

Luiz Carlos Begosso

**CIÊNCIA COGNITIVA: UMA ABORDAGEM
INFORMACIONAL DA MENTE**

UNESP - Universidade Estadual Paulista
Marília - 1999



LUIZ CARLOS BEGOSSO

1410000977



**CIÊNCIA COGNITIVA: UMA ABORDAGEM
INFORMACIONAL DA MENTE**



LUIZ CARLOS BEGOSSO

**CIÊNCIA COGNITIVA: UMA ABORDAGEM
INFORMACIONAL DA MENTE**

Dissertação apresentada à Faculdade de
Filosofia e Ciências da Universidade
Estadual Paulista "Júlio de Mesquita Filho",
Campus de Marília, para a obtenção do título
de Mestre em Filosofia (Linha de Pesquisa:
Ciência Cognitiva e Filosofia da Mente).

Orientadora: Prof^ª. Dr^ª. Maria Eunice Quilici Gonzales

MARÍLIA
1999



LUIZ CARLOS BEGOSSO

**CIÊNCIA COGNITIVA: UMA ABORDAGEM
INFORMACIONAL DA MENTE**

COMISSÃO JULGADORA

DISSERTAÇÃO PARA A OBTENÇÃO DO TÍTULO DE MESTRE

Presidente e Orientador: Profª Drª Maria Eunice Quilici Gonzales

2º Examinador : Profª Drª Plácida Leopoldina V. A. da Costa Santos

3º Examinador : Profª Drª Mariana Cláudia Broens

1º Suplente : Prof Dr Elias Humberto Alves

2º Suplente : Prof Dr Jônatas Manzolli

Marília, 11 de Novembro de 1999



DADOS CURRICULARES**LUIZ CARLOS BEGOSSO**

NASCIMENTO 17.05.1962 - CÂNDIDO MOTA/SP

FILIAÇÃO Luiz Begosso
Nézia Cavaca Begosso

1985-1988 Graduação: Tecnologia em Processamento de Dados
Universidade Estadual de Maringá.

1991-1992 Curso de Especialização em Metodologia do Ensino Superior,
na Universidade Estadual de Londrina.

1989- Professor da Área de Computação, do Instituto Municipal de
Ensino Superior de Assis (IMESA/FEMA).

Para Izabel, Rafael e Renata

Que me deram forças e apoio na
hora certa para a realização deste
trabalho.

Agradecimentos

O meu agradecimento especial é dedicado à Profª Drª Maria Eunice Quilici Gonzales, pela orientação, dedicação e paciência durante todos estes anos, tornando possível a elaboração desta dissertação.

Agradeço a todos os meus colegas do curso de Ciência Cognitiva, que sempre me apoiaram.

ao Maurizio, pela amizade e incentivo;

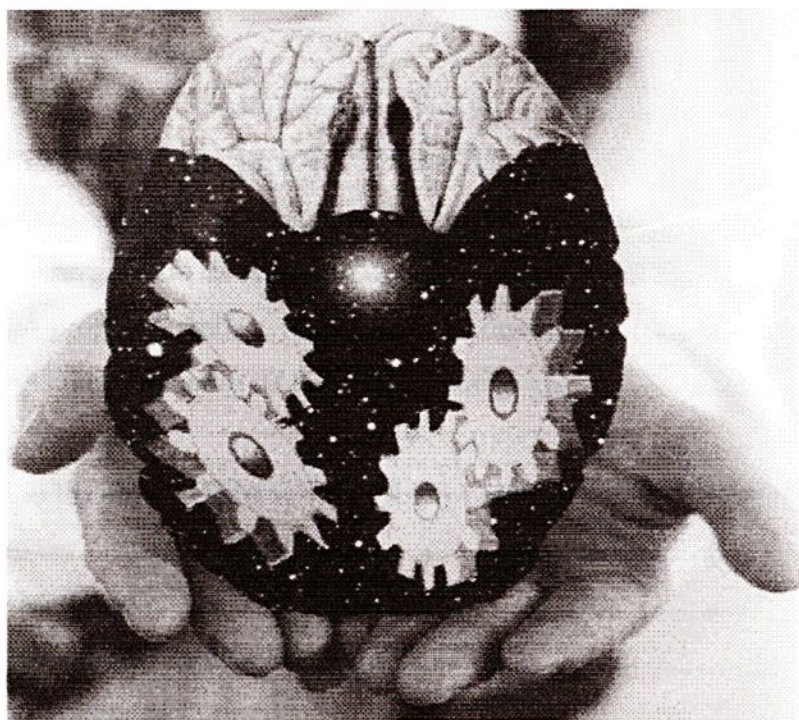
a todos os professores do curso com quem convivi estes últimos três anos;

aos meus companheiros de trabalho na Fundação Educacional do Município de Assis, pelo incentivo e apoio;

ao meu irmão, Nino, que sempre me incentivou nesse desafio;

aos meus pais, que sempre me estimularam e não pouparam esforços para minha formação;

agradeço a Deus pela minha existência.



É de se esperar que as máquinas acabem por competir com o homem em todos os campos puramente intelectuais... Podemos avistar só um pequeno trecho do caminho à nossa frente, mas ali já vemos muito do que precisa ser feito.

Alan Mathison Turing (1973, p.82)

SUMÁRIO

LISTA DE FIGURAS.....	10
INTRODUÇÃO	11
1 - O LUGAR DA MENTE NA CIÊNCIA COGNITIVA	
Introdução	13
Seção 1 - O surgimento da Ciência Cognitiva	13
Seção 2 - A proposta explicativa da Ciência Cognitiva	16
Seção 3 - As duas vertentes da Ciência Cognitiva	23
2 - TÉCNICAS DE MODELAGEM E MODELOS COMPUTACIONAIS	
Introdução	34
Seção 1 - Modelos	35
Seção 2 - O conceito de Algoritmo	46
Seção 3 - Máquina de Turing	51
Seção 4 - Máquina Registradora Ilimitada	58
3 - TEORIA DA INFORMAÇÃO	
Introdução	65
Seção 1 - Histórico da Teoria da Informação	65
Seção 2 - Teoria da Informação	73
Seção 3 - O elo entre a Teoria da informação e a Ciência Cognitiva.....	83
4 - REDES NEURAIS ARTIFICIAIS	
Introdução	89
Seção 1 - Principais características das Redes Neurais Artificiais.....	90
Seção 2 -Tipos de Aprendizagem nas Redes Neurais Artificiais	96

Seção 3 - Smolensky e o modelo da Harmonia no Sistema Cognitivo.....	103
CONCLUSÃO	110
BIBLIOGRAFIA	116
RESUMO	122
ABSTRACT	123

LISTA DE FIGURAS

Figura 1 - Três planos de análise no estudo da percepção visual	17
Figura 2 - Reconhecimento de um padrão	29
Figura 3 - Representação analógica de uma regra mecânica	49
Figura 4 - Máquina de Turing	52
Figura 5 - Máquina Universal	54
Figura 6 - Sistema de Comunicação	74
Figura 7 - Simulação do jogo Perguntas e Respostas	79
Figura 8 - Árvore genealógica de um crime.....	86
Figura 9 - Neurônio biológico	91
Figura 10 - Neurônio artificial com n entradas e uma saída	92
Figura 11 - Típica rede neural artificial com três camadas	93
Figura 12-Um nódulo computando entradas e a função transferência	95
Figura 13-Rede neural artificial de Kohonen	100
Figura 14-Um modelo harmônico	104
Figura 15-Esquema da rede neural da harmonia	108

INTRODUÇÃO

Na presente dissertação temos como objetivo investigar aspectos gerais do comportamento inteligente relacionados à organização informacional da memória cognitiva. Partimos dos pressupostos da Ciência Cognitiva: (a) de que seres humanos e máquinas, apesar de diferirem em inúmeros aspectos, possuem em comum a habilidade de processar informação; (b) e de que os computadores podem servir como bons modelos explicativos de aspectos do comportamento humano inteligente.

Para atingirmos tal objetivo, discutimos no *Capítulo 1* as raízes e os fundamentos da Ciência Cognitiva. Na *Seção 1*, apresentamos um breve histórico da Ciência Cognitiva e de seu programa de pesquisa, na tentativa de explicar aspectos relevantes sobre o funcionamento da mente. Na *Seção 2*, destacamos a importância da utilização de modelos, como recursos explicativos, na discussão sobre a natureza da mente humana. Finalmente, na *Seção 3*, ilustrando a natureza das explicações (em geral) na Ciência Cognitiva, apresentamos as suas duas vertentes principais: Inteligência Artificial e Redes Neurais Artificiais.

No *Capítulo 2*, discutimos o papel dos *modelos* na Ciência Cognitiva. Na *Seção 1*, enfatizamos o papel dos modelos como recursos explicativos. Na *Seção 2*, analisamos a noção de algoritmo segundo a visão de: Cutland e Rogers. Na *Seção 3*, apresentamos a Máquina Universal de Turing e estudamos algumas de suas aplicações na Ciência Cognitiva. Finalmente, na *Seção 4*, apresentamos a Máquina Registradora Ilimitada como uma variação da Máquina Universal. Preparamos, neste



capítulo, as bases sobre as quais se aplicarão alguns princípios da Teoria da Informação na Ciência Cognitiva.

No *Capítulo 3, Seção 1*, desenvolvemos um breve histórico da Teoria da Informação, sugerindo algumas possíveis contribuições para a Ciência Cognitiva. Na *Seção 2*, discutimos a Teoria Matemática da Comunicação proposta por Shannon & Weaver, procurando caracterizar o conceito de Informação que acreditamos estar presente nos modelos elaborados pelos cientistas cognitivos. Na *Seção 3*, sugerimos um elemento de ligação entre a teoria da informação e a Ciência Cognitiva, reforçando o pressuposto cognitivista de que seres humanos e máquinas possuem em comum um elemento fundamental, que consiste na capacidade de processar informação.

No *Capítulo 4*, investigamos as Redes Neurais Artificiais, que experimentaram um grande avanço na década de 80. Na *Seção 1*, estudamos as principais características das Redes Neurais. Na *Seção 2*, ressaltamos os mecanismos de aprendizagem e treinamento nessas redes. Finalmente, na *Seção 3*, apresentamos uma teoria das representações mentais, proposta por Smolensky, chamada de Teoria da Harmonia, que estabelece um elo entre a Teoria da Informação e o paradigma das redes neurais artificiais.

Com estes quatro capítulos concluímos a presente dissertação sugerindo que sistemas cognitivos inteligentes buscam constantemente descobrir relações informacionais no meio ambiente, característica esta que encontra um terreno fértil para as investigações na Ciência Cognitiva.

CAPÍTULO 1

O LUGAR DA MENTE NA CIÊNCIA COGNITIVA

INTRODUÇÃO

Neste Capítulo, discutimos as raízes e os fundamentos propostos por esta nova ciência chamada Ciência Cognitiva. Na Seção 1, faremos um breve histórico da Ciência Cognitiva e de seu programa de pesquisa, na tentativa de explicar aspectos relevantes sobre o funcionamento da mente. Para isso, na Seção 2, destacaremos, resumidamente, a importância do emprego de modelos na discussão sobre a natureza da mente humana. Finalmente, na Seção 3, mostramos as duas vertentes surgidas na Ciência Cognitiva: a Inteligência Artificial e as Redes Neurais Artificiais.

SEÇÃO 1 - O SURGIMENTO DA CIÊNCIA COGNITIVA

O início da década de 40 foi marcado pelo desejo de um grupo de cientistas em fundar uma ciência que, mais tarde, ficou conhecida como ciência da mente. Em 1943, Rosembueth, Wiener e Bigelow (Dupuy, 1996) publicaram um



artigo na revista *Philosophy of Science* chamado “Behavior, Purpose and Teleology”. O artigo mostrava as relações comportamentais de *input*, *output* e *feedback* de fenômenos naturais, ressaltando um dos primeiros elementos que deram início à Ciência Cognitiva. Ainda nesse mesmo ano, o neuropsiquiatra Warren McCulloch e o matemático Walter Pitts publicaram um segundo artigo que lançou, em definitivo, juntamente com o trabalho de Roseblueth, Wiener e Bigelow, a pedra fundamental da Ciência Cognitiva (Dupuy, 1996, p.45). O artigo de McCulloch e Pitts descrevia as relações entre as células nervosas, que poderiam, em sua opinião, ser descritas em termos da lógica. De uma maneira geral, McCulloch e Pitts modelizaram o neurônio natural num operador lógico que transforma inputs em outputs.

Dupuy (1996) ressalta que, com o término da II Guerra, McCulloch ficou encarregado de organizar um ciclo de conferências, patrocinadas pela Fundação Macy, sobre as questões que suscitaram debates relacionados ao funcionamento da mente, cérebro, comportamento etc. A partir de 1946 até 1953, ocorreram, nos Estados Unidos, dez encontros denominados de *Conferência Macy*. Também no ano de 1948 aconteceu o Simpósio Hixon.

Para Gardner (1995), tais encontros tinham em comum a *interdisciplinariedade*, a continuidade das discussões a respeito de mecanismos cerebrais, comportamentais e computacionais, além de reunir questionamentos sobre o modelo comportamentalista vigente na época. O Simpósio Hixon foi determinante para mostrar que o estudo behaviorista do comportamento bloquearia os avanços da nova ciência que começava a surgir. O neurofisiologista Karl Lashley, por exemplo,

foi, segundo Gardner (1995, p.26), um dos grandes críticos da visão behaviorista, no Simpósio. Uma das idéias que permeava esse simpósio era a vontade explícita, entre alguns participantes, de descrever, através da matemática e da física, fenômenos mentais e não apenas comportamentais.

Assim, o primeiro momento, que caracterizou, neste século, o início da Ciência Cognitiva, foi o que Soares (1993, p.13) chamou de "...vontade de simular informaticamente os processos mentais e/ou cerebrais". Porém, antes de tentar qualquer simulação, faz-se necessário definir, no contexto científico, os fundamentos teóricos dessa nova ciência.

Gardner (1995) caracteriza a Ciência Cognitiva como uma área de estudo interdisciplinar, cujas raízes filosóficas se encontram nos antigos debates gregos sobre a natureza do conhecimento humano. Ele ressalta, na Ciência Cognitiva, o esforço contemporâneo, fundamentado empiricamente, para responder a essas questões, propondo a seguinte caracterização para o programa de pesquisa dessa ciência:

- 1º Para compreender as atividades cognitivas, deve-se levar em consideração o estudo das Representações Mentais;
- 2º O computador será utilizado como ferramenta para modelar a mente humana;
- 3º Fatores afetivos ou emotivos não serão considerados, nesse momento, no estudo cognitivo;
- 4º O estudo interdisciplinar contribuirá em muito nessa investigação;



5º As questões sobre a natureza da mente, que preocupavam os antigos filósofos, devem ser consideradas no estudo da Ciência Cognitiva.

Um dos aspectos fundamentais da Ciência Cognitiva é a tentativa de explicação dos fenômenos inteligentes que operam nos organismos e nas máquinas. Para isso, acreditam os cognitivistas, é fundamental a construção de modelos para garantir o entendimento e explicação desses fenômenos. Na próxima Seção, apresentamos a proposta explicativa da Ciência Cognitiva apoiada, principalmente, nos três primeiros pressupostos dessa ciência, quais sejam: o estudo das atividades cognitivas deve levar em consideração as representações mentais, os computadores constituem uma ferramenta apropriada para modelar a mente e, finalmente, os fatores emocionais não serão considerados, nesse momento, nos estudos cognitivos.

SEÇÃO 2 - A PROPOSTA EXPLICATIVA DA CIÊNCIA COGNITIVA

Um dos pressupostos centrais da Ciência Cognitiva aponta na direção de que os computadores constituem instrumentos adequados para a construção de modelos que permitam entender o funcionamento da mente.

A modelização da mente, através dos computadores, origina-se no pressuposto, da Ciência Cognitiva, segundo o qual a mente deve ser estudada a partir de sua capacidade computacional e representacional (Gonzales, 1991, p.95). Desse modo, os computadores poderiam ser utilizados para verificar ou ilustrar alguns

aspectos do funcionamento da mente concernentes às formas de produção e organização das representações mentais.

Um dos primeiros e mais importantes modelos computacionais explicativos da Ciência Cognitiva foi elaborado por David Marr. Ele desenvolveu um modelo computacional, incorporando uma descrição sobre o sistema visual, que considerava os problemas relacionados à formação de imagem e conseqüente percepção de objetos. Nesse modelo, três planos de análise deveriam ser considerados: os planos da teoria computacional, do algoritmo e representação e, finalmente, sua implementação. A seguir, reproduzimos, na figura 1, a idéia de Marr (1982, p. 25) a respeito desses três planos de análise:

Teoria Computacional	Algoritmo e Representação	Implementação do Hardware
Qual é o objetivo que se deseja atingir e qual é a estratégia apropriada para atingi-lo?	Como a teoria computacional pode ser implementada? Qual é a representação para as entradas e saídas? Qual é o algoritmo para essa transformação?	Como a representação e o algoritmo podem ser avaliados fisicamente?

Figura 1 - Três planos de análise no estudo da percepção visual.

Dada a classificação explicativa de Marr, concebemos três etapas na investigação de um fenômeno. A primeira, que concebe a origem das idéias a partir

de uma perspectiva mecanicista, é chamada de *Teoria Computacional*. Nela apresentam-se as questões e idéias apropriadas para alcançar um determinado objetivo. Na segunda, *Algoritmo e Representação*, procura-se o algoritmo apropriado que implemente as idéias bem definidas no plano 1, ou seja, responde-se à questão: Como implementar, algoritmicamente, as hipóteses formuladas em (1)? A terceira etapa, de *Implementação do Hardware*, indaga: Como as idéias e os algoritmos definidos nos planos anteriores podem ser realizáveis fisicamente?

Na etapa de representação algorítmica (plano 2), elaboramos modelos explicativos dos problemas claramente concebidos na 1ª etapa. No que diz respeito à preocupação da Ciência Cognitiva, podemos desejar explicar, por exemplo, os mecanismos de aprendizagem envolvidos nos sistemas inteligentes. Nesse caso, como aplicaríamos os três planos de explicação de Marr?

No plano 1, que diz respeito às idéias e metas, um objetivo a ser alcançado poderia ser entender a natureza da inteligência e os mecanismos de aprendizagem nela envolvidos, por exemplo, numa tarefa de reconhecimento inteligente de padrões.

Tendo claramente enunciado qual a meta a ser atingida, passamos para o plano 2 (que diz respeito aos algoritmos), buscando algum algoritmo que envolva soluções para problemas de aprendizagem, por exemplo, a Lei de Hebb (apresentaremos essa Lei no Capítulo IV).

Uma vez estruturado o algoritmo, temos a construção de um modelo que, segundo Dupuy (1996, p. 23), "... é como uma forma abstrata que vem encarnar-

se ou realizar-se nos fenômenos" (no Capítulo 2, apresentamos uma discussão sobre a importância dos modelos na Ciência Cognitiva). O plano 3, após a criação do modelo, é responsável pela implementação e verificação da validade das hipóteses incorporadas no plano 2.

Marr argumenta que não é suficiente, para o entendimento de como funciona o sistema visual, concentrar-se apenas nos estudos dos processos cerebrais. A construção de modelos computacionais deve ser um bom caminho a ser seguido, pois permite não apenas a manipulação de dados, mas, acima de tudo, possibilita o teste das hipóteses explicativas a serem propostas sobre os mecanismos organizadores da visão.

O modelo elaborado por Marr, ajustado por nós, para a consecução do objetivo proposto no início desta dissertação, aplica-se da seguinte forma:

Objetivo Geral	Representação	Implementação do Hardware
Procurar entender alguns aspectos da natureza da inteligência e os mecanismos de aprendizagem neles envolvidos.	A construção de modelos computacionais, através de Redes Neurais Artificiais (RNA), ajudará no entendimento de alguns aspectos do comportamento inteligente.	Construção de uma RNA sustentada na Teoria da Harmonia proposta por Smolensky.

No mesmo contexto, Soares (1993, p. 46-7) aponta para a computação como auxiliar no estudo e no entendimento do conhecimento humano. A autora sugere que o aparato computacional deve ser utilizado na construção de modelos, o que evitaria, em alguns casos, a interferência da subjetividade humana, preservando-se a objetividade tão desejada nas explicações científicas.

Como veremos no Capítulo 2, os modelos tendem, na ciência em geral, a garantir um controle explicativo e preditivo de situações reais. O cientista, ao elaborar modelos, representa o mundo real na tentativa de entendê-lo. O conhecimento através da técnica de modelagem pode ser entendido como uma forma de representação do mundo, na qual a Ciência Cognitiva fundou sua sustentação.

É justamente com esse intuito, de tentar modelar aspectos do conhecimento humano, que Turing (1973) propõe a questão: "Podem as máquinas pensar?". Discutiremos a partir deste ponto algumas questões relativas à proposta de modelagem da inteligência e ao Jogo da Imitação, decorrentes dessa proposta.

O trabalho de Turing resultará na tentativa de manter uma resposta positiva à pergunta exposta no parágrafo anterior, qual seja: "Podem as máquinas pensar?". Esse trabalho serviu para que muitos cientistas concentrassem esforços na tentativa de explicar alguns aspectos do comportamento humano inteligente, através de instrumentos mecânicos.

Para desenvolver o Jogo da Imitação, que mais tarde ficou conhecido como Teste de Turing, Turing (1973, p.49-54) estabelece um jogo do qual participam um Homem (A), uma Mulher (B) e um Interrogador (C) de qualquer sexo. Os dois



primeiros jogadores devem permanecer numa sala, separados do Interrogador, em que se comunicam (A e C ou B e C) de maneira impessoal (de tal modo que o Interrogador não saiba com quem esteja se comunicando). O objetivo do Interrogador (C), no jogo, é descobrir quem é o Homem e quem é a Mulher. O jogador A (o Homem) tentará induzir o Interrogador a pensar que ele (A) seria B. Finalmente, para o jogador B (a Mulher) o objetivo do jogo é ajudar o Interrogador, dando sempre respostas verdadeiras.

Em seguida, Turing muda o jogo e pergunta: O que acontecerá quando uma máquina ocupar o lugar do homem? No mínimo, as diferenças entre as capacidades físicas e intelectuais do homem e da máquina tornar-se-ão evidentes. Porém, Turing acredita que o objetivo do jogo agora é de obter, da máquina, respostas tão naturais quanto se fossem dadas por um homem. Se a máquina puder substituir perfeitamente o homem neste jogo e, sendo o homem inteligente, a máquina estará desempenhando um papel igualmente inteligente.

Com relação ao tipo de máquina que poderá fazer parte do jogo, Turing responde que sua preocupação reside no fato de os computadores terem ou não capacidade para um bom desempenho no Teste. Não importa a constituição física da máquina; se ela está revestida de pele humana artificial ou de qualquer outro artifício. O interessante é que máquinas digitais realizem algo que possa ser descrito como uma atividade abstrata da inteligência.

Ao explicar, resumidamente, o funcionamento dos computadores digitais, Turing ressalta que essas máquinas são planejadas para executar qualquer

operação realizável por um homem (em seu texto, ele as chama de 'computador humano'). Para que um computador realize uma tarefa qualquer, esta deve, primeiramente, ser traduzida num conjunto claro e objetivo de passos a serem executados e programados no computador digital. No Capítulo 2, discutiremos detalhadamente a noção de algoritmo que aqui chamamos de *conjunto claro e objetivo de passos/etapas*.

O Teste de Turing influenciou o programa de pesquisa da Ciência Cognitiva, na tentativa de buscar algumas explicações científicas a respeito da natureza e estrutura do pensamento humano inteligente. Pensava-se, segundo Gardner (1995), que, a partir das idéias de Turing, fosse possível descrever detalhadamente alguns aspectos do comportamento inteligente humano. Tal descrição, quando incorporada a um computador, permitiria a ele agir e pensar inteligentemente. A lógica por trás desse raciocínio seria a seguinte: se a máquina passar no Teste de Turing, ou seja, se ela comportar-se como um ser humano inteligente, então seu programa poderá constituir um bom modelo explicativo para entender o comportamento inteligente por ela simulado.

Em resumo, as tentativas de modelizar aspectos do comportamento trouxeram consigo inúmeros recursos da ciência natural e da computação, resultando nesta nova ciência denominada de Ciência Cognitiva. Na próxima Seção, veremos duas tendências surgidas no estudo cognitivo, as quais são conhecidas como Inteligência Artificial (IA) e Redes Neurais Artificiais (RNA).



SEÇÃO 3 - AS DUAS VERTENTES DA CIÊNCIA COGNITIVA

No percurso histórico da Ciência Cognitiva, temos notado diversas tentativas de clarear as idéias no que diz respeito à formação e aquisição do conhecimento. Como vimos anteriormente, uma alternativa proposta para estudar os fenômenos cognitivos é a construção de modelos. A partir dos pressupostos iniciais dessa ciência, que inclui a crença nas capacidades representacional e computacional da mente, discutiremos as duas vertentes da Ciência Cognitiva: a Inteligência Artificial (que denotaremos de IA) e o Conexionismo (também conhecido como Redes Neurais Artificiais-RNA ou Sistemas de Processamento Paralelo e Distribuído - PDP).

Com o objetivo de caracterizar essas duas linhas de pesquisa (IA e PDP), Gonzales (1998) ressalta que o ponto em comum entre elas é a construção de modelos computacionais buscando-se o entendimento da estrutura dos estados mentais, perspectiva esta entendida a partir de uma concepção mecanicista.

O objetivo central da Inteligência Artificial, segundo Gonzales (1994, 1998), é buscar nas regras abstratas e nos algoritmos a explicação de algumas atividades de nosso universo cognitivo, utilizando-se de técnicas de processamento seqüencial de informação em computadores digitais. A IA caracteriza os estados cognitivos como entidades abstratas, desvinculando o estudo dessas entidades do seu

substrato físico. Nessa vertente, são ressaltados os aspectos racionais da cognição, tais como, atividade de resolução de problemas, demonstração de teoremas etc.

Essa desvinculação dos estados cognitivos com o seu substrato físico indica que os modelos computacionais construídos pela IA não simulam, necessariamente, os processos mentais e comportamentais, exclusivos do homem. O fator importante, aqui, é aquilo que a mente ~~inteligente~~ inteligente, humana, animal ou maquinal, pode realizar sem levar em consideração os processos físicos em geral que desencadearam as tarefas realizadas pela mente.

Nesse contexto, a IA tem experimentado, por exemplo, um grande desenvolvimento nas seguintes áreas:

- Sistemas especialistas;
- Identificação e composição de melodias musicais;
- Traduções em vários idiomas;
- Reconhecimento de padrões;
- Resolução de problemas;
- Identificação de caracteres;
- Demonstração de teoremas matemáticos;
- Jogos etc.

Alguns cientistas procuraram tirar da IA o foco na resolução de problemas matemáticos, como foi o caso de Weizenbaum (1976). Ele construiu um programa, denominado ELIZA, que simula a conduta de um terapeuta durante uma

sessão com um paciente. A seguir, ilustramos um exemplo de *conversação* entre o programa e o *paciente* (as frases em maiúsculas foram emitidas pelo programa):

- OLÁ MEU NOME É ELIZA
- Eu preciso de muito dinheiro
- VOCÊ POSSUI ALGUM PROBLEMA PSICOLÓGICO?
- Não!
- O QUE VOCÊ SUGERE?
- Eu já disse, preciso de dinheiro
- ENTENDO
- O que você sugere para mim?
- EU NÃO ESTOU CERTA. SE EU ENTENDO VOCÊ
COMPLETAMENTE.
- .
- .
- .

Conforme ilustra a sequência acima, o ELIZA pode *dialogar*, de forma simplificada, com qualquer pessoa que não conheça a estrutura em que o programa foi desenvolvido. O segredo está na busca, realizada pelo programa, de palavras-chave presentes nas sentenças do interlocutor. Uma vez localizadas as palavras-chave (para o programa, palavra-chave é aquela que está presente na frase imediatamente anterior do interlocutor e também em sua estrutura interna), o programa emite respostas fixas, ou seja, previamente armazenadas.

O trecho de programa abaixo ilustra as possíveis opções que ELIZA teria num diálogo típico. É importante notar que na linha 10 existe uma instrução de escolha aleatória, por parte do programa, com o objetivo de possuir alternativas de diálogo. Essas alternativas estão entre as linhas 12 e 20:

```
1. function GetReply (Question : string) : string;
2.   var
3.     I : integer;
4.     TempReply : string;
5.   begin
6.     if (Pos('HOW', Question) > 0) or (Pos('WHO', Question) > 0) or
7.       (Pos('WHAT', Question) > 0) or (Pos('WHEN', Question) > 0) or
8.       (Pos('WHY', Question) > 0) or (Pos('WHERE', Question) > 0) then
9.       begin
10.        I := Trunc(Random(9));
11.        case I of
12.          0 : GetReply := 'Porque você pergunta?';
13.          1 : GetReply := 'Você possui algum problema psicológico?';
14.          2 : GetReply := 'O que você sugere?';
15.          3 : GetReply := 'O que você acha?';
16.          4 : GetReply := 'Estas questões são freqüentes?';
17.          5 : GetReply := 'O que você realmente quer saber?';
18.          6 : GetReply := 'Você já perguntou isto para mais alguém?';
```

19. 7 : GetReply := 'Você já perguntou estas questões anteriormente?';
20. 8 : GetReply := 'O que você pensa quando pergunta isso?';
21. end;
22. Exit;
23. end;

No exemplo de conversação citado anteriormente, notemos que a palavra *dinheiro* não deve ser uma palavra-chave para o ELIZA, pois o *diálogo* não se desenvolveu por parte do programa. Talvez uma possível resposta pudesse ser: "Porque o dinheiro é importante para você?" A saída encontrada pelo programa foi a tentativa de mudança no *diálogo*.

Em resumo, através de regras, seja da estrutura lógica do diálogo, seja da demonstração de teoremas, procura-se simular nos modelos de IA os mecanismos e as regras responsáveis pelo comportamento inteligente.

Numa outra forma de investigação dos fenômenos inteligentes, as Redes Neurais Artificiais (RNA) (que estudaremos detalhadamente no Capítulo 4) não direcionam, segundo Gonzales (1996), totalmente seus estudos aos aspectos racionais do sistema cognitivo, mas sim aos aspectos naturais (do substrato físico) e adaptativos dos sistemas de processamento de informação. Em um outro trabalho, Gonzales (1998, p.132) salienta que as RNA enfatizam as leis naturais de formação e desenvolvimento das representações mentais. Segundo aponta a autora,

... os modelos conexionistas são desenvolvidos admitindo- se que, antes de mais nada, se trata de explicar a relação entre a macro e a micro estrutura de sistemas físicos que se adaptam às mais variadas situações do meio ambiente.

As RNA modelam os sistemas físicos de forma que estes possam direcionar, com os resultados das pesquisas neurológicas, novas investigações sobre o aparato cognitivo.

Lippmann (1987) reforça a idéia da importância de discutirmos a estrutura física do cérebro, lembrando que a estrutura das RNA foi inspirada no nosso conhecimento atual do sistema nervoso. Segundo o autor, algumas pesquisas produziram resultados interessantes, como, por exemplo, nas áreas do reconhecimento da fala e da imagem. Essas propriedades naturais (fala e imagem) requerem, na simulação artificial, um grande potencial de processamento paralelo e distribuído. As RNA aproximam-se, com a implementação de modernas tecnologias, dessas propriedades naturais.

Como mencionamos acima, uma das áreas mais desenvolvidas nas RNA é o estudo sobre reconhecimento de padrões. Neste estudo, as redes passam por um processo de *aprendizagem*, qual seja: através da apresentação de uma entrada, a RNA gera uma saída que é comparada ao que se deseja reconhecer. No caso, por exemplo de padrões como aquele da figura 2, a rede será treinada para reconhecer a entrada desejada. Caso sua saída seja diferente do padrão apresentado (a entrada), os mecanismos de adaptação e correção de erros da rede serão acionados, de forma a

tentar corrigir as distorções apresentadas pela saída (reduzindo as diferenças entre entrada e saída).

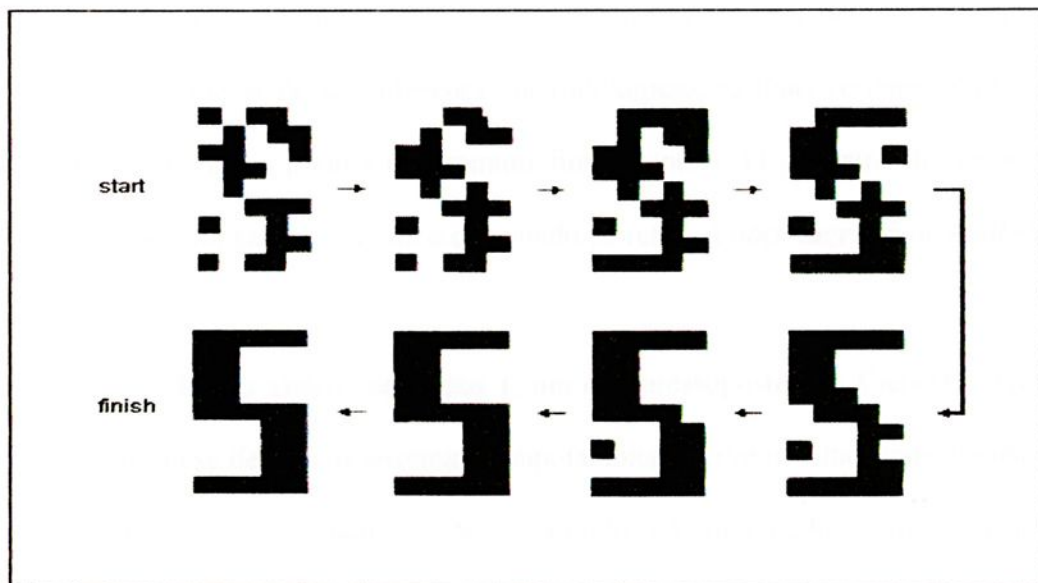


Figura 2 - Reconhecimento de um padrão (Beale & Jackson, 1994, p.144).

No Capítulo 4, mostraremos em detalhe como se dá esse processo de correção de erros numa rede neural. O que nos interessa ressaltar aqui é que tais modelos de RNA, diferentemente dos modelos de IA, *aprendem* a reconhecer padrões informacionais estabelecidos no meio ambiente. Eles não possuem de antemão a representação de tais padrões na memória. Além disso, o processo de aprendizagem ocorre com base em *mecanismos físicos* de ajustes de pesos nas conexões que se estabelecem entre as unidades dinamicamente conectadas.

A metodologia adotada na elaboração desses modelos é conhecida como *bottom-up*, ou ascendente, pois a partir de elementos (ou conceitos) particulares, busca o entendimento de conceitos gerais.

Já nos modelos de IA, a metodologia utilizada é a *top-down*, ou descendente, pois percorre o caminho inverso ao da abordagem anterior, ou seja, a partir de conceitos gerais busca-se a caracterização de conceitos mais particulares.

Apesar dessas diferenças metodológicas, as duas vertentes da Ciência Cognitiva apresentam pontos em comum fundamentais. O primeiro diz respeito à *abordagem mecanicista da mente* e o segundo se refere à *abordagem funcionalista* de explicação.

Como vimos na Seção 1, um dos pressupostos da Ciência Cognitiva reside na hipótese de que os sistemas computacionais servem como bons modelos de funcionamento da mente humana. Nesse sentido, os computadores amparam a tese funcionalista do sistema cognitivo, segundo a qual os estudos da mente podem ser realizados, independentemente de uma estrutura física particular.

O funcionalismo aponta para a capacidade das máquinas e seres humanos de processarem informação, apesar das diferenças materiais existentes entre eles. O que interessa ao funcionalista são as estruturas funcionais existentes nos sistemas de processamento de informação, estruturas estas que decorrem das relações causais estabelecidas entre as partes desses sistemas e os elementos que os circundam. Para o funcionalista, o material *específico* que instancia tais sistemas (seja ele metal, silício ou tecido orgânico) é irrelevante para a sua capacidade de processamento de informação. Essa capacidade, para o funcionalismo, pode explicar a organização das máquinas e dos seres humanos na produção dos estados cognitivos.



Gonzales (1993) categoriza o funcionalismo como Lógico Computacional (FLC) e Neurocomputacional (FNC). Essa classificação é explicada abaixo:

- De acordo com o funcionalismo lógico computacional (FLC), os processos cognitivos devem ser tratados como entidades simbólicas, abstratas, independentes de suas características físicas. O funcionalista computacional parte do princípio segundo o qual processar informação é manipular símbolos através de regras, sem a preocupação com o significado embutido nessa simbologia. Como já vimos nesta Seção, os cientistas da IA adotam o FLC em seus estudos do sistema cognitivo.
- O funcionalismo neurocomputacional (FNC) enfatiza alguns tipos de mecanismos físicos no estudo dos sistemas de processamento de informação. Assim, de acordo com esta abordagem, as funções cognitivas emergem do processamento paralelo e distribuído de informação a partir de conexões entre as unidades neurônio-símile (veremos essa estrutura, detalhadamente, no Capítulo 4).

Sem especificar um tipo particular de substrato físico que instanciaria tais mecanismos, o funcionalismo sugere uma analogia entre mentes e máquinas computadoras: assim como nas máquinas podem-se executar programas em diferentes substratos físicos, da mesma forma a mente humana pode ser entendida, no plano do *software*, como um conjunto de regras para a organização de representações. Tais representações poderiam ser materializadas em diferentes substratos físicos (entre os quais, o cérebro é considerado um excelente candidato).



A abordagem funcionalista permite uma resposta ao tema deste Capítulo. A pergunta: "Qual é o lugar da mente na Ciência Cognitiva?" poderá ser respondida em dois momentos.

No primeiro momento, considerando-se a mente (na visão funcionalista) como um conjunto dinâmico de representações estruturadas através de regras, ela não se situará em lugar algum, pois as regras e as representações são entidades abstratas que não ocupam lugar no espaço. De acordo com o funcionalismo, essas regras, ao serem instanciadas no cérebro, produzem o comportamento humano inteligente. Sua materialização em máquina daria lugar, segundo Turing, ao comportamento mecânico inteligente.

No segundo momento, a mente não é entendida como um conjunto de símbolos e regras abstratas que reproduziriam o comportamento humano inteligente. Ao contrário, ela é entendida como uma rede interconectada de unidades (neurônios no cérebro e neurônios-símile nas redes neurais artificiais) que produz comportamentos inteligentes, conhecimentos e representações mentais. Os Conexionistas, que representam esta segunda fase da Ciência Cognitiva, acreditam que a partir das conexões entre as unidades neurônio-símile emergem as atividades consideradas inteligentes. Para o conexionismo, a mente está nessas conexões.

A diferença entre as duas vertentes, segundo Gonzales (1996), é que enquanto a IA focaliza a manipulação de símbolos e regras para representar os mais altos níveis da estrutura da mente, a abordagem conexionista direciona seus esforços no entendimento dos níveis sub-simbólicos das representações, a partir de leis físicas,



mecanismos de aprendizagem e processos adaptativos, para a compreensão dos processos mentais. Ela, entretanto, não faz qualquer menção a respeito de tipos de materiais específicos que deveriam ser bons candidatos para a modelagem. Como veremos no Capítulo 4, tudo o que necessitamos são nódulos físicos, em geral, que recebam, acumulem e emitam pulsos elétricos respondendo aos estímulos do mundo externo.

Em resumo, analisamos neste Capítulo o surgimento da Ciência Cognitiva, bem como seus pressupostos básicos para existir enquanto ciência. Indicamos, também, a proposta dessa Ciência na tentativa de explicar aspectos do comportamento humano através da elaboração de modelos. A partir dessa tentativa, surgiram, dentro da Ciência Cognitiva, duas vertentes que buscam clarear as hipóteses sobre a formação e aquisição do conhecimento humano: IA e RNA. No próximo capítulo apresentaremos, em mais detalhes, a tentativa de modelização computacional da mente humana.



CAPÍTULO 2

TÉCNICAS DE MODELAGEM E MODELOS COMPUTACIONAIS

INTRODUÇÃO

O computador eletrônico foi um dos maiores inventos produzidos neste século. Estudos que levaram a essa prodigiosa invenção datam das primeiras décadas do século XIX, como os do matemático inglês Charles Babbage, que projetou a *máquina analítica*. Esta máquina executava as quatro operações aritméticas, possuía um dispositivo de armazenamento de dados e executava a impressão dos mesmos (Breton 1991).

Também não são recentes os esforços que tentam explicar o pensamento humano, desde os estados mais primitivos da intenção ou propósito da comunicação até a expressão de uma idéia ou um símbolo.

Neste capítulo, interessa-nos discutir o papel dos modelos na Ciência Cognitiva, porém se faz necessário definir o que seja um *modelo* ou, de outro modo, identificá-lo no domínio científico. Estudaremos a tentativa de modelização da mente com o trabalho proposto por Alan Mathison Turing e caracterizaremos a noção de



computabilidade, apresentando uma variação da máquina abstrata proposta por Turing, a Máquina Registradora Ilimitada.

Para estudar os esforços na tentativa de explicação do pensamento humano, apresentaremos: na Seção 1, o papel dos modelos na Ciência Cognitiva; na Seção 2, a noção de algoritmo e seu conceito, segundo a visão de alguns autores, tais como: Cutland (1980) e Rogers (1987); na Seção 3, a Máquina Universal (conhecida como Máquina de Turing) e algumas de suas aplicações; na Seção 4, a Máquina Registradora Ilimitada (M.R.I.), uma variação da Máquina Universal de Turing, que será de grande importância na caracterização de processos determinísticos.

SEÇÃO 1 - MODELOS

Nesse estudo sobre modelos, fazemos referência à mente e ao cérebro, os quais são tratados, pelos filósofos materialistas, como sinônimos. Como indicamos no Capítulo 1, dada a complexidade do estudo da mente, descontextualizaremos, temporariamente, na Ciência Cognitiva, fatores emocionais e atitudes subjetivas e culturais, enfatizando, principalmente, o conceito de representação mental.

De acordo com Gardner (1995), na década de 40, com a realização do Simpósio Hixon, os estudos relacionados às Representações Mentais (RM) tiveram um avanço significativo. O psicólogo Karl Lashley apresentou a idéia de que as



Representações Mentais existem e fazem parte de um processo evolutivo da comunicação humana. Para Lashley, os novos estudos sobre as RM deveriam descartar a posição comportamentalista que tratava a mente humana como uma *caixa-preta*. Os comportamentalistas (ou behavioristas) acreditavam que expressões cognitivas surgiam de *fora para dentro*, ou seja, o indivíduo era estimulado a se comportar de alguma maneira conforme o ambiente o estimulava a agir. Os estudiosos interessados nessa ciência eram estimulados a apresentarem resultados quantitativos e qualitativos dos seus experimentos. Entendia-se como estudo behaviorista aquele que pudesse ser aplicado e verificado por qualquer outro cientista. O behaviorismo descartou o introspeccionismo e o subjetivismo. Temas como pensamento, mente e imaginação não deveriam compor objetos de estudo para os comportamentalistas. Fenômenos internos, mentais e não-observáveis diretamente deveriam ser, segundo esses estudiosos, desconsiderados. O abandono do paradigma da *caixa-preta* trouxe à tona as discussões filosóficas e psicológicas referentes ao cérebro e à mente humana.

Os primeiros computadores surgiram nas décadas de 40-50 e, para uma boa parte dos cientistas da época, a invenção não era muito confiável. Apesar disso, alguns deles chamaram a atenção para o paralelo existente entre os símbolos utilizados pelo homem na comunicação e os sistemas simbólicos utilizados pelo computador para processar informação.

Segundo Dupuy, podemos dizer que a Ciência Cognitiva tem suas raízes no ano de 1943, com a publicação de dois artigos: o primeiro, de autoria de



Roseblueth, Wiener e Bigelow, que introduz o conceito de *feedback* descrito "...como um retorno da energia do output sobre o input" (Dupuy, 1996, p.45). Esse conceito traduz a observação do comportamento de um sistema a partir do meio ambiente em que ele está inserido. A saída produzida pelo sistema, quando interage com o ambiente, poderá produzir mudanças na entrada, modificando o comportamento do sistema, no caso o ser humano.

O que difere tal visão da visão comportamentalista é a inclusão da noção de *feedback circular*, que dá uma certa *ação* ao objeto em estudo. O sistema não responde à entrada simplesmente com uma resposta frágil e automática, mas, sim, de acordo com a saída é capaz de mudar a relação recebida entre a entrada e a saída.

O segundo artigo já mencionado no Capítulo 1, de grande impacto no estudo da mente, de autoria de McCulloch e Pitts, procura buscar formas lógicas e matemáticas para as questões relacionadas ao comportamento inteligente.

McCulloch e Pitts mostraram, segundo Gardner (1995), que as conexões entre as células nervosas (uma rede neural biológica) e suas operações podem ser modeladas em termos da lógica: os neurônios poderiam ser comparados aos enunciados lógicos e as operações dos impulsos (ou não impulsos) poderiam ser o cálculo proposicional. Assim, tal modelo permitia que se pensasse num neurônio ativado e ativando outros neurônios, da mesma forma que uma proposição lógica pode implicar outras proposições. A analogia entre células nervosas e circuitos lógicos pode ser considerada em termos de impulsos elétricos que passam (ou não) num determinado circuito. Um conjunto de circuitos lógicos ("neurônios") pode então



formar o que ficou conhecido, mais tarde, como Rede Neural Artificial, que será amplamente utilizada na técnica de modelagem da atual Ciência Cognitiva.

Dupuy (1996, p.23) apresenta uma discussão sobre modelos e indica que o modelo

...é como uma forma abstrata que vem encarnar-se ou realizar-se nos fenômenos... O modelo científico é uma imitação humana da natureza que o cientista logo toma como “modelo” - no sentido comum - desta.

O cientista, ao lidar com fatos do mundo material, tenta entender o fenômeno natural, de maneira que possa ter controle sobre a situação. Os modelos têm garantido esse controle explicativo e preditivo, e a matemática constitui-se numa poderosa ferramenta capaz de auxiliá-lo na sua construção. As simulações¹ feitas sobre o modelo traduzem as idéias de analogia, com o mundo real, em que o cientista impõe as diversas formas de modelização. Os modelos têm vida e estrutura próprias que se distanciam da realidade formal. O cientista projeta sua mente no mundo real durante a atividade da modelização. Tais fatos de simulação científica parecem sugerir que “a mente é ao mesmo tempo distinta da matéria e adequada a ela” (Dupuy, 1996, p. 25). A mente não pode produzir a própria mente, e a consciência não se faz por si mesma. Para que a mente conheça a própria mente, sendo ela capaz de criar um modelo de si mesma, será preciso que novas descobertas lógicas

¹ No contexto científico, segundo Dupuy (1996), a simulação é um caso particular de modelização e consiste em reproduzir as funções de um sistema no computador.



questionem tal manipulação. Parece-nos, pois, que o que está embutido no projeto científico do conhecimento é a produção (ou reprodução) de modelos de um fato, de tal modo que a representação da realidade possa ser feita através de repetição e de simulações.

Como já mencionamos no Capítulo 1, sob a luz do 2º aspecto que caracteriza a Ciência Cognitiva, o computador pode servir como um modelo que ilustra o funcionamento do pensamento humano, e como instrumento capaz de testar a eficiência desse modelo como sendo relevante para o entendimento da mente humana.

A modelização da mente ganhou força com o trabalho teórico de Alan Turing. De maneira bastante simplificada, veremos, na Seção 3, que a Máquina Universal proposta por ele computa funções matemáticas, determinísticas, através de uma linguagem formal (algoritmo). Segundo Dupuy (1996), o objetivo de Turing, quando apresentou seu trabalho, era resolver um problema proposto por Hilbert, o “problema da decisão”, que pode ser assim enunciado: “dada uma fórmula qualquer do cálculo de predicados, existe um procedimento sistemático, geral, efetivo, que permita determinar se essa fórmula é demonstrável ou não?” (Dupuy, 1996, p.28). Os estudos de Turing mostraram que toda função calculável de maneira mecânica é calculável por uma máquina de Turing, ou seja, que uma máquina de Turing pode modelizar qualquer função computável.

Turing (1973) afirma que os computadores digitais podem imitar, de modo muito aproximado, as ações do que ele chamou de “computador humano”. O



computador humano deve seguir regras fixas, pré-definidas, e não tem poder para questioná-las ou desviar-se delas. Tal situação nos leva a refletir se uma máquina é capaz de realizar somente aquilo que lhe fora ordenado a executar. Nesse mesmo artigo, Turing descreve as objeções de Lady Lovelace, em relação ao seu trabalho, que enfatiza o fato de as máquinas realizarem apenas o que lhes ordenamos. Lady Lovelace afirma que a máquina não pode produzir nada de novo e que esta jamais pode nos pegar de surpresa. Turing rebate a primeira afirmação, perguntando sobre quem poderá garantir que a obra produzida por Lovelace não foi fruto de um amadurecimento de idéias. Tal obra não poderá ser única, no sentido de ter concebido apenas uma idéia, sem sucessivas reflexões. Nesse sentido, tal obra também não nos pegaria de surpresa, uma vez que se apóia em idéias bem conhecidas.

A segunda reflexão é tomada por Turing como um descuido de Lovelace: "...a máquina jamais pode nos pegar de surpresa" (Turing, 1973, p. 71). Pelo contrário, diz ele, muitas vezes executamos algumas operações, apressadamente, sem nos preocupar com o rigor necessário àquela situação, enquanto a máquina nos surpreenderá com tal resolução adequada. O autor questiona tais reflexões, fazendo as seguintes analogias: se um homem pudesse colocar uma idéia numa máquina, esta responderia e retornaria ao seu estado inicial de imobilidade, como um martelo acionando a corda de um piano, ou uma idéia inserida numa pilha atômica corresponderá a um neutron entrante na pilha. Esse neutron poderá causar tamanha perturbação, reagindo em cadeia, excitando outros neutrons e por fim toda a pilha estará ativada em função da excitação inicial, provocando a destruição da pilha.



Turing pergunta se existe um fenômeno para máquinas e outro para mentes. Ele responde, em seguida, que existe um mecanismo para a mente, capaz de, a partir de uma idéia, produzir uma teoria e encadeando uma série de outras idéias simples ou complexas. Para o fenômeno da inteligência na máquina, Turing diz que o problema é principalmente de programação e, em segundo plano, tecnologia que permita mais avanços nessa área. Cerca de 10^{10} a 10^{15} bits apontam para a capacidade de memória do cérebro. Porém, apenas uma pequena parte desse volume é realmente utilizada por nós e, desta porção em uso, armazenamos impressões visuais.

Para Turing, na tentativa de imitar a mente humana, encontramos algumas máquinas que trabalham cerca de 1000 vezes mais velozes que as células nervosas e devemos, então, refletir acerca de três importantes aspectos:

- a) No nascimento, o estado inicial da mente;
- b) A educação recebida;
- c) Experiência à qual a mente foi submetida, não considerada como educação.

Em se tratando da mente, no seu estado inicial, poderíamos tentar simular numa máquina essa situação de mente infantil, ao invés de nos concentrarmos num programa que simule a mente adulta. O autor prossegue fazendo uma analogia da mente de uma criança com um caderno novo: o caderno tem muitas folhas em branco e pouco mecanismo (que seria a educação da criança). Dessa forma, o problema de simular a mente humana - de uma criança - estaria dividido em duas partes intimamente ligadas: a construção do programa infantil e o processo de educação. No início, não se deve esperar uma boa máquina-criança, ao contrário,

deve-se ensiná-la repetidas vezes e esperar por seu melhor comportamento. Ensinar uma outra máquina-criança é uma boa estratégia para poder comparar o desempenho das duas máquinas. Turing (1973) descreve três aspectos que devemos considerar ao treinarmos uma máquina-criança:

estrutura da máquina- criança = material hereditário

mudanças na máquina- criança = mutação

seleção natural = juízo do experimentador (Turing, 1973, p.78).

O experimentador, utilizando-se de sua própria inteligência, poderá identificar algum problema na educação da máquina-criança, produzindo mudanças de modo a corrigir o seu comportamento e indicar qual dos dois modelos estará em melhores condições de aprendizado.

Turing sugere que, para que a máquina-criança efetivamente aprenda, podemos programá-la considerando tais questões (de punição e recompensa): quando alguma situação levar a máquina-criança ao erro, imediatamente, após esse fato, deve-se repreendê-la de tal forma que isso não se repita. De outro modo, reforçar a situação para que ela ocorra com mais frequência, caso não inclua erros. Ele enfatiza que o uso de punições e recompensas² não pressupõe quaisquer fatores emotivos, e constitui uma parte importante do processo ensino-aprendizagem. Dessa forma, a

² Punição e recompensa, neste caso, aplicados à máquina-criança, podem ser valorizados através de uma linguagem simbólica interpretada pela máquina, atribuindo pontos positivos a situações de acerto e pontos negativos a situações de erro ou indesejáveis.

linguagem utilizada para transmitir as punições, recompensas e as ordens dadas à máquina-criança, poderia ser do tipo simbólico, através de um canal não emotivo.

Neste ponto, é necessário refletirmos sobre as proposições formuladas por Turing. Ao enveredar por caminhos da educação infantil e, analogamente, da educação da máquina-criança, Turing espera dotar de inteligência humana um computador digital. Para isso, ele propõe a utilização de regras simbólicas, as quais chama de mudanças de comportamento na máquina-criança. Por punição e recompensa podemos entender os diversos tipos de ajustes na programação computacional.

A programação da máquina-criança pode ser composta de um sistema de inferência lógica: definições e proposições. As proposições, como o próprio Turing (1973, p.79) sugere, podem ser

... fatos bem estabelecidos, conjeturas, teoremas matemáticos demonstrados, enunciados de autoridade, expressões que tenham a forma lógica de proposição mas não de valor de crença. Certas proposições podem ser descritas como imperativas.

Para ele, a construção da máquina-criança deve prever que tão logo encontre um imperativo, este deva ser executado imediatamente. O imperativo deve fazer o papel daquele que impõe uma certa hierarquia entre as regras a serem executadas, pois, quando se utiliza o sistema lógico, existe um grande número de caminhos para serem seguidos e a diferença se faz pela escolha do caminho mais

apropriado. Algumas proposições podem ser dadas às máquinas como autoridade ou a própria máquina pode produzi-las por indução.

Neste ponto, devemos refletir: desejamos construir uma máquina que seja inteligente ou apenas que simule alguns aspectos inteligentes? Como pode uma máquina aprender ou produzir qualquer outra ação, se as regras embutidas nela são imutáveis? Turing (1973, p.81) escreve: "A explicação do paradoxo está em que as regras que mudam, nos processos de aprendizagem, são de caráter menos pretensioso, aspiram apenas a uma validade efêmera".

Para Turing, o professor da máquina-criança pode, até certo ponto, prever o seu comportamento, apesar de ignorar o que se passa em seu interior. A atitude inteligente é aquela que promove um pequeno desvio de comportamento disciplinado, que resulta em computação, porém não sendo aleatório e sem objetivo. A aleatoriedade é um elemento bastante útil na resolução de problemas, e pode ser um fator muito interessante para ser incluída na máquina que está aprendendo, ao passo que o método sistemático sequencial não se traduz numa boa escolha, pois podemos experimentar um grande leque de situações em que não exista a solução desejada.

O processo de ensino-aprendizagem pode ser considerado como uma busca de mudança de comportamento da máquina, que satisfaça ao instrutor, ou algum outro critério. Esse processo, ao ser modelado por uma máquina, estará sendo apresentado como um candidato à explicação do comportamento inteligente. É

justamente a técnica de elaboração de modelos, como a ilustrada acima, que interessa à Ciência Cognitiva.

Em resumo, a construção de modelos tem desempenhado um papel muito importante para a ciência, e mais precisamente para a Ciência Cognitiva. Nesse caso, o objeto de estudo é a mente e não se conhece outra maneira de estudá-la a não ser construindo modelos.

Duas posições relacionadas à modelização da mente se contrapõem: a primeira, com o trabalho de McCulloch e Pitts mostrando que as conexões entre as células nervosas poderiam ser modeladas em termos da lógica e das operações de impulsos e não-impulsos, que poderiam ser comparadas às operações do cálculo proposicional. Dessa forma, podemos afirmar que a mente e o cérebro teriam a mesma caracterização, para os dois autores. A segunda, no trabalho apresentado por Turing, em 1950, em que ele propõe “ensinar” uma máquina, como se estivesse ensinando uma criança, considerando apenas:

- o estado inicial da mente, no nascimento;
- a educação recebida;
- a experiência à qual a mente foi submetida.

Para Turing, o processo de ensino da máquina-criança não tem por objetivo concretizar ações por parte da máquina e sim dotá-la de uma certa inteligência, de modo que possa simular alguns aspectos da mente humana. Segundo o autor, mente e cérebro apresentam aspectos distintos: no estudo da inteligência, apenas o seu substrato lógico, imaterial, é relevante.



Voltaremos a tratar dos processos de aprendizagem no Capítulo 4, quando discutiremos, particularmente, as técnicas e mecanismos de treinamento e aprendizagem nas Redes Neurais Artificiais.

Na próxima Seção, estudaremos o conceito de algoritmo e sua contribuição para a modelização de alguns aspectos considerados inteligentes.

SEÇÃO 2 - CONCEITO DE ALGORITMO

Começaremos esta Seção analisando um conceito fundamental relacionado à Máquina de Turing que é o de *algoritmo*. Na linguagem natural, algoritmo pode ser definido como:

1. Mat. Processo de cálculo, ou de resolução de um grupo de problemas semelhantes, em que se estipulam, com generalidade e sem restrições, regras formais para a obtenção do resultado, ou da solução do problema.
2. Proc. Dados. Conjunto predeterminado e bem definido de regras e processos destinados à solução de um problema, com um número finito de etapas. (Ferreira, 1988, p.30)

Segundo Hodges (1992), Alan Mathison Turing nasceu na cidade de Londres em 23 de junho de 1912. Aluno aplicado, interessou-se pela ciência desde o início de sua juventude. Turing consagrou-se como matemático aos 24 anos de idade, quando mostrou ser possível executar operações computacionais sobre a teoria dos



números por meio de uma máquina que tivesse nela embutida as regras de um sistema formal. Na Seção 3, estudaremos a Máquina por ele concebida e suas implicações para a computação.

Animado com os resultados obtidos pelos projetos dos primeiros computadores (Eniac, em 1944, e Colossus, em 1944-45) Turing escreve um texto em 1946 sobre a proposta de construção do Automatic Computing Engine-ACE. Nesse texto, ele apresenta aos patrocinadores do projeto as características e composições da máquina, bem como uma comparação do trabalho desta com o do ser humano. Quanto às máquinas, ele sugere que processos mais complexos podem ser realizados com maior rapidez e com menor probabilidade de erros, quando comparados com o trabalho desenvolvido pelo homem.

A máquina ACE teria uma memória mecânica com os seguintes requerimentos:

- a) Um dispositivo para imprimir os resultados;
- b) As instruções que, quando “carregadas” na memória, processariam o programa;
- c) As tabelas de funções aritméticas, que dão condições à máquina de realizar as operações matemáticas.

Ainda nesse mesmo texto, Turing escreve que a máquina deverá ter uma grande quantidade de memória e que estudos posteriores deveriam contemplar métodos de alocação e economia de memória.

Como vimos até aqui, Turing tinha uma boa idéia dos problemas que a ciência da computação iria enfrentar nos próximos anos. Um desses problemas era encontrar dispositivos teóricos e físicos que possibilitassem o armazenamento e recuperação de grandes quantidades de informação na memória dos computadores. Isso implicaria uma das limitações dos computadores atuais para servirem como modelos de atividades cognitivas que exijam armazenamento e recuperação de grande quantidade de informação.

As características das duas vertentes computacionais existentes dentro da Ciência Cognitiva já foram estudadas no Capítulo 1; agora, interessa-nos estudar a caracterização dos mecanismos que levaram ao desenvolvimento teórico da computabilidade, qual seja o conceito de algoritmo.

Através do exemplo abaixo, podemos ter uma idéia da aplicação do conceito de algoritmo: quando fazemos uma operação matemática, uma adição, por exemplo, estamos utilizando regras ou métodos para encontrar a soma de dois números. Tais métodos são exemplos de algoritmos. A operação realizada para somar dois números quaisquer pode ser feita desta forma:

- a) Obtenha o primeiro número;
- b) Obtenha o segundo número;
- c) Some os dois números;
- d) Mostre o resultado.

A partir desse exemplo, podemos delimitar a definição de algoritmo.



Um algoritmo ou procedimento efetivo é definido e exemplificado por Cutland (1980), como uma regra mecânica ou um método automático:

“More generally, an algorithm or effective procedure is a mechanical rule, or automatic method, or programme for performing some mathematical operation. Some more examples of operations for which easy algorithms can be given are

(a) given n , finding the n^{th} prime number;

(b) differentiating a polynomial;

(c) finding the highest common factor of two numbers (the Euclidean algorithm);

(d) given two numbers x, y deciding whether x is a multiple of y ”
(Cutland, 1980, p.7).

Ainda segundo Cutland, um algoritmo pode ser informalmente representado como segue:

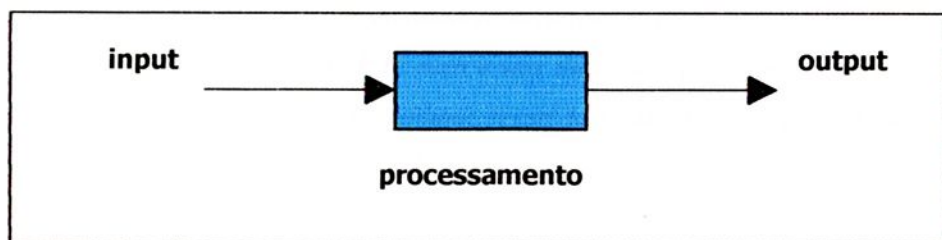


Figura 3 - Representação analógica de uma regra mecânica (Cutland, 1980, p. 2).

Onde: Input $\Rightarrow$ é o dado de *entrada* ;

Output $\Rightarrow$ é o produto obtido pelo *processamento* a partir do *input*;

Processamento $\Rightarrow$ contém o algoritmo que transforma o *input* em *output*.

Uma outra definição de algoritmo é dada por Rogers (1987), que o define como um procedimento determinístico composto por um número finito de regras, executadas sequencialmente, num espaço finito de tempo:

“ Roughly speaking, an algorithm is a clerical (i.e., deterministic, book-keeping) procedure which can be applied to any of a certain class of symbolic inputs and which will eventually yield, for each such input, a corresponding symbolic output.

...

**1. An algorithm is given as a set of instructions of finite size.*

**2. There is a computing agent usually human, which can react to the instructions and carry out the computations.*

**3. There are facilities for making, storing, and retrieving steps in a computation.*

**4. Let P be a set of instructions as in *1 and L be a computing agent as in *2. Then L reacts to P in such a way that, for any given input, the computation is carried out in a discrete stepwise fashion, without use of continuous methods or analogue devices.*

**5. L reacts to P in such a way that a computation is carried forward deterministically, without resort to random methods or devices. ...”*

(Rogers, 1987, p. 1-3)

As definições de algoritmo apresentadas acima: aquelas obtidas através do senso-comum, bem como as de Cutland e Rogers, de certa forma, convergem para o mesmo ponto. Elas mostram que o termo algoritmo é um procedimento mecânico, determinístico e regido por regras.

Veremos, na próxima Seção, a Máquina proposta por Turing e sua relação com o algoritmo.



SEÇÃO 3 - MÁQUINA DE TURING

Turing construiu uma máquina teórica que, metaforicamente, pode ser descrita como consistindo de uma fita e uma “cabeça” de leitura e escrita, que “obedece” aos comandos de um algoritmo, tal como descrito na Seção 2. A fita é infinita (para a esquerda e para a direita) e unidimensional, e é dividida em uma sequência de quadrados, cada um capaz de conter símbolos de um alfabeto β . Para efeito de ilustração, consideraremos o alfabeto $\beta = \{0,1\}$.

A “cabeça” de leitura/escrita é um dispositivo que a cada momento está visualizando um único quadrado da fita e obedece a uma das instruções do programa, sequencialmente.

A fita, apesar de ser conceitualmente infinita em comprimento, conterá um número finito de quadrados não vazios. Então, ela pode ser descrita de uma maneira finita, e a parte não vazia pode crescer ou diminuir durante a computação. As “ações” da máquina são determinadas pelo seu programa (algoritmo).

Arbib (1987) caracteriza a Máquina de Turing da seguinte maneira:

- a) A máquina possui um **controle**, que armazena um programa com um número finito de instruções;
- b) Uma fita, infinita, dividida em partes pequenas, é utilizada pela máquina para armazenar as entradas, os resultados de operações intermediárias e as saídas;

c) Um dispositivo de procura e escrita que percorre toda a extensão da fita, atendendo às instruções do controle.

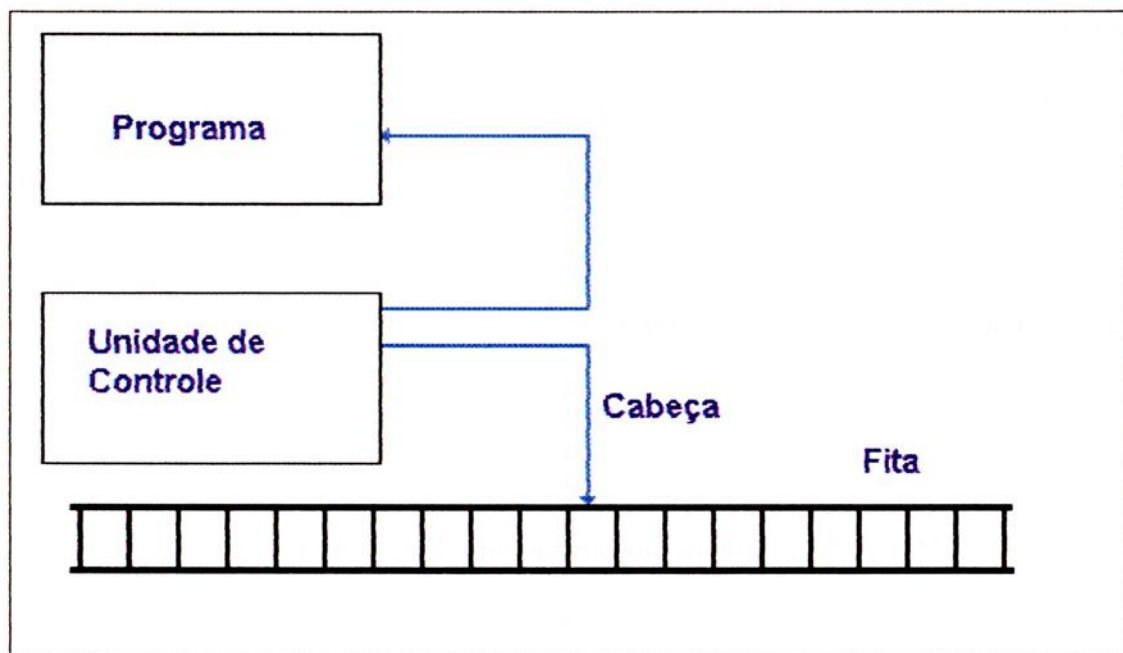


Figura 4 - Máquina de Turing (Arbib, 1987, p.128).

Nesta Seção, estudamos, primeiramente, a caracterização da Máquina de Turing proposta por Rogers (1987), que informalmente consiste também de uma fita unidimensional, potencialmente infinita nas duas direções (esquerda e direita), dividida em quadros. A máquina pode desenvolver as seguintes operações:

- 1) Escrever “1” no quadro atual, se nele não existir o elemento “1”;
- 2) Escrever “0” no quadro atual, se nele não existir o elemento “0” (originalmente Rogers utiliza um “espaço em branco”, preferimos o 0 (zero) por uma questão didática);
- 3) Movimentar a “cabeça” de leitura e escrita, um quadro para a esquerda;

4) Movimentar a “cabeça” de leitura e escrita, um quadro para a direita.

A máquina é capaz de assumir um número finito de estados internos distintos, de acordo com a instrução executada por ela num dado momento. Para Rogers (1987), o conjunto de regras que determina o comportamento da máquina num dado momento é o *conjunto de quádruplas*. O conjunto que determina uma quádrupla é composto de:

$Q_i \rightarrow$ Estado interno inicial (de um conjunto finito de estados) que a máquina pode assumir

$$i = 1, 2, 3, \dots, n;$$

$S_j \rightarrow$ Conteúdo do quadro visualizado, pertencente ao alfabeto $\beta = \{0, 1\}$;

$S_k \rightarrow$ Desloca um quadro para a direita se $S_k = R$

Desloca um quadro para a esquerda se $S_k = L$

Escreve 0 se $S_k = 0$, ou escreve 1 se $S_k = 1$;

$Q_f \rightarrow$ Estado interno final (de um conjunto finito de estados) que a máquina pode assumir

$$i = 1, 2, 3, \dots, n.$$

Desse modo, temos a seguinte quádrupla genérica: $Q_i S_j S_k Q_f$.

Originalmente, Turing (1965) definiu a computação da Máquina Universal da seguinte forma, representada na figura 5:

Configuration		Behavior	
m-config.	symbol	operations	final m-config
b	None	P0, R	c
c	None	R	e
e	None	P1, R	f
f	None	R	b

Figura 5 – Máquina Universal (Turing, 1965, p.119)

Como assumimos que apenas um quadrado da fita é visualizado a cada instante, a operação seguinte a ser executada é determinada pela configuração interna e pelo símbolo visualizado no momento.

Cutland (1980) mostra também que cada quádrupla especifica a ação a ser executada pela máquina.

A computação tem início quando a máquina, partindo de um estado inicial, posiciona sobre a fita a "cabeça" de leitura/escrita e visualiza um símbolo do alfabeto β . A partir desse momento, um conjunto de quádruplas determinará as "ações" que a máquina deverá executar. A condição de parada da máquina será satisfeita quando não existir quádrupla que especifique a próxima "ação".

Para caracterizarmos uma computação executada por Máquina de Turing, convencionaremos que, para a entrada (input) de valores:

o número 0 equivale a 1

o número 1 equivale a 11

o número 2 equivale a 111

o número 3 equivale a 1111, etc

e para as saídas (output), o número de 1's corresponde a:

nenhum 1	equivale a 0
1	equivale a 1
2	equivale a 11
3	equivale a 111, etc.

Seja um conjunto de quádruplas P:

$Q_1 0 R Q_1$

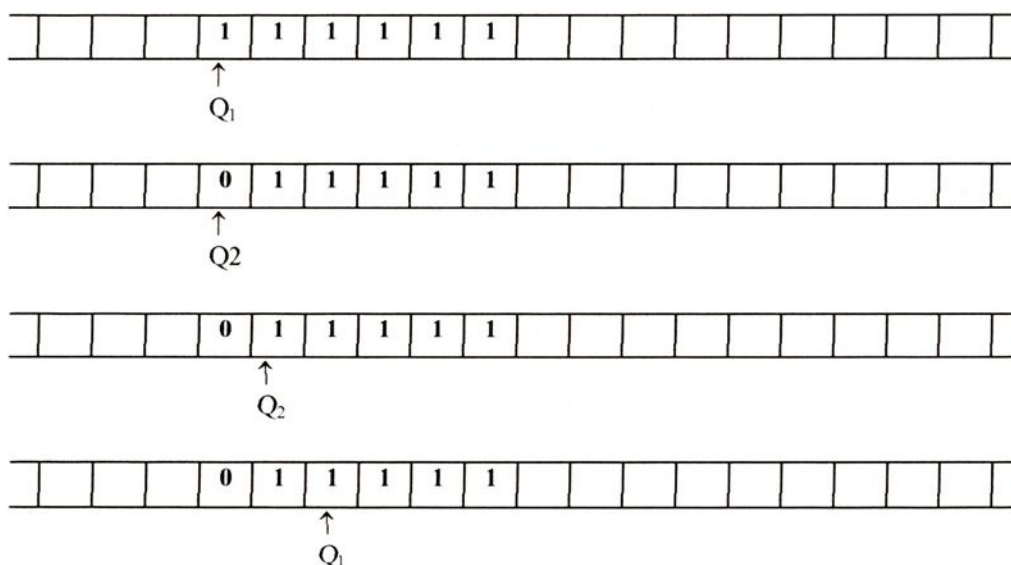
$Q_1 1 0 Q_2$

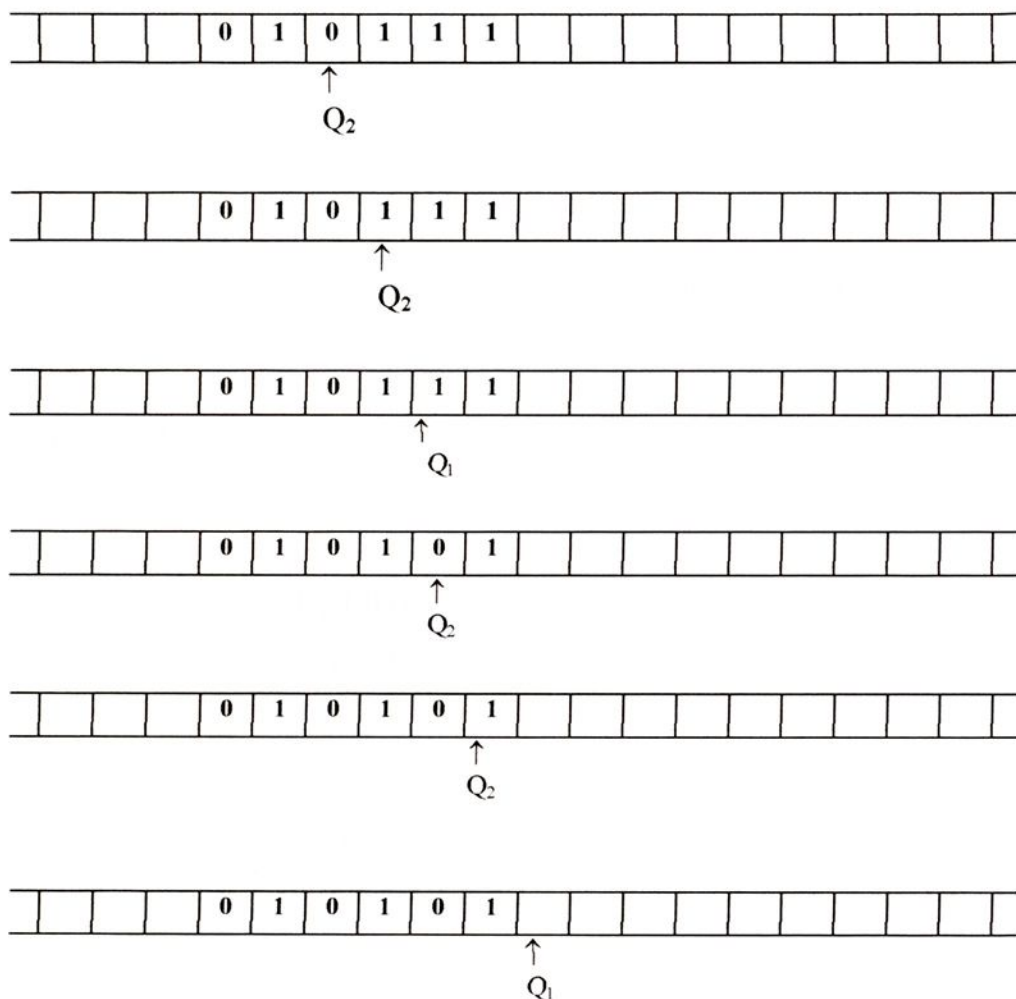
$Q_2 0 R Q_2$

$Q_2 1 R Q_1$

Extraído de Cutland (1980, p.55)

que especifica a seguinte sequência: 0 1 0 1 0 1 etc. Suponha a fita inicial:





A máquina irá parar quando encontrar o primeiro quadro em branco e as quádruplas não especificarem essa situação. Ela não pararia, neste exemplo, se a fita inicial tivesse sido especificada da seguinte forma: 0 1 0 1 0 1 etc.; neste caso, diríamos que a execução da máquina estaria num *loop* infinito.

Vejamos um outro exemplo, onde a máquina P soma dois números quaisquer:

$$Q_1 1 0 Q_1$$

$$Q_1 0 R Q_2$$

$$Q_2 1 0 Q_3$$

$$Q_2 0 R Q_2$$

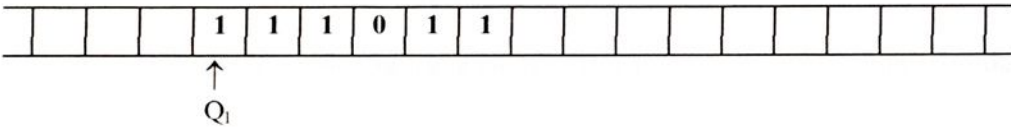
Somaremos, por exemplo, 2 e 1. Utilizaremos o número 0 entre as parcelas como separador.

Entrada: → 111

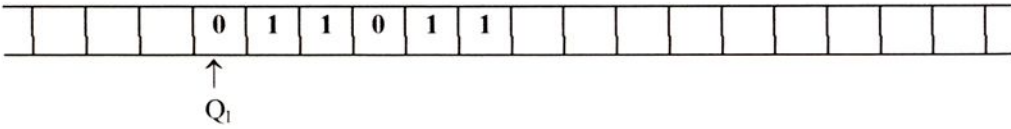
→ 11

Na fita:

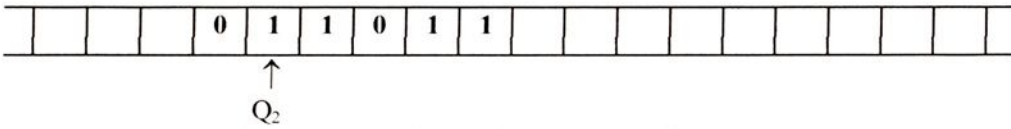
1ª quádrupla: $Q_1 1 0 Q_1$



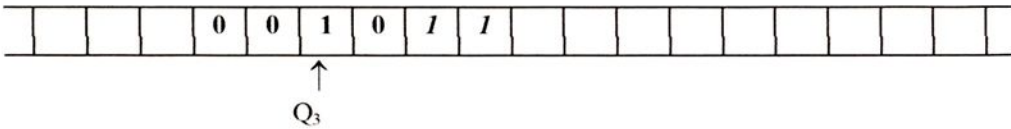
2ª quádrupla: $Q_1 0 R Q_2$



3ª quádrupla: $Q_2 1 0 Q_3$



resultado da 3ª quádrupla



Saída: 111 equivale a 3

A máquina pára, pois não existe instrução de quádrupla que especifique a próxima ação.

Através desse exemplo, procuramos salientar o conceito de Máquina Universal proposta por Turing, que tem como elemento fundamental a resolução de problemas. Essa máquina teórica traz consigo a proposta de que alguns problemas podem ser convertidos numa forma clara e bem definida, de modo a determinar um resultado.

Na Seção 4, a seguir, discutiremos a noção da Máquina Registradora Ilimitada (URM) de modo, agora, a caracterizarmos a noção de computabilidade.

SEÇÃO 4- MÁQUINA REGISTRADORA ILIMITADA

Após a idéia original de Turing, outros autores também desenvolveram *máquinas* abstratas que, manipuladas por algoritmos, representavam a noção de computabilidade.

Cutland (1980, p.9) apresentou a URM (Unlimited Register Machine), que possui a seguinte configuração:

R₁ R₂ R₃ R₄ R₅ R₆ ...

r ₁	r ₂	r ₃	r ₄	r ₅	R ₆ ...
----------------	----------------	----------------	----------------	----------------	--------------------

Cada quadro, também chamado de registro, possui uma identificação representada pela letra R (maiúscula) seguida de um índice. O quadro R_n individualiza o elemento representado pela letra r (minúscula), seguida de um índice.

A fita, infinita numa direção, é manipulada por um conjunto de instruções (algoritmo) que determina o *comportamento da máquina*. Convencionaremos que cada linha de instrução, que compõe o algoritmo, deverá ser numerada a partir do elemento I_1 até o elemento I_n .

Quatro tipos de instruções, em geral, caracterizam um algoritmo que executa uma URM:

(i) Instrução Zero (Z): tem por finalidade trocar o conteúdo do registro n por 0.

Abreviadamente, podemos escrever essa instrução através da seguinte forma geral:
 $Z(n)$

(ii) Instrução Sucessor (S): tem por finalidade incrementar o conteúdo do registro n de uma unidade.

Abreviadamente, podemos escrever essa instrução através da seguinte forma geral:
 $S(n)$

(iii) Instrução Transferência (T): tem por finalidade transferir o conteúdo do registro R_n para o registro R_m , não alterando o conteúdo de R_n .

Abreviadamente, podemos escrever essa instrução através da seguinte forma geral:
 $T(n,m)$

(iv) Instrução Salto (J): tem por finalidade desviar a execução sequencial das instruções para uma outra parte do algoritmo, obedecendo:

- Se $r_m = r_n$ desviar para a instrução I_q do algoritmo;
- Se $r_m \neq r_n$ executar a próxima instrução.

Abreviadamente, podemos escrever essa instrução através da seguinte forma geral:
 $J(m,n,I_q)$

Cutland (1980, p. 21) analisa várias funções que são computáveis pela URM. A seguir, descreveremos duas delas. Convencionaremos que o resultado determinado pela máquina estará em R_1 :

(i) $f(x) = 5$

- I_1 $J(1, 2, 8)$
- I_2 $S(2)$
- I_3 $J(1, 2, 8)$
- I_4 $J(1, 1, 2)$

R_1	R_2	R_3	R_4	R_5	R_6	...
5	0	0	0	0	0	
5	1	0	0	0	0	
5	2	0	0	0	0	
5	3	0	0	0	0	
5	4	0	0	0	0	
5	5	0	0	0	0	

Enquanto o conteúdo de R_1 for diferente de R_2 o algoritmo incrementa o registro R_2 de 1. Quando isto acontecer (instrução I_3), o processamento é desviado para a instrução I_8 . Como essa instrução I_8 não existe, a máquina pára.

$$(ii)f(x,y) = \begin{cases} 0 & \text{se } x = y \\ 1 & \text{se } x \neq y \end{cases}$$

$I_1 \quad J(1, 2, 5)$

$I_2 \quad Z(1)$

$I_3 \quad S(1)$

$I_4 \quad J(1, 1, 6)$

$I_5 \quad Z(1)$

Para $x \neq y$: Suponha $x = 4$ e $y = 2$

R_1	R_2	R_3	R_4	R_5	R_6	...
4	2	0	0	0	0	
0	2	0	0	0	0	
1	2	0	0	0	0	

Para $x = y$: Suponha $x = 5$ e $y = 5$

R1	R2	R3	R4	R5	R6	...
5	5	0	0	0	0	
0	5	0	0	0	0	

Assumindo qualquer valor para x em R_1 e qualquer valor para y em R_2 , o algoritmo verifica se x é igual a y em I_1 . Em caso afirmativo a instrução I_5 troca o conteúdo de R_1 por 0 e a URM pára.

Se x for diferente de y , a instrução I_2 coloca 0 em R_1 , I_3 soma um em R_1 e I_4 desvia o processamento para a instrução I_6 . Como ela não existe, a URM pára.

Esses exemplos ilustram de forma intuitiva a computabilidade que, segundo Cutland (1980), pode ser definida como tendo um algoritmo que é utilizado para calcular valores numéricos sobre funções:

“When an algorithm or effective procedure is used to calculate the values of a numerical function then the function in question is described by phrases such as effectively calculable, or algorithmically computable, or effectively computable, or just computable” Cutland (1980, p.8).

Computabilidade é definida, também pelo senso-comum (Ferreira, 1988, p.165), como: “Computável. o que pode ser computado. Computar. 1. Fazer o cômputo de; contar. 2. Calcular, orçar”.

Essas duas definições mostram que *computabilidade* é calcular, ou melhor, encontrar o resultado daquilo a que foi proposta a verificação.

Em resumo, as definições de algoritmo apresentadas na Seção 2 da presente dissertação convergem para o mesmo ponto, pois mostram o mecanicismo e o determinismo característicos desse termo.

A Máquina Universal, proposta por Turing e estudada na Seção 3, tem como elemento fundamental a resolução de problemas. Essa máquina, teórica, traz consigo a proposta de que alguns problemas podem ser convertidos numa forma clara e bem definida, de modo a determinar um resultado.

No artifício teórico, proposto por Cutland, na Seção 4, a Máquina Registradora Ilimitada (URM) é equivalente à Máquina de Turing, pois também tem por finalidade a resolução de problemas, determinando a computabilidade ou não de funções matemáticas.

As “máquinas” estudadas neste Capítulo podem realizar, em princípio, qualquer tipo de cálculo, desde que previamente preparadas para isso.

A contribuição do trabalho de Turing para toda a ciência da computação foi mostrar que existe um mecanismo universal, que pode simular qualquer situação matemática, desde que haja um formalismo descrito para tal simulação.

Consideramos importante ressaltar que a elaboração de técnicas de modelagem, através da Máquina de Turing na Ciência Cognitiva, só foi possível devido à hipótese de que há um elemento comum, fundamental entre máquinas e seres humanos. Esse elemento é justamente a habilidade de processar informação. Os humanos e as máquinas, nas suas atividades inteligentes, manipulam informação.



Mas o que deve ser entendido por informação nesse contexto? Esta questão será objeto de estudos do próximo Capítulo.



CAPÍTULO 3

TEORIA DA INFORMAÇÃO

INTRODUÇÃO

Neste Capítulo, desenvolvemos um breve histórico da Teoria da Informação. Não pretendemos, entretanto, propor uma definição do que seja Informação, pois, como veremos nas próximas páginas, trata-se de um conceito complexo, cuja análise nos leva invariavelmente a caminhos repletos de imprecisões e controvérsias.

O capítulo divide-se em três seções. Na primeira, apresentamos um histórico da Teoria da Informação. Na segunda, estudamos a Teoria da Informação proposta por Shannon & Weaver. Finalmente, na Seção 3, sugerimos um elo entre a Teoria da Informação e a Ciência Cognitiva.

SEÇÃO 1 - HISTÓRICO DA TEORIA DA INFORMAÇÃO

Entre as inúmeras tentativas de definir o conceito de *informação*, encontramos algumas caracterizações desenvolvidas por cientistas e filósofos. Assim, Ferreira (1988, p.361) define informação, no senso-comum, como “Ato ou efeito de informar (-se); informe. Dados acerca de alguém ou de algo. Conhecimento, participação. Instrução, direção”.



Uma segunda caracterização, proposta por Wiener (1961, p.132), está preocupada com a natureza ontológica da informação: "Information is information, not matter or energy. No materialism which does not admit this can survive at the present day". Aqui, Wiener está sugerindo que a informação possui uma natureza própria, que não se reduz à matéria ou à energia.

Um terceiro sentido do termo informação relaciona-o com mensagem carregada de significado:

"Frequently the messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities" (Shannon, 1949, p.31).

Existe um quarto sentido do conceito de informação, que não envolve aspectos semânticos, que constitui o núcleo da Teoria da Comunicação:

"These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one selected from a set of possible messages. The system must be designed to operate for each possible selection, not just the one which will actually be chosen since this is unknown at the time of design" Shannon (1949, p.31).

Nesse sentido, o elemento fundamental é a possibilidade de se selecionar uma mensagem entre um conjunto de possíveis sinais.



O interessante é que o conceito *informação* é explicado por outros conceitos, entendidos apenas vagamente, que por sua vez podem também ser ilustrados por sinônimos ou por aquilo que não seria a própria definição de informação. Assim é que Küppers citado por Gonzales (1996, p.54) refere-se à noção de informação cujo entendimento nos escapa:

...o entendimento da noção de informação talvez tenha que ser deixado para o futuro. Assim como o homem da idade do bronze e do ferro, lidavam com estes elementos mas não dispunham de estruturas conceituais apropriadas para defini-los, também nós, habitantes da era da informação, teremos que aguardar o desenvolvimento das ciências para podermos ir além das metáforas na descrição deste conceito.

Quaisquer que sejam as perspectivas da noção de informação, estudaremos sua abordagem quantitativa, levando em consideração as definições propostas por Shannon & Weaver.

Segundo Breton (1991), as preocupações sobre as origens da noção de informação aparecem sob três aspectos:

- a) Nas tentativas de distinguir entre a forma e o sentido;
- b) Na necessidade de aprimorar técnicas de transmissão de mensagem e;
- c) Nas pesquisas sobre a natureza do raciocínio correto e sobre as considerações acerca da verdade dos enunciados.

O ponto chave destas preocupações foi apresentado neste trabalho, no Capítulo 2, onde discutimos a noção de algoritmo formulada por Alan Turing.



No que diz respeito à distinção entre Forma e Sentido, Breton (1991) ressalta que é comum se tentar distinguir *sentido e forma* de uma mensagem sem a preocupação com os fatos. Essa distinção possibilitará um conhecimento técnico e matemático a respeito da *forma*, transformando o conhecimento moderno de "processamento de informação" e comunicação à distância.

A informação, entendida como mensagem, pode ser caracterizada como um conjunto de sinais ou símbolos transmitidos a alguém, ou a um grupo de pessoas. Um bom exemplo para ilustrar essa situação é quando uma pessoa deseja transmitir um telegrama. O funcionário dos correios faz a leitura do texto a ser transmitido, sem interessar-se pelo sentido do que está escrito. A fim de estabelecer o preço do serviço prestado, ele considera apenas os símbolos que a mensagem contém. Os símbolos, independentemente de sua significação, deverão ser processados e transformados em sinais telegráficos. Tais símbolos constituem a *forma* que compõe a mensagem.

Atualmente ao termo *informação* foi atribuído um sentido amplo e ao mesmo tempo vago. O termo designa *informação* no sentido jornalístico e *informação* no sentido técnico, o que pode causar, confusão algumas vezes. A informação nos meios de comunicação está carregada de sentido, sendo freqüentemente pouco objetiva. A informação processada pela técnica (computador, telefone, telégrafo etc.) é, em geral, destituída de sentido, pelo menos enquanto está sendo processada pela máquina; apenas sua forma é relevante para o técnico.



Um outro fator que ilustra a distinção entre *forma e sentido* é o fato de que, como suporte físico de uma mensagem, uma *forma* pode ser decomposta em símbolos (forma da mensagem) e sinais (transformação da mensagem para transmissão); um sentido, por sua vez, não pode ser facilmente decomposto.

Quanto à transmissão de mensagem, a história da humanidade mostra os esforços para estabelecer a comunicação à distância, desde o tempo em que para tal comunicação utilizavam-se os sinais de fumaça, os tambores ou os mensageiros a pé.

O quadro abaixo, compilado de Breton (1991), faz um recorte da evolução da comunicação à distância entre os séculos XVIII - XX, bem como a contribuição de alguns cientistas para a teoria da informação.

ANO	INVENTOR	INVENTO/CONTRIBUIÇÃO	FUNCIONAMENTO
1794	Claude Chappe	Telégrafo aéreo	Braços móveis montados sobre torres. Eficaz para mensagens breves.
1832	Samuel F. B. Morse	Telégrafo elétrico	Pontos e traços representavam números que correspondiam a palavras de um dicionário.
1838	Samuel F. B. Morse	Telégrafo elétrico	Correspondência direta dos traços e pontos com as letras do alfabeto.
1874	Thomas A. Edson	Aperfeiçoamento do Telégrafo	Possibilitou transmitir simultaneamente duas mensagens num mesmo circuito elétrico.
1875	Alexandre G. Bell	Telefone	Transmissões de sinais através de corrente elétrica.
1888	Heinrich Hertz	Geração de ondas elétricas por meio de descargas oscilatórias de um condensador.	~ ~ ~
1900	Landell de Moura	Primeira transmissão de voz humana por irradiação de onda eletromagnética.	~ ~ ~
1927	R.V.L. Hartley	Propôs uma medida exata de informação.	~ ~ ~
1949	Claude Shannon e Warren Weaver	Teoria Matemática da Comunicação.	~ ~ ~

Duas situações contribuíram para o avanço dos estudos e da tecnologia na área de transmissão da informação:

- O domínio e a utilização da eletricidade, que permitiu rapidez na transmissão de mensagens;
- Os estudos e pesquisas na articulação entre o sinal e o símbolo, resultando no entendimento das condições físicas para transmissão de mensagens.

Hartley apresentou, no Congresso Internacional de Telegrafia e Telefonia, em 1927, um estudo pioneiro intitulado *Transmissão de Informação*. No ano seguinte, quando foi publicado seu trabalho, Hartley (1928) propôs estabelecer uma medida quantitativa de informação, segundo a capacidade de transmissão dos sistemas telegráficos, telefônicos, radiofônicos e televisivos. Essa medida de informação deve auxiliar na melhoria de desempenho ou verificar falhas de operação dos equipamentos citados anteriormente.

Um fator importante, que age, em geral, negativamente na transmissão de mensagem e que deve ser considerado, é a interferência externa ou *ruído*. Essa interferência acaba influenciando, de alguma forma, o desempenho do aparato de transmissão. O emissor assume estar desconsiderando o papel da interferência durante a transmissão de mensagem e concentra-se no desempenho do sistema em si. Aqui, Hartley não está interessado no papel do ruído durante a transmissão de uma mensagem e sim na medida da quantidade de informação transmitida por um sistema.

Hartley (1928) entende que a capacidade de um sistema transmitir mensagem está diretamente relacionada à sua capacidade de traduzir a transmissão numa medida quantitativa de informação. Dessa forma, ele faz algumas considerações sobre a comunicação de informação:

- existem símbolos (pontos, palavras, traços etc.) que fazem parte de um vocabulário convencionado, transmitem significado entre as partes comunicantes. Por exemplo, numa situação de comunicação, o emissor, enquanto fala, utiliza-se de artifícios, como alteração da voz, para chamar a atenção do receptor num dado momento;
- através de sucessiva seleção, uma sequência de símbolos é transmitida ao receptor da mensagem. Por exemplo, Hartley (1928, p. 536) escreve: "Apples are red". Nessa sentença, outros tipos de fruta e objetos são eliminados. A atenção está dirigida para as maçãs e suas propriedades;
- a possibilidade de investigação sobre outras cores de maçã (a não ser a cor vermelha) é descartada. A informação torna-se mais precisa quando a seleção é feita continuamente e sequências de símbolos são eliminados.

Entretanto, no exemplo acima, existe a possibilidade de se questionar sobre tamanhos diferentes de maçãs, o que poderá acontecer com novas seleções de informação. Após sucessivas seleções, espera-se encontrar um número final de sentenças que deverá exprimir a medida quantitativa de informação. Porém, o número de símbolos disponíveis depende de vocabulários ou línguas existentes. Por exemplo, para duas pessoas que falam línguas distintas, o número de símbolos significativos é



desprezível, quando comparadas com pessoas que falam o mesmo idioma. Dessa forma, Hartley (1928) propõe eliminar os fatores psicológicos envolvidos nessa situação e estabelecer, como foi dito no início da apresentação do trabalho deste autor, uma medida puramente física de informação. Ele ressalta que para estimar a capacidade de transmissão de informação devemos desprezar qualquer tipo de interpretação pessoal e proceder à seleção arbitrária de mensagem. Isso significa que fatores psicológicos devem ser eliminados na medida quantitativa de informação. A definição, segundo Hartley (1928, p. 540), da quantidade de informação associada à seleção de uma mensagem é o logaritmo do número de símbolos disponíveis:

$$H = \log S$$

onde, $H \Leftrightarrow$ quantidade de informação

$S \Leftrightarrow$ número de símbolos de uma mensagem

Por exemplo, se o número de símbolos disponíveis, num sistema qualquer, é 2, o conteúdo de informação associado a sua seleção é:

$$H = \log 2$$

$$H = 1$$

Numa situação semelhante, agora com 8 seleções, teremos:

$$H = \log 8$$

$$H = 3$$

Assim, nessa abordagem proposta por Hartley, a informação é uma medida quantitativa que está presente (ou não) numa mensagem. No que se segue, indicaremos como as idéias de Hartley foram desenvolvidas por Shannon.

SEÇÃO 2 - TEORIA DA INFORMAÇÃO

Nesta Seção, apresentamos a Teoria Matemática da Comunicação, conhecida como Teoria da Informação, trabalho pioneiro de Shannon & Weaver. A obra de Claude Shannon e Warren Weaver, *The Mathematical Theory of Communication*, foi primordial para assentar a teoria da informação, na primeira metade deste século.

A teoria da informação seria, provavelmente, menos difundida entre os cientistas se Weaver não a tivesse reformulado e tornado mais acessível no estilo e situando-a no contexto científico, relacionando-a ao conceito de comunicação. Inicialmente, Weaver utiliza o termo *comunicação* num sentido amplo, sugerindo que ele inclui procedimentos nos quais a mente de uma pessoa pode afetar a mente da outra, envolvendo não somente a escrita, mas todo tipo de comportamento humano (Weaver, 1949, p.3).

Segundo Weaver (1949, p.4), a comunicação apresenta três tipos de problemas:

- 1–O problema técnico: focaliza a exatidão com que a informação é transmitida;
- 2–O problema semântico: focaliza a precisão do significado da informação entre o remetente e o receptor;
- 3–O problema de eficácia: focaliza a eficácia da informação sobre o comportamento do receptor.



Os problemas 2 e 3, apresentados acima, podem fazer uso da informação apenas quando ela tiver sido analisado 1. A importância da análise do Problema Técnico revela que este se sobrepõe os demais, pois não existirá informação em 2 e 3, se 1 perder a integridade.

De acordo com Shannon & Weaver (1949), a informação será concebida em termos de transmissão de mensagens. Para que as mensagens possam ser transmitidas, precisamos de um sistema de comunicação. Tal sistema é representado pela figura 6:

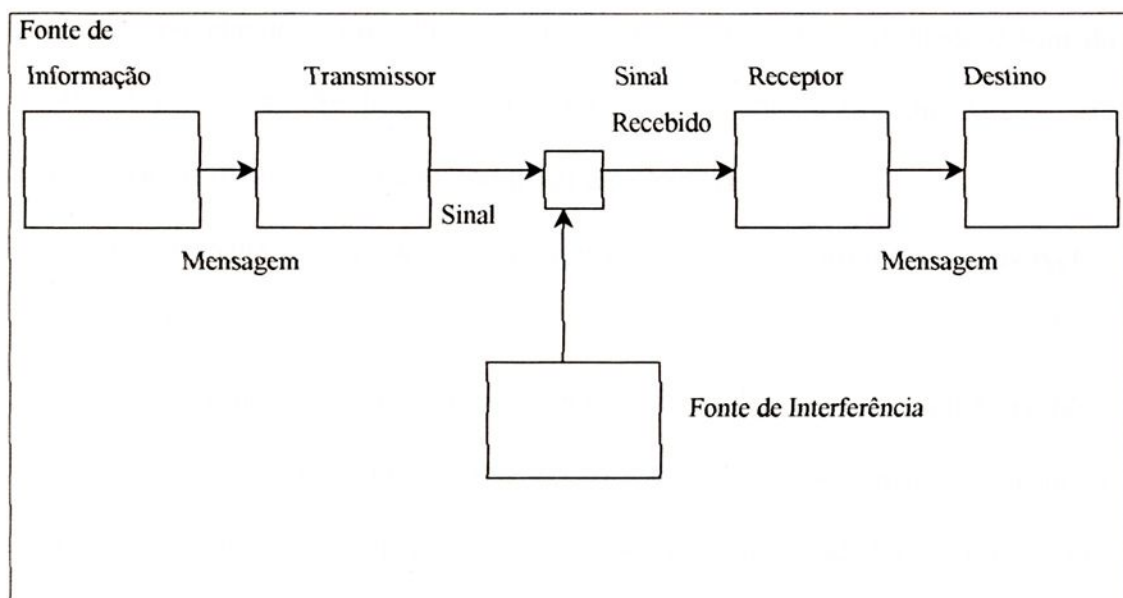


Figura 6 - Sistema de Comunicação (Weaver, 1949, p.7).

Os autores descrevem os elementos que compõem o sistema de comunicação da seguinte forma:

1. Fonte de Informação: é aquela que produz um conjunto de mensagens para serem comunicadas a um destino em particular. A mensagem é selecionada a partir de um

conjunto de possíveis mensagens. Ela pode ser composta por palavras, sons, imagens etc.;

2. Transmissor: produz uma sequência de sinais, de modo que possam ser transmitidos através de um canal. Por exemplo, no caso particular da comunicação entre duas pessoas, quando a primeira diz algo, a fonte de informação é o seu cérebro e o transmissor é sua voz;

3. Canal: transporta o sinal a partir do emissor até o receptor. Durante a transmissão, a informação pode ser perturbada e distorcida por outra fonte, chamada “fonte de ruído”. Completando o exemplo anterior, o canal, por onde se propaga o som do interlocutor, é o ar. Se houver distorções ou erros da mensagem durante sua transmissão, dizemos que o ruído perturbou essa informação;

4. Receptor: executa a operação inversa do transmissor, reconstruindo, se possível, a mensagem original;

5. Destino: é o endereço da mensagem e pode ser uma pessoa ou uma outra entidade. Ela requer conhecimento sobre a fonte de informação e deve ser capaz de entender a mensagem transmitida. Em qualquer caso (pessoa ou outra entidade), o destino deve conhecer o conjunto de sinais disponíveis na fonte.

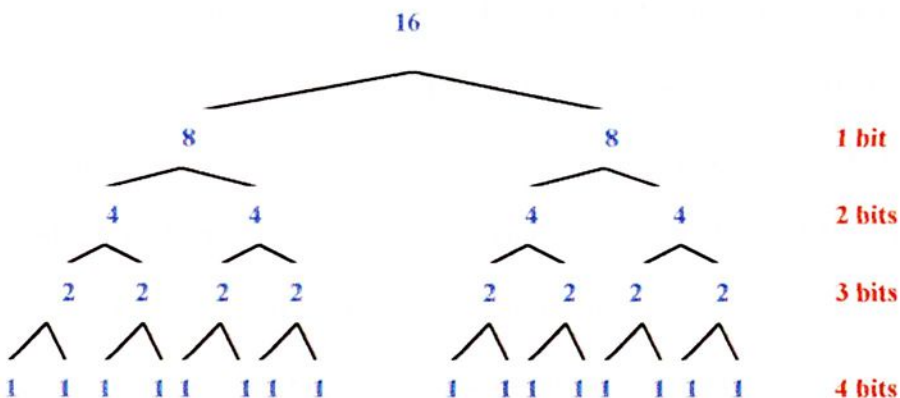
No início da Seção 1, deste Capítulo, definimos o termo informação, segundo o senso- comum, ressaltando que este está sempre ligado ao significado. Shannon & Weaver desvincularam, em seus estudos, a informação da noção de significado. Eles associaram informação ao grau de liberdade de escolha para selecionar uma mensagem. A definição de informação, segundo eles, não está

relacionada a uma mensagem significativa em particular, mas sim à situação em que existe a probabilidade de selecionar mensagens. Para Shannon (1949, p.31), o sistema de comunicação deve ser “construído” de tal maneira que consiga operar para qualquer informação selecionada a partir de um conjunto de possíveis mensagens. Esse conjunto possui, em certos casos, um número finito de mensagens, com a mesma probabilidade de serem selecionadas. Dessa forma, Shannon mostra que é conveniente utilizar o logaritmo de M na base 2, para medir a quantidade de informação numa certa situação. Cada unidade de informação foi denominada “*bit*” (contração do termo em inglês “*binary digit*”), sendo que os números são expressos no sistema binário (0 e 1). Tais elementos podem ser vistos, simbolicamente, como duas alternativas de escolha de informação. O número de alternativas de escolha de mensagens reflete o grau de liberdade dessa escolha, numa certa situação. Quanto maior for a quantidade de situações de escolha possíveis, num dado contexto, maior será a nossa liberdade para escolher alternativas. Resumindo, o *bit* é a unidade utilizada para contar o número de procedimentos seletivos na escolha de alternativas (de igual probabilidade), em que os resultados são reduzidos à metade, a fim de eliminar a incerteza.

Podemos ilustrar uma situação em que o professor deseja que um de seus 16 alunos resolva uma questão de matemática, na lousa. Mas qual aluno ele deve escolher para realizar essa tarefa? Nesse caso, o professor decidiu dividir a turma em dois grupos de oito alunos cada. Consideraremos que existam oito meninos e oito meninas na sala, e que quatro meninos e quatro meninas sempre tiveram bom

desempenho em matemática, e os outros dois grupos de ambos os sexos tiveram de médio a fraco desempenho. Consideraremos, também, que dois meninos e duas meninas de bom desempenho são de baixa estatura e que os outros quatro são altos. Metade dos meninos e meninas de fraco desempenho praticam esporte e a outra metade não pratica esporte algum. Do grupo de meninos e meninas de baixa estatura, metade deles são os mais velhos e a outra metade os mais novos de toda a turma. No grupo dos alunos mais altos, metade deles são bons alunos também em português e a outra metade não é. No grupo dos alunos que praticam esporte, metade deles pratica natação e a outra metade pratica ginástica. Finalmente, nos grupos daqueles alunos que não praticam esporte, metade deles são filhos de pais desempregados e a outra metade possui problemas disciplinares.

Notemos que as opções de escolha, em cada grupo, são sempre binárias e o número de procedimentos necessários para identificar o aluno, que resolverá o problema na lousa, é quatro. Dessa forma, podemos dizer que nessa situação existem 16 possíveis alternativas de escolha de mensagem e que 4 *bits* de informação caracterizam tal situação, conforme ilustra o esquema abaixo:



Seja a fórmula para o número de bits num contexto probabilisticamente igual para os resultados:



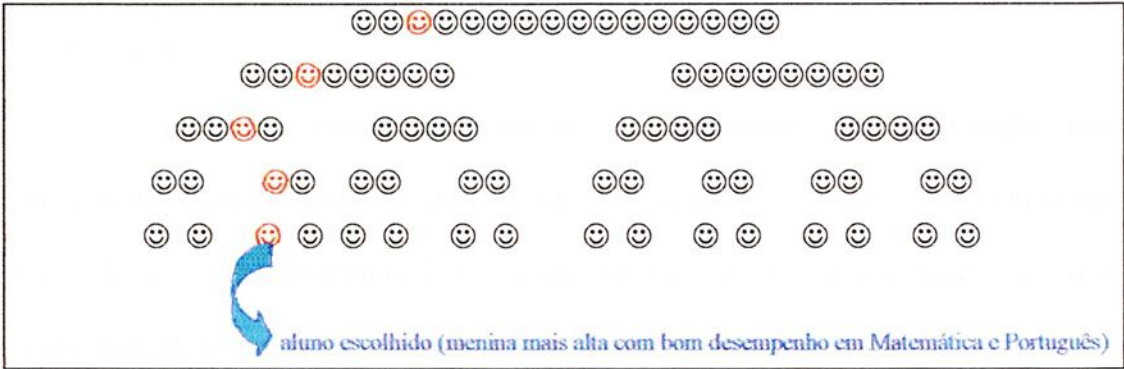
$$\text{Log}_2 M$$

Onde, M é o número de alternativas.

Aplicando ao exemplo temos:

$$\text{Log}_2 16 = 4$$

Esquemáticamente, podemos representar o exemplo acima do seguinte modo:



Um outro exemplo, construído por Edwards (1971), ilustra o jogo de Perguntas e Respostas entre duas pessoas. O jogo pode ocorrer da seguinte forma: a primeira pessoa tem por objetivo descobrir qual das oito primeiras letras do alfabeto (de A a H) o segundo jogador escolheu. O primeiro jogador, por exemplo, deve fazer perguntas do tipo "A letra está antes do C?" e as respostas do segundo jogador devem ser "SIM" ou "NÃO". Existem várias maneiras de descobrir a letra escondida. Uma delas consiste em dividir o conjunto de oito letras ao meio, resultando em dois grupos de quatro letras. A estratégia será perguntar em qual das duas metades a letra se

encontra e, a partir daí, subdividir os conjuntos até encontrar a solução. A figura 7 mostra como encontrar a solução do jogo, supondo que a letra escolhida seja o F:

Pergunta	Divisão	Resposta
1. Antes do E?	A B C D - E F G H	Não
2. Antes do G?	E F - G H	Sim
3. Antes do F?	E - F	Não

Figura 7 - Simulação do jogo Perguntas e Respostas (Edwards, 1971, p.46)

Na última pergunta, "Antes do F?", a resposta é "Não", pois a solução é o próprio F.

Como vimos anteriormente, a quantidade de informação está diretamente ligada ao tamanho do conjunto. Isto significa dizer que, no exemplo das letras de A a H, necessitamos de três perguntas para identificar a solução, ou seja, a quantidade de informação associada a essa situação é de três unidades.

O trabalho de Shannon & Weaver mostrou sua preocupação com a quantidade de incerteza presente num certo conjunto. A essa incerteza os dois autores chamaram de H. O valor de H, já definido anteriormente como

$$H = \text{Log}_2 M$$

mostra a grandeza de um conjunto versus a quantidade de incerteza associada a ele. Um conjunto com apenas uma unidade apresenta incerteza igual a zero, pois conhecemos todo o seu conteúdo. Por exemplo, Maria deseja saber quantas crianças de 3 a 5 anos e quantos jovens entre 18 e 25 anos estão presentes numa sala e, ao

adentrar a sala, Maria descobre que lá está apenas um jovem de 19 anos. Notamos que, dos exemplos apresentados acima, no primeiro, tínhamos um conjunto de 16 pessoas e fizemos 4 perguntas para reduzir a incerteza e encontrar a resposta. No segundo exemplo, tínhamos oito letras e fizemos três perguntas para encontrar o resultado.

Um paralelo entre a grandeza do conjunto e o número de perguntas a serem feitas para encontrar o valor de H é mostrado por Edwards (1971):

Grandeza do conjunto (n)	Número de Perguntas (H)
1	0
2	1
4	2
8	3
16	4

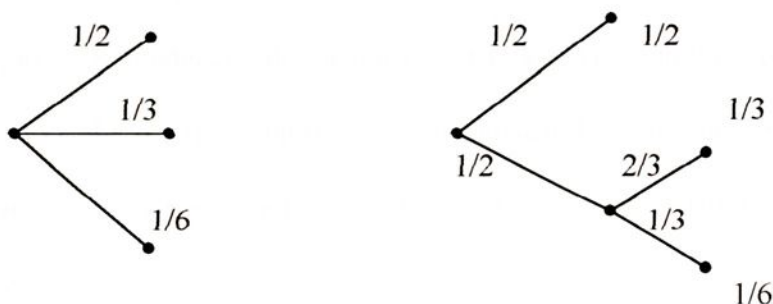
(Extraído de Edwards, 1971, p.49)

Em resumo, o valor da quantidade de informação está ligado a um valor surpresa, descoberto a partir de uma situação. A esse valor surpresa ou inesperado, Shannon & Weaver chamaram de entropia.

Shannon (1949, p.48-49) questiona se podemos medir, de algum modo, a quantidade de informação que é produzida numa certa situação, ou melhor, em qual proporção a informação é produzida. Inicialmente, podemos sugerir uma situação em que exista um número finito de eventos e que a probabilidade de

ocorrência deste é $p_1, p_2, p_3, \dots, p_n$. A escolha de um evento qualquer, como indicamos anteriormente, o autor chamou de H e atribuiu as probabilidades de escolha a essa variável, com as seguintes propriedades:

- H deve ser uma função contínua em p_i ;
- Se todo p_i é igual a $p_i = 1/n$, então H deve ser uma função monotônica crescente de n . Em outras palavras, numa situação de eventos igualmente prováveis, a incerteza, ou escolha, será maior na medida em que aumentar o número de eventos;
- O valor da incerteza H , presente numa situação, é igual à soma das incertezas dessa mesma situação. Por exemplo:



(Extraído de Shannon, 1949, p.49)

Existem situações nas quais nem todas as alternativas de escolha de informação são consideradas igualmente prováveis. Pode ocorrer que algumas alternativas tenham maior probabilidade de ocorrer do que outras. Assim, quanto maior o número de alternativas de escolha, maior a incerteza sobre uma determinada

situação e, conseqüentemente, maior a quantidade de informação. A fórmula para o cálculo da incerteza é definida por Shannon (1940, p.50) da seguinte maneira:

$$H = -k \sum_{i=1}^n p_i \text{Log } p_i$$

Onde: $\sum_{i=1}^n p_i$ é o somatório de todas as alternativas = 1

p_i é a probabilidade de ocorrência de uma dada alternativa.

Paralelamente a essa definição proposta por Shannon, Weaver em seu prefácio, recorre à termodinâmica para relacionar o conceito de entropia à informação. Entropia, na ciência natural, está associada ao grau de desordem numa situação e à tendência do sistema tornar-se mais randômico ou desorganizado. Weaver (1949, p.13) exemplifica: “Uma situação bem organizada não é caracterizada por um alto grau de randomicidade ou de escolha – isto é – a informação (ou entropia) é baixa”.

Em suma, quando, numa situação qualquer, aumentar a probabilidade de algumas alternativas, aumentará a entropia. Em outras palavras, quanto maior o número de mensagens a serem escolhidas, num certo contexto, maior será a quantidade de informação associada a essa situação.

Neste estudo, gostaríamos de ressaltar a importância do terceiro problema na comunicação apontado por Shannon & Weaver (1949), *O problema da eficácia*, relatado no início desta Seção, desde que os problemas do tipo 1 e 2 tenham

sido atendidos. O problema 3 focaliza a eficácia da informação sobre o comportamento do receptor, na medida em que este (o receptor) mantenha conexões regulares entre os padrões informacionais recebidos do emissor. Acreditamos que esse problema do tipo 3 possa indicar um bom caminho para traçarmos, na próxima Seção, um elo entre a Teoria da Informação e a Ciência Cognitiva.

SEÇÃO 3 - O elo entre a Teoria da Informação e a Ciência Cognitiva

Um dos pressupostos básicos da Ciência Cognitiva, como vimos no Capítulo 1, indica que os computadores são utilizados como ferramentas na tentativa de modelizar aspectos do comportamento inteligentes. Dessa suposição, deriva uma outra, segundo a qual, aponta Gonzales (1991, p.95), que nesta ciência "...os computadores podem ser usados como instrumentos para *testar* a precisão das hipóteses e teorias acerca do funcionamento da mente humana" (itálico nosso).

A partir desses pressupostos relativos à construção e teste de teorias/modelos explicativos da mente, que fundamentam a Ciência Cognitiva, devemos ressaltar que a Teoria Matemática da Comunicação, estudada na Seção anterior, mostra o aspecto quantitativo da informação num certo contexto. Tal aspecto possibilita uma modelagem matemática das características físicas para a seleção e transmissão de mensagens.

A Teoria da Informação indica, no nosso entender, três aspectos importantes para a proposta de modelagem ou simulação do comportamento

cognitivo: a) a definição da informação como processo dinâmico de seleção de sinais; b) a transmissão de informação; c) a alteração, mudança ou transformação promovida no comportamento do receptor.

Pereira Junior & Gonzales (1995, p.276) ressaltam que processos informacionais e aspectos cognitivos estão intimamente relacionados na transformação ou execução de comportamentos inteligentes. Esses autores argumentam que após a mensagem ter sido selecionada por um organismo ou sistema artificial (no caso das RNA), eles apresentarão uma estrutura interna organizada, com conexões neurais reforçadas, capazes de discriminar futuros padrões informacionais similares àquela estrutura. Para ilustrar a idéia acima, consideraremos a seguinte situação: imaginemos um sapo que possua em seu universo cognitivo uma relação de situações S em caso de fome. Onde S contém

S_1 = mosca parada	S_2 = mosca voando
S_3 = mosca andando	S_4 = mosca sem vida

Em qualquer uma das situações S , o sapo tentará saciar sua fome buscando padrões regulares do ambiente que condizem com seu sistema cognitivo.

É importante ressaltar, de um lado, que tal situação reforça a seguinte idéia: para sistemas cognitivos simples existe uma forte previsibilidade a respeito dos seus possíveis eventos, os quais a Teoria da Informação lida muito bem, como estudamos nesse Capítulo.



Por outro lado, sabemos que sistemas cognitivos complexos, como o do ser humano, não se comportam como máquinas finitas, excluindo a previsibilidade de todas as suas atividades cognitivas. O comportamento humano inteligente é dinâmico, complexo e envolve aspectos sociais/culturais, o que inviabiliza algumas das tentativas de explicação de suas nuances através da Teoria da Informação estudada aqui. Seria difícil explicar, através desta teoria, a qualidade da sensação do gosto do chocolate prazerosamente experimentado por uma criança. Contudo, existem situações que parecem ser perfeitamente explicáveis através da Teoria da Informação. No exemplo criado por Littlejohn (1982, p.154), na figura 8, a árvore genealógica é apresentada com o intuito de descobrir qual dessas pessoas cometeu um crime. Nos termos da Teoria da Informação, podemos perguntar: qual é a quantidade de informação associada a esta situação?

Durante as investigações, a polícia chegou a conclusão que:

- 1º O Pai, o Filho (B) e os Netos (D) e (E) estavam de férias do outro lado do mundo;
- 2º No momento do crime foi constatado que o Filho (C) e o Neto (F) brigavam em casa;
- 3º Há aproximadamente um ano faleceu o Neto (G).

Notamos que, nesse exemplo, existem oito alternativas possíveis para identificar o autor do crime (o Neto H). Logo, como vimos no Capítulo 3, a quantidade de informação associada à essa situação é de três unidades, considerando todas as alternativas igualmente prováveis.

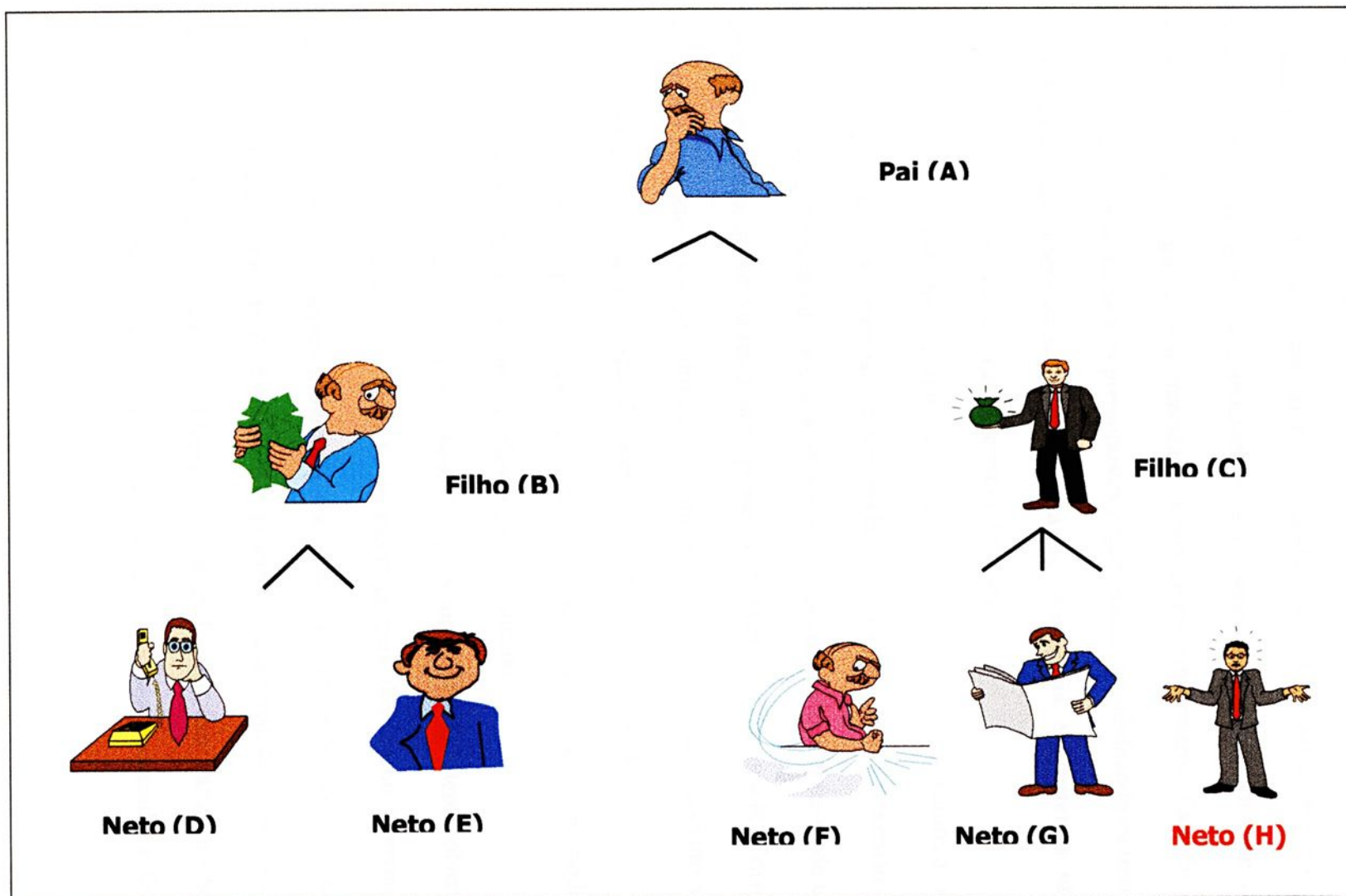


Figura 8 - Árvore genealógica de um crime, de Littlejohn (1982, p.154), ilustração nossa.

Exemplos como esse mostram a possibilidade de emprego da Teoria da Informação para auxiliar na elaboração de modelos computacionais de algumas atividades cognitivas. Nestes casos, conforme ressaltamos no Capítulo 2, apesar das diferenças estruturais entre os computadores e os seres humanos, o elo que os conecta, na Ciência Cognitiva, é a capacidade de processarem informação. Esse elo é muito importante, pois ampara os pressupostos cognitivistas, possibilitando os testes de verificação das hipóteses sobre a natureza da inteligência em sistemas não complexos que possuam um certo grau de previsibilidade.

No final do Capítulo 2, levantamos a questão da natureza da informação presentes nos sistemas inteligentes (humanos ou não). Sem a pretensão de propor uma nova definição do que seja informação, pois como vimos trata-se de um conceito complexo, queremos ressaltar as propostas de Hartley e Shannon & Weaver. No sentido proposto por esses últimos, o termo informação não está associado aos aspectos semânticos, mas à redução de incerteza presente numa situação.

Por exemplo, no caso do crime apresentado nesta Seção, um modelo inteligente poderia chegar à mesma conclusão que um humano.

Pesquisadores como Newell & Simon (1968) desenvolveram programas aplicando a idéia de redução de incerteza para demonstrar teoremas e jogos inteligentes. Um desses programas era o General Problem Solver - GPS (Programa Solucionador de Problemas), que possuía a característica de demonstrar teoremas, jogar xadrez, resolver problemas de criptografia, etc. Sua estrutura interna não era simples e pode ser resumida como segue: uma vez enunciada a forma



desejada para a resolução do problema, comparava o estágio atual do procedimento de solução com o objetivo final. Quanto menor a diferença desta comparação, mais o GPS estava próximo da solução.

Com o propósito, assumido no início dessa dissertação, de investigarmos aspectos do sistema cognitivo relacionados à organização informacional da memória, destacaremos no próximo Capítulo as principais características das redes neurais artificiais. Tais características nos levarão ao estudo da aprendizagem nessas redes e em especial na Rede Neural da Harmonia proposta por Smolensky (1986).

O ponto importante a ser discutido no próximo Capítulo diz respeito à verificação da possibilidade de modelos artificiais explicarem mecanismos de *aprendizagem*, ainda que elementares.



CAPÍTULO 4

REDES NEURAIIS ARTIFICIAIS

INTRODUÇÃO

Os estudos sobre as Redes Neurais Artificiais-RNA experimentaram um avanço até o início da década de 60, quando, então, encontraram dificuldades teóricas e foram ofuscados pela computação digital, que fez crescer a Inteligência Artificial.

No início da década de 80, cientistas, como McClelland & Rumelhart (1988), Kohonen (1987), entre outros, trouxeram à tona a discussão sobre as Redes Neurais Artificiais, com importantes publicações para a Ciência Cognitiva.

Neste Capítulo, na Seção 1, investigamos as principais características das Redes Neurais e como elas aprendem. Para isso, na Seção 2, estudamos os mecanismos de aprendizagem e treinamento nessas redes. Finalmente, na Seção 3, focalizamos uma teoria das representações mentais, proposta por Smolensky, chamada de Teoria da Harmonia (Harmony Theory), que apresenta um elemento de ligação entre a Teoria da Informação (estudada no Capítulo 3) e as RNA.



SEÇÃO 1 - Principais características das Redes Neurais Artificiais

Como mencionamos anteriormente, o recente interesse pelo estudo das RNA se deve ao trabalho de alguns pesquisadores e também a fatores, como: tecnologia disponível para fabricação de hardware especializado, computação paralela e um grande progresso teórico na área de computabilidade.

Antes de discutirmos as modernas estruturas das RNA, é necessário fazer referência àqueles que, na década de 40, como vimos no Capítulo 1, estiveram profundamente envolvidos em pesquisas que originaram o Conexionismo. McCulloch & Pitts (1943) identificaram o neurônio biológico a uma formalização matemática, que mais tarde ficou conhecida como *Rede Neural Artificial*. A inspiração no neurônio biológico indicava a tentativa de simulação de alguns aspectos humanos inteligentes.

O sistema nervoso é constituído por uma rede de neurônios, onde cada um possui um grande número de vias de entrada denominadas *dendritos*, o corpo celular denominado *soma* e um *axônio*, que é a via de saída. As sinapses, que são as junções dos neurônios, estão entre o axônio de um e o soma do outro. A figura 9 ilustra a composição esquemática de um neurônio biológico.



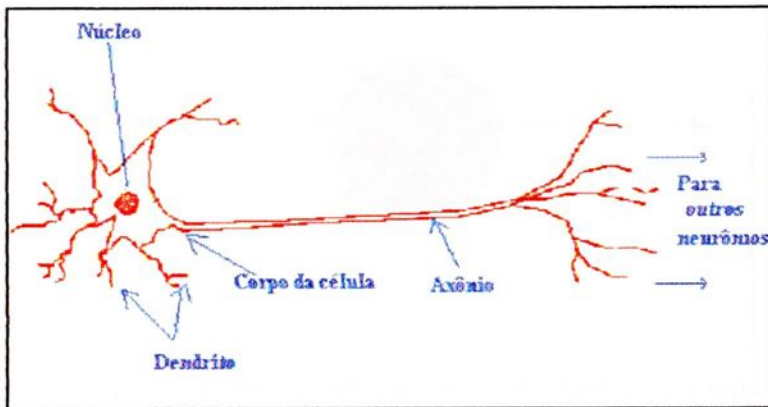


Figura 9 - Neurônio Biológico (Cheng & Titterington, 1994, p.3)

Os neurônios interagem entre si a partir do instante em que a excitação neural ultrapassa um certo limiar. De acordo com McCulloch & Pitts (1943), essas sinapses são chamadas de excitatórias, mas o inverso também pode ocorrer. As sinapses inibitórias, como são chamadas, podem ser de dois tipos:

- a) Quando um neurônio, ou grupo de neurônios, deixam de ser excitados por outros neurônios;
- b) Quando um neurônio tem seu limiar aumentado.

Simbolicamente, o neurônio artificial proposto McCulloch & Pitts pode ser representado de acordo com a figura 10:

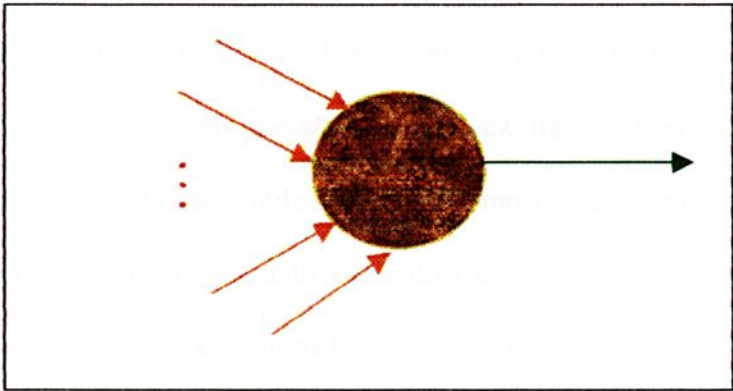


Figura 10 - Neurônio Artificial com n entradas e uma saída.

Uma RNA é um sistema de processamento paralelo de informação que, segundo parte dos seus defensores, como, por exemplo, McClelland & Rumelhart (1986), não é algorítmico. Como vimos no Capítulo 1, essas redes não são construídas como uma seqüência de regras, claramente definidas, as quais sempre determinam um resultado. Ao contrário, elas formam um conjunto físico de *neurônios* (que denominaremos nódulos ou unidade neurônio-símile) conectados, análogos aos neurônios naturais.

Um paralelo entre máquinas digitais e máquinas conexionistas está ilustrado na tabela abaixo:

Máquinas de Turing	Máquinas Conexionistas
Executam programas	Aprendem
Executam operações lógicas	Executam operações de transformação e comparação entre padrões informacionais
Dependem do modelo ou do programador	Descobrem as relações ou regras dos dados e exemplos
Verificam uma hipótese por vez	Verificam várias hipóteses em paralelo

Para Caudill & Butler (1992), um nódulo é um dispositivo que recebe estímulos (de entrada), do meio ambiente ou de outros nódulos, e traduz esses estímulos de entrada em saídas que são, por sua vez, transmitidas ao longo da RNA para outras unidades neurônio-símile. A figura 11 mostra as conexões físicas de uma rede neural artificial típica. Segundo esses dois autores, o ponto onde as entradas encontram os nódulos é denominado *sinapse* (termo inspirado nas conexões sinápticas entre os neurônios naturais). É importante ressaltar que uma RNA pode estar estruturada com um grande número de unidades neurais (nódulo), por exemplo: uma camada de entrada; uma de saída; e pelo menos uma (ou nenhuma) camada intermediária (ou escondida).

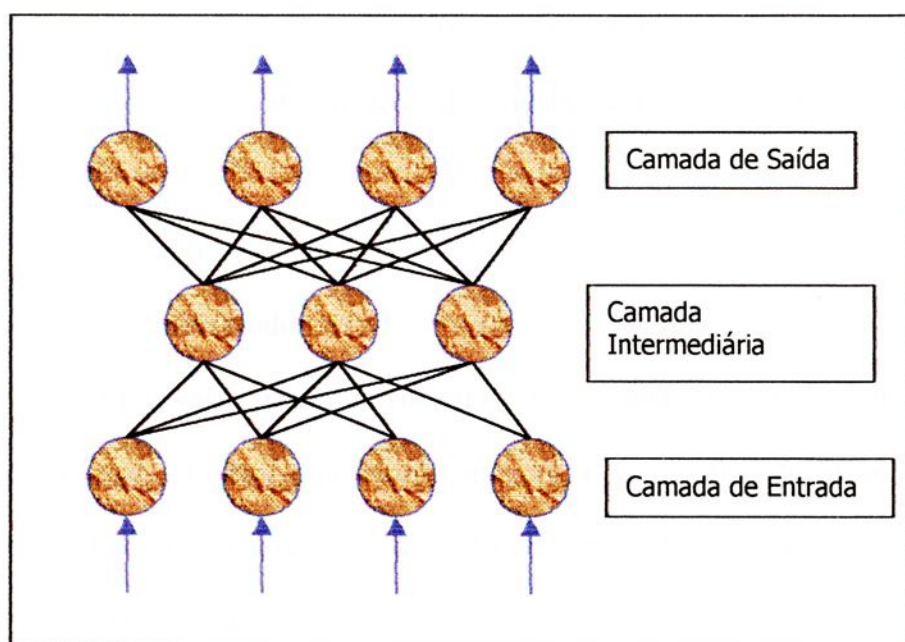


Figura 11 - Típica Rede Neural Artificial com 3 camadas

De acordo com Caudill & Butler (1992, p.5-6), o processo de organização de uma rede neural artificial obedece a três etapas:

- 1) Computar as entradas líquidas recebidas de outros nódulos, da seguinte forma:

$$I_i = \sum_{j=1}^n w_{ij} \cdot x_j$$

Onde: I_i é a entrada líquida recebida do nódulo i do total de n nódulos da rede

x_j é a intensidade da entrada a partir do nódulo j

w_{ij} é o peso da conexão a partir do nódulo j até o nódulo i

- 2) Converter a entrada líquida para um nível de ativação do nódulo (equivalente ao nível de excitação do neurônio biológico), da seguinte forma:

$$f(I) = 1 / (1 + e^{-I})$$

- 3) Converter o nível de ativação do nódulo para um sinal de saída, da seguinte forma:

$$y_i = \begin{cases} f(I), & \text{se } f(I) > T \\ 0, & \text{de outro modo} \end{cases}$$

Onde T é um valor *threshold* (limiar) e y_i é o sinal de saída denominado *função transferência*. O nódulo produzirá uma saída se o valor de ativação exceder o *threshold* dado (excitação neural); caso contrário, ele não produzirá saída (o valor computado é chamado de inibitório).

Em síntese, um valor de saída de uma unidade neurônio-símile pode ser obtido da seguinte maneira:

- Somando a entrada líquida total;
- Atualizando sua ativação;
- Relacionando o resultado obtido da função de ativação com o valor de excitação (limiar) da unidade.

Na figura 12, mostramos o esquema de um nódulo recebendo entradas e realizando a função de transferência:

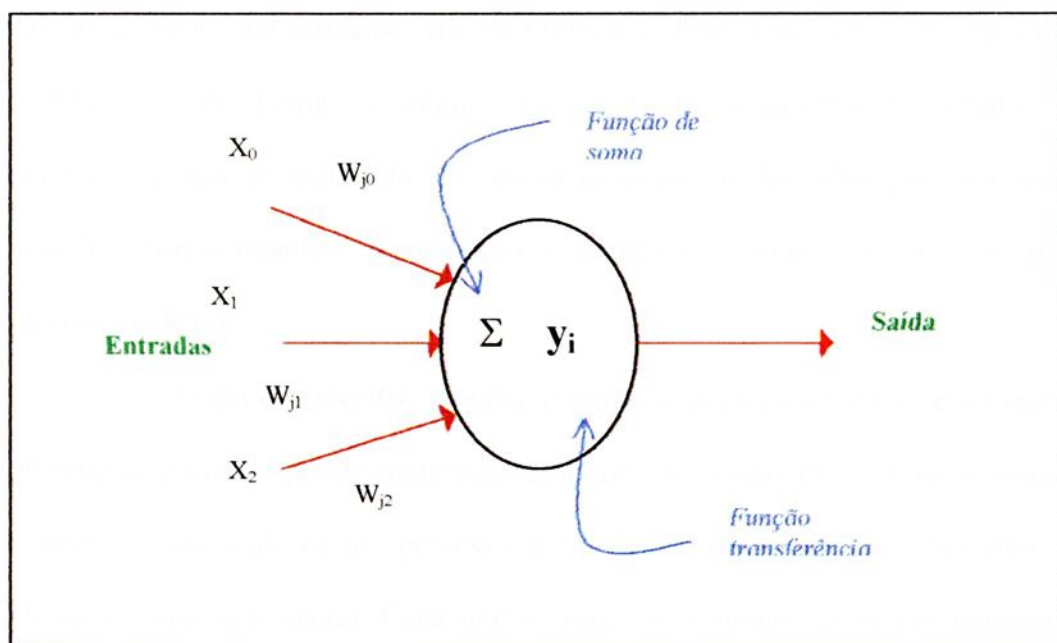


Figura 12-Um nódulo computando entradas e a função transferência

Conforme veremos a seguir, segundo os conexionistas, as RNA podem constituir um bom modelo explicativo de alguns aspectos do comportamento humano inteligente. Elas formam uma rede interconectada de unidades neurônio-símile que, ao se relacionarem, permitem que a rede aprenda. Esse processo associativo entre as

unidades faz emergir o comportamento inteligente da rede neural. Na próxima Seção, estudaremos o que significa e como é aprender para uma RNA.

SEÇÃO 2 - Tipos de Aprendizagem nas Redes Neurais Artificiais

Na Seção 1 deste Capítulo, expusemos um modelo de rede neural artificial proposto, inicialmente, por McCulloch & Pitts. Essa rede, diferentemente das Máquinas de Turing, constitui um sistema de processamento paralelo de informação e não se utiliza de algoritmos previamente definidos para apresentar resultados. Nosso objetivo, nesta Seção, é estudar a concepção de *aprendizagem* simulada nas RNA.

O ato de aprender, segundo o senso-comum, está ligado à experiência, à observação e à aquisição de conhecimento através do estudo. Para o Conexionismo, aprender é o resultado de um processo de ajuste dos pesos das conexões entre os núdulos de uma rede neural. Cada núdulo está num contínuo processo associativo com outros núdulos, inibindo ou excitando uns aos outros, de forma a representar um certo padrão. Gonzales (1991, p.99) ressalta:

O estado informacional de cada unidade é descrito por uma função de output, que soma todas as cargas das ativações das unidades intermediárias e de entrada do sistema, as quais são amplificadas ou inibidas em função dos pesos das conexões que existem entre essas unidades.



O processo de treinamento de uma rede neural foi formalizado, pela primeira vez, por Donald O. Hebb, em 1949. Para ele, o processo de aprendizagem ocorre a partir de eventos que se repetem, deixando marcas ou traços na memória. Por isso, tendem a fortalecer as conexões sinápticas.

Uma variação da Lei de Hebb é proposta por Rumelhart & McClelland (1986, p.36): "Quando a unidade A e a unidade B forem simultaneamente excitadas, a força da conexão entre elas deverá ser incrementada". O interessante nos modelos mais sofisticados de RNA é a capacidade de aprenderem sem a presença de uma regra geral que determine seu comportamento. Ainda segundo Rumelhart & McClelland (1986, p.39), os modelos neuronais, de uma maneira geral, tendem a produzir comportamentos similares aos daqueles apresentados a eles. Essa propriedade é caracterizada como *generalização espontânea*. Isso é garantido por mecanismos de aprendizado local, que podem possibilitar conexões fortes, entre os nódulos, *aprendendo* através da experiência. A experiência é, aqui, entendida como a interação entre a rede e as regularidades existentes no meio ambiente. Uma RNA possui a propriedade de detectar essas regularidades, reforçar as conexões entre os nódulos mais excitados, que representam tais regularidades e conduzir a um comportamento que evidencie aquele processo de reconhecimento de padrões.

Alguns autores, entre eles Caudill & Butler (1992, p.61), fizeram a adaptação da Lei de Hebb para uma linguagem computacional. A seguir, ilustramos essa adaptação de conexão e ativação entre dois nódulos:



Para cada tempo de t

 Compute a atividade de A

 Compute a atividade de A recebida por B

 Compute a atividade de B

 Se atividade de B > 0 e a atividade de A recebida por B > 0 então

 Aumente a força da conexão entre A e B

É importante ressaltar que essa lei de conexão e ativação é aplicada localmente, ou seja, apenas entre os dois nódulos simultaneamente excitados num certo instante, deixando um *traço ou marca* de aprendizado entre eles.

A lei de aprendizagem estabelecida por Hebb (1949) indica que os estímulos que ocorrem com maior frequência fortalecem as conexões neuronais. Em consequência disso, uma rede neural pode ser treinada e aprender a reconhecer um padrão visual, por exemplo.

No início desta Seção, dissemos que nosso objetivo era o de estudar o processo de aprendizagem nas RNA. Discutiremos os dois principais processos de treinamento de redes: o treinamento supervisionado e o não-supervisionado.

Uma das atividades mais desenvolvidas na pesquisa das RNA é o *reconhecimento de padrões*. Assim, quando apresentamos um padrão qualquer à rede, esperamos que ela faça o reconhecimento do padrão apresentado-o em seu *output* com um mínimo de fidelidade. Para uma atividade qualquer proposta à rede, esta deverá ser treinada para determinar seu objetivo (que é o aprendizado). Até este

momento, utilizamos os termos treinamento e aprendizado como sinônimos. É importante diferenciá-los. Caudill & Butler (1992) ressaltam que *treinamento* é o procedimento pelo qual a rede neural aprende (é um procedimento externo à rede); e aprendizado (processo ou atividade interna à rede) é o resultado final desse procedimento empregado.

No treinamento supervisionado, segundo Caudill & Butler (1992), é apresentado um padrão, como entrada, à rede e espera-se que ela mostre como saída o padrão da entrada. Um supervisor verifica o quanto o padrão de saída atual difere do padrão desejado. Seu objetivo é modificar os pesos (utilizando o erro encontrado) das entradas e verificar se a rede irá convergir, reconhecendo o padrão inicialmente apresentado.

Quanto ao treinamento não-supervisionado ou auto-organizado, os autores escrevem que à rede é direcionado apenas um conjunto de padrões de entrada e nenhuma informação (de *feedback*) sobre seu desempenho no reconhecimento. A rede faz seu próprio ajuste de pesos. Rumelhart & McClelland (1986, p.55) enfatizam que, nesse tipo de treinamento, a rede recebe (como entrada) uma parte do padrão e a parte ausente ela reconhece por associação com o padrão a ela apresentado.

Um estudo sobre RNA que enfatiza as atividades de treinamento não-supervisionado foi desenvolvido por Kohonen, na década de 80, situando suas pesquisas numa categoria denominada de redes que se auto-organizam. Kohonen (1987) destaca que seu propósito é analisar como estruturas neurais hipotéticas

podem ser afetadas por sinais sensoriais e que tipos de estruturas e propriedades funcionais são suficientes para a implementação da *auto-organização*.

Uma descrição da rede neural artificial de Kohonen é apresentada por Caudill & Butler (1992, p.83). Esta possui duas camadas de nódulos interconectados: a camada de entrada e a camada de Kohonen (que é, também, a de saída). A figura 13 ilustra a rede neural artificial de Kohonen.

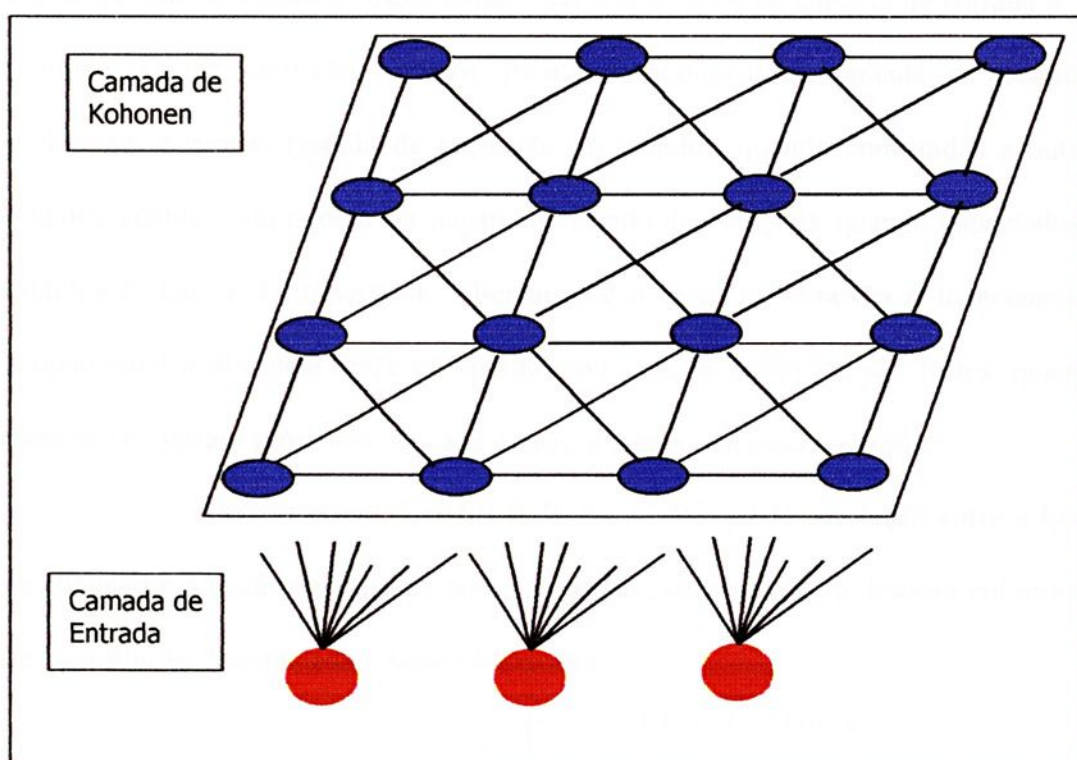


Figura 13 - Rede neural artificial de Kohonen

Nessa rede, a camada de entrada distribui o padrão apresentado para cada um dos nódulos da camada de Kohonen. Essa camada, que possui um grande

número de conexões entre os seus nódulos, estrutura-se como uma camada de saída, transmitindo o resultado de sua execução.

Amparado pela literatura sobre um mecanismo cerebral denominado de *inibição lateral*, Kohonen (1987) construiu seu modelo na tentativa de entender algumas características da formação do aprendizado humano.

De acordo com Caudill & Butler (1992, p.84), as unidades neurônio-símile da rede de Kohonen estão conectadas aos nódulos da camada de entrada e, na segunda camada, com eles mesmos. Para cada nódulo dessa camada, as conexões podem ser positivas (estado de excitação do nódulo), quando conectados a outros nódulos vizinhos, ou podem ser negativas (estado de inibição), quando conectados a nódulos distantes. É importante salientar que a força da conexão é inversamente proporcional à distância entre os nódulos, ou seja, as conexões são fortes, quando estes se encontram próximos uns aos outros, e fracas, em caso contrário.

Ainda segundo Caudill & Butler (1992, p.85), a relação entre a força da conexão e a distância entre os nódulos, na segunda camada, é descrita em termos de uma função denominada *Chapéu Mexicano*:



Para os autores, a interconexão dos nódulos estimula a competição entre eles. Quando um padrão é apresentado à rede, os nódulos da camada de entrada distribuem esse padrão (fenômeno denominado *fan-out*) para cada um dos nódulos da camada de Kohonen. Os pesos das conexões são modificados pela passagem do padrão de entrada para os nódulos da camada superior. O método da inibição lateral determina qual é o nódulo vencedor, ou seja, aquele que possui o padrão mais próximo daquele apresentado inicialmente à rede.

Em suma, durante o processo de treinamento da rede, o objetivo é encontrar a menor distância entre o vetor peso de um nódulo da camada de Kohonen e o padrão dado como entrada (vetor de entrada). Nesse processo, a vizinhança de cada nódulo vencedor tende a diminuir, a cada instante de tempo, até o ponto em que o número de iterações determinar, na melhor das hipóteses, o nódulo vencedor (aquele que contém o padrão mais próximo do vetor de entrada). A esse processo de ajuste (e correção) de pesos, que a rede faz sozinha, denomina-se *auto-organização*.

Este princípio de auto-organização, norteador das RNA, parece contribuir para a pesquisa da Ciência Cognitiva, no que diz respeito à simulação de aspectos humanos inteligentes, através do estudo da noção de treinamento e aprendizagem não dirigido. Na próxima Seção, estudaremos o modelo neural da harmonia proposto por Smolensky e a tentativa de medir o grau de ajuste entre o que um sistema possui de informação, em relação à informação disponível em seu mundo.

SEÇÃO 3 - Smolensky e o modelo da Harmonia no Sistema Cognitivo

Smolensky (1986) desenvolveu um trabalho denominado *Processamento de Informação em Sistemas Dinâmicos: Fundamentos da Teoria Harmônica*, em que propõe estudar classes de sistemas dinâmicos que possam servir como modelos da cognição humana e que permitam projetar sistemas cognitivos artificiais.

A idéia foi a de desenvolver uma rede neural considerando, na saída, um conjunto de inferências que forneça o melhor ajuste das entradas. Isso pode ser expresso através de uma função proposta por Smolensky (1988, p.89), chamada *função harmonia*, que, segundo ele: "...mede, para todo conjunto de inferências possíveis, o grau de ajuste na entrada relativamente às limitações leves estocadas nos pesos das conexões entre as unidades". Essa função, bem como o detalhamento da rede, veremos mais adiante. O objetivo da Teoria da Harmonia, segundo o autor, é criar uma linguagem para expressar as teorias cognitivas no paradigma conexionista (subsimbólico). Assim, ele fornecerá uma caracterização matemática precisa do padrão a ser reconhecido, sugerindo como resolvê-lo, utilizando uma rede neural artificial e, finalmente, mostrará um procedimento que permita à rede aprender com a experiência.

No modelo neural da Teoria da Harmonia, duas camadas distintas são representadas: a camada inferior é constituída por um conjunto de unidades de *aspectos representacionais* e a camada superior por *átomos de conhecimento*.

A figura 14 ilustra uma representação de um modelo particular harmônico:

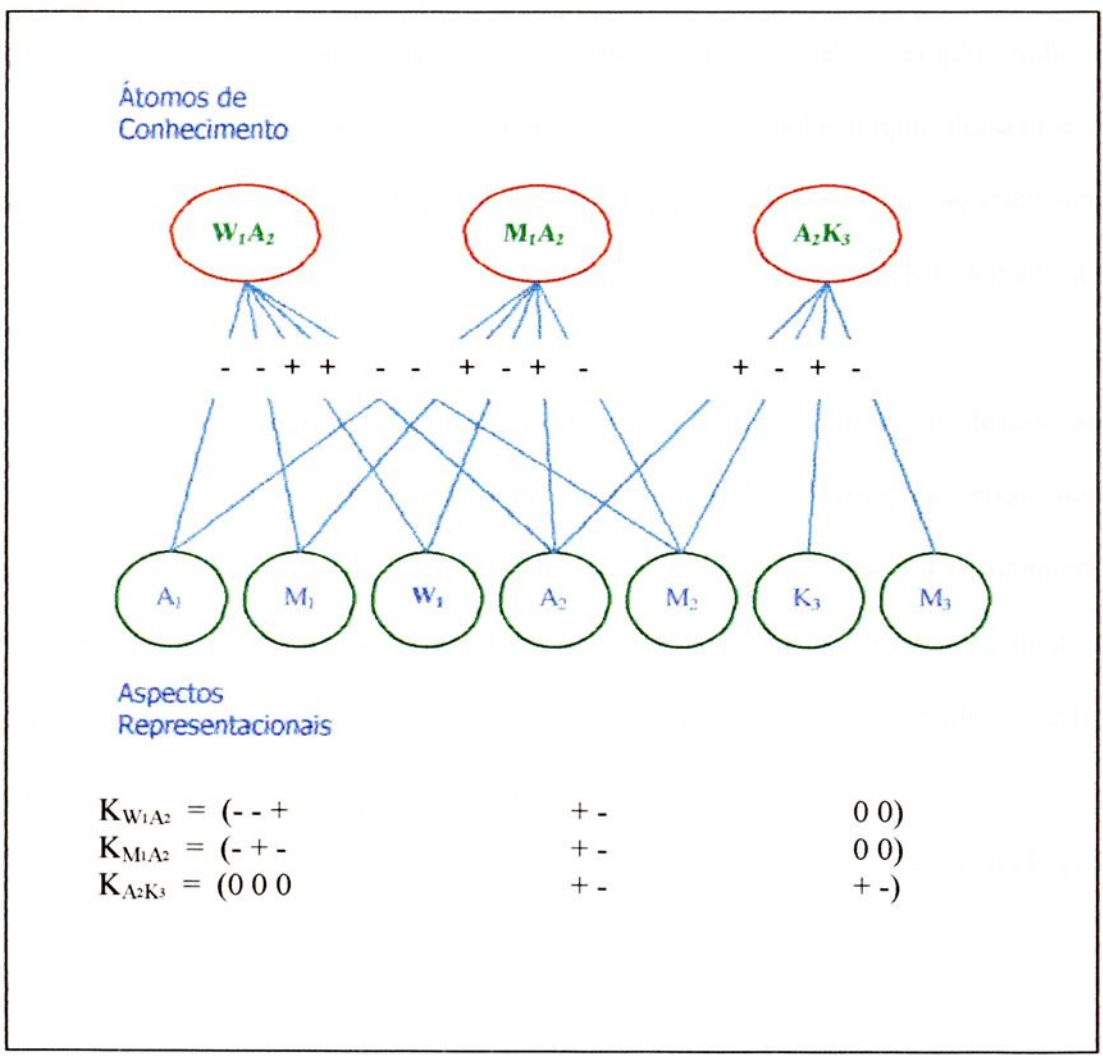


Figura 14 - Um modelo harmônico (Smolensky, 1986, p.205)

Nesse modelo, os *aspectos representacionais* indicam as letras em determinadas posições e os *átomos do conhecimento* representam combinações

comuns de letras. As linhas que ligam os aspectos representacionais aos átomos de conhecimento possuem um sinal $+$ ou $-$, que indica a força entre as unidades. Como veremos adiante, essa relação determina a formação do conhecimento ou da informação disponível à rede. Segundo Smolensky (1988, p.206), os átomos de conhecimento são, no caso humano, fragmentos de representações que se acumulam em nós com a experiência. No nosso sistema cognitivo, por exemplo, todo o conhecimento sobre aves está representado em átomos de conhecimento dessas aves. Na medida em que as representações (das aves) que temos do mundo não condizem com a realidade, ambos os sistemas (humano cognitivo e mundo) deixam de harmonizar-se.

Para o autor, a justificativa do nome *Teoria da Harmonia* deve-se ao fato de que, para toda configuração de entrada, o sistema busca encontrar uma configuração dos átomos de conhecimento que sejam consistentes, ou harmônicos, com os aspectos representacionais (apresentados como entrada). A situação final, e esperada, dos átomos do conhecimento nos dá indícios de que a rede neural da Teoria da Harmonia interpreta os aspectos informacionais presentes no mundo.

A função harmonia de Smolensky é expressa pela equação citada por ele (in: McClelland & Rumelhart, 1988, p.77) como :

$$\sum_i \sigma_i a_i h_i$$

Onde: $i \rightarrow$ é a quantidade de átomos do conhecimento

$\sigma_i \rightarrow$ é um valor que exprime a importância da associação entre o átomo de conhecimento e a unidade de entrada

$a_i \rightarrow$ é o valor de ativação do átomo de conhecimento, que pode ser:

$$a_i = \begin{cases} 0, \text{ átomo do conhecimento não está ativado} \\ 1, \text{ átomo do conhecimento está ativado} \end{cases}$$

$h_i \rightarrow$ é um valor que mede o grau de consistência entre o aspecto representacional e o átomo de conhecimento i . A variável h_i é determinada por:

$$\frac{\sum_j r_j k_{ij}}{n_i} - k$$

Sendo que j é o número de aspectos representacionais, r_j é a ativação da unidade representacional j , e n_i é o número de conexões, diferentes de zero, com o átomo i . À variável k_{ij} , que é o peso da conexão entre o aspecto representacional j e o átomo i , podem ser atribuídos os seguintes valores:

$$k_{ij} = \begin{cases} 1 \text{ se a conexão} > 0 \\ -1 \text{ se a conexão} < 0 \\ 0 \text{ se não existir conexão} \end{cases}$$

Se o átomo de conhecimento está ativado ($a_i = 1$), então ocorrerá o produto de σ_i (que é a importância da associação entre o átomo de conhecimento e a unidade de entrada)

e h_i , que representa o termo de consistência do átomo atual com as características representacionais.

O termo h_i indica a proporção de relevância dos aspectos representacionais que são consistentes, subtraídos daqueles que são inconsistentes (na fórmula, denominado k).

Quando não existir inconsistência ($k=0$), o átomo i contribuirá com uma quantia positiva para a harmonia de todo o sistema. Valores de k próximos a 1 influenciarão o sistema a não se harmonizar completamente.

A representação do conhecimento na rede neural da harmonia, proposta por Smolensky (1986, p.213), tem como ponto de partida os aspectos representacionais r . Esses aspectos constituem a representação de sistemas cognitivos de possíveis estados do meio com o qual se relacionam. Um estado representacional de um sistema cognitivo é determinado por um conjunto de valores r . Tal conjunto pode ser um vetor composto por + (característica presente) ou por - (característica ausente).

Os aspectos representacionais, segundo Smolensky (1988, p.215), funcionam como sustentação para que o sistema cognitivo realize sua tarefa de computação. O *conhecimento/informação* que direciona essa computação está associado à segunda camada da rede da harmonia: os átomos do conhecimento. Cada átomo i é caracterizado por um vetor a_i , que traduz um *pedaço* de conhecimento em r_j (+1, -1 ou 0). Os átomos de conhecimento armazenam padrões dos aspectos representacionais que representam o meio ambiente. Para o autor, a presença de

padrões informacionais, no mundo, é detectada pela rede da harmonia com o auxílio da medida da informação proposta por Shannon (como já estudamos no Capítulo 3). A relevância do conhecimento a ser armazenado, num certo átomo do conhecimento, é dada pela força da conexão, σ_i , desse átomo com o aspecto representacional.

A figura 15 ilustra uma rede harmonia genérica:

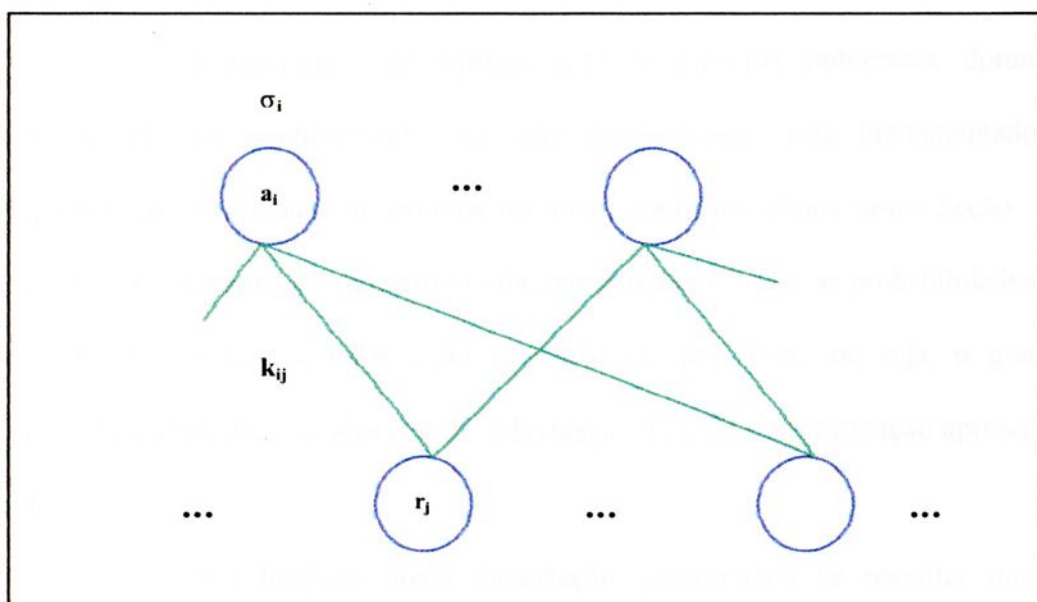


Figura 15 - Esquema da rede neural harmonia. Adaptado de Smolensky (1986, p.231)

Aos valores iniciais de r_j são atribuídos **+1** ou **-1** e os valores dos átomos de conhecimento são inicializados com **0**. De acordo com Smolensky (1986, p.231), o valor de cada nódulo é atualizado segundo:

$$\text{Prob (valor=1)} = \frac{1}{1 + e^{-I/T}}$$

Onde T é um parâmetro global e I é a entrada de um nódulo vindo de outros nódulos. Durante o processo de atualização dos nódulos, T inicia com algum valor positivo e

irá gradualmente diminuindo. Quando atingir 0, o sistema terá completado sua máxima harmonia.

Em resumo, na Teoria da Harmonia, o conhecimento está armazenado nos átomos que compõem a segunda camada da rede, representando padrões informacionais descritos pelas unidades representacionais. Essas unidades são compostas por frequências de regularidades presentes no meio ambiente.

O mecanismo de representação de aspectos ambientais, durante a formação de um conhecimento, na rede da harmonia, está fundamentado na frequência de ocorrências de eventos no meio, conforme vimos nessa Seção. Essa frequência é *captada* pela rede através da distribuição de todas as probabilidades que são consistentes com a informação já *conhecida* pela rede, ou seja, o grau de homogeneidade entre o conhecimento/informação da rede e a informação apresentada a ela.

Para finalizar nossa dissertação, gostaríamos de ressaltar que um aspecto interessante no trabalho de Smolensky é que o conhecimento é adquirido na medida em que o sistema busca relações informacionais dentro de um contexto. No caso do nosso sistema cognitivo, parece-nos que buscamos, na natureza, constantemente padrões regulares que concordam ou estão em harmonia com a informação que dispomos sobre eventos, situações, etc. Tal correlação reflete a contribuição da Teoria da Harmonia para o entendimento de aspectos inteligentes do comportamento humano.



CONCLUSÃO

O ponto inicial desta dissertação foi a investigação de aspectos gerais do comportamento inteligente, relacionados à organização informacional da memória cognitiva.

Iniciamos nosso trabalho revisando as raízes e os fundamentos da Ciência Cognitiva amparados principalmente pelos três primeiros pressupostos dessa ciência: a) o estudo das atividades cognitivas deve levar em consideração as representações mentais; b) os computadores constituem uma ferramenta apropriada para modelar a mente humana; c) os fatores emocionais não serão considerados, neste momento, nos estudos cognitivos. Neste contexto, ressaltamos a importância da utilização de modelos para a tentativa de entendimento sobre a natureza do conhecimento humano.

Seguindo a linha de modelização de aspectos do comportamento inteligente, discutimos a abordagem mecanicista da mente, proposta pela Ciência Cognitiva, através de regras abstratas e algoritmos para a resolução de problemas. Apresentamos a Máquina de Turing como um instrumento para elaboração de modelos explicativos da atividade cognitiva. Indicamos também que as máquinas apresentadas podem realizar, em princípio, os mais variados tipos de cálculo, desde que preparadas para isso. Conforme ressaltamos, insistentemente, a elaboração de técnicas de modelagem, na Ciência Cognitiva, através da Máquina de Turing, só foi



possível devido à hipótese de que há um elemento comum, fundamental entre máquinas e seres humanos, qual seja, a habilidade de processar informação.

Com o objetivo de aprofundar o nosso entendimento da hipótese que os computadores podem contribuir na explicação de alguns aspectos do comportamento inteligente, analisamos a Teoria Matemática da Comunicação, proposta por Shannon & Weaver.

A novidade da presente dissertação reside no esclarecimento do elo entre a Teoria da Informação e a Ciência Cognitiva. Nesse sentido, indicamos três características básicas da Teoria da Informação: a definição da informação, sua transmissão e, finalmente, a transformação do comportamento do receptor de uma mensagem em função da mensagem recebida. Assim, conforme procuramos enfatizar, nos casos discutidos no Capítulo 3, um receptor (seja ele uma RNA ou um cérebro humano) direcionará o seu comportamento no sentido de minimizar a incerteza e conseqüentemente maximizar a sua eficiência. É justamente na habilidade de maximizar a eficiência do seu comportamento que, julgamos, reside a inteligência de um sistema inteligente.

Uma vez de posse da informação, o comportamento do indivíduo é alterado, possibilitando tomadas de decisões mais eficientes. No decorrer da história da humanidade, a informação sempre desempenhou um papel decisivo no plano do indivíduo ou da coletividade. No plano individual, por exemplo, se Napoleão tivesse obtido, antecipadamente, informações sobre o clima na Rússia, certamente não teria autorizado a sua invasão, poupando milhares de vidas e, quem sabe, mudando o rumo

da história. De outro modo, no plano coletivo, com a quebra da cifra secreta dos alemães na 2ª Guerra, por Turing, os ingleses conseguiram antecipar-se aos eventos de ataque que contribuíram para a vitória dos aliados.

Julgamos assim que a capacidade de processar informação contribui para que seres humanos e máquinas sejam concebidos, apesar das diferenças estruturais, como compartilhando de um mesmo elemento decisivo na estruturação do seu processo de aquisição e desenvolvimento de habilidades inteligentes.

Nos casos dos sistemas que processam informação na Ciência Cognitiva, a informação é o elemento que aproxima, e ao mesmo tempo diferencia, a Máquina de Turing da Máquina Conexionista. Enquanto a Máquina de Turing processa informação destituída de qualquer significado, seguindo regras pré-estabelecidas, seqüencialmente, a Máquina Conexionista processa informação através de diversos núdulos em paralelo, não possuindo uma regra central exclusiva que determine, de modo absoluto, um resultado. Além disso, a RNA traz em si, segundo seus defensores, entre eles Rumelhart & McClelland (1986), um conteúdo informacional significativo. Os padrões de conectividade que representam, por exemplo uma laranja, contêm em si uma informação específica sobre laranjas, uma vez que eles são gerados causalmente pelos dados informacionais que caracterizam laranja no meio ambiente.

Finalmente, apresentamos o modelo da harmonia, proposto por Smolensky, como uma contribuição para a explicação da habilidade inteligente de



processar informação. O trabalho de Smolensky é uma caracterização matemática que tenta descrever e analisar o comportamento de um sistema inteligente.

Constatamos que o modelo neural da harmonia *busca* padrões de regularidades no meio ambiente, procurando ajustar o padrão informacional, criado na rede, com a informação disponível no mundo. Tal ajuste reflete a tentativa, de Smolensky, de formalizar a noção de esquemas ou idéias que emergem no processo de interação da rede com o meio ambiente, durante a aprendizagem.

Para a detecção dos padrões informacionais no ambiente, Smolensky enfatizou a busca, pela rede, de regularidades no mundo. Conforme procuramos mostrar, Smolensky tem como hipótese central de sua teoria a idéia de que sistemas cognitivos inteligentes buscam diminuir a incerteza a respeito dos padrões apresentados no meio. No caso particular de um tipo de informação já retida na memória, por um sistema cognitivo, as chances de identificação de uma informação, no meio ambiente, evidenciam a semelhança da rede com nosso sistema cognitivo. Em outras palavras, a hipótese que apresentamos aqui, em concordância com Smolensky, é de que sistemas cognitivos inteligentes buscam constantemente descobrir padrões informacionais que fortaleçam ou complementem o conhecimento já adquirido. É justamente isto que faz a rede elaborada por Smolensky: ela procura distinguir os padrões informacionais do ambiente que sejam *consistentes* com os seus padrões ou esquemas já adquiridos. Ao reforçar esse procedimento, tais esquemas são ressaltados.



O processo de harmonização entre os aspectos informacionais e os átomos de conhecimento nos forneceria uma medida do ajuste do sistema com o mundo, que, de acordo com nossa visão, possibilitaria medir o grau de inteligência desse sistema.

Nessa mesma linha, Ashby (1962) enfatiza que sistemas inteligentes (organismos ou máquinas) são aqueles que possuem a capacidade de adaptação ao meio no qual estão inseridos, de maneira a maximizar sua existência ou a permanência de seus parâmetros fundamentais.

Julgamos que a Teoria da Harmonia nos fornece bons elementos para o desenvolvimento de modelos explicativos do funcionamento dos sistemas cognitivos inteligentes. Pois, se admitirmos que os sistemas cognitivos são estruturados através de átomos de conhecimento e que estes se desenvolvem através das relações informacionais estabelecidas entre o meio ambiente e as redes neurais, então o instrumental matemático fornecido por Smolensky nos indica um caminho fértil para o estudo de tais sistemas.

Até aqui indicamos duas vertentes *representacionalistas* da Ciência Cognitiva: a Inteligência Artificial e as Redes Neurais Artificiais. Entretanto, uma nova vertente não-representacionista surgiu no início desta década, com os trabalhos de Van Gelder (1998), denominada de Ciência Cognitiva Dinâmica. Em linhas gerais, essa vertente tem como hipótese a idéia de que agentes cognitivos são sistemas dinâmicos e não sistemas meramente representacionais, como supõem as duas primeiras. Ele questiona a eficiência e a necessidade dos modelos

representacionistas da cognição. Nesse sentido, a Teoria da Harmonia não seria descartada do contexto dessa nova vertente da Ciência Cognitiva, pois, uma das características da dinâmica dos agentes cognitivos envolvidos é a habilidade de processamento de informação.

Após todos esses anos de incansáveis pesquisas, os cientistas cognitivos têm buscado e apresentado diversos caminhos que levam à *explicação de comportamentos inteligentes*. Uma alternativa para tal explicação foi apresentada aqui, entretanto, percebemos que o terreno para essas discussões é fértil e necessita de mais trabalhadores e idéias para lavrá-lo em toda a sua extensão.



BIBLIOGRAFIA

ARBIB, M. A. *Brains, machines and mathematics*. New York: Springer Verlag, 1987.

ASHBY, W.R. Principles of the self-organizing system. In: VON FOERSTER, H., ZAPF, G.W. (Org). *Principles of self-organization*. Oxford: Pergamon, 1962. p.255-78.

BEALE, R., JACKSON, T. *Neural computing: an introduction*. Bristol: Institute of Physics Publishing, 1994.

BRETON, P. *História da informática*. São Paulo: Ed. Unesp, 1991.

BUTTON, G., COULTER J., LEE, R. R. J., SHARROCK, W. *Computadores, mentes e conduta*. São Paulo: Ed. Unesp, 1998.

CAUDILL, M., BUTLER, C. *Understanding neural networks: Computer Explorations*. Cambridge: MIT Press, 1992. v.1.

CHENG, B., TITTERINGTON, D.M. Neural networks: a review from a statistical perspective. *Statistical Science*. (New York), v. 9, p.2-54, 1994.

CUTLAND, N. *Computability: an introduction to recursive function theory*. Cambridge: University Press, 1980.

DUPUY, J.P. *Nas origens das ciências cognitivas*. São Paulo: Ed. Unesp, 1996.



EDWARDS, E. *Introdução à teoria da informação*. São Paulo: Cultrix, 1971.

FERREIRA, A.B.H. *Dicionário Aurélio básico da língua portuguesa*. Rio de Janeiro: Nova Fronteira, 1988.

GARDNER, H. *A nova ciência da mente: uma história da revolução cognitiva*. São Paulo: Edusp, 1995.

GONZALES, M.E.Q. Redes neurais e representação mental: um ensaio sobre a harmonia e racionalidade. *Transformação*. (São Paulo), v.14, p.93-108, 1991.

_____. Um estudo cognitivo- informacional das representações mentais. In: ABRANTES, P. C. (Org). *Epistemologia e cognição*. Brasília: Ed. Universidade de Brasília, 1993. p.127-46.

_____. Informações e cognição: uma proposta de (dis)solução do problema mente-corpo. In: ENCONTRO BRASILEIRO/INTERNACIONAL DE CIÊNCIAS COGNITIVAS, 2, 1996, Campos dos Goytacazes. *Anais...* Campos dos Goytacazes, 1996. p.53-60.

_____. Ação, causalidade e ruído nas redes neurais auto-organizadas. In: DEBRUN, M., GONZALES, M.E.Q., PESSOA JUNIOR, O (Org). *Auto-Organização: Estudos interdisciplinares em filosofia, ciências naturais e humanas, e artes*. Campinas: UNICAMP - Centro de Lógica, Epistemologia e História da Ciência, 1996. Coleção CLE.

_____. Auto-organização e perspectivismo: algum acréscimo à ciência cognitiva? In: GONZALES, M. E. Q., BROENS, M. C. (Org). *Encontro com as Ciências Cognitivas*. Marília. Faculdade de Filosofia e Ciências, 1998. p.3-13.



HARTLEY, R. V. L. Transmission of information. *System Technical Journal*. v.7, p.535-63, 1928.

HEBB, D. O. *Organization of behavior*. New York: John Wiley, 1949.

HODGES, A. *Alan Turing: the enigma*. London: Vintage, 1992.

KOHONEN, T. *Self-organization and associative memory*. New York: Springer Verlag, 1987.

LIPPMANN, R. P. An introduction to computing with neural nets. *ASSP Magazine*, p.4-23, Apr. 1987.

LITTLEJOHN, S. W. *Fundamentos teóricos da comunicação humana*. Rio de Janeiro: Zahar, 1982.

MARR, D. *Vision: a computational investigation into the human representation and processing of visual information*. New York: W.H. Freeman and Company, 1982.

MCCLELLAND, J. L., RUMELHART, D. E. *Explorations in parallel distributed processing: a handbook of models, programs and exercises*. Cambridge: MIT Press, 1988.

MCCULLOCH, W. S., PITTS, W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, v.5, p.115-33, 1943.



NEWELL, A. ., SIMON, H. GPS a program that simulates human thought. In: FEINGEMBAUM & FELDMAN (Ed). *Computers and thought*. New York: McGraw-Hill, 1968. p.279-93.

PEREIRA JUNIOR, A., GONZALES, M.E.Q. Informação, organização e linguagem. In: ÉVORA, F. R. R. (Org). *Espaço e tempo*. Campinas: UNICAMP - Centro de Lógica, Epistemologia e História da Ciência, 1995. v.15 255-89. (Coleção CLE).

ROGERS. JUNIOR H. *Theory of recursive functions and effective computability*. New York: McGraw-Hill, 1987.

RUMELHART, D. E., MCCLELAND, J. L. et al. *Parallel distributed processing: explorations in the microestruure of cognition*. Cambridge: MIT Press, 1986. v.1.

SHANNON, C. The mathematical theory of communication. In: SHANNON, C., WEAVER, W. *The mathematical theory of communication*. Urbana: University of Illinois Press, 1949. p.29-125.

SIMONS, G. L. *Introdução à inteligência artificial*. Lisboa: Clássica, 1986.

SMOLENSKY, P. Information processing in dynamical systems: foundations of harmony theory. In: RUMELHART, D. E., MCCLELAND, J. L. et al. *Parallel distributed processing: explorations in the microestruure of cognition*. Cambridge: MIT Press, 1986. v.1.

_____. IA Conexionista, IA Simbólica e Cérebro. In: ANDLER, D. (Org). *Introdução às Ciências Cognitivas*. São Leopoldo: Ed. Unisinos, 1988.



SOARES, A. *O que são ciências cognitivas*. São Paulo: Brasiliense, 1993.

TEIXEIRA, J. F., GONZALES, M.E.Q. Inteligência artificial e teoria de resolução de problemas. *Transformação*. (São Paulo), v.6, p.45-52, 1983.

TURING, A. M. On computable numbers, with an application to the entscheidungsproblem. In: DAVIS, M. (Ed.). *The Undecidable*. New York: Raven, p.116-51, 1965.

_____. Computadores e inteligência. In: EPSTEIN, I. (Org.). *Cibernética e comunicação*. São Paulo: Cultrix, 1973. p.45-82.

_____. Proposal for development in the mathematics division of an automatic computing engine (ACE). In: CARPENTER, B. E., DORAN, R.W. (Ed.). *A. M. Turing's ACE Report of 1946 and other papers*. Cambridge: The MIT Press, 1988. p.20-105.

VAN GELDER, T. The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences*, v.21, n.5, p.615-28, 1998.

WEAVER, W. Recent contribuitions to the mathematical theory of communication. In: SHANNON, C., WEAVER, W. *The mathematical theory of communication*. Urbana: University of Illinois Press, 1949. p.3-28.

WIENER, N. *Cybernetics, or control and communication in the animal and the machine*. Cambridge: MIT Press, 1961.



WEIZEMBAUM, J. *Computer power and human reason: from judgment to calculation*. San Francisco: Freeman, 1976.



BEGOSSO, L. C. *Ciência Cognitiva: uma abordagem informacional da mente*. Marília, 1999. 123p. Dissertação (Mestrado em Filosofia - Linha de Pesquisa: Ciência Cognitiva e Filosofia da Mente) - Faculdade de Filosofia e Ciências, Campus de Marília, Universidade Estadual Paulista "Júlio de Mesquita Filho".

RESUMO

Neste estudo, investigamos aspectos do aparato cognitivo a partir dos pressupostos centrais da Ciência Cognitiva, adotando, em particular, a hipótese de que os computadores podem constituir bons modelos para explicar alguns aspectos da inteligência humana. Para isso, apresentamos uma análise sobre técnicas de modelagem e modelos computacionais. Estudamos, também, a Teoria da Informação, investigando sua possível contribuição para a Ciência Cognitiva. Argumentamos que o modelo de redes neurais artificiais, desenvolvido por Smolensky, fornece subsídios para uma discussão sobre os elos entre a teoria da informação e os modelos neurais informacionais aplicados ao estudo cognitivo humano.

Palavras-Chave: Ciência cognitiva; Inteligência artificial; Redes neurais; Teoria da informação; Teoria da harmonia.



BEGOSSO, L. C. Cognitive Science: an approach informational of mind. Marília, 1999. 123p. Dissertação (Mestrado em Filosofia - Linha de Pesquisa: Ciência Cognitiva e Filosofia da Mente) - Faculdade de Filosofia e Ciências, Campus de Marília, Universidade Estadual Paulista "Júlio de Mesquita Filho".

ABSTRACT

This study investigates aspects of cognitive system from the Cognitive Science perspective, considering the hypothesis that computers can provide good models to explain some aspects of human intelligence. For that, we present an analysis of modeling techniques and computational models. We also study Information Theory investigating its possible contribution to Cognitive Science. We argue that the artificial neural network developed by Smolensky provides subsidy for a discussion about possible links between Information Theory and neural networks applied to human cognitive studies.

Keywords: Cognitive science; Artificial intelligence; Neural networks; Information Theory; Harmony Theory.



