

Aprendizado Incremental *Fuzzy*: uma Abordagem Baseada na Teoria de Exemplos Generalizados

Maria do Carmo Nicoletti
Flávia Oliveira Santos

Departamento de Computação
Universidade Federal de São Carlos - UFSCar
Caixa Postal 676 13565-905 São Carlos SP - Brazil
{carmo,flavia}@dc.ufscar.br

Abstract: This paper initially discusses the main ideas underneath one of the most representative learning systems based on the generalization of exemplars, the NGE system. Next it presents the *Fuzzy NGE* system which learns by generalizing exemplars described as attribute-fuzzy_value pairs and a crisp class.

Resumo: Este artigo discute, inicialmente, as principais idéias que subsidiam um dos mais representativos sistemas de aprendizado baseado na generalização de exemplares, o algoritmo NGE. E, então, apresenta o sistema *Fuzzy NGE* que aprende generalizando exemplares descritos por pares atributo-valor_*fuzzy* e uma classe *crisp*.

1 Introdução

Aprendizado pode ser caracterizado como um processo multidimensional que, via de regra, ocorre através da aquisição de conhecimento declarativo, do desenvolvimento de habilidades motoras e cognitivas através de instrução e prática, da organização do conhecimento existente em representações mais efetivas, da descoberta de novos fatos e/ou teorias através de observação e experimentação ou, então, através da combinação e/ou composição dessas dimensões [Michalski (1993)].

Devido, principalmente, à sua natureza multidimensional e interdisciplinar, à existência ou não de conhecimento prévio para o aprendizado de novos conhecimentos, à não obediência a uma cronologia predefinida, a diferentes graus de especialização das informações disponíveis, às diferentes possibilidades — bem como às diferentes granularidades — de expressão do conhecimento aprendido, à sua natureza muitas vezes incremental e acumulativa, e em outras vezes não monotônica, aprendizado se evidencia como um processo altamente complexo e ainda não completamente entendido [Nicoletti (1994)].

Entre os vários paradigmas existentes para aprendizado tratado como uma área de pesquisa em Inteligência Artificial, o chamado aprendizado indutivo simbólico é um dos que tem sido mais largamente pesquisado e tem contribuído

efetivamente para a implementação de sistemas de aprendizado de máquina. A partir de exemplos, vários tipos de tarefas podem ser aprendidas como, por exemplo, diagnóstico de doenças, previsão meteorológica, predição do comportamento de novos compostos químicos, predição de propriedades mecânicas de metais com base em algumas de suas propriedades químicas, etc. Técnicas de aprendizado de máquina têm sido utilizadas em todas essas tarefas e em muitas outras, principalmente naquelas passíveis de serem enquadradas como aprendizado de conceitos a partir de exemplos. A aplicação mais óbvia de aprendizado de máquina está na articulação de sistemas de aprendizado de máquina a mecanismos de aquisição de conhecimento, em sistemas baseados em conhecimento.

Para a expressão de qualquer paradigma de aprendizado são necessárias linguagens que descrevam objetos assim como linguagens que descrevam os conceitos aprendidos. Vários formalismos lógicos têm sido usados em sistemas de aprendizado indutivo para a representação de exemplos e conceitos. Em geral, distinguem-se dois tipos de descrição: descrição baseada em atributos e descrição relacional. Em uma descrição baseada em atributos, objetos são descritos em termos de atributos e valores desses atributos. Em uma descrição relacional (também chamada de descrição estrutural) um objeto é descrito em termos de seus componentes e de relações entre eles.

Alguns dos algoritmos/sistemas de aprendizado indutivo que têm sido utilizados com relativo sucesso em

um número razoável de aplicações — tais como o ID3 [Quinlan (1986)] e o AQ [Michalski (1993)] — usam linguagens baseadas em atributos, também conhecidas como linguagens proposicionais, para a representação de exemplos e conceitos. Na realidade, as linguagens de descrição de conceitos utilizadas por esses sistemas para a expressão de suas hipóteses induzidas, que tipicamente são árvores de decisão e regras de produção respectivamente, podem ser tratadas como variantes de linguagens baseadas em atributos.

Com poucas exceções, os sistemas de aprendizado indutivo existentes são não incrementais, i.e., o conjunto de treinamento deve estar disponível ao sistema no começo do processo de aprendizado; a expressão do conceito é induzida considerando todos os exemplos de uma vez. Se por acaso novos exemplos de treinamento se tornam disponíveis após o processo de aprendizado ter começado, a única maneira possível de incorporá-los à expressão do conceito é reiniciar todo o processo de aprendizado com o conjunto de treinamento atualizado.

É importante notar, entretanto, que muitas situações de aprendizado podem ser caracterizadas pela maneira incremental através da qual exemplos de treinamento se tornam disponíveis. Num ambiente incremental, um sistema ideal de aprendizado deve ser capaz de modificar *online* a expressão do conceito à medida que novos exemplos de treinamento se tornam disponíveis. Um novo exemplo pode, potencialmente, provocar um rearranjo da expressão do conceito obtida até então. É fato, entretanto, que restrições com relação à extensão com que tais rearranjos devam acontecer deve ser objeto de estudo e avaliação. Incrementabilidade em um sistema de aprendizado, via de regra, está associada à maior versatilidade e à maior adaptabilidade do sistema. É, conseqüentemente, uma característica desejável, desde que eficientemente incorporada.

As versões incrementais do ID3, conhecidas como ID4 [Schlimmer (1986)] e ID5 [Utgoff (1988)] têm alguns inconvenientes com relação ao desempenho. Aparentemente, a descrição de conceito que usa árvore de decisão como linguagem representacional não é apropriada à abordagem incremental. Essas versões incrementais necessitam armazenar todos os exemplos na memória, com o objetivo de reusá-los quando um rearranjo da expressão corrente do conceito for necessário. Na realidade, o ID4 e ID5 implementam uma pseudo-incrementabilidade.

A busca por outras linguagens representacionais que pudessem expressar conceitos induzidos a partir de exemplos de maneira satisfatória, que fossem adequadas ao modelo

incremental de aprendizado e que permitissem a incorporação de tratamento *fuzzy*, evidenciou a teoria de exemplares como uma alternativa viável a ser explorada, e o algoritmo NGE como uma possível concretização/generalização daquelas idéias.

Este trabalho está organizado da seguinte maneira: na próxima seção são descritas as principais idéias que subsidiam o modelo de aprendizado baseado na generalização de exemplares, através da descrição do NGE (Nested Generalized Exemplar), um de seus principais representantes. A Seção 3 apresenta e discute o algoritmo *Fuzzy NGE* que realiza a generalização de exemplares descritos por atributos com valores *fuzzy*. Nessa seção são também propostos e discutidos o uso da medida de possibilidade *fuzzy* para direcionar a escolha do melhor exemplar a ser generalizado e o uso da operação de união *fuzzy* para gerar a generalização do exemplar. A Seção 4 descreve um exemplo simples de aplicação do *Fuzzy NGE* dando ênfase às operações de escolha e generalização do exemplar. A Seção 5 apresenta as conclusões desse trabalho e linhas de pesquisa para a sua continuidade.

2 Aprendizado Indutivo Incremental via Generalização de Exemplares — o NGE

A teoria de exemplares generalizados NGE, proposta por Salzberg em [Salzberg (1991)], é uma forma de aprendizado indutivo incremental supervisionado a partir de exemplos, que se baseia no modelo de aprendizado humano chamado *exemplar-based learning* [Medin (1978)] e é uma descendente da classificação de padrões *nearest neighbour* — NN [Cover (1967)].

Recentemente tem-se notado um crescente investimento em técnicas que têm como base o algoritmo NN, devido principalmente à natureza incremental deste algoritmo. A teoria NGE é uma teoria de aprendizado indutivo que pode ser vista como uma hibridação da classificação NN e da lógica de primeira ordem limitada a cláusulas de Horn [Wettschereck (1995)].

Na teoria NGE os conceitos induzidos assumem a forma de hiper-retângulos em um espaço n -dimensional. Esse espaço é definido pelos n atributos usados para a descrição dos exemplos. Num paralelo a processos de generalização convencionais que substituem fórmulas simbólicas por fórmulas simbólicas mais gerais, o algoritmo NGE modifica hiper-retângulos, através de seu crescimento e reestruturação (e em algumas situações, sua retração).

A entrada para o sistema NGE é um conjunto de exemplos de treinamento apresentados incrementalmente, cada um descrito como um vetor atributos/valores numéricos e uma classe associada. Os n atributos usados para a descrição dos exemplos definem o espaço Euclidiano no qual o conceito será representado. Atributos podem ter valores *crisp* que variam de 2 a infinito (valores reais) e classes podem ser binárias, discretas ou contínuas.

Na primeira fase (a de inicialização), o NGE aceita os n (número definido pelo usuário) primeiros exemplos que são fornecidos ao sistema como sendo o conjunto inicial de hiper-retângulos (chamados de triviais, já que são pontos), cada um deles considerado um exemplar, a partir dos quais generalizações podem ser feitas. Apenas quando o n -ésimo+1 exemplo se torna disponível é que a fase de inicialização termina e a de treinamento começa. Nessa segunda fase, onde o aprendizado efetivamente ocorre, à medida que novos exemplos vão sendo apresentados, o NGE generaliza os hiper-retângulos do conjunto inicial, dispostos no espaço Euclidiano n -dimensional, expandindo-os (e em alguns situações específicas, encolhendo-os), ao longo de uma ou mais dimensões. Como será visto a seguir, o número de hiper-retângulos no espaço pode aumentar, à medida que o aprendizado prossegue.

A escolha de qual hiper-retângulo generalizar depende de uma métrica de distância. Em um universo onde os atributos têm valores *crisp*, tal métrica é uma distância Euclidiana ponderada, ponto-a-ponto (hiper-retângulo trivial) ou ponto-a-hiper-retângulo.

Para cada novo exemplo de treinamento E , o NGE encontra entre todos os hiper-retângulos (exemplares) existentes até então no espaço n -dimensional, o hiper-retângulo mais próximo, $H_{\text{próximo}1}$, e o segundo mais próximo, $H_{\text{próximo}2}$, do novo exemplo E . Após encontrar os dois exemplares mais próximos, o sistema prossegue com a comparação das classes (a do exemplo com a de cada um dos dois exemplares) que permite escolher qual exemplar será generalizado. Se E e $H_{\text{próximo}1}$ têm a mesma classe, $H_{\text{próximo}1}$ é generalizado, processo onde o hiper-retângulo é expandido de maneira a incluir o novo exemplo. Se, no entanto, a classe de E difere da classe de $H_{\text{próximo}1}$, o sistema compara a classe de E com a classe de $H_{\text{próximo}2}$. Se ambos têm a mesma classe, o NGE vai generalizar $H_{\text{próximo}2}$. Entretanto, se E também não pertencer à mesma classe de $H_{\text{próximo}2}$, esse novo exemplo torna-se um novo exemplar, assumindo a forma de um hiper-retângulo trivial no espaço n -dimensional.

Um mecanismo de ajuste de pesos é adotado pelo NGE como uma forma de reforçar a relevância de atributos e exemplares no processo de classificação. Tal reforço pode ser positivo ou negativo, dependendo da contribuição de cada atributo e de cada exemplar quando da classificação de cada um dos exemplos de treinamento. O peso w_H do exemplar H é inicializado com 1 e à medida que o treinamento prossegue, $w_H=U/C$ vai sendo atualizado, onde U é o número de vezes que H foi usado e C é o número de vezes que H fez a

classificação correta. Como $U \geq C$, w_H é inversamente proporcional à confiança que se tem em H como um “descriptor” do conceito, i.e., quanto maior w_H , menos provável é que o exemplar H desempenhe um papel relevante na expressão final do conceito. Política semelhante é adotada com relação aos atributos f_i , que podem ter o seu peso w_{f_i} incrementado ou decrementado de Δf (taxa global de ajuste do peso do atributo, valor *default* adotado em [Salzberg (1991)] é 0.2).

Um outro aspecto importante que caracteriza o algoritmo NGE é que tal algoritmo permite, durante o processo indutivo, a criação de hiper-retângulos aninhados; o hiper-retângulo mais interno é tratado como uma exceção ao mais externo. Essa forma de representação de exceções é, de certa forma, não muito usual uma vez que não reflete a idéia intuitiva do que seja um conceito.

3 Aprendizado NGE em domínios *fuzzy*

O sistema *Fuzzy NGE* é uma versão do sistema NGE original, baseada no algoritmo proposto em [Nicoletti (1996)], que aceita exemplos de treinamento descritos por atributos que possuem valores *fuzzy* e uma classe *crisp* associada, chamados nesse trabalho de exemplos *fuzzy*. O seu processo de inicialização é o mesmo do NGE original, onde um número de pontos determinado pelo usuário são transformados em exemplares *fuzzy*. Com o término da fase de inicialização, e com o espaço de hipóteses formado pelos n primeiros exemplos, o sistema inicia sua fase de treinamento. Cada novo exemplo *fuzzy* é, então, comparado com cada um dos exemplares *fuzzy* existentes até então no espaço n -dimensional de atributos.

Com o objetivo de escolher o exemplar a ser generalizado, o sistema *Fuzzy NGE* avalia a proximidade do novo exemplo de treinamento com relação a todos os exemplares existentes no espaço e elege os dois mais próximos. Para a avaliação da proximidade entre um exemplo e um exemplar é usada uma medida de distância ponderada, baseada na noção *fuzzy* de possibilidade [Klir (1995)] entre os conjuntos *fuzzy* que descrevem o exemplo e o exemplar, para cada atributo existente.

Para encontrar a proximidade ponderada de um exemplo E a um exemplar H o sistema executa três procedimentos: • encontra a medida de proximidade atributo-a-atributo entre E e H , através da medida de possibilidade entre os conjuntos *fuzzy* associados a cada atributo que descreve E e H ; • calcula a média ponderada da proximidade atributo-a-atributo usando o peso associado a cada atributo; • pondera o valor obtido no passo anterior pelo peso do exemplar H . A proximidade ponderada entre E e H obtida após a execução desses três procedimentos é um número real.

Ao contrário do mecanismo de ponderação do NGE original, o *Fuzzy NGE* assume que quanto maior o peso associado a um atributo, mais relevante é o papel desse

atributo na indução da expressão do conceito. A mesma regra se aplica ao peso de exemplares. Todos os pesos de atributos e exemplares são inicializados com 1. O funcionamento do algoritmo de ajuste de pesos é mostrado na Tabela 1.

	Classe $H_{\text{próximo1}}$	Classe $H_{\text{próximo2}}$	Ajuste de w_{f_i}	Ajuste de w_H
Classe de E	=	—	$H_{\text{próximo1}}$: $w_{f_i} = w_{f_i} + 0.05$ $H_{\text{próximo2}}$: w_{f_i} não altera	$H_{\text{próximo1}}$: $w_H = w_H + 0.05$ $H_{\text{próximo2}}$: w_H não altera
	≠	=	$H_{\text{próximo1}}$: $w_{f_i} = w_{f_i} - 0.05$ $H_{\text{próximo2}}$: $w_{f_i} = w_{f_i} + 0.05$	$H_{\text{próximo1}}$: $w_H = w_H - 0.05$ $H_{\text{próximo2}}$: $w_H = w_H + 0.05$
	≠	≠	$H_{\text{próximo1}}$: $w_{f_i} = w_{f_i} - 0.05$ $H_{\text{próximo2}}$: $w_{f_i} = w_{f_i} - 0.05$	$H_{\text{próximo1}}$: $w_H = w_H - 0.05$ $H_{\text{próximo2}}$: $w_H = w_H - 0.05$

Tabela 1. Mecanismo de ajuste de pesos adotado pelo *Fuzzy NGE*

O cálculo da proximidade ponderada é repetido para todos os exemplares existentes no espaço n-dimensional. Usando as medidas obtidas de proximidade ponderada do novo exemplo a todos os exemplares, o *Fuzzy NGE* define o primeiro e o segundo exemplares mais próximos do novo exemplo. Uma vez identificados esses dois exemplares mais próximos ($H_{\text{próximo1}}$ e $H_{\text{próximo2}}$), segue-se o processo de generalização.

Para escolher qual dos dois será generalizado, o *Fuzzy NGE* utiliza o mesmo procedimento do NGE original. Dependendo do resultado da comparação entre as classes *crisp* de E e de cada um dos dois exemplares escolhidos, uma das situações descritas na Tabela 2 ocorrerá.

	Classe $H_{\text{próximo1}}$	Classe $H_{\text{próximo2}}$	<i>Fuzzy NGE</i>
Classe de E	=	—	Generaliza $H_{\text{próximo1}}$
	≠	=	Generaliza $H_{\text{próximo2}}$
	≠	≠	E torna-se um novo exemplar

Tabela 2. Possíveis situações quando da generalização

Assim como no NGE, o processo de generalização de um exemplar H usando o exemplo E, no *Fuzzy NGE*, pode ser descrito como uma absorção de E por H, a qual é realizada expandindo-se o exemplar H para que este inclua E. O *Fuzzy NGE* generaliza um exemplar através da união dos conjuntos *fuzzy* associados aos atributos usados para descrever E e H.

Os dois operadores implementados no sistema *Fuzzy NGE*, i.e., a medida de possibilidade usada na

determinação de proximidade e a união de conjuntos *fuzzy* usada na generalização de exemplares, são apresentados considerando uma situação de aprendizado onde:

- n é o número de atributos *fuzzy*
- para cada atributo *fuzzy* F_k ($1 \leq k \leq n$), existem i_k conjuntos *fuzzy* associados correspondentes aos i_k possíveis valores *fuzzy* desse atributo, notados por v_{jp_j} , onde $1 \leq j \leq n$ e $1 \leq p_j \leq i_j$
- E e H são descritos respectivamente por:

$$\left[v_{1p'_1}, v_{2p'_2}, v_{3p'_3}, \dots, v_{np'_n} \right]$$

e

$$\left[v_{1p''_1}, v_{2p''_2}, v_{3p''_3}, \dots, v_{np''_n} \right]$$

onde p'_j e p''_j são duas instâncias de $1 \leq p_j \leq i_j$, para $1 \leq j \leq n$

A medida de proximidade atributo-a-atributo entre E e H, definida como a medida de possibilidade, é calculada usando-se:

$$\text{poss}[v_{jp'_j} | v_{jp''_j}] = \max_x [v_{jp'_j} \wedge v_{jp''_j}], \quad j = 1, \dots, n \quad \text{e}$$

a generalização de um exemplar H dado um exemplo E, é calculada através da união dos conjuntos *fuzzy* em E e H, associados a cada atributo, como se segue:

$$\left[v_{1p'_1} \vee v_{1p''_1}, v_{2p'_2} \vee v_{2p''_2}, \dots, v_{jp'_n} \vee v_{jp''_n} \right]$$

4 Aplicando o *Fuzzy NGE*

Considere uma situação simples representada por dois exemplares *fuzzy* (os pontos iniciais), onde cada exemplar é descrito por dois atributos identificados como *Atributo_1* (*altura*) e *Atributo_2* (*peso*) e uma classe *crisp* associada. Seja o *Exemplar_1* = $[v_{11}, v_{21}]$ pertencente à classe 0 e *Exemplar_2* = $[v_{12}, v_{22}]$ pertencente à classe 1.

Os conjuntos de valores *fuzzy* associados ao *Atributo_1* são definidos no conjunto base $X = \{150, 160, 170, 180, 190, 200\}$. Os conjuntos *fuzzy* $v_{11} = \{1/150 + 0.8/160 + 0.2/170 + 0/180 + 0/190 + 0/200\}$ e $v_{12} = \{0/150 + 0/160 + 0.2/170 + 0.5/180 + 1/190 + 1/200\}$ representam, respectivamente, os valores lingüísticos *baixo* e *alto*.

Os conjuntos de valores *fuzzy* associados ao *Atributo_2* são definidos no conjunto base $Y = \{40, 50, 60, 70, 80, 90, 100\}$. Os conjuntos *fuzzy* $v_{21} = \{1/40 + 1/50 + 0.8/60 + 0.5/70 + 0.1/80 + 0/90 + 0/100\}$ e $v_{22} = \{0/40 + 0/50 + 0/60 + 0.5/70 + 1/80 + 0.3/90 + 0/100\}$ representam os valores lingüísticos *leve* e *meio pesado*, respectivamente.

Dado que o espaço de hipóteses tem estes dois exemplares *fuzzy* ($n=2$), suponha que um novo exemplo E seja fornecido, definido como $E = [v_{13}, v_{23}]$, pertencente à classe 1, cujos conjuntos *fuzzy* $v_{13} = \{0/150 + 0/160 +$

$0/170+0.2/180+1/190+1/200$ e $v_{23} = \{0/40+0/50+0/60+0/70+0.1/80+0.8/90+1/100\}$ representam os valores lingüísticos *muito alto* e *pesado*. É importante notar que os valores lingüísticos foram introduzidos para facilitar o entendimento e não apresentam, necessariamente, uma informação fundamental.

O *Fuzzy NGE* calcula, então, a proximidade atributo-a-atributo entre E e cada um dos exemplares:

$$\begin{aligned} \text{Exemplar}_1 & \begin{cases} \text{poss}[v_{11}|v_{13}] = \max_x [v_{11} \wedge v_{13}] = 0 \\ \text{poss}[v_{21}|v_{23}] = \max_x [v_{21} \wedge v_{23}] = 0.1 \end{cases} \\ \text{Exemplar}_2 & \begin{cases} \text{poss}[v_{12}|v_{13}] = \max_x [v_{12} \wedge v_{13}] = 1 \\ \text{poss}[v_{22}|v_{23}] = \max_x [v_{22} \wedge v_{23}] = 0.3 \end{cases} \end{aligned}$$

As Figuras de 1 a 4 mostram, graficamente, a medida de possibilidade entre os conjuntos *fuzzy* para o *Exemplar_1* e o *Exemplar_2* obtidos acima.

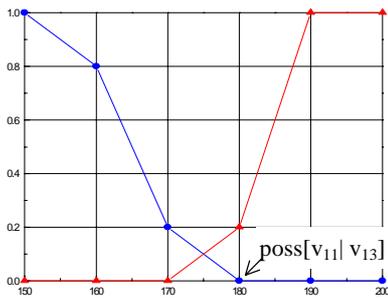


Figura 1. Possibilidade de v_{11} dado v_{13}

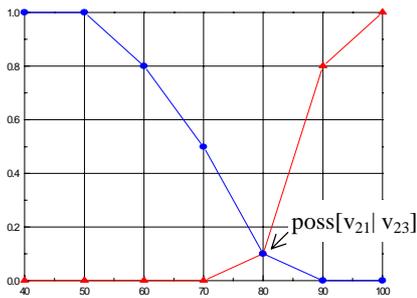


Figura 2. Possibilidade de v_{21} dado v_{23}

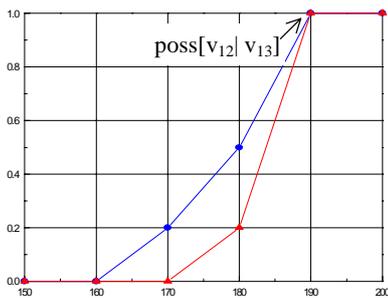


Figura 3. Possibilidade de v_{12} dado v_{13}

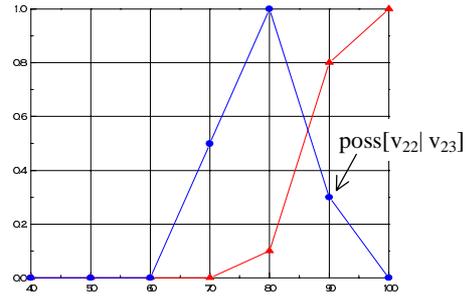


Figura 4. Possibilidade de v_{22} dado v_{23}

O próximo passo é combinar estas proximidades obtidas atributo-a-atributo em um único número. Aplicando os pesos dos atributos mostrados na Tabela 3, têm-se as proximidades ponderadas atributo-a-atributo para o *Exemplar_1* e *Exemplar_2* que são, respectivamente, 0.035 e 0.36, como mostra o cálculo na Tabela 3. Após o cálculo da proximidade ponderada atributo-a-atributo, pondera-se este valor pelo peso do exemplar, obtendo-se os valores para o *Exemplar_1* e o *Exemplar_2*, respectivamente, 0.021 e 0.108, como apresentado na Tabela 4.

	Atributo_1		Atributo_2		Proximidade atributo-a-atributo
	Prox	Peso	Prox	Peso	
<i>Exemplar_1</i>	0	0.2	0.1	0.7	$(0*0.2 + 0.1*0.7) / 2 = \mathbf{0.035}$
<i>Exemplar_2</i>	1	0.6	0.3	0.4	$(1*0.6 + 0.3*0.4) / 2 = \mathbf{0.36}$

Tabela 3. Proximidades ponderadas atributo-a-atributo entre E e cada um dos exemplares

	Peso do exemplar	Proximidade ponderada por exemplar
<i>Exemplar_1</i>	0.6	$0.035 * 0.6 = \mathbf{0.021}$
<i>Exemplar_2</i>	0.3	$0.36 * 0.3 = \mathbf{0.108}$

Tabela 4. Proximidades ponderadas por atributos e exemplares

Desta forma, o *Exemplar_2* é escolhido como o exemplar mais próximo do novo exemplo E pois possui a maior proximidade ponderada (0.108) e pertence à mesma classe de E (classe 1). Este exemplar será, então, generalizado, tornando-se $[v_{12}, v_{24}]$ (como mostra a Figura 5), onde v_{12} representa o valor *alto*. Nota-se que a versão *Fuzzy NGE* permite que “novos valores de atributos” sejam dinamicamente criados durante o processo de aprendizado, ao se generalizarem exemplares via união *fuzzy*. É importante notar que o conjunto *fuzzy* $v_{24} = \{0/40+0/50+0/60+0.5/70+1/80+0.8/90+1/100\}$ foi criado a partir da união dos conjuntos v_{22} e v_{23} e não tem associado a ele um valor lingüístico.

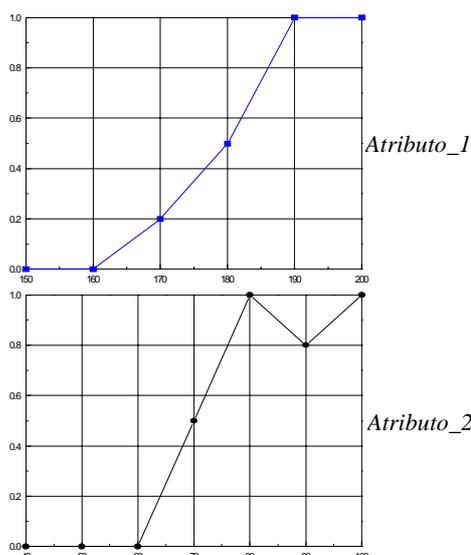


Figura 5. *Exemplar_2* generalizado através da união dos conjuntos *fuzzy* associados aos seus atributos (*Atributo_1* e *Atributo_2*)

Ao final da fase de aprendizado, todos os exemplares *fuzzy* existentes constituem a expressão do conceito. Num sistema de aprendizado de máquina, ao final da fase de aprendizado segue-se a fase de classificação, quando o conceito aprendido é usado para classificar novos exemplos. Na fase de classificação, o *Fuzzy NGE* classifica um novo exemplo medindo sua proximidade a cada um dos exemplares (hiper-retângulos triviais ou não) que compõem a expressão do conceito aprendido na fase anterior e assume como classe daquele exemplo a classe do exemplar mais próximo.

5 Conclusões

É fato que algoritmos de aprendizado indutivo que sejam incrementais são os candidatos mais indicados a implementarem o processo de aprendizado em domínios reais. A característica principal desses algoritmos de “adaptar” a expressão do conceito quando novos exemplos de treinamento se tornam disponíveis, de maneira a representá-los também, permite que o aprendizado possa ser implementado como um processo dinâmico. Esse trabalho descreve as principais características de um desses algoritmos incrementais, o *NGE*, e apresenta e discute o sistema *Fuzzy NGE*, em fase de implementação, que realiza aprendizado incremental baseado em exemplares em domínios *fuzzy*. É intenção como continuidade deste trabalho, investigar a potencialidade de outros possíveis operadores *fuzzy* para a implementação da proximidade e generalização. Além disso, é objetivo investir em uma modificação do *Fuzzy NGE* para que exemplos de treinamento possam também ser descritos por classes *fuzzy*. Pretende-se com isso, além de tornar

o sistema mais geral, que cada exemplar *fuzzy* que define o conceito seja transformado em uma regra de produção *fuzzy*, com n entradas (dimensão do espaço) e uma saída e , então, utilizar um método de inferência *fuzzy* disponível (como o método Mamdani) [Klir (1995)]. Como a saída da regra de produção, pelo método Mamdani, deve ser um conjunto *fuzzy* e, neste caso, a saída é a própria classe, esta não pode ter um valor *crisp*.

Agradecimentos: À Fapesp pelo apoio financeiro Proc. Nº 1996/10119-2 e Proc. Nº 95/9392-9.

Referências

- T. Cover, P. Hart. Nearest Neighbour Pattern Classification. *IEEE Transactions on Information Theory* 13, 1967, pp 21-27.
- G. J. Klir, B. Yuan. *Fuzzy Sets and Fuzzy Logic: theory and applications*. Prentice-Hall International, 1995.
- D. Medin, M. Schaffer. Context Theory of Classification Learning. *Psychological Review* 85, 1978, pp 207-238.
- R. S. Michalski, G. Tecuci. Multistrategy Learning. Tutorial T15, *IJCAI 1993*, 1993.
- M. C. Nicoletti. Ampliando os Limites do Aprendizado Indutivo de Máquina através das Abordagens Construtiva e Relacional. *Tese de Doutorado*, IFSC/USP, 1994, 195 págs.
- M. C. Nicoletti, F. O. Santos. Learning Fuzzy Exemplars through a Fuzzified Nested Generalized Exemplar Theory. *Proceedings of the European Workshop on Fuzzy Decision Analysis for Management, Planning and Optimization*, Dortmund, Germany, May 1996, pp 140-145.
- J. R. Quinlan. Induction of Decision Trees. *Machine Learning* 1, 1986, pp 81-106.
- S. L. Salzberg. A Nearest Hyperrectangle Learning Method. *Machine Learning* 6, 1991, pp 251-276.
- J. C. Schlimmer, D. Fisher. A Case Study of Incremental Concept Induction. *Proceedings of the Fifth National Conference on Artificial Intelligence*, Morgan Kaufmann Publishers, 1986, pp 496-501.
- P. E. Utgoff. ID5: An Incremental ID3. *Proceedings of the Fifth National Conference on Machine Learning*, University of Michigan, June 1988, pp 107-120.
- D. Wettschereck, T. G. Ditterich. An Experimental Comparison of the Nearest-Neighbour and Nearest-Hyperrectangle Algorithms. *Machine Learning* 19, 1995, pp 5-27.