funcionariam de uma maneira semelhante.

Pressupõem também os estudiosos cognitivistas que no cérebro humano possa existir internalizada alguma teoria que habilite as pessoas a aprender uma língua, por exemplo. Relacionando o programa de computador com o computador propriamente dito, encontram um modelo para fazer um paralelo com o cérebro humano e explicar as relações entre a mente e o cérebro.

Ora, tal questão pode ser esclarecida de um modo muito simples: no sentido em que os seres humanos seguem regras (e, incidentalmente, os seres humanos seguem regras bem menos do que pretendem os cognitivistas), nesse sentido, os computadores de nenhum modo seguem . Apenas atuam de acordo com certos procedimentos formais.

O programa do computador determina os vários passos que uma máquina deve realizar; determina o modo como um estado será transformado num estado subsequente. E podemos falar metaforicamente como se se tratasse do seguimento de regras. Mas, no sentido literal em que os seres humanos seguem regras, os computadores não seguem regras, apenas atuam como se estivessem seguindo.

Podemos falar metaforicamente de qualquer sistema como se ele seguisse regras, por exemplo, o sistema solar. A metáfora só se torna prejudicial, quando se confunde com o sentido literal.

Pode-se utilizar uma metáfora psicológica para explicar o computador. A confusão surge quando se compreende a metáfora no sentido literal e se usa o sentido metafórico do computador de seguir regras para tentar explicar o sentido psicológico do seguimento de regras.

Por mais que tentemos clarear a distinção da metáfora do computador com o sentido psicológico do seguimento de regras, tentamos de certa maneira encontrar as mais profundas fraquezas do argumento cognitivista.

A partir de fato de que Searle faz o processamento de informação ao pensar, e do fato de que o computador faz o processamento da informação (o mesmo processamento de informação que pode simular as características formais do pensamento), não significa que existe algo de psicologicamente relevante relacionado ao programa de computador.

Provavelmente, o computador não é uma metáfora para o cérebro, melhor ou pior do que as metáforas mecânicas criadas anteriormente, como, por exemplo, a de que o cérebro trabalhava como um sistema telegráfico, segundo Sherrington, neurocientista britânico.

Entretanto, o cognitivismo, tendência recente da Ciência Cognitiva, defende que o computador apresenta uma imagem correta da natureza do mental, e não deve ser visto como metáfora, apenas. Tal visão não afirma, necessariamente, que computadores têm, literalmente, estados mentais, mas, sim, que o cérebro efetua processamento de informação, ou seja, pensar, por exemplo, seria processar informação.

Ora, se processar informação é, justamente, manipular símbolos, e os computadores digitais efetuam processamento de informação, então, a melhor maneira de se estudar a mente seria através de modelos computacionais. Para os cognitivistas, a tese geral é que: se podemos conceber computadores que seguem regras quando processam informações, e se os seres humanos também seguem regras ao pensar, então, existe algum sentido em que cérebro e computador funcionem de uma maneira semelhante e, quem sabe, talvez idêntica.

Com relação à afirmação de que seres humanos e computadores seguem regras, Searle observa que existe uma enorme diferença entre ambos. No caso dos seres humanos, eles são guiados pelo efetivo conteúdo da regra, são os significados que produzem comportamento. Para que uma regra seja seguida, o significado da regra tem de ter uma causa que leve a uma conduta. Nesse sentido os computadores de nenhum modo seguem regras, eles apenas seguem procedimentos formais, ou seja, atuam como se estivessem seguindo regras.

Capítulo III

3. Inteligência Artificial (IA) ou entendimento automático?

A crescente expansão dos computadores modernos, com sua velocidade de processar dados, tem permitido a aplicação de tarefas que teriam sido quase impossíveis há alguns anos. Por exemplo, reconhecimento de voz, de escrita e de imagens; a criação de robôs cada vez mais complexos, etc.

Todo esse progresso leva-nos a pensar nos limites de atuação dos computadores, ou seja, será que um dia as máquinas poderão substituir toda a atividade humana intelectual?

Será que as máquinas vão ser capazes de revelar comportamento inteligente e substituir os seres humanos em tarefas criativas? Será que desenvolverão algum tipo de pensamento e sentimento que os seres humanos têm?

Tais questionamentos permeiam discussões nos meios científicos e fornecem vasto material para produção de filmes de ficção científica e outras produções no meio artístico e científico.

3.1. O que é um computador?

Ao conceituar um computador, procede-se do ponto de vista lógico, ou seja, caracterizam-se suas formas de processar dados e o fato de tratar-se de um processamento sintático e não semântico.

Entretanto, os computadores modernos digitais são máquinas matemáticas, lógico-simbólicas, algorítmicas. Isto é, o processamento e o efeito de qualquer instrução interpretada em linguagem da máquina (minuciosamente, um computador nunca executa uma instrução, ele a interpreta) podem ser matematicamente descritos, quer dizer, representam uma função matemática.

Além disso, a matemática utilizada é bastante restrita: só trabalha símbolos tirados de um conjunto finito, discreto, para o qual sempre se pode atribuir um sistema numérico. Isto é, o espaço de trabalho de um computador é sempre quantificado.

Ora, um programa é uma seqüência de instruções, podendo sempre ser associado a uma função matemática que leva elementos de um conjunto de dados de entrada a um conjunto de dados de saída. Ou seja, dados são representações de símbolos quantificados ou qualificáveis.

Por exemplo, se uma figura é varrida por um scanner e introduzida no computador, que então a imprime, a figura resultante parecerá a mesma que a original. Tal figura é quantificável, porque dentro do computador todo objeto é representado usando-se um sistema numérico, isto é, por meio de quantidades. Outros exemplos de dados são textos, sons gravados e animação.

Porém, os computadores são máquinas algorítmicas, um algoritmo é uma seqüência finita de ações matematicamente bem definidas.

Um programa é uma seqüência de regras matemáticas sobre como transformar, transportar e armazenar dados. Dentro do computador, os dados são representados como cadeias de símbolos quantificados. Sendo assim, as regras podem ser consideradas como uma sintaxe que é aplicada a essas cadeias. As próprias cadeias sempre seguem uma determinada estrutura.

Por exemplo, uma cadeia representando um endereço pode ser composta de três partes: logradouro (incluindo o número do imóvel), CEP e cidade. O CEP deve seguir uma regra de formação: uma cadeia de cinco dígitos decimais, seguida de um hífen e de mais uma cadeia com dígitos. Como programas e dados seguem regras sintáticas, e os programas são executados rigidamente, cada instrução segue regras específicas para sua interpretação pela máquina, podendo dizer-se que, neste caso, o computador é uma máquina sintática.

Sendo assim, podemos dizer que computadores são máquinas que processam dados. Não são máquinas que processam informações, porque eles não têm nenhum entendimento do que processam. Searle, conforme expusemos minuciosamente no capítulo II, desenvolveu uma interessante experiência mental para ilustrar esse aspecto.

3.2 Homens versus máquinas

Embora se criem muitas fantasias em torno das possibilidades de uma máquina, o fato é que as máquinas apenas realizam tarefas determinadas por nós, seres humanos.

“ [...] quando os computadores finalmente conseguem realizar com competência um objetivo determinado, é usual que passemos a considerar aquela atividade como não sendo “inteligente” o suficiente, simplesmente porque agora entendemos melhor o seu funcionamento. Como na época em que lavadoras de prato automáticas eram denominadas, por seus criadores, robôs. Quando se tornaram comuns, foram “rebaixadas” a simples, eletrodomésticos.”³

Os eletrodomésticos e as outras máquinas do dia-a-dia não passam de artefatos pendurados nas portas de outros artefatos maiores e mais complexos. A rede elétrica da cidade, o sistema telefônico; também as usinas, as fábricas, as indústrias e assim por diante. Não nos é dado nem ver nem atinar muito bem com o funcionamento desses sistemas.

Quando alguém se refere às maravilhas das máquinas modernas, está especialmente se referindo aos computadores e família. Ou, o que dá no mesmo, aos engenhos que contam unidades processadoras de informação fazendo as vezes do elemento controlador.

A origem do computador eletrônico pode ser fixada em 1946, com a construção do Eniac.⁴ Daí para cá, o que se viu foi um surto espantoso de tais máquinas, que proliferaram principalmente graças ao desenvolvimento da microeletrônica. Da prodigiosa família dos computadores, todos conhecem seus membros mais corriqueiros, as calculadoras portáteis. Há, hoje, maior número delas espalhadas por aí do que jamais houve dos antigos artefatos de fazer contas: os ábacos, as tabuadas, as régua de cálculo,

³ Disponível em http://pt.wikipedia.org/wiki/conhecimento_humano Acesso 20 Abril 2007.

⁴ Disponível em <http://pt.wikipedia.org/wiki/ENIAC> acessado em 23 Janeiro 2009. Electrical Numerical Integrator and Calculator.

etc. Tal disseminação prova que o advento das calculadoras não é meramente uma febre temporária. Ao contrário, o domínio da tecnologia é definitivo e elas vieram para ficar.

A principal característica das máquinas computadoradas é seu potencial para acumular e assimilar informações. Essa capacidade permite que ofereçam mais recursos do que os programados originalmente.

Ora, é comum creditar-se a capacidade de inovação dos seres humanos à existência de uma inteligência ou à faculdade de pensar e criar. Daí, dizer-se que as máquinas computadoradas são capazes de pensar ou exibir um comportamento inteligente.

Classificar o desempenho de uma máquina como inteligente pressupõe que não se conheça o preciso significado do vocábulo inteligente. Entretanto, há pesquisadores que se empenham na busca de fatores que venham a distinguir, ou confundir, homens e máquinas.

3.3 A máquina permitindo ao homem ocupar-se de si

A disputa homens versus máquinas não se dá apenas no que diz respeito às suas potencialidades intelectivas. Há também o problema das relações sócio-econômicas.

A substituição da ação humana pela de uma máquina é, antes de tudo, uma moeda de duas faces. De um lado pode ser vista com bons olhos: o homem fica desobrigado de realizar um trabalho possivelmente desagradável e pode se dedicar a uma atividade mais prazerosa, compensado pela máquina. Do outro lado, o homem deixa de fazer um trabalho, ao qual tinha tanto o direito como a necessidade, porque apareceu uma máquina mais eficiente.

De certa forma, isso acaba virando um dilema: as máquinas criadas pelo homem para servi-lo podem, de repente, se voltar contra ele na disputa por empregos, ou seja, as máquinas desalojam mão de obra.

Ainda assim, parece um contra senso privarmo-nos do seu uso por causa disso, pois o homem, utilizando a máquina, poderia ocupar-se com assiduidade das coisas que, como ser humano, lhe dizem mais respeito.

As máquinas foram e são construídas com a única finalidade de substituir a força de trabalho humano. Algumas delas são chamadas de robôs, porém nem toda a máquina é um robô.

No entanto, muitas das que estão hoje ao nosso redor são, sim, pequenos robôs. Existem alguns que possuem sensores que lhes permitem rastrear o terreno à sua volta; outros têm a capacidade de se locomoverem; muitos têm braços mecânicos que lhes permitem agarrar e deslocar objetos. Mas, robôs não reproduzem necessariamente a figura humana. Apenas amplificam nossa aptidão para realizar trabalho.

A idéia de robô é bastante antiga e sua origem divide pesquisadores de diversas áreas. Apesar da falta de consenso sobre quando surgiu o primeiro robô, sabe-se que o termo nasceu primeiro na ficção científica e, mais tarde, foi aplicado à ciência.

3.4 A invenção da cibernética

A cibernética é a ciência que estuda o sistema de controle e de comunicação nas máquinas e nos organismos vivos, de homens e de animais.

Em 1948, o matemático Nobert Wiener, começou a apresentar as hipóteses e os fundamentos da cibernética, resultado de vários anos de pesquisa e interação com pesquisadores de diversas áreas científicas, incluindo as ciências sociais.

A idéia fundamental, desenvolvida por Wiener com seus principais colaboradores, era a de que certas funções de controle e de processamento de informações semelhantes em máquinas, em seres vivos e, também, de alguma forma, na sociedade são, de fato, equivalentes e redutíveis aos mesmos modelos e às mesmas leis matemáticas.

Ele entendia que a cibernética seria uma teoria das mensagens mais ampla que a “teoria da transmissão de mensagens de engenharia elétrica”,

[...] um campo mais vasto que inclui não apenas o estudo da linguagem mas também o estudo das mensagens como meios de dirigir a maquinaria e a sociedade, o desenvolvimento de máquinas computadoras e outros autômatos [...], certas reflexões acerca da psicologia e do sistema nervoso, e uma nova teoria conjectural do método científico. Wiener (1984, p.15).

O campo que Wiener designou como “cibernética” teve início durante a II Grande Guerra, quando ele realizou pesquisas com programação de máquinas computadoras e com mecanismos de controle para artilharia antiaérea.

Tanto em uma como em outra pesquisa, Wiener engajou-se no que descreveu como “estudo de um sistema eletro-mecânico, desenhado para exercer indevidamente uma função especificamente humana”: a “execução de um complicado padrão de cálculo” em um caso, e a “previsão do futuro”, em outro, no qual se deveria prever a

trajetória de uma aeronave, para que o projétil de canhão antiaéreo encontrasse seu alvo em algum momento do futuro. . Wiener (1948, pp. 11,13).

A Cibernética envolve diversas áreas, tais como a Matemática, a Medicina, a Psicologia, a Antropologia, a Filosofia e a Sociologia.

3.5 Modelo antropológico e resíduo cultural

Nas ciências sociais, encontramos Bateson como um dos fundadores do pensamento cibernético e sua enorme contribuição pode ser encontrada na obra de autores como Goffman, 1983, Strathern, 1992, Rapport, 2000 e Lévi-Strauss, 2000.

Também encontramos componentes cibernéticos no pensamento de Geertz, 1989, que, por sua vez, vê na relação entre a evolução cultural e a evolução biológica princípios da cibernética que levam a um processo contínuo de realimentação e influências recíprocas e condicionadas:

“À medida que a cultura, num passo a passo infinitamente pequeno, acumulou-se e se desenvolveu, foi concedida uma vantagem seletiva àqueles indivíduos da população mais capazes de levar vantagem [...] Australopiteco proto-humano, de cérebro pequeno, tornou-se o Homo Sapiens, de cérebro grande, totalmente humano. Entre o padrão cultural, o corpo e o cérebro foi criado um sistema de realimentação (*feedback*) positiva, no qual cada um modelava o progresso do outro, um sistema no qual a interação entre o uso crescente das ferramentas, a mudança da anatomia da mão e a representação expandida do polegar no córtex é apenas um dos exemplos mais gráficos. Submetendo-se ao governo de programas simbolicamente mediados para a produção de artefatos, organizando a vida social ou expressando emoções, o homem determinou, embora inconscientemente, os estágios culminantes do seu próprio desenvolvimento biológico. Literalmente, embora inadvertidamente, ele próprio se criou.” (Geertz, 1989, p. 60).

Apesar de ter estimulado hipóteses, teorias e pesquisas, tanto na antropologia como em diversos outros campos científicos, e ter dado origem a novas áreas, como as Ciências Cognitivas, ainda assim, a cibernética foi esquecida como a “vasta teoria das mensagens” desejada por Wiener, e, hoje, quase ninguém se auto-intitula um ciberneticista.

Os modelos teóricos da Cibernética se desgastaram e, mesmo no campo do controle artificial, onde se consolidaram sólidas disciplinas “cibernéticas”, como a Informática e a Robótica, a proposta de Wiener esvaziou-se na prática.

É importante notar que, se por um lado, a cibernética não se consolidou no plano científico, ela influenciou de forma determinante a cultura moderna que poderíamos chamar hoje de “cibercultura”. com resíduos de seus modelos explicativos.

Tais resíduos são certas noções e valores provenientes do discurso técnico e científico que, deslocados para o plano do senso comum, introduzem novas distinções nos antigos esquemas interpretativos, para que eles possam fazer frente às propriedades de um mundo no qual as fronteiras entre os domínios do orgânico, do tecno-econômico e do textual tornaram-se permeáveis:

[...] produzindo sempre montagens e misturas de máquina, corpo e texto: enquanto natureza, os corpos e os organismos certamente possuem uma base *orgânica* eles são cada vez mais produzidos em conjunção com as *máquinas*, e esta produção é sempre mediada por narrativas científicas [...] e pela cultura em geral. (Escobar, 2000, p. 61).

Um dos resíduos mais importantes que a cibernética transmitiu à cibercultura foi a visão de que os seres vivos e as máquinas não são essencialmente diferentes. Essa noção se manifestou, em especial, nas tecnologias especializadas em modificar a vida (tecnologia da informação, robótica, biônica e nanotecnologia) e nas tecnologias especializadas em manipular a vida (as biotecnologias).

A relação entre organismo e máquina depende exclusivamente do texto, não só na forma de narrativa científica, mas também na forma dos códigos que determinam o funcionamento tanto das máquinas (softwares) como dos seres vivos (o código genético).

Os produtos – reais e imaginários – de tais tecnologias podem contradizer certas noções de classificação fundamentais, tais como a oposição entre natureza e cultura, entre orgânico e inorgânico, entre o homem e a máquina, dentre outras.

Segundo Lévi-Strauss (2002, p. 25), a exigência de ordem “constitui a base de todo pensamento”. É por isso que seres ambíguos são, com frequência, objeto de restrições e tabus: são sinais de desordem, contradizem as fronteiras estabelecidas entre as categorias classificatórias e, assim, ameaçam as próprias convicções acerca da ordem do mundo.

De acordo com Mary Douglas (1991, p. 54),

“a cultura divide-se com a experiência dos indivíduos. Fornece-lhes, em contra partida, algumas categorias básicas, uma esquematização positiva na quais idéias e valores se encontram dispostos de forma ordenada”.

O universo não é um aglomerado de “objetos em si”, mas um repertório organizado de objetos significantes que portam significados socialmente compartilhados.

Conforme Sahlins (1990, p. 10-11) observa, além dos consensos que as sociedades elaboram serem resultados da interação de perspectivas diversas, os significados das coisas e suas relações estruturais são reavaliados na realização prática e, freqüentemente, repensados criativamente dentro de certos limites – dados pelo sentido coletivo, em resposta às contingências apresentadas pela experiência prática.

Assim, podemos, por exemplo, entender que o consenso social acerca do que é correio eletrônico (e-mail) está dentro dos limites de significação de “eletrônico” e “correio” (electronic e mail), sobre os quais já havia um consenso social.

O mesmo ocorre com ciberespaço: os termos que sintetizam o discurso técnico-científico adquirem novas conotações e produzem significados inéditos na sua conjunção com antigos significantes (mail, space, organismo), projetando o sistema antigo de interpretação da realidade sob novas formas, dentro de dadas possibilidades históricas e culturais de significação.

O que ultimamente tem se chamado de “cibercultura” é uma resposta positiva da cultura na criação de uma “nova ordem do real”, frente aos novos contextos práticos que desafiam as categorias tradicionais de interpretação da realidade.

Entre o homem de lata mecanizado e o corpo humano, ou entre uma máquina de calcular programável à válvula e a mente humana, existem discontinuidades gigantescas, de tal forma que aqueles dificilmente passam de representações caricaturadas do homem, chegando, em muitos casos, a reafirmar a oposição das categorias que separam o humano da máquina.

Nesse sentido não são, ainda, “cibernéticos”, pois a principal característica enunciada pela “cibernética” é a de que não existe discontinuidade entre máquinas e organismo.

O futuro cibernético implica uma nova ordem do real, porque, enfim, a permuta é apenas uma questão de compatibilidade funcional.

3.6 O corpo pós-humano

Partindo-se da idéia de que os grandes avanços tecnológicos na atualidade são extraordinariamente surpreendentes, o que será que os pesquisadores pretendem ainda verificar na área das ciências da vida? Uma vez que a tecnologia pode modificar o comportamento humano, será que um dia afetará os valores da humanidade?

“As máquinas do final do século XX tornaram completamente ambígua a diferença entre o natural e o artificial, entre a mente e o corpo, entre aquilo que se auto cria e aquilo que é externamente criado, podendo-se dizer o mesmo de muitas outras distinções que se costumavam aplicar aos organismos e às máquinas. Nossas máquinas são perturbadoramente vivas, assustadoras e inertes.” (Haraway, 2000, p. 46).

O ciborgue é o primeiro produto cultural dessa “nova ordem do real”, baseada na cibernética, e conjuga as promessas da biônica com as perspectivas anunciadas pela cibernética. A “biônica” é uma área relacionada com a biomimética, que pode ser definida como a “ciência de sistemas que têm alguma função copiada da natureza, ou que representa características de sistemas naturais ou seus análogos” (Vincent, 2004, p. 1).

Já, o termo cyborg nasceu da contração de cybernetics organism e foi apresentado, em 1960, por Manfred Clynes e Nathan Kline em um simpósio sobre os aspectos psico-fisiológicos do vôo espacial. Na ocasião, eles apresentaram a idéia de se ligar ao ser humano um sistema de monitoramento e regulação das funções físico-químicas, a fim de deixá-lo dedicado apenas às atividades relacionadas com a exploração espacial.

Em 1972, Martin Caidin lançou a ficção científica cyborg, que conta a história de um piloto de testes da Força Aérea americana, Steve Austin, que após um grave acidente é reconstruído com partes biônicas pelo laboratório cibernético do Dr. Killian.

“Uma nova raça. Um casamento da biônica (biologia aplicada à engenharia de sistemas eletrônicos) e cibernética. Um organismo cibernético. Chame-o de ciborgue.” (Caidin, 1972, p. 55-56).

O ciborgue que Caidin nos legou é produto de uma biônica reinventada que, sob a inspiração da idéia de Clynes e Kline, não é mais uma simples técnica de mimese da natureza, mas um meio de reconstruí-la e superá-la.

Produto do pensamento utilitarista aplicado sem limites à carne e ao aço, o ciborgue anuncia a imagem de um homem “melhorado” com a acoplagem da tecnologia e cada vez mais além das limitações de desempenho ditadas pela natureza.

A “performance” é a noção fundamental para a reformulação da imagem do ser humano na direção da imagem do “pós-humano”. O desenvolvimento de próteses também está intimamente ligado à superação de limites.

Originalmente, tais limites eram os impostos àqueles cuja natureza do corpo fora mutilada, por nascença ou acidente. Mas, hoje, acoplados em próteses de competição, os para-atletas velocistas agregam muita tecnologia. E eles são capazes de ultrapassar, e muito, a velocidade das pessoas comuns e chegam próximo às de recordistas mundiais olímpicos:

“Tony Volpentest inspira admiração e, quem sabe, até despeito. Munido de duas pernas mecânicas, o atleta americano, de 26 anos, faz 100 metros rasos em impressionantes 11 segundos e 36 centésimos de segundo – apenas um segundo e meio atrás do recordista mundial, o canadense Donovan Bailey, que nasceu com tudo no lugar. Medalha de ouro nos Jogos Paraolímpicos de Atlanta, em 1996, Tony veio ao mundo sem os pés e sem as mãos.” (Dias, 1999, p. 136)

Exibindo próteses de alta tecnologia, desenhadas sob medida para competições, a imagem de para-atletas tem sido explorada em propagandas e desfiles de moda. No discurso da mídia e da propaganda, onde exibem ostensivamente o seu corpo híbrido, os para-atletas corredores materializam hoje as aspirações do futuro do corpo pós-humano, o homem redesenhado para uma “melhor performance”.

Na perspectiva da “estética” da performance, as máquinas de musculação, os programas planejados de modelagem muscular, as próteses estéticas, as técnicas cirúrgicas são apenas meios que a tecnologia disponibiliza para se atingir a imagem do corpo de alto desempenho, a imagem na direção do corpo pós-humano.

3.7 Os limites do ciberespaço

Em seu livro de não-ficção, *The Hacker Crackdown – Law and Disorder on the Eletronic Frontier*, Bruce Sterling comenta que o termo *cyberspace* surgiu em 1982, na literatura cyberpunk (Sterling, 1992, p. XI). Naquele ano, Willian Gibson lançou *Neuromancer*, um clássico da literatura cyberpunk que, além do termo *cyberspace*, também introduziu o termo *matrix*, para se referir ao ciberespaço como uma rede global de simulação.

Sterling, 1992, acrescenta que o “ciberespaço” não é uma fantasia de ficção científica, mas um “lugar” onde temos experiências genuínas e que existe há mais de um século:

Mas o território em questão, a fronteira eletrônica, tem cerca de 130 anos. Ciberespaço é o “lugar” onde a conversação telefônica parece ocorrer. Não dentro do seu telefone real, o dispositivo de plástico sobre sua mesa. [...] [Mas] O espaço entre os telefones. O lugar indefinido fora daqui, onde dois de vocês, dois seres humanos, realmente se encontram e se comunicam. [...] Apesar de não ser exatamente “real”, o “ciberespaço” é um lugar genuíno. Coisas acontecem lá e têm conseqüências muito genuínas. [...] Este obscuro submundo elétrico tornou-se uma vasta e florescente paisagem eletrônica. Desde os anos 60, o mundo do telefone tem se cruzado com os computadores e a televisão, e [...] isso tem uma estranha espécie de fisicalidade agora. Faz sentido hoje falar do ciberespaço como um lugar em si próprio. [...] Porque as pessoas vivem nele agora. Não apenas um punhado de pessoas [...] mas milhares de pessoas, pessoas tipicamente normais. [...] Ciberespaço é hoje uma “Rede”, uma “Matriz”, internacional no escopo e crescendo rapidamente e constantemente. (Sterling, 1992, pp.11,12).

A preocupação de Sterling em estabelecer o ciberespaço como realidade tem a ver com a natureza do ciberespaço, atualmente conhecida como “virtual”.

Entretanto, o reconhecimento de que a realidade é “uma qualidade pertencente a fenômenos independentes de nossa própria vontade”. (Berger; Luckmann, 1998, p. 11) basta para ver que essa oposição “virtual” versus “real” é ilusória e bastante confusa.

Nos bancos de dados de corporações e governos, já somos seres “virtuais”, queiramos ou não, e cada vez mais temos o conhecimento de que o ciberespaço, apesar

de virtual, é bastante “real”. Os crimes virtuais são exemplo de que a virtualidade do ciberespaço possuem uma problemática realidade.

Durante muito tempo, foi o texto, na forma de complexos códigos mnemônicos e textuais, e não a imagem visual, a mediação, por excelência, entre as máquinas computadoradas e o homem.

O texto introduziu, aos antigos computadores baseados em cartões perfurados, o teclado e o display alfanuméricos. A mediação derradeira entre o homem e a máquina são os bits: pequenos sinais físicos que podem assumir apenas dois valores, convencionalmente representados por “um” e “zero”. O bit é o átomo da informação eletrônica que é armazenado, processado e intercambiado dentro dos computadores.

O que temos no disco rígido, no CD, no *pendrive* ou na Internet são, em última instância, apenas cadeias binárias. O que trafega pelo cabo da impressora, pela linha telefônica, ligada ao modem, ou pelo cabo da rede são bits. A própria indexação das cadeias corretas que compõem um arquivo ou um programa está em outras cadeias binárias.

A imagem gerada pelo computador é resultado de simulações de modelos que reformulam de modo sensível os conceitos lógicos e matemáticos, contidos nos dados e nos programas de computador.

Toda imagem eletrônica é um mosaico matricial de pequenos pontos, os pixels, cada qual com uma graduação de luz e cor. Ao contrário da televisão, em que o pixel é resultado de um processo mimético de “contágio” da luz, através dos vários suportes ópticos e eletrônicos, o computador domina cada ponto da imagem: ele substitui “o automatismo analógico das técnicas televisuais pelo automatismo calculado, resultante

da informação relativa à imagem”. [...] “Cada pixel é um permutador minúsculo entre imagem e número” (Couchot, 1993, p. 38-39).

O que chamamos de realidade virtual é a camada de interação sensível entre o homem e o ciberespaço. Mas as representações imagéticas da informação digital implicam uma descontinuidade entre aquilo que vemos e aquilo que realmente está por trás da simulação.

A realidade virtual opera em dois sentidos, um que cria mundos sensoriais da informação digital e outro que trabalha ocultando a estrutura essencial e material do ciberespaço. São movimentos indissociáveis e, por mais perfeito que venha a ser um modelo de simulação, ele será sempre ambíguo: o mesmo poder de simular mundos é o poder de falsificar e mascarar (Taussig, 1993, pp. 42-43).

A tecnologia para que o computador passasse a ser o lugar por excelência de um “espaço virtual” foi inicialmente desenvolvida pela Xerox, em 1971, com um protótipo de interface gráfica: era o primórdio daquilo que viria ser chamado de interface amigável, popularizada com o lançamento do Macintosh uma década depois (Negroponte, 1995, p. 90).

O nosso computador de hoje é um desktop virtual, um plano de signos organizados seletivamente, para construir uma imitação visual, mediadora da nossa relação com o ciberespaço.

No plano dessa realidade sensorial, o cursor do mouse é o nosso “dedo virtual”, que arrasta outros “objetos virtuais” para o “lixo virtual” ou aciona o “telefone” (modem ligado à linha telefônica) que nos conecta a outros sistemas e à Internet.

Enfim, quanto mais humanizamos a nossa relação com o ciberespaço, por meio de simulações que imitam a nossa realidade não-virtual, mais nos tornamos

cibernéticos. A contrapartida da naturalização do ciberespaço é que nos tornamos, também, extensão dele.

Daí, nossa fascinação e temor em relação aos hackers, que tanto trafegam como sujeitos virtuais do ciberespaço como manipulam os códigos “secretos” por trás das realidades virtuais. O hacker é, pois, uma figura que simboliza a cibercultura.

O hacker é a própria síntese da defesa ao mundo sintético como extensão do homem, ele incorpora a imagem daquele que transcende a condição de objeto virtualizado, tornando-se assim, o sujeito. capaz de superar nosso próprio cotidiano cibernetizado.

3.8 Inteligência Artificial (IA) na ficção científica

Um tema bastante recorrente em histórias de ficção científica, a Inteligência Artificial (IA) está presente em livros, desenhos animados e filmes. Um dos autores de grande destaque nessa área é o russo Isaac Asimov, autor de histórias de sucesso como: “O Homem Bicentenário” e “Eu, Robô”, dentre outras.

Essas histórias foram transformadas em filmes. Outro filme que merece ser citado é “Inteligência Artificial”, de Steven Spielberg. e Stanley Kubrick, de 2001.

Nessa história, um casal com um filho praticamente desenganado, congelado pelo método criogênico, resolve acolher em casa a última novidade produzida pela Cybertronics Manufacturing: um Meca filho, robô idêntico a uma criança, programado para amar seus pais adotivos, assim que eles tiverem lido sete palavras, previamente definidas pela empresa que o construiu.

Quando o filho verdadeiro do casal supera as previsões médicas e se recupera, ele passa a disputar com o Meca filho o amor da mãe, que após diversos incidentes entre os dois, decide devolver o robô para a Cybertronics.

No entanto, a empresa só o receberia de volta na condição de destruí-lo. Com pena da criatura que chegou a acolher como filho, a mãe adotiva resolve abandoná-lo em uma floresta.

Daí em diante, a exemplo do personagem infantil Pinóquio, o Meca filho passa a tentar descobrir, de forma incansável, como fazer para se tornar humano. A clássica história infantil, “As aventuras de Pinóquio”, sobre o boneco de madeira que ganha vida após ser construído pelo mestre Gepeto, foi publicada, pelo italiano Carlo Collodi, em 1893.

Quatro séculos antes, seu conterrâneo Leonardo da Vinci, que viveu entre 1452 e 1519, já desenvolvia pesquisas sobre a anatomia humana que ajudariam na criação de articulações mecânicas. A partir dos estudos de Leonardo da Vinci, surgiram bonecos que moviam mãos, olhos e pernas e conseguiam realizar ações simples como escrever ou tocar certos instrumentos musicais.

A idéia de bonecos mecânicos, com funcionamento previamente programado, como o robô do filme “Inteligência Artificial”, só se desenvolveria após 1940, quando foi apresentado ao mundo o primeiro computador digital da história.

Entretanto, a ficção científica tratou de muitos filmes referentes a robôs quase humanos. Isaac Asimov, em seu livro “Histórias de robôs”, criou as seguintes leis fundamentais da robótica (2007 pp.80-81):

“1ª lei: Um robô não pode ferir um ser humano ou, por omissão, permitir que um ser humano sofra algum mal.

2ª lei: Um robô deve obedecer às ordens que lhe sejam dadas por seres humanos, exceto nos casos em que tais ordens contrariem a Primeira Lei.

3ª lei: Um robô deve proteger sua própria existência, desde que tal proteção não entre em conflito com a Primeira e Segunda Leis.”

As leis propostas por Asimov sintetizam algumas inquietações humanas apontadas por Tucherman, 2004, expressas na ficção, diante de novas possibilidades científicas e tecnológicas que não se restringem à área da robótica.

Conforme Tucherman (2004), no romance Frankenstein, publicado pela inglesa Mary Shelley, em 1816, o personagem criado a partir da união de partes humanas retiradas de diversos corpos, acaba se voltando contra o seu criador.

Em, “A ilha do Dr. Moreau”, de 1896, o também inglês H.G.Wells conta a história de um cientista que faz experimentos de hibridação de espécies animais com

humanos, e as criaturas bestiais resultantes se revelam incontroláveis e também destroem quem as criou.

Ainda neste contexto, podemos falar do filme “Blade Runner”, de Ridley Scott, de 1982, no qual andróides projetados por uma empresa de biotecnologia lutam para prolongar a sua vida, previamente programada para se extinguir em poucos anos. O líder dos andróides, diante da impossibilidade de conseguir seu objetivo, assassina o dono da empresa que o projetou.

Porém, nem tudo é tão fascinante na visão daqueles que levam a Inteligência Artificial (IA) para a ficção. Baseando-se em histórias fictícias, como as citadas anteriormente, não é difícil imaginar o caos que poderá ser causado por seres de metal, com um enorme poder físico e de raciocínio, agindo independentemente da vontade humana. Guerras desleais, escravidão e até mesmo a extinção da humanidade estão no rol das conseqüências imaginadas pelos autores de ficção científica.

Por outro lado, robôs inteligentes podem ser de grande utilidade na medicina, na exploração de outros planetas, no resgate de pessoas soterradas por escombros, além de sistemas inteligentes para resolver cálculos e realizar pesquisas que poderão encontrar a cura de doenças. Como se pode notar, a Inteligência Artificial (IA) é um tema complexo. São diversos os pontos a favor e contra, e cada lado tem justificativas para suas afirmações.

De qualquer forma, o progresso das invenções e descobertas tecnológicas é irreversível e, acreditamos, continuará a trazer enormes benefícios para a humanidade, como já vem acontecendo.

Capítulo IV

4. Inteligência Artificial (IA): interagindo com robôs de conversação brasileiros

A perspectiva deste capítulo é compreender o funcionamento e conceituar os robôs de conversação, por meio de seus pesquisadores, mais especificamente Primo e Roth Coelho, (2001). Eles afirmam que os *chatterbots* brasileiros são experimentos recentes e requerem estudos mais profundos, uma vez que precisam de uma linguagem natural para interagirem.

Embora o objetivo de um *chatterbot* seja responder as perguntas de tal forma que tenhamos a impressão de estar falando com outra pessoa, e não com uma máquina, isso ainda não acontece. Após o envio de perguntas em linguagem natural, o programa consulta sua base de conhecimento e, em seguida, fornece uma resposta que tenta imitar o comportamento humano.

Também para Primo e Roth Coelho, (2001) o questionamento, “Pode uma máquina pensar?”, continua desafiando a ciência. No decorrer deste capítulo, estaremos enfocando as críticas feitas à Inteligência Artificial (IA) forte, a partir da Comunicação e das perspectivas utilizadas frequentemente pela Inteligência Artificial (IA), a saber: Teoria da Informação e Behaviorismo. Além disso, reproduzimos uma breve discussão a respeito do potencial interativo da robô Cybelle, do robô Ed, e da robô Sete Zoom.

4.1 Vôo de reconhecimento

A Inteligência Artificial (IA) é uma ciência que recebe contribuições da Neurociência, Informática, Linguística, Psicologia, Filosofia, entre outras. Todas elas pretendem oferecer uma contribuição interdisciplinar, porém encontram enormes dificuldades, pois cada uma delas quer respeitar e preservar seus paradigmas, considerados como certezas.

A partir desse cenário, pode-se logo concluir a fertilidade dos temas aí semeados e a necessidade por novos e constantes insumos que permitam às idéias, que aí florescem, acompanhar as variações do ambiente em questão.

Os debates que se travam no âmbito da Inteligência Artificial (IA) são acalorados e dificilmente conseguem chegar a um consenso. Os produtos dos projetos experimentais de alta tecnologia provocam o ceticismo dos críticos da Inteligência Artificial (IA) e, às evidências apresentadas por defensores e críticos da real potencialidade da Inteligência Artificial (IA), somam-se defensores apaixonados.

Se, há cem anos, não se podia imaginar o estágio tecnológico do início do século XXI, então não se deve duvidar do que pode vir a acontecer tecnologicamente nos próximos séculos, embora as dificuldades enfrentadas pela Robótica e Inteligência Artificial (IA) não encontrem soluções fáceis. Pode-se dizer que muitas vezes tais dificuldades mais parecem barreiras gigantescas, cujas alternativas de ultrapassagem normalmente levam a outras novas barreiras.

Existe hoje uma renovada demanda por pesquisas sobre comunicação interpessoal, tendo em vista que o ramo da Inteligência Artificial (IA) dedicada ao estudo da “linguagem natural” trabalha basicamente com simulação de diálogos.

Como linguagem e interação são objetos de estudo da comunicação, os robôs que simulam diálogos (*chatterbots*) particularmente nos interessam e fazem parte da investigação desta dissertação.

Portanto, pretendo apresentar o projeto do primeiro robô de conversação da web brasileira, a Cybelle, dentre outros robôs de conversão, como o robô Ed e a robô virtual Sete Zoom, encontrados no sites:

Robô Ed <http://www.ed.conpet.gov.br>

Robô Sete Zoom <http://www.inbot.com.br/sete>

Robô Cybelle <http://www.cybelle.com.br>, que já não se encontra mais disponível para acesso. Somente na versão americana no site <http://www.agentland.com/>

E o que são *chatterbots*? A palavra “robô” teve origem na peça “R.U.R.” de Karel Capek, escrita em 1921. A sigla era uma abreviatura para “*Rossum’s Universal Robots*”, onde *robota* quer dizer em tcheco “trabalho”. *Bot* é uma simplificação da palavra *robot*.

Para Simon Laven, 2001, um *chatterbot* é um programa com o objetivo de simular conversação, com o intuito de, pelo menos temporariamente, enganar um ser humano, fazendo-o pensar que está falando com outra pessoa.

A primeira experiência com robô de conversação, desenvolvida entre 1964 e 1966, no MIT, por Joseph Weizenbaum, foi o robô Eliza, um dos programas de Inteligência Artificial (IA) mais antigo e mais conhecido no mundo. Pode-se também dizer que é um dos programas mais estudados na história da informática, embora sua versão somente apareça na língua inglesa.

O objetivo deste pequeno programa, de apenas 204 linhas de código, é simular uma conversação entre uma psicóloga de estilo rogeriano e seu paciente (sendo assim, normalmente responde às perguntas com outras perguntas). Eliza é bastante compreensiva com seu paciente, mas não se lembra de nada do que foi dito na interação.

Por outro lado, mesmo sendo uma implementação pioneira, Eliza tem uma das personalidades mais bem definidas entre os robôs de conversação, apesar de sua simplicidade.

Segundo Hutchens, 1996, programas como Eliza e outros, baseados em seu modelo, têm eficácia em seu objetivo pois as pessoas tendem a ler muito mais significado nas curtas respostas dos *chatterbots* do que realmente existe. Isto é, há uma tendência em ludibriar-se com a simulação, em ler estrutura no caos em ser bastante tolerante com as evasivas.

Até mesmo muitos psiquiatras acreditaram que Eliza poderia ser desenvolvida até alcançar um sistema quase totalmente automatizado de psicoterapia. Algumas pessoas chegaram a acreditar que o programa demonstrava uma solução geral para o problema computacional de compreensão.

Até pouco tempo atrás, os robôs que “falavam” apareciam apenas nos filmes de ficção científica e podiam ser testados apenas por pessoas ligadas à informática. Atualmente, estão acessíveis, via web, a qualquer internauta, através de links, tais como: news bots, fun bots, commerce bots, etc.

4.2 Cybelle, a primeira robô de experiência brasileira na web

Considerada como uma pioneira, em se tratando de *chatterbots* em língua portuguesa, Cybelle, criada por Alex Primo e Roth Coelho, em 2001, apresenta um mecanismo de análise e inferência e uma base de conhecimento. É um sistema baseado na interação por estímulo-resposta e apresenta uma estrutura semelhante à de Eliza.

O mecanismo (engine) é um programa que analisa os inputs do internauta, cuja relação entre mecanismo e conhecimento é regida por uma lógica estímulo-resposta. Isto é, o input do internauta é analisado, buscando-se por estímulos previstos ou suas combinações. Para um mesmo estímulo, existe mais de uma resposta, possibilitando o sorteio entre diversas alternativas, evitando que a robô se repita caso o mesmo tópico seja abordado mais de uma vez.

Se o internauta escrever algo que a robô não saiba responder adequadamente, uma mensagem padrão é sorteada (muitas vezes, trata-se de uma evasiva, um convite para discutir outro assunto, uma mensagem irônica ou uma menção a uma história ficcional da qual Cybelle participe).

Existem outras características que qualificam o desempenho da *chatterbot*. A interface gráfica do *site* apresenta uma ilustração sensual de Cybelle, que pode ser manipulada, e um arquivo de áudio com a “voz” da robô, que diz: “Oi! Meu nome é Cybelle. Eu preciso muito de sua ajuda”.

A partir daí, ela demonstra sensações de tristeza por estar presa ao computador e percebe-se que ela não pode criar nenhuma resposta, apenas oferece aquelas que estejam em sua programação. Ela também não armazena informações, portanto não pode lembrar-se de uma questão anterior à que estiver respondendo.

Desse modo, ela pode se repetir e oferecer respostas descontextualizadas. Para minimizar isso, que prejudica a simulação, várias são as respostas que não são específicas, ainda que tratem do tema proposto; muitas vezes ela responde com uma pergunta, ou convida o internauta a conversar sobre um assunto que ela possui cadastrado; ou lamenta sua condição de robô e pede ajuda ao visitante.

Na verdade, as dificuldades encontradas por *chatterbots* e que os levam a falar coisas sem sentido é que para seu funcionamento, eles utilizam-se de uma lógica que dispensa o compreender. Por exemplo o robô Ed apresenta-se como um robô de conversação, como demonstra a simulação abaixo:

Ed: Olá! Meu nome é Ed, estou no site do conpet para ajudar na preservação de energia e de outros recursos naturais. Sobre o que você quer conversar?

Você: Quem você é?

Ed: Eu sou Ed, o robô. Vim de outro planeta com a missão de ajudar na preservação de energia e de outros recursos naturais. Sobre o que quer conversar?

Você: Nós nos conhecemos?

Ed: Ficha não possui resposta

Você: Você vem sempre aqui?

Ed: Sempre. E você ?

Você: Podemos ser amigos?

Ed: Claro que sim, adoro fazer amigos!

Você: Você gosta de mim?

Ed: Claro, você é uma pessoa maravilhosa.

Você: Podemos nos falar mais tarde?

Ed: Mais tarde acho que ainda vou estar aqui.

Os usuários reagem de maneiras diferentes à interação com o robô Ed. Alguns se iludem com a possibilidade de o robô ter autonomia para dar respostas, outros, não.

Conforme Lévy, 1998, todas as atividades de construção e exploração de modelos mentais que se dão no raciocínio espontâneo são trocadas pela execução de regras formais sobre proposições. A partir disso, pode-se lembrar do trocadilho de computadores “estúpidos, mas perfeitamente lógicos”.

Para ilustrar o caráter artificial de tal lógica recorre-se a um exemplo:

- (1) “Se nevar amanhã, iremos esquiar”
- (2) “Se formos esquiar ficaremos contentes”
- (3) “Nós não ficaremos contentes”

Segundo essas proposições, podem-se deduzir a partir de (2) e (3) uma quarta proposição: (4) “Nós não iremos esquiar”. Finalmente, a partir de (1) e (4) pode-se deduzir (5): “Não nevará amanhã”.

O raciocínio lógico é correto, mas não se podem defender com segurança as deduções retiradas. Pode haver muitas outras razões para que o descontentamento ocorra além de uma possível falta de neve. Isto é, muitos são os modelos mentais que podem ser construídos a partir da terceira proposição.

Porém, o logicismo apresentado se reduz a (4) e (5), já que se resume às premissas explícitas. Já o raciocínio espontâneo recorre a todos os conhecimentos que tem sobre a situação. Assim, extrapola as premissas explícitas relacionando um conjunto de conhecimentos muito mais vasto.

Mesmo Lévy, 1990, um otimista confesso, sugere que a lógica é uma tecnologia intelectual datada, que é baseada na escrita e não no pensamento natural. De fato, a maioria dos raciocínios humanos não se utiliza de formalismos lógicos e regras de dedução.

Daí a impossibilidade de se chegar a uma simulação profunda da inteligência humana, pois esta não está apenas e simplesmente baseada na lógica formal. O que a Inteligência Artificial (IA) pôde produzir é uma nova tecnologia intelectual, como os sistemas especialistas, embora ainda não tenha se chegado a uma réplica do pensamento humano.

A questão da autonomia também é de grande importância para os estudos de Inteligência Artificial (IA). Na verdade, muitos são os programas cujos desenvolvedores alegam ser autônomos.

Baseando-se em Maturana e Varela, 1997, pode-se afirmar que a autonomia apresenta uma complexidade muito maior do que supõe grande parte dos programadores em informática.

Para tanto, vale acompanhar, ainda que brevemente, a diferenciação que os autores fazem entre máquinas autopoéticas, que são autônomas e apresentam auto-criação e máquinas alopoiéticas que não têm autonomia, como o automóvel, por exemplo.

As primeiras (por exemplo, o homem) apresentam autonomia, pois têm todas as suas mudanças subordinadas à conservação de sua própria conservação organizacional (independente da profundidade das transformações). Já as máquinas alopoiéticas não são autônomas, pois as mudanças que sofrem em seu funcionamento subordinam-se à produção de algo diferente delas mesmas.

Nas máquinas autopoieticas a criação das próprias fronteiras define o sistema como uma unidade e especifica o domínio das operações da rede. Por outro lado, as fronteiras das máquinas alopoieticas são determinadas por fatores independentes.

O desempenho de um *chatbot* (uma máquina alopoietica) depende daquilo que seu programador permite e limita.

Na verdade, muitos debates sobre Inteligência Artificial (IA) contaminam-se de metáforas por demais otimistas e imprecisas, gerando fantasias que passam a ser vistas como reproduções perfeitas do comportamento humano.

Questões como autonomia e aprendizado são alguns dos conceitos tratados de forma superficial por pesquisadores de Inteligência Artificial (IA).

Um internauta escreveu algo como “Fulano cujo nome é opcional é um idiota”. Quando uma pessoa com tal nome começou a dialogar com Cybelle, a robô disparou: “Você é um idiota”. Isso demonstra que o mecanismo teria “aprendido” a informação, mas não tinha como contextualizar, discernir ou ponderar sobre o conteúdo “aprendido”, ou seja, aprender não é apenas armazenar novos inputs e relacioná-los por associação simples a determinados outputs.

Porém, a lógica behaviorista reduz o processo de aprendizado a um automatismo mecânico, deixando de lado a construção ativa de significados e a dinâmica cognitiva.

Se, por um lado, esse encaminhamento viabiliza a produção de sistemas especialistas em Inteligência Artificial (IA), por outro, pouco pode contribuir para o estudo da cognição humana.

Para Piaget, 1996, nenhum conhecimento, mesmo que através da percepção, é uma simples cópia do real. O conhecimento tampouco se encontra totalmente determinado pela mente do indivíduo. É, na verdade, o produto de uma interação entre

elementos. Sendo assim, conhecer é agir sobre o real, transformando-o (em aparência ou na realidade). Percebe-se então que o conhecimento não é apenas uma acumulação mecânica de dados.

Sabe-se que a memória é condição necessária para a inteligência, porém, a memória não é como um baú que mantém intactos e estáticos os objetos que ali se depositam. Se assim fosse, bastaria recorrer ao baú para encontrar determinado objeto-memória ali depositado, com suas características e seu lugar mantidos. No entanto, o significado da memória é recriado a todo instante.

Outra questão é o estatuto de verdade que possuem nossas lembranças. Piaget (1990, p. 241) comenta que veio a descobrir aos quinze anos que uma de suas lembranças de infância mais antigas e mais vivas (uma tentativa de seqüestro, quando estava sendo levado para passear em um carrinho de bebê), era na verdade uma mentira que sua babá lhe tinha contado e que ele sempre teve como fato real.

Diante disso, pode-se perguntar como pode decidir um robô sobre o que é verdade, relevante, fantasia, etc?

Na prática, a implementação de robôs não possui tais dinâmicas. Todas as associações contidas no “cérebro” de Cybelle são, na verdade, estáticas. Quando um internauta fizer perguntas que foram previstas pela equipe de programação, as respostas de Cybelle parecerão sempre adequadas e até contextualizadas. Quando não tiverem sido previstas, a robô não tem como criar de forma autônoma respostas ainda não cadastradas e nem mesmo aprender com o internauta.

Atualmente, o mecanismo e o conhecimento da robô estão sendo aperfeiçoados.

Futuramente, a robô poderá reconhecer o sujeito de uma frase indicado por pronomes e, também, poderá permanecer falando sobre um mesmo tópico, enquanto o internauta desejar, sem que ele precise ser explícito sobre isso.

Esses novos recursos permitirão que a simulação seja ainda mais eficiente, qualificando a fantasia interativa criada.

Pode-se especular que com essa progressão, em algum momento, será difícil para um internauta reconhecer se do outro lado da linha encontra-se uma outra pessoa ou um *chatbot*. Contudo, isso não significará que esse *chatbot* é inteligente.

4.3. Informação, caixa preta e interação

O projeto Cybelle contribuiu principalmente para demonstrar que a mente e a comunicação humana são muito mais do que estímulo-resposta e transmissão de sinais.

Além disso, seus autores buscaram criticar a Inteligência Artificial (IA) Forte através de um projeto de Inteligência Artificial (IA) real e de resultados, no mínimo, curiosos.

Com esse direcionamento, pode-se detectar que grande parte dos projetos em Inteligência Artificial (IA) tem em sua base a inspiração da Teoria Matemática, da Teoria da Informação e do Behaviorismo. Embora a Teoria da Comunicação Contemporânea tenha criticado tais orientações, o debate ainda não se encerrou. Pelo contrário, parece que uma nova etapa acaba de começar.

Nesse sentido, faz-se hoje mais do que necessário atualizar as críticas a esses paradigmas, pois como as metáforas se tornam hoje cada vez mais sofisticadas tecnologicamente, o debate sobre o conhecimento e a comunicação humana pode mais uma vez ser novamente simplificado em demasia.

Implementar a mecanização da comunicação é bastante criticável, uma vez que inúmeras vezes se levantam contra as teorias mecânicas da comunicação. Tal empreitada se alicerça na Teoria da Informação e no Behaviorismo. Quando se pensava ter tal discussão ultrapassada, ela volta renovada e com nova vestimenta tecnológica, enfeitada com animações em 3D e sintetização de voz.

No entanto, o modelo transmissionista da Teoria da Informação de Shannon e Weaver, 1962, (desenvolvido em 1949) não visava demonstrar o processo de comunicação humana e sim processos eletrônicos. Shannon estava preocupado em

estudar como transmitir sinais de forma correta e eficiente, principalmente através da telefonia.

Era, pois, uma descrição de interesse da engenharia de telecomunicações, com a intenção de maximizar a eficiência da transmissão (medida em bits por segundo) entre o emissor e o receptor, e de garantir a fidelidade da informação. Esse corpo teórico foi generalizado para outros contextos comunicativos.

Isso se deve, inicialmente, ao último capítulo do livro “Teoria Matemática da Comunicação”, onde Weaver pretendia expandir o poder explicativo da teoria em questão para outros eventos comunicativos.

Por outro lado, Shimith (1970, p.18) diz que o modelo de Shannon e Weaver não pode ser visto como um modelo de processo de comunicação humana já que ele somente descreve “a cadeia transmissora de informação nos termos de suas partes componentes estacionárias”. Isto é, o modelo proposto estuda os sinais em seu estado físico, ignorando os níveis semânticos e pragmáticos da comunicação.

Portanto, o projeto de simular o diálogo humano a partir de uma perspectiva desenvolvida para o estudo técnico da transmissão telefônica é, em sua gênese, problemático.

A Teoria Matemática da Comunicação e o Behaviorismo aproximaram a Comunicação e a Psicologia dos métodos das ciências duras, pagando o preço de se afastar da espontaneidade do comportamento humano, ao procurar a delimitação de leis e redução do processo sistêmico não-somativo e interdependente a relações formais, discretas, quantificáveis e previsíveis.

A tradição Behaviorista (comportamental) em Psicologia defende o estudo de padrões comportamentais entre estímulos e respostas. E é essa a lógica por trás de

chatbots como os robôs Ed, Sete Zoom e Cybelle, pois se trabalha através da associação de certas respostas a certos estímulos.

O Behaviorismo trabalha com a previsão de comportamentos a partir de condições anteriores. Ou seja, desconsidera o poder criativo do sujeito. Na verdade, o que se passa entre um estímulo e uma resposta não interessa aos behavioristas, é uma caixa preta na qual não interessa ver o seu interior.

Segundo Del Nero, 1993:

“A tentativa behaviorista de eliminação das variáveis intermediárias entre o estímulo e a resposta poderia ter conseguido modelar o comportamento não fosse a forte possibilidade de existência de não-linearidades e eventualmente caos no processo intermediário. A modelização dos dois extremos da cadeia, isto é, do estímulo e da resposta, se ganha em observabilidade e objetividade, só pode vingar quando o sistema for linear e, ainda assim, bem comportado.” Del Nero (1997, p. 157).

Esse autor observa ainda o problema semântico enfrentado pela tradição comportamental. Se o positivismo busca matematizar os objetos e processos em estudo, a questão da representação traz uma nova problemática.

Inicialmente, representar é entendido como rerepresentar o mesmo conteúdo apresentado sob outra forma. Porém, a atividade cognitiva humana pode expressar objetos que não tenham ligação com o mundo físico, isto é, não tenham existência concreta. Essa instância constitui uma barreira intransponível para uma tradução radical do vocabulário mental e de suas categorias a similares físicos ou quantificáveis.

Portanto, a inventividade e a criação não podem ser contempladas por esse modelo.

Por outro lado, Cybelle é uma personagem virtual que desperta a curiosidade de quem a conhece. É interessante perceber (através dos diálogos gravados) como, mesmo sendo um programa “frio”, ela mexe com as pessoas com quem conversa, despertando emoções.

A robô Cybelle mantém uma conversação sobre diferentes assuntos, simulando um comportamento dialógico humano. Além disso, diferentemente de outros *chatterbots* americanos, Cybelle simula ter emoções, é deprimida, pois se sente presa ao computador; se irrita com comentários maldosos e, muitas vezes, se apaixona por quem conversa com ela.

A implementação da *chatterbot* Cybelle e a análise dos resultados apurados levantam uma série de questões relevantes para a comunicação, como as que seguem.

Pode um robô pensar e interagir como um ser humano? A interação homem-máquina através da linguagem natural se iguala a um encontro interpessoal humano? Pode um robô compreender um texto? O cérebro e ou a mente humana podem ser reproduzidos artificialmente?

Se assumirmos uma postura defensora da Teoria da Informação e do Behaviorismo (tão comum em Inteligência Artificial (IA)) e entendermos que o teste de Turing é um método adequado para verificar a inteligência e o “pensamento” de uma máquina, ficaríamos tentados a inferir que, tendo em vista o sucesso de alguns robôs de conversação, o computador pode ou brevemente poderá pensar e interagir exatamente como um ser humano.

Nesse caso, estaríamos supondo que a mente humana (um grande banco de dados de armazenamento cumulativo e estático) não passa de um mecanismo previsível de associação técnica que reage a inputs, oferecendo outputs relacionados a essas entradas, e que a mente é determinada de fora para dentro. Sendo assim, o cérebro não seria nada mais que um computador.

Contudo, a questão e sua solução não são tão simples. Se Cybelle pode, durante certo tempo, parecer fazer sentido nos diálogos que conduz, não quer dizer que seja

inteligente ou que pense. Como aponta Searle (1997, p. 36), o problema não é “mesmo-comportamento-portanto-mesmos-fenômenos-mentais”.

Com relação à interação, pode-se dizer que Cybelle é uma interface de hipertexto diferenciada. Trata-se pois de uma interação reativa, já que a relação internauta/site encontra-se determinada, isto é, sempre que se clicar no link X o internauta é conduzido à página Y. No diálogo com Cybelle não existem links para serem clicados, mas as palavras usadas (e suas combinações) servem de âncoras para certos textos pré-determinados.

Logo, Cybelle é, em última instância, um hipertexto onde o internauta envia novos textos que se associam aos estímulos que ele dá entrada. Com esse perfil, entendemos que a interação mantida com a *chatterbot* brasileira é reativa, pois acontece via previsão e programação. Mesmo a entrada de frases totalmente desconhecidas pelo internauta é prevista .

A conversa de um internauta com Cybelle não configura uma interação do tipo mútua, pois não há implicações e transformações recíprocas. As respostas de Cybelle afetam o internauta, mas ela jamais se transforma frente ao comportamento daquele que a argui.

Se o internauta repetir a mesma seqüência de perguntas, obterá (exceto nas situações que mais de uma resposta foi programada para dado estímulo) as mesmas respostas. O comportamento de Cybelle foi previsto e definido mesmo antes da interação com o internauta começar.

Finalmente, a *chatterbot* não compreende o que lhe escrevem. Suas reações são automatizadas por uma relação estímulo-resposta. E como não existe interpretação ou compreensão, Cybelle também não lembra o que foi dito anteriormente. Sendo assim, a

conversa com Cybelle está sempre começando, pois não existe evolução ou construção cooperada. Por mais que se trate de uma interação, ela é de um tipo restritivo e limitado.

Concluimos, então, que ainda existem muitas limitações nas questões relacionadas à Inteligência Artificial (IA). Uma das limitações é o próprio desconhecimento que temos sobre os mecanismos da inteligência natural humana, seja dos processos cerebrais ou dos processos mentais e da consciência. Bittencourt (2006, pp.170-175).

Ainda não se sabe quais são os limites da capacidade de simulação e reprodução, nos computadores, do funcionamento dos mecanismos humanos de inteligência, pois o cérebro possui estruturas completamente diferentes e de complexidade muito maior do que os computadores atuais.

A maior limitação atual da Inteligência Artificial (IA) é não conseguir realizar tarefas que não podem ser expressas por modelos matemáticos ou lógicos, como tarefas que requerem intuição ou capacidade de relacionar informações aparentemente sem conexão. Bittencourt (2006, p.220)

Por meio de uma técnica conhecida como “Processamento de Linguagem Natural”, aliada a uma gigantesca base de conhecimento sobre assuntos do mundo real, os *chatterbots* atuais podem ser utilizados para entretenimento, suporte online, portais corporativos, jogos, projetos educacionais, culturais, treinamentos, call centers e auxílio no ensino a distância. A aplicação depende apenas do conteúdo criado por redatores e programadores que ensinam ao personagem a base de conhecimento.

Dentre os benefícios que, atualmente, um *chatterbot* proporciona, destacamos:

- Baixo custo por dispensarem atendentes reais;

-facilidade para os visitantes de um site encontrarem informações (basta perguntar, ao invés de precisar navegar por todo o site e ler grandes quantidades de texto);

-navegação mais excitante, estimulando os visitantes a conhecerem melhor outras áreas do site.

CONSIDERAÇÕES FINAIS

O objetivo da presente dissertação foi seguir uma perspectiva de resumir e compreender a questão da Inteligência Artificial (IA), na Filosofia da Mente, como ciência pertencente à natureza da cognição.

Os teóricos da Inteligência Artificial (IA), a partir de 1956, fundamentaram-se na idéia de que é possível modelar o funcionamento da mente humana através do computador, de tal forma que o cérebro fosse visto como uma máquina e os neurônios, como processadores de informação. Considerando tal enfoque, seria perfeitamente possível modelá-lo dentro do computador. No entanto, ainda hoje, pesquisadores se defrontam com essa complexa questão, tentando desvendar seus mistérios.

No início, os estudos sobre Inteligência Artificial (IA) buscavam apenas uma forma de reproduzir a capacidade humana de pensar, porém, percebendo que esse ramo da ciência poderia evoluir consideravelmente, pesquisadores e cientistas abraçaram a idéia de fazer com que uma máquina pudesse reproduzir não só a capacidade de um ser humano pensar, como também a capacidade de sentir, de ter criatividade, e se aperfeiçoar utilizando a linguagem natural.

Diante desse desafio e bastante incentivados por ele, fomos em busca das muitas respostas encontradas pelos cientistas, pelos filósofos da mente e demais estudiosos da Inteligência Artificial (IA), do cérebro, da consciência e da cognição, para tentar responder aos seguintes questionamentos: “Será que as máquinas vão ser capazes de revelar comportamento inteligente e substituir os seres humanos em tarefas criativas?” “Será que desenvolverão algum tipo de pensamento e sentimento que os seres humanos têm?” “Pode uma máquina pensar?”

Cabe dizer que os estudiosos da Inteligência Artificial (IA) não querem simplesmente imitar o que o homem faz e, sim, estudar minuciosamente as propriedades da inteligência no homem, no animal e no robô, objetivando alcançar benefícios para a humanidade, nas diferentes áreas do conhecimento.

Ao longo desta pesquisa, fundamentando e conceituando os princípios da Inteligência Artificial (IA) e suas razões históricas, compreendemos que até mesmo o próprio conceito de inteligência permanece em estudo, pois ainda não existe uma definição precisa para ele; existem, sim, formas de contribuir cada vez mais com o seu significado.

A Inteligência Artificial (IA) é subdividida em Inteligência Artificial (IA) fraca e Inteligência Artificial (IA) forte.

A Inteligência Artificial (IA) fraca pressupõe que uma máquina não é capaz de verdadeiramente raciocinar e resolver problemas. Essa máquina, com características de inteligência, age como se fosse inteligente, mas não tem autoconsciência, nem noção de si. Atualmente, tais máquinas são os nossos computadores, amplamente utilizados em todas as áreas do conhecimento humano.

A Inteligência Artificial (IA) forte pressupõe a criação de inteligência, baseada em computador, que consiga raciocinar e resolver problemas. Ela é classificada como auto-consciente ou, pode-se dizer, categorizada em sistemas que pensam como um ser humano, agem como um ser humano, pensam e agem racionalmente. Nas décadas de 1960 e 1970, alguns pesquisadores trabalharam intensamente, acreditando em tal possibilidade.

Da atualidade, embora tenhamos pesquisado vários autores, escolhemos expor os pensamentos de Searle e João de Fernandes Teixeira, por considerar que eles se aproximam mais daquilo que acreditamos e inferimos de todas as nossas leituras.

Para Searle, os computadores não são capazes de apreender a natureza significativa própria dos estados mentais. Segundo ele, o computador é simplesmente sintático e não compreende sequer os significados da sua própria linguagem, limitando-se a manipular símbolos.

A seguir, reproduziremos as conclusões de Searle (1984, pp.48-51), que entram em confronto com a pretensão da teoria computacional, que acredita ser possível reproduzir a mente, produzida por qualquer artefato físico, desde que bem montado.

“Conclusão I – Nenhum programa de computador é, por si mesmo, suficiente para dar uma mente a um sistema. Os programas, em suma, não são mentes e por si mesmos não chegam para ter mentes.”

“Conclusão II – A maneira como a função cerebral causa mentes não pode ser apenas em virtude da ativação de um programa de computador.”

“Conclusão III – Tudo ou mais que causou mentes deveria ter poderes causais, pelo menos, equivalentes aos do cérebro.”

“Conclusão IV – Para qualquer artefato que pudéssemos construir, o qual tivesse estados mentais equivalentes aos estados mentais humanos, a realização de um programa de computador não seria por si só suficiente. Antes, o artefato deveria ter poderes equivalentes aos poderes do cérebro humano.”

Para João de Fernandes Teixeira, desvendar a complexidade que envolve as múltiplas conexões do cérebro humano, é uma tarefa ainda incompleta. Nesse caso, como representar e reproduzir aquilo que ainda não compreendemos?

Em seu livro “Como ler a filosofia da mente” (2008, p.62), ele diz que:

“Não há dúvida de que modelar a mente ou descobrir os mecanismos que produzem a consciência, por meio de um *software*, constitui uma estratégia inviável. Esta parece ter sido a lição imediata da neurociência cognitiva e do movimento em direção à redescoberta do cérebro, que se iniciou na década de 1990. Será, porém, que a neurociência cognitiva poderá abandonar completamente a utilização de modelos computacionais para estudar o cérebro? A resposta é negativa. Segundo a neurociência cognitiva, não são os modelos computacionais que devem ser abandonados, mas a pretensão de, a partir deles, achar que podemos replicar a mente humana. Isto nos leva, contudo, a uma outra questão polêmica: serão as

características biológicas do cérebro impossíveis de ser replicadas por uma máquina? Consciência é um tipo de experiência subjetiva. Não sabemos ainda quais são seus correlatos no nível cerebral. Não sabemos se sua produção está relacionada com algum tipo específico de arquitetura neural ou com algum tipo de frequência oscilatória de alguns grupos de neurônios que permite um tipo especial de codificação de informação.”

A presente investigação também se preocupou em estudar os *chatterbots* que simulam diálogos tal como os chats de conversação, ao quais estamos acostumados.

Os *chatterbots* surgiram como alternativa aos sistemas computacionais convencionais, que distanciavam o usuário do meio virtual, por não fornecerem um ambiente de comunicação que fosse próximo ao utilizado na linguagem natural.

Dessa forma, seu uso na educação tornou-se bastante interessante. A possibilidade de prover continuamente a capacidade de atendimento a dúvidas dos alunos surgiu com o uso de *chat robots*, que tentam replicar o papel de um tutor virtual, oferecendo novas possibilidades de suporte ao aprendizado.

Neste trabalho, apresentamos o resultado da pesquisa com os *chatterbots*, na área de Inteligência Artificial (IA), destacando suas características e sua importância nos dias atuais, por meio de um experimento de integração de um *bot* (robô) a um ambiente virtual de ensino.

Salientamos também o potencial elevado dos *chatterbots*, em se tratando de formas de utilização, principalmente na internet.

Pretendemos, ao escolher e levar adiante o estudo das novas tecnologias, reverenciar os pesquisadores, cientistas e filósofos, que empenharam e empenham suas vidas em transcender os limites do conhecimento para superar as dificuldades e melhorar a vida dos seres deste planeta.

Referências Bibliográficas

- ASIMOV, I. (2007) **Histórias de robôs**. Porto Alegre. L&PM.
- BERGER, P. L. T. (1998) **A construção social da realidade**. Petrópolis: Vozes.
- BITTENCOURT, G. (2006) **Inteligência Artificial Ferramentas e Teorias**. Santa Catarina. UFSC.
- CAIDIN, M. (1972) **Cyborg**. New York: Arbor House.
- COELHO, H. (1999) **“Sonho e Razão”, 2. a. ed. Lisboa: Editora Relógio d’ água.**
- COUCHOT, E. (1993) **Da representação à simulação: evolução das técnicas e das artes de figuração**. In: PARENTE, André (org). **Imagem máquina**. São Paulo: Editora 34.
- DAMÁSIO, A. (1996) **O erro de descartes: emoção, razão e cérebro humano**. São Paulo: Companhia das letras.
- DEL NERO, H. C. (1997) **O Sítio da mente: pensamento, emoção, e vontade no cérebro humano**. São Paulo: Collegium Cognitio.
- DIAS, C. (1999) **Quase melhor que o original**. Super Interessante, São Paulo, ano 13, n.1, p. 42-46.
- DOUGLAS, M. (1991) **Pureza e perigo**. Lisboa: Edição 70.
- DREYFUS, H.L. (1997) **What Computers Still Can’t Do – A critique of artificial reason**. The MIT Press, Cambridge, Massachusetts.
- DURKIN, J. (1994) **Expert Systems Design and Development**. Prentice Hall, Durkin.
- ESCOBAR, A. (2000) **Welcome to cyberia – notes on the antropology of cybercultura** In Bell, David; Kennedy, Bárbara M. **The cybercultures reader**. London: Routledge.

- GARDNER, H. (1994(a)) **Estruturas da mente: a teoria das inteligências múltiplas**. Porto Alegre: Artes Médicas Sul.
- GARDNER, H. (1995) **Inteligências múltiplas: a teoria na prática**. Porto Alegre: Artes Médicas.
- GEERTZ, C. (1989) **A interpretação das culturas**. Rio de Janeiro: Guanabara Koogan.
- GOFFMAN, E. (1983) **A representação do eu na vida cotidiana**. Petrópolis: Vozes.
- HARAWAY, D. (2000) **Manifesto ciborgue: ciência, tecnologia e feminismo-socialista no final do século XX**. In: SILVA, Thomas Tadeu da (org.). *Antropologia do ciborgue*. Belo Horizonte: Autêntica.
- HUTCHENS, J. L. (1996) **How to pass the Turing Test by cheating**. Disponível em <http://www.ciips.ee.uwa.edu.au/hutch/research/papers/> acessado em 03 de Maio 2008.
- KAERNBACH, C. (2006) **No Virtual Mind in the Room**. Institut Fur Psychologie, Karl- Franzens- Universitat Graz, Schubertstr. 51a. 8010 Graz, Áustria. Disponível em <http://assc.caltech.edu/assc10/>
- LAVEN, S. (2001) Disponível em: www.cin.ufpe.br/~tg/2002-1/abs2-proposta.doc acesso em 20/12/2007.
- LEVI-STRAUSS, C. (2000) **Mito e significado**. Lisboa: Edições 70.
- LEVI-STRAUSS, C. (2002) **O pensamento selvagem**. Campinas: Papirus.
- LÉVY, P. (1990) **As Tecnologias da Inteligência**. Editora: Instituto Piaget, Portugal. 1ª Edição.
- LÉVY, P. (1998) **A ideografia dinâmica: rumo a uma imaginação artificial?** São Paulo: Loyola.

- LÉVY, P. (1998) **A máquina universo: criação, cognição e cultura informática.** Porto Alegre: Artes Médicas.
- LIMA, C. M. ; LABIDI, S. (1999) **Introdução à inteligência artificial.** Disponível em <http://www.elo.com.br> acessado 10 fevereiro de 2007.
- MATURANA, H. e Varela, F. (1997) **De máquinas e seres vivos: autopoiesi – a organização do vivo.** 3 ed. Porto Alegre: Artes Médicas.
- MINSKY, M. (1985) **The Society of Mind.** New York, USA: Touchstone.
- NEGROPONTE, N. (1995) **A vida digital.** São Paulo: Schwarez.
- NEISSER, U. (1967) **Cognitive Psychology.** New York: Appleton- Century-Crofts.
- PIAGET, J. (1990) **A formação do símbolo na criança: imitação, jogo e sonho, imagem e representação.** Rio de Janeiro: LTC.
- PIAGET, J. (1996) **Biologia e conhecimento.** 2 ed. Petrópolis: Vozes.
- PRIMO, A. F. T. e COELHO, L. R. (2001) **A chatterbot Cybelle: experiência pioneira no Brasil. In: Mídia, Textos & Contextos.** Porto Alegre: Edipucrs.
- PUTNAM, H. (1961) **Minds and Machines.** Em: Hook, S. (org). *Dimensions of mind.* New York: Coolier.
- RAPPORT, N. e OVERING, J. (2000) **Cybernetics.** In: RAPPORT, Nigel; OVERING, Joanna. **Social and cultural anthropology: the key concepts.** London: Routledge.
- SAHLINS, M. (1990) **Ilhas de história.** Rio de Janeiro: Zahar.
- SCOTT, R.(1982) **Blade Runner- O caçador de andróides.**
- SEARLE, J. R. (1984) **Mente, Cérebro e Ciência.** Lisboa. Edições 70.
- _____. (1992) **The Rediscovery of the Mind.** Cambridge: MIT Press.

- _____. (1997) **A redescoberta da mente**. São Paulo: Martins Fontes.
- SFEZ, L. (1994) **Crítica da comunicação**. São Paulo, SP: Loyola.
- SHANNON, C. & Weaver, W. (1962) **The mathematical theory communication**. Urbana, IL: University of Illinois.
- SHIMITH, R.G. (1970) **Speech Communication: theory and models**. New York: Harper & Row.
- SPIELBERG, S. (2001) **A.I. Inteligência Artificial**, Filme de Ficção Científica.
- STERLING, B. (1992) **The hacker crackdown: law and disorder on the electronic frontier**: New York: Bantam Books.
- STERNBERG, R. J. (1996) **Psicologia Cognitiva**. Porto Alegre: Artes Médicas Sul.
- STRATHERN, M. (1992) **Reproducing the future: anthropology, kinship and the new reproductive technologies**. Manchester: Manchester University Press.
- TAUSSIG, M. (1993) **Mimesis and alterity: a particular history of senses**. New York: Routledge.
- TEIXEIRA, J. de F. (1990) **O que é Inteligência Artificial**. São Paulo. Ed. Brasiliense.
- _____. (1994) **O que é Filosofia da Mente**. 1ª Edição. São Paulo: Brasiliense.
- _____. (1995) **Cognitivism e Teorias da Consciência**. In: **Carvalho, MCM (Org). A filosofia analítica no Brasil**. Campinas, SP: Papirus.
- _____. (1998) **Mentes e Máquinas. Uma introdução a Ciência Cognitiva**. Porto Alegre. Artes Médicas.
- _____. (2000) **Mente, cérebro e cognição**. 2ª Edição. Rio de Janeiro: Petrópolis.
- _____. (2004) **Filosofia e Ciência Cognitiva**. Petrópolis, RJ: Vozes.

_____. (2008) **Como ler a filosofia da mente**. São Paulo, Paulus.

_____. **Filosofia na era digital**. O teste de Turing. São Paulo: Escala. Filosofia ciência&cognição nº18- 2008.

TUCHERMAN, I. (2004) Disponível em <http://www.fazendogenero.ufsc.br/artigos> acessado em 02/12/2007.

VINCENT, J. F. V. (2004) **Stealing ideas from nature**. Disponível em <http://www.bath.ac.uk/mech-eng/biomimetics.pdf> acessado em 23 Janeiro 2008.

WEBER, R. (1998) **Intelligent Jurisprudence Research**. Florianópolis: UFSC.

WIENER, N. (1948) **Cybernetics: or the control and communication in the animal and the machine**. Massachusetts Institute of Technology.

WIENER, N. (1984) **Cibernética e sociedade: o uso humano de seres humanos**. São Paulo: Cultrix.

Disponível em: http://pt.wikipedia.org/wiki/conhecimento_humano acesso 02 Maio 2008.

Disponível em: http://pt.wikipedia.org/wiki/inteligencia_artificial acesso em 05 Abril 2008.

Disponível em: <http://www.wikipedia.org/wiki/sintaxe> acesso em 17/11/2008.

Disponível em: Robô Ed <http://www.ed.conpet.gov.br> acesso em 19/01/2009.

Disponível em: Robô Sete Zoom <http://www.inbot.com.br/sete> acesso em 19/01/2009.

Disponível em: Robô Cybelle <http://www.cybelle.com.br>, <http://www.agentland.com/> acesso em 19/01/2009.

Disponível em: <http://pt.wikipedia.org/wiki/ENIAC> acesso em 23 Janeiro 2009.

Disponível em: http://eis.bris.ac.uk/~plajb/research/papers/Kuhn_for_DLB.pdf acesso

13 Fevereiro 2009