Estudos teóricos e práticos sobre sistemas fuzzy baseados em regras e métodos de ativação

Renato Lopes Moura

Orientador: Peter Sussner

Universidade Estadual de Campinas

ra163050@ime.unicamp.br

14 de setembro de 2017

Organização do trabalho

Organização do trabalho

Introdução

Este trabalho pretende fazer um estudo sobre sistemas de inferência *fuzzy* observando os modelos conjuntivo e implicativo de equações relacionais *fuzzy* e diferentes regras de composição de relações *fuzzy*.

Também serão estudados diferentes métodos de ativação de regras nos sistemas de inferência, baseado nos trabalhos de Peter Sussner e Estevão Laureano em Memórias Associativas Morfológicas. Mais especificamente S-FAMs e E-FAMs, que utilizam medidas de subsethood e equivalência respectivamente.

Os axiomas de Moser-Navara fornecem uma fundamentação matemática para avaliar a coerência de um sistema de inferência. Posteriormente este trabalho foi expandido por Martin Stepnicka *et al.* para avaliar outros modelos de inferência além do tradicional proposto por Mamdani-Assilian. Aqui, faremos uma revisão teórica destes resultados e analisaremos o comportamento dos sistemas de inferência em problemas práticos de classificação e regressão.

Organização do trabalho

Operações fuzzy

Estendendo os operadores lógicos booleanos para o domínio *fuzzy* podemos obter uma família de operações entre conjuntos *fuzzy* que modelam as operações de conjunção (AND) e disjunção (OR)[Pedrycz e Gomide, 2007][Klir e Yuan, 1995].

Uma conjunção *fuzzy* é uma função $C : [0,1] \times [0,1] \to [0,1]$ tal que $C(0,0) = C(1,0) = C(0,1) = 0$ e $C(1,1) = 1$.

Se além disso, esta função for <u>comutativa</u>, <u>associativa</u> e satisfizer $C(1,a) = a$ para todo $a \in [0,1]$, então C é chamada <u>t-norma</u>.

Da mesma forma, pode ser definida uma disjunção *fuzzy* como uma função $D : [0,1] \times [0,1] \to [0,1]$ tal que $D(0,0) = 0$ e $D(1,0) = D(0,1) = D(1,1) = 1$. Similarmente, se D for uma função <u>comutativa</u>, <u>associativa</u> e satisfizer $D(0,a) = a$ para todo $a \in [0,1]$, então D é chamada <u>t-conorma</u> ou <u>s-norma</u>.

Operações fuzzy

Outro operador que pode ser estendido da lógica booleana é o operador de implicação. Uma implicação *fuzzy* é uma função $I : [0,1] \times [0,1] \to [0,1]$ onde $I(0,0) = I(0,1) = I(1,1) = 1$ e $I(1,0) = 0$.

Um tipo especial de implicação pode ser construído utilizando-se t-normas que sejam contínuas à esquerda, ou seja:

$$\lim_{n\to\infty} T(x_n, y) = T(\lim_{n\to\infty} x_n, y)$$

Então é possível definir a implicação associada à t-norma T como:

$$I_T(a,b) = \bigvee\{c \in [0,1] | T(a,c) \leq b\}$$

Esta implicação I_T é chamada <u>implicação residual *fuzzy*</u> (ou R-implicação).

Relações fuzzy

Uma relação binária pode ser vista como a associação (ou a ausência dela) entre elementos de dois conjuntos. Uma relação binária *fuzzy* é uma extensão deste conceito para conjuntos *fuzzy*.

Formalmente, segundo [Barros e Bassanezi, 2010] temos:

Definição

Uma relação binária *fuzzy* sobre $\mathcal{F}(X)$ é uma relação $\mathcal{R}$ sobre $\mathcal{F}(X) \times \mathcal{F}(X)$ definida por uma função de pertinência:

$$\mathcal{R} : \mathcal{F}(X) \times \mathcal{F}(X) \rightarrow [0, 1]$$

Medida de Subsethood

Um exemplo importante de relação binária *fuzzy* é a medida de *subsethood*. Uma medida de *subsethood* avalia o grau de "inclusão" de um conjunto *fuzzy* em outro, em outras palavras é uma generalização do operador $\subseteq$ da teoria de conjuntos clássica.

Neste trabalho adotaremos a seguinte definição de medida de *subsethood* proposta por [Fan *et al.*, 1999]:

Definição

Seja uma função $S : \mathcal{F}(X) \times \mathcal{F}(X) \to [0, 1]$. $S(A, B)$ é uma medida de *subsethood* se para $A, B, C \in \mathcal{F}(X)$ temos:

1. $S(A, B) = 1 \Leftrightarrow A \subseteq B$
2. $S(X, \emptyset) = 0$
3. Se $A \subseteq B \subseteq C$, então $S(C, A) \leq S(B, A)$ e $S(C, A) \leq S(C, B)$

Medida de Similaridade

Outra importante relação entre conjuntos *fuzzy* é a medida de similaridade, que generaliza a igualdade de conjuntos da teoria de conjuntos clássica. Neste trabalho utilizaremos a definição mais comum de medida de similaridade encontrada na literatura (que também pode ser encontrada como a definição de uma medida de equivalência).

Definição

Uma função $SM : \mathcal{F}(X) \times \mathcal{F}(X) \to [0, 1]$ é dita uma medida de similaridade se SM satisfaz as seguintes propriedades para todos $A, B, C \in \mathcal{F}(X)$:

1. $SM(A, B) = SM(B, A)$
2. $SM(X, \emptyset) = 0$
3. $SM(A, A) = 1$
4. Se $A \subseteq B \subseteq C$, então $SM(A, B) \geq SM(A, C)$ e $SM(B, C) \geq SM(A, C)$

Modelos de relações fuzzy

Modelo conjuntivo

$$R_*^{\vee}(x, y) = \bigvee (A_i(x) * B_i(y))$$

onde $*$ é uma t-norma contínua à esquerda.

Modelo implicativo

$$R_{\rightarrow}^{\wedge}(x, y) = \bigwedge (A_i(x) \rightarrow B_i(y))$$

onde $\rightarrow$ é uma implicação residual.

Composições fuzzy

Composições fuzzy

Regra de Composição de Inferência (composição sup-$*$) [Zadeh, 1973]

$$B'(y) = (A' \circ R)(y) = \bigvee (A'(x) * R(x,y))$$

Subproduto de Bandler-Kohout (composição inf-$\longrightarrow$) [W.Bandler, L.J.Kohout, 70s]

$$B'(y) = (A' \lhd R)(y) = \bigwedge (A'(x) \to R(x,y))$$

Composições fuzzy

Composição inf-$\rightarrow$ (Subproduto de Bandler-Kohout)

A composição inf-$\rightarrow$ foi sugerida como um método de inferência por W. Pedrycz em [Pedrycz, 1985].

Entretanto, devido à popularidade alcançada pelo modelo de Mamdani-Assilian em aplicações práticas, ela acabou sendo pouco estudada.

Mais recentemente, esta composição teve suas propriedades estudadas em [Stepnicka e Jayaram, 2010] e se mostrou tão adequada quanto a Regra de Composição de Inferência de Zadeh, desde que corretamente combinada com o modelo de relações *fuzzy* utilizado.

Organização do trabalho

Modus Ponens

Na lógica proposicional, o *modus ponens* é uma forma de raciocínio que utiliza como base uma regra de inferência para gerar conclusões a partir de fatos. Ele pode ser representado da seguinte forma:

Regra:	"Se x é A, então y é B"
Fato:	"x é A"
Conclusão:	"y é B"

Modus Ponens Generalizado

Em contrapartida, o *modus ponens generalizado* utiliza as ferramentas da teoria de conjuntos *fuzzy* para aumentar o poder de inferência do *modus ponens* tradicional dando origem ao chamado raciocínio aproximado. Ele pode ser representado da seguinte forma:

Regra:	"Se x é A, então y é B"
Fato:	"x é A' "
Conclusão:	"y é B' "

Interpretação da base de regras

Considerando o *modus ponens generalizado* onde cada regra é vista como uma relação entre conjuntos fuzzy, cabem duas interpretações para uma base de regras: uma para o modelo conjuntivo e uma para o modelo implicativo de relações *fuzzy*.

Relembrando o modelo conjuntivo de relações *fuzzy* é dado por:

$$R_*^{\vee}(x,y) = \bigvee (A_i(x) * B_i(y))$$

Desta forma a base de regras pode ser interpretada como:

$$\begin{array}{ll} R1\textbf{:} & "x \text{ é } A_1 \textbf{ e } y \text{ é } B_1" \\ & \textbf{ou} \\ & \ldots \\ & \textbf{ou} \\ Rn\textbf{:} & "x \text{ é } A_1 \textbf{ e } y \text{ é } B_n" \end{array}$$

Interpretação da base de regras

Citando [Dubois *et al.*, 1997]:

> *It seems that fuzzy rules modeled this way are not viewed as constraints, but are considered to be pieces of data. Then, the maximum expresses accumulation of data.*

Que em tradução livre quer dizer:

> *Parece que as regras fuzzy modeladas desta maneira não são vistas como restrições, mas sim como pedaços de informação. Então, a agregação pelo máximo expressa acumulação de informação.*

Interpretação da base de regras

Por outro lado, o modelo implicativo de relações *fuzzy* é dado por:

$$R^{\wedge}_{\rightarrow}(x, y) = \bigwedge(A_i(x) \rightarrow B_i(y))$$

Desta forma a base de regras pode ser interpretada como:

$R1$**:** "Se x é A_1 **então** y é B_1"
e
...
e
Rn**:** "Se x é A_1 **então** y é B_n"

Interpretação da base de regras

Novamente citando [Dubois *et al.*, 1997]:

> *In the above view, each piece of information (fuzzy rule) is viewed as a constraint. This view naturally leads to a conjunctive way of merging the individual pieces of information since the more information, the more constraints and the less possible values to satisfy them.*

Que em tradução livre quer dizer:

> *Nesta visão, cada pedaço de informação (regra fuzzy) é vista como uma restrição. Esta visão naturalmente leva a uma maneira conjuntiva de agregação dos pedaços individuais de informação uma vez que quanto mais informação, mais restrições e menos valores possíveis que as satisfazem.*

Organização do trabalho

Métodos de ativação de regras fuzzy

Recentemente, surgiram alguns trabalhos sugerindo métodos de ativação de regras *fuzzy* baseados em medidas de Subsethood e Similaridade [Esmi *et al.*, 2015] [Esmi *et al.*, 2016] [Wagner *et al.*, 2016] [Kaburlasos e Kehagias, 2014].

A discussão de tais métodos faz mais sentido no contexto em que as entradas do sistema de inferência são conjuntos *fuzzy*. Neste trabalho, o foco será no estudo deste caso em particular, deixando o caso de entradas *crisp* não fuzzificadas para trabalhos futuros.

Métodos de ativação de regras fuzzy

Além disso, o caso *crisp* independe do modelo de inferência adotado pois para um dado $x' \in X$ e A' um singleton em x' temos:

$$(A' \circ R)(y) = (A' \lhd R)(y) = R(x', y), y \in Y$$

Método de ativação tradicional

Relembrando o esquema clássico de um sistema de inferência *fuzzy*:

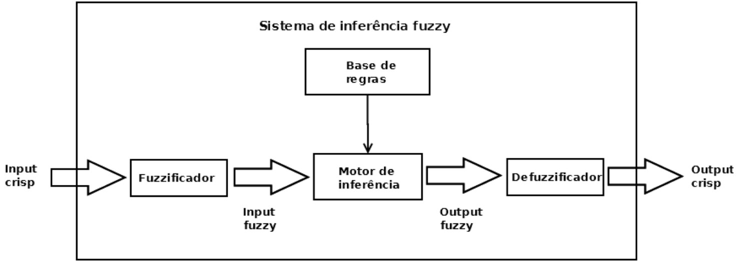

Figura: Sistema de inferência *fuzzy*.

Método de ativação tradicional

Dada uma entrada *crisp* $x' \in X$ no sistema de inferência, temos a seguinte sequência de passos:

1. A entrada x' é fuzzificada para um conjunto *fuzzy* $A'(x) \in \mathcal{F}(X)$ com $A'(x') = 1$
2. Para cada regra R_i com antecedente A_i é calculado $\alpha_i = \bigvee_{x \in X} [A'(x) \wedge A_i(x)]$
3. O valor de α_i representa o grau de ativação da regra R_i

Obs: para simplificar a notação, os passos descritos acima descrevem um cenário onde as regras possuem um antecedente e um consequente apenas.

Casos onde a ativação por similaridade é superior

Como é possível observar na figura a seguir, temos um caso onde entradas claramente diferentes produzem o mesmo grau de ativação de uma determinada regra no sistema de inferência *fuzzy* de Mamdani-Assilian.

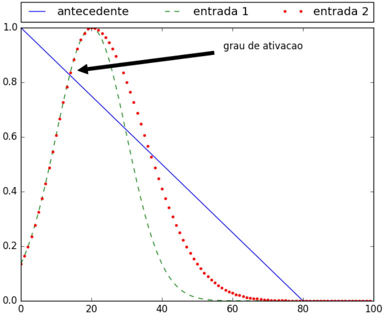

Figura: Problema de ativação de regras no sistema de Mamdani-Assilian.

Ativação de regras fuzzy utilizando medidas de similaridade

Este problema foi a principal motivação para a busca por outras abordagens, como a ativação por centróide [Pourabdollah *et al.*, 2015] [Pourabdollah *et al.*, 2015] e por similaridade [Wagner *et al.*, 2016].

Desta forma, a abordagem de ativação por uma medida de similaridade $SM(A, B) \in \mathcal{F}(X) \times \mathcal{F}(X)$ pode ser expressa matematicamente por:

$$B'(y) = \bigvee(SM(A'(x), A_i(x)) \wedge B_i(y))$$

ou

$$B'(y) = \bigwedge(SM(A'(x), A_i(x)) \rightarrow B_i(y))$$

dependendo da composição de inferência escolhida.

A ativação de regras por medida de similaridade apesar de ser intuitivamente melhor do que o modelo mais comumente utilizado, ainda apresenta seus problemas, como é o caso exposto na figura a seguir.

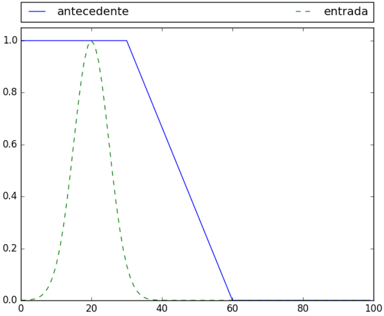

Figura: Problema de ativação de regras por similaridade.

Ativação de regras fuzzy utilizando medidas de subsethood

Similar à abordagem de ativação por uma medida de similaridade, a ativação por uma medida de subsethood $S(A, B) \in \mathcal{F}(X) \times \mathcal{F}(X)$ pode ser expressa matematicamente por:

$$B'(y) = \bigvee (S(A'(x), A_i(x)) \wedge B_i(y))$$

ou

$$B'(y) = \bigwedge (S(A'(x), A_i(x)) \rightarrow B_i(y))$$

dependendo da composição de inferência escolhida.

Organização do trabalho

Axiomas de Moser-Navara

Como ressaltado em [Stepnicka e Jayaram, 2010], a escolha dos objetos envolvidos (modelo de relação *fuzzy*, conjuntos *fuzzy*, inferência, operações matemáticas) não deve ser arbitrária. Tendo em vista os resultados empíricos obtidos ao longo dos anos, deve haver uma axiomatização natural para esta combinação.

Tal axiomatização foi proposta em [Moser e Navara, 2002].

Axiomas de Moser-Navara

1. Para todo $i \in 1, \ldots, n$
$$A_i \circ R_*^{\vee} = B_i$$

2. Para cada *input* normal $A' \in \mathcal{F}(X)$ existe um índice i tal que
$$A' \circ R_*^{\vee} \not\subseteq B_i$$

3. O *output* $A' \circ R_*^{\vee}$ pertence à união dos consequentes B_i de regras ativadas
$$A' \circ R_*^{\vee} \subseteq \bigcup_{i \in F} B_i$$
$$F = \{i | Supp(A_i) \cap Supp(A') \neq \emptyset\}, (B_i \cup B_j)(y) = B_i(y) \vee B_j(y)$$

Interpretação dos axiomas

Uma interpretação literal destes axiomas dada em [Stepnicka e Jayaram, 2010] é a seguinte:

1. A relação fuzzy $R(x, y)$ deve apresentar *recall* (preservação do *modus ponens* ou interpolação *fuzzy*)
2. Para *inputs* normais, o sistema deve apresentar *outputs* significativos com informação não-trivial
3. Se o *input* do sistema está contido no universo dos antecedentes, então deve existir um *output* não-nulo para ele

Axiomas de Moser-Navara

Com base nesta interpretação, foram propostos axiomas similares para o caso da composição inf-$\rightarrow$ [Stepnicka e Jayaram, 2010]

1. Para todo $i \in 1, \ldots, n$
$$A_i \lhd R^{\wedge}_{\rightarrow} = B_i$$

2. Para cada *input* normal $A' \in \mathcal{F}(X)$ existe um índice i tal que
$$A' \lhd R^{\wedge}_{\rightarrow} \not\supseteq B_i$$

3. O *output* $A' \lhd R^{\wedge}_{\rightarrow}$ contêm a interseção dos consequentes B_i de regras ativadas
$$A' \lhd R^{\wedge}_{\rightarrow} \supseteq \bigcap_{i \in F} B_i$$
$$F = \{i | Supp(A_i) \cap Supp(A') \neq \emptyset\}, (B_i \cap B_j)(y) = B_i(y) \wedge B_j(y)$$

Combinação de $R^\vee$ e $\circ$

Relembrando um resultado de [Klawonn, 2000], temos o seguinte teorema:

Teorema

Sejam A_i, para $i = 1, ..., n$ conjuntos *fuzzy* normais. Então $R^\vee$ é uma solução de

$$A_i \circ R = B_i, i = 1, ..., n$$

se, e somente se, a seguinte condição é válida para $i, j \in \{1, ..., n\}$ arbitrários

$$\bigvee_{x \in X} (A_i(x) * A_j(x)) \leq \bigwedge_{y \in Y} (B_i(y) \leftrightarrow B_j(y))$$

Combinação de $R^{\vee}$ e $\circ$

Proposição [Moser e Navara, 2002]

Seja $*$ uma t-norma sem divisores de zero. Seja todos $A_i, i = 1, ..., n$ contínuos e normais e sejam $B_i, i = 1, ..., n$ conjuntos *fuzzy* com **suportes mutuamente diferentes**. Então a combinação de $R^{\vee}$ e $\circ$ não satisfaz os axiomas 1 e 2 simultaneamente.

Suportes mutuamente diferentes de $B_i \Rightarrow \bigwedge_{y \in Y} (B_i(y) \leftrightarrow B_j(y)) = 0$

$\bigvee_{x \in X} (A_i(x) * A_j(x))$ deve ser igual a 0 $\Rightarrow A_i$ devem ser disjuntos!

Combinação de $R^\wedge$ e $\lhd$

Um resultado semelhante para a combinação de $R^\wedge$ e $\lhd$ foi proposto em [Stepnicka e Mandal, 2015].

Proposição [Stepnicka e Mandal, 2015]

Seja $*$ uma t-norma sem divisores de zero. Seja todos $A_i, i = 1, ..., n$ contínuos e normais e sejam $B_i, i = 1, ..., n$ conjuntos *fuzzy* com **suportes mutuamente diferentes**. Então a combinação de $R^\wedge$ e $\lhd$ não satisfaz os axiomas 1 e 2 simultaneamente.

A demonstração segue a mesma linha uma vez que a idéia original dos axiomas não foi modificada para o caso da composição inf-$\rightarrow$.

Consequências

Diante dos problemas apresentados, as combinações mais seguras de modelos de relações *fuzzy* e métodos de inferência são:

- $R^{\wedge}$ (modelo implicativo) e $\circ$ (composição sup-t)
- $R^{\vee}$ (modelo conjuntivo) e $\triangleleft$ (composição inf-$\rightarrow$)

Apesar disso, essas combinações são pouco utilizadas.

Organização do trabalho

Previsão de séries temporais

Sensor de estacionamento

Organização do trabalho

Conclusão

Referências

Martin Stepnicka, Bernard de Baets and Lenka Noskova (2010)
Arithmetic Fuzzy Models
IEEE Transactions on Fuzzy Systems Vol. 18, No. 6

Martin Stepnicka and Balasubramaniam Jayaram (2010)
On the Suitability of the Bandler-Kohout Subproduct as an Inference Mechanism
IEEE Transactions on Fuzzy Systems Vol. 18, No. 2

L.C.Barros and R.C.Bassanezi (2010)
Tópicos de Lógica Fuzzy e Biomatemática
Campinas, SP, Brasil 2a ed.

D.Dubois, H.Prade and L.Ughetto (1997)
Checking the coherence and redundancy of fuzzy knowledge bases
IEEE Transactions on Fuzzy Systems Vol. 5, No. 3

Referências

G.J.Klir and B.Yuan (1995)
Fuzzy Sets and Fuzzy Logic: Theory and Applications
Upper Saddle River, N.Y. Prentice Hall

W.Pedrycz and F.Gomide (2007)
Fuzzy Systems Engineering: Towards Human-Centric Computing
New York Wiley, IEEE Press

J.Fan, W.Xie and J.Pei (1999)
Subsethood measure: new definitions
Fuzzy Sets and Systems 106.2, pp. 201-209

C.Wagner, A.Pourabdollah, J.McCulloch, R.John and J.M.Garibaldi (2016)
A Similarity-based Inference Engine for Non-Singleton Fuzzy Logic Systems
IEEE International Conference on Fuzzy Systems 2016 Vancouver, Canada

Referências

V.G.Kaburlasos and A.Kehagias (2014)
Fuzzy Inference System (FIS) Extensions Based on Lattice Theory
IEEE Transactions on Fuzzy Systems vol. 22, no. 3, pp. 531-546

E.Esmi, P.Sussner, H.B.Sola and J.Fernandez (2015)
θ-Fuzzy Associative Memories (θ-FAMs)
IEEE Transactions on Fuzzy Systems vol. 23, pp. 1-1

E.Esmi, P.Sussner and S.Sandri (2016)
Tunable Equivalence Fuzzy Associative Memories
Fuzzy Sets and Systems vol. 292, pp. 242-260

A.Pourabdollah, C.Wagner and J.Aladi (2015)
Improved uncertainty capture for non-singleton fuzzy systems
IEEE Transactions on Fuzzy Systems (accepted for)

Referências

A.Pourabdollah, C.Wagner and J.Aladi (2015)
Changes under the hood-a new type of non-singleton fuzzy logic system
IEEE Conference on Fuzzy Systems 2015 pp. 1-8

Bernhard Moser and Mirko Navara (2002)
Fuzzy controllers with conditionally firing rules
IEEE Transactions on Fuzzy Systems 2002

L.A.Zadeh (1973)
Outline of a New Approach to the Analysis of Complex Systems and Decision Processes
IEEE TSMC 1973

M.Stepnicka and S.Mandal (2015)
Conditionally Firing Implicative Rules
Proc. of IFSA-EUSFLAT 2015 Conference Neuveden, 2015, pp. 42-48

Referências

W.Pedrycz (1985)
Application of fuzzy relational equations for methods of reasoning in presence of fuzzy data
FSS 1985

F.Klawonn (2000)
Fuzzy Points, Fuzzy Relations and Fuzzy Functions
Discovering the world with fuzzy logic Heidelberg, Germany: Springer-Verlag/Physica-Verlag, 2000, pp. 431-453