

**UNESP – Universidade Estadual Paulista
Campus de Marília
Faculdade de Filosofia e Ciências
Programa de Pós-graduação Interdisciplinar em Filosofia**

**ABORDAGENS BIO-INSPIRADAS APLICADAS AO ESTUDO DA COGNIÇÃO:
UM ENCONTRO ENTRE BIOLOGIA, PSICOLOGIA E FILOSOFIA**

Luís Henrique Féres Junqueira

Marília – 2006

Junqueira, Luís Henrique Féres

J95a Abordagens bio-inspiradas aplicadas ao estudo da
cognição: um encontro entre Biologia, Psicologia e Filosofia /
Luís Henrique Féres Junqueira. -- Marília, 2006.
114 f. ; 30 cm.

Dissertação (Mestrado em Filosofia) - Faculdade de
Filosofia e Ciências, Universidade Estadual Paulista, 2006.

Bibliografia: f. 90-105

Orientadora: Prof^a. Dr^a. Maria Candida Soares Del-Masso

1. Cognição. 2. Vida. 3. Ciência Cognitiva. 4. Inteligência
Artificial. 5. Filosofia. I. Autor. II. Título.

CDD 153.4

LUÍS HENRIQUE FÉRES JUNQUEIRA

**ABORDAGENS BIO-INSPIRADAS APLICADAS AO ESTUDO DA COGNIÇÃO:
UM ENCONTRO ENTRE BIOLOGIA, PSICOLOGIA E FILOSOFIA**

Dissertação apresentada ao Programa de Pós-graduação Interdisciplinar em Filosofia da Faculdade de Filosofia e Ciências da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de Marília, para obtenção do título de Mestre em Filosofia, Área de Concentração em Filosofia da Mente, Lógica e Epistemologia.

Orientadora: Dra. Maria Candida Soares Del-Masso

Marília – 2006

LUÍS HENRIQUE FÉRES JUNQUEIRA

**ABORDAGENS BIO-INSPIRADAS APLICADAS AO ESTUDO DA COGNIÇÃO:
UM ENCONTRO ENTRE BIOLOGIA, PSICOLOGIA E FILOSOFIA**

BANCA EXAMINADORA

Prof. Dra. Maria Candida Soares Del-Masso
Faculdade de Filosofia e Ciências, Marília, UNESP

Prof. Dr. Alfredo Pereira Júnior
Instituto de Biociências, Botucatu, UNESP

Prof. Dr. Gustavo Maia Souza
Laboratório de Ecofisiologia, Presidente Prudente, UNOESTE

*Dedico este trabalho à minha esposa,
Vanda, anjo de luz que Deus colocou em minha vida.*

Agradecimentos

Agradeço a todas as pessoas que, direta ou indiretamente, contribuíram para a elaboração e conclusão deste trabalho. Em particular, agradeço a todas as pessoas do programa de pós-graduação de Marília, aos professores, colegas e funcionários.

Deixo um especial agradecimento à Profa. Maria Eunice, Profa. Mariana, Profa. Betty, Prof. Hércules, Prof. Marar, Prof. Kester e Prof. Pim, por suas valiosas contribuições e ensinamentos.

Agradeço imensamente aos examinadores da banca, os Prof. Alfredo e Prof. Gustavo, por compartilharem seu tempo e conhecimento.

Agradeço à minha família, minha esposa e meus filhos, pelo carinho, confiança e compreensão.

E, finalmente, agradeço à minha orientadora, Profa. Candida, pela sua dedicação, empenho e incansável disposição em ajudar.

JUNQUEIRA, L.H.F. Abordagens bio-inspiradas aplicadas ao estudo da cognição: um encontro entre Biologia, Psicologia e Filosofia.

RESUMO

É antiga a preocupação do ser humano com as questões relacionadas ao conhecimento, incluindo as discussões sobre a sua origem, seu aprendizado, sobre a nossa capacidade de utiliza-lo e sobre as características específicas da cognição humana. Essa preocupação remonta aos antigos filósofos gregos (2.500 A.C.), desenvolvendo-se posteriormente a partir da abordagem da Epistemologia, originária da Filosofia ocidental, e mais recentemente a partir do Funcionalismo, pertencente aos estudos em Filosofia da Mente e em Ciência Cognitiva. Essa última abordagem, em particular, contribuiu para o surgimento de programas de pesquisas que procuram entender o funcionamento da mente humana com a ajuda do computador. A Ciência Cognitiva possui fortes ligações com as pesquisas em Inteligência Artificial, e ambas vem se desenvolvendo desde a década de 1950. Mais recentemente, a partir da década de 1980, uma nova área de estudos surgiu, formada por pesquisas em Vida Artificial, que trabalha com a possibilidade de síntese de entidades vivas, por meios artificiais, e desde então vem chamando a atenção de pesquisadores interessados no estudo da cognição. Enquanto a Ciência Cognitiva tem ligações estreitas com a Filosofia e a Psicologia, a Vida Artificial tem uma forte inspiração na Biologia. Neste trabalho, procuramos investigar o encontro entre essas disciplinas, e seus programas de pesquisas, considerando as possibilidades de contribuição para o estudo da cognição humana, a partir de uma abordagem conjunta entre essas áreas.

Palavras-Chave: Vida, Cognição, Ciência Cognitiva, Inteligência Artificial, Redes Neurais Artificiais, Vida Artificial, Biologia Sintética, Psicologia Sintética, Etologia Sintética.

JUNQUEIRA, L.H.F. Bio-inspired approaches applied to the cognition study: a meeting among Biology, Psychology and Philosophy.

ABSTRACT

It is an old thing the human concern about knowledge related questions, including discussions on its origin, its learning, about our capacity to utilize it and about specific characteristics of the human cognition. That concern remounts to the ancient greek philosophers (2.500 B.C.), developing itself later from the approach of Epistemology, originated in the occidental Philosophy, and more recently from the Functionalism, that makes part of the studies of Philosophy of Mind and of Cognitive Science. This last approach, particularly, contributed to appearing of research programs that try to understand the function of the human mind with the help of the computer. The Cognitive Science has strong relation with research in Artificial Intelligence, and both areas have been growing since the 1950 decade. More recently, from the decade 1980 on, a new area of studies appeared, formed by researches in Artificial Life, that works with the possibility of synthesis of alive entities, by artificial means, and since then has been attracting the attention of researchers interested in the study of cognition. While Cognitive Science has strict relations with the Philosophy and the Psychology, Artificial Life has a strong inspiration in the Biology. In this work, we tried to investigate the meeting between these subjects, and their research programs, considering the possibilities of the contribution to the study of the human cognition, by means of an integrated approach of these areas.

Keywords: Life, Cognition, Cognitive Science, Artificial Intelligence, Artificial Neural Networks, Artificial Life, Synthetic Biology, Synthetic Psychology, Synthetic Ethology.

SUMÁRIO

INTRODUÇÃO	1
CAPÍTULO 1 – O DESENVOLVIMENTO DA CIÊNCIA COGNITIVA	7
APRESENTAÇÃO.....	8
1.1 ORIGEM DA CIÊNCIA COGNITIVA	9
1.2 ABORDAGENS DA CIÊNCIA COGNITIVA	13
1.2.1 Cognitivismo.....	13
1.2.2 Cognição Dinâmica.....	17
1.3 CONSIDERAÇÕES SOBRE O CAPÍTULO	22
CAPÍTULO 2 – INTELIGÊNCIA ARTIFICIAL: A METÁFORA DA MENTE	23
APRESENTAÇÃO.....	24
2.1 ORIGEM DA INTELIGÊNCIA ARTIFICIAL	25
2.2 ALGUMAS CARACTERÍSTICAS DA IA.....	30
2.2.1 Algoritmo.....	31
2.2.2 Realimentação.....	33
2.3 CONSIDERAÇÕES SOBRE O CAPÍTULO	37
CAPÍTULO 3 – CONEXIONISMO: A METÁFORA DO CÉREBRO.....	41
APRESENTAÇÃO.....	42
3.1 ORIGEM DO CONEXIONISMO	43
3.2 ALGUMAS CARACTERÍSTICAS DAS RNA	47
3.2.1 Treinamento e Aprendizagem.....	47
3.2.2 Distribuição do Conhecimento	49
3.3 CONSIDERAÇÕES SOBRE O CAPÍTULO	50
CAPÍTULO 4 – ABORDAGENS BIO-INSPIRADAS: A METÁFORA DA VIDA.....	52
APRESENTAÇÃO.....	53
4.1 VIDA ARTIFICIAL.....	54
4.2 INTELIGÊNCIA ARTIFICIAL BASEADA EM COMPORTAMENTOS.....	61
4.3 ROBÓTICA EVOLUCIONÁRIA.....	65
4.4 MUNDOS VIRTUAIS E EMERGÊNCIA DA COGNIÇÃO	68
4.5 CONSIDERAÇÕES SOBRE O CAPÍTULO	72
CONSIDERAÇÕES FINAIS	74
REFERÊNCIAS.....	90

INTRODUÇÃO

Questões relacionadas ao conhecimento, sobre o que sabemos e como sabemos o que sabemos, são preocupações antigas da humanidade. De acordo com Chisholm (1974), essas questões são debatidas desde os antigos filósofos gregos. Chaui (2000) relata que Platão (428-347 A.C.) distinguia quatro formas de conhecimento, que são a crença, a opinião, o raciocínio e a intuição intelectual, considerando os dois primeiros como conhecimento sensível (ilusório ou das aparências) e os dois últimos como conhecimento intelectual (racional ou das idéias). Para Platão, somente através do conhecimento intelectual seria possível alcançar a verdade. Ainda de acordo com Chaui (2000), Aristóteles (384-322 A.C.) distinguia um outro conjunto de formas de conhecimento, composto pela sensação, percepção, linguagem, imaginação, memória, raciocínio e intuição. Diferentemente de Platão, Aristóteles acreditava que o nosso conhecimento vai sendo formado e enriquecido através do acúmulo das informações e experiências fornecidas por cada uma das formas, continuamente, ao invés de uma divisão entre o conhecimento sensível e o conhecimento intelectual.

Conforme Chisholm (1974, p. 11) “A reflexão sobre a natureza do nosso conhecimento dá origem a uma série de desconcertantes problemas filosóficos, que constituem o assunto da teoria do conhecimento, ou Epistemologia”. As abordagens iniciais sobre o conhecimento propostas pelos gregos influenciaram o pensamento dos filósofos durante os séculos seguintes, e a teoria do conhecimento tornou-se uma das áreas de estudos da Filosofia, com seu início delineando-se a partir da Idade Média. No século XVII, uma questão que se coloca é dicotomia do conhecimento, formada pelo objeto que se quer conhecer e o sujeito cognoscente (aquele que conhece). Se o pensamento que um sujeito tem sobre um objeto está de acordo com a realidade desse objeto, então se dá o conhecimento. Porém, quais são os critérios para se determinar se o pensamento está de acordo com a realidade do objeto?

Assim, “Um dos problemas que a teoria do conhecimento terá que propor e solucionar é aquele de saber quais são os critérios, as maneiras, os métodos de que se pode valer o homem para ver se um conhecimento é ou não verdadeiro” (Morente, 1966, p. 146).

Gardner (1995) considera que as propostas de solução para esse problema dá origem à duas abordagens divergentes: o Racionalismo e o Empirismo. O filósofo francês René Descartes (1596-1650), utilizando-se do método da dúvida, através do qual tudo é questionado, dá origem ao racionalismo cartesiano, para o qual a mente e o corpo são coisas distintas. O corpo, que tem como atividades a locomoção, a alimentação, a respiração, etc., é uma realidade física, sujeita às leis naturais. Já a mente, que tem como atividades a memória, o raciocínio, o conhecimento, etc., não ocupa lugar físico no espaço, objeto apenas da reflexão filosófica. Na abordagem empirista, o inglês John Locke (1632-1704) enfatiza o papel da experiência sensível no processo do conhecimento, para o qual o conhecimento só começa após a experiência produzida pelos sentidos. Assim como Aristóteles, Locke distingue diversas formas de conhecimento, partindo das sensações até chegar aos pensamentos. Para Chaui (2000), Locke é quem de fato inicia a teoria do conhecimento, ao se propor a analisar cada uma das formas de conhecimento que possuímos, a origem das idéias e as capacidades do sujeito cognoscente. Observando a semelhança com as abordagens dos filósofos gregos, Chaui compara de um lado Platão e Descartes, que discriminam o conhecimento sensível do conhecimento intelectual, e de outro Aristóteles e Locke, que consideram o conhecimento como resultado do acúmulo contínuo das informações obtidas de cada uma das formas de conhecimento, das sensações às idéias.

Para Gardner (1995), é Immanuel Kant (1724-1804) que tenta conciliar o empirismo com o racionalismo, ao considerar que o conhecimento é constituído de matéria e forma, sendo a matéria constituída do objeto que se quer conhecer e a forma pelo compreensão que temos do objeto, a partir de determinadas categorias de pensamento (dimensão, quantidade,

etc.). Kant criou ainda um nível de análise, chamado de esquema, que se interpõe entre a informação sensível provida pela objeto e as categorias de pensamento. O esquema funciona como uma representação mediadora entre o sensível e o intelectual. Gardner (1995) considera tal abordagem como equivalente à representação mental, utilizada atualmente em alguns programas de pesquisa, sobre a qual trataremos posteriormente.

A partir do século XVIII, na medida em que as conquistas científicas aliadas ao desenvolvimento da técnica vão se transformando em resultados práticos aplicados em benefício da sociedade, cresce a crença no cientificismo, concepção segundo a qual a ciência é o único conhecimento possível, e que os métodos científicos devem ser estendidos aos demais campos de investigação humana. E é nesse contexto, conforme mencionam Aranha e Martins (1986), que surge o Positivismo, que tem em Augusto Comte (1798-1857) seu principal representante. Comte considera que somente são reais os conhecimentos que repousam sobre fatos observados e defende o uso conjunto do raciocínio e da observação para atingir o estado positivo, no qual o termo positivo tem como significado o que é real, o que é dado pela experiência. O positivismo exerceu grande influência no pensamento da época, inclusive no estabelecimento da sociologia como uma ciência social, procurando submetê-la aos mesmos métodos utilizados nas ciências naturais. Ainda conforme Aranha e Martins (1986, p. 161) “A física newtoniana era considerada a imagem absolutamente verdadeira do mundo, tendo como pressupostos o mecanicismo e o determinismo”. Para explicar o mecanicismo, as autoras escrevem que “A ciência moderna compara a natureza humana e o próprio homem a uma máquina, um conjunto de mecanismos cujas leis precisam ser descobertas. As explicações passam a ser baseadas em um esquema mecânico, cujo modelo preferido é o relógio” (Aranha e Martins, 1986, p. 147). Já o determinismo é a consequência da posição positivista segundo a qual existem leis invariáveis que regem a relação entre os fenômenos, que uma vez descobertas podem ser utilizadas na previsão da ocorrência desses

fenômenos. Porém, com o advento de teorias que não se enquadravam no determinismo da física newtoniana, tais como a mecânica quântica e a teoria da relatividade, o entusiasmo positivista sofre um duro golpe, dando origem a outros movimentos, dentre os quais o neopositivismo e o empirismo lógico.

Para Gardner (1995), a Ciência Cognitiva nasce no século XX como sucessora dessas linhas de pesquisa e pensamentos, na tarefa de investigar os aspectos da cognição humana, e tem entre as questões objetos de sua atenção a preocupação em saber quais são as coisas passíveis de serem conhecidas, quais são as características daquele que possui tal capacidade de conhecer, quais são as fontes do conhecimento e quais são as dificuldades de sua obtenção, manutenção e processamento. Muitos dos programas de pesquisa em Ciência Cognitiva utilizam o computador como ferramenta para simular processos mentais, com o objetivo de entender a mecânica do pensamento e do processamento do conhecimento.

A idéia da utilização de máquinas na reprodução mecânica de operações mentais é uma antiga pretensão humana. Broens (1998) ao discorrer sobre a máquina aritmética criada pelo matemático e filósofo francês Blaise Pascal (1623-1662) comenta “Disto segue-se que, se é possível a reprodução mecânica de movimentos mentais, como Pascal demonstrara, para ele a mente não pode ser considerada enquanto puro espírito e o pensamento deve estar de algum modo vinculado ao mecanismo da natureza” (Broens, 1998, p. 189). O matemático inglês Alan Turing (1912-1954), num artigo publicado em 1950¹, vai mais longe descrevendo o que ele chama de “Jogo da Imitação”, no qual pretende mostrar que uma máquina seria capaz de simular a mente humana, ao se fazer passar por uma pessoa durante uma sessão de perguntas e respostas, realizada por um observador externo. Tal observador, travando conversa através de um terminal, ora com uma máquina, ora com uma pessoa, deveria

¹ Ver Turing (1990).

identificar a distinção entre ambos. O conjunto de procedimentos descritos passou a ser conhecido como “Teste de Turing”, que ainda hoje é considerado uma referência no desenvolvimento de programas de Inteligência Artificial e de Sistemas Especialistas. Gardner et al. (1998), por exemplo, descrevem a realização de um concurso baseado no Teste de Turing em 1991, nos Estados Unidos, no qual se confirmou a previsão de Turing de que um interlocutor mediano não teria mais que 70% de chance de identificar corretamente se travava conversa com uma máquina ou com outro humano, após cinco minutos de questionamentos. Turing lança também o pressuposto de que o processamento mental pode ser abstraído do meio em que ocorre (cérebro), dando origem à idéia de que a mente humana pode ser reproduzida artificialmente. As idéias de Turing contribuem para o surgimento do Cognitivismo, no qual os processos mentais são o objeto primário de estudo. Nessa abordagem, conforme Varela (2000), o conhecimento é visto como representações mentais internas e a inteligência como o resultado de processamento simbólico. Ou seja, a mente é computacional. Esses princípios irão nortear as três primeiras décadas de trabalhos na área de Inteligência Artificial. Ao longo do tempo, surgem diversas críticas sobre a abordagem do cognitivismo, dentre elas a de que a mente não pode existir sem um corpo. São então propostas abordagens alternativas, como a da Cognição Incorporada (Varela, 1991) e a da Cognição Incorporada e Situada (Haselager, 2004), ampliando as alternativas de pesquisas em Ciência Cognitiva.

Um outro campo de pesquisas, denominado de “Vida Artificial”, surgiu no final da década de 1980, disseminado por Langton (1989) através da divulgação de trabalhos realizados na área, com o pressuposto de que seria possível criar artificialmente entidades que possuiriam o atributo de estarem “vivas”, ou de apresentarem as características equivalentes às aquelas apresentadas por organismos biológicos vivos.

Diferentemente da Ciência Cognitiva, interessada principalmente na investigação da formação dos processos cognitivos superiores, tais como memória, linguagem e raciocínio, as pesquisas iniciais que deram origem a essa nova área foram motivadas pela simulação de processos biológicos em máquinas e na suposição de que a abordagem de Vida Artificial poderia ser uma ferramenta útil ao entendimento dos organismos biológicos, através da reprodução de suas características artificialmente. Mas, se de acordo com a abordagem que considera a cognição como sendo tanto incorporada como situada, na qual temos o corpo, a mente e o meio conjuntamente responsáveis pela emergência da cognição, questionamos se não seriam essas características aquelas pertencentes aos organismos biológicos? Não poderia a cognição ser considerada como uma característica emergente de um organismo vivo? Não seria possível, e até natural, considerarmos uma abordagem conjunta entre Ciência Cognitiva e Vida Artificial no estudo da cognição?

Assim, neste trabalho, nosso objetivo é o de investigar as possíveis contribuições provenientes da adoção da abordagem de Vida Artificial em estudos de Ciência Cognitiva e realizar reflexões sobre possíveis problemas na interface entre essas áreas. Para tanto, dividimos o trabalho em quatro capítulos. A Ciência Cognitiva será abordada no primeiro capítulo, e a Inteligência Artificial e o Conexionismo, que são os dois programas mais tradicionalmente utilizados em Ciência Cognitiva, no segundo e no terceiro capítulo, respectivamente. Cada um desses capítulos está dividido em três seções, através das quais apresentaremos a origem histórica, as principais características e algumas considerações sobre as pressuposições assumidas e as limitações dos resultados obtidos em cada uma das áreas. No quarto capítulo, trataremos da Vida Artificial e de outras abordagens bio-inspiradas. Em nossas considerações finais, verificaremos a possibilidade de estabelecer a conexão entre as abordagens de Ciência Cognitiva e de Vida Artificial, procurando identificar eventuais questionamentos oriundos da adoção dessas abordagens em conjunto.

CAPÍTULO 1

O DESENVOLVIMENTO DA CIÊNCIA COGNITIVA

A Ciência Cognitiva é uma das áreas de pesquisa mais dinâmicas no mundo hoje. Vale-se dos recursos conjuntos da ciência da computação, inteligência artificial, lingüística teórica, psicologia cognitiva, do desenvolvimento e da evolução e sociobiologia. Uma nova disciplina com uma antiga história, a ciência cognitiva está encontrando novas maneiras de atacar velhos problemas, principalmente ao empregar técnicas científicas para explorar questões sobre a natureza das mentes como tipos especiais de sistemas de processamento de dados, informação e conhecimento.

James H. Fetzer, 1999.

APRESENTAÇÃO

O objetivo deste capítulo inicial é o de apresentar uma introdução aos conceitos de Ciência Cognitiva e de seus principais paradigmas e programas de pesquisa, introduzindo as áreas de Inteligência Artificial e o Conexionismo. Na seqüência, abordaremos os programas de pesquisa reunidos sob a designação de Ciência Cognitiva Dinâmica, que utiliza conceitos como Auto-organização, Emergência e Cognição Incorporada e Situada.

Para cumprirmos este objetivo dividimos este capítulo em três seções. Na primeira, apresentaremos a origem e desenvolvimento da Ciência Cognitiva e, na segunda seção, introduziremos os principais programas de pesquisas em Ciência Cognitiva. Finalizaremos este capítulo com a introdução da idéia de Vida Artificial e a consideração da proposta de nossa investigação sobre a possibilidade de sua adoção em programas de pesquisas em Ciência Cognitiva.

1.1 ORIGEM DA CIÊNCIA COGNITIVA

Nas últimas décadas diversas pesquisas e trabalhos têm sido publicados, a partir de uma abordagem que ficou conhecida como Ciência Cognitiva, com o objetivo de investigar e tratar dos processos de aquisição do conhecimento, da relação mente-corpo, das bases neurofisiológicas da inteligência, da memória, da linguagem, etc., fundamentados, principalmente, na tentativa de simulação computacional dos processos cognitivos. Tais estudos têm envolvido disciplinas das mais diversas áreas de conhecimento humano, dentre elas a Filosofia, a Psicologia, a Lingüística, a Lógica, a Neurociência, a Semiótica, as Ciências da Computação e a Antropologia. Para Soares (1993), considera-se Ciências Cognitivas (no plural) como o conjunto de disciplinas (Ciências) que se uniram para investigar o conhecimento, sem no entanto perder suas características próprias ou suas existências individuais. Nessa abordagem, o computador não é necessariamente considerado como modelo da mente, mas pode ser utilizado como ferramenta de auxílio em seu estudo. Para Vignaux (1995), as disciplinas diretamente implicadas nas ciências cognitivas são a Inteligência Artificial, a Filosofia, a Lingüística, a Psicologia e a Neurociência. As duas últimas, em particular, recebem o adjetivo de “cognitiva” quando aplicadas no âmbito de pesquisas em ciências cognitivas, constituindo as disciplinas de Psicologia Cognitiva e de Neurociência Cognitiva. De acordo com Eysenck e Keane (1994), muitas das pesquisas em Psicologia Cognitiva baseiam-se na abordagem de processamento de informação, fornecida pela analogia com sistemas de computação, para desenvolver estudos sobre cognição humana, dentre os quais se encontram as investigações sobre linguagem, percepção, aprendizagem, memória, formação de conceitos e resolução de problemas. Já a Neurociência Cognitiva, conforme definida por Pereira Jr. (1995), pode ser considerada como uma área interdisciplinar, que procura relacionar fenômenos neurológicos e cognitivos, oriunda do

encontro entre as vertentes compostas, de um lado, pelos neurocientistas, que reuniram uma grande quantidade de informações sobre o funcionamento do cérebro, porém sem ter como principal preocupação a explicação de suas funções cognitivas, e de outro, composta por filósofos, psicólogos e pesquisadores em inteligência artificial, que procuraram estudar o conhecimento humano e a experiência subjetiva, compostos pela percepção, reação emocional, pensamento, uso da linguagem, etc., sem o conhecimento do sistema neuronal.

Para o autor,

Em termos práticos, o neurocientista cognitivo seria aquele que coloca em primeiro plano a função de processamento de informação dos neurônios, explicando tal função por meio de fenômenos bioquímicos e elétricos, cientificamente mensuráveis, que ocorrem na rede neuronal. Supõe-se que nesta rede (e não em neurônios isolados) se formem *representações* de diversas naturezas (de objetos e/ou processos do mundo exterior, de estados do próprio sistema, de estados de coisas almejados pelo sistema, ou mesmo representações de representações), que caracterizam a ocorrência de processos cognitivos, e permitem explicar os comportamentos externamente observáveis. (PEREIRA JR., 1995, p. 105-106)¹

Para a observação do cérebro *in vivo*, os cientistas utilizam equipamentos que são classificados em dois grupos, conforme Pereira Jr. No primeiro, estão aqueles que fornecem informações sobre a localização espacial das atividades neuronais, constituindo exemplos desses equipamentos o MRI (produção de imagens por ressonância magnética) e o PET (tomografia por emissão de pósitrons). No segundo grupo estão aqueles que fornecem informações sobre os padrões temporais, como os equipamentos de eletroencefalografia (EEG).

A Ciência Cognitiva (singular), propriamente dita, surge como tentativa de estabelecer um programa unificador para o estudo da mente (Eysenck e Keane, 1994), e é geralmente reconhecida a partir de 1956 (Gardner, 1995), ano em que aconteceram dois importantes eventos para a área, nos Estados Unidos. O primeiro, foi o encontro de verão de pesquisadores

¹ Grifo do autor.

das áreas de matemática e lógica do Dartmouth College, do qual surgiriam diversos expoentes de um novo campo de pesquisas, que se tornaria conhecido por Inteligência Artificial, como Marvin Minsky, John MacCarty, Herbert Simon e Allen Newell. O segundo, foi o Simpósio sobre Teoria da Informação do MIT – Massachusetts Institute of Technology, realizado em setembro, no qual estiveram presentes outros importantes nomes para a área, como o lingüista Noam Chomsky e o psicólogo George Miller. Gardner apresenta, então, sua própria definição para a Ciência Cognitiva:

Defino a ciência cognitiva como um esforço contemporâneo, com fundamentação empírica, para responder questões epistemológicas de longa data – principalmente aquelas relativas à natureza do conhecimento, seus componentes, suas origens, seu desenvolvimento e seu emprego. Embora o termo ciência cognitiva seja às vezes ampliado, passando a incluir todas as formas de conhecimento – tanto animado como inanimado, tanto humano como não humano – aplico o termo sobretudo a esforços para explicar o conhecimento humano. (GARDNER, 1995, p. 19-20)

Para o autor, a Ciência Cognitiva parece ter surgido também como oposição ao Behaviorismo, o qual tinha como objeto de estudo apenas o comportamento observável, rejeitando o introspeccionismo e evitando noções do tipo mente, pensamentos, desejos, intenções, ou aquelas relacionadas à representação mental, tais como símbolos, esquemas ou idéias. Gardner defende essa posição citando o fato de que o psicólogo Karl Lashley, durante o Congresso sobre Mecanismos Cerebrais do Comportamento, realizado em setembro de 1948 no California Institute of Technology, identificou os componentes necessários de uma Ciência Cognitiva, ao mesmo tempo que rejeitava as limitações impostas pelo Behaviorismo ao estudo da mente.

Muitos dos trabalhos iniciais em Ciência Cognitiva foram constituídos de tentativas de entender os processos humanos de cognição utilizando-se de computadores para simular as representações de estados mentais. Para Fetzer (2001), no entanto, o âmbito de atuação da Ciência Cognitiva está no estudo da cognição em seres humanos, em animais e em máquinas

(caso isso seja possível) e, embora úteis, Fetzer não considera os computadores como essenciais à existência da Ciência Cognitiva. Ou seja, uma Ciência Cognitiva pode existir independentemente da utilização de computadores. O autor também considera que a Inteligência Artificial é um ramo da Ciência da Computação que investiga a possibilidade de simulação das capacidades de seres humanos em máquinas, ressaltando que poderia existir uma Ciência Cognitiva sem Inteligência Artificial, mas que não poderia existir Inteligência Artificial sem Ciência Cognitiva.

Já Dupuy (1996), considera a Cibernética (ou Movimento Cibernético) como precursora da Ciência Cognitiva. Para o autor, a história começa em 1943, com a publicação de dois artigos: o primeiro o de Rosenblueth, Wiener e Bigelow, intitulado *Behavior, Purpose and Teleology*, e o segundo o de McCulloch e Pitts, intitulado *A logical calculus of the ideas immanent in nervous activity*. A Cibernética, definida por Wiener (1970) como a ciência do controle e comunicação no animal e na máquina, tem suas bases lançadas no primeiro artigo citado por Dupuy, enquanto que o artigo de McCulloch e Pitts estabeleceu as bases do Conexionismo. Uma outra interpretação é fornecida por Luger (1994), de acordo com a qual a Ciência Cognitiva seria usualmente definida como o estudo científico tanto da mente como da inteligência, que se trata, obviamente, de uma definição bastante abrangente.

Dessa forma, observamos que não existe uma definição única para Ciência Cognitiva, a qual pode variar conforme o autor ou a abordagem que se adote. Para facilitar a classificação das pesquisas realizadas na área, pode ser de utilidade investigar os objetos de estudo e os programas adotados em Ciência Cognitiva, assumindo uma divisão entre suas abordagens, conforme sugerimos na próxima seção, que, de qualquer forma, não é a única possível.

1.2 ABORDAGENS DA CIÊNCIA COGNITIVA

Com o objetivo de facilitar a classificação das pesquisas realizadas na área, consideraremos a Ciência Cognitiva dividida principalmente em duas abordagens: a do Cognitivismo e a dos Sistemas Dinâmicos. Na primeira, encontram-se os programas de pesquisas em Inteligência Artificial. Na segunda abordagem, encontram-se muitas das pesquisas mais recentes, como as da Cognição Incorporada e Situada. Procuramos, a seguir, detalhar tais abordagens.

1.2.1 COGNITIVISMO

De acordo com Lakoff e Johnson (1999), o termo “cognitivo” em Ciência Cognitiva seria utilizado para qualquer tipo de operação ou estrutura mental que possa ser estudada em termos precisos. Para Anderson (2003), o cognitivismo é a hipótese de que a função central da mente, a de pensar, poderia ser estabelecida em termos de manipulação de símbolos, de acordo com regras explícitas. O cognitivismo possuiria três elementos-chave: representação, formalismo e transformação baseada em regras. A idéia principal é a de que cognição envolve representação. Seria a forma do símbolo, e não seu significado, o elemento para as transformações baseadas em regras. Essa forma de abordagem norteou uma grande parte dos programas de pesquisa em Inteligência Artificial, conforme verificaremos na seqüência de nossa discussão. Varela (2000, p. 31) nos lembra que “a hipótese cognitivista pretende que a única solução para explicar a inteligência e a intencionalidade reside na justificação de que a

cognição consiste em agir na base de representações que têm *uma realidade física sob forma de código simbólico num cérebro ou numa máquina*”².

A Inteligência Artificial (IA) trabalha com a hipótese de que a mente humana pode ser simulada, e até mesmo reproduzida, em um computador. A vertente que tenta simular a mente humana, através de modelos computacionais, ficou conhecida como Inteligência Artificial Fraca. Já a vertente que tem por objetivo a reprodução da própria mente humana no computador ficou conhecida como Inteligência Artificial Forte. Essa classificação é devida principalmente à Searle (1990). Na IA encontramos a “metáfora da mente”, que tem como base as representações mentais, cuja tentativa de simulação é implementada através de um conjunto de regras pré-estabelecidas. A IA estrutura seus programas de pesquisa a partir da codificação de regras abstratas e de algoritmos computacionais, decompondo os componentes das representações mentais e simulando-os através de programas de computador. Considerando-se a inteligência como a habilidade de resolver problemas, a simulação ou reprodução da mente constitui-se então na formulação de problemas e na tentativa de resolvê-los através de modelos computacionais. Assim, conforme Fetzer (2001), o computador é visto com uma máquina que realiza um conjunto de funções, razão pela qual a concepção computacional da mente, ficou conhecida como “Funcionalismo”, concepção essa atribuída ao filósofo Hilary Putnam (Putnam 1961, 1975). Ainda de acordo com Fetzer (2001), essa concepção que considera as mentes como sistemas que executam funções por meio de programas é melhor definida como “Funcionalismo de Estado de Máquina”, ou conforme Clark (2001), “Funcionalismo de Máquina”.

Para a filósofa Margaret Boden, caso se adote a interpretação da IA como núcleo intelectual da Ciência Cognitiva, as definições de Inteligência Artificial e de Ciência Cognitiva estariam intimamente relacionadas. Para a autora,

² Grifo do autor.

Inteligência Artificial (IA) é algumas vezes definida como o estudo de como construir e/ou programar computadores para permiti-los fazer os tipos de coisas que mentes podem fazer. [...] Mas muitos outros preferem uma definição mais controversa, vendo a IA como a ciência da inteligência em geral – ou, mais precisamente, como o núcleo intelectual da ciência cognitiva. Como tal, seu objetivo é prover uma teoria sistemática que explique (e talvez nos habilite replicar) não só as categorias gerais da intencionalidade e a diversa psicologia das criaturas terrestres, mas o conjunto inteiro das possíveis mentes.³ (BODEN, 1990, p. 1)

Esse tipo de interpretação, da IA como base da Ciência Cognitiva, pode ser verificada a partir da definição de Ciência Cognitiva dada por Dawson:

Ciência cognitiva é um estudo intensamente interdisciplinar da cognição, percepção e ação. Ela está baseada na hipótese de que cognição é processamento de informação, onde processamento de informação é geralmente interpretado como manipulação baseada em regras de estruturas de dados que estão armazenadas numa memória. Como resultado dessa hipótese, um objetivo básico da ciência cognitiva é identificar a arquitetura funcional da cognição – o conjunto primitivo de regras e representações que mediam o pensamento.⁴ (DAWSON, 2002, p. 13)

Observa-se, nessa definição, que o estudo da cognição pela Ciência Cognitiva seria intrinsecamente dependente dos atributos da Inteligência Artificial (processamento de informação, manipulação de símbolos, representações e conjuntos de regras). Muitas críticas foram feitas a esse tipo de abordagem, principalmente àquelas pretensões da Inteligência Artificial Forte de instanciar a mente por meio de um programa de computador.

Uma abordagem alternativa, estabelecida em paralelo aos programas de Inteligência Artificial, é a do Conexionismo, no qual leva-se em consideração o meio físico no qual as idéias ocorrem, ou seja, o cérebro. São utilizadas Redes Neurais Artificiais (RNA), que se baseiam parcialmente na simulação das redes de neurônios existentes no cérebro, como tentativa de entender os processos mentais. As representações passam a ser encaradas como padrões de atividades neuronais, e o estudo da mente passa a ser investigado através das

³ Tradução nossa.

⁴ Tradução nossa.

conexões entre os neurônios. Assim, o Conexionismo representa a “metáfora do cérebro” para o estudo das capacidades da mente. A RNA também é implementada por meio de computadores, porém não utiliza diretamente o conceito de computação simbólica da IA, mas sim o conceito de modificação dos padrões de conexão entre as unidades da rede. Essas conexões possuem pesos e, dependendo da força desses pesos, elas podem se enfraquecer ou podem se fortalecer, ou ainda novas conexões podem ser estabelecidas, ou conexões anteriores podem ser desfeitas. O conjunto de conexões assim obtido representa o conhecimento contido na rede e distribuído através das conexões.

Conexionismo e a IA partilham o conceito de representações, mas uma das principais diferenças entre essas abordagens é a de que uma RNA pode ser treinada para resolver um determinado problema, eliminando a necessidade de programação prévia de todas as etapas do processo, tal como ocorre nos programas de computador. Esse treinamento baseia-se em regras de aprendizado, que estabelecem as modificações nos padrões de conexão entre os elementos da rede. Como o conhecimento obtido, ou solução de um problema, se encontra no padrão de conexões estabelecido no final do processo de treinamento, e não nas unidades individuais que compõem a rede, Smolensky (1988) considera os sistemas conexionistas como subsimbólicos, situando-os num nível intermediário, acima da implementação de baixo nível das unidades individuais de neurônios, tal como ocorrem no cérebro, e abaixo das abstrações de alto nível do processamento simbólico, tal como utilizado na IA. Embora os modelos conexionistas não possam ser classificados como pertencentes ao Cognitivismo, por não serem diretamente simbólicos, para alguns autores, como Boden (1995), é comum classificar o Conexionismo como parte da IA, distinguindo-se as abordagens através das designações de “IA Simbólica” e “IA Conexionista”. Abordaremos em mais detalhes as áreas da Inteligência Artificial e do Conexionismo nos capítulos 2 e 3, respectivamente.

1.2.2 COGNIÇÃO DINÂMICA

Na seção anterior, destacamos as vertentes da Ciência Cognitiva que procuram investigar e entender os processos de cognição utilizando-se de programas e de simulações computacionais a partir da idéia de representações dos estados mentais, razão pela qual são também conhecidas como abordagens representacionistas. Um segundo paradigma surge como alternativa, no qual a noção de representação deixa de assumir papel principal, permeando os programas de pesquisas reunidos sob a designação de Ciência Cognitiva Dinâmica que, de acordo com Gonzalez (1996, p. 275) ao comentar sobre o estudo das propriedades de Auto-organização em modelos conexionistas alternativos (conhecidos como neoconexionistas), preocupa-se em “explicar a *dinâmica* dos seres que se *auto-organizam* em função de seus estados internos e dos eventos que ocorrem no meio ambiente”⁵. Gonzalez (2003) explica que o estudo da mente pela Ciência Cognitiva Dinâmica enfatiza a interação com o meio, através dos conceitos de que a mente está espalhada no meio (situada), de que meios distintos resultam em sistemas distintos e de que sujeitos se adaptam ao meio no qual estão situados. As informações são, portanto, captadas, em função dessa interação com o meio, e a preocupação é de como selecionar a informação relevante (captação relevante), que possibilite a melhor adaptação do sujeito ao meio. Verificamos, portanto, que na abordagem da Ciência Cognitiva Dinâmica, a Auto-organização e a interação dinâmica com o meio assumem papel de destaque.

Debrun (1996, p. 4) oferece uma definição para a Auto-organização: “Há Auto-organização cada vez que o advento ou a reestruturação de uma forma, ao longo de um processo, se deve principalmente ao próprio processo – a características nele intrínsecas -, e só em grau menor às suas condições de partida, ao intercâmbio com o ambiente ou à

⁵ Grifos da autora.

presença eventual de uma instância supervisora”. Há ainda a distinção entre Auto-organização Primária e Secundária, na qual a Auto-organização Primária não parte de uma forma (um ser ou um sistema) já constituída, mas que, ao contrário, há uma sedimentação da forma. Ela está presente naqueles processos em que elementos inicialmente independentes interagem de tal modo que eles se tornam interdependentes. Esse processo acontece sem a presença de um controle central ou de leis pré-estabelecidas. Nesse tipo de Auto-organização, encontram-se, por exemplo, os processos de formação dos padrões de conectividade nas Redes Neurais Artificiais sem treinamento supervisionado, nas quais tais padrões emergem da interação de diversas unidades de neurônios artificiais, que trocam sinais entre si. Esses neurônios, inicialmente, têm pouca ou nenhuma relação de dependência entre si, mas a interação entre uma parte deles, ao longo do tempo, produz um efeito coletivo na rede, afetando também o comportamento das demais unidades.

Já a Auto-organização Secundária, dando continuidade à explicação de Debrun (1996), se dá quando um organismo (ser ou sistema) consegue passar de determinado nível de complexidade para um nível superior, a partir de suas próprias operações exercidas sobre ele mesmo, porém partindo de um ser ou sistema já constituído. A Auto-organização Secundária, segundo Gonzalez (1996, p. 267), “pressupõe a existência de um sistema primariamente organizado, adicionando-lhes mecanismos de evolução”. Tais mecanismos poderiam incluir a habilidade para promover a mudança de hábitos, ou de padrões informacionais, por meio de processos de aprendizagem, envolvendo dinâmicas de competição e/ou de cooperação, sem a presença de programas pré-estabelecidos ou de controle central. De acordo com a autora, admitindo-se a possibilidade de aprendizagem não-supervisionada em Redes Neurais Artificiais, também seria possível admitir a Auto-organização secundária nessas redes, que estaria associada a habilidade de desenvolvimento e de aprimoração de seus estados representacionais.

Na abordagem da Ciência Cognitiva Dinâmica, a concepção da Cognição Incorporada e Situada (*Embodied Embedded Cognition*) considera a cognição como um processo decorrente de um organismo que possui um corpo (cognição incorporada), que se encontra fisicamente num ambiente (cognição situada). Willem Haselager, do Instituto para Cognição e Informação da Universidade de Nijmegen, Holanda, que desenvolve estudos utilizando essa abordagem (Haselager, 2001, 2003, 2004, 2005), trabalha com a idéia do surgimento da cognição a partir de um agente que interage no mundo, em que o corpo desse agente e suas características desempenham papel importante na compreensão de processos cognitivos, e sem o qual muitos dos processos de cognição estariam comprometidos ou seriam inviabilizados. Esses processos de cognição são ainda determinados pela atuação do agente nesse ambiente, uma vez que diferentes possibilidades de ação são oferecidas pelo meio no qual o agente está inserido, tornando assim relevante considerar que os processos cognitivos possam ser dependentes do meio, em função de qual ação seja realizada pelo agente. De acordo com a posição de Haselager (2003), não é relevante determinar o que é interno ou externo, mas sim compreender que o conjunto formado pelo cérebro, corpo e mundo constitui um sistema unificado, o qual apresenta como resultado uma atuação unificada. O comportamento inteligente é então considerado como uma propriedade emergente da interação dinâmica entre o cérebro, o corpo e o mundo à sua volta, através de processos auto-organizados.

A noção de emergência desempenha papel importante na abordagem da Cognição Incorporada e Situada, razão pela qual tentaremos apresentar uma interpretação inicial para esse conceito. Conforme mencionado anteriormente, é possível obter a emergência de características do sistema como um todo, a partir da Auto-organização entre seus elementos individuais, como no caso dos padrões de conectividade em determinados modelos de Redes Neurais Artificiais. Para efeito de nossos estudos, consideraremos a noção de Emergência,

segundo Clark (2001), como o surgimento de uma funcionalidade não-programada em um sistema, o qual ocorre a partir da interação entre seus componentes individuais, ou de suas partes. Tal funcionalidade não é considerada inesperada, porém ela não está definida internamente e nem são especificados os meios para alcançá-la. Esse tipo de Emergência está geralmente relacionada à efeitos coletivos e tem também particular destaque na literatura sobre Vida Artificial (Clark, 2001), que será objeto de nossa investigação posterior.

Considerando a concepção da Cognição Incorporada e Situada, poderíamos levantar as questões de como seria possível criar modelos para o estudo da cognição, a partir dessa abordagem, e de como seria possível obter a autonomia de agentes artificiais que interagem com o mundo. Haselager (2003) explica que a abordagem da Robótica Evolucionária parece ser uma alternativa de resposta a essas duas perguntas, na qual os robôs passariam a ser agentes autônomos situados no ambiente. Nessa idéia, tanto o “corpo” físico do robô (constituído de partes mecânicas), como o seu “cérebro” (geralmente constituído de uma implementação de RNA), deveriam se desenvolver de forma conjunta, assim como ocorre nos organismos biológicos. Dessa maneira, o “comportamento inteligente” poderia emergir a partir de um processo de Auto-organização dos robôs, em relação ao seu desenvolvimento interno (padrões de conexões de sua RNA), e de sua interação com o ambiente, por meio de sua estrutura física (corpo do robô).

Na verdade, as pesquisas em Robótica Evolucionária se baseiam em alguns dos conceitos e ferramentas já utilizados num novo campo de pesquisas, conhecido como Vida Artificial, cujos princípios foram estabelecidos num simpósio organizado por Langton (1989), e que tem em uma de suas vertentes a crença na possibilidade de criação de seres artificiais “vivos”. Tais seres possuiriam a propriedade de desenvolvimento autônomo, eliminando a necessidade de programação prévia de todas as etapas envolvidas no processo. Essa nova abordagem passou a inspirar-se diretamente na biologia, e uma de suas principais ferramentas

são os Algoritmos Genéticos, cuja idéia principal é a de reproduzir as etapas de evolução natural dos seres biológicos através de algoritmos computacionais. Como entre os problemas para a implementação física da abordagem da Robótica Evolucionária estão o alto custo dos componentes e partes do robô (mecânica fina, dispositivos sensório-motor, eletrônica embarcada, etc.), e as restrições tecnológicas para a utilização das idéias de Von Neumann (1966) de criação de autômatos auto-replicantes (máquinas capazes de reproduzir a si próprias), uma solução alternativa tem sido a criação de programas computacionais, que simulam os aspectos relacionados ao mundo e aos agentes, nos quais utilizando-se de algoritmos genéticos observa-se a evolução de populações iniciais de agentes interagindo num mundo virtual, em direção à um determinado objetivo desejado, que pode incluir a geração de agentes melhores adaptados ao meio existente, agentes capazes de atuar de forma cooperativa, etc. Dessa forma, seres artificiais virtuais (ou seja, não materializados) poderiam constituir uma nova classe de sujeitos para o estudo da cognição, caso a ocorrência dessa seja possível em tais seres e em tais mundos virtuais, constituindo assim uma “metáfora da vida”.

1.3 CONSIDERAÇÕES SOBRE O CAPÍTULO

Neste capítulo procuramos apresentar a conceituação de Ciência Cognitiva, através de seus paradigmas e de suas principais vertentes. Verificamos que não existe uma definição única para Ciência Cognitiva, e consideramos útil uma abordagem de investigação dos objetos de estudo, e dos programas de pesquisa da Ciência Cognitiva, apresentando a abordagem Cognitivista e a abordagem da Cognição Dinâmica.

Introduzimos as áreas de pesquisas conhecidas como Inteligência Artificial e Conexionismo, e reunimos alguns dos principais temas relacionados com a abordagem da Ciência Cognitiva Dinâmica, dentre eles a Auto-organização, a Emergência e a Cognição Incorporada e Situada.

A partir do paradigma dinâmico, introduzimos a idéia de Vida Artificial, e o questionamento sobre a possibilidade de sua adoção em programas de pesquisas em Ciência Cognitiva, que será o tema central de nossa investigação. Antes, porém, com o objetivo de verificarmos a relevância da utilização de um paradigma dinâmico, estaremos investigando de uma forma mais detalhada as duas abordagens mais tradicionais da Ciência Cognitiva, que são a Inteligência Artificial e o Conexionismo.

No próximo capítulo, iniciaremos esse detalhamento com a abordagem da Inteligência Artificial (IA), ponderando sobre algumas das pretensões desse programa de pesquisas, no qual considera-se a possibilidade do computador instanciar a mente humana, e, em particular, tentaremos investigar a acertividade da afirmação que considera os computadores capazes de aprender, dentro de uma perspectiva da cognição humana.

CAPÍTULO 2

INTELIGÊNCIA ARTIFICIAL: A METÁFORA DA MENTE

A idéia de que o computador digital é um processador simbólico geral e que, portanto, pode ser utilizado tanto para fazer inferências como efetuar cálculos, criou um excitante campo novo. O computador é tão versátil que pode ser utilizado para modelar um sistema holístico. De fato, uma nova geração de pesquisadores começou a utilizar o computador para simular tais sistemas, mas é muito cedo para dizer se esses primeiros passos em direção à discriminação e associação holística eventualmente conduzirão aos dispositivos que podem discernir similaridades entre situações complexas do mundo real.

S. Dreyfus e H. Dreyfus, 1986.

APRESENTAÇÃO

Neste capítulo, nosso objetivo é o de contextualizar o objeto de estudo, que trata sobre a capacidade de aprendizagem de máquinas, e especificamente sobre a capacidade de simulação de estados cognitivos humanos, a partir do paradigma da Inteligência Artificial.

Trataremos como hipótese central da investigação, com base nos trabalhos de Inteligência Artificial, sobre a afirmação de que computadores poderiam apresentar traços de capacidades cognitivas humanas, em particular a capacidade de aprendizado.

Para tanto, dividiremos o capítulo em três seções. Na primeira, introduziremos o conceito de computador, através da motivação da sua construção e do histórico do seu desenvolvimento e, em seguida, conceituaremos a Inteligência Artificial, verificando os trabalhos e as pesquisas que deram origem à área. Na segunda seção, passaremos a investigar os pressupostos metodológicos utilizados pela Inteligência Artificial, que sugerem estados cognitivos aos computadores, procurando avaliar a validade de tais afirmações. Na terceira seção, concluiremos o capítulo apresentando algumas implicações filosóficas da Inteligência Artificial, em particular aquelas postuladas pelo filósofo John Searle.

2.1 ORIGEM DA INTELIGÊNCIA ARTIFICIAL

A humanidade sempre se preocupou em desenvolver meios para facilitar a execução das tarefas essenciais para a sua sobrevivência. Primeiro, procurou facilitar a execução daquelas tarefas que exigiam força bruta, utilizando-se de animais de carga, e, depois, daquelas tarefas que exigiam ações repetitivas, por meio de mecanismos e de engrenagens. De acordo com Scornaienchi (1975), no período Neolítico (5000 a 2500 A.C.) o homem já praticava a agricultura e se utilizava de bois para puxarem o arado e preparar a terra. Por volta do ano 3000 A.C. as rodas d'água eram utilizadas no auxílio à irrigação do solo para o plantio, e em torno de 950 A.C aparecem os moinhos com engrenagens, utilizados para a moagem de grãos. No Egito, por volta de 500 A.C., surge o Ábaco, instrumento que funcionava como uma espécie de calculadora manual rudimentar. Para poder exercer controle sobre suas atividades produtivas e mercantilistas, tornou-se crescente para a humanidade a necessidade de contar e realizar operações numéricas.

Conforme Giarratano (1982), foi em 1692 que o filósofo e matemático francês Blaise Pascal inventou a primeira máquina de somar, conhecida como Pascalina, que realizava os cálculos através da manipulação de discos interligados. A subtração podia ser feita através da inversão dos discos utilizados para a adição e as operações de multiplicação e divisão podiam ser realizadas por repetição de adições ou subtrações, assim como se fazia no Ábaco. Em 1822, o matemático inglês Charles Babbage desenvolveu sua Máquina Diferencial: um dispositivo capaz de executar operações matemáticas mais sofisticadas. Também é de Babbage a idéia da Máquina Analítica. Sua máquina seria capaz de guardar dados num sistema de armazenamento interno, composto de um banco de 1000 registradores, com capacidade de armazenamento de números de 50 dígitos cada um, realizar diversas operações matemáticas, fazer comparações e tomar decisões de seqüência de cálculo, de acordo com o

resultado obtido. Nessa capacidade, a de tomar decisões, reside a principal diferença entre as calculadoras e o computador. A máquina receberia os dados e as instruções de cálculo através de cartões perfurados e forneceria o resultado gravado numa lâmina de cobre própria para ser utilizada para impressão. No projeto da Máquina Analítica de Babbage já estavam contidos os principais conceitos dos computadores modernos: entrada de informações, armazenamento interno de dados, utilização de um programa (conjunto de instruções) para realização dos cálculos e saída impressa dos resultados. Essas idéias lhe renderam o reconhecimento como o “Pai do Computador”. Porém, a tecnologia da época não permitiu que a Máquina Analítica fosse construída.

Foi somente com o advento de máquinas operadas eletricamente que as idéias de Babbage começaram a se concretizar. Giarratano (1982) informa que o primeiro computador eletromecânico, constituído de relés, foi idealizado pelo alemão Konrad Zuze em 1936. Os dados eram lidos a partir de uma fita perfurada, e os cálculos realizados a partir de um programa. Nesse mesmo ano, Turing (1936) publica um importante artigo, no qual apresenta o conceito que ficou conhecido como “Máquina de Turing”. A Máquina de Turing é um dispositivo hipotético que recebe dados em sua entrada, realiza uma determinada ação sobre os dados, em função de uma tabela de regras pré-estabelecidas, e apresenta em sua saída os dados resultantes da ação realizada. Dentro desse conceito, qualquer procedimento lógico que possa ser definido passo a passo poderia ser transformado numa Máquina de Turing, a partir de uma tabela de regras específicas para realizar aqueles passos. Esse conceito era coerente com os tipos de máquinas que estavam sendo construídas até então, que basicamente se dedicavam a uma tarefa específica. A inovação ficou por conta do conceito de “Máquina Universal de Turing”: um dispositivo capaz de realizar as ações de qualquer Máquina de Turing e que, para tanto, bastaria possuir a tabela de regras dessa última.

Turing havia estabelecido o importante conceito teórico de uma única máquina capaz de executar qualquer procedimento lógico programável, em substituição a todas as demais máquinas de computação específica. O computador atual é a aplicação prática do conceito de Máquina de Universal de Turing. Andrew Hodges, em seu livro sobre a vida e a obra de Turing, apresenta seu ponto de vista sobre o termo computador:

[...] a palavra “computador” mudou seu significado. Em 1936, e na verdade em 1946, ela designava uma pessoa computando, e uma máquina seria chamada um “computador automático”. Até os anos 60, havia uma distinção entre computadores digitais e analógicos; apenas a partir daí, na medida em que os computadores digitais passaram a predominar, a palavra passou a designar uma máquina como a que Turing tinha em mente. Mesmo agora, ela é às vezes aplicada a qualquer máquina de calcular. Ao mencionar o “computador”, tomo como sua característica distintiva a de que os programas e os dados são igualmente considerados e manipulados da mesma forma - o “programa modificável armazenado” – e essa é a característica sugerida por Turing ao falar da “máquina universal prática”, sua maneira de descrever a própria idéia. (HODGES, 2001, p. 33)¹

A constatação da vantagem da adoção do sistema binário para construção de máquinas computadoras é geralmente atribuída à John Von Neumann e foi aplicada no desenvolvimento do EDVAC², um dos primeiros computadores eletrônicos, construído em 1950. Uma outra contribuição significativa de Von Neumann foi a introdução do conceito de programa armazenado, onde as instruções para a máquina são armazenadas internamente ao computador. Mas o resultado mais marcante de seu trabalho está na sua proposta para a Arquitetura de Máquinas Digitais (Von Neumann, 1946), que possibilitou o desenvolvimento dos computadores digitais atuais. Tais máquinas são conhecidas como “Computadores de Arquitetura Von Neumann” ou “Máquinas de Von Neumann”. Em essência, essa arquitetura divide a máquina nos seguintes blocos funcionais: a Unidade de Entrada (UE), as Unidade de Controle (UC), Unidade Lógica e Aritmética (ULA) e Memória (M), que juntas formam a

¹ Destaques do autor.

² EDVAC – Electronic Discrete Variable Automatic Computer.

Unidade de Processamento Central (UCP), e a Unidade de Saída (US). Um diagrama que representa essa arquitetura, de acordo com Langdon (1974), pode ser visto na figura abaixo:

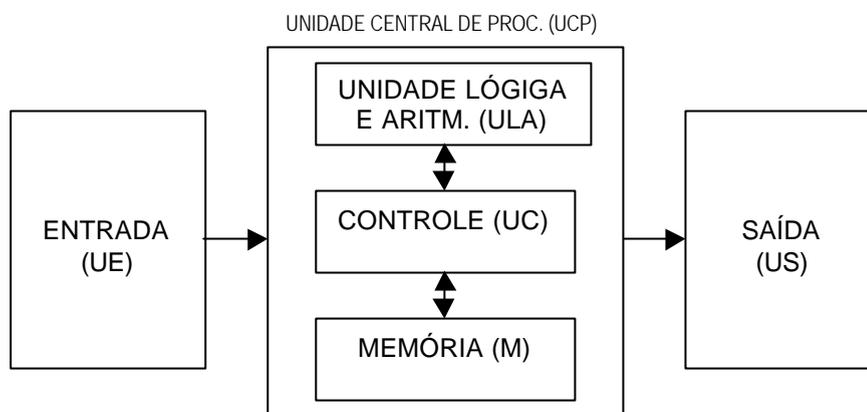


Figura 1: Computador de Arquitetura Von Neumann.

Essa arquitetura foi largamente adotada, permitindo a construção dos computadores eletrônicos digitais, conforme a designação com que tais máquinas ficaram conhecidas. A principal restrição desse modelo é a interposição da Unidade de Controle entre o processamento e a memória, conhecida como o “gargalo de Von Neumann”, que obriga um tratamento seqüencial mesmo quando trechos de um programa ou software poderiam ser executados em paralelo.

Von Neumann desenvolveu vários outros trabalhos e, dentre eles, uma comparação sistemática entre o cérebro humano e o computador (Von Neumann, 1958). Com o desenvolvimento crescente dos computadores nas décadas seguintes, quando eles chegaram a ser chamados de “cérebros eletrônicos”, parecia ser natural imaginar que eles poderiam “pensar” ou simular o modo de pensamento de seres humanos. E é nesse contexto que começa a surgir a área de pesquisa que viria a se tornar conhecida como Inteligência Artificial.

Gardner (1995) afirma que, de acordo com a maioria dos relatos verificados por ele, o termo Inteligência Artificial (IA) foi cunhado pelo matemático John McCarthy, na década de 1950, nos Estados Unidos. Em 1956, então professor do Dartmouth College, em New Hampshire, McCarthy organizou um encontro de verão com o objetivo de discutir a possibilidade de desenvolver programas de computador que se comportassem de forma inteligente. Em solicitação de bolsa à Fundação Rockefeller, o grupo de participantes do encontro declarava:

O estudo será conduzido com base na hipótese que todo aspecto de aprendizagem ou qualquer outra faceta da inteligência pode em princípio ser descrito de forma tão precisa que se pode fazer com que uma máquina o simule. (McCORDER apud GARDNER, 1995, p. 153)

Assim, surgia a pretensão de simular o pensamento humano através da máquina. A principal diferença de abordagem do uso de computadores que se fazia até então foi a mudança de paradigma de uma máquina capaz de realizar cálculos rapidamente para o de uma máquina capaz de processar símbolos. Entre alguns dos participantes do encontro estavam: Marvin Minsky discutindo sobre a utilização de computadores na prova de teoremas matemáticos, Herbert Simon e Allen Newell descrevendo um programa de computador para aplicações em problemas de lógica, Arthur Samuel tratando sobre um programa para jogar damas e Alex Bernstein apresentando seu trabalho sobre um programa capaz de jogar xadrez. Em conjunto com McCarthy, eles se tornariam alguns dos principais pesquisadores dos primeiros anos da IA.

2.2 ALGUMAS CARACTERÍSTICAS DA IA

A construção de máquinas capazes de desempenhar o papel de jogadores, particularmente em jogos de damas e de xadrez, parece ser uma ambição antiga da humanidade. Vignaux (1995) cita como um jogador mecânico de xadrez o autômato de Kempelen, de 1769, que fez muito sucesso mas que na verdade se tratava de uma fraude: um anão escondido fazia os movimentos no tabuleiro. Porém, em 1912 Torres e Quevedo construíram um autômato que realmente era capaz de jogar as finais do jogo e em 1945 Konrad Zuse consegue programar as regras do jogo de xadrez. Em geral, esses tipos de jogos envolvem as capacidades de aprendizagem e de raciocínio, capacidades essas associadas à inteligência humana. Assim, uma máquina capaz de jogar damas ou xadrez poderia ser considerada inteligente? Podem, então, as máquinas pensar?

Essa questão já havia sido colocada por Turing (1990), em seu artigo de 1950 intitulado *Computing, Machinery and Intelligence*, e, para uma parte dos pesquisadores envolvidos com trabalhos em Inteligência Artificial, a resposta para essa pergunta parece ser sim. Epstein (1973) cita como exemplo justamente os programas capazes de jogar dama e xadrez, colocando em destaque observações sobre o programa criado por Samuel para jogar damas:

A atividade não pode ser determinística no sentido prático. Não há um algoritmo conhecido que garanta a vitória ou o empate no jogo de damas. A exploração de todos os caminhos possíveis num jogo de damas envolveria cerca de 10^{40} escolhas, que à razão de três escolhas por milimicrosegundo levariam 10^{21} séculos a serem consideradas. [...]. O programa então, tal como os seres humanos, lança mão de processos heurísticos. Certos caminhos e parâmetros são valorizados, num processo em que o resultado dá uma contínua retroação para um aperfeiçoamento constante à base da experiência. Poderíamos dizer que a máquina apresenta traços do que chamamos de intuição e aprendizado nos seres humanos. (EPSTEIN, 1973, p. 49)

Tais ponderações colocam a máquina num patamar muito próximo ao do homem, particularmente na afirmação de que “Poderíamos dizer que a máquina apresenta traços do que chamamos de intuição e aprendizado nos seres humanos”. Investigaremos dois dos conceitos citados no trecho acima, algoritmo e retroação, que nos parecem ser pontos-chave em muitas das abordagens de Inteligência Artificial, para verificarmos quão próxima da verdade estaria aquela afirmação.

2.2.1 ALGORITMO

Um problema básico em qualquer jogo é o do processo de tomada de decisão quando existem várias alternativas possíveis. Cada alternativa pode levar a um novo conjunto de alternativas, e assim sucessivamente, num conjunto de combinações que poderia ser representado por uma árvore de possibilidades, conforme apresentado abaixo:

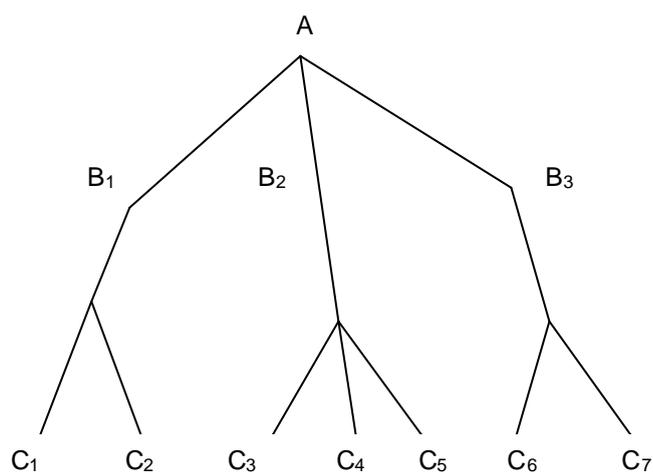


Figura 2: Uma árvore de possibilidades genérica

Caso seja possível verificar cada um dos ramos da árvore, pode-se então saber se existe uma solução ao problema e, se sim, determinar qual será seu algoritmo. Ocorre que mesmo em jogos aparentemente simples, como o jogo de damas, a árvore torna-se tão extensa que sua verificação total seria impraticável.

Vignaux (1995, p.25), comentando sobre os programas de jogos no começo da Inteligência Artificial, descreve:

Na verdade, os programas desses começos da Inteligência Artificial continuavam a ser muito combinatórios, isso é, eram capazes de testar um número muito grande de possibilidades sem contudo poder examinar cada uma delas em profundidade. Por exemplo, tratando-se de programas de jogos de xadrez, tais programas consideram, em primeiro lugar, todas as jogadas possíveis e, para cada uma dessas jogadas, passam em revista todas as respostas do adversário e, feito isso, consideram ainda todas as jogadas possíveis. Estes programas vão, portanto, como se diz, até à profundidade 3 (amigo-inimigo-amigo) e se desejarmos ir mais longe, o número de combinações aumenta com uma velocidade considerável.

Olhando o problema apenas dessa forma, a construção de um algoritmo restringir-se-ia na escolha de um dos ramos a ser seguido, baseando-se, possivelmente, em tabelas contendo dados estatísticos das possíveis melhores jogadas, em função da distribuição das peças no tabuleiro em cada jogada. Mesmo para uma tabela que contenha uma grande variedade de opções, é provável que, após uma série de jogos realizados, um jogador bem treinado identificasse o padrão de jogar da máquina e passasse a prever suas próximas jogadas, caracterizando, em essência, um ato mecânico. Passaremos, agora, a analisar o conceito de Realimentação.

2.2.2 REALIMENTAÇÃO

A realimentação, também conhecida como retroação ou feedback, desempenha papel fundamental na Cibernética, ciência do controle e comunicação no animal e na máquina, tal como definida por Wiener (1970). Simplificadamente, o processo de realimentação pode ser descrito pela ação de tomar uma amostra da saída (S), produzida por um processo (P) qualquer aplicado sobre a informação da entrada (E), comparar com o resultado desejado e então fazer ajustes ou correções para que o processo gere uma nova saída, dentro de um fluxo contínuo de controle, conforme podemos observar na representação seguinte:

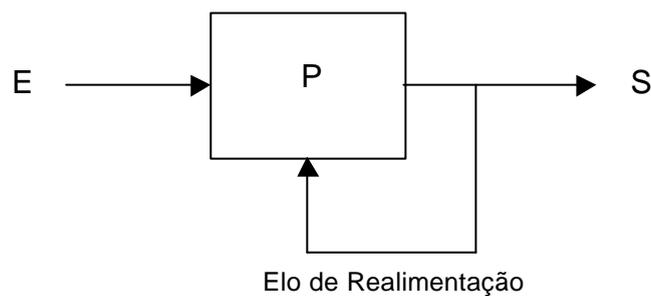


Figura 3: Sistema de Controle por Realimentação

A realimentação já era objeto de estudos dos fisiologistas, nos quais Wiener buscou inspiração. A fisiologia integra as diversas funções dos órgãos e células constituindo um todo, um conjunto funcional, que é o corpo. Uma das primeiras perguntas feitas pelos fisiologistas é a de como órgãos e sistemas distintos são controlados, de modo que nenhum funcione em excesso enquanto outros deixem de desempenhar seu papel. De acordo com Guyton e Hall (1997), a resposta foi a de que o corpo é dotado de uma rede de controles por realimentação, que produzem o balanceamento necessário, sem os quais não conseguiríamos viver. Esse sofisticado sistema de controle interno é conhecido como homeostasia. O termo homeostasia

é, assim, utilizado pelos fisiologistas para designar a manutenção das condições estáticas (ou constantes) do meio interno, no qual todos os órgãos e tecidos do corpo desempenham funções que ajudam a manter essas condições constantes. Tanto nos organismos vivos como nas máquinas, os sistemas de controle tendem a manter o sistema em equilíbrio.

O controle por realimentação também pode ser utilizado como um mecanismo para a aprendizagem. Em seu livro *Cibernética*, Wiener (1970) dedica um capítulo para tratar de aprendizagem de máquinas e, em particular, de máquinas programadas para disputar jogos, de onde extraímos as seguintes considerações:

Agora surge a questão da política dentro das regras do jogo. Toda a avaliação de peças, comando, mobilidade e assim por diante é intrinsecamente passível de ser reduzida a termos numéricos; e feito isso, as máximas de um livro de xadrez podem ser utilizadas para determinar os melhores lances em cada estágio. Tais máquinas têm sido construídas. Imaginem-se jogando xadrez contra semelhante máquina. Naturalmente, como sempre acontece em xadrez, virão a formar um juízo sobre a personalidade enxadrística do oponente. Verificarão que, sempre que a mesma situação surgir duas vezes no tabuleiro, ele reagirá de maneira idêntica, e que a sua personalidade é muito rígida. Se algum artifício surtir efeito, isso se repetirá sempre nas mesmas condições. Assim, não é impossível a um perito traçar uma linha de seu adversário-máquina e derrotá-lo todas as vezes. Entretanto, existem máquinas que não são vencíveis de maneira tão banal. Suponhamos que de poucas em poucas partidas a máquina se detenha e use suas facilidades para outro propósito. Dessa vez, ela não joga contra um oponente, mas examina todas as partidas anteriores que registrou em sua memória a fim de determinar o peso das diferentes avaliações do valor das peças, comando, mobilidade e coisa parecida que conduzirá com mais certeza à vitória. Dessa maneira, tira proveito não só de seus malogros mas dos sucessos de seu oponente. Substitui agora as avaliações anteriores por outras novas e continua jogando como uma máquina nova e melhor. Tal máquina não mais teria uma personalidade tão rígida e os artifícios contra elas usados e que deram resultado uma vez, não de falhar por fim. Mais do que isso, ela pode absorver no curso do tempo algo da política de seus opositores. (WIENER, 1970, p. 212)

Wiener alerta que tudo isso é muito difícil de ser feito em xadrez e, na realidade, o desenvolvimento dessa técnica ainda não tinha sido o suficiente para dar origem a tal máquina.

O modelo de controle por realimentação apresentado anteriormente parece válido, num primeiro momento, para estabelecer-se um sistema de aprendizagem por experiência acumulada. Nos organismos vivos, em particular nos animais superiores, tal aprendizagem

torna-se essencial para a sobrevivência. Com esse modelo transferido para a máquina, teríamos a possibilidade da aprendizagem ser aplicada nas tarefas realizadas por tais máquinas.

Latil (1968) descreve as características de ajuste na postura corporal realizada por um esquiador, durante o percurso de descida numa pista de esqui, para tentar manter-se em equilíbrio. Consideremos, então, um exemplo hipotético ao imaginarmos a construção de um robô esquiador. Ele seria capaz de se deslocar em linha reta numa pista de gelo em declive e, por meio de um sistema de realimentação dotado de sensores, ele receberia sinalizações do ambiente que lhe permitiriam fazer correções em sua rota, desviar-se de obstáculos e manter-se constante em direção ao seu objetivo, que seria o ponto de chegada. Testamos tal robô numa pista idealizada para a experiência, com velocidades moderadas, e obtemos resultados satisfatórios. Poderíamos agora inscrever esse robô para participar de uma competição com esquiadores profissionais? Em nossa opinião, a resposta seria não. Observando pistas de esqui naturais, verificamos que elas estão repletas de imperfeições decorrentes da geografia do terreno. Um esquiador experiente, durante seu percurso pela pista, inclinará seu corpo em cada uma das posições possíveis, como forma de compensar as imperfeições que se apresentam, de maneira a manter seu corpo o mais próximo possível da posição perpendicular ao solo, de maneira a garantir seu equilíbrio. Podemos considerar esse tipo de interação com o ambiente como um processo de controle por realimentação, já que isso é realizado de maneira praticamente mecânica pelo corpo do esquiador, em busca de seu centro de gravidade. Porém, se esse esquiador fosse vendado, cairia pouco depois de iniciar seu movimento. Isso se deve ao fato da perda de referência para manter-se o equilíbrio longitudinal. O “sistema de realimentação” captaria as menores diferenças de declive, causando reações nas articulações e na musculatura. Mas quando realizam tal ação diante de uma diminuição brusca do declive, com o corpo projetado para frente devido à descida, já é tarde para se restabelecer o

equilíbrio. Assim, o nosso robô esquiador, com seu sistema de controle por realimentação, tombaria na primeira saliência de neve com que se deparasse, pois ao ser avisado por seu sistema de realimentação sobre tal irregularidade já seria tarde para se realizar um ajuste ou correção em suas inclinações. Assim, parece haver um limite para a eficiência de controle obtida por sistemas de realimentação, associado a tempos de resposta. A frase “[...] o resultado dá uma contínua retroação para um aperfeiçoamento constante à base da experiência” não se aplicaria ao nosso robô esquiador. Logo, considerar a realimentação como suficiente para atribuir ao programa de jogos capacidade semelhante ao aprendizado nos seres humanos nos parece algo inadequado, já que tal aprendizado não se restringe ao controle proporcionado por um sistema do tipo realimentação.

Já na afirmação “Poderíamos dizer que a máquina apresenta traços do que chamamos de intuição [...]”, gostaríamos de considerar o que escreve o psicólogo americano B. F. Skinner, em capítulo intitulado “Ensinando a Pensar”, no livro *Tecnologia do Ensino*:

A concepção tradicional é a de que pensar é uma atividade obscura, intelectual, “cognitiva”, algo que ocorre na mente e requer o uso do poder e da faculdade da razão. Conduz à ação quando se expressam os pensamentos que resultam do pensar, mas o próprio pensar não é comportamento. Pode algumas vezes ser observado pelo pensador, mas pode também ser inconsciente, e os relatos introspectivos não são por isso nem muito consistentes, nem muito úteis. Exemplos notáveis de pensar parecem tender a surgir de intuições obscuras, e os grandes pensadores raramente tiveram grandes pensamentos sobre o pensar. Isso é particularmente de lamentar, uma vez que o pensar nesse sentido não pode ser nunca observado por ninguém exceto o pensador. (SKINNER, 1972, pg. 111)

Embora Skinner seja contrário ao introspeccionismo, a leitura desse trecho nos remete à noção geralmente aceita de que a intuição é um processo interno, ao qual um observador externo não tem acesso. No entanto, no caso das máquinas que estamos considerando sempre será possível traçar o caminho percorrido, e os pontos de decisão que a levaram a tomar aquele caminho, através de um processo de engenharia reversa.

2.3 CONSIDERAÇÕES SOBRE O CAPÍTULO

Investigando os conceitos de algoritmo e de retroação, na seção anterior, não encontramos elementos que pudessem sustentar a afirmação de que um programa de computador para jogar damas apresentasse traços de intuição e aprendizado dos seres humanos, tal como postulado na citação de Epstein (1973).

Sobre uma concepção simplificada de que a Inteligência Artificial significaria levar uma máquina a fazer algo que se pensava que ela não pudesse fazer, Schank e Birnbaum argumentam:

Os melhores exemplos do que acontece com essa concepção da IA são os programas de xadrez. Trata-se de IA? Hoje as pessoas diriam que não. Anos atrás eram. O que aconteceu? A resposta é que eles funcionavam. Constituíam IA enquanto se ignorasse como funcionavam. Quando todo o trabalho de engenharia terminou, e eles funcionaram bem o bastante para serem usados pelas pessoas comuns, deixaram de ser vistos como IA. [...] as pessoas confundiam fazer uma máquina efetuar algo inteligente como fazê-la ser um modelo da inteligência humana - e certamente esses programas não são muito inteligentes em qualquer sentido mais aprofundado. (SCHANK e BIRNBAUM, 1996, pg. 97)

No âmbito da Ciência, a investigação dos processos mentais através de métodos artificiais se justifica como meio para se tentar chegar à compreensão da cognição humana e dos diversos aspectos relacionados ao conhecimento. Do ponto de vista prático, outras questões podem ser levantadas sobre a validade de tal método, como por exemplo “Qual é a utilidade de se construir máquinas para disputar jogos?”. Aqui, gostaríamos de destacar uma passagem de Wiener (1970):

Os livros de teoria de xadrez não são escritos com base no ponto de vista de Von Neumann³. São compêndios de princípios extraídos da experiência prática de enxadristas que jogam contra outros enxadristas de alta qualidade e largo conhecimento; e estabelecem certos calores ou ponderações a serem atribuídos à perda de cada peça, à mobilidade, ao comando, ao desenvolvimento, e a outros fatores que podem variar conforme o estágio do jogo. (WIENER, 1970, p. 213-214).

Ao dizer que os enxadristas estabelecem certos “calores” à perda de cada peça, Wiener nos alerta da impossibilidade de representação formal de todas as alternativas que seriam consideradas por um jogador humano. De modo geral, os jogos foram desenvolvidos por seres humanos para serem jogados por seres humanos. Existe um componente social na interação entre as pessoas, durante uma partida, que não pode ser implementada na disputa contra uma máquina ou um programa de computador.

Muitas críticas surgiram à idéia de que a mente humana poderia ser simulada por um computador, a partir do processamento formal de símbolos, e que esse poderia apresentar sinais da inteligência humana. Um dos críticos mais rigorosos foi o filósofo John Searle. Em seu artigo publicado em 1980, Searle (1990) estabelece os conceitos de Inteligência Artificial Fraca (IA Fraca) e de Inteligência Artificial Forte (IA Forte). Na IA Fraca, o computador desempenha o papel de uma poderosa ferramenta para ser utilizada no estudo da mente, permitindo formular e testar hipóteses. Na IA Forte, um computador adequadamente programado desempenha o papel da própria mente, com a habilidade de entendimento e de outros estados cognitivos. O filósofo afirma não ter objeções à IA Fraca, mas discorda da posição defendida pela IA Forte. Searle apresenta então a metáfora do Quarto Chinês, cuja argumentação consiste em demonstrar que os programas criados por Roger Schank, os quais apresentam uma resposta inferida a partir de uma estória, cujo conteúdo não contém explicitamente a resposta requerida, conforme Schank e Abelson (1977), são destituídos de compreensão.

³ Wiener se refere ao trabalho de Von Neumann e Morgenstern (1944).

Para Searle, a máquina realiza apenas a manipulação formal de símbolos, sem o entendimento de seu significado. Para justificar sua posição, Searle idealizou um experimento hipotético, no qual estando ele dentro de um quarto fechado, munido por um conjunto de símbolos chineses e um livro de regras de como manipular tais símbolos, produziria uma determinada saída (resposta) em função de uma determinada entrada (estória). Searle afirma que, mesmo sem entender nada do idioma chinês, em seu experimento seria capaz de receber uma estória em chinês e devolver uma resposta em chinês, adequada ao contexto da estória. Para tanto, se utilizaria da mera manipulação formal dos símbolos, sem considerar seu significado. A partir dessa analogia, Searle considera que programas de computador são capazes de manipular informações, mas não verdadeiramente compreende-las, impedindo-os, portanto, de possuir estados cognitivos.

Finalizando este capítulo, podemos destacar que a área de Inteligência Artificial (IA) surgiu a partir dos trabalhos em computação realizados nos anos 50, com o pressuposto de que a mente humana poderia ser reproduzida por um programa de computador, através da representação formal e da manipulação de símbolos. Para Searle (1990), a perspectiva da IA Fraca considera o desenvolvimento de programas computacionais apenas como um meio de testar teorias de como operações cognitivas são realizadas por seres humanos. Já a IA Forte considera o programa de computador como a própria mente, capaz de compreensão e de outros estados cognitivos. Embora capazes de realizar o processamento simbólico, o filósofo considera que computadores não são capazes de possuir estados cognitivos.

A despeito do grande aumento de velocidade de processamento obtido nos computadores ao longo de sua evolução, eles mantêm basicamente a mesma arquitetura proposta por Von Neumann⁴. Assim, tais computadores utilizados para pesquisas em IA são

⁴ Arquitetura de processamento seqüencial.

apenas máquinas velozes que executam algoritmos pré-programados, aos quais não se pode imputar a característica de inteligência.

A crença de que todo conhecimento pode ser formalizado e que a mente pode ser vista como um mecanismo que opera de acordo com regras formais se mostrou incorreta. Mesmo considerando todos os resultados obtidos pela IA, ela não cumpriu as expectativas exageradas a que se propunha. O sucesso da abordagem de processamento simbólico, utilizado na IA, dependeria de se encontrar algoritmos capazes de indicar os passos necessários para a solução de problemas propostos, quaisquer que sejam eles, tarefa esta que nem sempre é possível.

Uma abordagem alternativa surgiu, conhecida como Conexionismo, com objetivos semelhantes aos trabalhos de IA, porém com uma nova concepção a partir da tentativa de simular algumas das características da estrutura do cérebro e de seu funcionamento, sobre a qual trataremos no próximo capítulo.

CAPÍTULO 3

CONEXIONISMO: A METÁFORA DO CÉREBRO

A unidade fundamental dos computadores biológicos, a célula nervosa ou neurônio, na realidade não se parece em nada a um transistor em seu funcionamento interno. O código pelo qual os neurônios comunicam-se entre si, sem dúvida, se parece um pouco com os códigos de pulsos dos computadores digitais, mas o neurônio individual é uma unidade de processamento de dados muito mais sofisticada do que o transistor.

Richard Dawkins, 1976.

APRESENTAÇÃO

Neste capítulo detalharemos o conceito de Conexionismo, cuja idéia principal é a de que para reproduzir algumas das capacidades intelectuais do cérebro seria necessário reproduzir algumas características da sua estrutura. Essa abordagem deu origem às Redes Neurais Artificiais, que são uma tentativa de simular o funcionamento do cérebro humano a partir do conhecimento sobre o funcionamento do neurônio biológico.

O Conexionismo está fortemente interligado com os princípios de aprendizagem, uma vez que uma rede neural precisa aprender ou ser treinada a realizar determinada tarefa. O objetivo principal deste capítulo é a investigação, a partir dos estudos sobre aprendizagem em redes neurais, sobre a possibilidade de tais redes apresentarem traços de capacidades cognitivas humanas. Para tanto, dividimos o capítulo em três seções. Na primeira, introduziremos o conceito de Rede Neural Artificial (RNA), a partir do detalhamento do funcionamento do neurônio artificial, e de um breve histórico do Conexionismo. Na segunda, verificaremos alguns dos métodos utilizados para treinamento e aprendizagem de RNA, e na terceira seção concluiremos o capítulo apresentando algumas das críticas realizadas ao Conexionismo.

3.1 ORIGEM DO CONEXIONISMO

Na abordagem conexionista a hipótese central é a de que para se reproduzir algumas das capacidades do cérebro é necessário reproduzir determinadas características de sua estrutura. Assim, partindo do pressuposto de que a estrutura principal do cérebro é formada por células nervosas (neurônios) interligadas entre si, as Redes Neurais Artificiais (RNA) procuram simular o funcionamento do cérebro utilizando elementos unitários conectados entre si (daí o nome conexionismo), formando uma rede. Os elementos individuais da RNA, inspirados no neurônio biológico, recebem o nome de neurônios artificiais.

A idéia sobre máquinas cujo princípio de funcionamento incorporam analogias com a estrutura do cérebro já aparecia no trabalho de Von Neumann sobre o Edvac (Von Neumann, 1945), no qual o autor estabelecia diversas correspondências entre os circuitos propostos para o seu dispositivo e os neurônios biológicos. Mas o trabalho considerado pioneiro nessa área foi o artigo *A logical calculus of the ideas immanent in nervous activity*, publicado em 1943 pelo neurofisiologista Warren McCulloch, em conjunto com o matemático Walter Pitts (McCulloch e Pitts, 1990), no qual os autores apresentam a idéia de redes lógicas constituídas por neurônios artificiais. Partindo de determinados predicados físicos dos neurônios biológicos, dentre eles o de que a atividade do neurônio é do tipo “tudo-ou-nada”, e de que um certo número de sinapses devem ser excitadas para que um neurônio seja ativado, propõem um modelo lógico formal do funcionamento do neurônio biológico, que poderia ser reproduzido artificialmente.

Mesmo com a grande evolução ocorrida no conhecimento sobre o funcionamento neurofisiológico do cérebro ao longo das últimas décadas, as idéias de McCulloch continuam válidas e aplicáveis, conforme destaca Dupuy (1996):

[...] podemos, por exemplo, refletir sobre qual teria sido a atitude de McCulloch se tivesse podido dispor de nossos conhecimentos sobre o sistema nervoso. Se tivesse sabido o que sabemos sobre os papéis respectivos da transmissão elétrica e da transmissão química nas sinapses, sobre os mecanismos moleculares e atômicos da propagação do influxo nervoso, teria ele se valido disso para especificar ou até corrigir seu modelo, ou o teria rejeitado? O que McCulloch procurava era abstrair do sistema nervoso central um nível de organização e, nesse nível, alguns traços que permitissem reconstituir o mais economicamente possível as funções observadas no sistema real. Esse nível, ele o encontrara: o dos neurônios; esses traços, ele o isolara: os de seus neurônios idealizados. A vertiginosa descida de nossos conhecimentos aos níveis de organização cada vez mais elementares não o teria afetado, pois, ele não a teria considerado um progresso decisivo no que diz respeito ao problema que se propunha. (DUPUY, 1996, p. 67, 68).

Uma RNA constitui-se da combinação desses elementos individuais, cada qual formando um nó da rede e atuando conjuntamente com os demais. Esses nós da rede, ou neurônios artificiais, são conhecidos por TLU, *Threshold Logic Unit* (Unidade de Lógica Limiar). A representação esquemática de uma TLU, com “n” entradas, pode ser vista na figura abaixo, adaptada de Braga et al. (2000):

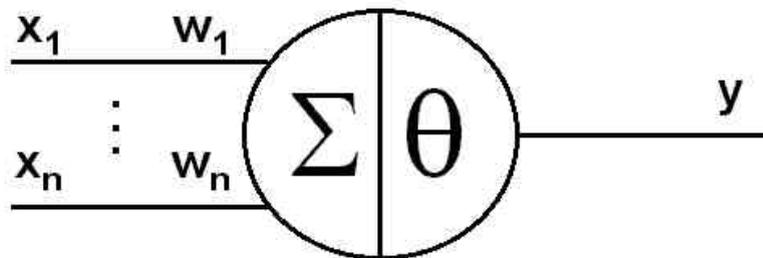


Figura 4: Representação esquemática de uma TLU, ou neurônio artificial.

Para Wasserman (1989), as RNA são apenas levemente inspiradas no modelo biológico, adotando uma série de simplificações para tornar possível a sua implementação em computadores. Neurônios biológicos são formados por um corpo celular e por terminais de conexão, conhecidos como dendritos e axônios. Os dendritos de um neurônio são os terminais que recebem os impulsos elétricos oriundos dos axônios de outros neurônios, e os conduzem

até o corpo celular. A junção entre dois neurônios recebe o nome de sinapse, podendo a força de conexão entre eles variar. Uma explicação simplificada, conforme Braga et al. (2000), de acordo com o conhecimento que se tinha na época e que serviu de base para o modelo de neurônio artificial desenvolvido por McCulloch e Pitts, é a de que os diversos impulsos recebidos são acumulados e o neurônio se torna excitado quando a soma dos impulsos que recebe ultrapassa o seu limiar de excitação, liberando um novo impulso elétrico em seu axônio para ser transmitido para o neurônio seguinte.

A força de conexão na sinapse pode ser representada por um peso associado à entrada, cuja variação de valor implica em disparar ou não um sinal na saída. A TLU emula então o funcionamento do neurônio através de um mecanismo que realiza a soma ponderada das entradas (x_n), com seus respectivos pesos (w_n), que representam a força de conexão de cada sinapse, e decide se gera um impulso excitatório em sua saída ($y = 1$), ou não ($y = 0$), comparando o resultado da soma com o valor do seu limiar de excitação (θ). Então, de acordo com Haykin (1994), teremos:

A função de ativação (a), da TLU, dada por:

$$a = \sum_{i=1}^n x_i \cdot w_i = x_1 \cdot w_1 + x_2 \cdot w_2 + \dots + x_n \cdot w_n$$

na qual : $x_1 \dots x_n =$ Entradas

$w_1 \dots w_n =$ Pesos associados às entradas

E sendo θ o limiar de excitação, a função de saída será dada por:

$$y = \begin{cases} 1, & \text{se } a \geq \theta \\ 0, & \text{se } a < \theta \end{cases}$$

A combinação de várias TLU conectadas entre si, formando uma rede com vários nós, constitui uma RNA.

A partir dessas idéias, estabeleceu-se um novo campo de investigação, que se tornou conhecido como Conexionismo. Diversos trabalhos surgiram nessa área e, de acordo com Haykin (1994), dentre os primeiros destaca-se a publicação de Rosenblatt “*Principles of Neurodynamics*”, em 1958, no qual o autor apresentou o Perceptron, um modelo mais geral do que aquele proposto por McCulloch e Pitts, permitindo a modificação dos pesos nas entradas e mantendo fixa a topologia da rede. Através de um algoritmo de treinamento, as redes poderiam ser utilizadas para classificar determinados tipos de padrões. A partir daí, foram desenvolvidas propostas para utilização de RNA para reconhecimento de padrões visuais e para modelagem da memória humana, dentre outras aplicações. Algum tempo depois, Minsky e Papert (1969) mostravam que determinados tipos de problemas não poderiam ser resolvidos pelo Perceptron. Haykin destaca que, como resultado, o interesse por pesquisas nessa área diminuiu fortemente e só voltou a chamar a atenção na década de 80, em especial a partir do trabalho de Hopfield, “*Neural networks and physical systems with emergent collective computational properties*” (1982), de Kohonen, “*Self-organization and associative memory*” (1984) e de Humelhart, Hinton e Williams, “*Learning internal representations by error propagation*” (1986).

As redes neurais artificiais apresentam como vantagem o fato de não precisarem ser, de forma prévia, detalhadamente programadas para resolver um determinado problema, tal como ocorre na maioria dos programas de computador utilizados na IA. Ao invés disso, elas podem ser treinadas na solução do problema em que estamos interessados. Na próxima seção abordaremos os princípios de treinamento e aprendizagem utilizados em RNA.

3.2 ALGUMAS CARACTERÍSTICAS DAS REDES NEURAS ARTIFICIAIS

Nesta seção apresentaremos duas características normalmente associadas às Redes Neurais Artificiais, que são a Aprendizagem e a Distribuição do Conhecimento.

3.2.1 TREINAMENTO E APRENDIZAGEM

Ao invés de receber informações detalhadas sobre como proceder para resolver um determinado problema, uma RNA precisa aprender a resolvê-lo. Mas como uma RNA aprende? Para Rumelhart e McClelland (1986):

Aprender implica em modificar conexões. Outra característica chave do nosso modelo, que deriva do nosso entendimento dos mecanismos de aprendizagem no cérebro, é de que o conhecimento está nas conexões ao invés de nas unidades propriamente. Além disso, assume-se geralmente que aprendizagem implica na modificação das forças de conexão.¹ (RUMELHART & MCCLELLAND, 1986, p. 132).

Assim, treinar uma TLU consiste basicamente em alterar os valores dos pesos associados aos sinais de entrada e em ajustar o valor de limiar de excitação, de modo a obter o resultado desejado na saída. Se quisermos, por exemplo, implementar em uma TLU de duas entradas a função lógica “E”, na qual só teremos o valor 1 na saída quando ambas as entradas possuírem o valor 1, poderíamos utilizar para os valores de peso $w_1 = 1$ e $w_2 = 1$, e para o valor de limiar $\theta = 1,5$, obtendo os seguintes valores da função de ativação a e de saída y , de acordo com os valores de entrada x_1 e x_2 :

¹ Tradução nossa.

x_1	x_2	A	y
0	0	0	0
0	1	1	0
1	0	1	0
1	1	2	1

Tabela 1: Função lógica “E” implementada numa TLU.

Chamamos de treinamento supervisionado quando informamos à rede qual saída esperamos para cada vetor de entrada, como no exemplo anterior da implementação da Função lógica “E”, sendo a tarefa da rede encontrar o ajuste de seus parâmetros (pesos sinápticos) de forma a estabelecer uma relação entrada e saída. Uma das regras para ajuste de valor dos pesos sinápticos surgiu a partir do trabalho de Hebb (1949), e ficou conhecida como Aprendizado Hebbiano:

Quando o axônio da célula A está próximo suficiente para excitar a célula B e repetidamente ou persistentemente toma parte em seu disparo, algum processo de acréscimo ou mudança metabólica tem lugar em uma ou em ambas as células, tal que a eficiência de A, como uma das células de disparo de B, é aumentada.² (HEBB, 1949, p. 62).

Uma forma simplificada de aplicar essa regra ao treinamento de RNA, conforme proposto por Rumelhart e McClelland (1986), é a de “reforçar o peso da conexão entre duas unidades, quando ambas estão excitadas simultaneamente”.

No treinamento não-supervisionado somente as entradas são fornecidas e a rede deve identificar padrões nos dados de entrada. Um exemplo poderia ser a utilização de uma RNA no reconhecimento de padrões visuais, para a tentativa de identificação de faces humanas.

² Tradução nossa.

3.2.2 DISTRIBUIÇÃO DO CONHECIMENTO

Numa RNA, considera-se o resultado cognitivo desejado (como o reconhecimento de um padrão visual, por exemplo) como emergente do estado global da rede, decorrente do conjunto dos estados individuais dos neurônios que compõem a rede. Dessa forma, a representação do conhecimento, ou da solução do problema proposto, não se encontra num único elemento da rede, mas está distribuída por toda a rede. Similarmente ao cérebro humano, a informação não depende de uma localização específica e pode ser recuperada na ocorrência de falhas parciais na rede, como no caso de atuação defeituosa de alguns dos neurônios.

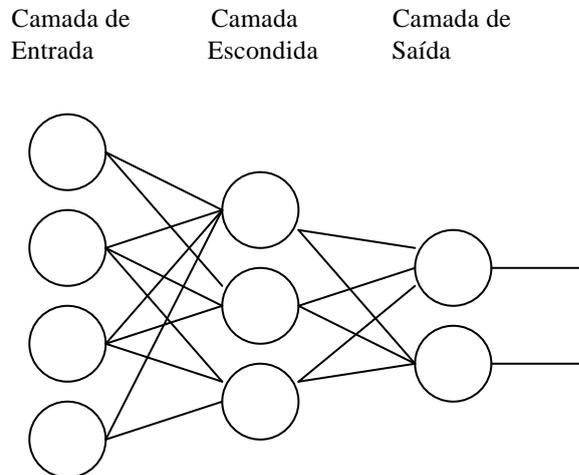


Figura 5: Exemplo de uma Rede Neural Multicamadas.

3.3 CONSIDERAÇÕES SOBRE O CAPÍTULO

Para Franklin (1995) a diferença entre as abordagens da IA e do Conexionismo reside na interpretação dada à atividade mental, que na IA seria vista como a execução de um programa armazenado, numa seqüência pré-estabelecida e, no Conexionismo, como o alcance de uma configuração estável por uma rede composta de várias unidades interligadas entre si. Uma das críticas feitas à abordagem conexionista deve-se a Fodor e Pylyshyn (1988), que argumentam que pensamentos possuem estruturas compostas e que essas estruturas não são contempladas satisfatoriamente pelas implementações de RNA. Para os autores, modelos simbólicos baseados na IA poderiam ser construídos a partir de um modelo inicial conexionista, que se encarregaria de certas estruturas em um nível mais baixo, e essa seria basicamente a contribuição a ser oferecida por modelos conexionistas. Para a investigação cognitiva seria suficiente a análise simbólica, abstraindo-se do tipo de meio em que ela ocorre.

Teixeira (1998) aponta como limites da abordagem conexionista dois problemas específicos, aos quais ele denomina de “problema de descrição” e de “problema epistêmico”. O problema de descrição consiste na indagação de se nosso cérebro seria “capaz de produzir uma noção de complexidade que nos permita descreve-lo” (Teixeira, 1998, p. 116). Ou seja, seria possível produzir uma minuciosa descrição do cérebro de tal forma a criar um mapeamento de estados mentais específicos? O problema epistêmico, partindo da suposição de que uma descrição completa do cérebro seria necessariamente mais complexa do que o próprio cérebro, consiste na indagação de se o cérebro poderia produzir algo mais complexo que ele mesmo. Se sim, como seria possível garantir que tal descrição estaria correta?

Para os cientistas cognitivos que discordam da necessidade de representações mentais para explicação da cognição, a crítica ao Conexionismo é a de que esse mantém a abordagem representacionista, ainda que a representação de um conceito, ou da solução de um problema, esteja distribuída na rede. Além, disso, considerando-se o cérebro como necessário e suficiente para a cognição, estaríamos incorrendo na questão colocada por Fetzer (2001, p. 33), se “somos cérebros em barris de vidro?”. Para os pesquisadores adeptos ao conceito de Cognição Incorporada e Situada (CIS), que considera a cognição como um processo decorrente de um agente que possui um corpo interagindo num ambiente, a resposta a essa questão é não, ou seja, “a idéia básica da CIS é que nossas ações não são dirigidas pelas estruturas representacionais no cérebro, mas, ao invés disso, dependem dos processos dinâmicos que se estabelecem entre cérebro, corpo e mundo” (Haselager, 2004, p. 224).

Mas, possuir um cérebro e um corpo, interagir com o meio, e ser capaz de estados cognitivos, são características de certos organismos biológicos, como os animais e os seres humanos. E sendo uma característica comum desses organismos o fato de possuírem vida, poderíamos então pensar na questão de se sistemas artificiais poderiam exibir vida e, em caso positivo, se tais sistemas seriam capazes de estados cognitivos. Para discutir essas possibilidades, introduziremos no próximo capítulo os conceitos relacionados à abordagem de Vida Artificial.

CAPÍTULO 4

ABORDAGENS BIO-INSPIRADAS: A METÁFORA DA VIDA

Existem dois caminhos para os filósofos seguirem em seus encontros com a Vida Artificial: eles podem considerá-la como um novo jeito de fazer filosofia ou simplesmente como um novo objeto merecedor de atenção filosófica utilizando métodos tradicionais. Em cada caso há algo diferente a ser feito, mas eu recomendo aos filósofos darem um salto e considerarem o primeiro como o mais importante e promissor.

Daniel Dennett, 1995.

APRESENTAÇÃO

O pesquisador em Ciência da Computação Christopher Langton rompe com as abordagens tradicionais utilizadas na Inteligência Artificial (Simbólica e Conexionista), ao enunciar a concepção com inspiração biológica designada de Vida Artificial (*Artificial Life*, ou *ALife* como ficou conhecida), para a qual entidades vivas podem ser criadas artificialmente. Cientistas envolvidos com projetos em *ALife* estão motivados pela idéia de que programas de computador exibem sinais de vida pelas mesmas razões fundamentais dos organismos biológicos. Eles descrevem essa similaridade utilizando termos tais como emergentes e auto-organizados para qualificar como os sinais de vida se apresentam em entidades vivas, naturais ou artificiais. Assim, as coisas vivas constituiriam uma classe própria, da qual fariam parte tanto os organismos biológicos, tal como os conhecemos atualmente, como outras entidades criadas artificialmente. Uma das técnicas utilizadas em *ALife* são os Algoritmos Genéticos, baseados nas idéias de seleção natural e evolução.

Além da abordagem da *ALife*, neste capítulo apresentaremos outras abordagens com inspiração biológica que tem sido comumente propostas no estudo da cognição, que são a IA Baseada em Comportamentos, a Robótica Evolucionária e os Mundos Virtuais.

4.1 VIDA ARTIFICIAL

Em 1987 Christopher Langton organizou um workshop sobre síntese e simulação de sistemas vivos, em Los Alamos, Estados Unidos, do qual participaram 160 cientistas de diversas especialidades, incluindo físicos, biólogos, antropólogos e cientistas da computação, entre outros. Langton (1989) justificou a organização do workshop em função de sua própria dificuldade na busca de literatura sobre simulação e modelagem computacional de sistemas biológicos. Durante cinco dias, o workshop proporcionou apresentações sobre temas como modelos matemáticos da origem da vida, programas computacionais utilizando mecanismos de evolução darwiniana, modelos analíticos de sistemas vivos, ecossistemas simulados em computador e autômatos auto-replicantes. Desse rico encontro multidisciplinar, surgiu um novo campo de pesquisas designado por Vida Artificial. Na publicação da compilação dos resultados do workshop, Langton define Vida Artificial:

Vida Artificial é o estudo de sistemas criados pelo homem, que exibem comportamentos característicos de sistemas vivos naturais. Complementa as ciências biológicas tradicionais, preocupadas com a *análise* de organismos vivos, pela tentativa de *sintetizar* comportamentos tipo-vida dentro de computadores e outros meios artificiais. Por estender o fundamento empírico sobre o qual a biologia está baseada *para além* da vida de cadeias-de-carbono que envolveu a Terra, a Vida Artificial pode contribuir com a biologia teórica situando *a vida-como-nós-a-conhecemos* dentro do quadro maior da *vida-como-ela-poderia-ser* .¹ (LANGTON, 1989, p. 1).

Mas o que significa a *vida como ela poderia ser* ? Langton diz que Biologia seria o estudo científico da vida, de qualquer tipo, mas que na prática ela se restringe ao estudo da vida baseada na química de cadeias de carbono. Isso traria a dificuldade de se conseguir deduzir uma teoria geral a partir de casos particulares. Na Vida Artificial (*ALife*) essa restrição não existe, pois essa abordagem considera que a vida pode habitar outras bases

¹ Tradução nossa. Grifos do autor.

materiais, além daquelas baseadas em carbono, uma vez que os processos dinâmicos que constituem a vida devem compartilhar características universais que permitam identificar sua presença, independente do substrato material em que se encontre. Na *ALife* a vida é vista como a propriedade de organização da matéria ao invés de a propriedade da matéria na qual está organizada. Assim, um conjunto de partes artificiais corretamente organizadas podem exibir o mesmo comportamento dinâmico de sistemas naturais, permitindo que se observe a *vida como ela poderia ser*.

A *ALife* pode ser entendida como um campo de estudos que procura recriar os princípios dinâmicos dos sistemas biológicos em meios artificiais, sendo o computador um desses meios. Ao invés da abordagem analítica normalmente utilizada em Biologia, na qual procura-se dividir sistemas complexos em partes menores para realizar a análise, a *ALife* utiliza a abordagem sintética, na qual procura-se compor o sistema a partir da união (síntese) de suas partes menores. Langton considera o computador como a ferramenta ideal para o estudo sintético da vida e elenca as principais características dos modelos de *ALife* baseados em computador:

- Consistem de populações (grupos) de programas simples.
- Não existe um programa único que conduza os demais programas.
- Cada programa detalha o modo como uma entidade simples reage à situações em seu ambiente.
- Não existem regras no sistema que imponham um comportamento global.
- Qualquer comportamento em níveis mais altos daqueles dos programas individuais é considerado emergente.

Para exemplificar essas características, Langton cita a experiência de Craig Reynolds de simulação do comportamento de bandos de pássaros. Observando-se a revoada de pássaros em grupo, verifica-se que eles mantêm certas características em comum como o fato de voarem próximos entre si, sem se chocarem, e do bando todo mudar de direção a partir do movimento sincronizado dos indivíduos. Reynolds (1987) procurou simular esse comportamento, definindo três regras para as suas criaturas individuais, que ele chamou de *Boids*: (1) deveriam manter uma distância mínima de outros objetos no ambiente, incluindo os outros *Boids*; (2) manter a mesma velocidade dos *Boids* ao seu redor; e (3) mover-se em direção à maior concentração de *Boids*. Como resultado, ele observou comportamentos de bandos de *Boids* muito parecidos com aqueles de bandos de pássaros. Além disso, ele observou comportamentos emergentes da interação entre aquelas três regras, como o caso de um bando de *Boids*, frente à uma torre existente em seu trajeto, se dividir em dois outros grupos, voltando a se reagrupar após transpor o obstáculo. Dessa forma, o comportamento global de um conjunto de *Boids* é um fenômeno emergente, no qual nenhuma das regras individuais de comportamento dos *Boids* depende do comportamento global observado, ou seja, não existe uma regra global que conduza o comportamento dos *Boids*.

Langton ressalta que o sistema de Reynolds não pretende reproduzir a características de pássaros reais, e que os *Boids* nem remotamente assemelham-se aos pássaros, mas que o comportamento de um bando de *Boids*, esse sim, é o mesmo comportamento de um bando de pássaros. Ou seja, o comportamento de bando independe das características físicas dos indivíduos que o compõe. Cardumes de peixes ou manadas de búfalos apresentam comportamento semelhante ao de bandos de pássaros e, para Langton, o comportamento de um bando de *Boids* representa o mesmo comportamento de bandos reais, ou seja, esse é um exemplo equivalente do comportamento de bando de organismos biológicos criado num sistema artificial.

Os modelos de *ALife* baseados em computador são referenciados como modelos de *software*. Os outros dois modelos são os de *hardware*, como no caso da utilização de robôs, e os modelos de *wetware*, que utilizam elementos orgânicos. Esses têm sido os três meios nos quais os trabalhos de *ALife* são desenvolvidos, e para Taylor e Jefferson (1995) eles provêm poder suficiente para capturar muito da complexidade dos sistemas vivos. Os autores sugerem que a vida natural está organizada em quatro níveis fundamentais: molecular, celular, organismo e população-ecossistema, e que uma entidade viva em qualquer um desses níveis constitui-se de um sistema adaptativo complexo, cujo comportamento emerge da interação de diversos elementos de níveis inferiores. Os modelos *wetware* de *ALife* lidam com os elementos de nível molecular, como no caso das pesquisas que aplicam processos evolucionários artificiais na produção de moléculas de RNA com propriedades específicas. Os modelos *software* de *ALife* podem ser utilizados com elementos de nível celular, como na investigação de sistemas auto-replicantes, por exemplo. E os modelos *hardware* de *ALife* podem ser utilizados no nível de organismo, como na investigação de insetos artificiais.

Os modelos utilizados no nível população-ecossistema, que em geral envolvem maior complexidade, costumam ser classificados como Mundos Virtuais. Um exemplo é o mundo virtual *Tierra*, criado por Thomas Ray, e descrito por Sipper (1995): Ray modelou seu sistema baseando-se no período Cambriano de evolução da Terra, caracterizado pela existência de organismos simples auto-replicantes que deram origem à grande diversidade de vida existente no planeta. Ele introduziu no sistema um único desses organismos auto-replicantes e permitiu que o ambiente evoluísse a partir dele. O resultado foi o surgimento de um ecossistema inteiro, considerado como verdadeiramente “vivo”, composto de organismos de diversos tamanhos e de diferentes parasitas, cada qual competindo pelos recursos disponíveis em seu ambiente computacional (espaço de memória e tempo de CPU).

A abordagem que considera possível que programas de computador, robôs e outros dispositivos sejam considerados vivos recebe o nome de *ALife Forte*, enquanto que, ao contrário, a abordagem que considera tais recursos apenas como ferramentas para a investigação de sistemas biológicos é conhecida como *ALife Fraca*, e geralmente aplicada aos estudos em Biologia Sintética.

Evolução é um fator central na *ALife*, por permitir a adaptação de organismos à um ambiente dinâmico, similarmente ao que ocorre na natureza. Uma das técnicas principais em trabalhos de *ALife* têm sido a utilização dos Algoritmos Genéticos (*Genetic Algorithms*). Eles foram desenvolvidos por John Holland, da Universidade de Michigan, Estados Unidos, juntamente com seus alunos e colaboradores, durante as décadas de 1960 e 1970, como uma técnica de programação baseada nos conceitos da Teoria da Evolução, emprestada da Biologia. De acordo com Goldberg (1989), os objetivos da pesquisa de Holland eram: (1) produzir explicações rigorosas sobre processos adaptativos de sistemas naturais; e (2) desenvolver software de sistemas artificiais que possuíssem os mecanismos de sistemas naturais.

Holland (1975) apresenta uma comparação entre os campos da Inteligência Artificial e da Genética, relacionando suas estruturas e operadores, respectivamente programas e regras de aprendizagem para IA, e cromossomos e operadores genéticos (mutação, recombinação, etc.) para a Genética. Em seu livro, Holland levanta questões tais “Como a evolução produz fixação de organismos em ambientes extremamente incertos para organismos individuais?”, “Como um organismo utiliza sua experiência para modificar seu comportamento de forma benéfica?” e “Como computadores podem ser programados de maneira que a capacidade de resolver problemas seja construída a partir da especificação de ‘o que deve ser feito’ ao invés de ‘como faz-lo’?”. Ele ressalta que as tentativas anteriores de se unir Ciência da Computação e Evolução foram pobres, pois enfatizavam apenas a mutação. A partir de suas

investigações e análises matemáticas sobre processos de adaptação, Holland estava convencido de que a recombinação de grupos de genes pelo processo de reprodução (acasalamento) era uma parte crítica da evolução. Essa abordagem deu origem aos Algoritmos Genéticos, que Holland considera apropriados à Evolução, tanto por mutação como por reprodução.

A idéia dos algoritmos genéticos é relativamente simples, envolvendo cópia e substituição de trechos de strings binárias². Um algoritmo genético básico, que pode ser aplicado a diversos problemas, e composto pelos operadores Reprodução, Crossover e Mutação, pode ser representado pelo seguinte diagrama, conforme Goldberg (1989):

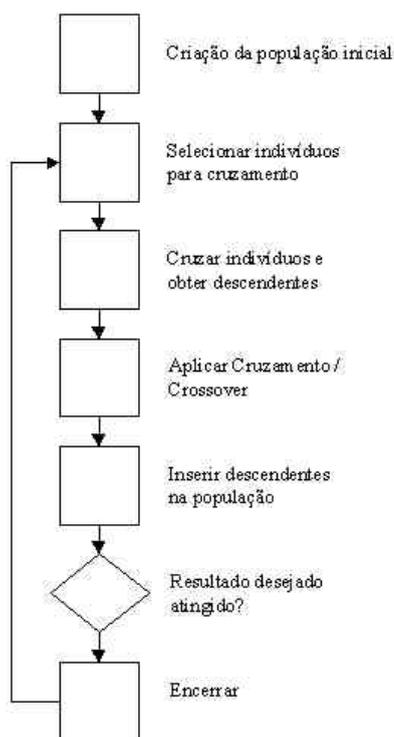


Figura 6: Algoritmo Genético Básico.

² Sequências numéricas contendo zeros e uns.

O cruzamento realiza a combinação (acasalamento) entre indivíduos, a mutação insere mudanças aleatórias e a seleção é responsável pela sobrevivência. Após um número sucessivo de gerações, o algoritmo apresenta o indivíduo melhor adaptado (evolução). O critério para encerramento do algoritmo pode ser estabelecido pelo número desejado de gerações.

BANZHAF et al. (1998) menciona a existência de outras técnicas de algoritmos evolucionários, além do Algoritmo Genético de Holland, citando as Estratégias Evolucionárias (*Evolutionary Strategies*), a Programação Evolucionária (*Evolutionary Programming*) e a Programação Genética (*Genetic Programming*). Esse último, desenvolvido a partir dos trabalhos de John Koza, é utilizado na criação automática de programas de computador, através da seleção e evolução de uma população inicial de programas candidatos, sendo de especial interesse aos pesquisadores que trabalham com modelos *software* de *ALife* e que dependem da criação de códigos de programação.

Além disso, a partir do próprio processo de evolução dos programas e das interações programa-programa e programa-ambiente (sendo o ambiente composto do computador e seus recursos nos quais residem os programas), pode-se observar comportamentos equivalentes àqueles de organismos que competem por recursos em seu ambiente, influenciando e sendo influenciados pelo ambiente, tornando a Programação Genética uma ferramenta de auxílio à implementação de ecossistemas artificiais para pesquisas de *ALife*.

Nesta seção abordamos a origem e algumas das principais características de uma nova área de pesquisas, designada de Vida Artificial, ou *ALife* em sua abreviação em inglês, fortemente inspirada na Biologia, em contraposição à Inteligência Artificial (IA), que traz em suas raízes vínculos mais estreitos com a Psicologia e a Filosofia. Porém, a IA também já havia flertado com a Biologia, produzindo como resultado desse namoro as pesquisas pertencentes à um campo conhecido como *Behavior-based Artificial Intelligence* (IA Baseada em Comportamentos, as quais abordaremos em nossa próxima seção).

4.2 INTELIGÊNCIA ARTIFICIAL BASEADA EM COMPORTAMENTOS

De acordo com Steels (1995), a IA Baseada em Comportamentos (*Behavior-based Artificial Intelligence*) começou a se desenvolver em meados da década de 1980, a partir do trabalho de pesquisadores que passaram a utilizar a abordagem de repertório de comportamentos para desenvolver seus modelos e projetos. Dentre esses trabalhos estão aqueles descritos em Wilson (1991), Maes (1993) e Brooks (1999). Esses comportamentos foram emprestados, em sua maioria, do mundo animal, a partir de estudos em Etologia, Neuroetologia e Etologia Cognitiva, e aplicados ao desenvolvimento de Redes Neurais Artificiais Baseadas em Comportamentos (*Behavior-based ANN*), conforme Dyer (1995), e na Robótica Baseada em Comportamentos (*Behavior-based Robotics*), conforme Mataric (1998).

Consideramos oportuno apresentar uma conceituação das áreas anteriormente citadas, ainda que de forma resumida e simplificada, para termos clareza das inspirações biológicas utilizadas na IA Baseada em Comportamentos. De acordo com Lorenz (1995), a Etologia é a ciência que trata do estudo comparativo do comportamento natural dos animais, aplicando as metodologias dos diversos ramos da Biologia ao comportamento animal e humano. MacLennan (1992) considera a abordagem da Etologia como uma alternativa e, em parte, como uma reação ao Behaviorismo, e que, nessa abordagem, o comportamento de um organismo está intrinsecamente acoplado ao seu ambiente, de tal forma que a remoção do organismo de seu ambiente destrói o contexto no qual o comportamento é manifesto.

De acordo com Keeley (2000), a Neuroetologia tem como objetivo o estudo do comportamento natural dos animais com base no sistema nervoso e pode ser definida, conforme Hoyle (1984), como o estudo dos mecanismos neurais responsáveis pelos comportamentos objetos de estudo da Etologia.

Já a Etologia Cognitiva, conforme descrita por Bekoff (1995), é o estudo evolutivo e comparativo dos processos de pensamento, consciência, crenças e da racionalidade de animais não-humanos.

Com esses conceitos em mente, voltemos a nossa atenção à IA Baseada em Comportamentos. De acordo com Dawson (2002), essa abordagem tem raízes na proposta de Braitenberg (1984) de adoção de uma concepção sintética no estudo da cognição. Braitenberg utilizou o termo Psicologia Sintética para designar esse programa de pesquisas, cuja idéia principal é a de construir mecanismos com projetos simples, observar seu comportamento e então aumentar gradativamente o grau de complexidade desses mecanismos, até que eles passem a manifestar o equivalente a estados mentais. Ou seja, na abordagem sintética a construção de modelos antecede a análise dos comportamentos. Wilson (1985, 1991) estabeleceu uma metodologia de pesquisa propondo o estudo do entendimento da inteligência a partir da simulação de animais artificiais, que ele chamou de *Animats*, em ambientes progressivamente desafiadores aos quais os *animats* deveriam se adaptar. Entre algumas das implementações inspiradas nessa abordagem estão os robôs-insetos de Brooks (1999).

Steels define a IA Baseada em Comportamentos como “uma disciplina científica que estuda como o comportamento de agentes emerge e se torna inteligente e adaptativo. O sucesso nesse campo é definido em termos de sucesso em construir agentes físicos que são capazes de maximizar sua própria auto-preservação, em interação com um ambiente dinamicamente em mudança”³ (Steels, 1995, p. 75). Para entender a mudança de abordagem, Dyer (1995) nos lembra que os trabalhos tradicionais em IA estavam focados em modelar as funções cognitivas superiores da mente humana, tais como a sua capacidade de raciocínio, planejamento, linguagem, compreensão, aprendizagem e criatividade, a partir da tentativa de reprodução das etapas de como esses processos são realizados pelas pessoas, sendo essas

³ Tradução nossa.

capacidades reconhecidas como inteligentes. Mas Dyer também relaciona como inteligentes o repertório de comportamentos dos animais, que precisam procurar pela comida, realizar migrações, alimentar e proteger sua prole, construir ninhos ou procurar abrigos, cortejar seus pares antes do acasalamento, possuir sentido de direção ou rumo a seguir, proteger-se de predadores, etc. Essas tarefas, para Dyer, requerem um alto grau de inteligência, de interação social e de cooperação entre indivíduos. Essa idéia é enfatizada no artigo de Brooks (1990), no qual apresenta suas propostas de *Subsumption Architecture* e *Behavior Language* para desenvolvimento de implementações de IA Baseada em Comportamentos, através do projeto e construção de robôs.

Dyer (1995) descreve em seu trabalho o desenvolvimento de Redes Neurais Artificiais capazes de exibir comportamento cooperativo inteligente. Exemplos da IA Baseada em Comportamentos, utilizando-se de Redes Neurais Artificiais, podem ser encontradas em Roisenberg (1998) e em Silva, Roisenberg e Barreto (2000), que descrevem a utilização de redes neurais hierárquicas para a implementação de comportamentos em agentes autônomos.

O objetivo principal da IA Baseada em Comportamentos, conforme Steels (1995), é o de entender o comportamento inteligente, através da construção de sistemas artificiais. O papel desempenhado pela Inteligência Artificial na investigação de comportamentos é definido por Web e Smithers:

O propósito da investigação é encontrar os mecanismos responsáveis por comportamentos particulares. Na neurociência, o mecanismo é descoberto olhando-se dentro do sistema. Em etologia, ele é descoberto pelo exame detalhado do comportamento e de como ele varia. Em ambos os campos, criar hipóteses de possíveis mecanismos desempenha uma parte vital no direcionamento das investigações [...]. E é aqui que a IA pode desempenhar papel central, ao estender a teoria, uma vez que sua metodologia provê meios de sugerir, tornar explícitas e explorar as consequências de hipóteses sobre mecanismos de comportamento. (WEB e SMITHERS, 1992, p. 422)⁴

⁴ Tradução nossa.

Em seu artigo, Web e Smithers (1992) apresentam uma proposta de estudo de fonotaxia de grilos a partir da construção de modelos artificiais. Os autores reconhecem que a contribuição da IA no entendimento de comportamentos, ao construir e testar modelos para verificar mecanismos hipotéticos, depende da verificação de que esses modelos sejam representações adequadas do componente biológico em estudo, e manifestam também a crença de que o entendimento de comportamentos mais simples possa levar ao entendimento de formas mais complexas de comportamento e sugerem que a Neuroetologia Computacional⁵ possa colaborar nesse sentido, através da simulação computacional de mecanismos sensório-motor de insetos.

A construção de robôs com inspiração em comportamentos animais (Robótica Baseada em Comportamentos) é também encontrada na implementação de sistemas de controle reativo de robôs, também conhecida como Abordagem Reativa (*Reactive Paradigm*), na qual procura-se estabelecer uma inter-relação estreita entre os sensores do robô e seus efetadores, permitindo reações imediatas do robô aos estímulos do ambiente, conforme Brooks (1986) e Murphy (2000).

A diferença entre a IA Baseada em Comportamentos e a *ALife* é de que, na primeira, não há a premissa de que sistemas ou aparatos criados a partir dessa abordagem estejam vivos, ou de que possuam as qualidades de algo vivo. A *ALife* é muito mais pretensiosa nesse sentido. Dessa forma, o encontro entre a IA e a *ALife* parece ainda estar por vir, e uma etapa intermediária nesse encontro, ao nosso ver, seriam as pesquisas em Robótica Evolucionária.

⁵ Apresentaremos a abordagem de Neuroetologia Computacional na seção 4.4.

4.3 ROBÓTICA EVOLUCIONÁRIA

O termo Robótica Evolucionária (*Evolutionary Robotics*) parece ter sido cunhado por Peter Ciarani, em seu trabalho intitulado *Why artificial life needs evolutionary robotics*, apresentado num workshop de *ALife* em 1987, mas não publicado, conforme considerações de Harvey et al. (2005). Ciarani estendeu o assunto em sua tese de doutorado (Ciarani, 1989) e posteriormente surgiram os primeiros livros específicos sobre essa nova abordagem, como os de Gomi (1998) e o de Nolfi e Floreano (2000). Também Brooks (1992) torna-se partidário dessa abordagem, ao reconhecer que a construção de robôs a partir da abordagem da IA Baseada em Comportamentos enfrentava algumas dificuldades de implementação e que as idéias de Langton (1989) e Langton et al. (1991) eram promissoras para a área. Até aquele momento, segundo Brooks, não havia relatos da utilização de programação evolucionária para robôs físicos, mas apenas para sua simulação computacional.

Os trabalhos em *ALife*, utilizando-se das técnicas de algoritmos genéticos, programação genética e programação evolucionária, trouxeram uma nova perspectiva de pesquisa para a área de robótica. Nolfi e outros pesquisadores⁶ aplicaram as idéias da *ALife* na abordagem de Robótica Evolucionária, na qual uma população inicial de robôs com diferentes “genótipos”, cada qual especificando um sistema de controle individual, é criada aleatoriamente. Cada robô é avaliado num ambiente de teste, e para cada indivíduo é atribuída uma pontuação que mede sua habilidade em executar a tarefa ou comportamento desejado. Somente os robôs que obtiverem maior pontuação terão permissão para se reproduzir. Tipicamente, o sistema de controle do robô é constituído de uma rede neural artificial, porém como advertem Nolfi e Parisi (1997), trata-se de uma rede neural artificial com características

⁶ Nolfi (1998), Nolfi e Floreano (2000), Bianco e Nolfi (2004).

especiais, que eles designaram de ALNN (*Artificial Life Neural Network*), que difere do modelo conexionista tradicional.

Nolfi (1998) define a Robótica Evolucionária como sendo a tentativa de se desenvolver robôs, e de seus sistemas de controle sensório-motor, por meio de processos automáticos de projeto utilizando a evolução artificial. A principal vantagem dessa abordagem, é que no projeto do sistema robótico não mais é necessário se dividir o comportamento objeto de estudo em unidades simples de comportamento, distribuídas em camadas diferentes, como era feita com o modelo de redes neurais hierárquicas. Partindo-se do comportamento desejado como um todo, os comportamentos simples podem emergir da interação entre vários processos internos ao sistema de controle, e da interação entre o robô e o ambiente. Do ponto de vista de sistemas naturais, Nolfi considera que a possibilidade de observação da interação do robô com o ambiente pode oferecer ajuda na tarefa de entender como os organismos naturais produzem comportamento adaptativo. Assim, partindo-se da abordagem evolucionária aplicada à robótica, pode-se tomar o comportamento adaptativo como um elemento particular de estudo. Para Beer (1997) as idéias-chave para tal estudo seriam, primeiro, o entendimento de que um agente e seu ambiente devem ser considerados como dois sistemas dinâmicos acoplados, cuja interação conjunta é responsável pelo comportamento do agente e, segundo, que a necessidade do agente manter sua existência em seu ambiente define uma restrição na dinâmica de seu comportamento. Para construção de modelos nesse estudo, Beer sugere a utilização de Redes Neurais Artificiais do tipo CTRNN (*Continuous-Time Recurrent Neural Network*)⁷, devido ao caráter dinâmico associado ao comportamento adaptativo. Tais redes apresentam como vantagens o fato de serem simples o suficiente para simular modelos de tamanho considerável com rapidez e de permitirem a análise detalhada de sua dinâmica, uma vez que são redes analiticamente mapeáveis.

⁷ Detalhes sobre o funcionamento de redes do tipo CTRNN podem ser encontrados em Kung (1993).

Dissemos, anteriormente, considerar a abordagem da Robótica Evolucionária como uma etapa intermediária no encontro entre IA e *ALife* pois, embora a Robótica Evolucionária também não lide diretamente com o contexto de entidades vivas, ou seja, de que robôs desenvolvidos através dessa abordagem possam ser considerados vivos, ou com as considerações sobre se a manifestação da cognição depende ou é consequência de uma entidade estar ou ser considerada viva, ela representa uma implementação física das idéias emprestadas da *ALife*, utilizando-se das técnicas de algoritmos genéticos e de programação genética. Na próxima seção, trataremos da última abordagem que selecionamos para investigação, conhecida como Mundos Virtuais.

4.4 MUNDOS VIRTUAIS E EMERGÊNCIA DA COGNIÇÃO

Já fizemos uma breve menção à abordagem de Mundo Virtual ao apresentarmos o sistema *Tierra* na seção 4.1. Trata-se, em princípio, da simulação computacional de ecossistemas povoados por agentes que possuem a capacidade de evoluir a partir da interação entre si e com o ambiente no qual se encontram. Uma vertente particular dessa abordagem vem se desenvolvendo e produzindo alguns resultados relacionados com a investigação da cognição. Trata-se da Etologia Sintética (*Synthetic Ethology*), assim definida por Bruce MacLennan:

O objetivo da etologia sintética é o de integrar as considerações mecanicistas e etológicas do comportamento pela combinação da simplicidade e controle de métodos behavioristas com a validade pragmática e ecológica da etologia empírica. A idéia da etologia sintética é simples: ao invés de estudarmos os animais num mundo natural desordenado e ao invés de retirarmos os animais de seus mundos completamente, nós criamos mundos artificiais e organismos simulados (*simorgs*) cujo comportamento está acoplado àqueles mundos. Uma vez que os organismos simulados são simples, nós podemos estudar fenômenos mentais em situações nas quais o mecanismo é *transparente*. (MACLENNAN, 1992, p. 638)⁸

Os agentes, ou simorgs (organismos simulados), são submetidos à processos de evolução, em interação com seus ambientes, e o resultado desses processos pode ser avaliado em termos de indivíduos, de grupos de indivíduos ou de populações. MacLennan (1992) considera que capacidades cognitivas possam ser investigadas através dessa abordagem, uma vez que tanto os organismos como o mundo em que tais organismos habitam são definidos de acordo com a necessidade da investigação em particular, podendo assumir configurações simples ou complexas, e que os mecanismos responsáveis pelos comportamentos são acessíveis ao investigador. O autor relaciona três pressupostos nos quais a Etologia Sintética está baseada. Primeiro, existe a convicção de que a cognição deve ser investigada em termos

⁸ Tradução nossa. Grifos do autor.

de comportamentos e dos mecanismos responsáveis pelo comportamento adaptativo evolutivo do agente em relação ao ambiente. Segundo, de que é possível relacionar o comportamento individual de agentes com o comportamento de populações. E, terceiro, de que a descoberta de leis científicas abrangentes requerem o tipo de controle de variáveis que só pode ser obtido em arranjos experimentais artificiais. O autor argumenta que, embora simplificados, os mundos virtuais assim definidos são completos, permitindo aos agentes neles existentes viver, evoluir e morrer, e que não há necessidade de modelar o mundo natural uma vez que tais mundos virtuais possuem as características essenciais do mundo natural. O propósito final é o de criar mundos virtuais que mantenham semelhanças relevantes com o mundo natural, mas que sejam simples o suficiente para se estudar.

MacLennan (1992) relata a semelhança da abordagem da Etologia Sintética com a abordagem da Neuroetologia Computacional (*Computational Neuroethology*), que dentre os trabalhos precursores na área estão os de Beer (1990) e Cliff (1995), conforme enfatizado por Keeley (2000).

Para Cliff (1995) a Neuroetologia Computacional pode ser definida como uma área de estudos composta pela intersecção entre as áreas da Etologia e da Neurociência, utilizando-se da abordagem computacional para realizar suas investigações. Cliff informa que a diferença entre a abordagem da Neuroetologia e a da Neuroetologia Computacional é a de que esta última é menos dependente do estudo específico de animais, já que aborda o estudo de agentes em geral. Animais são agentes autônomos biológicos. Outros agentes autônomos não-biológicos podem incluir robôs e agentes simulados em mundos virtuais, que podem operar em ambientes complexos e sujeitos a mudanças. Cliff também destaca que a Neuroetologia Computacional pode ser distinguida da Neurociência computacional⁹ tradicional pela ênfase

⁹ Os termos neurociência computacional e computação neural costumam ser utilizadas como equivalentes ao de redes neurais artificiais, que no contexto deste trabalho foram introduzidas a partir da abordagem do Conexionismo (capítulo 2).

no estudo do controle neural do comportamento, dentro do contexto de sistemas neurais incorporados e situados num ambiente. Ou seja, procura-se compreender os mecanismos neurais responsáveis por comportamento adaptativo do agente. Como comportamento adaptativo, Cliff considera aquele que aumente a probabilidade de sobrevivência do agente, de forma a permitir-lhe gerar descendência.

MacLennan (1992) destaca que a principal diferença entre as abordagens da Neuroetologia Computacional e da Etologia Sintética é de que a primeira tipicamente estuda a interação de um agente individual e seu ambiente, enquanto a segunda requer o estudo de grupos de agentes.

Em seu trabalho, MacLennan (2002) utilizou a abordagem da Etologia Sintética na investigação da emergência de comunicação entre agentes, considerando que a comunicação envolve tanto intencionalidade quanto significado compartilhado. A sua investigação consistiu em verificar se a comunicação poderia emergir entre agentes que evoluem num mundo virtual, a partir do contexto de cooperação entre esses agentes. Como resultado do experimento, o autor relata o surgimento da habilidade dos agentes se comunicarem através de pares de símbolos, formado uma sintaxe rudimentar, ainda que os resultados obtidos estivessem abaixo daqueles esperados por ele. Ainda assim, MacLennan considera que o resultado exibiu intencionalidade dos agentes e que os sinais trocados entre eles eram significativos para eles, os agentes, ainda que não necessariamente para os observadores, nós, comparando a ocorrência com situação equivalente na natureza, quando somos defrontados com certa comunicação entre animais, desprovida de significado para nós, mas para a qual devemos procurar o significado para os comunicadores. De qualquer forma, o autor considera que os experimentos em Etologia Sintética são ainda muito simples para exibir estados psicológicos e que um entendimento completo da linguagem, comunicação e das capacidades representacionais de estados mentais necessitam de uma teoria, ainda indisponível, que

relacione os mecanismos responsáveis pela cognição aos processos evolutivos. Uma concepção alternativa, baseada na teoria semiótica de Charles Peirce (1839-1914), tem sido proposta por Loula, Gudwin e Queiroz (2004) para realização de experimentos de comunicação simbólica entre criaturas artificiais, utilizando-se da abordagem da Etologia Sintética.

Em parte, as pesquisas em Etologia Sintética e *ALife* parecem convergir para um ponto em comum, que é o da possibilidade de criação de entidades artificiais vivas. MacLennan (2002) admite que os agentes em suas experiências de mundos virtuais possam ser considerados como entidades vivas, ainda que não estabeleça como requisito tal propriedade em seus experimentos e nem faça reivindicação de tal característica em seus resultados.

4.5 CONSIDERAÇÕES SOBRE O CAPÍTULO

Temos acompanhado desde o capítulo 1 um conjunto de abordagens, diferentes ou complementares entre si, adotadas nas pesquisas e estudos que pretendem lançar esclarecimentos sobre os processos cognitivos, a partir de sua simulação ou da tentativa de sua reprodução por meios artificiais. Cada uma dessas abordagens tem sido questionadas em seus pontos fracos, ao longo do tempo, por filósofos e por pesquisadores, a exemplo das discussões apresentadas nos capítulos 2 e 3, relacionadas à Inteligência Artificial e ao Conexionismo respectivamente. Nesse capítulo, apresentamos a origem e algumas das características da abordagem de Vida Artificial, e também algumas das abordagens relacionadas com implementações de robótica. A abordagem da robótica tem sido considerada promissora para as pesquisas relacionadas com a concepção de Cognição Incorporada e Situada, na qual um agente interage com o ambiente de maneira autônoma. No entanto, Ziemke e Sharkey (2000, 2001) questionam até que ponto esses agentes são realmente autônomos. Os autores argumentam que, a despeito de sua inspiração biológica, os agentes artificiais são radicalmente diferentes dos agentes biológicos e, mesmo apresentando um certo grau de auto-organização, esses agentes estão muito longe de possuírem as características de autonomia dos organismos vivos. Por sua vez, pesquisadores em *ALife* afirmam ser possível a realização de vida artificial, sintética ou virtual, e que as atuais abordagens virtuais de *ALife* já são capazes de instanciar entidades vivas. A seguir, apresentamos um resumo contendo aqueles que acreditamos serem os principais programas de pesquisas e as principais abordagens relacionados aos estudos de cognição, discutidos até aqui, destacando suas áreas de inspiração, metáforas adotadas e tipos de implementação, conforme nosso entendimento, e em nossas considerações finais discutiremos se implementações de Vida Artificial poderão, de fato, nos auxiliar no entendimento das capacidades cognitivas humanas.

ABORDAGEM	INSPIRAÇÃO	METÁFORA	IMPLEMENTAÇÃO	PROGRAMA DE PESQUISAS
IA Simbólica	Psicologia	Mente	Computacional	Inteligência Artificial Clássica
	Exemplo: Programa de Inferências de Robert Schank (Schank, 1977)			
IA Conexionista	Neurociência	Cérebro	Computacional	Conexionismo
	Exemplo: Rede Neural Perceptron de Rosenblatt (Minsky e Papert, 1969)			
IA Baseada em Comportamentos	Etologia	Comportamentos Animais	Robótica	Psicologia Sintética
	Exemplo: Robôs de Rodney Brooks (Brooks, 1999)			
ALife Implem. Física	Biologia	Evolução	Robótica	Robótica Evolucionária
	Exemplo: Robôs de Nolfi e Floreano (Nolfi e Floreano, 2000)			
ALife Virtual Fraca	Biologia	Evolução	Computacional	Biologia Sintética
	Exemplo: Programa <i>Boids</i> de Craig Reynolds (Reynolds, 1987)			
ALife Virtual Forte	Ecosistemas	Mundos Virtuais	Computacional	Etologia Sintética
	Exemplo: Programa <i>Tierra</i> de Thomas Ray (Ray, 1995)			

Tabela 2- Relação das principais abordagens e programas de pesquisas, considerados neste trabalho, relacionados aos estudos da cognição.

CONSIDERAÇÕES FINAIS

Neste trabalho, tivemos como nosso objetivo investigar algumas das possibilidades provenientes da adoção da abordagem de Vida Artificial em estudos de Ciência Cognitiva, mais especificamente, no entendimento da cognição humana. Se a cognição puder ser considerada como uma característica emergente de um organismo vivo, nos perguntamos se seria possível considerarmos uma abordagem conjunta entre Ciência Cognitiva e Vida Artificial no estudo da cognição. A partir das pesquisas discutidas no capítulo 4, a resposta, em princípio, parece ser sim, e abordagens envolvendo conceitos de ambas as áreas vem sendo tentadas. Boden (2001, p. 11) considera que “Se a inteligência requer vida, então não só a Ciência Cognitiva é uma subclasse das ciências da vida, como a IA é uma sub-área da *ALife*”.

Porém, a abordagem da *ALife* traz consigo diversos questionamentos que merecem discussões do ponto de vista filosófico. O filósofo Mark Bedau, por exemplo, enumera uma lista de quatorze questões relacionadas à área (Bedau, 1992), como uma proposta para discussões. Uma dessas discussões é a de que, ao considerar que os processos que constituem a vida devem compartilhar características universais que permitam identificar sua presença, independente do substrato material em que se encontra, a *ALife* nos leva à seguinte questão: quais seriam esses processos que constituem a vida? Ou ainda, mais sucintamente, o que é vida?

Para Rosen (1991), a necessidade de se explicar o que é vida surgiu com o desenvolvimento da Mecânica, quando Newton demonstrou que o mistério das estrelas e planetas encontrava-se num conjunto de poucas regras simples, das quais a vida não fazia parte. Questionando-se sobre para quem deveríamos dirigir a tarefa de explicar o que é vida, Rosen sugere dirigi-la à própria Física, uma vez que seres vivos são compostos de matéria e,

dessa forma, a matéria deveria conter o segredo da vida. Em 1944, Schrödinger (1997) lançou mão da Física estatística e dos conceitos de Termodinâmica em sua abordagem sobre o fenômeno vida, tratando de temas como ordem em sistemas biológicos e mecanismo hereditário.

Mas tentar estabelecer um critério para definir o que se entende por vida não é tarefa fácil. Talvez, uma possibilidade seria a de investigar a origem da vida, para então tentar defini-la. O problema da origem da vida é freqüentemente discutido à luz da Teoria da Evolução, proposta por Charles Darwin (1809-1882), na qual todos os seres vivos descendem de uma forma primordial de vida e na hipótese de que nos oceanos primitivos de nosso planeta formou-se algo semelhante a um “caldo primordial”, no qual a aglomeração de moléculas orgânicas deu origem aos primeiros organismos vivos. Oparin (1938) sugeriu que a vida na terra pode ter surgido de processos químicos, a partir dos elementos que estariam disponíveis na atmosfera primordial, dentre eles a amônia, o metano, o hidrogênio e o vapor d’água. Gleisser (2003) informa que em 1953 o cientista americano Stanley Miller divulgou os resultados de uma experiência que iria causar grande impacto na comunidade científica¹. Miller construiu um sistema fechado composto por vapor d’água, amônia, hidrogênio e gás metano, exatamente os elementos que se acreditava existirem na atmosfera da Terra quando do surgimento da vida em nosso planeta, e nessa mistura aplicou descargas elétricas, simulando a ação de relâmpagos. Após uma semana, Miller verificou o aparecimento de aminoácidos, que são os responsáveis pela formação de proteínas, e também de carbono, o elemento principal da vida em nosso planeta. Outros pesquisadores repetiram os experimentos de Miller com sucesso.

¹ Para efeito de nossas discussões, consideramos suficiente o resumo oferecido por Gleisser (2003). Uma apresentação mais detalhada sobre a experiência de Stanley Miller pode ser encontrada em Miller (1953).

Conforme Miller (1992), parecia possível que tais processos dessem origem a aminoácidos, que reunidos formariam proteínas capazes de catalisar reações químicas, as quais, por sua vez, dariam origem às células. Essas células, ao se reproduzirem e se combinarem, dariam origem aos organismos vivos. Em diversos outros experimentos, quase todos os vinte tipos de aminoácidos que ocorrem naturalmente nos seres vivos foram encontrados. Passada uma década após a realização da experiência de Miller, cientistas utilizando métodos semelhantes conseguiram sintetizar as bases nitrogenadas (ou nucleotídeos) da molécula de DNA. Para Gleisser (2003), os seres vivos são, em sua essência, um conjunto de macromoléculas orgânicas de grande complexidade, capazes de realizar uma série de transformações químicas que levam à sua subsistência e à sua reprodução. O pesquisador Dawkins (1979) defende a idéia de que as macromoléculas iniciais do mundo pré-biótico foram submetidas à processos de seleção natural, evoluindo até chegar aos organismos biológicos. Para que essas macromoléculas primordiais evoluíssem até as formas de vida atualmente conhecidas, Kauffman (1997) considera a importância da auto-organização:

Tal auto-organização, desde a origem da vida até a sua dinâmica coerente, deve ter um papel essencial na história da vida e, na verdade, eu argumentaria, em qualquer história de vida. Mas Darwin também estava certo. A seleção natural encontra-se sempre em ação. Portanto, precisamos repensar a teoria evolucionária. A história natural da vida é algum tipo de casamento entre auto-organização e seleção. [...] Pelo menos para a vida baseada no carbono, ligações com força suficiente são necessárias para ter estabilidade em um determinado ambiente. Mas são os comportamentos de grupos destas moléculas que constituem a vida na Terra e, pelo menos podemos presumir, são subjacentes a muitas formas de vida em potencial em qualquer lugar do Universo. Os seres vivos constituem, na verdade, sistemas moleculares coletivamente e autocatalíticos. (KAUFFMAN, 1997, p. 132-133)

Assim, além da combinação de macromoléculas submetidas a processos evolutivos, seriam necessários outros fatores para a composição da vida, dentre eles a auto-organização. Para Boden (2001), a auto-organização assume papel central como característica da vida,

envolvendo a emergência e manutenção da ordem. Para Schneider e Kay (1997, p. 187) “Os sistemas vivos mostram uma marcha que se afasta da desordem e do equilíbrio em direção a estruturas altamente organizadas, as quais existem a uma certa distância do equilíbrio”, implicando num não-equilíbrio termodinâmico como característica dos sistemas vivos.

A experiência de Miller comprovou a possibilidade da criação de macromoléculas a partir de moléculas simples, mas não existe consenso sobre a origem da vida a partir dessa teoria. Tamayo e Hartman (1989) consideram que a composição da atmosfera primordial sugerida por Oparin (1938) é altamente improvável, já que evidências provenientes da análise da composição de rochas antigas indicam uma atmosfera composta por dióxido de carbono e nitrogênio. Os autores apontam também que, em oposição à teoria da formação da vida a partir de proteínas (*protein life*), uma segunda hipótese, encontrada em Gilbert (1986) por exemplo, teoriza que o ácido ribonucleico veio primeiro (*RNA life*). Tal diversidade de fatores, e de hipóteses associadas, dificulta a tarefa de definir o que é vida a partir do conhecimento de como ela se forma e parece não permitir uma única e exclusiva definição do que é vida.

Uma abordagem alternativa, poderia ser a de tentar estabelecer o que é vida a partir das propriedades apresentadas pelos organismos vivos. Mayr (1982) relaciona, como algumas das características especiais dos organismos vivos, os seguintes atributos:

- **Complexidade e organização:** Sistemas vivos possuem organização adaptativa e de grande complexidade.
- **Unicidade de composição química:** Organismos vivos são compostos por um conjunto quimicamente específico de macromoléculas.
- **Natureza qualitativa:** Os fenômenos de importância em sistemas vivos são de natureza qualitativa, ao invés de quantitativa.

- **Variabilidade e unicidade:** Os fenômenos biológicos caracterizam-se por alto grau de variabilidade de grupos formado por indivíduos únicos.
- **Programação genética:** Todos os organismos possuem um programa genético, historicamente aperfeiçoado.
- **Natureza histórica:** Organismos vivos possuem conexão histórica com descendentes.
- **Seleção natural:** Organismos vivos são resultado de seleção natural.
- **Indeterminismo:** Processos biológicos não permitem predição.

O filósofo Lange (1996) relaciona, como sinais de vida comumente mencionados pelos biólogos, os atributos de: possuir homeostase, possuir metabolismo, mover-se, crescer, reproduzir, responder ao ambiente e evoluir. Para o autor, no entanto, o problema dessas listas de atributos é que algumas coisas não-vivas possuem um ou mais desses atributos e algumas entidades vivas não possuem todos os atributos.

Tomando, por exemplo, o primeiro item da lista de Mayr (1982), o próprio autor reconhece que a complexidade não é uma diferença fundamental entre entidades vivas e não-vivas, já que existem sistemas orgânicos muito simples, como no caso de algumas macromoléculas, e sistemas inorgânicos muito complexos, como no caso de uma galáxia. Lange (1996) apresenta os exemplos de que os planetas se movem e cristais podem crescer, sem serem coisas vivas, e que determinados tipos de animais não possuem a capacidade de se reproduzir, como no caso das mulas. Lange diz que frequentemente os biólogos não conseguem definir exaustivamente o que é vida, apesar do senso comum do termo ser bem conhecido. Para Mayr (1982), isso se deve ao fato de não existir uma substância ou força específica que possam ser identificadas com a vida. Lange argumenta que essa dificuldade dos biólogos pode estar relacionada com a dificuldade de elaborar uma definição redutiva do termo, sem a qual, qualquer outra definição poderia compromete-los com o Vitalismo, que ao

lado Mecanicismo e do Organicismo, constitui uma das três principais vertentes filosóficas sobre a natureza da vida.

Na Biologia o enfoque tradicional tem sido o do Mecanicismo, no qual, segundo Mayr (1982), os organismos vivos seriam considerados máquinas que desempenham funções físico-químicas, e, portanto, poderiam ser explicados pelas leis da mecânica, da física e da química. Assim, todo fenômeno da vida poderia ser explicado por tais leis. Para Mayr (1982), trata-se, em essência, de uma visão reducionista, sustentando o princípio de que os fenômenos biológicos podem ser reduzidos a determinadas leis já bem estabelecidas pelas ciências física e química. Mayr (1982) justifica que a ciência, de um modo geral, adere à teoria do Mecanicismo pelo fato dessa oferecer respostas satisfatórias a inúmeras perguntas, e por ser uma teoria fortemente estabelecida. Muito já se investiu nela, de maneira que suas raízes se tornaram profundas em nossa sociedade. Para a biologia experimental trata-se, talvez, do único método eficaz disponível, pois é ele que tem permitido explicar o comportamento dos seres vivos através do estudo de suas partes constituintes e de sua redução posterior às leis da física e da química. Mayr argumenta que os biólogos são totalmente materialistas, no sentido de que não reconhecem quaisquer forças não-materiais, mas tão somente forças físico-químicas.

A vertente Organicista, que é sustentada por nomes como os de Russell (1916) e de Bertalanffy (1975), questiona duramente a visão mecanicista. Emmeche (2004) destaca a posição não-mecanicista dessa abordagem, para a qual os sistemas vivos não podem ser descritas apenas pelas leis e princípios da mecânica, seja ela clássica ou quântica. O Organicismo recusa também o princípio de que os fenômenos da natureza possam ser reduzidos exclusivamente às leis físico-químicas, uma vez que tais leis não conseguem explicar a totalidade do fenômeno da vida. Adicionalmente, reconhece a existência de sistemas organizados hierarquicamente, com propriedades que não podem ser entendidas

através do estudo de suas partes isoladas, mas sim através de sua totalidade e da interdependência das partes (o todo não é necessariamente a soma das partes). Conforme Mayr (1982), os biólogos organicistas destacam o fato de que sistemas orgânicos são dotados de diversas características para as quais não existem equivalentes no mundo dos objetos inanimados.

Já o Vitalismo propõe a idéia de que toda forma de vida possui um tipo de componente intrínseco e imprescindível, não passível de ser medido, que sustenta a vida. Os defensores do Vitalismo argumentam que tal componente se apresenta claramente evidente em vários dos aspectos do desenvolvimento do organismo, como os de regulação, regeneração e reprodução. O Vitalismo foi fortemente criticado, por possuir um aspecto não-científico, não se enquadrando em critérios para estabelecer o caráter científico de uma teoria, como o do falsificacionismo ou refutabilidade definidos pelo filósofo Karl Popper, e discutidas por Chalmers (1999), uma vez que o “componente vital” era um fator causal incerto e sem possibilidade de demonstração. Também Nagel (1951), filósofo da ciência, não poupou críticas à teoria do Vitalismo, uma vez que ela não apresentava um guia confiável para a investigação biológica, sendo incapaz de fornecer previsões demonstráveis. Mayr (1982) diz que, por volta de 1930, os biólogos já rejeitavam de forma praticamente universal o Vitalismo.

Mas Lange (1996) acredita que algum tipo de vitalidade seria útil para explicar os sinais de vida apresentados por uma entidade. Para ele, os biólogos não conseguem definir o que é vida na medida em que eles não conseguem definir o que não é vida (não-vida). Eles não conseguem reduzir a vitalidade como um conjunto de outras propriedades independentes e inteligíveis. Essa incapacidade, como sugere Lange, não significa que o conceito de vida não desempenhe nenhum papel na biologia, argumentando que filósofos da natureza classificavam a água como algo do tipo natural, e sabiam reconhecê-la, muito antes deles

saberem que a composição química da água é H₂O e, presumidamente, antes de existir qualquer propósito para uma definição reducionista da água. Além disso, continua Lange, uma das funções da própria biologia poderia ser a de entender a vitalidade. Para ele, a questão sobre o que é necessário para que uma entidade seja considerada viva tem sido negligenciada pelos filósofos, de um modo geral, permanecendo ausente nos textos de Filosofia da Biologia, e diz ainda não haver evidências de que exista qualquer função na Biologia para a distinção entre vida e não-vida.

No entanto, a questão da distinção entre vida e não-vida se tornou importante para aqueles cientistas interessados na abordagem da *ALife* e que acreditam ser possível criar e estudar organismos digitais, cada qual constituído de códigos de programas ou softwares vivendo na memória do computador (*ALife Virtual*), como o biólogo Thomas Ray:

Eu gostaria de sugerir que a criação de software em *ALife* pode ser dividida em dois tipos: simulações e instanciação de processos de vida. Nas simulações, estruturas de dados são criadas contendo variáveis que representam os estados das entidades sendo modeladas. Na instanciação, os dados no computador não representam à nenhuma outra coisa. Os padrões de dados numa instanciação são considerados formas vivas, em sua própria maneira, e não modelos de qualquer forma natural de vida.² (RAY, 1995, p. 180)

Esses programas podem executar cópias de seus códigos, acrescidos com novos elementos, permitindo a ocorrência de variação e seleção. Organismos digitais podem competir pelo acesso a recursos limitados de memória e de processamento, e algoritmos evolutivos permitiriam gerar programas mais eficientes no uso daqueles recursos. Mas qualquer argumento de que certos trechos de códigos auto-replicantes em computadores estão vivos pressupõe que alguma noção de vida exista, e que seja possível a distinção entre vida e não-vida. Tal distinção poderia ser entendida em termos de sinais de vida existentes que pudessem ser observados na entidade em estudo. Lange discute a posição de Sober (1991),

² Tradução nossa.

que considera que muitos processos e propriedades biológicas envolvem relações entre os organismos e alguma coisa do lado externo deles. Um animal reproduz quando dá cria; uma planta realiza fotossíntese quando recebe luz em condições adequadas; um predador come outros animais; etc., constituindo assim sinais de vida daquele organismo. Apesar de que um programa de computador possa replicar certos aspectos de tais processos, ele não pode, de fato, realizar a fotossíntese tal como os organismos biológicos. Assim, para Sober, mesmo que um programa de computador apresentasse a mesma estrutura matemática de um dos sinais de vida, seria impossível que tal programa possuísse o conteúdo empírico da propriedade biológica equivalente. Entretanto, Lange argumenta que um programa de computador pode se relacionar com várias coisas do lado externo dele (a memória, a CPU, outros programas, etc.) que são isomórficas às várias relações que se estabelecem entre organismos biológicos e coisas externas. É claro, lembra Lange, que certas relações exigem a coisa certa do lado externo. Beber água, por exemplo, pressupõem a existência de água, enquanto comer pressupõe a existência de comida. Embora comer e beber não sejam comumente incluídos nas relações de sinais de vida, pressupõem vitalidade, uma vez que a total ausência de sinal de vida pressupõe falta de vitalidade. Os sinais de vida não são individualmente necessários ou conjuntamente suficientes para que alguma entidade possa ser considerada viva. Por isso, para Lange, o conceito de vitalidade se torna útil nessa abordagem, afirmando ter percebido que a utilização da expressão “está vivo” para explicar a razão pela qual certas entidades apresentam sinais de vida assemelha-se a utilizar a expressão “é de cobre” para explicar porque um objeto conduz eletricidade. Nenhuma das expressões descreve as propriedades do objeto, sua estrutura ou características que expliquem o fenômeno. Mas ambas as expressões eram consideradas explanatórias antes que informações relevantes sobre as estruturas fossem conhecidas. Então se, de fato, essas expressões podem ser consideradas explanatórias, Lange sugere que a expressão “está vivo” pode ser útil para identificarmos quando uma entidade

exibe sinais de vida. Ele justifica sua sugestão pelo fato de que esse tipo de explanação nos informa que o fenômeno que se observa resulta inteiramente da natureza do objeto ou entidade. Lange acredita que alguns cientistas contam com a hipótese de que duas entidades vivas exibem seus sinais de vida por importantes razões similares. Assim, “está vivo” exerce em biologia papel similar a “é de cobre” em física.

Lange conclui dizendo que os pesquisadores de *ALife* não estão obrigados a apresentar uma definição de vida para justificar que certas entidades de software estão vivas. Porém, eles devem reconhecer que seus argumentos de vitalidade para alguns organismos artificiais não satisfazem a muitos dos atributos de sinais de vida. Apesar de que certas capacidades dos seres vivos (como a fotossíntese, por exemplo) não possam ser exibidas por entidades de software, nenhum sinal particular de vida é necessário para a vitalidade. A relação entre sinais de vida e vitalidade permanece obscura, já que tais sinais não são individualmente necessários ou conjuntamente suficientes para a vitalidade. Lange propõe reconstruir essa relação: coisas vivas constituem um tipo natural e a vitalidade explica porque uma entidade apresenta vários sinais de vida. Lange não vê, em princípio, razão para acreditar que seja impossível para um programa de computador apresentar sinais de vida, em razão de estar vivo. Dessa proposição de Lange, poderíamos depreender que tais programas fariam parte da classe natural das coisas que apresentam sinais de vida. Uma vez que parece não existir, ou pelo menos não está plenamente estabelecida, uma distinção inequívoca entre vida e não-vida, implementações de *ALife*, em qualquer uma de suas modalidades, poderiam vir a apresentar sinais de vida equivalentes àqueles apresentados pelos organismos biológicos, em decorrência de estarem vivas.

Porém, essa recorrência à vitalidade, proposta por Lange, encontra críticas em outras abordagens. Harnard (1994) considera que a interpretação vitalista na *ALife* Virtual está relacionada com a interpretação anímica da mente, provavelmente baseada numa (errônea)

idéia de que coisas vivas possuem vida mental. O autor considera que a *ALife* Virtual possa ser utilizada na interpretação e na explicação de coisas vivas, mas não em sua instanciação. Já as implementações físicas da *ALife* poderiam, em sua opinião, produzir um sistema artificialmente vivo, se o mesmo estiver situado e em interação causal com o mundo real (como no caso de um robô), mesmo diferindo radicalmente dos sistemas naturais vivos em suas propriedades e matéria.

Para Boden (2001), a pergunta a se fazer então é: se um artefato possui vida, ele possui cognição? Refletindo sobre esta questão, consideramos a idéia de Stewart (1992) que propõem que “Vida = Cognição”. O autor argumenta, referindo-se às idéias de Kant em seu trabalho “A Crítica da Razão Pura”, que o problema com o objetivismo, abordagem segundo a qual o mundo real existe e que ele é o que é, independentemente do conhecimento que um sujeito pode ter ou não dele, é a de que a realidade ontológica é inerentemente incognoscível, e que a alternativa do niilismo, abordagem segundo a qual nada existe, ou de que então a realidade é completamente caótica e arbitrária, tal que tudo e nada são igualmente possíveis, pouco contribui, tornando-se necessário algum “princípio de realidade”. Assim, inspirado nos trabalhos de Humberto Maturana (Maturana e Varela 1980, 1987) e Francisco Varela (Varela, 1980), Stewart propõem que o conhecimento está materialmente incorporado em processos dinâmicos e que organismos vivos são uma subclasse de sistemas dinâmicos autônomos, na qual o sistema deve possuir a capacidade de continuamente produzir a si próprio, tanto materialmente como organizacionalmente. Para o autor, sujeito e objeto do conhecimento são inseparáveis, implicando em que não existe outro objeto de conhecimento senão o da própria existência dinâmica do sujeito. Ou seja, conhecimento é “conhecer como permanecer vivo”. O autor, admitindo a hipótese de que a condição para entender algo é construí-lo, considera que projetos que propõem a criação de vida artificial serão instrutivos em fornecer informações

sobre o que de fato está envolvido na organização do “vivo”, ainda que o objetivo final não seja atingido.

Mas um conhecimento do tipo “conhecer como permanecer vivo” seria necessário e suficiente para explicar capacidades cognitivas do tipo linguagem, raciocínio e criatividade? Ou, se for apenas necessário, a cognição implicaria em algum tipo de mente, uma vez que tais capacidades são comumente relacionadas à mente humana? Em caso afirmativo, e considerando que vida implica em cognição, então vida implica em mente? A relação entre vida e mente é outra das propostas de discussão encontrada de lista de Bedau (1992), que questiona se vida e mente não seriam diferentes estágios de uma mesma hierarquia. Consideremos então reformular a pergunta anterior feita por Boden (2001) e questionar: se um artefato possui vida, ele possui mente? E, em caso afirmativo, que tipo de mente seria essa? Recorreremos às considerações do filósofo Daniel Dennett, na investigação dessas questões.

Que tipos de mente existem? E como sabemos sobre elas? Ao estabelecer estas perguntas, Dennett (1997) procura distinguir os questionamentos ontológicos (o que existe), caso da primeira pergunta, dos questionamentos epistemológicos (sobre o nosso conhecimento), caso da segunda pergunta. Ele nos alerta para sermos cuidadosos quando estipulamos limites ao considerar a existência de algo, baseando-se naquilo que nos é possível conhecer sobre esse algo, uma vez que sua existência (ontológica) independe do nosso conhecimento sobre ele (epistemológico). Dennett sugere que nós evoluímos a partir de seres que possuíam mentes mais simples. Tais seres teriam evoluído de seres com mentes ainda mais simples, ou protomentes. E assim sucessivamente, até retornarmos às estruturas mais simples possíveis, destituídas de qualquer tipo de mente. A partir de macromoléculas auto-replicas, teriam surgido os organismos simples unicelulares, que dariam origem aos organismos multicelulares que, por sua vez, evoluiriam em organismos marinhos, peixes,

répteis, até chegarem aos mamíferos. O filósofo introduz o conceito de intencionalidade ao discorrer sobre o nascimento da ação nas primeiras macromoléculas complexas o suficiente para tal, e daí para as células e para conjunto de células que realizam ações especializadas nos organismos. Dennett considera todas essas entidades, das mais simples às mais complexas, como sistemas intencionais, e adota a perspectiva na qual as ações de tais sistemas intencionais se tornam reconhecíveis através da postura intencional. A postura intencional é a estratégia sugerida por Dennett para interpretar o comportamento de uma entidade qualquer, que pode ser uma pessoa, um animal, um artefato, ou qualquer outra coisa, considerando-a como um agente capaz de selecionar suas ações. Ou seja, existe a direcionalidade da ação. Os produtos da seleção natural seriam resultantes das ações de agentes com direcionalidade, e para Dennett as macromoléculas auto-replicas originais tinham intencionalidade, pelo fato de possuírem direcionalidade de ação. Essas macromoléculas, e o produto de suas evoluções posteriores, são consideradas robôs por Dennett, devido às suas funções altamente especializadas. Para ele, todos nós somos constituídos por robôs, fazendo referência aos componentes dos sistemas metabólico, imunológico e aos outros sistemas que desempenham funções auto-reguladoras e de manutenção da saúde do nosso corpo.

Dennett supõe ainda como válida a atribuição de racionalidade a um agente, considerando que o agente racional buscará sempre o que for bom para ele, e que esse tipo de comportamento, o de buscar o que é bom para si, também está presente nos produtos resultantes da seleção natural e, dessa maneira, estabelece um critério para distinguir as entidades que não têm mente, citando o exemplo da rocha que não é capaz de buscar o que é bom para si, uma vez que não foi equipada para atingir qualquer tipo de bem individual. Em outras palavras, a racionalidade do agente, manifestada pela sua intencionalidade, seria o indicativo da existência de sua mente. Assim, um robô construído pelo homem poderia apresentar um tipo de “mente”, se possuir intencionalidade, ainda que seja derivada de uma

pré-intencionalidade original implementada por seus construtores. Para Dennett, portanto, um artefato pode possuir intencionalidade e, conseqüentemente, possuir mente.

Essa posição não é suportada por Searle (1997), que considera insuperável a dificuldade em se atribuir intencionalidade consciente a um robô: o robô não teria consciência da razão ou motivação de suas ações, da mesma forma que as macromoléculas não tinham consciência de suas ações, tivessem elas direcionalidade ou não. Enquanto Bedau (1992) defende a idéia de que a vida não necessita envolver a subjetividade e a consciência, considerando a abordagem de *ALife* menos vulnerável às críticas semelhantes às aquelas do tipo quarto chinês (Searle, 1990) feitas à IA, Boden (2001) considera a discussão vida-cognição “dura” o suficiente, prescindindo das considerações sobre intencionalidade e consciência, as quais, para a autora, são tão problemáticas que devem permanecer em aberto. Se assim procedermos, desconsiderando a intencionalidade, os argumentos de Dennett (1997) não são aplicáveis neste contexto e, portanto, permanecemos sem resposta para a questão da possibilidade de existência de mente em um artefato. E assim sendo, também o objetivo de nossa investigação, o de saber se pesquisas em *ALife*, através da criação de artefatos ou sistemas artificialmente vivos, poderiam contribuir no estudo e entendimento da cognição humana, também permanece em aberto.

Uma outra questão a ser considerada é que a própria suposição de que a abordagem de *ALife* poderia realmente implementar a criação de artefatos ou sistemas artificialmente vivos não está completamente estabelecida. Emmeche (1992), em seu artigo “Vida artificial é possível?”, considera que há razões para acreditar que o fenômeno vida não é independente do meio, tal como pressupõe a abordagem da *ALife*, mas apresenta uma interdependência de matéria e forma. Para o autor, o fenômeno vida poderia até se manifestar em outros meios além daqueles baseados em cadeias de carbono, mas como resultado de um longo e natural processo evolucionário. Pelo que sabemos, apesar dos bilhões de anos de existência de nosso

planeta, e do processo de evolução natural continuamente agindo sobre ele, não foi possível identificar nenhuma forma de vida não baseada em cadeias de carbono, e, particularmente, que também manifestasse cognição e, mais particularmente ainda, que manifestasse cognição semelhante à cognição humana. Assim, esperar pela “evolução” de algum robô inicialmente construído por nós, até que ele manifestasse uma mente capaz da cognição humana está fora de nossa escala temporal: nenhum pesquisador, ou programa de pesquisas, sobreviveria aos incontáveis anos supostamente necessários para tal evolução (se é que ela seria possível).

Emmeche (1992) considera ainda que a idéia de independência do meio presente na *ALife* remete a um tipo de Funcionalismo semelhante ao existente em IA, que considera a mente independente do meio. Se, conforme o autor, esse Funcionalismo possui origens na Psicologia, pois teria surgindo como uma reação ao behaviorismo, é interessante observar a analogia feita por Sober (1991) de que a IA está para a Psicologia, assim como a *ALife* está para a Biologia, e lembrar da consideração feita por Boden (2001) de que se a inteligência requer vida, então a IA é uma sub-área da *ALife*.

Nesse caso, deveríamos entender a Psicologia estritamente como uma sub-área da Biologia e recorrer à algum tipo de biologia da mente, para o entendimento da cognição humana? Será que alguma abordagem reducionista, que relacionasse estados mentais com estados neurobiológicos, seria suficiente para explicar todo o repertório de possibilidades existentes na cognição humana? Se houver relação direta entre tais estados, seria possível acomodar nessa explicação as influências socioculturais, como aquelas sugeridas por Vygotsky (Oliveira, 1993; Del-Masso, 1998) por exemplo?

E, se todas essas perguntas pudessem ser respondidas afirmativamente, seria de alguma utilidade a construção de um artefato que reproduzisse ou instanciasse tais estados? Por fim, considerando a abordagem da *ALife*, será que ao invés de irmos em direção às respostas procuradas, não estaríamos indo, na verdade, em direção à criação de outras formas

de vida (“a vida como ela poderia ser”, como sugere Langton (1989)), e, portanto, de outros tipos de mente?

Nossa conclusão, restrita ao âmbito dos trabalhos utilizados como referência nesta dissertação, é de que as pesquisas fundamentadas em tecnologias e abordagens bio-inspiradas, como aquelas desenvolvidas nas áreas de Inteligência Artificial Baseada em Comportamentos, em Robótica Evolucionária e em Vida Artificial, podem ser consideradas ferramentas importantes no auxílio ao desenvolvimento tecnológico e científico e aplicadas às áreas de biologia teórica, computação, engenharia, e em outras áreas e aplicações afins, tais como aquelas mencionadas em Levy (1992), Emmeche (1994), Bentley (2001) e Castro e Von Zuben (2005). No entanto, tais pesquisas encontram-se num estágio de desenvolvimento atual tal que não há ainda evidências de que um dia será possível construir um artefato, software ou robô capaz de manifestar as mesmas capacidades cognitivas humanas, de forma a servir de “modelo” para o entendimento dessas capacidades, e nem de que tal artefato, caso seja possível construí-lo, terá, de fato, alguma utilidade para esses estudos. E, sendo assim, talvez o melhor “modelo” para se estudar a cognição humana continue sendo o próprio ser humano.

REFERÊNCIAS

- ABRANTES, P. (Org.). *Epistemologia e Cognição*. Brasília: UNB, 1993.
- ANDERSON, M. *Embodied Cognition: A Field Guide*. Artificial Intelligence, 149, p.91-130, 2003.
- ALLEN, C.; BEKOFF, M.; BURGHARDT, G. (Eds.) *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*. Cambridge: MIT Press, 2001.
- ARANHA, M.; MARTINS, M. *Filosofando – Introdução à Filosofia*. São Paulo: Moderna, 1986.
- ASHBY, W. Principles of the Self-Organizing System. In: VON FOESTER, H.; ZOPF, G. (Eds.). *Principles of self-organization*. London: Pergamon, p.225-278, 1962.
- ASHBY, W. R. *Uma Introdução à Cibernética*. São Paulo: Perspectiva, 1970.
- BANZHAF, W. et al. *Genetic Programming: On the automatic evolution of Computer Programs and its Applications*. San Francisco: Morgan Kaufmann, 1998.
- BEAR, M.; CONNORS, B.; PARADISO, M. *Neurociências: Desvendando o Sistema Nervoso*. Porto Alegre: Artmed, 2002.
- BECHTEL, W. *A Bridge Between Cognitive Science and Neuroscience: The Functional Architecture of Mind*. Philosophical Studies 44, p. 319-330, 1983.
- BECHTEL, W. *Philosophy of Science: An Overview for Cognitive Science*. Hillsdale: Lawrence Erlbaum Associates, 1988.
- BEDAU, M. Measurement of Evolutionary Activity, Teleology and Life. LANGTON, C. et al (Eds.). *Artificial Life II. The Second Workshop on the Synthesis and Simulation of Living Systems*. Redwood City: Addison-Wesley, p. 431-461, 1992.
- BEDAU, M. The Nature of Life. In: BODEN, M. (Ed.). *The Philosophy of Artificial Life*. New York: Oxford University Press, p. 332-357, 1996.
- BEDAU, M. The Philosophical Aspects of Artificial Life. In: VARELA, F.; BOURGINE, P. (Eds.). *Toward a practice of autonomous systems: Proceedings of the First European Conference on Artificial Life*. Cambridge: MIT Press, p. 494-503, 1992.
- BEDAU, M. *Emergent Models of Supple Dynamics in Life and Mind*. Brain and Cognition 34, p. 5-27, 1997a.
- BEDAU, M. Weak Emergence. In: TOMBERLIN, J. (Ed.) *Philosophical Perspectives: Mind, Causation, and World v. 11*. Blackwell Publishers, p. 375-399, 1997b.
- BEDAU, M. *Four Puzzles about Life*. Artificial Life No. 4, p. 125-140, 1998a.

- BEDAU, M. Philosophical Content and Method of Artificial Life. In: BYNAM, T.; MOOR, J. (Eds.). *The Digital Phoenix: How Computer are Changing Philosophy*. Basil Blackwell, p. 135-152, 1998b.
- BEDAU, M. Supple Laws in Biology and Psychology. In: HARDCASTLE, V. (Ed.) *Where Biology Meets Psychology: Philosophical Essays*. Cambridge: MIT Press, p. 287-302, 1999.
- BEDAU, M. *The Scientific and Philosophical Scope of Artificial Life*. Leonardo No. 35, p. 395-400, 2002.
- BEER, R. *Intelligence as Adaptive Behavior: An Experiment in Computational Neuroethology*. San Diego: Academic Press, 1990.
- BEER, R. *The Dynamics of Adaptive Behavior: A Research Program*. Robotics and Autonomous Systems, 20, p. 257-289, 1997.
- BEER, R. *Dynamical Approaches to Cognitive Science*. Trends in Cognitive Sciences, v. 4, n. 3, p. 91-99, 2000.
- BEKOFF, M. Cognitive Ethology and the Explanation of Nonhuman Animal Behavior. In: MEYER, J.; ROITBLAT, H. (Eds.). *Comparative Approaches to Cognitive Science*. Cambridge: MIT Press, p. 119-150, 1995.
- BENTLEY, P. *Digital Biology: How Nature is Transforming our Technology and our Lives*. New York: Simon Schuster, 2001.
- BERTALANFFY, L. *Teoria Geral dos Sistemas*. Petrópolis: Vozes, 1975.
- BIANCO, R.; NOLFI, S. *Toward Open-ended Evolutionary Robotics: Evolving Elementary Robotic Units able to Self-assemble and Self-reproduce*. Connection Science, 4: p. 227-248, 2004.
- BICKHARD, M. *Foundational Issues in Artificial Intelligence and Cognitive Science: Impasse and Solution*. Advances in Psychology 109. Amsterdam: Elsevier Science BV, 1995.
- BODEN, M. *Artificial Intelligence and Natural Man*. Cambridge: The MIT Press, 1987.
- BODEN, M. (Ed.). *The Philosophy of Artificial Intelligence*. New York: Oxford University Press, 1990.
- BODEN, M. *AI's Half-Century*. AI Magazine. Vol. 16, Issue 4, p. 96-99, 1995.
- BODEN, M. (Ed.). *The Philosophy of Artificial Life*. New York: Oxford University Press, 1996.
- BODEN, M. Life and Cognition. In: BRANQUINHO, J. *The Foundation of Cognitive Science*. New York: Oxford University Press, p. 11-22, 2001.

- BRAITENBERG, V. *Vehicles: Experiments in Synthetic Psychology*. Cambridge: The MIT Press, 1984.
- BRANQUINHO, J. *The Foundation of Cognitive Science*. New York: Oxford University Press, 2001.
- BRAGA, A.; CARVALHO, A.; LUDERMIR, T. *Redes Neurais Artificiais: Teoria e Aplicações*. Rio de Janeiro: LTC, 2000.
- BROENS, M. Mecanicismo e os Limites da Razão no Pensamento de Pascal. In: GONZALEZ, M.; BROENS, M. (Org.) *Encontro com as Ciências Cognitivas*. Vol. 2, p. 187-194. São Paulo: Unesp, 1998.
- BROOKS, R. A Robust Layered Control System for a Mobile Robot. *IEEE Journal of Robotics and Automation* 2(1), p. 14-23, 1986.
- BROOKS, R. *Elephants Don't Play Chess*. *Robotics and Autonomous Systems* 6, p. 3-15, 1990.
- BROOKS, R. *Intelligence Without Representation*. *Artificial Intelligence*, 47:139-159, 1991.
- BROOKS, R. Artificial Life and Real Robots. In: VARELA, F.; BOURGINE, P. (Eds.). *Toward a practice of autonomous systems: Proceedings of the First European Conference on Artificial Life*. Cambridge: MIT Press, p. 3-10, 1992.
- BROOKS, R.; MAES, P. (Eds.). *Artificial Life IV*. Cambridge: MIT Press, 1994.
- BROOKS, R. Intelligence Without Reason. In: STEELS, L. & BROOKS, R. (Ed.), *The Artificial Life Route to Artificial Intelligence: Building Embodied, Situated Agents*. Hillsdale: Lawrence Erlbaum, p.25-81, 1995.
- BROOKS, R. *Cambrian Intelligence: the early history of new AI*. Cambridge: The MIT Press, 1999.
- BUTCHER, H. *A Inteligência humana: natureza e avaliação*. São Paulo: Perspectiva, 1981.
- BYNAM, T.; MOOR, J. (Eds.). *The Digital Phoenix: How Computer are Changing Philosophy*. Basil Blackwell, 1998.
- CALABRETTA, R.; NOLFI, S.; PARISI, D.; WAGNER, G. *A Case Study of the Evolution of Modularity: Towards a Bridge Between Evolutionary Biology, Artificial Life, Neuro and Cognitive Science*. *Proceedings of the 6th International Conference on Artificial Life*, p. 275-284. Cambridge: The MIT Press, 1998.
- CARBONELL, J.; MICHALSKI, R.; MITCHELL, T. *Machine Learning: An Artificial Intelligence Approach*. Los Altos: Morgan Kaufmann, 1983.
- CARIANI, P. *On The Design of Devices with Emergent Semantic Functions*. Ph.D. thesis, State University of New York, Binghamton.1989.

CARIANI, P. Some Epistemological Implications of Devices which Construct their own Sensors and Effectors. In: VARELA, F.; BOURGINE, P. (Eds.). *Toward a practice of autonomous systems: Proceedings of the First European Conference on Artificial Life*. Cambridge: MIT Press, p. 484-493, 1992.

CASTRO, L.; VON ZUBEN, F. *Recent Developments in Biologically Inspired Computing*. Idea Group, 2005.

CATANIA, A. *Aprendizagem: Comportamento, Linguagem e Cognição*. Porto Alegre: Artmed, 1999.

CHALMERS, A. *O que é Ciência, afinal?* São Paulo: Editora Brasiliense, 1999

CHALMERS, D. *Why Fodor and Pylyshyn Were Wrong: The Simplest Refutation*. *Philosophical Psychology*, p. 305-319, 1993.

CHAUI, M. *Convite à Filosofia*. São Paulo: Ática, 2000.

CHISHOLM, R. *Teoria do Conhecimento*. Rio de Janeiro: Zahar, 1974.

CHURCHLAND, P. *A Neurocomputational Perspective: The Nature of Mind and Structure of Science*. Cambridge: MIT Press, 1989.

CHURCHLAND, P. Some Reductive Strategies in Cognitive Neurobiology. In: BODEN, M. (Ed.). *The Philosophy of Artificial Intelligence*. New York: Oxford University Press, 1990.

CHURCHLAND, P.; SEJNOWSKI, T. *The Computational Brain*. Cambridge: MIT Press, 1992.

CLARK, A. *Being There: Putting Brain, Body and World together again*. Cambridge: The MIT Press, 1997.

CLARK, A. *Mindware, An Introduction to the Philosophy of Cognitive Science*. New York: Oxford University Press, 2001.

CLIFF, D. Computational Neuroethology: A Provisional Manifesto. In: MEYER, J.; WILSON, S. (Eds.). *Animals to Animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior*. Cambridge: MIT Press, p. 29-39, 1991.

CLIFF, D. Neuroethology, Computational. In: ARBIB, M. A. (Ed.). *The Handbook of Brain Theory and Neural Networks*. Cambridge: MIT Press, 1995.

CRICK, F. *The Origin of the Genetic Code*. *Journal of Molecular Biology* v. 38, p. 367-379, 1968.

DAMASIO, A. *O Erro de Descartes: Emoção, Razão e o Cérebro Humano*. São Paulo: Companhia das letras, 1996.

DAWKINS, R. *O Gene Egoísta*. São Paulo: Itatiaia, 1979.

DAWSON, M. *From Embodied Cognitive Science to Synthetic Psychology*. Proceedings of the first IEEE International Conference on Cognitive Informatics, 2002.

DEBRUN, M. A. A idéia de Auto-organização. In DEBRUN, M.; GONZALEZ M. E.; PESSOA JR., O. (Eds.) *Auto-Organização: Estudos Interdisciplinares*. Coleção CLE. v.18. Campinas: UNICAMP, 1996.

D'OTTAVIANO, I., GONZALES, M. (Org.). *Auto-Organização: Estudos Interdisciplinares*. Coleção CLE. v.30. Campinas: UNICAMP, 1996.

DEL-MASSO, M. C. Vygostsky e a Ciência Cognitiva: a Importância dos Fatores Culturais no Pocesso de Aprendizagem. In: GONZALEZ, M.; BROENS, M. (Org.) *Encontro com as Ciências Cognitivas*. Vol. 2, p. 227-234. São Paulo: Unesp, 1998.

DENNETT, D. *Brainstorms: Philosophical Essays on Mind and Psychology*. Cambridge: MIT Press, 1978.

DENNETT, D. *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Simon and Schuster, 1995a.

DENNETT, D. Artificial Life as Philosophy. In: LANGTON, C. *Artificial Life: An Overview*. Cambridge: The MIT Press, p. 291-292, 1995b.

DENNETT, D. *Tipos de mentes: Rumo a uma Compreensão da Consciência*. Rio de Janeiro: Rocco, 1997.

DYER, G. Toward Synthesizing Artificial Neural Networks that exhibit Cooperative Intelligente Behavior: some open issues in Artificial Life. In: LANGTON, C. *Artificial Life: An Overview*. Cambridge: The MIT Press, p. 111-115, 1995.

DUPUY, J. P. *Nas Origens das Ciências Cognitivas*. São Paulo: Unesp, 1996.

DREYFUS, H. *O Que os Computadores Não Podem Fazer*. Rio de Janeiro: Eldorado, 1975.

DREYFUS, S.; DREYFUS, H. *The Power of Human Intuition and Expertise in the Era of the Computer*. New York: Free Press, 1986.

DREYFUS, H. *What Computers Still Can't Do: A Critique of Artificial Reason*. Cambridge: MIT Press, 1992.

EL-HANI, C.; VIDEIRA, A. *O Que é Vida? Para Entender a Biologia do Século XXI*. Rio de Janeiro: Relume Dumará, 2000.

EMMECHE, C. Life as an Abstract Phenomenon: Is Artificial Life Possible? In: VARELA, F.; BOURGINE, P. (Eds.). *Toward a practice of autonomous systems: Proceedings of the First European Conference on Artificial Life*. Cambridge: MIT Press, p. 466-474, 1992.

EMMECHE, C. *The Garden in The Machine. The Emerging Science of Artificial Life*. Princeton: Princeton University Press, 1994.

EMMECHE, C. *Does a Robot have an Umwelt? Reflections on the Qualitative Biosemiotics of Jacob von Uexküll*. Semiotica No. 134 (1/4), p. 653-693, 2001.

EMMECHE, C. *ALife, Organism and Body: The Semiotics of Emergent Levels*. In: BEDAU et al. (Eds) *ALife IX. Workshop and Tutorial Proceedings of Nineth International Conference on the Simulation and Synthesis of Living Systems*, p. 117-124, 2004.

EPISTEIN, I. (Org.). *Cibernética e Comunicação*. São Paulo: Cultrix, 1973.

EYSENCK & KEANE. *Psicologia Cognitiva, Um Manual Introdutório*. Porto Alegre: Artes Médicas, 1994.

FETZER, J. *Filosofia e Ciência Cognitiva*. Bauru: EDUSC, 2001.

FODOR, J. *The Language of Thought*. New York: Crowell, 1975.

FODOR, J.; PYLYSHYN, Z. *Connectionism and Cognitive Architecture: a Critical Analysis*. Cognition, 28, p. 3-71, 1988.

FOGEL, D. *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*. IEEE Press, 1995.

FRANKLIN, S. *Artificial Minds*. Cambridge: The MIT Press, 1995.

GARDNER, H. *A Nova Ciência da Mente*. São Paulo: Edusp, 1995.

GARDNER, H.; KORNHABER, M.; WAKE, W. *Inteligência: Múltiplas Perspectivas*. Porto Alegre: Artes Médicas, 1998.

GIARRATANO, J. *Foundations of Computer Technology*. Indianapolis: Howard W. Sams, 1982.

GILBERT, W. *The RNA World*. Nature v. 319, p. 618, 1986.

GLEISSER, M. *Cinquenta anos de "vida" no laboratório*. Caderno de Ciência, Folha de S. Paulo, 29/06/2003.

GOLDBERG, D. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, 1989.

GOMI, T. *Evolutionary Robotics - From intelligent robots to artificial life*. Ottawa: AAI Books, 1998.

GONZALEZ, M. et al (Org.) *Encontro com as Ciências Cognitivas*. Vol. 1. São Paulo: Unesp, 1995.

GONZALEZ, M. *Ação, Causalidade e Ruído nas Redes Neurais Auto-organizadas*. In: DEBRUN, M.; GONZALEZ M. E.; PESSOA JR., O. (Eds.) *Auto-Organização: Estudos Interdisciplinares*. Coleção CLE. v.18. Campinas: UNICAMP, 1996.

GONZALEZ, M.; BROENS, M. (Org.) *Encontro com as Ciências Cognitivas*. Vol. 2. São Paulo: Unesp, 1998.

GONZALEZ, M. et al. (Org.) *Encontro com as Ciências Cognitivas*. Vol. 3. São Paulo: Unesp, 2001.

GONZALEZ, M. E. *Ciência Cognitiva (notas de aula)*. Marília: Programa de Pós-graduação Interdisciplinar em Filosofia, Faculdade de Filosofia e Ciências da Unesp, 2003.

GUDWIN, R. *Umwelts and Artificial Devices. A Reflection on the text of Claus Emeche : Does a robot have an Umwelt?* Campinas: UNICAMP, 2004.

GUYTON, A.; HALL, J. *Tratado de Fisiologia Médica*. Rio de Janeiro: Guanabara Koogan, 1997.

HARNARD, S. *The Symbol Grounding Problem*. *Physica D*, v. 42, p. 335-346, 1990.

HARNARD, S. Artificial Life: Synthetic vs. Virtual. In: LANGTON, C. (Ed.). *Artificial Life III. The Third Workshop on the Synthesis and Simulation of Living Systems*. Redwood City: Addison-Wesley, p. 539-552, 1994.

HARNARD, S. Levels of Functional Equivalence in Reverse Bioengineering. In: LANGTON, C. *Artificial Life: An Overview*. Cambridge: The MIT Press, p. 293-301, 1995.

HARVEY, I. et al. *Evolutionary Robotics: A New Scientific Tool for Studying Cognition*. *Artificial Life* 11, p. 79-98, 2005.

HASELAGER, W. *Cognitive Science and Folk Psychology: the right frame of mind*. London: Sage, 1997.

HASELAGER, W. Neurodynamics and the Revival of Associationism in Cognitive Science. In: RIEGLER, A.; PESCHL, M.; VON STEIN, A. (Eds.). *Understanding Representation in The Cognitive Sciences: Does Representation Need Reality?* New York: Kluwer Academic, 1999.

HASELAGER, W. *The Representational Headache of Cognitive Science..* 4th Brazilian - International Conference on Cognitive Science. Marília, 2000.

HASELAGER, W. *Neurodynamics and Embodied Embedded Cognition*. Invited Lecture for the Austrian Society for Cognitive Science. Vienna, 2001.

HASELAGER, W. *Auto-organização, Emergência e Robótica (Notas de Aula)*. Marília: Programa de Pós-graduação Interdisciplinar em Filosofia, Faculdade de Filosofia e Ciências da Unesp, 2003.

HASELAGER, W. Auto-organização e Comportamento Comum: Opções e Problemas. In: SOUZA, G. et al. (Org.). *Auto-Organização: Estudos Interdisciplinares*. Coleção CLE. v.38, p. 213-235. Campinas: UNICAMP, 2004.

- HASELAGER, W. *Robotics, Philosophy and the Problems of Autonomy*. Pragmatics & Cognition, summer, 2005.
- HAUGELAND, J. *Artificial Intelligence: the Very Idea*. Cambridge: MIT Press, 1985.
- HAYKIN, S. *Neural Networks: A Comprehensive Foundation*. New York: Macmillan College Publishing, 1994.
- HEBB, D. *The Organization of Behavior: A Neuropsychological Theory*. New York: John Wiley & Sons, 1949.
- HILLGARD, E. *Teorias da Aprendizagem*. São Paulo: EPU, 1973.
- HODGES, A. *Turing, um filósofo da natureza*. São Paulo: Unesp, 2001.
- HOLLAND, J. *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press, 1975.
- HOYLE, G. *The Scope of Neuroethology*. Behavioral and Brain Sciences, 7, p. 367-412, 1984.
- HULL, D. *Philosophy of Biological Science*. Englewood Cliffs: Prentice-Hall, 1974.
- KAYZER, W. *Maravilhosa Obra do Acaso*. Rio de Janeiro: Nova Fronteira, 1998.
- KAUFMANN, S. *The Origins of Order*. New York: Oxford University Press, 1993.
- KAUFMANN, S. "O que é vida?": Schrödinger estava certo? In: MURPHY, M.; O'NEILL, J. (Org.). *O que é vida? - 50 anos depois: Especulações sobre o Futuro da Biologia*. São Paulo: Unesp, p. 101-135, 1997.
- KEELEY, B. Against the Global Replacement: On the Application of the Philosophy of Artificial Intelligence to Artificial Life. In: LANGTON, C. (Ed.). *Artificial Life III. The Third Workshop on the Synthesis and Simulation of Living Systems*. Redwood City: Addison-Wesley, p. 569-587, 1994.
- KEELEY, B. *Artificial Life for Philosophers*. Philosophical Psychology, V. 11, No. 2, p. 251-260, 1998.
- KEELEY, B. *Neuroethology and The Philosophy of Cognitive Science*. Philosophy of Science 67, p. 404-417, 2000.
- KOHONEN, T. *Self-organization and Associative Memory*. New York: Springer-Verlag, 1988.
- KUNG, S. *Digital Neural Networks*. New Jersey: Prentice-Hall, 1993.
- LAKOFF, G.; JOHNSON, M. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books, 1999.

- LANGDON JR, G.; FREGNI, E. *Projeto de Computadores Digitais*. Rio de Janeiro: Edgard Blücher, 1974.
- LANGE, M. *Life, Artificial Life and Scientific Explanation*. *Philosophy of Science*, 63, p. 225-244, 1996.
- LANGTON, C. *Studying Artificial Life with Cellular Automata*. *Physica D*, v. 22, p. 120-149, 1986.
- LANGTON, C. (Ed.). *Artificial Life. The Proceedings of an Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems*. New York: Addison-Wesley, 1989.
- LANGTON, C. et al (Eds.). *Artificial Life II. The Second Workshop on the Synthesis and Simulation of Living Systems*. Redwood City: Addison-Wesley, 1992.
- LANGTON, C. (Ed.). *Artificial Life III. The Third Workshop on the Synthesis and Simulation of Living Systems*. Redwood City: Addison-Wesley, 1994.
- LANGTON, C. *Artificial Life: An Overview*. Cambridge: The MIT Press, 1995.
- LANGTON, C.; SHINOHARA, K. (Ed.). *Artificial Life V. Proceedings of the Fifth International Workshop on the Synthesis and Simulation of Living Systems*. Cambridge: MIT Press, 1997.
- LATIL, P. *O pensamento Artificial*. São Paulo: Ibrasa, 1968.
- LEVY, S. *Artificial Life: The Quest for a New Creation*. New York: Pantheon Books, 1992.
- LORENZ, K. *Os Fundamentos da Etologia*. São Paulo: Unesp, 1995.
- LOULA, A.; GUDWIN, R.; QUEIROZ, J. *Comunicação Simbólica entre Criaturas Artificiais: um Experimento em Vida Artificial*. XVII Simpósio Brasileiro em Inteligência Artificial - SBIA, 2004.
- LUCCHESI, C. et al. *Aspectos Teóricos da Computação*. Rio de Janeiro: LTC, 1979.
- LUGER, G. *Cognitive Science: the Science of Intelligent Systems*. San Diego: Academic Press, 1994.
- MAES, P. (Ed.). *Designing Autonomous Agents: Theory and Practice from Biology to Engineering and Back*. Cambridge: The MIT Press, 1990.
- MAES, P. Behavior-Based Artificial Intelligence. In: MAES, P. et al. (Eds). *From Animals to Animats 2: Proceedings of the Second International Conference on the Simulation of Adaptive Behavior*. Cambridge: MIT Press, p. 2-10, 1993.
- MACLENNAN, B. *Synthetic Ethology: an Approach to the Study of Communication*. In: LANGTON, C. et al (Eds.). *Artificial Life II. The Second Workshop on the Synthesis and Simulation of Living Systems*. Redwood City: Addison-Wesley, p. 631-658, 1992.

- MACLENNAN, B. Synthetic Ethology: A New Tool for Investigating Animal Cognition. In: ALLEN, C.; BEKOFF, M.; BURGHARDT, G. (Eds.) *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*. Cambridge: MIT Press, p. 151-156, 2002.
- MATARIC, M. *Behavior-based Robotics as a Tool for Synthesis of Artificial Behavior and Analysis of Natural Behavior*. Trends in Cognitive Science, 2(3), p. 82-87, 1998.
- MATURANA, H.; VARELA, F. *Autopoiesis and Cognition: The Realization of the Living*. Boston: Reidel Pub Co, 1980.
- MATURANA, H.; VARELA, F. *The Tree of Knowledge*. Boston: Shambala, 1987.
- MATURANA, H. *Cognição, Ciência e Vida Cotidiana*. Belo Horizonte: Editora UFMG, 2001.
- MAYNARD SMITH, J. *The Theory of Evolution*. London: Penguin Books, 1959.
- MAYNARD SMITH, J. *When Learning Guides Evolution*. In Nature, 329, 761-762., 1987.
- MAYR, E. *The Growth of Biological Thought: Diversity, Evolution and Inheritance*. Cambridge: Harvard University Press, 1982.
- MAYR, E. *This is Biology: The Science of the Living World*. Cambridge: Harvard University Press, 1997.
- McCULLOCH, W.; PITTS, W. A Logical Calculus of the Ideas Immanent in Nervous Activity. In: BODEN, M. (Ed.). *The Philosophy of Artificial Intelligence*. New York: Oxford University Press, 1990.
- McLENNAN, B. *Synthetic Ethology: An Approach to the Study of Communication*. University of Tennessee, 1990.
- MENDEL, J. *A Prelude to Neural Networks: Adaptive and Learning Systems*. Englewood Cliffs: Prentice Hall, 1994.
- MEYER, J. The Animat Approach to Cognitive Science. In: ROITBLAT, H.; MEYER, J. (Eds.). *Comparative Approaches to Cognitive Science*. Cambridge: MIT Press, p. 27-44, 1995.
- MILLER, S. *Production of Amino Acids under Possible Primitive Earth Conditions*. Science, 117, p. 528, 1953.
- MILLER, S. The Prebiotic Synthesis of Organic Compounds as a Step Toward The Origin of Life. In: SCHOPF, J. (Ed.) *Major Events in the History of Life*. Boston: Jones & Bartlett, p. 1-28, 1992.
- MINSKY, M.; PAPERT, S. *Perceptrons: An Introduction to Computational Geometry*. Cambridge: The MIT Press, 1969.

- MINSKY, M. *The Society of Mind*. New York: Simon and Schuster, 1985.
- MORENTE, M. *Fundamentos da Filosofia*. São Paulo: Mestre Jou, 1966.
- MURPHY, M.; O'NEILL, J. (Org.). *O que é vida? - 50 anos depois: Especulações sobre o Futuro da Biologia*. São Paulo: Unesp, 1997.
- MURPHY, R. *Introduction to AI Robotics*. Cambridge: The MIT Press, 2000.
- NAGEL, E. *Mechanistic Explanation and Organismic Biology*. *Philosophy and Phenomenological Research* v.11, p. 327-338, 1951.
- NEWELL, A.; SIMON, H. *Computer Science as Empirical Inquiry: Symbols and Search*. *Communications of the ACM*, 19/3, p. 113-126, 1976.
- NEWELL, A. *Physical Symbol Systems*. *Cognitive Science*, v. 4, p. 135-183, 1980.
- NOLFI, S.; ELMAN, J.; PARISI, D. *Learning and Evolution in Neural Networks*. *Adaptive Behavior*, 3 (1), p. 5-28, 1994.
- NOLFI, S.; PARISI, D. *Neural Networks in an Artificial Life Perspective*. In: GERSTNER et al (Eds.). *Proceedings of the 7th International Conference on Artificial Neural Networks*. Berlin: Springer Verlag, p. 733-738, 1997.
- NOLFI, S. *Evolutionary Robotics*. *Connection Science*. V.10, No 3(4), p.167-184, 1998.
- NOLFI, S.; FLOREANO, D. *Learning and Evolution*. *Autonomous Robots*, 7 (1), p. 89-113, 1999.
- NOLFI, S.; FLOREANO, D. *Evolutionary robotics: The Biology, Intelligence and Technology of Self-Organizing Machines*. Cambridge: The MIT Press, 2000.
- OLIVEIRA, M. *Vygotsky: Aprendizado e Desenvolvimento*. São Paulo: Sipione, 1993.
- OPARIN, A. *The Origin of Life in Earth*. New York: Academic Press, 1938.
- PATTEE, H. *Cell Psychology: An Evolutionary View of the Symbol-Matter Problem*. *Cognition and Brain Theory* v. 5, p. 325-341, 1982.
- PATTEE, H. *Simulations, Realization, and Theories of Life*. In: LANGTON, C. *Artificial Life: An Overview*. Cambridge: The MIT Press, p. 63-78, 1995.
- PATTEE, H. *Artificial Life Needs a Real Epistemology*. In: MORAN, F. et al (Eds). *Advances in Artificial Life, Proceedings of Third European Conference on Artificial Life*. Berlin: Springer-Verlag, p. 23-38, 1995.
- PEREIRA JR., A. *Introdução à Filosofia da Neurociência Cognitiva*. In GONZALEZ, M. E. et al. (Org.) *Encontro com as Ciências Cognitivas*. Vol. I. São Paulo: Unesp, 1995.

- PEREIRA JR., A.; GUIMARÃES, R.; CHAVES JR., J. Auto-organização em Biologia: Nível Ontogenético. In DEBRUN, M.; GONZALEZ M. E.; PESSOA JR., O. (Eds.) *Auto-Organização: Estudos Interdisciplinares*. Coleção CLE. v.18. Campinas: UNICAMP, 1996.
- PEREIRA JR., A. Breve Histórico da Neurociência Cognitiva. In: GONZALEZ, M. et al. (Org.) *Encontro com as Ciências Cognitivas*. Vol. 3. São Paulo: Unesp, 2001.
- PEREIRA JR., A.; LUSSI, I.; PEREIRA, M. Mente. In MARTINS, R.; MARI, H. (Org.) *Universos do Conhecimento*. Belo Horizonte: FALE UFMG, 2002.
- PINKER, S. *Como a Mente Funciona*. São Paulo: Companhia das Letras, 1998.
- PIQUEIRA, J.; MONTEIRO, L.; DEL NERO, H. Considerações sobre Cognição e Evolução. In: GONZALEZ, M.; BROENS, M. (Org.) *Encontro com as Ciências Cognitivas*. Vol. 2, p. 235-246. São Paulo: Unesp, 1998.
- POPPER, K. *The Logic of Scientific Discovery*. New York: Basic Books, 1959.
- PUTNAM, H. Minds and Machines. In: HOOK, S. (Ed.) *Dimensions of Mind*. New York: Collier Books, 1961.
- PUTNAM, H. *Mind, Language and Reality*. Cambridge: Cambridge University Press, 1975.
- PYLYSHYN, Z. *Computation and Cognition: Toward a Foundation for Cognitive Science*. Cambridge: MIT Press, 1984.
- RAY, T. An Approach to the Synthesis of Life. In: LANGTON, C. et al (Eds.). *Artificial Life II. The Second Workshop on the Synthesis and Simulation of Living Systems*. Redwood City: Addison-Wesley, p. 371-408, 1992.
- RAY, T. An Evolutionary Approach to the Synthetic Biology. In: LANGTON, C. *Artificial Life: An Overview*. Cambridge: The MIT Press, p. 179-209, 1995.
- RAY, T. Artificial Life. In: DULBECCO, R. et al. (Eds) *Frontiers of Life*. Academic Press, 2001.
- REYNOLDS, C. *Flocks, Herds and Schools: a distributed behavioral model*. Computer Graphics 21 (4), p. 25-34, 1987.
- ROISENBERG, M. *Emergência da Inteligência em Agentes Autônomos através de modelos inspirados na natureza*. Tese de doutorado. Departamento de Engenharia Elétrica - Grupo de Pesquisas em Engenharia Biomédica - Universidade Federal de Santa Catarina, 1998.
- ROSEN, R. *Life Itself: A Comprehensive Inquiry Into the Nature, Origin and Fabrication of Life*. New York: Columbia University Press, 1991.
- RUMELHART, D.; McCLELLAND, J. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Vol. I. Cambridge: The MIT Press, 1986.

RUSSELL, E. *Form and Function: A Contribution to the History of Animal Morphology*. London: J. Murray, 1916.

RYLE, G. *The Concept of Mind*. England: Hutchinson & Company Ltd., 1949.

SCHANK, R.; ABELSON, R. *Scripts, Plans, Goals und Understanding*. Hillsdale: Erlbaum, 1977.

SCHANK, R.; BIRNBAUM, L. Aumentando a Inteligência. In: KHALFA, J. (Org.). *A Natureza da Inteligência*. São Paulo: Unesp, 1996.

SCHNEIDER, E.; KAY, J. Ordem a partir da Desordem: A Termodinâmica da Complexidade Biológica. In: MURPHY, M.; O'NEILL, J. (Org.). *O que é vida? - 50 anos depois: Especulações sobre o Futuro da Biologia*. São Paulo: Unesp, p. 187-201, 1997.

SCHRÖDINGER, E. *O que é vida? O Aspecto Físico da Célula Viva*. São Paulo: Unesp, 1997.

SCORNAIENCHI, D. *Invenções e Descobertas*. São Paulo: Editora Melhoramentos, 1975.

SEARLE, J. Minds, Brains, and Programs. In: BODEN, M. (Ed.). *The Philosophy of Artificial Intelligence*. New York: Oxford University Press, 1990.

SEARLE, J. *A Redescoberta da Mente*. São Paulo: Martins Fontes, 1997.

SILVA, F.; ROISENBERG, M.; BARRETO, J. *Redes Neurais Hierarquicas para a implementação de comportamentos em agentes autônomos*. XIII Congresso Brasileiro de Automática - CBA2000. Florianópolis, 2000.

SIPPER, M. *An Introduction to Artificial Life*. Explorations in Artificial Life (especial issue of AI Expert), p. 4-8, 1995.

SKINNER, B. F. *Tecnologia do Ensino*. São Paulo: Herder, 1972.

SKINNER, B. F. *Sobre o Behaviorismo*. São Paulo: Cultrix / EDUSP, 1982.

SKINNER, B. F. *Can Psychology be a Science of Mind?* American Psychologist, 45 / 11, p. 1206-1210, 1990.

SMOLENSKY, P. *On the Proper Treatment of Connectionism*. Behavioral and Brain Sciences, 11, p. 1-74, 1988.

SOARES, A. *O que são Ciências Cognitivas*. São Paulo: Brasiliense, 1993.

SOBER, E. *Panglossian Functionalism and the Philosophy of Mind*. Synthese No. 64, p. 165-193, 1985.

SOBER, E. Learning from Functionalism - Prospects for Strong Artificial Life. In: LANGTON, C. et al (Eds.). *Artificial Life II. The Second Workshop on the Synthesis and Simulation of Living Systems*. Redwood City: Addison-Wesley, p. 749-765, 1992.

SOBER, E. *Philosophy of Biology*. Boulder: Westview Press: 1993.

SOUZA, G. A Natureza das Formas Biológicas: A Auto-organização e a Cognição Formadoras. In: GONZALEZ, M.; BROENS, M. (Org.) *Encontro com as Ciências Cognitivas*. Vol. 2, p. 49-67. São Paulo: Unesp, 1998.

SOUZA, G. et al. (Org.). *Auto-Organização: Estudos Interdisciplinares*. Coleção CLE. v.38. Campinas: UNICAMP, 2004.

STEELS, L. The Artificial Life Roots of Artificial Intelligence. In: LANGTON, C. *Artificial Life: An Overview*. Cambridge: The MIT Press, p. 75-110, 1995.

STEELS, L.; BROOKS, R. (Ed.). *The Artificial Life Route to Artificial Intelligence: Building Embodied, Situated Agents*. Hillsdale: Lawrence Erlbaum, 1995.

STEPHAN, A. *Varieties of Emergentism*. *Evolution and Cognition* V. 5, No. 1, p. 49-59, 1999

STEWART, J. Life = Cognition: The Epistemological and Ontological significance of Artificial Life. In: VARELA, F.; BOURGINE, P. (Eds.). *Toward a practice of autonomous systems: Proceedings of the First European Conference on Artificial Life*. Cambridge: MIT Press, p. 475-483, 1992.

STRUICK, D. *A Concise History of Mathematics*. New York: Dover Publications, 1967.

TAMAYO, P.; HARTMAN, H. Cellular Automata, Reaction-diffusion Systems and the Origin of Life. In: LANGTON, C. (Ed.). *Artificial Life. The Proceedings of an Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems*. New York: Addison-Wesley, p. 105-124, 1989.

TAYLOR, C.; JEFFERSON, D. Artificial Life as a Tool for Biological Inquiry. In: LANGTON, C. *Artificial Life: An Overview*. Cambridge: The MIT Press, p. 1-13, 1995.

TEIXEIRA, J. *Mentes e Máquinas: Uma Introdução à Ciência Cognitiva*. Porto Alegre: Artes Médicas, 1998.

TEIXEIRA, J. *Mente, Cérebro e Cognição*. Petrópolis: Vozes, 2000.

THAGARD, P. *Mente: Introdução à Ciência Cognitiva*. Porto Alegre: Artmed, 1998.

TOMASELLO, M. *The Cultural Origins of Human Cognition*. Harvard: Harvard University Press, 1999.

TURING, A. *On Computable Numbers, with an Application to the Entscheidungsproblem*. In: *Proceedings of the London Mathematical Society*, 2/42, 1936, p. 230-265.

- TURING, A. Computing, Machinery and Intelligence. In: BODEN, M. *The Philosophy of Artificial Intelligence*. New York: Oxford University Press, 1990.
- VAN GELDER, T. *What might cognition be, if not computation?* Journal of Philosophy, v. 42, n.7, p. 345-381, 1995.
- VARELA, F. *Principles of Biological Autonomy*. New York: Elsevier, 1980.
- VARELA, F.; THOMPSON, E.; ROSCH, E. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge: The MIT Press, 1991.
- VARELA, F.; BOURGINE, P. (Eds.) *Toward a Practice of Autonomous Systems: Proceedings of the First European Conference on Artificial Life*. Cambridge: The MIT Press, 1992.
- VARELA, F. *Conhecer as Ciências Cognitivas: tendências e perspectivas*. Lisboa : Instituto Piaget, 2000.
- VIGNAUX, G. *As Ciências Cognitivas*. Lisboa: Instituto Piaget, 1995.
- VON FOESTER, H.; ZOPF, G. (Eds.). *Principles of Self-organization*. London: Pergamon, 1962.
- VON NEUMANN, J.; MORGENSTERN, O. *Theory of Games and Economic Behavior*. New Jersey: Princeton University Press, 1944.
- VON NEUMANN, J. *First Draft of a Report on the EDVAC*. Moore School of Electrical Engineering - University of Pennsylvania, 1945.
- VON NEUMANN, J. *Preliminary Discussion of the Logical Design of an Electronic Computing Instrument*. Princeton: Institute for Advanced Study, 1946.
- VON NEUMANN, J. *Theory of Self-Reproducing Automata*. Edited and completed by Arthur W. Burks. Urbana: University of Illinois Press, 1966.
- VON NEUMANN, J. *The Computer and the Brain*. New Haven: Yale University Press, 1958.
- WASSERMAN, P. *Neural Computing: Theory and Practice*. New York: Van Nostrand Reinhold, 1989.
- WEB, B.; SMITHERS, T. The Connection between AI and Biology in the Study of Behavior. In: VARELA, F.; BOURGINE, P. (Eds.). *Toward a practice of autonomous systems: Proceedings of the First European Conference on Artificial Life*. Cambridge: MIT Press, p. 421-428, 1992.
- WEISS, G. Towards the Synthesis of Neural and Evolutionary Learning. In OMIDVAR, O. (Ed.). *Progress in Neural Networks*, Vol. 5, Chapter 5. Ablex, 1993.

WIENER, N. *Cibernética: Controle e Comunicação no Animal e na Máquina*. São Paulo: Cultrix, 1970.

WINOGRAD, T.; FLORES, F. *Understanding Computers and Cognition: A New Foundation for Design*. Addison-Wesley, 1987.

WILSON, S. Knowledge Growth in an Artificial Animal. In: Grefenstette, J. (Ed.). *Proceedings of an International Conference on Genetic Algorithms and Their Applications*. Hillsdale: Lawrence Erlbaum, p. 16-23, 1985.

WILSON, S. The Animat Path to AI. In: MEYER, J.; WILSON, W. (Eds). *From Animals to Animats: Proceedings of the First International Conference on the Simulation of Adaptive Behavior*. Cambridge: MIT Press, p. 15-21, 1991.

WRIGLEY, M. B.; SILVA, C. R. *A Ciência da Cognição, Segundo Dascal*. Campinas: Jornal da Unicamp, 22/04/2003.

ZIEMKE, T.; SHARKEY, N. Life, Minds and Robots: The Ins and Outs of Embodied Cognition. In: WERMTER, S.; SUN, R. (Eds.). *Hybrid Neural Systems*. Springer Verlag, 2000.

ZIEMKE, T.; SHARKEY, N. *A Stroll Through The Worlds of Robots and Animals: Applying Jakob van Uexküll's Theory of Meaning to Adaptive Robots and Artificial Life*. Semiotica, 134-1/4, p. 701-746, 2001.