

# CAPÍTULO IV

## OUTRAS APLICAÇÕES UTILIZANDO REDES NEURAIS

### IV.1 - Determinação do Caminho Mínimo pelo Modelo de Hopfield

#### IV.1.1: Introdução:

Redes neurais recorrentes e estáveis (na qual se inclui o modelo de Hopfield), representam uma alternativa interessante para problemas clássicos de otimização combinatória em função de seu paralelismo inerente e da possibilidade de implementação por meios analógicos. Além disso, tem a capacidade de convergir para mínimos locais (associados ao problema) tão satisfatórios quanto se queira.

Mostraremos uma aplicação do modelo de Hopfield ao problema do caminho mínimo entre dois vértices de um grafo (ref. [16]). A maior dificuldade neste caso, reside na grande quantidade de parâmetros utilizados e na atribuição de valores a estes parâmetros. Apesar disso, como demonstrado em [16], a convergência para soluções ótimas é garantida com uma aceitável margem de erro (isto ocorre em função do elevado número de mínimos locais associados à função de energia).

#### IV.1.2: A Rede Neural:

Seja um grafo  $G=(V,E)$  conexo com  $n$  vértices. Analogamente ao problema do caixeiro viajante, definiremos uma rede com  $n^2$  neurônios, onde cada linha corresponde a um grupo e cada coluna a um neurônio desse grupo. Sempre que um neurônio  $a_{ij}$  estiver ativado, estaremos fazendo uma conexão entre os nós  $i$  e  $j$ .

No exemplo abaixo, representamos (em negrito) um caminho qualquer entre os nós 4 e 8:

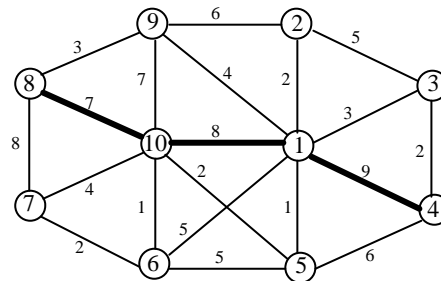


Figura IV.1: Grafo valorado com 10 nós

Teremos então a seguinte configuração em nossa rede de neurônios:

	1	2	3	4	5	6	7	8	9	10
1	0	0	0	0	0	0	0	0	0	<b>1</b>
2	0	0	0							
3	0	0	0							
4	<b>1</b>			0						
5	0				0					
6	0					0				
7	0						0			
8	0							0	0	0
9	0							0	0	0
10	0	0	0	0	0	0	0	<b>1</b>	0	0

Figura IV.2: Matriz de Conexões representando um caminho entre os nós 4 e 8.

Para que possamos trabalhar com o modelo de Hopfield devemos definir uma função de energia associada ao problema de caminho mínimo. Nesta função algumas características devem ser observadas:

- (a) Cada nó deve ser conectado apenas a outro nó. Ou seja, cada linha na matriz de conexões terá somente um neurônio ativado.
- (b) Nenhum nó poderá ser percorrido duas vezes. Assim, teremos apenas um nó ativado em cada coluna.

Estas duas condições são idênticas às utilizadas no problema do caixeiro viajante. Obtemos portanto as seguintes parcelas:

$$\frac{A}{2} \sum_x \sum_i \sum_{j \neq i} a_{xi} a_{xj} + \frac{B}{2} \sum_i \sum_x \sum_{y \neq x} a_{xi} a_{yi}$$

- (c) Como o nó fonte (por exemplo  $s$ ), deve ser conectado apenas a um outro nó, teremos apenas um neurônio ativado na linha  $s$ . Assim, definimos a seguinte parcela:

$$\frac{C}{2} \left( \sum_i a_{si} - 1 \right)^2$$

- (d) Se  $d$  é o nó destino, nenhum nó estará ativado na linha  $d$ . Logo, podemos definir a parcela:

$$\frac{D}{2} \left( \sum_i a_{di} \right)^2$$

- (e) Como um nó qualquer  $i$  não pode ser conectado a ele próprio, os neurônios  $a_{ii}$  presentes na diagonal da matriz de permutações, deverão ser todos nulos. Assim:

$$\frac{E}{2} \sum_i \sum_{j \neq i} a_{ii} a_{jj}$$

- (f) Nenhum nó poderá estar ligado ao nó fonte  $s$ . Segue então que todos os neurônios na coluna  $s$  devem estar desativado. Teremos portanto a seguinte parcela:

$$\frac{G}{2} \left( \sum_x a_{is} \right)^2$$

- (g) Finalmente, devemos elaborar uma parcela responsável pela minimização da distância entre o nó fonte e o nó destino. Obviamente neste caso, nem todos os nós serão ativados no "menor" caminho entre a fonte e o destino. Devemos nos preocupar em descartar todos os caminhos fechados (ou seja, todas aquelas conexões onde o nó fonte é conectado a 2 nós distintos), descartar a repetição de vértices em um mesmo caminho e finalmente descartar a ocorrência de "loops" ( $a_{ii} = 1$ ). Temos portanto, a seguinte expressão:

$$\frac{F}{2} \sum_x \sum_{i \neq s} \sum_{j \neq i} d_{xi} \cdot a_{xi} \cdot a_{ij} + \frac{F}{2} \sum_x d_{xd} \cdot a_{xd}$$

$i \neq x \quad j \neq x$

onde cada termo  $d_{xi} \cdot a_{xi} \cdot a_{ij}$  (na primeira parcela) representa uma componente de cada possível caminho entre os nós  $s$  e  $d$ . Este termo não pode levar em consideração o último arco no caminho entre a fonte e o destino já que  $a_{dj} = 0$  para todo  $j$ . A última parcela será responsável pela última aresta presente no menor caminho. Note também que ela corresponde à parcela linear presente na função de energia.

Se reagrupamos todas as parcelas da função de energia acima teremos uma função quadrática do tipo  $x^T W x + d^T x$ , onde  $W$  representa a matriz das sinápses existentes entre os neurônios da rede. Como discutido em [16] e [18], a matriz  $W$  é assimétrica e possui elementos não nulos na diagonal principal, segue portanto que as condições de estabilidade da rede, como analisadas por Cohen e Grossberg [18] não serão satisfeitas.

Como não podemos alterar as sinápses de maneira que soluções estáveis correspondam a padrões solução, adicionamos à função de energia uma nova parcela como abaixo:

$$\frac{F}{2} \sum_y \sum_{j \neq s} \sum_{i \neq j} d_{yj} \cdot a_{ji} \cdot a_{yj}$$

$j \neq y \quad i \neq y$

Desta forma, reagrupando novamente as parcelas da função de energia obtemos uma matriz  $W$  simétrica sem que as condições relativas às distâncias e restrições do problema sejam alteradas. Apesar das condições de estabilidade ainda não serem satisfeitas, pois ainda temos elementos não nulos na diagonal, a parcela acima proporciona uma contribuição substancial na determinação da solução ótima (vide [16]).

As sinápses presentes na matriz de conexões  $W$ , considerando-se a introdução da parcela responsável pela simetrização será:

$$W_{(xi,yj)} = \begin{cases} -A, & \text{se } x = y \text{ e } i \neq j \\ -B, & \text{se } x \neq y \text{ e } i = j \\ -C, & \text{se } x = y \text{ e } x = s \text{ e } i = j \\ -2C & \text{se } x = y \text{ e } x = s \text{ e } i \neq j \\ -D & \text{se } x = y \text{ e } x = d \text{ e } i = j \\ -2D & \text{se } x = y \text{ e } x = d \text{ e } i \neq j \\ -E & \text{se } x \neq y \text{ e } x = i \text{ e } y = j \\ -F \cdot d_{xy} & \text{se } x \neq i \text{ e } i \neq s \text{ e } j \neq i \text{ e } y = i \text{ e } j \neq x \\ -F \cdot d_{xy} & \text{se } x = j \text{ e } j \neq s \text{ e } j \neq i \text{ e } y \neq j \text{ e } i \neq y \\ -G & \text{se } i = j \text{ e } i = s \text{ e } x = y \\ -2G & \text{se } i = j \text{ e } i = s \text{ e } x \neq y \end{cases}$$

Como discutido em [16], o tempo de processamento do modelo de Hopfield em circuitos VLSI são praticamente independentes de sua complexidade, ou seja, do número total de nós presentes no grafo.

## VI.2 - O Problema das Quatro-Cores:

### IV.2.1 - Introdução:

Suponha um mapa M qualquer com n regiões distintas. Desejamos colorir este mapa com o menor número possível de cores de maneira que regiões adjacentes não tenham a mesma cor. Em outubro de 1977, Appel e Kaken [19], apresentaram uma demonstração (com o auxílio de computador) para este problema afirmando que são necessárias apenas 4 cores para se colorir qualquer grafo planar. Existe ainda uma certa controvérsia entre os pesquisadores quanto a este tipo de demonstração pois baseia-se principalmente na enumeração de casos particulares (grafos planares com uma certa estrutura), e na análise exaustiva desses casos.

Apresentamos um algoritmo paralelo baseado no modelo lógico (neurônio de McCulloch-Pitts) e no modelo de Hopfield.

#### IV.2.2 - A Rede Neural:

Utilizaremos uma rede bidimensional 4xn para a solução do problema das 4-cores (onde n representa o número de regiões).

Quatro cores podem ser facilmente representadas por 4 neurônios de nossa rede. Por exemplo: vermelho, amarelo, verde e azul podem ser representados por 1000, 0100, 0010 e 0001 respectivamente. Dessa forma, a cada região estarão associados 4 neurônios representando uma determinada cor. Para a coloração de todo o mapa necessitaremos portanto 4xn neurônios. Vejamos o seguinte exemplo:

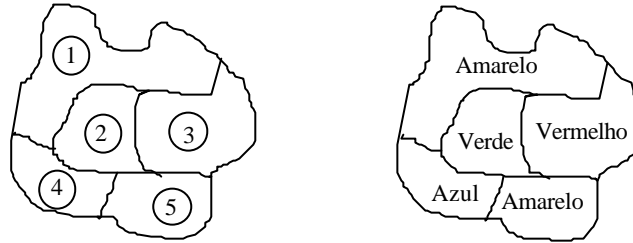


Figura IV.3: Um mapa com 5 regiões e uma possível coloração utilizando 4 cores.

As relações de vizinhança (matriz de adjacência) e a rede neural são representadas abaixo:

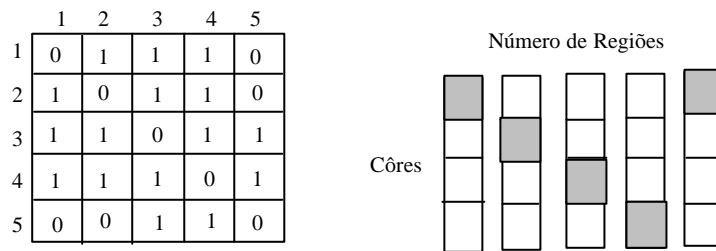


Figura IV.4: Matriz de Adjacência e a Rede Neural

Na construção da função de energia devemos nos preocupar basicamente com dois aspectos: cidades vizinhas não possuem a mesma cor, e apenas um neurônio deverá ser ativado em cada coluna da rede. Assim:

$$E = \frac{A}{2} \sum_{x=1}^n \left( \sum_{i=1}^4 a_{ix} - 1 \right)^2 + B \sum_{i=1}^n \sum_{y=1, y \neq i}^n \sum_{x=1}^4 d_{xy} a_{ix} a_{iy}$$

onde  $d_{xy}$  é 1 se as regiões x e y são adjacentes entre si, e 0, caso contrário. Note que a primeira parcela é responsável pela ativação de apenas um neurônio em cada coluna. A segunda

parcela penaliza regiões vizinhas de mesma cor. Note que se  $a_{ix} = a_{iy} = 1$  (para  $x \neq y$ ) e  $d_{xy} = 1$  (as regiões x e y são vizinhas) teremos um incremento na função de energia decorrente da violação das restrições do problema.

Derivando a função de energia em relação a  $a_{ix}$  temos:

$$\frac{\partial E}{\partial a_{ix}} = \frac{dU_{ix}}{dt} = -A \left( \sum_{j=1}^4 a_{jx} - 1 \right) - B \sum_{\substack{y=1 \\ y \neq x}}^n d_{xy} a_{iy}$$

Como discutido em [20], podemos introduzir uma nova parcela na derivada da função de energia que permite sair de mínimos locais caminhando em direção ao mínimo global da função de energia:

$$\frac{\partial E}{\partial a_{ix}} = \frac{dU_{ix}}{dt} = -A \left( \sum_{j=1}^4 a_{jx} - 1 \right) - B \sum_{\substack{y=1 \\ y \neq x}}^n d_{xy} a_{iy} \sum_{k=1}^n d_{yk} + Ch \left( \sum_{j=1}^4 a_{jx} \right) \left( C_1 \sum_{k=1}^n d_{xk} + C_2 \frac{\sum_{k=1}^n \sum_{y=1}^n d_{xy} d_{yk}}{\sum_{k=1}^n d_{xk}} \right)$$

onde  $C, C_1, C_2$  são constantes e  $h(x)=1$  se  $x=0$  e 0, caso contrário. Esta parcela provoca uma força excitatória somente quando todos os  $a_{jx}$  são nulos.

Vejamos então o seguinte algoritmo:

**Algoritmo:** Coloração;

**Início**

-  $t:=0; \Delta t:=1; A = B = C = C_1 = C_2 = 1;$

- Selecionar  $U_{ix}(t)$  aleatoriamente onde  $x=1, \dots, n$  e  $i=1, \dots, 4$ .

- **Repita**

- Avaliar:

$$a_{ix}(t) = f(U_{ix}(t)) = \begin{cases} 1, & \text{se } U_{ix}(t) > 0 \\ 0, & \text{caso contrário} \end{cases}$$

- Calcular:

$$\frac{dU_{ix}}{dt} = -A \left( \sum_{j=1}^4 a_{jx} - 1 \right) - B \sum_{\substack{y=1 \\ y \neq x}}^n d_{xy} a_{iy} \sum_{k=1}^n d_{yk} + Ch \left( \sum_{j=1}^4 a_{jx} \right) \left( C_1 \sum_{k=1}^n d_{xk} + C_2 \frac{\sum_{k=1}^n \sum_{y=1}^n d_{xy} d_{yk}}{\sum_{k=1}^n d_{xk}} \right)$$

- Calcular:

$$U_{ix}(t+1) := U_{ix}(t) + \Delta U_{ix}(t) \Delta t;$$

-  $t:=t+1;$

- **Até** "Critério de Parada";

**fim.**

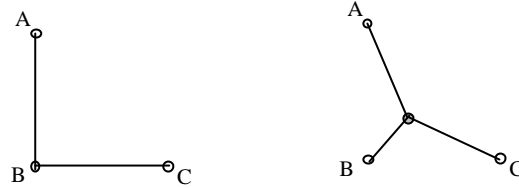
Como discutido em [20], o algoritmo termina se pelo menos um entre os valores  $\Delta_{1x}; \Delta_{2x}; \Delta_{3x}$  e  $\Delta_{4x}$  for nulo para  $x=1, 2, \dots, n$ .

### IV.3 - O Problema da Árvore Mínima de Steiner

#### IV.3.1 - Introdução:

O problema de encontrar a menor árvore conectando um conjunto de pontos no plano (ou  $\mathbb{R}^n$ ) é conhecido como problema da árvore mínima de Steiner (AMS) e tem aproximadamente 300 anos. Apresentaremos uma rede neural que se auto-organiza para encontrar a árvore mínima.

No problema da árvore geradora mínima (AGM), buscamos conectar os pontos do plano, de maneira que o somatório das arestas (distâncias) presentes na árvore seja mínimo. Este problema foi resolvido em tempo polinomial por Kruskal 1956, e Prim 1957 (ref.[24] e [25]). Entretanto, é possível construir árvores com um menor comprimento incorporando novos pontos ao problema (árvore mínima de Steiner). Abaixo visualizamos um exemplo onde a adição de um novo ponto diminui o comprimento total da árvore:



**Figura IV.5: Árvore Geradora Mínima e Árvore de Steiner**

O problema da árvore mínima de Steiner euclidiano é encontrar a menor árvore conectando  $n$  pontos, utilizando, se conveniente, pontos extras  $\{s_1, s_2, \dots, s_k\}$  que chamaremos *pontos de steiner* (teremos sempre  $k \leq n-2$ ). Chamaremos os  $n$  pontos (dados do problema) de *pontos fixos*  $\{p_1, p_2, \dots, p_n\}$ .

Este problema, como demonstrado por Garey e Johnson [21] em 1977 é NP-Completo, logo, o esforço computacional exigido para resolvê-lo cresce exponencialmente no tamanho do problema.

Duas dificuldades principais estão presentes na obtenção da árvore mínima de Steiner:

- (a) não sabemos o número de pontos de steiner a priori;
- (b) para um número dado de pontos de steiner, devemos determinar a topologia correta com respeito ao conjunto de pontos fixos.

Obviamente poderíamos resolver o problema enumerando árvores mínimas para cada topologia, variando  $k$  de 1 até  $n-2$ . Entretanto, para  $n=6$  o número de árvores ultrapassa a 5600 e para  $n=8$  teremos mais de 3 milhões de possibilidades (Hwang and Richards, 1992 / [22])!

Como veremos a seguir, a rede neural se auto-organiza (self-organization) para encontrar a árvore mínima de steiner, ou seja, o número de pontos de steiner e a topologia são determinadas através de um processo de auto-modificação.

### **IV.3.2 - A Rede Neural para o Problema de Steiner:**

Seja  $p_i = (p_{i1}, p_{i2}, \dots, p_{iD})$  as coordenadas de um ponto fixo em  $\mathbb{R}^D$  e  $x_i = (x_{i1}, \dots, x_{iD})$  as coordenadas de pontos de steiner (para  $i=1, 2, \dots, m=n-2$ ). Com a finalidade de facilitar a visualização, consideramos  $D=2$ . Vejamos então a seguinte figura:

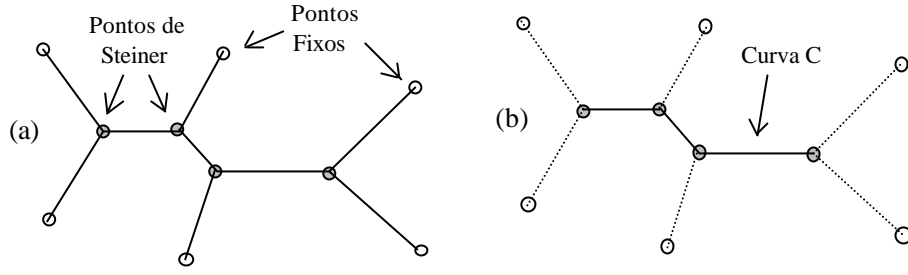


Figura IV.6: (a) Árvore de Steiner e (b) Curva C formada por pontos de Steiner

Como discutido acima, nossa maior dificuldade será na determinação do número de pontos de steiner e na localização desses pontos (que formarão uma curva C), ou seja, buscamos construir uma função de energia que determine a localização exata dos pontos de steiner (variáveis  $x_i$ ). Trabalharemos com  $m=n-2$  pontos de steiner em nossa curva C. Como veremos, isto não impede soluções com um número menor de pontos de steiner já que alguns pontos presentes na curva C poderão ser coincidentes. A medida que chegamos ao mínimo da função de energia a topologia vai sendo definida através das coordenadas de  $x_i$ .

Consideremos então a seguinte função de energia:

$$E = \sum_{i=1}^{m-1} \text{dist}(x_{i+1}, x_i) + \sum_{i=1}^n \sum_{j=1}^m d_{ij} h(d_{ij})$$

onde:

$$\begin{aligned} \text{dist}(x_{i+1}, x_i) &= \left( (x_{i+1,1} - x_{i,1})^2 \right) + \left( (x_{i+1,2} - x_{i,2})^2 \right)^{1/2}; \\ d_{ij} &= \text{dist}(p_i, x_j); \\ h(d_{ij}) &= \begin{cases} 1, & \text{se } d_{ij} = \min_k (d_{ik}) \\ 0, & \text{caso contrário} \end{cases} \end{aligned}$$

Calculando a derivada da função de energia (na direção contrária ao gradiente) temos:

$$\frac{-\nabla E}{\nabla x_{ij}} = \frac{(x_{i+1,j} - x_{i,j})}{\text{dist}(x_{i+1}, x_i)} + \frac{(x_{i-1,j} - x_{i,j})}{\text{dist}(x_i, x_{i-1})} + \sum_{k=1}^n h(d_{ki}) \frac{(p_{k,j} - x_{i,j})}{d_{ki}}$$

onde  $i=1, 2, \dots, m$  e  $j=1, 2$ . A derivada de  $h(d_{ij})$  é calculada através do gradiente generalizado.

Note que a função de energia representa a soma do comprimento da curva C com o comprimento das ligações entre os pontos fixos e a curva C.

Para ilustrar um exemplo trivial, considere 3 pontos A, B e C localizados em (0,0); (1,0) e (0,1) respectivamente. O número de pontos de steiner pode ser no máximo  $n-2=1$ . Neste caso, a curva C se resume a um único ponto (que chamaremos D). A função de energia portanto dependerá apenas das variáveis  $x_{D1}$  e  $x_{D2}$ . Note que o primeiro termo da energia é nulo e  $h(d_{ij})=1$  para  $i=1,2,3$  e  $j=1$ . Logo, a função de energia dependerá apenas das distâncias entre os pontos A, B, C e D respectivamente.

Considere agora o caso de 4 pontos fixos A, B, C e D com coordenadas (0,0); (2,0); (2,1) e (0,1) respectivamente ( neste caso:  $n=4$ ). Como discutido acima, poderemos ter no máximo dois pontos de steiner representados por E e F:

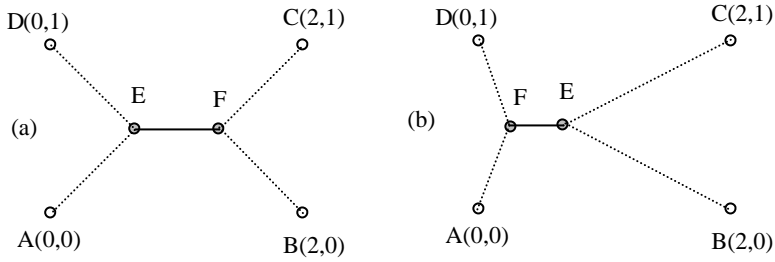


Figura IV.7: Problema de Steiner (n=4 e m=2)

Suponha por exemplo que E seja fixo e F tenha um deslocamento livre (na realidade ambos os pontos podem se deslocar). Na figura acima estão representadas duas situações distintas: em (a) temos um mínimo local obtido quando colocamos F próximo dos pontos C e B; em (b) representamos uma situação inicial onde F é colocado próximo aos pontos A e D, neste caso, F se deslocará em direção a A ou D, correspondendo cada um a um mínimo local. Teremos portanto, 3 mínimos locais. Como demonstrado em [23], a situação (a) corresponde ao mínimo global do problema.

Observando novamente a função de energia, podemos aproximar a função descontínua  $h(d_{ij})=1$  por uma sequência de funções do tipo:

$$\bar{h}(d_{ij}) = \frac{\exp(-d_{ij}^2 / \mathbf{b}_t^2)}{\sum_{k=1}^m \exp(-d_{ik}^2 / \mathbf{b}_t^2)}$$

onde  $\mathbf{b}_t \rightarrow 0$  quando  $t \rightarrow \infty$ . Em Bhaumick [23] são apresentadas outras opções na aproximação da função h. Assim, para cada i, calculamos  $\bar{h}(d_{ij})$  para  $\mathbf{b}_t \rightarrow 0$ . Logo, para cada ponto fixo i teremos a menor distância de i até um ponto j (de steiner) na curva C (vide [23]).

Conectar um ponto fixo a um ponto mais próximo na curva C nem sempre leva à menor árvore. Na verdade, pode ocorrer que em um mínimo local, algum ponto fixo esteja mais próximo de outro ponto fixo que um ponto de steiner. Neste caso, será mais conveniente fazer a conexão dos pontos fixos entre si. Vejamos o seguinte exemplo:

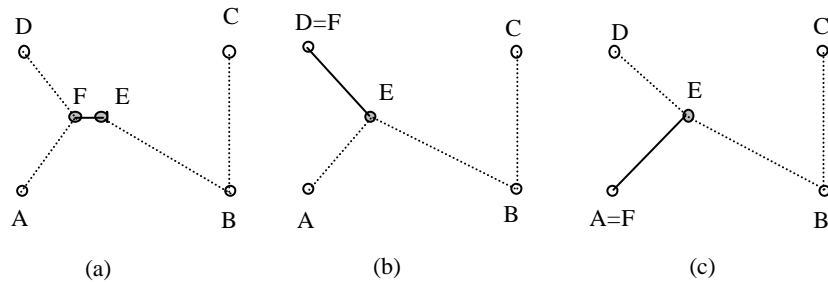


Figura IV.8: Árvore de steiner com menor comprimento

Na figura (a) colocamos inicialmente o ponto F (mantendo E fixo) próximo aos pontos A e D. Note que o segmento EC é maior que BC, logo será conveniente conectarmos os pontos B e C (árvore de menor comprimento). Em (b) e (c) estão representados dois mínimos locais decorrentes da minimização do caso (a). Em ambas as situações teremos  $BC < EC$ !

Como demonstrado em [23], a função de energia possui mecanismo para escapar de mínimos locais. Na verdade, a função de energia não decresce monotonamente com o tempo. Como demonstrado por Bhamick [23], em um ponto de equilíbrio estável (ponto de mínimo



local) temos  $\frac{\mathcal{H}^2}{\mathcal{H}_{ij}^2} > 0$ . Quando  $\mathbf{h} \rightarrow 0$ ,  $\frac{\mathcal{H}^2}{\mathcal{H}_{ij}^2}$  decresce em magnitude até se tornar negativa. Ou seja, a natureza dos pontos de equilíbrio, muda de estável para instável a medida que o tempo varia. Dessa forma, a melhor configuração (árvore "mínima" de steiner) é sempre armazenada para um instante  $t$  qualquer.

Este mecanismo é bastante diferente do que ocorre em Simulated Annealing (vide [27]), onde a probabilidade de mudança de estado da rede é baseada em algum critério de aceitação.

#### IV.4 - O Modelo de Hopfield e Problemas de Otimização Combinatória com Restrições de Desigualdade

##### IV.4.1 - Introdução:

Como discutido na seção III.3 (equação V), podemos representar a função de energia no Modelo de Hopfield através da seguinte função quadrática:

$$E = \frac{1}{2} x^T W x + d^T x \quad (\text{I})$$

onde:  $x = (x_1, \dots, x_n)^T$  é vetor de incógnitas  
 $b = (b_1, \dots, b_n)^T$  é vetor de entrada e

$$W = \begin{pmatrix} w_{11} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots \\ w_{n1} & \cdots & w_{nn} \end{pmatrix}$$

é a matriz de sinápses.

Cada elemento  $x = (x_1, \dots, x_n)^T$  deverá pertencer a um hipercubo de lado 1 contido em  $\mathcal{R}^n$ . O vetor  $x$  representa os estados de ativação de todos os neurônios da rede. Devemos ter portanto  $0 \leq x_i \leq 1$  para  $i = 1, \dots, n$ . A matriz  $W$  (matriz das sinápses), deverá ser quadrada e simétrica (ou seja,  $W = W^T$ ).

Minimizar a função  $E$  restrita a um hipercubo contido em  $\mathcal{R}^n$  será equivalente a resolver o seguinte conjunto de equações diferenciais:

$$\begin{aligned} \frac{du}{dt} &= \frac{-\mathcal{H}}{\mathcal{H}_x} = -Wx - d \\ x_i &= \frac{1}{2} (1 + \tanh(u_i)) \\ 0 &\leq x_i \leq 1 \quad \text{para } i = 1, 2, \dots, n \end{aligned}$$

Para resolver esse sistema de equações, devemos eliminar o vetor  $\mathbf{u}$  e calcular:  $\frac{dx}{dt} = \frac{dx}{du} \cdot \frac{du}{dt}$ . Assim:

$$\frac{dx}{dt} = - \begin{pmatrix} 2(1-x_1)x_1 & & 0 \\ & \ddots & \\ 0 & & 2(1-x_n)x_n \end{pmatrix} (Wx - b) \quad (\text{II})$$

Ao integrarmos (II) para um conjunto inicial de valores, caminhamos em direção a um mínimo local do problema. Não podemos entretanto, garantir convergência global.

A solução  $x(t)$  poderá convergir para:

- (a) um vértice do hipercubo;
- (b) um ponto em uma aresta do hipercubo;
- (c) um ponto interior do hipercubo.

Esperamos que a solução ótima convirja para um vértice do hipercubo (caso (a)). São discutidos em Abe [28], algumas condições de convergência do problema através de uma análise de autovalores.

Note que  $E=E(x(u(t)))$  (vide capítulo I). Logo:

$$\frac{dE}{dt} = \frac{\nabla E}{\nabla x} \cdot \frac{dx}{dt} = - \sum_{i=1}^n 2(1-x_i)x_i \left( \frac{\nabla E}{\nabla x_i} \right)^2 \leq 0$$

ou seja, a energia decresce, a medida que o tempo aumenta.

O modelo de Hopfield da maneira apresentada acima é bem aplicado a problemas de otimização combinatória que envolvam restrições de igualdade. Note por exemplo, que no problema do caixeiro viajante, caminho mínimo, problema das quatro cores entre outros, são adicionadas à função de energia parcelas correspondentes a restrições de igualdade (verifique)!! O que ocorre entretanto se trabalhamos com restrições de desigualdade? Como veremos, nestes casos, devemos fazer uma extensão do modelo de Hopfield que permita trabalhar com esta nova situação. Em Kawakami [29], são discutidas características de convergência, baseado também em uma análise de autovalores.

Suponha então que tenhamos a função de energia juntamente com novas restrições de igualdade e desigualdade como abaixo:

$$E_1 = \frac{1}{2} x^T W x + d^T x \quad (\text{III})$$

$$r_i^T x = s_i, \quad \text{para } i = 1, 2, \dots, l \quad (\text{IV})$$

$$p_i^T x \leq q_i \quad \text{ou} \quad p_i^T x \geq q_i \quad \text{para } i = 1, \dots, k \quad (\text{V e VI})$$

onde  $r_i^T = (r_{i1}, r_{i2}, \dots, r_{in})$  e  $p_i^T = (p_{i1}, \dots, p_{in})$  são vetores de  $\Re^n$  e  $q_i$  e  $s_i$  pertencem aos reais.

Como o modelo original de Hopfield não pode trabalhar com restrições de desigualdade, introduziremos uma variável  $y_i$  que converta (V) e (VI) em restrições de igualdade. Assim, para a inequação (V), teremos:

$$q_i y_i - w_i^T x = 0 \quad \text{onde } y_i \leq 1 \quad (\text{VII})$$

Da mesma forma, para (VI) temos:

$$q_i y_i - w_i^T x = 0 \quad \text{onde } y_i \geq 1 \quad (\text{VIII})$$

Para evitar confusões entre as restrições de igualdade (IV) e as restrições de igualdade (VII) e (VIII) (convertidas a partir de restrições de desigualdade), chamaremos (VII) e (VIII) de restrições de desigualdade.

Podemos portanto, formular a função de energia da seguinte maneira:

$$E = AE_1 + BE_2 + CE_3 \quad (\text{IX})$$

onde  $E_2$  é a energia correspondendo às restrições (IV) e  $E_3$  correspondendo às restrições (VII) e (VIII). Para ver como isso ocorre, vejamos inicialmente a segunda parcela: elevando o termo  $(r_i x - s_i)$  ao quadrado (equação IV) e dividindo por 2 temos:

$$E_2 = \frac{(r_i^T x - s_i)^2}{2} = \frac{x^T R x - 2s_i r_i x + s_i^2}{2} = \frac{1}{2} x^T R x + s_i^T x + \text{const} \quad (\text{X})$$

onde:  $R = \sum_{i=1}^l r_i r_i^T$  e  $s = -\sum_{i=1}^l s_i r_i$ .

Da mesma forma, repetindo o processo para a terceira parcela da equação (IX) e observando as desigualdades (VII) e (VIII) temos:

$$\begin{aligned} E_3 &= \frac{1}{2} (q_i y_i - w_i^T x)^2 = \frac{1}{2} ((q_i y_i)^2 - 2(q_i y_i)(w_i^T x) + (w_i^T x)^2) = \\ &= \frac{1}{2} ((q_i y_i)^T (q_i y_i) - y_i q_i w_i^T x - x^T q_i w_i y_i + (w_i^T x)(w_i^T x)) = \\ &= \frac{1}{2} (y_i d_i^2 y_i - y^T (q_1 w_1^T, \dots, q_l w_l^T) x - x^T (q_1 w_1, \dots, q_l w_l) y + x^T w_i w_i^T x) = \\ &= \frac{1}{2} (y^T Q y - y^T V^T x - x^T V y + x^T W x) \end{aligned}$$

onde:  $W = \sum_{i=1}^k w_i w_i^T$ ,  $V = -(q_1 w_1, \dots, q_l w_l)$  e  $Q = \begin{pmatrix} q_1^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & q_l^2 \end{pmatrix}$

Assim:  $E_3 = \frac{1}{2} \begin{pmatrix} x^T w + y^T V^T & x^T V + y^T Q \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$ . Ou ainda:

$$E_3 = \frac{1}{2} \begin{pmatrix} x^T & y^T \end{pmatrix} \begin{pmatrix} W & V \\ V^T & Q \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (\text{XI})$$

Portanto, das equações (III), (X) e (XI) temos:

$$E = \frac{A}{2} x^T W x + d^T x + \frac{B}{2} x^T R x + s^T x + \frac{C}{2} \begin{pmatrix} x^T & y^T \end{pmatrix} \begin{pmatrix} W & V \\ V^T & Q \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Reagrupando novamente as parcelas e fazendo:

$$\begin{aligned} T &\leftarrow AW + BR; \\ b &\leftarrow Ad + Bs; \\ W &\leftarrow CW; \\ V &\leftarrow CV; \\ Q &\leftarrow CQ. \end{aligned}$$

obtemos:

$$E = \frac{1}{2} \begin{pmatrix} x^T & y^T \end{pmatrix} \begin{pmatrix} W & V \\ V^T & Q \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} b^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (\text{XII})$$

O resultado seguinte (demonstrado em [29]), garante a transformação da função  $E_1$  com restrições (equações IV,V e VI) para a função E acima:

**Proposição IV.1:** Se existir constantes A,B e C tais que  $(x^*,y^*)$  minimiza a função de energia E, então ela também minimiza  $E_1$  sujeita às restrições de igualdade (equação IV) e desigualdade (equações V e VI).

Em seu trabalho, Abe [29] mostra os ajustes necessários para implementar o modelo de Hopfield utilizando a nova função de energia E acima. Através de uma análise de autovalores ele procura estabelecer parâmetros que tornem todas as soluções viáveis pontos de equilíbrio estável. Desta forma, a solução será pelo menos viável, senão ótima.

Como exemplo de transformação, vejamos o seguinte exemplo:

#### IV.4.2 - Problema da Mochila 0/1:

Esperamos encher uma mochila com no máximo n objetos. Desejamos maximizar o valor dos objetos presentes em seu interior sem ultrapassar a capacidade da mochila. Assim, temos o seguinte modelo:

$$\begin{aligned} \max \quad & \sum_{i=1}^n a_i x_i \\ \text{s.a.} \quad & \sum_{i=1}^n w_i x_i \leq d \end{aligned}$$

onde:  $a_i > 0$  (valor do objeto i);  $x_i = 0$  ou 1 (variável de decisão);  $d > 0$  (capacidade da mochila) e  $w_i > 0$  (peso do objeto i).

Transformando o problema de maximização em um problema de minimização e, as restrições de desigualdade em restrições de igualdades temos:

$$\begin{aligned} \min \quad & \frac{1}{2} \left( s - \sum_{i=1}^n a_i x_i \right)^2 \quad \text{onde } s = \sum_{i=1}^n a_i \\ \text{s.a.} \quad & dy - \sum_{i=1}^n w_i x_i = 0 \quad \text{para } 0 \leq y \leq 1. \end{aligned}$$

Segue portanto que a função de energia será dada por:

$$\begin{aligned} E &= \frac{1}{2} \left( s - \sum_{i=1}^n a_i x_i \right)^2 + \frac{A}{2} \left( dy - \sum_{i=1}^n w_i x_i \right)^2 \\ &= \frac{1}{2} \begin{pmatrix} x^T & y \end{pmatrix} \begin{pmatrix} T+W & V \\ V^T & Ad^2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + b^T x \end{aligned}$$

$$\text{onde: } T = \begin{pmatrix} a_1^2 & \cdots & a_1 a_n \\ \vdots & \ddots & \vdots \\ a_n a_1 & \cdots & a_n^2 \end{pmatrix}; \quad W = A \begin{pmatrix} w_1^2 & \cdots & w_1 w_n \\ \vdots & \ddots & \vdots \\ w_n w_1 & \cdots & w_n^2 \end{pmatrix}; \quad V = -A \begin{pmatrix} dw_1 \\ \vdots \\ dw_n \end{pmatrix}; \quad b = - \begin{pmatrix} sa_1 \\ \vdots \\ sa_n \\ 0 \end{pmatrix}$$

Temos portanto, uma função de energia quadrática (com uma matriz simétrica), necessária para a utilização do modelo de Hopfield.