

PS-908

ARTIFICIAL NEURAL NETWORKS A METHOD OF BEHAVIORAL REPRESENTATION

Jamir Mendes Monteiro (Universidade Paulista – UNIP, São Paulo, Brasil)
jamirmonteiro@hotmail.com

Anthropological researches, along the time, have pointed out that the main characteristic of human being is their thinking capacity, the solution procurement for their problems. This fact allied to the modern technologies represent, for the first time in history, that the knowledge keeps growing despite of controls, political differences, culture and region diversities. This paper represents a theoretical essay with the purpose of showing the evolution trends of carrying principles of human intelligence to the machines (Artificial Intelligence), whose main topic are the Artificial Neural Networks. The results of this study suggest that the transference evolution speed from human tasks to robots might be amplified, chiefly by Molecular Genetic and Quantum Physics researches. In contrast, the more this evolution grows, the more people be excluded from the process. Find out ways to attenuate that seems to be the challenge of the millenium!

Keywords: Artificial Neural Networks; Artificial Intelligence; Knowledge Representation

REDES NEURAIS ARTIFICIAIS UM MÉTODO DE REPRESENTAÇÃO COMPORTAMENTAL

Pesquisas antropológicas, ao longo do tempo, têm enfatizado ser a capacidade de pensar a principal responsável pela evolução dos seres humanos, de procurar soluções para seus problemas. Esse fato, aliado às modernas tecnologias, representa, pela primeira vez na história, que o conhecimento evolui a despeito de controles, diversidades políticas, culturais e de região. Esse artigo representa um ensaio teórico com o objetivo de mostrar as tendências de evolução da transmissão dos processos de inteligência humanos para as máquinas (Inteligência Artificial), cujo tópico principal são as *Redes Neurais Artificiais*. O resultado deste estudo sugere que a velocidade da transferência de tarefas humanas para robôs tende a ser ampliada, em função principalmente das pesquisas em Genética Molecular e Física Quântica. No entanto, quanto mais acelerado esse processo, maior a exclusão de grande parte da humanidade. Pesquisar caminhos que atenuem esse processo é o grande desafio do milênio!

Palavras-chave: Redes Neurais Artificiais; Inteligência Artificial; Representação do Conhecimento.

REDES NEURAIIS ARTIFICIAIS

Um método de representação comportamental

1. INTRODUÇÃO

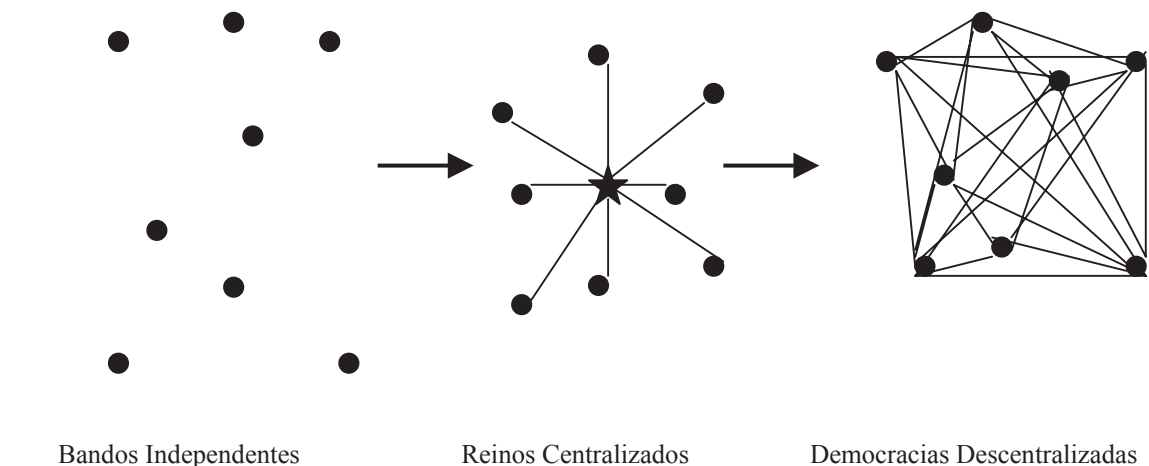
Substituir o homem pela máquina, desde os tempos primitivos, sempre foi um caminho para a busca de melhor eficiência nas atividades do dia-a-dia que sustentavam a existência da espécie. No princípio foram as ferramentas, que podiam servir de utensílios para a confecção de armas de caça; depois, com a descoberta do fogo, a transformação desse ferramental para o tratamento do ferro, surgindo então a indústria primitiva, cuja dedicação, no início, se orientava para as armas de caça e de guerra. Ao longo dos tempos, com a evolução das estruturas sociais, surgiram as primeiras organizações de trabalho, dando origem à estrutura política que rege a sociedade como um todo.

Farto material antropológico sustenta que,

Há dois milhões de anos, a espécie humana vivia nas planícies da África. À semelhança dos seus ancestrais, viviam da caça aos animais selvagens e da coleta de frutas silvestres. Gradualmente, aqueles seres humanos foram se espalhando ao redor do globo, evoluindo, nos últimos cem mil anos, para a moderna raça humana. (MALONE, 2004, p. 16, tradução do autor).

A *Figura 1* demonstra, de forma simplificada, segundo esse mesmo autor, a sequência de mudanças ocorrida na organização das sociedades desde o aparecimento da raça humana.

Figura 1: Os principais caminhos trilhados pelas sociedades em evolução ao longo da história revelam padrões referenciais que podem estar direcionando a mudança dos negócios neste século



Fonte: MALONE (2004, p. 16, tradução do autor).

Da vida em isolamento, descentralizados em bandos de caçadores e coletores de frutas, nossos ancestrais vagarosamente se organizaram em grandes reinos centralizados, talvez impulsionados pelo advento da agricultura. Por quê, organizados em reinados por centenas

de anos, as sociedades evoluíram para organizações sociais democráticas, nos últimos duzentos anos, é uma questão a ser respondida (MALONE, 2004). Entre outras razões, esse autor admite como principal fator a evolução e a redução dos custos das comunicações ao longo do tempo.

À semelhança da evolução das sociedades, parece que as organizações trilham o mesmo caminho. Do trabalho artesanal independente para as organizações centralizadas, sob o domínio de chefes e supervisores, padronizaram-se os produtos, os procedimentos, reduziram-se os custos, e, em nome da sociedade de consumo, o homem foi paulatinamente perdendo sua liberdade de ir-e-vir (MALONE, 2004).

Com o advento da revolução industrial, a tecnologia foi direcionada para a produção em massa, com o intuito de reduzir custos e aumentar a produtividade. Pouco a pouco, as máquinas foram desenvolvidas para substituir o trabalho de rotina, em princípio onde se exigiam procedimentos rigorosos de produção ou onde as possibilidades de risco de acidentes eram maiores. Com a evolução da eletrônica e dos computadores, princípios de inteligência foram desenvolvidos no sentido de que as máquinas desenvolvessem atividades que antes somente os seres humanos tinham capacitação.

A esperança era que o tempo dedicado ao trabalho pelo homem fosse reduzido, aumentando-se o tempo a ser dedicado à família e ao entretenimento. No entanto, o que se viu foi o aumento gradativo do desemprego e a criação de situações de instabilidade social, como o aumento da criminalidade e a desestruturação do tecido social (DRUCKER, 2002).

Este artigo representa um ensaio teórico que contempla um tópico da tecnologia que visa dotar a máquina de inteligência, substituindo o ser humano nas suas atividades essenciais de trabalho: as **Redes Neurais Artificiais (RNAs)**. Além dessa introdução, a segunda seção trata do histórico evolutivo dessa forma de tecnologia, descrevendo o seu aparecimento e as tendências de evolução futura; a terceira seção relata os modelos de aprendizado identificados nos seres humanos e a forma como estão sendo representados pelas RNAs; a quarta seção descreve os principais tipos de redes e as suas características mais importantes. A quinta seção será dedicada às principais aplicações dessa tecnologia. A sexta e última seção contemplará as considerações finais deste estudo.

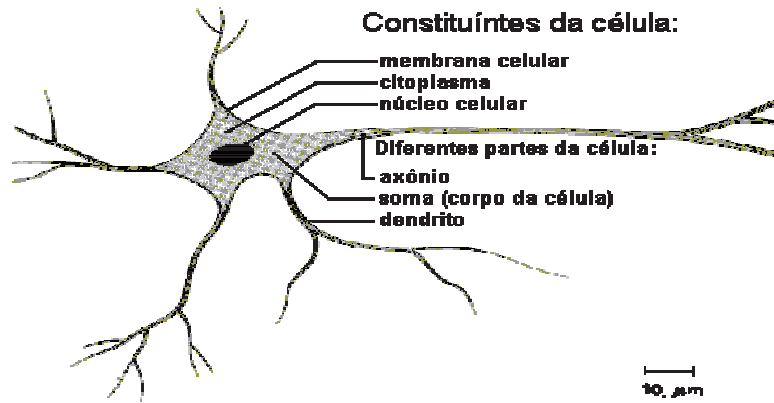
2. HISTÓRICO EVOLUTIVO DAS RNAs

Redes Neurais Artificiais (RNAs) são “Sistemas Paralelos Distribuídos, compostos por unidades de processamento simples (nodos) que calculam determinadas funções matemáticas, normalmente não lineares” (BRAGA; CARVALHO; LUDERMIR, 2000, p. 1). Constituem um tópico da Inteligência Artificial (IA), ramo de estudo que pretende desenvolver nas máquinas princípios de inteligência naturais dos seres humanos.

Elaborar um modelo de Rede Neural Artificial subentende estudar o modelo do cérebro humano no sentido da interpretação, reconhecimento, relacionamento e criação do “conhecimento”, o que torna essa tarefa bastante árdua, tendo em vista que a forma como opera o cérebro humano, ainda nos dias atuais, é bastante pouco conhecida pela ciência humana. O que se sabe a respeito do funcionamento do cérebro humano, a partir de

investigações biológicas realizadas, no decorrer do século passado, diz respeito mais à sua estrutura do que à sua forma de funcionamento. A *Figura 2.1* retrata com precisão a estrutura do cérebro humano e os seus principais componentes:

Figura 2.1 – Representação Gráfica do Neurônio Biológico



Fonte: PIAZZI (2001)

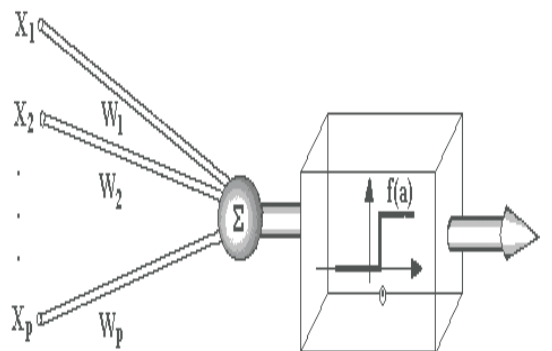
A idéia, representada pela figura, é a de que o cérebro humano funciona na modalidade *Entrada – Processamento – Saída* que caracteriza o funcionamento dos computadores e de todas as máquinas que se utilizam de processadores e memória, mais conhecidas como máquinas de controle numérico. Seus principais componentes são os seguintes:

- **Dendritos:** tem por função receber os impulsos nervosos, oriundos de outros neurônios, e conduzi-los até o corpo celular
- **Corpo Celular:** é o processador, por excelência, do neurônio, processando e gerando novos impulsos
- **Axônio:** é a saída da célula, o componente que a liga ao dendrito do neurônio seguinte
- **Sinapse:** representa o ponto de contato entre um axônio da célula anterior e o dendrito da célula seguinte. É através desse sinal que os neurônios se unem funcionalmente formando as redes neurais.

As sinapses funcionam como válvulas, controlando os impulsos entre os nodos. Como cada neurônio recebe impulsos de alguns outros, o processo de verificação de qual “sinal” passar para o neurônio seguinte se dá através de um processo de comparação “entre sinais”, passando para a célula seguinte o sinal mais forte. Naturalmente, isso significa que “a capacidade de realizar funções complexas surge com a operação em paralelo de todos os 10 elevados a 11 nodos do nosso cérebro” (BRAGA; CARVALHO; LUDERMIR, 2000, p. 6).

Após 20 anos de estudos, o psiquiatra e neuroanatomista Warren McCulloch, em trabalho conjunto com Walter Pitts, um matemático recém-graduado, que se associou a ele em 1942, publicou em 1943 o primeiro modelo artificial de um neurônio biológico. Esse trabalho, publicado em 1943 com o título de “*A Logical Calculus of the Ideas Immanent in Nervous Activity*” deu início às pesquisas decorrentes sobre as Redes Neurais Artificiais, daí para frente simplesmente chamadas de Redes MCP, em homenagem ao trabalho pioneiro desses autores. A *Figura 2.2* representa graficamente a estrutura do neurônio biológico artificial:

Figura 2.2 – O Neurônio Artificial



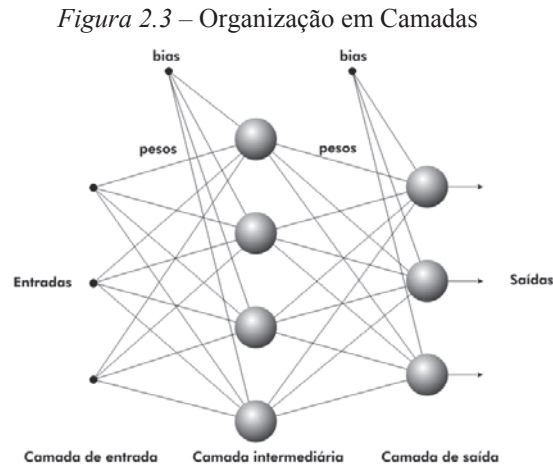
Onde, os X s representam os sinais de entrada, os quais são multiplicados um a um pelos seus pesos correspondentes W s, e totalizados em Σ . A soma correspondente é, então, comparada com um certo limite (*Threshold*), gerando um sinal de saída sempre que exceder o seu valor.

Fonte: IMAGENS(2006)

Representar artificialmente a estrutura de um neurônio, logo se notou, era insuficiente para inserir nas máquinas a capacidade de “pensar”. Estudos posteriores revelaram que os atributos mais importantes da capacidade mental do indivíduo são as técnicas de aprendizado, sem as quais de pouco adianta a estrutura cerebral. Em 1949, Donald Hebb publicou o primeiro trabalho relacionado com o aprendizado. Com seus estudos, Hebb “demonstrou como a plasticidade da aprendizagem de redes neurais é conseguida através da variação dos pesos de entrada dos nodos” (BRAGA; CARVALHO; LUDERMIR, 2000, p. 3). Sua proposta teórica explica o aprendizado em nodos biológicos através do reforço das ligações sinápticas entre nodos excitados. A Regra de Hebb, como é conhecida, foi interpretada do ponto de vista matemático e é, até hoje, utilizada em vários algoritmos de aprendizado.

Mais tarde, Widrow e Hoff sugeriram uma regra de aprendizado, conhecida como regra de Widrow-Hoff, ou *Regra Delta*, hoje a mais popular entre elas, e que se baseia no método do gradiente para minimização do erro na saída de um neurônio com resposta linear.

Em 1958, Frank Rosenblatt apresentou à comunidade acadêmica um novo modelo, o *Perceptron*, cuja principal característica era atingir a capacidade de treinamento para certos tipos de padrões, através do ajuste das sinapses nos nodos. Essa topologia de RNA apresenta três camadas distintas: a primeira recebe as entradas do exterior e possui conexões fixas; a segunda recebe impulsos da primeira, ajustáveis através de pesos, e as envia para a terceira camada, de respostas, conforme representação gráfica na *Figura 2.3*.



Fonte: IMAGENS (2006)

Em função da sua estrutura, esse tipo de *Perceptron* processa apenas padrões, linearmente separáveis, pois, embora inicialmente a saída da rede seja aleatória, pelo ajuste gradual dos pesos, a saída converge para o modelo do conjunto de treinamento.

Em 1969, dois pesquisadores, Minsky e Papert, demonstraram que o *Perceptron* original resolvia apenas problemas linearmente separáveis, deixando de atender um vasto campo de problemas, não linearmente separáveis, como aqueles com atributos de paridade, conectividade e simetria. O principal argumento desses pesquisadores era que o tempo e o espaço, despendidos para a solução de problemas complexos, cresceria com o tempo, vindo a inviabilizar o emprego das RNAs e do próprio *Perceptron* na solução desses problemas. A grande repercussão desse trabalho provocou o adormecimento das pesquisas em torno da abordagem conexionista, durante toda a década de 1970, embora alguns pesquisadores dessem continuidade aos seus trabalhos, como Igor Aleksander (Redes sem Pesos) na Inglaterra, Kunihiko Fukushima (Cognitron e Neocognitron) no Japão, Steven Grossberg (Sistemas Auto-adaptativos) nos EUA e Teuvo Kohonen (Memórias Associativas e Auto-organizadas) na Finlândia.

Em 1982, finalmente, John Hopfield publicou um artigo que restabeleceu a confiança sobre as propriedades associativas das RNAs, quando demonstrou a relação entre redes recorrentes auto-associativas e sistemas físicos, abrindo espaço para a utilização de teorias da física para estudar tais modelos. Alguns anos mais tarde, a descrição do algoritmo de treinamento *back-propagation* demonstrou que a visão de Minsky e Papert sobre o *Perceptron* era bastante pessimista, acentuando que as RNAs de múltiplas camadas são capazes de solucionar “problemas difíceis de aprender”.

A partir daí, restaurados os estudos sobre as RNAs, o desenvolvimento tecnológico, verificado a partir de meados da década de 1980, impulsionou a implementação de modelos de aprendizado em máquinas, capazes de substituir grande parte das atividades rotineiras executadas pelos seres humanos.

3. MODELOS DE APRENDIZADO

“O aprendizado em Redes Neurais é de relevante importância e alvo de intenso trabalho de pesquisa, desenvolvido sob ambos os pontos de vistas: o biológico e o artificial.” (KARTALOPOULOS, 1996, p. 42, tradução do autor)

Pesquisadores comportamentalistas do mundo inteiro tentam, através de estudos comparativos, encontrar respostas adequadas para as seguintes questões: Como aprendemos? Qual o processo mais eficiente para o aprendizado? Qual a quantidade de informações e rapidez com que conseguimos assimilar? Quais os bloqueios que impedem o aprendizado?

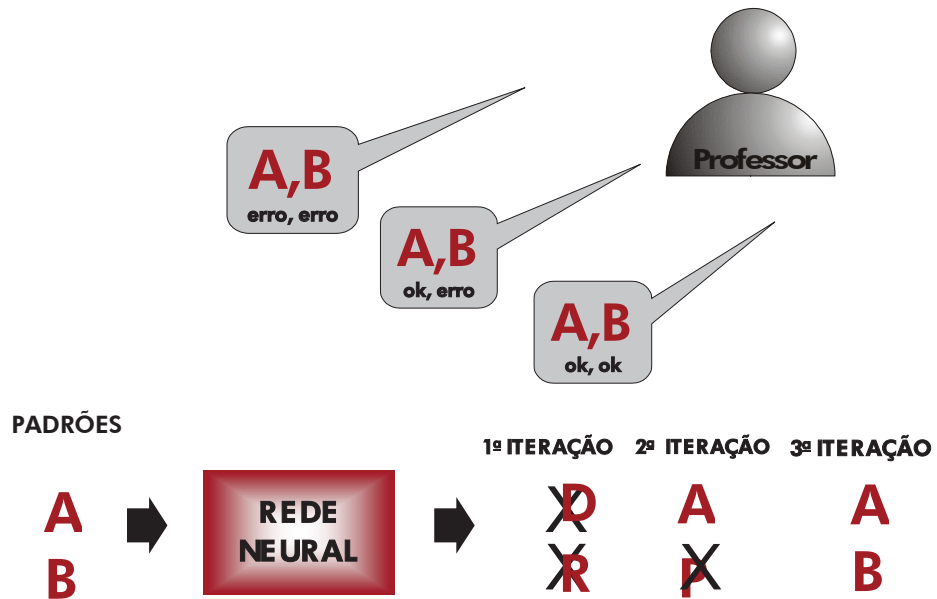
Esses estudos, conduzidos através de testes em animais de diferentes tipos de inteligência, em seres humanos com idade e maturidade diferenciadas, em seres primitivos e na vida biológica marinha, atestam que existem diferentes processos de aprendizagem, cada um deles próprios de cada espécie em particular. Além dessa dificuldade, de caracterizar propriamente os diferentes tipos de processos de aprendizagem, o maior desafio está na representação desses procedimentos em circuitos eletrônicos ou programas de computador, capazes de, através das redes neurais artificiais, fazer com que as máquinas se comportem como seres vivos pensantes nas atividades para as quais são desenvolvidas.

Aprender, como dizem os estudiosos, é um processo constante de ajuste das redes neurais aos estímulos recebidos, gerando, em consequência, uma determinada resposta. Em linhas gerais, é a extração de informações relevantes de padrões de informações apresentados, criando, assim, uma representação própria para o problema. “Na realidade, durante o processo de aprendizagem, a rede ajusta continuamente seus parâmetros, os *pesos sinápticos*, em resposta aos estímulos de entrada de tal forma que a resposta real encontrada converge para a resposta desejada para o problema.” (KARTALOPOULOS, 1996, p. 43, tradução do autor) Quando a resposta encontrada é a mesma que a resposta desejada para o problema, diz-se que a rede completou sua fase de aprendizagem; caso contrário, declara-se que apenas “adquiriu conhecimento”.

São os seguintes os principais tipos de modelos de aprendizagem:

- **Aprendizagem Supervisionada:** as entradas e saídas desejadas para a rede são fornecidas por um supervisor externo. Exemplo: Regra Delta e sua generalização para redes de múltiplas camadas; o algoritmo de retro-propagação. A *Figura 3.1* representa graficamente este modelo,

Figura 3.1 – Treinamento Supervisionado

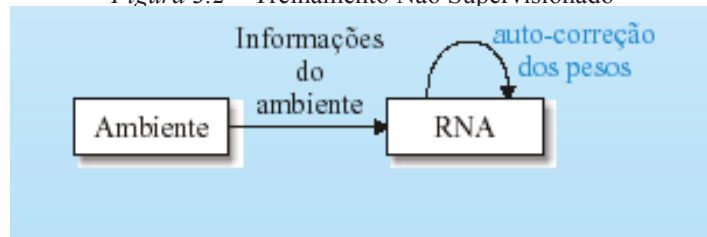


Fonte: PIAZZI (2001)

Durante a sessão de treinamento da Rede Neural, um sinal de entrada é aplicado, resultando num sinal de saída. Essa resposta é comparada com um sinal desejado, calculado a priori. Caso os sinais sejam divergentes, a Rede Neural gera um sinal de erro, o qual é utilizado para ajustar o peso sináptico aplicado, de tal forma que a nova saída convirja para a saída desejada. Em outras palavras, o erro é minimizado, tendendo no final a zero. A principal limitação desse modelo é que, na ausência do supervisor ou professor, a Rede não consegue aprender novas estratégias.

- **Aprendizagem por Reforço:** é um caso particular de Aprendizagem Supervisionada, cuja diferença está em que não cabe ao supervisor, durante o processo de aprendizagem, indicar quão perto da saída desejada está a saída calculada, mas sim declarar se a saída bate com o resultado esperado ou não. O supervisor não apresenta a saída desejada para a Rede, apenas indicando um resultado binário, do tipo passa-não passa. Como esse tipo de aprendizagem não indica a tendência do resultado em direção ao esperado, alguns parâmetros devem ser estabelecidos, no sentido de evitar que o treinamento nunca acabe.
- **Aprendizagem Não Supervisionada:** caracteriza-se pela ausência de supervisor durante o processo de treinamento. O melhor exemplo desse modelo são os estágios iniciais dos processos de visão e audição nos seres humanos. A Figura 3.2 apresenta graficamente esse modelo:

Figura 3.2 – Treinamento Não Supervisionado

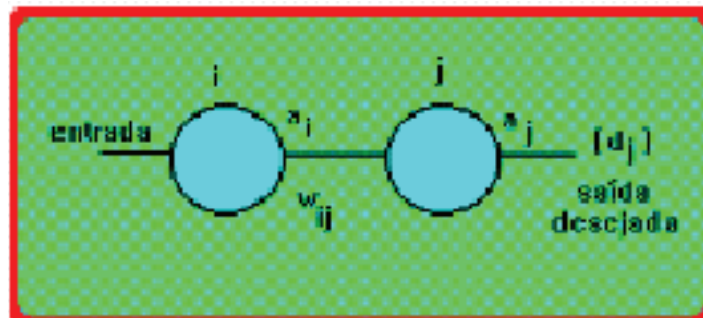


Fonte: IMAGENS (2006)

Durante a sessão de treinamento, a Rede Neural recebe os impulsos de entrada e os organiza, arbitrariamente, em categorias. Depois de treinada, quando a Rede recebe um estímulo de entrada ela o classifica indicando, como resposta, a categoria à qual pertence. Quando não encontra nenhuma categoria onde classificar o estímulo de entrada, a Rede imediatamente cria uma nova categoria. Embora esse modelo não exija a presença de um supervisor, mesmo assim não prescinde de informações de atributos que possibilitem a Rede formar as categorias.

- **Aprendizagem por Competição:** é um tipo particular de Aprendizagem Não Supervisionada, onde diversos neurônios competem entre si, comparando seus resultados de saída com o desejado, de forma a obter o que mais se aproxima dele, tornando-o sinal dominante e inibindo os demais neurônios. Para Redes muito extensas, devido a registros de conectividade, esse modelo pode não ser adequado, devendo-se tentar, então, o modelo de aprendizagem por reforço. A Aprendizagem por Competição é bastante empregada quando se tem um grupo de pessoas, onde cada uma delas é treinada numa atividade específica, dentro do princípio da pessoa certa, no lugar certo, no tempo certo.
- **A Regra Delta:** essa regra se baseia na idéia do ajustamento contínuo do valor dos pesos aplicados em cada nodo de forma que a diferença do erro (Delta) entre o valor desejado na saída e o valor real conseguido seja minimizada. Essa regra também é conhecida como Regra de Aprendizagem Widrow-Hoff [2] ou Regra dos Mínimos Quadrados, devido ela minimizar o erro mínimo quadrado. Na *Figura 3.3* temos uma representação gráfica dessa regra:

Figura 3.3 – Regra Delta



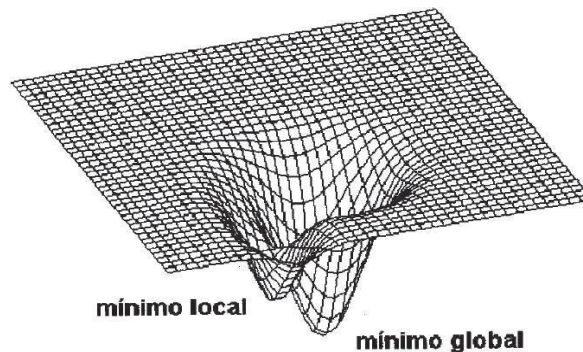
$$\Delta W_{ij} = (d_j - a_j) a_i$$

Fonte: IMAGENS (2006)

Dessa forma, obtém-se o seguinte esquema de treinamento:

- Utilizar, no início, em todas as conexões, pesos aleatórios
 - Para cada par de treinamentos, calcular a resposta obtida, aplicando os pesos ao valor de entrada
 - Calcular o erro, subtraindo da resposta desejada a resposta obtida
 - Se o erro não for satisfatoriamente pequeno, fazer: peso novo = peso anterior + taxa de aprendizado
 - Repetir o processo de cálculo até que o erro seja satisfatoriamente pequeno
- **A Regra do Gradiente Descendente:** neste método, os pesos são ajustados mediante um valor proporcional ao primeiro derivativo (gradiente) do erro encontrado entre a saída desejada e o resultado apurado para cada elemento da rede. O que se procura é minimizar a função Erro, procurando a melhor aproximação ao mínimo global, evitando-se sempre a utilização de mínimos locais. Na *Figura 3.4*, temos uma representação gráfica do conjunto solução para esse método de cálculo.

Figura 3.4 – Regra do Gradiente Descendente



Fonte: IMAGENS (2006)

O principal desafio nesse método é como convergir para o mínimo global, evitando-se a ocorrência de mínimos locais, num tempo razoável, racionalizando-se a utilização de espaços?

- **A Regra de Aprendizagem Hebbiana:** essa regra [2] baseia-se nos estudos de Donald Hebb que verificou que, no neurônio biológico, quando um axônio de uma célula *A* está próximo o bastante de uma célula *B*, a ponto de excitá-la, sua persistência nesse processo faz com que uma ou ambas as células mudem o seu metabolismo, de tal forma que a eficiência de *A*, como uma das células excitando *B*, seja aumentada. Dessa forma, a força do sinal sináptico, conhecida como peso *W*, entre a célula *A* e a célula *B*, é modificada de acordo com o grau de correlação entre a entrada e a saída da célula. O nome *Regra de Aprendizagem Anti-Hebbiana* é utilizado quando os contatos sinápticos são apenas inibitórios.

Generalizando, a *Rede Neural* comporta-se como um processo de comunicação que está continuamente adaptando-se a uma resposta desejada, através de uma função conhecida como “função de custo”.

Matematicamente falando,

Uma *Rede Neural* representa um sistema dinâmico que pode ser modelado como um conjunto de equações diferenciais conjugadas. Como ensina a Teoria do Controle por Realimentação (*Feedback Control Theory*), um sistema como esse depende dos valores que os parâmetros das equações podem assumir; pequenas mudanças nos parâmetros levam a resultados estáveis, instáveis, com pequenas oscilações, ou instáveis com oscilações catastróficas (caos). (KARTALOPOULOS, 1996, p. 47, tradução do autor)

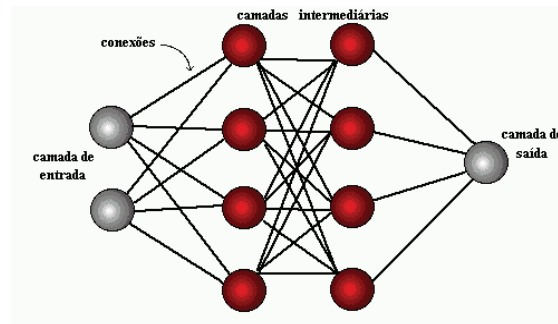
Portanto, Redes Neurais Biológicas têm como principais características:

- Poder de computação coletiva e sinérgica (neurocomputação), onde os programas são executados coletiva e sinérgicamente e as operações são descentralizadas, proporcionando processamento em paralelo
- Robustez, onde as operações são insensíveis a ocorrência de falhas, bem como insensíveis a estímulos de entrada parciais ou pouco representativas
- Aprendizagem, onde as redes apresentam poder de associação automática entre elas; a programação é criada pela própria rede durante o aprendizado e apresentam poder de adaptação com e sem supervisão, sem necessidade de intervenção do programador
- Operação Assíncrona. Não há indicações que assegurem que as Redes Neurais Biológicas utilizem o Relógio Biológico para sincronizar suas operações. Contrariamente, grande parte das Redes Neurais Artificiais utiliza relógios de sincronia para suas operações.

4. TIPOS DE REDES

Na literatura acadêmica, o desenvolvimento das Redes Neurais Artificiais, a partir dos aspectos comportamentais do mundo biológico, é conhecido como *Paradigma*. A pesquisa pelo paradigma mais representativo – aquele que melhor simule a Rede Neural Biológica – ainda não chegou ao seu fim. Um paradigma típico é estruturado em camadas de neurônios, onde alguns apresentam apenas uma camada: as Redes Neurais de Camada Única; outros, apresentam múltiplas camadas de neurônios: Redes Neurais de Múltiplas Camadas. A partir dessa estrutura topológica, à camada inicial é dado o nome de Camada de Entrada; à camada final, Camada de Saída, e às camadas intermediárias, Camadas Escondidas, devido a que as suas saídas não são facilmente legíveis. A *Figura 4.1* apresenta uma representação topológica de Rede:

Figura 4.1 – Representação topológica de Rede

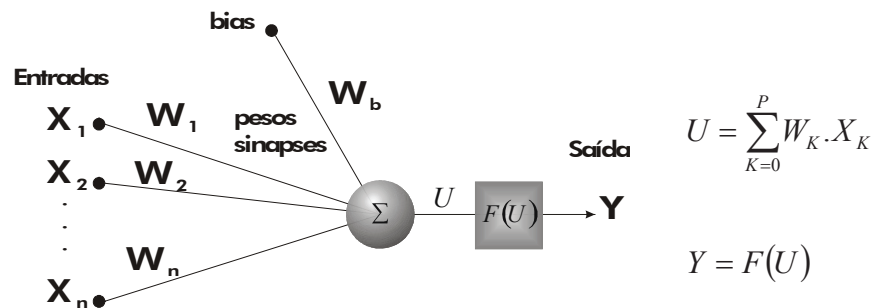


Fonte: IMAGENS (2006)

Dentro desses princípios, são os seguintes os principais tipos de Redes Neurais Artificiais:

- **Rede Perceptron:** o modelo de McCulloch-Pitts representava um sistema de realimentação aberto muito simples cujo único objetivo era a modulação de um único neurônio. Não havia nenhum mecanismo que proporcionasse a comparação do resultado esperado com a resposta calculada, bem como nenhuma ponderação por pesos, nem regras de aprendizagem. No entanto, mostrava um caminho, através do qual se conseguiria simular o processamento real dos neurônios biológicos, o que estimulou os pesquisadores a chegar ao modelo desenvolvido por Frank Rosenblatt, denominado de Rede Perceptron, que, em sua origem, trabalhava com uma camada de entrada, uma intermediária e outra, de saída. A Figura 4.2 apresenta um modelo dessa rede, com seus principais componentes:

Figura 4.2 – Rede Perceptron de Camada Única



Onde,

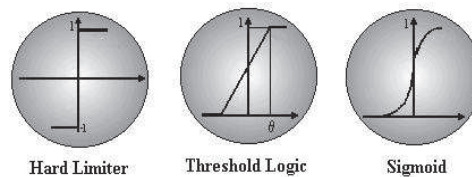
X representa os sinais de entrada;

W , os pesos a serem aplicados

$Bias$, a polarização que, conforme seu sinal, estimula ou inibe a sinapse;

U , a variável interna, que representa o resultado calculado e

F , que representa a Função de Ativação, cujas principais são as seguintes:

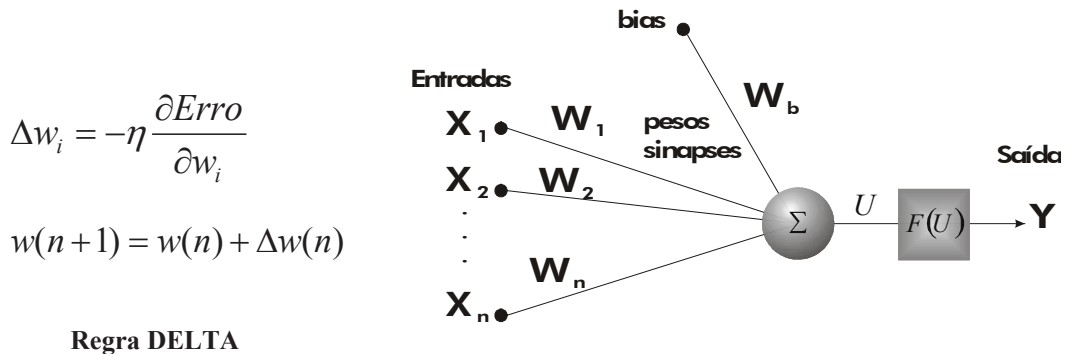


das quais a mais utilizada é a Função Sigmoidal.

Fonte: IMAGENS (2006)

- **Rede Adaline (ADaptive LINear Element** ou posteriormente **ADaptive LINear NEuron**): desenvolvido por Bernard Widrow quase na mesma época que a Rede Perceptron. Sua regra de treinamento é conhecida como Regra *Delta*, origem do desenvolvimento do Algoritmo de Retro Propagação, introduzido, em 1986, para tratamento do Perceptron de Múltiplas Camadas. Difere do modelo original do Perceptron porque a comparação com a saída desejada é realizada antes da aplicação da Função de Ativação, representando, portanto, uma função linear o que facilita o nível de compreensão. Na *Figura 4.3* temos uma representação desse modelo:

Figura 4.3 – Rede Adaline



Fonte: IMAGENS (2006)

- **Rede Perceptron de Múltiplas Camadas (MLP)**: representada por um conjunto de perceptrons de camada única, dentro de uma estrutura hierárquica, formando uma topologia do tipo *feedforward* (alimentação dianteira). O número de camadas intermediárias e o número de neurônios em cada camada não são fixos, devendo ser calculados por intuição ou por tentativa-e-erro. Cada aplicação, em função da sua complexidade, exige um determinado número de camadas e neurônios. O que se sabe é que uma Rede com uma camada intermediária é capaz de representar qualquer função contínua, sendo que, com duas camadas, pode ser implementada qualquer função mais complexa.

A funcionalidade dessa Rede decorre de:

- O processamento de cada nodo é definido pela combinação dos nodos anteriores conectados a ele
- À medida que se caminha em direção à camada de saída, as funções implementadas aumentam o seu grau de complexidade (essas funções definem como é realizada a divisão do espaço geométrico onde ocorre o conjunto de decisão)
- Numa rede com duas camadas intermediárias:
 - Na primeira camada, cada nodo traça retas no espaço geométrico onde figurará o conjunto solução
 - A segunda camada combina as retas traçadas pela camada anterior, formando uma região convexa, onde o número de lados é dado pelo número de nodos anteriores a ela conectados
 - A camada de saída organiza as regiões convexas de forma abstrata formando figuras muito difíceis de serem representadas matematicamente

- **Rede Neural Artificial FUZZY (FANN):** combina as propriedades principais das RNAs (associativa, adaptabilidade e aprendizado) com as regras de funcionalidade da Lógica *Fuzzy*. Uma das suas características mais importante é permitir que o algoritmo de treinamento gere automaticamente as funções de participação e a lógica *Fuzzy* durante e após o período de treinamento, agregando grande poder de adaptação ao processo. O neurônio *Fuzzy* se aproxima bastante do neurônio desenvolvido por McCulloch-Pitts, apresentando, no entanto, algumas diferenças: em lugar dos pesos escalares W e das funções de ativação, os neurônios *Fuzzy* utilizam conjuntos *Fuzzy*, trabalhando da seguinte maneira: quando os sinais de entrada são recebidos, o neurônio *Fuzzy* aplica a eles os pesos *Fuzzy*, os quais, após multiplicados por um fator de atenuação, são somados, representando o sinal cumulativo de entrada. Esse sinal é, então, mapeado para uma região de conjuntos *Fuzzy* onde são comparados com os níveis de ativação associados com o número de sinais ativos ou com os valores que representam a força do sinal cumulativo atenuado. Caso a força do sinal esteja acima de um valor mínimo, o neurônio é disparado; caso contrário, o mesmo é inibido. Se a função de ativação for um conjunto *Fuzzy*, a saída obtida não será o clássico sinal de disparo do neurônio, mas, sim, um sinal com graduação variável. Nesse caso, o valor indicativo do disparo (*threshold*) é também *Fuzzy*.

À medida que avançam as pesquisas, “o uso integrado das Redes Neurais Artificiais com a Lógica *Fuzzy* torna-se cada vez mais popular. [...] sendo encarado como um método versátil que integra o que de melhor existe em ambos os métodos, de forma a melhor solucionar os problemas para os quais a *FANN* é indicada como o meio mais eficiente de solução“. (KARTALOPOULOS, 1996, p. 155, tradução do autor)

5. PRINCIPAIS APLICAÇÕES

São vários os campos de aplicação das Redes Neurais Artificiais, associadas ou não à Lógica *Fuzzy*. Nos sistemas de controle de Metrô, redes especialmente treinadas são utilizadas para a determinação da velocidade do comboio, as partidas e as paradas eventuais; em máquinas de lavar roupa, são utilizadas para o cálculo da quantidade de água e a quantidade de ingredientes, como sabão em pó, alvejantes e amaciantes; em câmeras fotográficas e de vídeo, ajustam as cores, os contrastes, o brilho, foco e assim por diante. Em câmaras de vácuo, calculam o poder de sucção em função da quantidade e tamanho das partículas; em automóveis, com transmissão automática, determinam a sincronização; em sistemas veiculares inteligentes, calculam a melhor rota de trânsito e automaticamente guiam o veículo e, nos modernos sistemas de comunicação, processam sinais, encadeiam e roteiam canais e controlam todo o sistema.

Torna-se incalculável, em função da rápida evolução da tecnologia de comunicações e das pesquisas, desenvolvidas em Biologia Molecular, estimar-se o tamanho do mercado de aplicações das Redes Neurais Artificiais.

Além dos citados, pode-se ainda acrescentar os seguintes:

- **Modelagem:** aplicação em controles de processos, controles estatísticos de processos industriais, etc.

- **Otimização e Predição:** predição das vendas de um produto, otimização de roteiros, manutenção preditiva, análise de crédito, simulação de cotação de títulos e mercadorias, controle de tráfego, etc.
- **Processamento de Sinais e Imagens:** identificação de assinaturas, reconhecimento de faces, etc
- **Reconhecimento de Voz e Visão Computacional:** controladores robóticos, identificação de voz, etc.

6. CONSIDERAÇÕES FINAIS

O grande avanço, verificado na tecnologia de comunicações, a partir da segunda metade do século passado, fez da Internet o laboratório ideal para trabalhos multidisciplinares, utilizando recursos de inteligência do mundo inteiro. Organizações virtuais vem se especializando no sentido de aproximar a “necessidade” do recurso de inteligência que a irá desenvolver, esteja ele onde estiver. O trabalho, neste século XXI, caracterizar-se-á pela virtualidade, determinando mudanças na gestão dos negócios e na forma operacional das organizações (DRUCKER, 2002).

Construir autômatos que simulem perfeitamente o comportamento humano é um velho sonho da ciência. Entretanto fazer da ficção a realidade não será tarefa fácil. Nesse sentido, amplos esforços de pesquisa vêm sendo conduzidos no sentido de melhorar o entendimento da forma como os seres humanos interpretam a realidade e, através de mecanismos mentais, elaboram o conhecimento, o qual, depois de armazenado, serve de referência para interpretações e reinterpretações de novas percepções, criando um ciclo virtuoso que permite o despertar da sabedoria, como enfatizam os filósofos.

Nos dias atuais, fábricas que operam às escuras, totalmente operadas por robôs, são uma realidade, bem como supermercados onde o único elemento humano que se vê é o cliente. O incentivo mundial dado aos jovens estudantes de engenharia, através de campeonatos mundiais de robótica, estimula as pesquisas e fortalece o empenho em transferir para as máquinas modelos de inteligência humanos.

A consequência imediata, no entanto, desse “progresso” será o aumento gradual da legião dos excluídos, onde muitos não apresentam as mínimas condições de recuperação, criando uma crise social que já se percebe nos países tecnologicamente mais desenvolvidos. Qual o caminho a percorrer para atenuar tal disparidade é o grande desafio do milênio!

BIBLIOGRAFIA

- ABOUZEEDAN, Adli; BUSLER, Michael. **ASPEM as the New Topographic Analysis Tool for Small and Medium-Sized Enterprises (SMEs) Performance Models Utilization**. 2005. Disponível em: <<http://proquest.umi.com/pqdweb?cfc=1>> Acesso em: 21 de março de 2006.
- BISHOP, Christopher M. **Neural Networks for Pattern Recognition**. Great Britain: Oxford University, 2003. 482 p.
- BRAGA, Antonio de Pádua; CARVALHO, André Ponce de Leon F. de; LUDERMIR, Teresa Bernarda. **Redes Neurais Artificiais; teoria e aplicações**. RJ: LTC, 2000. 262 p.
- CHEN, Kuzn-Yu. **Using Back-Propagation Neural Network to Forecast the Production Values of the Machinery Industry in Taiwan**. Março de 2006. Disponível em: <<http://proquest.umi.com/pqdweb?cfc=1>> Acesso em: 21 de março de 2006.
- DRUCKER, Peter F. **Administrando em Tempos de Grandes Mudanças**. Tradução por Nivaldo Montingelli Jr. SP: Pioneira Thomson Learning, 2002.
- EDWARDS, David J.; CABAHUG, Ruel; LOVE, Peter E.D. **Intelligence and Maintenance Proficiency: an examination of plant operators**. 15 de fevereiro de 2005. Disponível em: <<http://proquest.umi.com/pqdweb?cfc=1>> Acesso em: 21 de março de 2006.
- IMAGENS. Disponível em: <<http://www.google.com>> Acesso em: 21 de março de 2006.
- KARTALOPOULOS, Stamatios V. **Understanding Neural Networks and Fuzzy Logic: basic concepts and applications**. New Jersey: IEEE Press, 1996. 205 p.
- MALONE, Thomas W. **The Future of Work: how the new order of business will shape your organization, your management style, and your life**. Boston: Harvard Business School Press, 2004. 225p.
- PIAZZI, Fábio Silva. **Introdução às Redes Neurais**. Novembro de 2001. Disponível em: <<http://www.google.com>> Acesso em: 21 de março de 2006.
- ROHDE, Jens. **Hierarchical Supply Chain Planning using Artificial Neural Networks to Anticipate Base-level Outcomes**. 2004. Disponível em: <<http://proquest.umi.com/pqdweb?cfc=1>> Acesso em: 21 de março de 2006.
- ROY Probir. **An Artificial Intelligence Stock Classification Model**. Setembro de 2005. Disponível em: <<http://proquest.umi.com/pqdweb?cfc=1>> Acesso em: 21 de março de 2006.
- SINANOGLU, Cem; BORKLU, H. Riza. **An Assembly Sequence-Planning System for Mechanical Parts using Neural Network**. 2005. Disponível em: <<http://proquest.umi.com/pqdweb?cfc=1>> Acesso em: 21 de março de 2006.
- SUKTHOMYA, Wimalin; TANNOCK, James. **The Training of Neural Networks to Model Manufacturing Processes**. Maio de 2003. Disponível em: <<http://proquest.umi.com/pqdweb?cfc=1>> Acesso em: 21 de março de 2006.