

CAPÍTULO II

MODELOS CONEXIONISTAS

Os modelos conexionistas se fundamentam nos estudos sobre a estrutura do cérebro humano para tentar emular sua forma inteligente de processar informação. Alguns estudos da neurofisiologia consideram que a riqueza computacional do cérebro humano vem do grande número de neurônios que estão interconectados por uma rede complexa de sinapses [CARV88].

Observando a *máquina inteligente* que é o cérebro humano, estima-se que a quantidade de neurônios existentes no mesmo está na casa dos bilhões. A velocidade de processamento destes componentes é baixa se comparada com a velocidade dos computadores tradicionais. Esta deficiência é superada pela imensa quantidade de neurônios existentes que operam em paralelo [SIMP90]. Estima-se que existam cerca de 10^{11} a 10^{14} neurônios operando em paralelo no cérebro humano. Cada um destes está conectado através de 10^3 a 10^4 sinapses em média [COTT85].

Tais características permitem ao cérebro humano executar rapidamente certas funções (por exemplo: reconhecer fisionomias). que os computadores convencionais não conseguem realizar com o mesmo desempenho. Na tabela 2, baseada em Cottrell [COTT85] e Simpson [SIMP90], é apresentada uma comparação das principais diferenças existentes entre os computadores tradicionais e o cérebro humano. Esta comparação permite ter uma idéia mais clara sobre a capacidade adaptativa do cérebro humano, em contraste com a rigidez e precisão dos computadores convencionais.

Tabela 2. Diferenças entre o computador tradicional e o cérebro humano

	COMPUTADOR TRADICIONAL	CÉREBRO HUMANO
Elementos Computacionais	processadores poderosos	neurônios simples
Velocidade de Processamento	10^{-9} s	10^{-3} s
Tipo de Processamento	serial	paralelo
Confiabilidade dos Elementos	confiável	não-confiável
Tolerância a Falhas	quase nenhuma	grande
Tipo de Sinal	precisos, simbólicos	imprecisos
Tipo de Controle	centralizado	distribuído
Armazenamento de Informação	substituível	adaptável

Uma definição, elaborada por Hecht-Nielsen [HECH88], de modelos connexionistas de computação, também chamados de redes neuronais artificiais (RNAs) redes neurais, ou sistemas de processamento paralelo distribuído (PDP), é apresentada a seguir:

Um modelo connexionista é uma estrutura de processamento de informação distribuída e paralela. Ela é formada por unidades de processamento, comumente chamadas de nós, neurônios ou células, interconectadas por arcos unidirecionais, também chamados de ligações, conexões ou sinapses. Os nós possuem memória local e podem realizar operações de processamento de informação localizada. Cada célula possui uma única saída (axônio), a qual pode se ramificar em muitas ligações colaterais (cada ramificação possuindo o mesmo sinal da saída do neurônio). Formalmente, o sinal de saída do nó pode ser equacionado de diversas maneiras. Todo o processamento que se realiza em cada unidade deve ser completamente local, isto é, deve depender apenas dos valores correntes dos sinais de entrada que chegam dos neurônios através das conexões. Estes valores atuam sobre os valores armazenados na memória local da célula.

Os principais elementos usados na descrição das RNAs são a representação distribuída, as operações locais e o processamento não-linear [SIMP90]. Estes

atributos especificam duas aplicações básicas dos modelos conexionistas: situações onde poucas decisões tem que ser tomadas a partir de uma grande quantidade de dados, e situações onde um complexo mapeamento não linear deve ser aprendido. Geralmente, o ferramental matemático usado nesta tecnologia inclui: equações diferenciais, sistemas dinâmicos, álgebra linear, probabilidade e estatística.

Neste capítulo é feita uma revisão sobre os modelos conexionistas. Na seção II.1 são descritos os componentes básicos destes modelos, e como se interrelacionam. Na seção subsequente, é dedicada uma atenção especial às técnicas de aprendizado utilizadas nos modelos conexionistas. Na seção 11.3 são apresentados os conceitos de convergência e estabilidade. Finalizando este capítulo, são tecidas considerações sobre alguns dos modelos clássicos encontrados na literatura, destacando suas aplicações e propriedades principais.

II.1 - COMPONENTES DOS MODELOS CONEXIONISTAS

De acordo com Rumelhart [RUME86A], um modelo conexionista pode ser descrito por oito elementos principais:

- um conjunto de unidades de processamento;
- um estado de ativação;
- uma função de saída;
- um padrão de interconexão;
- uma regra de propagação;
- uma regra de ativação;
- uma regra de aprendizado; e
- um ambiente onde o sistema deve funcionar.

Unidades de Processamento

As células são o meio de representação do conhecimento, por exemplo conceitos de um domínio, existente na rede neural. Os nós podem representar pontos ("pixels"), caracteres (letras, números), palavras, ou outros conceitos, dependendo da aplicação. Na figura 2 temos uma ilustração de um neurônio como unidade limiar. As entradas que chegam a ele representam os dentritos. Cada dentrito possui um sinal que é adicionado (Σ). Depois da adição, o sinal é processado através de uma função limiar $f()$, a qual produz um sinal de saída. Nesta figura, o neurônio limiar pode ser considerado como uma representação simplificada dos neurônios biológicos.

Nas RNAs existem dois tipos de representação do conhecimento possíveis: a localizada e a distribuída. A representação localizada corresponde a um único neurônio representando um conceito. Na representação distribuída, o padrão de ativação de um conjunto de unidades é que possui significado [RUME86A]. Neste trabalho, as unidades de processamento da rede são designadas pela letra u , seguida de um índice i que indica a posição que o neurônio ocupa na rede.

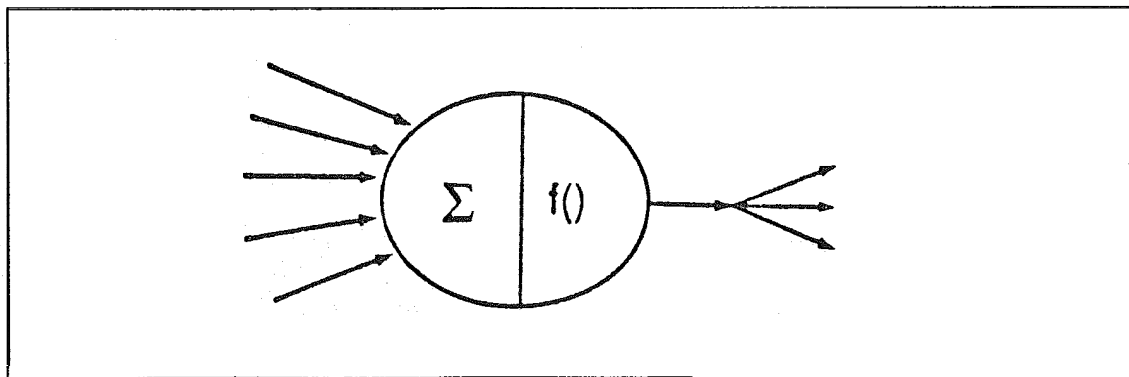


Figura 2. Neurônio como unidade limiar

Estado de Ativação

Cada célula u_i da rede computa um estado de ativação, que é um valor numérico líquido de saída. O cálculo desta ativação é computado a partir das ativações das células conectadas diretamente a este nó, e dos correspondentes pesos destas conexões.

O estado de ativação de todas as unidades da rede, ou seja, o estado de ativação do sistema, especifica o que está sendo representado na rede em um determinado instante t qualquer. Este estado de ativação do sistema pode ser representado por um vetor $a(t)$. Os valores das ativações existentes na rede podem ser discretos, por exemplo assumindo os valores $\{0,1\}$ ou $\{-1,0,+1\}$, como também podem ser contínuos, assumindo valores no intervalo $[0,1]$ ou $[-1,+1]$, que são computados pela regra de ativação a ser vista depois ([LIPP87], [GALLSS]).

Função de Saída

As unidades interagem entre si através de um valor que é transmitido pelas sinapses. Este valor é determinado pela ativação da unidade estimuladora. Formalmente, o valor de saída é dado por uma função do tipo $o_i(t) = g(a_i(t))$.

Padrão de Interconexão

Pode-se representar o padrão de interconexão da rede por uma matriz de pesos W , onde um elemento $w_{i,j}$ corresponde à influência da célula u_i sobre a célula u_j . Conexões com pesos positivos, chamadas de excitatórias, indicam reforço na ativação do neurônio u_j . Sinapses com pesos negativos, chamadas de inibitórias, indicam inibição na ativação da célula u_j . O conjunto das ligações excitatórias e inibitórias existentes na rede determina o comportamento da mesma.

Topologicamente, as RNAs podem ser organizadas em camadas. A camada de entrada da rede é instanciada externamente e não recomputa suas saídas. Portanto, não existem arcos de entrada em suas células. Os valores resultantes das células pertencentes a camada de saída são considerados os resultados finais da rede como um todo. Na figura 3, as unidades u_9 e u_{10} são consideradas as saídas da rede. As células que não pertencem nem à camada de entrada e nem à camada de saída são chamadas de intermediárias ou ocultas.

Observando a figura 3 quando u_{10} é ativado o valor de sua ativação é determinado pelas ativações de u_4 , u_5 , u_7 e u_8 e os pesos $w_{4,8}$, $w_{5,8}$, $w_{7,10}$ e $w_{8,10}$. Podem existir arcos que conectam nós da mesma camada, sendo chamados de sinapses intra-camadas. As ligações inter-camadas conectam células de camadas diferentes. No caso da rede da figura 3, os arcos que unem as unidades u_6 e u_7 são sinapses intra-camada inibitórias. A ligação que conecta o nó u_7 ao nó u_{10} é inter-camada excitatória.

Com relação aos arcos que conectam as unidades u_6 e u_7 , eles são chamados de ligações recorrentes. Estas conexões formam um ciclo, voltando para o neurônio de origem (aquele que foi ativado primeiro). Podem existir sinapses que ligam unidades de camadas diferentes, formando ciclos entre as mesmas. Por exemplo, uma ligação que conectasse a célula u_{10} a u_5 formaria um outro ciclo na rede da figura 3. Portanto, as RNAs podem ser classificadas em redes cíclicas, que possuem ciclos, e acíclicas, que não possuem ciclos [LIPP87].

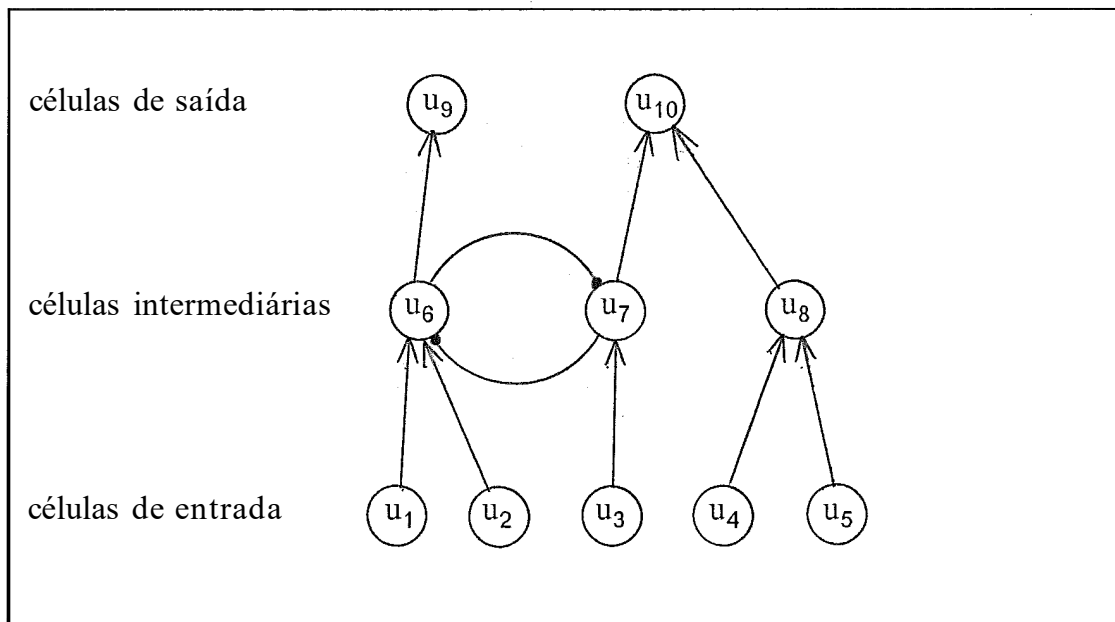


Figura 3. Exemplo de RNA

Regra de Propagação

Cada célula u_i computa sua nova ativação através de uma regra de propagação. Em geral, ela é definida como sendo uma função soma da

entrada líquida dos pesos ("net") das células u_j que estão diretamente conectadas a u_i conforme fórmula abaixo:

$$net_i = F(\sum_{j=1}^n w_{i,j}u_j - \Theta_i),$$

onde u_j é o estado da j -ésima unidade, $w_{i,j}$ é o peso da conexão da i -ésima para a j -ésima unidade e Θ_i é o limiar da i -ésima unidade. Este limiar, que pode ser nulo inclusive, deve ser superado para que ocorra a ativação da célula.

Existem variações da regra de propagação que utilizam os conceitos da lógica nebulosa proposta por Zadeh [ZADE65]. Geralmente, nestas definições emprega-se os operadores de máximo e mínimo. Estes operadores atuam sobre o chamado *produto "fuzzy"*, entre as entradas e os pesos dos nós, executando as operações de E e OU próprias dos conjuntos nebulosos. Neste caso, os pesos podem ser interpretados como graus de pertinência do conceito representado na entrada para o conceito representado na saída da célula ([ROMA88], [MACH89]).

Regra de Ativação

É necessário uma regra que calcule o valor de ativação de uma unidade no instante t . É preciso uma função f que calcule a nova ativação $a(t)$ utilizando as entradas líquidas (net). Geralmente, esta função possui a forma $a_i(t+1) = f(a_i(t), net_i(t))$, onde f é a função de ativação, também chamada de função limiar. Esta função mapeia os neurônios de entrada para um intervalo pré-especificado de saída. As quatro funções de ativação mais utilizadas são: linear, rampa, salto e sigmóide [SIMP90]. Uma ilustração das quatro é apresentada na figura 4.

A função linear, figura 4.(a), possui a seguinte equação: $f(x) = \alpha x$, onde α é uma constante real que regula a intensidade da atividade de x .

A figura 4.(b) representa a função limiar rampa. Ela é uma função linear limitada pelo intervalo $[-\gamma, +\gamma]$, definida pelas seguintes equações:

$$f(x) = \begin{cases} +\gamma & \text{se } x \geq \gamma \\ x & \text{se } |x| < \gamma \\ -\gamma & \text{se } x \leq -\gamma, \end{cases}$$

onde γ indica os valores de saída máximo e mínimo, sendo chamados de pontos de saturação.

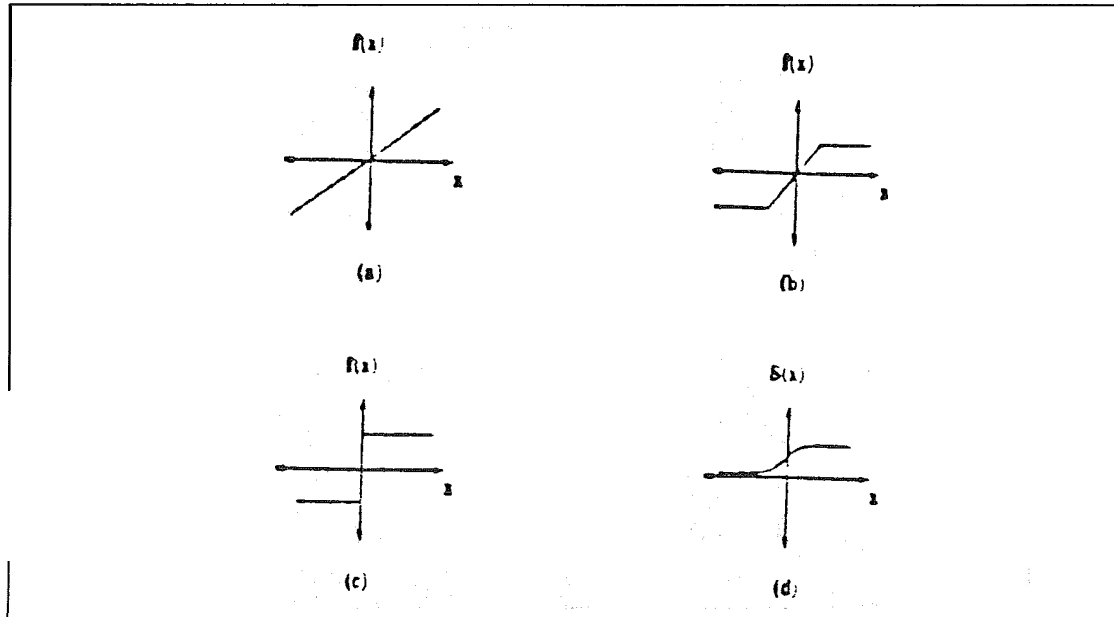


Figura 4. Exemplos de funções de ativação

A função salto, figura 4.(c), responde ao sinal de entrada emitindo o valor $+1$ se o sornatório for positivo, e -1 nos demais casos.

A figura 4.(d) representa a função sigmóide. Esta função limiar é monótona, não-decrescente, e sua resposta é gradual e não-linear. A função sigmóide mais conhecida é a função logística, cujos pontos de saturação são 0 e 1. A equação que a define é apresentada a seguir:

$$S(x) = (1 + e^{-x})^{-1}.$$

Regra de Aprendizado - modificando a conectividade através da experiência

A modificação do processamento ou da estrutura de conhecimento de uma rede neuronal envolve modificar seu padrão de interconexão [RUME86A]. Em princípio, isto pode ser feito de três maneiras:

1. Desenvolvimento de novas conexões;
2. Perda de conexões existentes na rede; e
3. Modificação dos pesos das conexões já existentes.

Quando o padrão de interconexão for uma matriz de pesos W , os itens (1) e (2) podem ser simulados através de (3). Tomando-se uma ligação com peso zero, e modificando-o para um valor positivo ou negativo, equivale a desenvolver esta sinapse. Da mesma forma, alterar o peso de uma conexão para zero significa desconectá-la. Portanto, as regras de aprendizado alteram os pesos das sinapses das redes através da experiência.

Em geral, as regras de aprendizado podem ser consideradas como uma variante da Regra de Hebb [HEBB49]. Ele estabeleceu o princípio de que a alteração da eficiência sináptica é a base do aprendizado, segundo o postulado apresentado a seguir:

Quando o axônio A se encontra próximo do axônio B de forma a poder excitá-lo, e o faz repetidas vezes, algum processo desconhecido provoca o crescimento de sinapses entre as células A e B, facilitando assim a excitação de B por A.

Especificamente, se uma unidade u_j recebe uma entrada de uma outra u_i , e ambas estão fortemente ativas, o peso $w_{i,j}$ (de u_i para u_j) deve ser fortalecido. Uma extensão desta idéia é apresentada na equação a seguir:

$$\Delta w_{i,j} = g(a_i(t), t_i(t)) \cdot h(o_j(t), w_{i,j}),$$

onde $t_i(t)$ é uma espécie de unidade de entrada *instrutora* (professor) da unidade u_i . Esta equação estabelece que a mudança no peso da conexão de u_i para u_j é resultado do produto da função $g()$ (ativação de u_i e sua entrada *instrutora* t_i) pela função $h()$ (valor de saída de u_j e peso $w_{i,j}$).

Uma variação desta regra apresenta $h(o_j(t), w_{i,j}) = o_j(t)$ e $g(a_i(t)) = \eta \cdot (t_i(t) - a_i(t))$, onde η é uma constante de proporcionalidade que representa o

rateio de aprendizado. Esta regra é chamada de regra delta, pois o aprendizado é proporcional a diferença (delta) entre a ativação realmente encontrada e a ativação fornecida pelo professor. Esta regra é uma generalização da regra de convergência do perceptron, para o qual o *teorema* de convergência do perceptron foi provado.

Devido a sua importância, a próxima seção é dedicada ao aprendizado no contexto dos modelos conexionistas. Considerações acerca da dinâmica dos modelos conexionistas (convergência e estabilidade) são apresentadas na seção 11.3.

Ambiente

O último componente das RNAs é o ambiente onde a rede deve funcionar. É necessário especificar a natureza do ambiente, estabelecendo os possíveis padrões de entrada e de saída. Em alguns modelos, por exemplo o PDP [RUME86A], o ambiente é representado como uma função estocástica que varia ao longo do tempo sobre um espaço de padrões de entrada.

Geralmente, o ambiente é caracterizado como uma distribuição de probabilidade estável sobre um conjunto de padrões de entrada. Esta distribuição pode ser independente, ou não, de entradas ou de respostas passadas (histórico) do ambiente.

II.2 - APRENDIZADO EM MODELOS CONEXIONISTAS

Na sua essência, o conceito de aprendizado envolve mudança associada a aperfeiçoamento [PESS90]. Carbonell [CARB89] define o conceito de aprendizado, dentro do contexto da IA, como a habilidade de realizar tarefas novas que não podiam ser realizadas anteriormente, ou melhorar a realização de tarefas antigas, como resultado de mudanças produzidas pelo processo de aprendizado.

Uma classificação, apresentada por Rumelhart e Zipser [RUME85], relacionou a função do aprendizado em modelos conexionistas. Eles distinguiram quatro paradigmas que são apresentados na tabela 3.

Posteriormente, Rumelhart [RUME86A] apresentou uma outra classificação que é semelhante à apresentada na tabela 3. Neste caso existem apenas dois tipos de aprendizado: o associativo e o detetor de regularidades.

Tabela 3. Funcionalidade do aprendizado em RNAs

PARADIGMA	DESCRIÇÃO GERAL
Auto-Associador	Um conjunto de padrões é repetidamente apresentado e o sistema o armazena. Posteriormente, um padrão, ou parte de um, semelhante aos originais é apresentado. O sistema retoma o padrão original através de uma espécie de procedimento de <i>término de padrão</i> . Este é um processo de auto-associação onde um padrão é associado consigo mesmo, de tal forma que uma versão modificada do original pode servir de <i>deixa</i> para o procedimento de recuperação.
Associador de Padrões	Inicialmente, um conjunto de pares de padrões é apresentado ao sistema. Posteriormente, quando um membro do par é apresentado, o sistema faz a associação, produzindo o elemento correspondente do par.
Classificador	Neste caso há um número fixo de categorias nas quais são classificados os padrões de entrada. O objetivo é ensinar a rede a classificar corretamente os padrões de entrada, de tal forma que, quando apresentado um padrão, mesmo parcialmente modificado, a rede saiba classificá-lo corretamente.
Detetor de Regularidades	Existe uma população de padrões estímulo, sendo que cada membro possui uma probabilidade associada. O sistema desenvolve uma representação das características dos estímulos desta população, a qual captura as propriedades mais salientes dos padrões de entrada. Não existe, a priori, um conjunto de categorias no qual os padrões desta população possam ser classificados.

Nesta classificação, os três primeiros da tabela acima foram agrupados sob o rótulo de aprendizado associativo. Neste aprendizado, dois padrões a serem associados são apresentados, e a rede deve aprender a mapeá-los (no caso do auto-associador, o objetivo é mapeá-lo em si mesmo). No caso do detetor de regularidades, não existe, a priori, um conjunto de classes determinado para separar os padrões de entrada. Não é fornecido um padrão de saída, ele deve ser descoberto.

A classificação de Rumelhart [RUME86A] se assemelha àquela que classifica as redes quanto ao tipo de controle realizado durante o aprendizado (supervisionado ou não-supervisionado), apresentado na introdução desta tese. A classificação de Rumelhart e Zipser [RUME85] é mais específica do ponto de vista de como são processados (armazenados e recuperados) os padrões durante o treinamento das redes.

Uma descrição geral das diversas técnicas de aprendizado utilizadas nas RNAs, inspirado em Simpson [SIMP90], é apresentado na tabela 4.

Tabela 4. Principais técnicas de aprendizado em RNAs	
APRENDIZADO	DESCRIÇÃO GERAL
Correção de Erros	Aprendizado supervisionado que ajusta os pesos das conexões entre os nós na proporção da diferença entre os valores desejados e computados de cada nó da camada de saída.
Reforço	Aprendizado supervisionado onde os pesos são recompensados quando o sistema executa ações apropriadas, e punidos caso ele não as execute.
Estocástico	Aprendizado supervisionado que usa processos aleatórios, probabilidade e relações de energia para ajustar os pesos dos arcos.
Sistemas "Hardwired"	As conexões e os respectivos pesos são pré-determinados, semelhante a um autômato de estados finito.
Regra de Hebb	Aprendizado onde o ajuste dos pesos das conexões é realizado em função da relação de valores dos dois nós que ela conecta. Pode ser aplicado tanto ao aprendizado supervisionado quanto ao aprendizado não-supervisionado.
Competitivo e Cooperativo	Aprendizado não supervisionado onde os processos competitivo e cooperativo são descritos em termos de redes com conexões recorrentes auto-excitáveis. Estes arcos podem ser inibidores dos nós vizinhos (competitivo), e/ou excitadores dos vizinhos (cooperativo).
Sistemas Conectados Aleatoriamente (SCA)	Aprendizado não supervisionado utilizado para suportar a teoria de que a mente é uma rede conectada aleatoriamente quando vista do nível macroscópico.

No aprendizado por Correção de Erros existia um problema devido a sua incapacidade de estender o aprendizado para uma rede com mais de duas camadas. Especificamente, o valor do erro acumulado em cada unidade intermediária, que deveria ser creditado para as unidades de saída, não era definido. Com o advento de modelos como o "backpropagation" [RUME86A] este problema foi solucionado. Este algoritmo consiste numa generalização da regra delta (ver seção 11.1, regra de aprendizado - modificando a conectividade através da experiência). Derivando-se a regra delta obtém-se [RUME86B]: $\Delta w_{i,j} = \eta \cdot \delta_j \cdot o_i$, onde

$$\delta_j = (t_j - o_j) \cdot f'_j(\text{net}_j) \text{ se a célula } j \text{ for de saída}$$

$$f'_j(\text{net}_j) \cdot \sum_k \delta_k \cdot w_{j,k} \text{ se a célula } j \text{ for intermediária,}$$

onde f é a função de ativação f' é a sua derivada. As unidades de saída usam a diferença entre os valores obtido e desejado, como na regra delta. As unidades intermediárias precisam utilizar os valores de δ previamente calculados. Para estas células, este cálculo é realizado de maneira recorrente. Desta forma, na primeira fase (a fase "forward"), a entrada é ativada e espera-se a ativação da saída, calculando-se δ_j para os nós de saída. Na segunda fase (a fase "backward"), a partir deste cálculo de δ_j , é feita a propagação dos sinais de erro para as camadas anteriores, efetuando-se as mudanças nos pesos. Apesar da pouca plausibilidade biológica, os resultados obtidos pelo "backpropagation" aumentaram consideravelmente o interesse em RNAs.

O aprendizado por Reforço pode ser modelado pela equação:

$$\Delta w_{i,j} = \alpha \cdot [r - \Theta_j] \cdot e_{i,j},$$

onde r é um valor escalar de sucesso/falha informado pelo ambiente, Θ_j é o valor do limiar de reforço para o j -ésimo neurônio de saída, $e_{i,j}$ é a elegibilidade canônica do peso da i -ésima para a j -ésima unidade de processamento, e α é uma constante ($0 < \alpha < 1$) que regula o rateio de aprendizado. A elegibilidade canônica é função da distribuição de probabilidade selecionada que é usada para determinar se o valor de saída computado é igual ao valor desejado. Ela é definida como $e_{i,j} = \partial \ln g_i / \partial w_{i,j}$, onde g_i é a probabilidade da saída desejada se igualar à saída computada.

O aprendizado Competitivo/Cooperativo pode ser descrito como um sistema de equações diferenciais do tipo $dx_i/dt = F_i(x_1, x_2, \dots, x_n) = F_i(x)$, $i = 1, 2, \dots, n$. Ele é competitivo se $\partial F_i / \partial x_j \leq 0 \forall j \neq i$, e cooperativo se $\partial F_i / \partial x_j > 0 \forall j \neq i$. A idéia fundamental deste aprendizado é possibilitar que os neurônios da rede disputem, de alguma maneira, para obtenção do direito de responder a determinados estímulos. Este tipo de aprendizado permite que as unidades se especializem, e passem a atuar como detetores de características ou classificadores de padrões sem auxílio externo [RUME85].

O aprendizado Estocástico permite criar RNAs que escapem do problema de *mínimos locais* (que é descrito na próxima seção), utilizando um procedimento de diminuição de energia chamado de "simulated annealing" [KIRK83].

Os Sistemas Conectados Aleatoriamente (SCA) se fundamentam em um princípio de organização dos caminhos de sensoramento na mente. Este princípio postula que o posicionamento dos neurônios é ordenado e, geralmente, reflete alguma característica física do estímulo que está sendo percebido (por exemplo áreas visuais, áreas auditivas) [LIPP87]. Desta forma, o cérebro humano pode ser considerado como uma coleção estruturada de neurônios [KOH088].

II.3 - DINÂMICA: CONVERGÊNCIA E ESTABILIDADE

Dois importantes conceitos regem a dinâmica das redes neurais: a convergência e a estabilidade global. Biologicamente, *convergência* é a formação de similaridades sucessivas entre organismos ou associações antes distintas. Dentro do contexto de modelos conexionistas, a convergência está relacionada com a minimização do erro (eventual) entre as saídas computadas e desejadas dos neurônios. Em geral, o conceito de convergência está associado ao aprendizado supervisionado.

O conceito de *estabilidade* está fortemente relacionado com a manutenção, ou retorno, a um estado de equilíbrio, sólido e estável. No contexto das RNAs, a

estabilidade global é a estabilização de todas as ativações das unidades que compõem a rede, independente de qualquer entrada. Geralmente, o conceito de estabilidade está relacionado com a chamada *recursiva* existente nas redes cíclicas. Durante a execução de uma determinada tarefa, estas redes executam, eventualmente, vários ciclos de processamento até atingir um estado estável.

Quando se estuda a questão da convergência e da estabilidade em sistemas dinâmicos é comum se utilizar o conceito de *atrator*. Associando a cada RNA uma *superfície* de energia, de modo que, quando excitada com um padrão de atividade inicial, representado por um ponto sobre a referida superfície, a rede evoluirá dinamicamente, reduzindo a cada instante o valor da sua energia computacional, tendendo para o ponto de menor energia mais próximo que possa atingir. O atrator seria justamente este ponto de mínimo da superfície. O ponto de mínimo mais baixo de toda a superfície é chamado de mínimo global. Os demais pontos de mínimo existentes na rede são chamados de mínimos locais. Uma representação de um mínimo local é apresentado na figura 5.

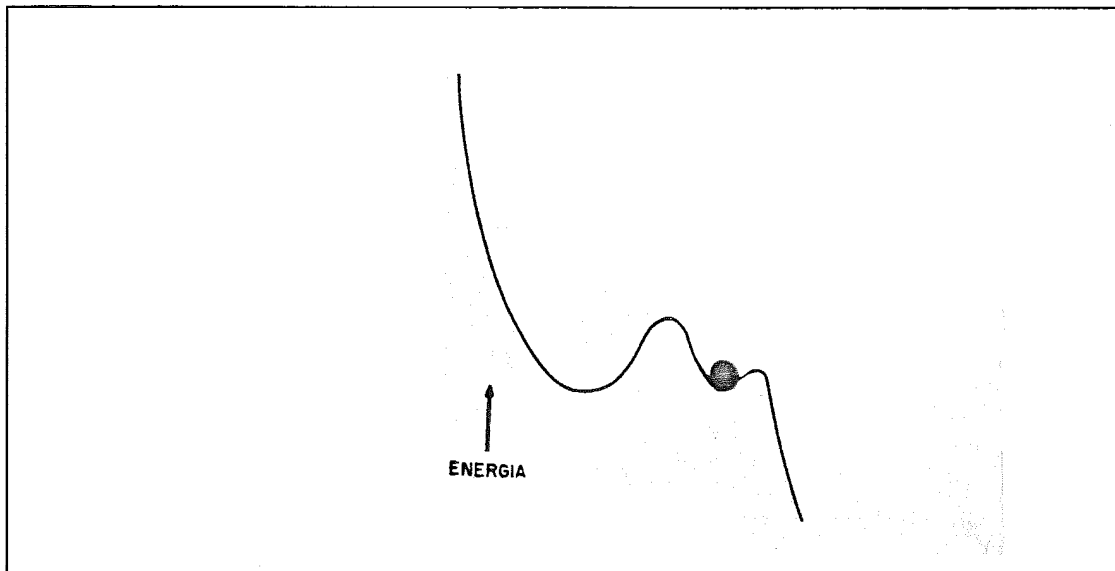


Figura 5. Representação de um mínimo local

Convergência

A convergência é definida como uma sequência de números x_n que tendem a um limite x , se dado um $\varepsilon > 0$ pode-se obter um inteiro n_0 tal que $|x_n - x| < \varepsilon$

para todo $n > n_0$. Existem dois métodos de convergência que se fundamentam neste princípio: *convergência com probabilidade 1*, e *convergência pela diferença dos quadrados*. A convergência com probabilidade 1 é definida como sendo

$$\lim_{n \rightarrow \infty} P\{x_n = x\} = 1,$$

onde $P\{x\}$ representa a probabilidade de x . A convergência pela diferença dos quadrados é definida como sendo

$$\lim_{n \rightarrow \infty} E\{|x_n - x|^2\} = 0,$$

onde $E\{x\}$ representa o valor esperado de x . A convergência é uma forma de analisar se um procedimento de aprendizado qualquer está capturando corretamente o mapeamento entre os dados apresentados. Tal afirmativa é verdadeira se o mapeamento convergir para um valor, ou para um conjunto, fixo. Um *grau de erro* é utilizado para medir o desempenho do procedimento de aprendizado durante a realização desta tarefa.

Estabilidade Global

As RNAs globalmente estáveis são definidas como sistemas dinâmicos que mapeiam rapidamente todas as entradas em atratores. Estes atratores (pontos de convergência, equilíbrio, limites, etc) são locais onde a informação pode ser deliberadamente armazenada. Isto não garante que os sistemas fazem o mapeamento da maneira desejada.

Existem três teoremas principais que são usados para provar a estabilidade em RNAs. Cada um destes teoremas utiliza o *método direto de Lyapunov* [SIMP90]. Este método estuda a estabilidade de encontrar certas funções das variáveis de sistemas dinâmicos cujas derivadas, em relação ao tempo, possuem propriedades específicas. Os três teoremas são os seguintes:

1. Teorema de Cohen-Grossberg - usado para mostrar a estabilidade em auto-associadores não-adaptativos. Este tipo de RNA opera em tempo contínuo, utiliza chamada recursiva, e modifica os pesos das conexões que são desprezíveis, em relação aos valores de ativação dos neurônios, podendo ser consideradas constantes [COHE83];
2. Teorema de Cohen-Grossberg-Kosko - utilizado para descrever a estabilidade em auto-associadores adaptativos. É uma extensão do primeiro teorema. Neste caso, os auto-associadores adaptativos são redes que não possuem a restrição de ser um sistema que não pode aprender quando estiver recebendo informação [KOSKSS];
3. Teorema ABAM - usado para descrever as ativações de redes neurais de duas camadas globalmente estáveis, que podem aprender e se adaptar ao mesmo tempo (heteroassociadores). Este teorema representa um caso específico do teorema de Cohen-Grossberg-Kosko [KOSK88].

II.4 - APLICAÇÕES E PRINCIPAIS MODELOS CONEXIONISTAS

As RNAs tem sido utilizadas em diversas aplicações. A relação que é apresentada a seguir, baseada em Simpson [SIMP90], representa as áreas em que as redes tem sido usadas, ou que possuem aplicação em potencial. A relação é a seguinte:

- Processamento de Sinais;
- Reconhecimento de Padrões;
- Compressão de Dados;
- Processamento de Imagens;
- Processamento de Fala;
- Controle;
- Sistemas Ecológicos;

- Projeto de Circuitos Integrados;
- Processamento de Conhecimento;
- Tarefas de Classificação (ex. diagnóstico);
- Estudo de Conexões Neurobiológicas;
- Memória Associativa;
- Recuperação de Informações em Banco de Dados;
- Processamento de Texto e Sentenças;
- Busca em Grafos;
- Coloração de Grafos;
- Problema do Caixeiro Viajante;
- Alocação de Recursos; e
- Escalonamento de Tarefas.

Existem muitos modelos conexionistas, e uma quantidade razoável de publicações que se dedicam a classificá-los ([LIPP87], [HINT89], [SIMP90]). Dentre estes modelos, foram selecionados cinco que são apresentados a seguir:

1. Perceptron
2. Perceptron de Multi-Camadas
3. Classificador de Carpenter/Grossberg
4. Rede de Kohonen
5. Rede de Hopfield

A tabela 5 apresenta os referidos modelos. A tabela 6 corresponde a uma descrição das principais características dos mesmos.

Tabela 5. Apresentação dos principais modelos conexionistas

MODELO	PESQUISADORES	INÍCIO DE PUBLICAÇÃO	CLASSE DE TAREFAS
Perceptron	F. Rosenblatt	1957	Reconhecimento de padrões (ex: caracteres impressos)
Perceptron de Multi-Camadas (Backpropagation)	P. Werbos, D. Parker e D. Rumelhart	1974	Reconhecimento de padrões (ex: controle adaptativo de braços de robôs) e processamento de fala.
Classificador de Carpenter/ Grossberg (Sistema ART)	G. Carpenter e S. Grossberg	1978	Reconhecimento de padrões (ex: reconhecimento de sinais de radar ou sonar) e processamento de imagens.
Rede de Kohonen	T. Kohonen	1980	Reconhecimento de padrões (ex: reconhecimento de fala) e aprendizado da distribuição de probabilidades dos dados (ex: auto-organização de mapas de características)
Rede de Hopfield	J. Hopfield	1982	Reconhecimento de padrões (ex: reconhecimento de dados ou imagens completas a partir de fragmentos) e memória associativa.

Observando a tabela 5.(B), pode-se notar que existe uma preferência pelo uso da função sigmóide como função de ativação. Esta função permite modelar a saída de cada neurônio, em relação a sua ativação de entrada, de uma forma gradual e não-linear.

O primeiro modelo conexionista desenvolvido foi o Perceptron de duas camadas [ROSE59] que pode ser usado com valores contínuos. Esta rede gerou muito interesse pela habilidade de aprender a reconhecer padrões linearmente separáveis. Um problema com este procedimento de convergência do perceptron é que os limites de decisão podem oscilar continuamente quando as

entradas forem não separáveis, e quando ocorrer uma sobreposição de distribuições [SIMP90].

Tabela 6. Características dos principais modelos conexionistas

MODELO	PROPRIEDADES DAS CÉLULAS	PROPRIEDADES DAS REDES	APRENDIZADO
Perceptron	As células possuem entradas binárias e saídas que assumem os valores $+1$ ou -1 . A função de ativação é a função salto.	Rede acíclica de duas camadas.	Utiliza a técnica de Reforço
Backpropagation	As células são do tipo perceptron, e possuem valores contínuos. A função de ativação é a função sigmóide.	Rede acíclica de três camadas no mínimo.	Utiliza a técnica de Correção de Erros com o uso da regra delta generalizada.
Sistema ART	As células possuem entradas binárias, podendo assumir valores contínuos. A função de ativação é a função sigmóide.	Rede cíclica de três camadas.	Utiliza a técnica de aprendizado Competitivo e Cooperativo que foi introduzida pelo próprio Grossberg.
Rede de Kohonen	As células possuem entradas contínuas. A função de ativação é a função sigmóide.	Rede cíclica de duas camadas.	Utiliza a técnica de SCA. O próprio Kohonen foi um dos seus introdutores.
Rede de I-opfield	As células possuem entradas binárias e saídas que assumem os valores $+1$ ou -1 . A função de ativação é a função sigmóide.	Rede cíclica de uma camada.	Os padrões são armazenados no início, se assemelhando a técnica utilizada pelos sistemas "hardwired".

Ele não é apropriado para classes que não podem ser separadas por um hiperplano. Por exemplo, as distribuições para as duas classes do problema do OU-exclusivo (XOR) não são linearmente separáveis, logo não podem ser separadas por uma simples linha. O OU-exclusivo é o problema difícil clássico

para as RNAs. Ele representa o problema não linearmente separável mais simples que existe. Este problema foi utilizado por Minsky e Papert [MINS69] para ilustrar a fraqueza do perceptron, que contribuiu para o ostracismo desta área por mais de uma década.

No começo da década de setenta foram desenvolvidos os Perceptrons de Multi-Camadas. Estes são redes acíclicas com uma ou mais camadas de neurônios intermediários entre as camadas de entrada e de saída. Um algoritmo capaz de treinar os perceptrons de multi-camadas é o "backpropagation" (ver seção 11.2) [RUME86C].

O algoritmo "backpropagation" foi testado em uma série de problemas clássicos, tais como o OU-exclusivo, e em problemas relacionados com reconhecimento de padrões visuais [RUME86C]. Geralmente, o "backpropagation" usa, para aprender a classificar o OU-exclusivo, cerca de 150 iterações sobre a tabela verdade do mesmo. Na maioria dos casos, ele encontrou boas soluções para os problemas propostos, apesar do algoritmo, às vezes, fornecer uma configuração de pesos correspondente a um mínimo local da função de erro. Isto é devido ao fato do backpropagation utilizar o método do gradiente e, a princípio, a superfície do erro possui uma forma qualquer.

Uma aplicação bastante conhecida do "backpropagation" é o sistema NETtalk [SEJN87]. Com o treinamento, este sistema é capaz de aprender a pronunciar palavras em inglês de uma maneira correta. Ele também possui a capacidade de generalizar para palavras novas, que ainda não haviam sido apresentadas ao mesmo.

No final da década de setenta, Carpenter e Grossberg [GROS86] projetaram uma rede capaz de formar aglomerados de informações ("clusters"), e de ser treinada sem supervisão, chamada de sistema ART. Nesta rede, as sinapses são bi-direcionais. As ligações "bottom-up" tentam categorizar os estímulos de entrada, e os arcos "top-down" tentam aprender estas categorias. Os nós da camada de saída funcionam como unidades "winner-take-all", ou seja, apenas a unidade que recebe a maior excitação convergirá para a ativação 1, as demais evoluirão para a ativação 0.

Em linhas gerais, o sistema ART funciona da seguinte maneira: O algoritmo principal seleciona a primeira entrada como um exemplo para o primeiro aglomerado. A entrada seguinte é comparada com este primeiro exemplo. Ela é agrupada com o mesmo se a distância para o primeiro for menor que um certo limiar, chamado de limiar de vigilância. Caso contrário, este exemplo formará um novo aglomerado. Este processo se repete para todas as entradas existentes. O número de aglomerados cresce em função do limiar e da métrica da distância usada para comparar os exemplos de entrada dos aglomerados.

No início da década de oitenta, Kohonen [KOH84] propôs uma rede onde se corroborou os estudos teóricos sobre a organização dos caminhos de sensorialidade na mente. Segundo esta teoria, conforme apresentado na seção 11.2, o cérebro humano foi considerado como uma coleção estruturada de neurônios. Esta *ordem espacial das unidades de processamento* [KOH82] permitiu elaborar uma rede neural dotada de mecanismos que permitem formar representações estruturadas dos estímulos de entrada. Após o aprendizado, as unidades respondem a diferentes estímulos de maneira ordenada, formando um sistema de coordenadas de características sobre a rede.

O algoritmo de auto-organização, proposto por Kohonen [KOH88], é um quantificador de vetores sequenciais, e funciona como um algoritmo de aglomeração. Se os estímulos formam aglomerados no espaço de entrada, o algoritmo faz com que os padrões de um aglomerado sejam mapeados para uma região localizada comum no mapa. Ele faz um arranjo destas regiões ao longo do mapa, de maneira a capturar o máximo possível da topologia geral dos aglomerados de entrada. Nesta rede, as células de entrada estão conectadas a todas as células de saída. As unidades de saída estão interconectadas em um arranjo bi-dimensional.

Kohonen demonstrou como utilizar o algoritmo de auto-organização para problemas de reconhecimento de fala [KOH88]. Neste experimento, a rede foi apresentada a conteúdos espectrais, em vez de fonemas, e observou-se a formação de um mapa fonotópico onde as distâncias entre os pontos no mapa

(unidades) são proporcionais às diferenças vetoriais entre os conteúdos espectrais originais.

Também na década de oitenta, o trabalho de Hopfield [HOPF82] contribuiu substancialmente para o ressurgimento das pesquisas em RNAs. A rede de Hopfield, bem como o classificador de Carpenter/Grossberg são mais apropriadas quando representações binárias permitem modelar a situação desejada. Por exemplo, imagens em preto e branco, onde os elementos de entrada podem ser representados pelos valores de cada ponto da imagem (0 = branco, 1 = preto) [LIPP87].

Esta rede pode ser usada como uma memória associativa ou para resolver problemas de otimização. Uma versão da rede original [HOPF82] pode ser usada como uma memória de acesso por conteúdo. Uma extensão da rede de Hopfield, com aplicação ao problema do caixeiro viajante e à representação do conhecimento, pode ser encontrada em Carvalho [CARV89].

Esta rede possui duas limitações quando usada como memória de acesso por conteúdo. Primeiro, apesar dos padrões armazenados, a rede pode convergir para um novo padrão diferente dos padrões exemplo existentes. Isto pode produzir a situação de não bntimento da rede, ou seja, a rede não casa com um padrão já existente.

Uma segunda limitação é que o padrão exemplo será considerado instável se ele compartilhar muitos bits com outro padrão exemplo, o que pode ocasionar uma convergência da rede para este outro exemplo. Este problema pode ser eliminado, consequentemente melhorando seu desempenho, utilizando-se procedimentos de ortogonalização [LIPP87].

CAPÍTULO III

SISTEMAS ESPECIALISTAS CONEXIONISTAS

Uma das principais características dos modelos conexionistas é a sua capacidade de aprendizado. Esta característica tem atraído alguns pesquisadores da área de SEs. Tais pesquisadores vêem nesta capacidade dos modelos conexionistas uma forma de superar o *gargalo* de aquisição de conhecimento apresentado por Feigenbaum [FEIG82].

Em relação ao desempenho apresentado pelos SEs tradicionais, existe um aspecto importante a ser analisado. Estes sistemas conseguem demonstrar uma capacidade de raciocínio semelhante ao apresentado por especialistas humanos dentro de domínios de conhecimento restritos. Esta restrição quanto ao tamanho do domínio é o principal motivo de sucesso dos SEs e também fonte de limitações. Neste sentido, estes sistemas são considerados *frágeis* por responderem apropriadamente *apenas* em domínios restritos. A intervenção humana torna-se necessária quando se deseja efetuar mudanças, mesmo que pouco significativas, no conhecimento representado no domínio [HOLLSGA].

Como resultado destas motivações, alguns trabalhos integrando SEs e modelos conexionistas começam a surgir ([GALL88], [SAMA88], [SAIT88], [ROMA89], [MACH89]). Gallant [GALL88] criou o termo *Sistema Especialista Conexionista* para classificar tal categoria de SEs. Neste capítulo são apresentados alguns trabalhos encontrados na literatura. Uma atenção especial é dedicada ao Modelo Neural Combinatório [MACH89], pois o mesmo serviu de base para esta pesquisa.

III.1 - APLICAÇÕES NA ÁREA DE SISTEMAS ESPECIALISTAS

Alguns trabalhos que utilizam o paradigma conexionista aplicado a SEs começam a surgir. Relacionamos cinco destes a saber:

1. O SE Conexionista (SEC) proposto por Gallant [GALL88];
2. O RUBICON proposto por Samad [SAMA88];
3. O FUZZNET proposto por Romaniuk [ROMA89];
4. O SE para diagnóstico baseado no PDP (SDPDP) proposto por Saito [SAIT88]; e
5. O Modelo Neural Combinatório (MNC) proposto por Machado [MACH89].

A tabela 7 apresenta os referidos SEs conexionistas. A tabela 8 descreve as principais características das redes dos mesmos.

Tabela 7. Apresentação dos SEs Conexionistas			
Modelo	Pesquisadores	Ano de Publicação	Exemplos de Aplicações
Sistema Especialista Conexionista (SEC)	S. Gallant	1988	Diagnóstico de Doença do Sarcófago e Seleção de Tratamento
RUBICON	T. Samad	1988	Identificação de Títulos de Filmes
FUZZNET	S. Romaniuk e L. Hall	1989	Diagnóstico de Febre e Identificação de Pedras Preciosas
Sistema de Diagnóstico baseado no PDP (SDPDP)	K. Saito e R. Nakano	1988	Diagnóstico de Dor de Cabeça por Contração Muscular
Modelo Neural Combinatório (MNC)	R. Machado e A. Rocha	1989	Diagnóstico de Doenças Renais

Tabela 8. Características das redes dos SEs Conexionistas.

MODELO	PROPRIEDADES DAS CÉLULAS	PROPRIEDADES DAS REDES	APRENDIZADO
SEC	O grau de ativação das células pode assumir os valores -1, 0 ou +1 para falso, desconhecido e verdadeiro respectivamente. A função de ativação é a função salto.	Rede acíclica de três camadas. A camada de entrada representa os sintomas, a intermediária as doenças e a de saída os tratamentos.	Rede treinada com supervisão, de acordo com a técnica de Reforço.
RUBICON	As células possuem entradas binárias e saídas que assumem os valores 0 ou 1. A função de ativação é a função salto.	Rede acíclica de seis camadas conectadas na seguinte ordem: entrada, antecedente, coletora, regra, conseqüente e saída.	Rede treinada com supervisão com uma técnica semelhante a utilizada pelos sistemas "hardwired".
FUZZNET	As células podem assumir um valor de ativação no intervalo [0, 1], que corresponde a um grau de possibilidade nebuloso. A função de ativação é a função sigmóide.	Rede acíclica de quatro camadas no mínimo (entrada, intermediária(s), extra e saída). A camada extra é formada por duas células (positiva e negativa) estando conectadas a todas as células intermediárias e de saída.	Rede treinada com supervisão com uma técnica semelhante à utilizada pelos sistemas "hardwired".
SDPDP	As células são do tipo perceptron, podendo assumir um valor de ativação no intervalo [0, 1]. A função de ativação é a função sigmóide.	Rede acíclica de três camadas. A camada de entrada representa os sintomas, a intermediária combinações de sintomas, e a de saída as doenças.	Rede treinada com supervisão, de acordo com a técnica de Correção de Erros.
MNC	O grau de ativação das células pode assumir um valor no intervalo [0, 1], e representa um grau de possibilidade nebuloso. As células podem ser do tipo E-nebuloso ou OU-nebuloso.	Rede acíclica de 3 ou mais camadas. As camadas intermediárias são formadas por células E-nebuloso e representam combinações de evidências. A camada de saída é formada por células OU-nebuloso e representam as hipóteses.	Rede treinada com supervisão. Punições e recompensas para as sinapses são calculadas para cada exemplo de acordo com a regra de Hebb. Os pesos sinápticos são calculados normalizando-se os valores acumulados de recompensas e punições recebidas pelas sinapses.

Observando a tabela 6.(A) pode-se notar que os exemplos de aplicações destes sistemas estão inseridos em uma classe de tarefas chamada de *classificação*. A classificação é uma tarefa de processamento de informação, onde entidades específicas são mapeadas em categorias mais gerais [CHAN88]. Esta tarefa é subjacente a muitos domínios de problema, tais como diagnose, predição, monitoração, interpretação, identificação, depuração e seleção, dentre outros.

O SEC [GALL88] é formado por uma base de conhecimento (composto por uma matriz de pesos, nomes de variáveis e questões), um motor de inferência, chamado de MACIE ("Matrix Controlled Inference Engine"), e por último uma interface com o usuário. O algoritmo de aprendizado utilizado no sistema é o *algoritmo de Bolso*.

O algoritmo de Bolso é uma modificação do algoritmo do Perceptron. Ele computa os vetores de peso, chamados de perceptron P , os quais substituem ocasionalmente os vetores de peso w já existentes no *bolso*. Esta substituição ocorre quando este algoritmo, ao classificar um exemplo de treinamento E qualquer, consegue uma taxa de acerto superior à obtida pelos vetores de peso w já existentes.

Uma falha no algoritmo de Bolso é que não existe um limite (número) de execuções necessárias para garantir que os pesos *embolsados* são os melhores possíveis. Entretanto, se não existir muitos exemplos de treinamento ($< 10^5$), é possível checar periodicamente os pesos embolsados contra o conjunto de todos os exemplos de treinamento com o objetivo de avaliar seu desempenho.

O motor de inferência utiliza o paradigma conexionista para as seguintes tarefas:

- Inferência baseada em informação parcial, usando o encadeamento progressivo ("Forward Chaining");
- Identificação das variáveis de entrada desconhecidas que são fundamentais para a realização de inferências adicionais; e
- Produção de justificativas para as inferências realizadas utilizando o encadeamento regressivo ("Backward Chaining").

O RUBICON é uma arquitetura conexionista usada para implementação de SEs baseados em regras [SAMA88]. Ela utiliza tanto representações distribuídas como localizadas. Algumas características importantes deste sistema são:

- Número arbitrário de antecedentes e conseqüentes em qualquer regra;