

## 3

### Modelo Neuro-Fuzzy Hierárquico BSP do Tipo 2

#### 3.1

##### Introdução

Este capítulo apresenta formalmente o novo modelo neuro-fuzzy hierárquico BSP do tipo 2 (NFHB-T2). Este modelo é da classe TSK do tipo 2, no qual os conseqüentes são conjuntos fuzzy *intervalares* do tipo 1 ou combinações lineares das entradas cujos parâmetros são conjuntos fuzzy *intervalares* do tipo 1. Este modelo, além de ter a capacidade de criar e expandir automaticamente sua estrutura, de reduzir a limitação quanto ao número de entradas e de extrair regras de conhecimento a partir de um conjunto de dados, é também capaz de lidar com incertezas existentes em sistemas reais, minimizando os efeitos destas e apresentando, em conseqüência, um melhor desempenho. Além disso, este modelo fornece um intervalo de confiança para a sua saída. Estas características resultam da utilização de conjuntos fuzzy do tipo 2 e do particionamento recursivo, discutido na seção 3.2.

O modelo NFHB-T2, detalhado neste capítulo, é baseado nos modelos Neuro-Fuzzy Hierárquicos NFHB [SOUZ99] que usa conjuntos fuzzy do tipo 1. O modelo NFHB-T2 é composto por células padrão, chamadas *neuro-fuzzy BSP do tipo 2*. Essas células são dispostas numa estrutura hierárquica de árvore binária, na qual a célula de maior hierarquia gera a saída. As saídas das células de menor hierarquia são os conseqüentes das células de maior hierarquia.

Neste novo modelo híbrido, as regras do tipo 2 são geradas por um processo automático de particionamento do espaço de entrada. A explosão combinatória de regras do tipo 2, presente em alguns SIFs do tipo 2, é minimizada com um particionamento concentrado em regiões do espaço que apresentam informações deficientes ou que exigem uma precisão maior. O aspecto hierárquico reside no fato de que cada partição do espaço de entrada define um subsistema do tipo 2, que, por sua vez, pode ter como conseqüente um subsistema do tipo 2 com a mesma estrutura (recursividade).

Neste trabalho, utiliza-se o particionamento recursivo BSP (*Binary Space Partitioning*) [CHIN89], [CHRY92], [FINK74], [JANI99] estendido para regiões fuzzy do tipo 2. Esta forma de particionamento do espaço é denominada recursiva porque usa um processo recursivo na geração das partições, e foi escolhida com o objetivo de preservar a independência das entradas, o que é de fundamental importância para a interpretabilidade das regras, embora o fato de o modelo ser hierárquico e do tipo TSK prejudique um pouco esta interpretabilidade. No particionamento Quadtree, o espaço bi-dimensional é dividido sucessivamente em quatro partições (quadrantes), que, por sua vez, podem ser divididos novamente em quadrantes, numa operação recursiva. No particionamento BSP, o espaço é dividido sucessivamente, de forma recursiva, em duas regiões.

A seguir, descreve-se o particionamento BSP, base do procedimento hierárquico que dá origem à estrutura do novo modelo NFHB-T2 proposto neste trabalho. Em seguida, apresentam-se os componentes principais do modelo: sua célula básica, a arquitetura hierárquica e o algoritmo de aprendizado.

### 3.2.

#### Particionamento BSP

Nos sistemas neuro-fuzzy, o particionamento do espaço de entrada indica a forma como as regras fuzzy são relacionadas no espaço. O particionamento do espaço de saída costuma ser mais simples e está associado aos consequentes das regras. O particionamento do espaço de entrada/saída tem grande influência no desempenho de sistemas neuro-fuzzy nos aspectos de precisão, generalização, grau de interpretabilidade e geração automática de regras.

No particionamento BSP, o espaço é dividido sucessivamente em duas regiões. Este particionamento pode ser representado por uma árvore BSP que ilustra as sucessivas sub-divisões do espaço n-dimensional em subespaços fechados. O processo de construção desta árvore toma um subespaço e o divide por um hiperplano de dimensão 'n-1' que passa pelo interior deste subespaço. Isto resulta em dois novos subespaços, não necessariamente iguais, que podem ser posteriormente particionados pela aplicação do mesmo método (recursividade). A Figura 3.1 ilustra, para o caso bidimensional, um exemplo deste tipo de particionamento e sua respectiva árvore representativa.

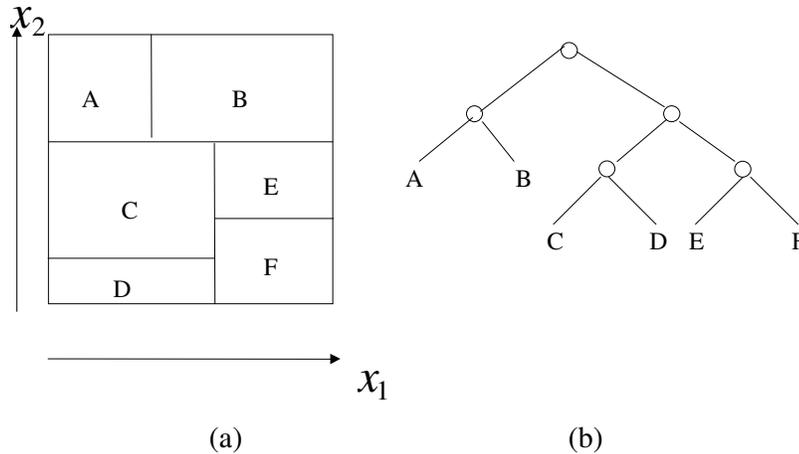


Figura 3.1 (a) Particionamento BSP. (b) Árvore BSP referente ao particionamento BSP.

A Figura 3.1 indica que o espaço foi dividido em duas partes na direção da dimensão vertical (variável  $x_2$ ); por exemplo: *baixo* e *alto*). A partição superior foi subdividida em duas novas partições A e B, segundo a direção da dimensão horizontal (variável  $x_1$ ). A partição inferior, por sua vez, foi subdividida sucessivamente, na direção horizontal e vertical, gerando finalmente as partições C, D, E e F. Cada partição final é representada por nós-folhas na árvore BSP; os nós interiores representam as diversas partições intermediárias realizadas.

O particionamento BSP é flexível e minimiza o problema do crescimento exponencial de regras, uma vez que cria apenas regras na(s) região(ões) que necessitam de mais informação. Sua principal vantagem é a construção automática de sua própria estrutura. Particionamentos recursivos resultam em modelos com hierarquia em sua estrutura e, conseqüentemente, em regras hierárquicas. A utilização de métodos recursivos de particionamento das entradas/saídas resultou em uma nova classe de SNF do tipo 1, denominada Sistemas Neuro-Fuzzy Hierárquicos que usam conjuntos fuzzy do tipo 1 [SOUZ99].

Uma extensão do particionamento BSP para regiões fuzzy do tipo 2 é utilizado para dividir o espaço de entrada e criar a estrutura hierárquica binária do modelo NFHB do tipo 2. O modelo NFHB-T2 é composto por uma ou várias células padrão, chamadas células neuro-fuzzy BSP do tipo 2 (NFB-T2). Estas células são dispostas numa estrutura hierárquica de árvore binária, e são descritas em detalhes na seção 3.3 a seguir.

### 3.3 Célula Básica Neuro-Fuzzy BSP do Tipo 2

Uma célula NFB do tipo 2 (Neuro-Fuzzy Binária do tipo 2 - NFB-T2) é um mini-sistema neuro-fuzzy intervalar do tipo 2 que realiza um particionamento fuzzy binário em um determinado espaço, através de sua variável de entrada  $x$ , segundo as funções de pertinência sigmóide do tipo 2,  $\tilde{\rho}$  (*baixo*) e  $\tilde{\mu}$  (*alto*), descritas na Figura 3.2. Nesta figura,  $\bar{\mu}_{\tilde{\rho}}(x) = \bar{\rho}(x)$ ,  $\underline{\mu}_{\tilde{\rho}}(x) = \underline{\rho}(x)$ ,  $\bar{\mu}_{\tilde{\mu}}(x) = \bar{\mu}(x)$ ,  $\underline{\mu}_{\tilde{\mu}}(x) = \underline{\mu}(x)$  são os graus de pertinência superior e inferior dos conjuntos fuzzy *intervalares* do tipo 2 *baixo*  $\tilde{\rho}$  e *alto*  $\tilde{\mu}$ , respectivamente.

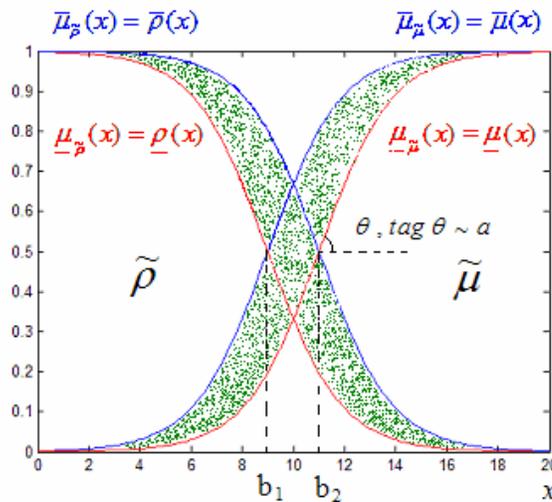


Figura 3.2. Perfil das funções de pertinência sigmóides do tipo 2 da célula NFB-T2

Deste modo, a célula NFB-T2 representa regras cujos antecedentes são definidos pelos dois conjuntos fuzzy sigmóides *intervalares* do tipo 2 associados à variável de entrada  $x$ . O valor da variável de entrada é avaliado nos conjuntos fuzzy sigmóides *intervalares* do tipo 2 dos antecedentes. Se o grau de disparo do antecedente é maior que zero, a regra é disparada. Os conseqüentes, denotados por  $\subset_i$ , são conjuntos *fuzzy intervalares* do tipo 1 e estão representados pelos seus pontos limites esquerdo  $c_{1_i}$  e direito  $c_{2_i}$ , onde  $i$  indica a  $i$ -ésima regra do sistema. Desta forma, as regras do tipo 2 neste modelo levam em consideração simultaneamente as incertezas nas funções de pertinência dos antecedentes e as incertezas nos valores dos parâmetros dos conseqüentes.

A célula NFB-T2 gera uma saída que é um conjunto fuzzy *intervalar* do tipo 1, representado por seus pontos limites esquerdo  $y_l$  e direito  $y_r$ . Este conjunto de saída é defuzzificado usando a média de  $y_l$  e  $y_r$  [MEND00-I].

$$y_{NFHBT2} = \frac{y_l + y_r}{2} \tag{Equação 3.1}$$

A Figura 3.3 ilustra o processamento de inferência da célula e o encadeamento dos conseqüentes. Similarmente aos sistemas de inferência fuzzy TSK do tipo 2, esta célula não precisa do processo de redução de tipo, uma vez que a saída do bloco de inferência é um conjunto fuzzy *intervalar* do tipo 1. Conforme ilustrado na Figura 3.3, a entrada 'x' produz os níveis de disparo das duas regras fuzzy do tipo 2 após terem sido computados os graus de pertinência superior e inferior dos conjuntos fuzzy *intervalares* do tipo 2 *baixo* e *alto*. Somente uma entrada pode estar associada a cada célula.

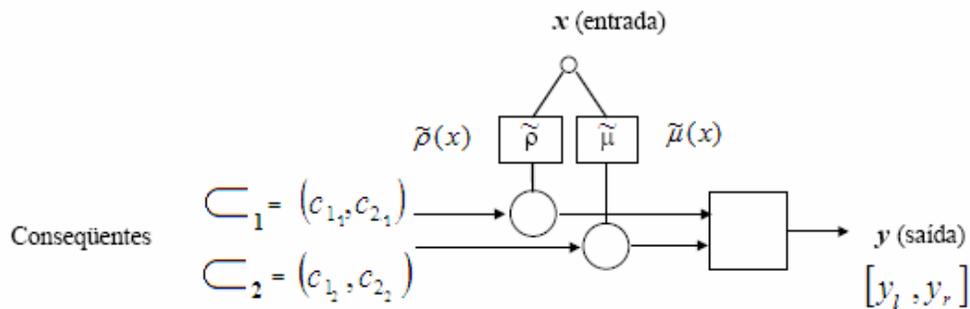


Figura 3.3. Célula Neuro-Fuzzy BSP do tipo 2

A Figura 3.4 apresenta a célula neuro-fuzzy BSP do tipo 2 de forma simplificada.

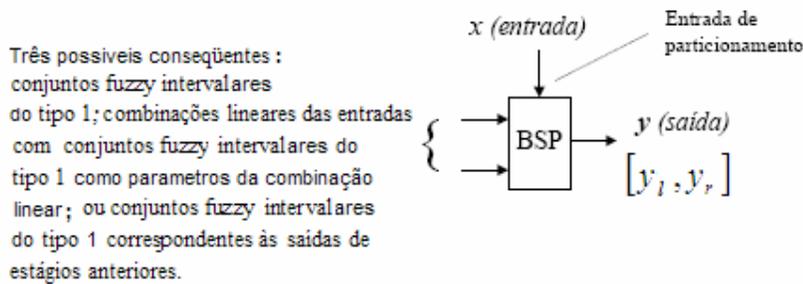


Figura 3.4. Célula simplificada NFB-T2

A interpretação lingüística do mapeamento implementado pela célula NFB-T2 é dada pelo seguinte conjunto de regras do tipo 2:

regra 1: Se  $x \in \tilde{\rho}$  então  $y = c_1$  ;  $c_1 = [c_{1_1}, c_{2_1}]$  . (partição 1)

ou

regra 2: Se  $x \in \tilde{\mu}$  então  $y = c_2$  ;  $c_2 = [c_{1_2}, c_{2_2}]$  . (partição 2)

Cada regra corresponde a uma das duas partições geradas pelo particionamento BSP. Quando o valor da entrada está contido no domínio correspondente a esta célula (este domínio é definido pelos valores do Limite Inferior-LI e Limite Superior-LS), as regras 1 e 2 são disparadas. Se o valor é inferior ao valor de ‘b<sub>1</sub>’, a regra 1 tem maior nível de disparo. Caso o valor seja superior a ‘b<sub>2</sub>’, então é a regra 2 que tem maior nível de disparo. Cada partição, por sua vez, pode ser subdividida em duas partes, através de outra célula NFB-T2.

As Figuras 3.5 e 3.6 apresentam a célula NFB-T2 sob o formato de rede neuro-fuzzy. Nestas figuras são mostradas as camadas de fuzzificação, regras do tipo 2 e o bloco de processamento da saída (defuzzificação); os  $F^i$  simbolizam o nível de disparo das regras do tipo 2, que neste modelo são conjuntos fuzzy *intervalares* do tipo 1, representados por seus pontos limites esquerdo  $\underline{f}^i$  e direito  $\bar{f}^i$ .

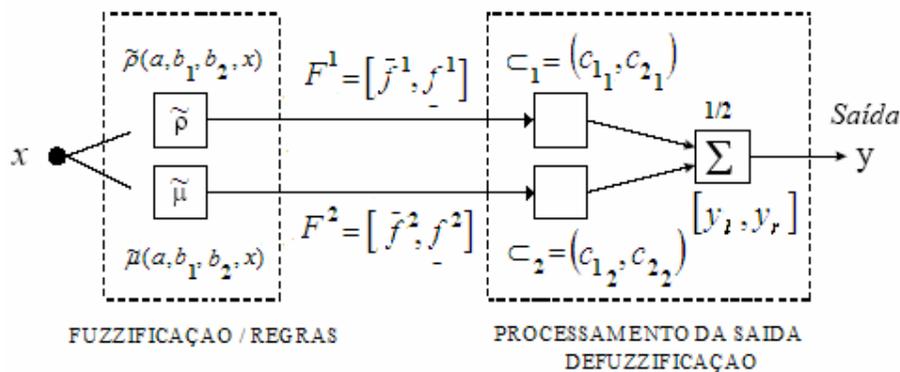


Figura 3.5. Representação da Célula NFB-T2 sob o formato de rede neuro-fuzzy.

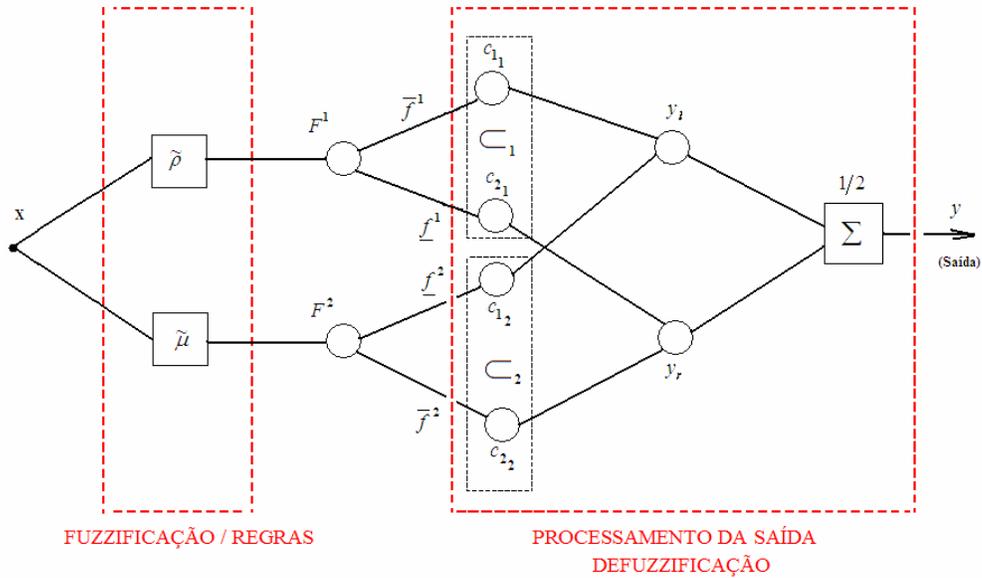


Figura 3.6. Detalhe da representação da célula NFB-T2 sob o formato de rede neuro-fuzzy.

As funções de pertinência primárias das funções de pertinência sigmóides do tipo 2 -  $\tilde{\rho}$  e  $\tilde{\mu}$  - são dadas pelas equações 3.2 e 3.3 a seguir:

$$\mu_{\rho}(x) = sig[a, b, x] = \frac{1}{1 + e^{-a(x-b)}}, \quad b \in [b_1, b_2] \quad \text{Equação 3.2}$$

$$\mu_{\mu}(x) = 1 - \mu_{\rho}(x) \quad \text{Equação 3.3}$$

Na célula NFB-T2, as funções de pertinência  $\tilde{\rho}$  e  $\tilde{\mu}$  do tipo 2 são implementadas de modo que:

$$\underline{\rho}(x) + \underline{\mu}(x) = 1 ; \quad \underline{\rho}(x) + \underline{\mu}(x) = 1.$$

A equação 3.2 descreve as funções de pertinência primárias sigmóide de  $\tilde{\rho}$  e  $\tilde{\mu}$  com inclinação 'a', de valor fixo, e ponto médio de transição com incerteza que assume valores em  $[b_1, b_2]$ . Na célula NFB-T2,  $\underline{\rho}(x)$  e  $\underline{\mu}(x)$  têm como ponto médio de transição 'b1', e  $\overline{\rho}(x)$  e  $\overline{\mu}(x)$  tem como ponto médio de transição 'b2', conforme mostrado na Figura 3.2.

O conjunto fuzzy *intervalar* do tipo 1 de saída  $y = [y_l, y_r]$  de uma célula NFB-T2 é calculado através das equações 3.4 e 3.5 e do teorema C.1 (apêndice C) [MEND00-I]:

$$y_l = \frac{\sum_{i=1}^{M=2} g_l^i y_l^i}{\sum_{i=1}^{M=2} g_l^i} = \frac{g_l^1 c_{1_1} + g_l^2 c_{1_2}}{g_l^1 + g_l^2}; \quad \text{Equação 3.4}$$

$$y_r = \frac{\sum_{i=1}^{M=2} g_r^i y_r^i}{\sum_{i=1}^{M=2} g_r^i} = \frac{g_r^1 c_{2_1} + g_r^2 c_{2_2}}{g_r^1 + g_r^2}; \quad \text{Equação 3.5}$$

onde:

M = número de regras do tipo 2 da célula. (M=2)

$g_l^i$  = ponto limite (esquerdo  $\underline{f}^i$  ou direito  $\bar{f}^i$ ) de  $F^i$  associado a  $c_{1_i}$  do conjunto conseqüente  $\subset_i$ , para a  $i$ -ésima regra. Os valores de  $g_l^i$  são calculados através do teorema C.1.

$g_r^i$  = ponto limite (esquerdo  $\underline{f}^i$  ou direito  $\bar{f}^i$ ) de  $F^i$  associado a  $c_{2_i}$  do conjunto conseqüente  $\subset_i$ , para a  $i$ -ésima regra. Os valores de  $g_r^i$  são calculados através do teorema C.1.

Este conjunto de saída é defuzzificado usando-se a média de  $y_l$  e  $y_r$  (equações 3.6 e 3.7), obtendo-se:

$$y = \frac{y_l + y_r}{2} \quad \text{Equação 3.6}$$

$$y = \frac{y_l + y_r}{2} = \frac{1}{2} \left( \frac{g_l^1 c_{1_1} + g_l^2 c_{1_2}}{g_l^1 + g_l^2} + \frac{g_r^1 c_{2_1} + g_r^2 c_{2_2}}{g_r^1 + g_r^2} \right) \quad \text{Equação 3.7}$$

onde cada conseqüente  $\subset_i = [c_{1_i}, c_{2_i}]$  corresponde a um dos três conseqüentes possíveis abaixo:

- a. um conjunto fuzzy *intervalar* do tipo 1 (para uma célula NFB-T2 *intervalar* TSK de ordem zero): caso em que  $\subset_i = [c_{1_i}, c_{2_i}]$ .
- b. um conjunto fuzzy *intervalar* do tipo 1 resultante da combinação linear das entradas com conjuntos fuzzy *intervalares* do tipo 1 como parâmetros da combinação linear (para uma célula NFB-T2 *intervalar* TSK de 1ª ordem): caso em que  $\subset_i = [c_{1_i}, c_{2_i}]$  é dado pelas equações 3.8 e 3.9. [MEND00-I].

$$c_{1_i} = \sum_{k=0}^n \varsigma_{1_{k_i}} x_k = \sum_{k=0}^n (\varsigma_{k_i} - \delta_{k_i}) x_k = \varsigma_{0_i} + \sum_{k=1}^n \varsigma_{k_i} x_k - \delta_{0_i} - \sum_{k=1}^n \delta_{k_i} |x_k|$$

Equação 3.8

$$c_{2_i} = \sum_{k=0}^n \varsigma_{2_{k_i}} x_k = \sum_{k=0}^n (\varsigma_{k_i} + \delta_{k_i}) x_k = \varsigma_{0_i} + \sum_{k=1}^n \varsigma_{k_i} x_k + \delta_{0_i} + \sum_{k=1}^n \delta_{k_i} |x_k|$$

Equação 3.9

onde:

$x_k$  =  $k$ -ésima entrada do sistema.

$n$  = número total de entradas do sistema.

$\varsigma_{k_i}$  = centro (média) do conjunto fuzzy *intervalar* do tipo 1 que representa o peso da entrada  $x_k$  na combinação linear para a  $i$ -ésima regra.

$\delta_{k_i}$  = extensão do conjunto fuzzy *intervalar* do tipo 1 que representa o peso da entrada  $x_k$  na combinação linear para a  $i$ -ésima regra.

$\varsigma_{0_i}$  = centro (média) do conjunto fuzzy *intervalar* do tipo 1 que representa o peso sem entrada na combinação linear para a  $i$ -ésima regra.

$\delta_{0_i}$  = extensão do conjunto fuzzy *intervalar* do tipo 1 que representa o peso sem entrada na combinação linear para a  $i$ -ésima regra.

$\varsigma_{1_{k_i}}$  = ponto limite esquerdo do conjunto fuzzy *intervalar* do tipo 1 que representa o peso da entrada  $x_k$  na combinação linear para a  $i$ -ésima regra.

$\varsigma_{2_{k_i}}$  = ponto limite direito do conjunto fuzzy *intervalar* do tipo 1 que representa o peso da entrada  $x_k$  na combinação linear para a  $i$ -ésima regra.

- c. um conjunto fuzzy *intervalar* do tipo 1 correspondente à saída de um estágio de nível anterior:

caso em que  $C_i = y_j$  ;

onde:

$$C_i = [c_{1_i}, c_{2_i}] ;$$

$y_j = [y_l^j, y_r^j]$  = saída da célula NFB-T2 genérica 'j' calculada pela equação 3.7.

No caso da célula NFB-T2 utilizar como conseqüentes combinações lineares das entradas, o conjunto fuzzy *intervalar* do tipo 1 de saída  $y = [y_l, y_r]$  é calculado através das equações 3.10 e 3.11 e do teorema C.1 [MEND00-I].

$$y_l = \left[ \frac{g_l^1 \sum_{k=0}^n (\zeta_{k_1} - \delta_{k_1}) x_k + g_l^2 \sum_{k=0}^n (\zeta_{k_2} - \delta_{k_2}) x_k}{g_l^1 + g_l^2} \right] \quad \text{Equação 3.10}$$

$$y_r = \left[ \frac{g_r^1 \sum_{k=0}^n (\zeta_{k_1} + \delta_{k_1}) x_k + g_r^2 \sum_{k=0}^n (\zeta_{k_2} + \delta_{k_2}) x_k}{g_r^1 + g_r^2} \right] \quad \text{Equação 3.11}$$

Portanto, a saída (crisp) y da célula será dada pela equação 3.12, onde  $\zeta_{k_i}$  e  $\delta_{k_i}$  são também parâmetros ajustáveis dos conseqüentes.

$$y = \frac{1}{2} \left[ \left[ \frac{g_l^1 \sum_{k=0}^n (\zeta_{k_1} - \delta_{k_1}) x_k + g_l^2 \sum_{k=0}^n (\zeta_{k_2} - \delta_{k_2}) x_k}{g_l^1 + g_l^2} \right] + \left[ \frac{g_r^1 \sum_{k=0}^n (\zeta_{k_1} + \delta_{k_1}) x_k + g_r^2 \sum_{k=0}^n (\zeta_{k_2} + \delta_{k_2}) x_k}{g_r^1 + g_r^2} \right] \right]$$

Equação 3.12

Como na célula NFB-T2 básica, as funções de pertinência sigmoidais do tipo 2 -  $\tilde{\mu}$  (*alto*) e  $\tilde{\rho}$  (*baixo*) - são tais que:

$$\bar{\rho}(x) + \underline{\mu}(x) = 1 \quad \text{Equação 3.13}$$

$$\underline{\rho}(x) + \bar{\mu}(x) = 1 \quad \text{Equação 3.14}$$

Do teorema C.1 [MEND00-I] sabe-se que (apêndice C):

$$\text{Se } c_{1_1} \langle c_{1_2} \text{ então } g_l^1 = \bar{f}^1 = \bar{\rho}(x) \text{ e } g_l^2 = \underline{f}^2 = \underline{\mu}(x)$$

$$\text{Se } c_{1_1} \rangle c_{1_2} \text{ então } g_l^1 = \underline{f}^1 = \underline{\rho}(x) \text{ e } g_l^2 = \bar{f}^2 = \bar{\mu}(x)$$

$$\text{Se } c_{2_1} \langle c_{2_2} \text{ então } g_r^1 = \underline{f}^1 = \underline{\rho}(x) \text{ e } g_r^2 = \bar{f}^2 = \bar{\mu}(x)$$

$$\text{Se } c_{2_1} \rangle c_{2_2} \text{ então } g_r^1 = \bar{f}^1 = \bar{\rho}(x) \text{ e } g_r^2 = \underline{f}^2 = \underline{\mu}(x)$$

Equação 3.15

Das considerações acima, obtém-se uma simplificação no cálculo da saída quando a célula NFB-T2 utiliza como conseqüente um conjunto fuzzy *intervalar* do tipo 1 (equações 3.4, 3.5 e 3.7), pois as somas dadas pelas equações 3.13 e 3.14 são iguais a 1 para qualquer valor de ‘x’. Assim, a equação de saída fica:

$$y = \frac{1}{2} [g_l^1 c_{1_1} + g_l^2 c_{1_2} + g_r^1 c_{2_1} + g_r^2 c_{2_2}] \quad \text{Equação 3.16}$$

De forma similar, obtém-se uma simplificação no cálculo da saída no caso da célula NFB-T2 utilizar como conseqüente um conjunto fuzzy *intervalar* do tipo 1 resultante das combinações lineares das entradas – com conjuntos fuzzy *intervalares* do tipo 1 como parâmetros da combinação linear (equações 3.10, 3.11 e 3.12) –, pois as somas dadas pelas equações 3.13 e 3.14 são iguais a 1 para qualquer valor de ‘x’. Assim, a equação de saída é:

$$y = \frac{1}{2} \left[ g_l^1 \sum_{k=0}^n (\zeta_{k_1} - \delta_{k_1}) x_k + g_l^2 \sum_{k=0}^n (\zeta_{k_2} - \delta_{k_2}) x_k + g_r^1 \sum_{k=0}^n (\zeta_{k_1} + \delta_{k_1}) x_k + g_r^2 \sum_{k=0}^n (\zeta_{k_2} + \delta_{k_2}) x_k \right]$$

Equação 3.17

Neste caso, na célula NFB-T2 os  $\bar{f}^1$  e  $\underline{f}^1$  se confundem com os  $\bar{\rho}(x)$  e  $\underline{\rho}(x)$ , e os  $\bar{f}^2$  e  $\underline{f}^2$  se confundem com os  $\bar{\mu}(x)$  e  $\underline{\mu}(x)$ , respectivamente.

As células NFB-T2 formam uma estrutura hierárquica que resulta nas regras do tipo 2 do sistema.

### 3.4

#### Arquitetura NFHB-T2

O modelo NFHB-T2 é criado a partir da interligação de várias células NFB-T2 na forma de uma árvore binária. Essas células são dispostas em uma estrutura hierárquica. A célula de maior hierarquia gera a saída; as células de menor hierarquia operam como conseqüentes das de maior hierarquia. Somente na célula raiz é efetuada a defuzzificação. A Figura 3.7 ilustra uma pequena arquitetura NFHB-T2; o seu respectivo particionamento BSP é mostrado na Figura 3.8. Na árvore binária, a raiz simboliza todo o espaço a ser particionado; cada partição não subdividida é chamada de bi-partição.

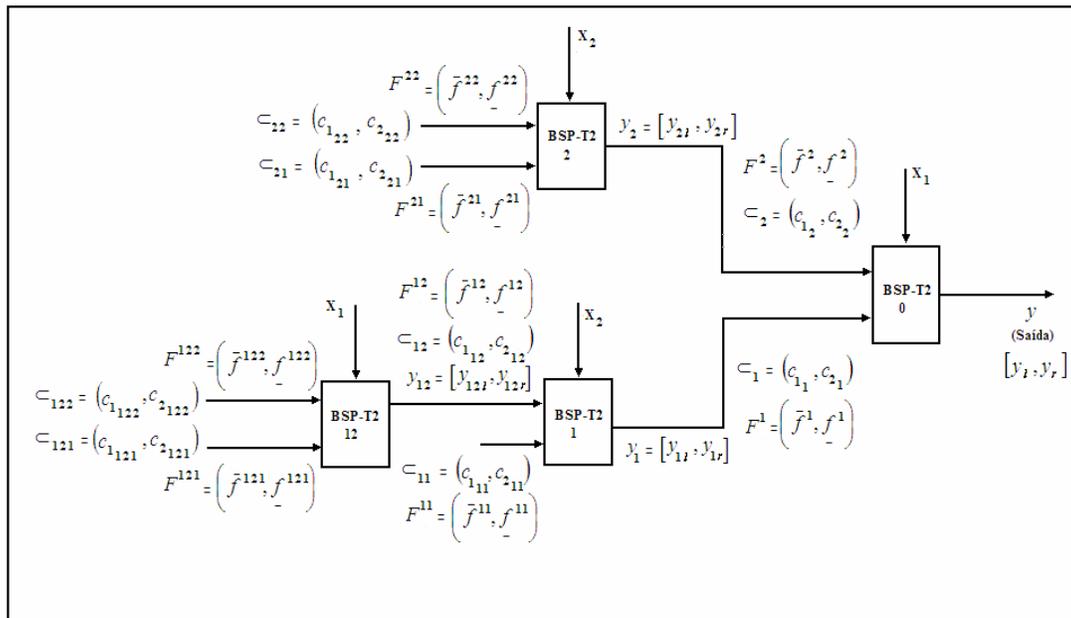


Figura 3.7 - Exemplo de uma arquitetura do modelo NFHB-T2

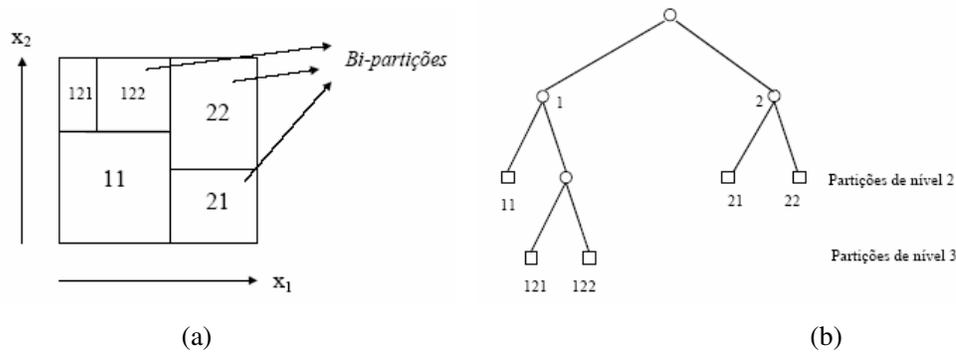


Figura 3.8 - (a) Particionamento do espaço de entrada do sistema NFHB-T2 descrito na Figura 3.7.(b) árvore representativa do particionamento BSP mostrado na Figura 3.8.(a).

Na arquitetura NFHB-T2 da Figura 3.7, as partições iniciais 1 e 2 (célula ‘BSP-T2 0’) foram subdivididas; portanto, os conseqüentes de suas regras do tipo 2 são as saídas (conjuntos fuzzy *intervalares* do tipo 1) dos subsistemas 1 e 2, isto é,  $y_1 = [y_{1l}, y_{1r}]$  e  $y_2 = [y_{2l}, y_{2r}]$ , respectivamente. Estes, por sua vez, têm como conseqüentes  $c_{11} = [c_{11l}, c_{11r}]$ ,  $y_{12} = [y_{12l}, y_{12r}]$ ,  $c_{21} = [c_{21l}, c_{21r}]$ ,  $c_{22} = [c_{22l}, c_{22r}]$ . O conjunto fuzzy *intervalar* do tipo 1, conseqüente  $y_{12} = [y_{12l}, y_{12r}]$ , é a saída da célula ‘BSP-T2 12’. Cada  $c_i = [c_{il}, c_{ir}]$  corresponde a um conjunto fuzzy *intervalar* do tipo 1, caso se faça uso de conseqüentes de Sugeno de ordem 0, ou a uma combinação linear das entradas com conjuntos fuzzy *intervalares* do tipo 1, caso se utilizem conseqüentes de Sugeno de 1ª ordem.

Durante o processo de aprendizado e geração da estrutura, o modelo NFHB-T2 usa uma estratégia de seleção de variáveis de entrada para distribuir estas variáveis em cada nível da árvore BSP de acordo com as suas relevâncias (seqüência de variáveis, seção 3.6). Dependendo do processo de convergência do modelo, os atributos de menor relevância podem não ser usados na geração da árvore. Quando todos os atributos são usados, a distribuição das variáveis de entrada em cada nível da árvore repete a seqüência de variáveis inicialmente obtida.

O conjunto fuzzy *intervalar* do tipo 1 de saída  $y = [y_l, y_r]$  do sistema da Figura 3.7 é calculado através das equações 3.18 e 3.19 e do teorema C.1 [MEND00-I] (mais detalhes encontram-se no apêndice A), que determinam os pontos limites deste conjunto fuzzy *intervalar* do tipo 1 de saída:

$$y_l = \frac{g_l^1 \left( g_l^{11} c_{1_{11}} + g_l^{12} \left( \frac{g_l^{121} c_{1_{121}} + g_l^{122} c_{1_{122}}}{g_l^{121} + g_l^{122}} \right) \right) + g_l^2 \left( \frac{g_l^{21} c_{1_{21}} + g_l^{22} c_{1_{22}}}{g_l^{21} + g_l^{22}} \right)}{g_l^1 + g_l^2}$$

Equação 3.18

$$y_r = \frac{g_r^1 \left( (g_r^{11} c_{2_{11}} + g_r^{12}) + g_r^2 \left( \frac{g_r^{121} c_{2_{121}} + g_r^{122} c_{2_{122}}}{g_r^{121} + g_r^{122}} \right) \right) + g_r^2 \left( \frac{g_r^{21} c_{2_{21}} + g_r^{22} c_{2_{22}}}{g_r^{21} + g_r^{22}} \right)}{g_r^1 + g_r^2}$$

Equação 3.19

Aplicando as equações 3.13, 3.14 e 3.15 nas equações 3.18 e 3.19 (apêndice A), tem-se:

$$y_l = g_l^1 (g_l^{11} c_{1_{11}} + g_l^{12} (g_l^{121} c_{1_{121}} + g_l^{122} c_{1_{122}})) + g_l^2 (g_l^{21} c_{1_{21}} + g_l^{22} c_{1_{22}}) \quad \text{Equação 3.20}$$

$$y_r = g_r^1 (g_r^{11} c_{2_{11}} + g_r^{12} (g_r^{121} c_{2_{121}} + g_r^{122} c_{2_{122}})) + g_r^2 (g_r^{21} c_{2_{21}} + g_r^{22} c_{2_{22}}) \quad \text{Equação 3.21}$$

Logo, a saída crisp do sistema da Figura 3.7 é dada por:

$$y = 1/2 [g_l^1 (g_l^{11} c_{1_{11}} + g_l^{12} (g_l^{121} c_{1_{121}} + g_l^{122} c_{1_{122}})) + g_l^2 (g_l^{21} c_{1_{21}} + g_l^{22} c_{1_{22}})] + \\ 1/2 [g_r^1 (g_r^{11} c_{2_{11}} + g_r^{12} (g_r^{121} c_{2_{121}} + g_r^{122} c_{2_{122}})) + g_r^2 (g_r^{21} c_{2_{21}} + g_r^{22} c_{2_{22}})]$$

Equação 3.22

O conjunto fuzzy *genérico intervalar* de saída do tipo 1 e a saída crisp genérica de um sistema NFHB-T2 de 3 níveis de hierarquia são dados pelas equações 3.23, 3.24 e 3.25 a seguir. As variáveis  $k_i$ ,  $k_{ij}$  e  $k_{ijk}$  assumem apenas valores iguais a ‘0’ ou ‘1’, indicando a existência ou não das bi-partições de ordem ‘i’, ‘ij’ e ‘ijk’, respectivamente.

$$y_l = \sum_{i=1}^{M=2} g_l^i k_i c_{1_i} + \sum_{i=1}^{M=2} \sum_{j=1}^{M=2} g_l^i g_l^{ij} k_{ij} c_{1_{ij}} + \sum_{i=1}^{M=2} \sum_{j=1}^{M=2} \sum_{k=1}^{M=2} g_l^i g_l^{ij} g_l^{ijk} k_{ijk} c_{1_{ijk}} \quad \text{Equação 3.23}$$

$$y_r = \sum_{i=1}^{M=2} g_r^i k_i c_{2_i} + \sum_{i=1}^{M=2} \sum_{j=1}^{M=2} g_r^i g_r^{ij} k_{ij} c_{2_{ij}} + \sum_{i=1}^{M=2} \sum_{j=1}^{M=2} \sum_{k=1}^{M=2} g_r^i g_r^{ij} g_r^{ijk} k_{ijk} c_{2_{ijk}} \quad \text{Equação 3.24}$$

$$y = \frac{y_L + y_r}{2} \quad \text{Equação 3.25}$$

onde:

$g_l^i, g_l^{ij}, g_l^{ijk}$  = pontos limites (esquerdos,  $\underline{f}^i, \underline{f}^{ij}, \underline{f}^{ijk}$ , ou direitos,  $\bar{f}^i, \bar{f}^{ij}, \bar{f}^{ijk}$ ) do conjunto fuzzy *intervalar* do tipo 1 do nível de disparo das regras do tipo 2 de cada bi-partição  $i, ij, ijk$ , respectivamente, que estão associados aos  $c_{1_i}, c_{1_{ij}}, c_{1_{ijk}}$ ; os valores de  $g_l^i, g_l^{ij}, g_l^{ijk}$  são calculados através do teorema C.1.

$g_r^i, g_r^{ij}, g_r^{ijk}$  = pontos limites (esquerdos,  $\underline{f}^i, \underline{f}^{ij}, \underline{f}^{ijk}$ , ou direitos,  $\bar{f}^i, \bar{f}^{ij}, \bar{f}^{ijk}$ ) do conjunto fuzzy *intervalar* do tipo 1 do nível de disparo das regras do tipo 2 de cada bi-partição  $i, ij, ijk$ , respectivamente, que estão associados aos  $c_{2_i}, c_{2_{ij}}, c_{2_{ijk}}$ ; os valores de  $g_r^i, g_r^{ij}, g_r^{ijk}$  são calculados através do teorema C.1.

$c_{1_i}, c_{1_{ij}}, c_{1_{ijk}}$  = pontos limites esquerdos dos conjuntos fuzzy *intervalares* do tipo 1 conseqüentes das regras do tipo 2 de cada bipartição  $i, ij, ijk$ .

$c_{2_i}, c_{2_{ij}}, c_{2_{ijk}}$  = pontos limites direitos dos conjuntos fuzzy *intervalares* do tipo 1 conseqüentes das regras do tipo 2 de cada bipartição  $i, ij, ijk$ .

$k_i, k_{ij}, k_{ijk}$  = variáveis cujo valor será igual a ‘1’ se a partição  $i, ij, ijk$ , existir, e ‘0’ caso contrario.

O conjunto de regras que traduzem, linguisticamente, o conhecimento da estrutura do exemplo da Figura 3.7 é:

$$\begin{aligned} & \text{if } x_1 \in \tilde{\rho}_0 \text{ then} \\ & \quad \{ \\ & \quad \text{if } x_2 \in \tilde{\rho}_1 \text{ then } y = [c_{1_{11}}, c_{2_{11}}] \end{aligned}$$

$$\begin{aligned}
 & \text{if } x_2 \in \tilde{\mu}_1 \text{ then} \\
 & \quad \{ \\
 & \quad \quad \text{if } x_1 \in \tilde{\rho}_{12} \text{ then } y = [c_{1_{121}}, c_{2_{121}}] \\
 & \quad \quad \text{if } x_1 \in \tilde{\mu}_{12} \text{ then } y = [c_{1_{122}}, c_{2_{122}}] \\
 & \quad \} \\
 & \} \\
 & \text{if } x_1 \in \tilde{\mu}_0 \text{ then} \\
 & \quad \{ \\
 & \quad \quad \text{if } x_2 \in \tilde{\rho}_2 \text{ then } y = [c_{1_{21}}, c_{2_{21}}] \\
 & \quad \quad \text{if } x_2 \in \tilde{\mu}_2 \text{ then } y = [c_{1_{22}}, c_{2_{22}}] \\
 & \quad \}
 \end{aligned}$$

onde:

$\tilde{\rho}_0, \tilde{\mu}_0, \tilde{\rho}_1, \tilde{\mu}_1, \tilde{\rho}_2, \tilde{\mu}_2, \tilde{\rho}_{12}, \tilde{\mu}_{12}$  são as funções de pertinência do tipo 2 que definem as partições de nível 0, 1, 2 e 12, correspondentes às células ‘BSP-T2 0’, ‘BSP-T2 1’, ‘BSP-T2 2’, ‘BSP-T2 12’, respectivamente.

Cada uma das funções de pertinência do tipo 2 acima possui 3 parâmetros ajustáveis ( $'a'$ ,  $'b_1'$ , e  $'b_2'$ ), os quais definem o perfil das funções *alto*  $\tilde{\mu}$  e *baixo*  $\tilde{\rho}$  de cada variável de entrada. O parâmetro  $'a'$  define a inclinação das funções de pertinência tipo 2 das células. Os parâmetros  $'b_1'$ , e  $'b_2'$ , definem a incerteza no ponto médio de transição das funções de pertinência tipo 2. Se a estrutura NFHB-T2 possui células em diferentes níveis com a mesma variável de entrada, estes três parâmetros de cada uma das funções de pertinência do tipo 2 de cada célula são fixados (no caso do modelo NFHB-T2 ser treinado com particionamento fixo) ou inicializados (no caso do modelo ser treinado com particionamento adaptativo) de acordo com a seguinte heurística:

- a. **Parâmetro  $'a'$ :** o valor do parâmetro  $'a'$  de uma célula com uma determinada variável de entrada pertencente a um nível inferior é o dobro do valor de  $'a'$  das funções de pertinência da célula com a mesma variável de entrada pertencente ao nível superior. Isto permite definir como um hiperparâmetro, no caso de particionamento fixo, o parâmetro  $'a'$  de uma célula com uma determinada entrada pertencente ao nível superior.

b. **Parâmetros ‘ $b_1$ ’ e ‘ $b_2$ ’:** são definidos de modo que as células filhas com a mesma variável de entrada são particionadas segundo uma das opções a seguir:

- **opção 1:** a média dos parâmetros ‘ $b_1$ ’ e ‘ $b_2$ ’ das funções de pertinência do tipo 2 das células com uma determinada variável de entrada pertencente a um nível inferior coincide com a metade das bipartições da célula com a mesma variável de entrada pertencente ao nível superior. A Figura 3.9 ilustra esta opção.

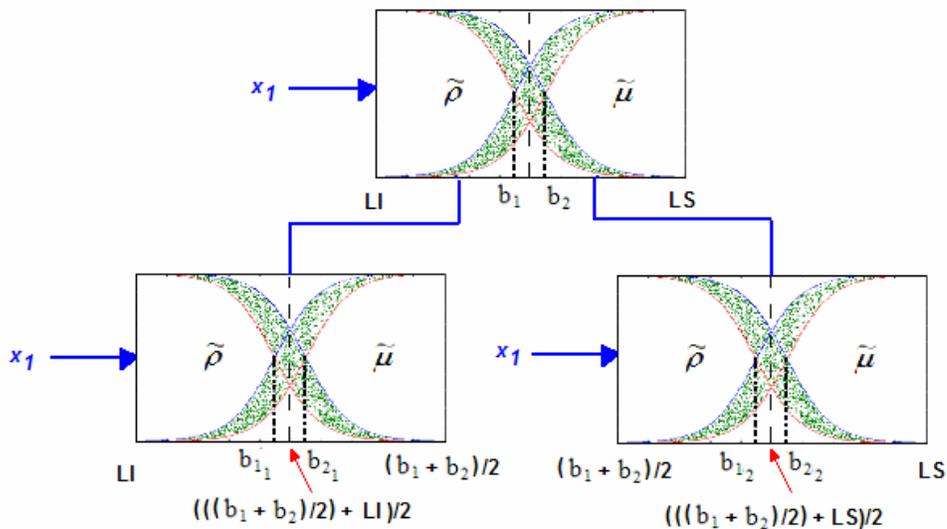


Figura 3.9 – Opção 1 para o modo de definição dos parâmetros  $b_1$ , e ‘ $b_2$ ’.

- **opção 2:**
  - o parâmetro ‘ $b_2$ ’ das funções de pertinência do tipo 2 da célula com uma determinada variável de entrada pertencente a um nível superior coincide com o limite superior do domínio dos conjuntos fuzzy *intervalares* do tipo 2 da célula com a mesma variável de entrada pertencente ao nível inferior resultante da decomposição do primeiro quadrante do nível superior;
  - o parâmetro ‘ $b_1$ ’ das funções de pertinência do tipo 2 da célula com uma determinada variável de entrada pertencente ao nível superior coincide com o limite inferior do domínio dos conjuntos fuzzy

*intervalares* do tipo 2 da célula com a mesma variável de entrada pertencente ao nível inferior resultante da decomposição do segundo quadrante do nível superior;

Assim, a média dos parâmetros ' $b_1$ ' e ' $b_2$ ' das funções de pertinência do tipo 2 das células com uma determinada variável de entrada pertencentes ao nível inferior deve coincidir com a metade do intervalo do universo de discurso da variável de entrada desta célula. A Figura 3.10 ilustra esta opção.

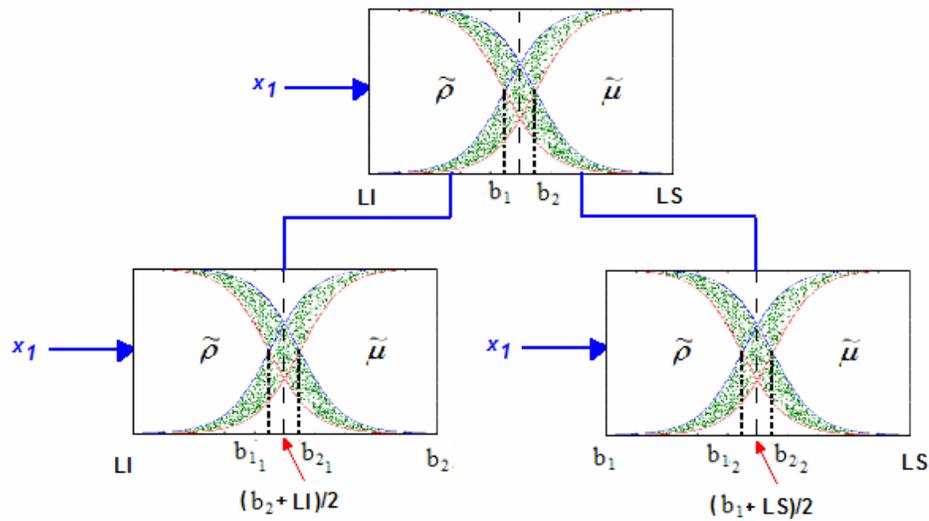


Figura 3.10 – Opção 2 para o modo de definição dos parâmetros ' $b_1$ ', e ' $b_2$ '.

Durante o treinamento, foram consideradas restrições para limitar os valores dos parâmetros ' $b_1$ ' e ' $b_2$ ' (apêndice B). A seguir descreve-se o algoritmo de aprendizado utilizado no modelo NFHB-T2.

### 3.5 Algoritmo de Aprendizado

O aprendizado de sistemas neuro-fuzzy é geralmente dividido em duas partes: a identificação da estrutura e o ajuste dos parâmetros. O modelo NFHB-T2 realiza essas duas tarefas de aprendizado em um único algoritmo. Assim, o modelo NFHB-T2 tem a capacidade de construir a sua própria estrutura de regras hierárquicas.

O processo de aprendizado do modelo NFHB-T2 é efetuado em oito passos, conforme apresentado no fluxograma da Figura 3.11. Este algoritmo considera 2 tipos de consequentes possíveis: conjuntos fuzzy *intervalares* do tipo 1 ou combinações lineares das entradas com conjuntos fuzzy *intervalares* do tipo 1. No modelo NFHB-T2, os  $c_i = [c_{1_i}, c_{2_i}]$  e os parâmetros 'a', 'b<sub>1</sub>' e 'b<sub>2</sub>' são encarados como sendo os pesos fuzzy do modelo.

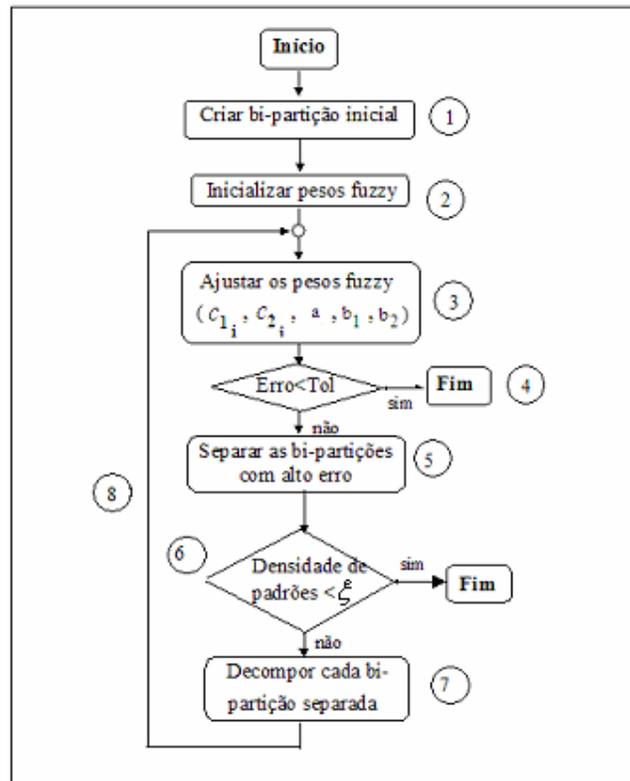


Figura 3.11 - Algoritmo de aprendizado do modelo NFHB-T2.

A seguir são detalhados os oito passos do algoritmo:

#### a. Passo 1:

**Cria-se a bi-partição inicial** dividindo-se em duas partes o espaço de entrada, através de dois conjuntos fuzzy *intervalares* do tipo 2, *alto* e *baixo*, da variável de entrada  $x$ . Neste passo, é criada a primeira célula NFB-T2, chamada de célula raiz, que tem como domínio dos seus conjuntos fuzzy *intervalares* do tipo 2 os valores mínimo e máximo da entrada  $x$  (Limite Inferior – LI e Limite Superior –LS).

**b. Passo 2:**

Neste passo, **os parâmetros ajustáveis são inicializados** da forma descrita abaixo:

- **Parâmetros**  $c_{1_i}, c_{2_i}$  (pesos fuzzy) correspondentes ao conseqüente  $\subset_i$  :

$$c_{1_i} = \zeta_i - \delta_i \quad \text{e} \quad c_{2_i} = \zeta_i + \delta_i ,$$

onde  $\zeta_i$  e  $\delta_i$  são o centro e a extensão do conjunto fuzzy *intervalar* do tipo 1 do conseqüente da regra do tipo 2 da *i-ésima* bipartição, respectivamente. Desta forma,  $\zeta_i$  e  $\delta_i$  são inicializados, de forma heurística, com valores pequenos. Esse processo se aplica aos conjuntos fuzzy *intervalares* do tipo 1 dos conseqüentes ou, no caso de os conseqüentes serem combinações lineares das entradas com conjuntos fuzzy *intervalares* do tipo 1, aos parâmetros que representam os pesos das entradas nas equações 3.8 e 3.9.

- **Parâmetros** ' $b_1$ ', e ' $b_2$ ' dos antecedentes das regras do tipo 2:

$$b_1 = b - \lambda \quad \text{e} \quad b_2 = b + \lambda$$

onde:  $b = (b_1 + b_2)/2$

$\lambda$  = porcentagem de extensão para fixar  $b_1$  e  $b_2$ .

Desta forma,  $b$  é inicializado com valor igual à metade do intervalo do universo de discurso da variável de entrada da célula.  $\lambda$  é inicializado com valor igual a 1% do dito intervalo.

- **Parâmetro** ' $a$ ' dos antecedentes das regras: é inicializado, de forma heurística, com o valor igual ao dobro do inverso do universo de discurso da variável de entrada da célula.

As Equações 3.26, 3.27, 3.28, 3.29 e 3.30 resumem a inicialização de ' $a$ ', ' $b_1$ ', e ' $b_2$ '.

$$a = 2/(LS - LI) \quad \text{Equação 3.26}$$

$$b_1 = b - \lambda \quad \text{Equação 3.27}$$

$$b_2 = b + \lambda \quad \text{Equação 3.28}$$

$$\lambda = 0.01(LS - LI) \quad \text{Equação 3.29}$$

$$b = (b_1 + b_2)/2 = (LS + LI)/2 \quad \text{Equação 3.30}$$

onde:  $LI$  e  $LS$  são, respectivamente, os limites inferior e superior do universo de discurso da variável de entrada do particionamento da célula.

**c. Passo 3:**

Este passo refere-se ao **ajuste dos pesos fuzzy**. Para este processo de ajuste, são sugeridas as seguintes heurísticas, baseadas no método de *Gradient Descent*:

- o método de *Gradient Descent* ajusta apenas os pesos fuzzy  $c_{1_i}$  e  $c_{2_i}$  correspondentes aos consequentes  $\subset_i$  (conjuntos fuzzy *intervalares* do tipo 1 ou parâmetros que representam os pesos das entradas na combinação linear das equações 3.8 e 3.9). Neste caso, utiliza-se o particionamento fixo, não havendo ajuste dos perfis dos antecedentes. Assim, mantêm-se como valores dos parâmetros antecedentes aqueles definidos durante o processo de inicialização.
- o método de *Gradient Descent* ajusta tanto os pesos fuzzy  $c_{1_i}$  e  $c_{2_i}$ , correspondentes aos consequentes  $\subset_i$ , quanto os parâmetros ' $a$ ', ' $b_1$ ', e ' $b_2$ ' das funções de pertinência do tipo 2 dos antecedentes. Neste caso, tem-se o particionamento adaptativo.

Através do método de *Gradient Descent*, cada parâmetro ajustável é atualizado segundo as equações abaixo.

$$\theta(n+1) = \theta(n) - \alpha_\theta \left. \frac{\partial e^{(t)}}{\partial \theta} \right|_n \quad \text{Equação 3.31}$$

$$e^{(t)} = \frac{1}{2} [y(x^{(t)}) - y(x^{d(t)})]^2 \quad \text{Equação 3.32}$$

$$y(x^{(t)}) = \frac{y_l(x^{(t)}) + y_r(x^{(t)})}{2} \quad \text{Equação 3.33}$$

onde:

$\theta$  = qualquer parâmetro dos antecedentes ou conseqüentes do modelo NFHB-T2.

$\alpha_\theta$  = taxa de aprendizado

$y(x^{(t)})$  = saída precisa do modelo NFHB-T2

$y(x^{d(t)})$  = saída precisa desejada do modelo NFHB-T2

$y_l(x^{(t)})$  = ponto limite esquerdo do conjunto fuzzy *intervalar* do tipo 1 de saída do modelo NFHB-T2.

$y_r(x^{(t)})$  = ponto limite direito do conjunto fuzzy *intervalar* do tipo 1 de saída do modelo NFHB-T2.

Derivando-se o erro médio quadrático  $e^{(t)}$  em relação ao parâmetro a ser ajustado, tem-se:

$$\frac{\partial e^{(t)}}{\partial \theta} = \frac{\partial e^{(t)}}{\partial y(x^{(t)})} \left[ \frac{\partial y(x^{(t)})}{\partial y_l^{(t)}} \cdot \frac{\partial y_l^{(t)}}{\partial \theta} + \frac{\partial y(x^{(t)})}{\partial y_r^{(t)}} \cdot \frac{\partial y_r^{(t)}}{\partial \theta} \right] \quad \text{Equação 3.34}$$

Finalmente, derivando-se:

$$\frac{\partial e^{(t)}}{\partial \theta} = \frac{1}{2} [y(x^{(t)}) - y^{d(t)}] \left[ \frac{\partial y_l^{(t)}}{\partial \theta} + \frac{\partial y_r^{(t)}}{\partial \theta} \right] \quad \text{Equação 3.35}$$

As equações desenvolvidas para a atualização destes pesos encontram-se no apêndice B.

#### d. Passo 4:

Neste passo o **erro total do sistema** é calculado para todo o conjunto de treinamento, de acordo com a expressão do erro médio quadrático dado pela equação 3.36.

$$E_{RMS} = \sqrt{\frac{1}{L} \sum_{n=1}^L (y_n - y_n^d)^2} \quad \text{Equação 3.36}$$

onde:  $L$  = número de padrões no conjunto de treinamento

$y_n, y_n^d$  = valor de saída precisa do sistema NFHB-T2 e valor desejado de saída para o padrão de índice “n”, respectivamente.

Caso esse erro esteja abaixo do mínimo desejado, o processo de aprendizado pára; caso contrário, o processo de aprendizado continua com o passo seguinte.

**e. Passo 5:**

Neste passo, cada bi-partição é avaliada em relação à sua contribuição para o erro total e em relação ao erro mínimo aceitável. **Cada bi-partição com erro inaceitável é separada** para a sua posterior decomposição; a avaliação do erro gerado pelo conjunto de dados que incidem sobre a partição  $ij$  é calculada pela equação 3.37 como segue (apêndice B):

$$E_{RMS} = \sqrt{\frac{1}{L} \sum_{n=1}^L \left( \frac{g_l^{i(n)} g_l^{i j(n)} + g_r^{i(n)} g_r^{i j(n)}}{2} \right) (y_n - y_n^d)^2} \quad \text{Equação 3.37}$$

De forma similar, a avaliação do erro gerado pelo conjunto de dados que incidem sobre a partição  $ijk$  é calculada pela equação 3.38 como segue:

$$E_{RMS} = \sqrt{\frac{1}{L} \sum_{n=1}^L \left( \frac{g_l^{i(n)} g_l^{i j(n)} g_l^{i j k(n)} + g_r^{i(n)} g_r^{i j(n)} g_r^{i j k(n)}}{2} \right) (y_n - y_n^d)^2} \quad \text{Equação 3.38}$$

onde:

$g_l^{i(n)}, g_l^{i j(n)}, g_l^{i j k(n)}, g_r^{i(n)}, g_r^{i j(n)}, g_r^{i j k(n)}$  = pontos limites (esquerdos,  $\underline{f}^i, \underline{f}^{ij}, \underline{f}^{ijk}$ , ou direitos,  $\bar{f}^i, \bar{f}^{ij}, \bar{f}^{ijk}$ ) dos conjuntos fuzzy *intervalares* do tipo 1 dos níveis de disparo das regras do tipo 2 para as bipartições  $i, ij, ijk$  respectivamente, que estão associadas aos  $c_{1_i}, c_{1_{ij}}, c_{1_{ijk}}, c_{2_i}, c_{2_{ij}}, c_{2_{ijk}}$ , para o padrão ‘n’.

**f. Passo 6:**

Para limitar o crescimento indefinido da estrutura do modelo NFHB-T2, utiliza-se um parâmetro de aprendizado denominado **taxa de decomposição** ( $\xi$ ). Este parâmetro é adimensional e atua impedindo que o processo de decomposição (crescimento da árvore binária) seja realizado indefinidamente. Seu valor situa-se, geralmente, entre 0,001 e 0,05. Ele é constantemente comparado, durante o aprendizado, com a população de padrões que incidem sobre um determinado quadrante. Quando a densidade populacional de padrões de um quadrante (razão entre o número de padrões que incidem sobre o quadrante e o número total de padrões) cai abaixo da taxa de decomposição, este quadrante não deve ser decomposto, o que limita o crescimento da estrutura.

**g. Passo 7:**

Neste passo é efetuada a **decomposição das partições separadas**. Para cada bi-partição separada é realizado um processo de decomposição que segue duas vertentes:

- Vertente 1: Nesta vertente utilizam-se somente conseqüentes que são conjuntos fuzzy *intervalares* do tipo 1. Aloca-se um novo nó (nova célula) na estrutura BSP para a bi-partição separada (ela é dividida em 2), utilizando-se conjuntos fuzzy *intervalares* do tipo 1 como conseqüentes. Esta divisão é feita segundo a direção da variável de entrada determinada pela seqüência de variáveis calculada para toda a estrutura NFHB-T2. Deste modo, são geradas duas novas funções de pertinência do tipo 2 que constituirão as duas partições recém-criadas. Os parâmetros antecedentes destas duas novas funções de pertinência serão inicializados segundo o passo 2 do algoritmo de aprendizado caso não exista nenhuma célula com a mesma variável de entrada em algum nível superior na estrutura NFHB-T2; caso contrário, procede-se à inicialização destes parâmetros como descrito na seção 3.4 A Figura 3.12 ilustra este processo.

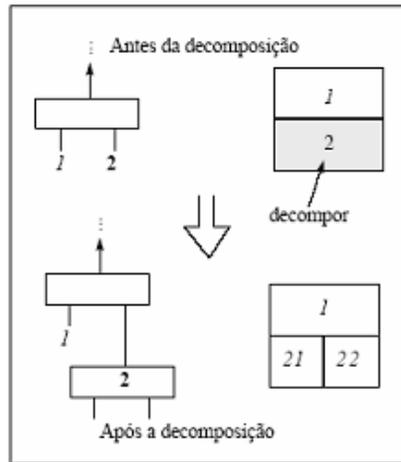


Figura 3.12. Vertente 1 de decomposição BSP-T2

- Vertente 2:** Nesta vertente utilizam-se consequentes mistos (consequentes que são conjuntos fuzzy *intervalares* do tipo 1 e consequentes combinações lineares das entradas com conjuntos fuzzy *intervalares* do tipo 1). Antes de se particionar uma bipartição, verifica-se se o seu consequente é do tipo *conjunto* fuzzy *intervalar* do tipo 1. Caso o seja, substitui-se este por uma combinação linear das entradas com conjuntos fuzzy *intervalares* do tipo 1 consequentes (neste caso não é efetuada a decomposição da bi-partição); caso contrário, efetua-se a decomposição descrita na vertente 1, utilizando-se inicialmente *conjuntos* fuzzy *intervalares* do tipo 1 como consequentes da nova célula alocada. A Figura 3.13 ilustra este processo.

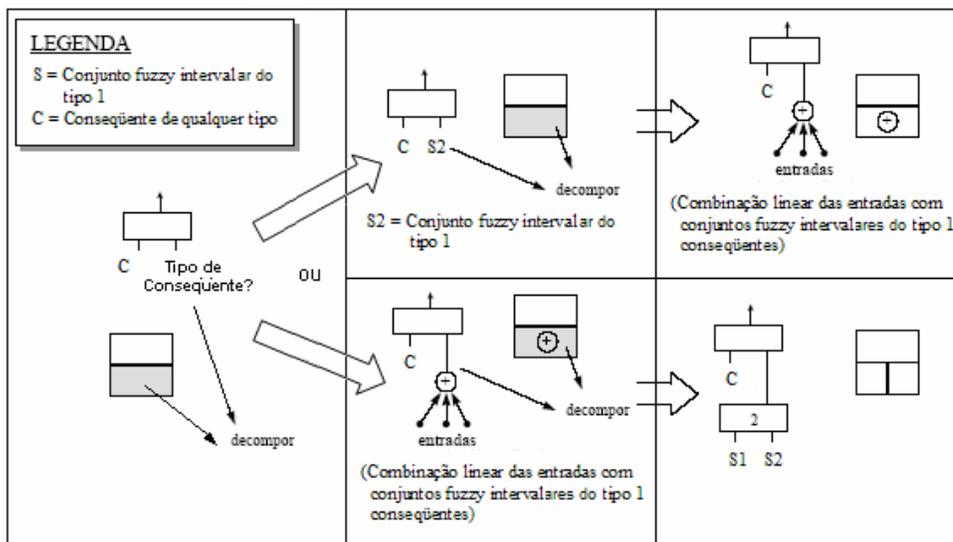


Figura 3.13 - Vertente 2 de decomposição BSP-T2

**h. Passo 8:**

Voltar ao passo “3” para continuar o aprendizado.

**3.6****Estratégias e Métodos de Seleção de Entradas no Modelo NFHB-T2**

O modelo NFHB-T2 necessita de um método de seleção de variáveis para distribuir, em cada nível da árvore BSP, as variáveis de acordo com suas relevâncias, já que durante o processo de aprendizado e geração da estrutura de árvore BSP no modelo NFHB-T2 são extraídas características nessa ordem de relevância (seqüência de variáveis), para serem usadas como entradas para cada nível da árvore BSP.

A seleção da seqüência de variáveis de entrada pode ser efetuada por meio das três opções a seguir:

- **opção 1:**

O método baseia-se na execução de mini-sistemas ANFIS, cada um com uma combinação de 2 entradas do total de características presentes no problema. Os mini-sistemas ANFIS são treinados por um número específico de ciclos e os pares de variáveis são ordenados por ordem crescente de erro no treinamento. A seqüência de pares de entradas obtida por este processo é utilizada na arquitetura NFHB-T2, escolhendo-se uma outra seqüência sem repetição, entre as variáveis de entrada dos primeiros pares listados na ordenação crescente de erro. Esta outra seqüência obtida por este processo é posteriormente usada pelo algoritmo de aprendizado do modelo NFHB-T2. Maiores detalhes deste método podem ser encontrados em [SOUZ99], [JANG96].

- **opção 2:**

O método baseia-se no algoritmo do Estimador de Mínimos Quadrados (LSE) para determinar a importância de cada variável.

Seja um sistema de  $n$  entradas e uma saída; o método LSE calcula a importância da  $i$ -ésima variável de entrada  $x_i$  estimando o  $i$ -ésimo parâmetro  $b_i$  da função  $F$  (equação 3.39) que descreve a variação da variável de saída  $\Delta y$  em relação à variação de cada  $i$ -ésima variável de entrada  $\Delta x_i$  sobre o conjunto completo de dados. A função  $F$  é descrita abaixo [CHUN00], [JANG95],[JANG97]:

$$F = \Delta y = b_1[\Delta x_1] + b_2[\Delta x_2] + b_3[\Delta x_3] + \dots + b_n[\Delta x_n] \quad \text{Equação 3.39}$$

onde os componentes do vetor  $\Delta y$  são obtidos subtraindo-se os valores da variável de saída nos padrões da base de dados em combinações de duas a duas, e os componentes do vetor  $\Delta x_i$  são obtidos subtraindo-se os valores da variável de entrada  $x_i$  nos padrões da base de dados em combinações de duas a duas.

Da Equação 3.39 pode-se dizer que cada parâmetro  $b_i$  representa a importância da  $i$ -ésima variável de entrada em relação à saída no sentido estatístico. O cálculo dos parâmetros  $b_i$  é feito mediante o algoritmo do Estimador de Mínimo Quadrado LSE.

As variáveis de entrada são ordenadas de forma decrescente segundo a sua importância; a seqüência de pares de entradas obtida por este processo é utilizada na arquitetura NFHB-T2. Maiores detalhes deste método podem ser encontrados em [JIME02], [JIME05].

- **opção 3:**

O método baseia-se no cálculo da Efetividade de uma Entrada Singular (SIE). Esta técnica obtém o ranking das entradas calculando o grau de efetividade de cada uma na saída [CAO97], [CHAN90].

Considerando o sistema linear  $y = Gu$  de  $m$  saídas e  $r$  entradas, a importância de cada variável de entrada é obtida expressando-se o vetor de entradas  $u$  do sistema como a soma de duas projeções ortogonais  $u_n$  e  $u_c$ , uma pertencente ao espaço nulo da matriz de transferência  $G$  do sistema e a outra pertencente ao complemento ortogonal do espaço nulo de  $G$ ,

respectivamente. Portanto, a efetividade de cada variável de entrada  $x_i$  é calculada como a norma da projeção da entrada no canal  $i$  sobre o complemento ortogonal do espaço nulo de  $G$ , já que a norma da projeção da entrada no canal  $i$  sobre o espaço nulo de  $G$  não tem efeito na variável de saída.

As variáveis de entrada são ordenadas de forma decrescente segundo a sua importância, a seqüência de pares de entradas obtida por este processo é utilizada na arquitetura NFHB-T2. Maiores detalhes deste método podem ser encontrados em [JIME02], [JIME05].

Por outro lado, foi implementada uma estratégia de aplicação destes três métodos para distribuir as variáveis selecionadas hierarquicamente de acordo com a árvore binária do modelo NFHB-T2. Esta estratégia é chamada de seleção fixa e utiliza a base de dados completa (original) para escolher as características mais relevantes. Ela consiste em determinar a ordem dos atributos (seqüência de variáveis de entrada) mediante uma das três opções descritas acima; posteriormente, durante o processo de aprendizado e construção da arquitetura NFHB-T2, cada uma destas características é escolhida e usada como entrada para cada nível da árvore BSP-T2. A mesma entrada (atributo) é utilizada para todos os nós do mesmo nível. Maiores detalhes destas estratégias podem ser encontradas em [LANA00] e [JIME02].

### 3.7

#### Conclusão

Neste capítulo foi apresentado o modelo neuro-fuzzy hierárquico BSP do tipo 2, com sua célula básica, sua arquitetura, seus tipos de consequentes e seu algoritmo de aprendizado. Foram também descritos os métodos de seleção de características utilizados por este modelo.

O próximo capítulo apresenta o estudo de casos, onde serão discutidos os resultados de desempenho do modelo Neuro-Fuzzy Hierárquico BSP do tipo 2 em problemas *benchmark* e reais de aproximação de funções e previsão.