UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E COMPUTAÇÃO CIENTÍFICA
DEPARTAMENTO DE MATEMÁTICA APLICADA

GISLAINE DE OLIVEIRA QUEIROS

Identificaçao de Anomalia em Sistemas de Base de Regras Fuzzy

Monografia apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos para obtenção de créditos na disciplina Projeto Supervisionado, sob a orientação do(a) Prof. Estevão Esmi Laureano.

Campinas
2019

Resumo

Sistemas baseados em regras fuzzy (SBRF) consistem em sistemas de entrada-saída que utilizam informações prévias descritas através de um conjunto de regras implicativas (do tipo "Se-Então"), fornecidas por especialistas e/ou conjunto de dados, para associar novas entradas à saídas desejadas. Aplicações de SBRFs podem ser encontradas em diversas áreas do conhecimento como medicina, biologia, química, engenharia, etc.

O comportamento e a acurácia do SBRF estão intimamente ligados a qualidade do conjunto de regras disponível. Se as base de regras contém inconsistencias, tais como contradições e/ou incompletude do domínio, então, isso pode comprometer o resultado fornecido por um sistema baseado nessas regras. Logo, é importante se utilizar de recursos de verificação de integridade das base de regras fuzzy para verificar se elas expressam corretamente o fenômeno de interesse.

As bases de regras, em geral, podem ser classificadas como conjuntivas ou implicativas. Na literatura temos diversos métodos para identificar anomalias em base de regras fuzzy, porém, poucos deles consideram regras conjuntivas que são as utilizadas pelos bem conhecidos SBRF do tipo Mamdani. Além disso, nehum método de identificação de anomalias que conhecemos considera a regras fuzzy ponderadas. Esse projeto objetiva-se a estudar o método proposto por Scarpelli $et\ al.$, que é baseado em medidas de similaridade fuzzy para identificação de regras conjuntivas possivelmente contraditórias, além de propor uma tratativa para estimar o limitante β que indica o grau de contradição permitido em um sistema de base de regras fuzzy.

Abstract

Fuzzy rule-based systems (FRBSs) are of input-output systems that use prior information described through a set of ("If-Then") rules, provided by experts and / or datasets, to associate new inputs with the desired outputs. Applications of FRBSs can be found in so many areas of knowledge such as medicine, biology, chemistry, engineering, etc.

The behavior and accuracy of the FRBSs are closely connected to the quality of the available rules. If the rule base contains inconsistencies, such as contradictions and/or incompleteness, then this may compromise the result provided by a system based on these rules. Therefore, it is important to use methods to check the integrity of the fuzzy rule base taking into account the phenomenon of interest.

Rule bases can generally be classified as conjunctive or implicative. In the literature, there are many methods for identifying anomalies on fuzzy rules, but few of them consider conjunctive rules such as those used by Mamdani FRBSs. This project focus on the method proposed by Scarpelli it et al. which is based on fuzzy similarity measures to identify possibly contradictory conjunctive rules. In addition, we propose a treatment for estimating the threshold β that indicates the degree of contradiction allowed in a fuzzy rule-based system.

Conteúdo

1 Introdução

A teoria de conjuntos fuzzy é uma teoria matemática introduzida por Lofti Zadeh para modelar e processar conceitos e objetos cujas fronteiras são difusas e/ou incertas [13]. Uma das razões da difusão e popularização dessa teoria em diversas áreas do conhecimento está, em parte, ligado ao desevenvolvimento dos chamados sistemas baseados em regras (SBRF). Um SBRF utiliza operadores da lógica fuzzy e um conjunto de informações sobre um determinado fenômeno descritas por regras do tipo "Se-Então" para elaborar um sistema de entrada-saída associado ao problema em questão [9]. A utilização de regras "Se-Então" facilita a interdisciplinaridade e interpretabilidade do sistema obtido, devido sua proximidade com a linguagem natural. De fato, muitas tarefas e fenômenos podem ser descritos de maneira natural através de regras de modo que sistemas baseados em regras fuzzy tem sido utilizados para resolver diversos problemas de engenharia, medicina, química, computação, etc [2, 5, 9, 10].

A base de regras tem papel central na elaboração do SBRFs. De maneira geral, quanto melhor e mais preciso for o conjunto de regras fuzzy melhor será a acuracia e comportamento geral do SBRF obtido. Esse tipo de relação entre a "qualidade" das regras com a "qualidade" do SBRF fica evidente em estudos que investigam a capacidade de aproximação universal desses tipos de sistemas [9]. Assim, faz-se necessário o estudo e o desenvolvimento de métodos que identifiquem a existência de possíveis anomalias em um conjunto de base de regras que possam produzir comportamento indesejados no SBRF construído. Do ponto de vista prático, a identificação dessas anomalias também permitem que o conjunto de dados seja revisto por especialistas, indicando quais regras requerem uma análise mais cuidadosa.

Neste projeto focaremos em SBRF cuja base de regras é composta por regras fuzzy conjuntivas, que representam parte de uma relação funcional entre entrada e saída, no entanto essa relação funcional não necessariamente corresponde a uma relação de causa-efeito tal como acontece em regras fuzzy implicativas. As regras conjuntivas, em geral, são representadas por relações fuzzy definidas em termos de t-normas que nada mais são do que extensões do conectivo Booleano "e" para a lógica fuzzy. Um exemplo de SBRF que considera base de regras conjuntivas é dada pelo método de Mamadani [9].

Neste estudo, vamos abordar três tipos de inconsistência para regras implicativas e conjuntivas: incompletude, redundância e inconsistência. Scarpelli *et al.* proporam um método para identificar possiveis incosisências entre regras fuzzy que pode ser aplicado tanto a base de regras conjuntivas quanto implicativas [11]. Este método visa comparar as regras fuzzy duas a duas com o intuito de identificar potenciais inconsistências e contradições entre elas. Isto é obtido através de um índice de consistência definido através da noção de medida de similaridade entre conjuntos fuzzy. Vale destacar que este método não leva em conta possíveis ponderações das regras tais como pode ocorre em alguns SBRFs.

2 Teoria de Conjuntos Fuzzy

Um (sub)conjunto fuzzy A de um conjunto arbitrário não vazio X é identificado como uma função $\mu_A : X \to [0,1]$ chamada de função de pertinência de A, onde $\mu_A(x)$ representa o grau com que x pertence ao conjunto fuzzy A [13]. Por simplicidade de notação, alternativamente, utilizaremos a notação $A(x)$ ao invés de $\mu_A(x)$. Dizemos que um subconjunto fuzzy A está contido em um subconjunto fuzzy B ($A \subseteq B$) se $A(x) \leq B(x)$ para todo $x \in X$. Além disso, a interseção e a união de A e B são dados respectivamente pelos conjuntos fuzzy $A \cap B$ e $A \cup B$ cujas função de pertinência são:

$$(A \cap B)(x) = \min\{A(x), B(x)\} \quad \text{e} \quad (A \cup B)(x) = \max\{A(x), B(x)\}$$

para todo $x \in X$. O complemento de A é o conjunto fuzzy $\bar{A}$ cuja função de pertinência é dada por $\bar{A}(x) = 1 - A(x),\ \forall x \in X$ [1].

A classe dos conjuntos fuzzy de X é denotado por $\mathcal{F}(X)$. Observe que $\mathcal{F}(X)$ estende a teoria clássica de conjuntos uma vez que cada subconjunto Y de X também pode ser unicamente identificado como o subconjunto fuzzy cuja função de pertinência é dada pela função característica χ_Y de Y dada por $\chi_Y(x) = 1$ se $x \in Y$ e $\chi_Y(x) = 0$ se $x \in Y$.

Um conjunto fuzzy trapezoidal $(a;b;c;d)$, $a \leq b \leq c \leq d$, é um conjunto fuzzy de R cuja função de pertinênica é dada por:

$$(a;b;c;d)(x) = \begin{cases} 1 & , \text{ se } b \leq x \leq c \\ \frac{x-a}{b-a} & , \text{ se } a \leq x < b \\ \frac{d-x}{d-c} & , \text{ se } c \leq x \leq d \\ 0 & , \text{ caso contrário} \end{cases} .$$

Se $b = c$, falamos de conjunto fuzzy triangular e denotamos por $(a; b; d)$ ao invés de $(a; b; b; d)$.

Cada conjunto fuzzy A de X pode ser identificado com uma família de subconjuntos clássicos de X chamados de α-níveis de A [6]. Precisamente, para todo $\alpha \in (0, 1]$ definimos o α-nível de A como sendo o subconjunto [8]:

$$[A]_\alpha = \{x \in X \mid \mu_A(x) \geq \alpha\}.$$

Se, adicionalmente, X é também um espaço topológico, então, definimos

$$[A]_0 = \overline{\{x \in X \mid \mu_A(x) > 0\}}$$

onde $\overline{Y}$ denota o fecho de $Y \subseteq X$. A relação de inclusão e de igualdade entre conjuntos fuzzy podem ser caracterizada em termos de α-níveis como se segue [6]:

(i) $A \subseteq B \Leftrightarrow [A]_\alpha = [B]_\alpha \; \forall \alpha \in (0, 1]$;

(ii) $A = B \Leftrightarrow [A]_\alpha = [B]_\alpha \; \forall \alpha \in (0, 1]$.

Diversos operadores sobre os conjuntos fuzzy podem ser definidos através de operadores elementares da lógica fuzzy [1]. A lógica fuzzy nada mais é do que a estensão da lógica Booleana para o domínio $[0, 1]$. Uma t-norma t é um operador de $[0, 1]^2$ para $[0, 1]$ que estende o conectivo Booleano "e" e satisfaz as seguintes propriedades para todo $a, b, c, d \in [0, 1]$:

(a) comutatividade: $a \; t \; b = b \; t \; a$;

(b) associatividade: $a \; t \; (b \; t \; c) = (a \; t \; b) \; t \; c$;

(c) monotocidade: $a \; t \; b \leq c \; t \; d$ se $a \leq c$ e $b \leq d$;

(d) 1 é elemento neutro: $1 \; t \; a = a$.

Um exemplo de t-norma muito utilizada é a do mínimo $a \wedge b = \min\{a, b\}$ para todo $a, b \in [0, 1]$.

Similarmente, uma s-norma (ou uma t-conorma) $s : [0, 1]^2 \to [0, 1]$ estende o conectivo Booleano "ou" e satisfaz as seguintes propriedades [1], para todo $a, b, c, d \in [0, 1]$:

(a) comutatividade: $a \; s \; b = b \; s \; a$;

(b) associatividade: $a \; s \; (b \; s \; c) = (a \; s \; b) \; s \; c$;

(c) monotocidade: $a \, s \, b \leq c \, s \, d$ se $a \leq c$ e $b \leq d$;

(d) 0 é elemento neutro: $0 \, s \, a = a$.

Um exemplo bem conhecido de s-norma é a do máximo: $a \vee b = \max\{a, b\}$ para todo $a, b \in [0, 1]$.

Uma relação fuzzy R entre um universo X e outro universo Y nada mais é do que um conjunto fuzzy em $X \times Y$, onde $R(x, y)$ representa o grau de relação entre $x \in X$ e $y \in Y$. De forma mais geral [1], um conjunto fuzzy R de $X = X_1 \times \ldots \times X_n$ é também chamado de uma relação n-ária em X. Um caso especial de relação fuzzy é o produto cartesiano de conjuntos fuzzy. Sejam t uma t-norma e A_i um conjunto fuzzy de X_i, $i = 1, \ldots, n$. O produto cartesiano fuzzy de $A_1, \ldots, A_n$ com respeito à t-norma t é a relação fuzzy $A_1 \times_t \ldots \times_t A_n$ cuja função de pertinência é dada por

$$(A_1 \times_t \ldots \times_t A_n)(x_1, \ldots, x_n) = A_1(x_1) \, t \ldots t \, A_n(x_n)$$

para todo $(x_1, \ldots, x_n) \in X_1 \times \ldots \times X_n$. Se a t-norma considerada é do mínimo, então, o correspondente produto cartesiano fuzzy é denotado simplesmente por $A_1 \times \ldots \times A_n$.

Uma relação fuzzy $R \in \mathcal{F}(X \times Y)$ e uma relação fuzzy $S \in \mathcal{F}(Y \times Z)$ podem ser combinadas através de composições relacionais [8] para produzir relações fuzzy em $X \times Z$. Seja t uma t-norma qualquer, a composição sup-t de $R \in \mathcal{F}(X \times Y)$ e $S \in \mathcal{F}(Y \times Z)$ é a relação fuzzy $R \circ_t S$ de $X \times Z$ cuja função de pertinência é dada por

$$(R \circ_t S)(x, z) = \sup_{y \in Y} R(x, y) \, t \, S(y, z), \quad \forall (x, z) \in X \times Z.$$

Similarmente, a composição sup-t do conjunto fuzzy $A \in \mathcal{F}(Y)$ e da relação fuzzy R é o conjunto fuzzy $A \circ_t R$ de Y cuja função de pertinência é dada por

$$(A \circ_t R)(y) = \sup_{x \in X} A(x) \, t \, R(x, y), \quad \forall y \in Y.$$

No caso em particular que t é a t-norma do mínimo, falamos de composição max-min e utilizamos simplesmente o símbolo $\circ$ ao invés de $\circ_\wedge$.

Outra composição relacional de interesse neste trabalho é a composição inf-i [9].

Seja i uma implicação fuzzy, isto é, uma função $i : [0,1]^2 \to [0,1]$ tal que $0 \; i \; 0 = 0 \; i \; 1 = 1 \; i \; 1 = 1$ e $1 \; i \; 0 = 0$, decrescente no primeiro argumento e crescente no segundo argumento, a composição inf-i de $R \in \mathcal{F}(X \times Y)$ e $S \in \mathcal{F}(Y \times Z)$ é a relação fuzzy $R \bullet_i S$ de $X \times Z$ cuja função de pertinência é dada por

$$(R \bullet_i S)(x,z) = \inf_{y \in Y} R(x,y) \; i \; S(y,z), \quad \forall (x,z) \in X \times Z.$$

Similarmente, a composição inf-i do conjunto fuzzy $A \in \mathcal{F}(X)$ e da relação fuzzy R é o conjunto fuzzy $A \bullet_i R$ de Y cuja função de pertinência é dada por

$$(A \bullet_i R)(y) = \inf_{x \in X} A(x) \; i \; R(x,y), \quad \forall y \in Y.$$

Uma negação fuzzy η é uma função de $[0,1]$ para $[0,1]$ que é decrescente e estende a negação da lógica Booleana: $\eta(0) = 1$ e $\eta(1) = 0$. Se adicionalmente η for estritamente crescente e é involutiva, isto é, $\eta(\eta(x)) = x \; \forall x \in [0,1]$, então, η é dita uma negação forte [1][9]. A negação usual é dada por $\eta(x) = 1 - x$ para todo $x \in [0,1]$ e que o complemento de um conjunto fuzzy $A \in \mathcal{F}(X)$ surge desta negação: $A^c(x) = \eta(A(x))$ para todo $x \in X$. De fato, para cada negação fuzzy podemos definir o complemento de um conjunto fuzzy com respeito à respectiva negação fuzzy. Contudo, para o propósito deste trabalho, iremos focar apenas na negação fuzzy usual.

3 Sistemas Baseados em Regras Fuzzy

Um sistema de base de regras fuzzy (SBRF) é um mapeamento Φ composto por três módulos, isto é, $\Phi(\cdot) = D(I(F(\cdot)))$, onde F é o módulo de fuzzificação, I é o módulo de inferência e D é o módulo de defuzzificação [1]. A Figura 1 ilustra a arquitetura de um sistema de base de regras fuzzy.

O módulo de fuzzificação F é um operador que mapea uma entrada em conjuntos fuzzy em domínios apropriados [1][9]. Um dos métodos de fuzzificação mais bem-conhecidos e utilizados é o método de inclusão canônica que consiste da função que associa cada vetor $x \in R^n$ com o subconjunto fuzzy $\{x\} \in \mathcal{F}(R^n)$.

O módulo de inferência [5] é o mapeamento I que associa os conjuntos fuzzy

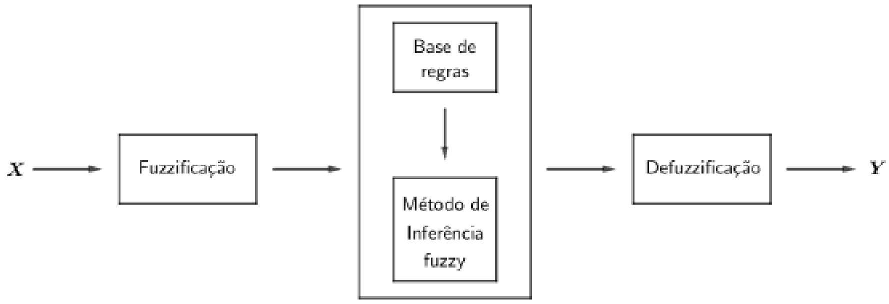

Figura 1: Arquitetura de um sistema de base de regras fuzzy (fonte: [10]).

$A \in \mathcal{F}(X)$, obtidos pelo módulo de fuzzificação F, a conjuntos de um domínio Y. O mapeamento I é obtido a partir de uma base de regras fuzzy que consiste de um conhecimento prévio do fenômeno considerado advindo de uma conjunto de dados e/ou por especificações de especialistas, e um sistema de inferência responsável por associar uma saída a novas entradas levando em conta a base de regras fuzzy.

Uma base de regras fuzzy [1] consiste de um conjunto de regras fuzzy da forma

$$\textbf{Se } x_1 \text{ is } A_{i1} \text{ e } \ldots \text{ e } x_n \text{ is } A_{in} \textbf{ então } y_1 \text{ is } B_{i1} \text{ e } \ldots \text{ e } y_m \text{ is } B_{im}, \quad i = 1, ..., k$$

onde A_{ij} e B_{ip} são conjuntos fuzzy de R associados respectivamente as variável linguísticas x_j e y_p. Note que, utilizando a noção de produto cartesiano fuzzy com respeito a uma t-norma t, cada regra pode ser reescrita da seguinte forma

$$\textbf{Se } x \text{ is } A_i \textbf{ então } y \text{ is } B_i, \qquad i = 1, ..., k$$

onde $A_i = A_{i1} \times_t \ldots \times_t A_{in} \in \mathcal{F}(R^n)$ e $B_i = A_{i1} \times_t \ldots \times_t A_{im} \in \mathcal{F}(R^m)$. As proposições fuzzy "$x$ is A_i" e "y is B_i" são respectivamente o antecedente e o consequente da i-ésima regra da base de regras fuzzy. Os conjuntos fuzzy A_i e B_i representam termos linguíticos associados respectivamente as varíaveis de entrada x e de saída y, tais como "alto", "médio", "quente", etc [9][1].

Cada regra pode ser representada por uma relação fuzzy R_i de $R^m \times R^n$ cuja função de pertinência é dada por

$$R_i(x, y) = f(A_i(x), B_i(y)), \qquad \forall\ (x, y) \in R^n \times R^m, \tag{1}$$

onde f denota uma implicação fuzzy ou uma t-norma. Se f for uma implicação fuzzy, então, dizemos que as regras são implicativas, interpretadas como relações de causa-efeito. Agora, se f for uma t-norma, então, dizemos que as regras são conjuntivas assumindo-se que as regras fuzzy representam o conhecimento parcial sobre a relação funcional entre a entrada e a saída que podem ser obtidos através do conhecimento de especialistas ou de um conjunto de dados [2, 9].

O processo de se unir as relações R_i induzidas por cada regra em uma única relação fuzzy R é chamada de agregação [5]. Considere uma base de regras fuzzy conjuntivas "**Se** x *is* A_i **então** y *is* B_i", $i = 1, ..., k$, e as respectivas relações fuzzy R_i dadas conforme a Equação (1) para alguma t-norma t dada. A relação fuzzy R é definida como a união das relações fuzzy R_i, isto é,

$$R_t = \bigcup_{i=1}^{k} R_i. \tag{2}$$

ou seja,

$$R_t(x, y) = \sup_{i=1,...,k} A_i(x)\ t\ B_i(y),\ \forall\ (x, y) \in R^n \times R^m. \tag{3}$$

A relação R_t é também chamada de *base de regras de sup-t*. Intuitivamente, a relação fuzzy de R_t consiste da união dos conhecimentos parciais sobre uma relação funcional (desconhecida) entre a entrada e a saída descritas por cada regra fuzzy. Se a t-norma t adotada para representar a base de regras fuzzy do tipo conjuntiva é a do mínimo, isto é, $t = \wedge$, então, obtemos a relação fuzzy que descreve a base de regra de Mamdani [7] que é denota por R_M ao invés de $R_\wedge$. Neste caso, a Equação (3) pode ser reescrita da seguinte forma

$$R_M(x, y) = \sup_{i=1,...,k} A_i(x) \wedge B_i(y),\ \forall\ (x, y) \in R^n \times R^m. \tag{4}$$

Agora, se a t-norma escolhida for a do produto, então, obtemos a base de regras de Larsen que denotaremos por R_L.

Um método de inferência fuzzy visa produzir uma resposta para uma determinada entrada utilizando a base de regras fuzzy dada [2]. Neste trabalho, focaremos na regra composicional de inferência [2, 9]. Mais precisamente, se R a relação fuzzy induzida por uma base de regras fuzzy do tipo "Se x is A_i, então, y is B_i para $i = 1, ..., k$", então, uma regra composicional de inferência é um mapeamento I que associa um conjunto fuzzy $A \in \mathcal{F}(R^n)$

a um conjunto fuzzy $B \in \mathcal{F}(R^m)$ dado por

$$B = I(A) = A \otimes R \tag{5}$$

onde $\otimes$ denota uma composição relacional entre a relação fuzzy R e o conjunto fuzzy A, tal como sup-t ou inf-i.

O método de inferência de Mamdani considera $\otimes$ como sendo a composição max-min na Equação (5). Neste caso, o método de Mamdani, que consiste da utilização da base de regras de Mamdani junto com o método de inferência de Mamdani, é dado por:

$$\begin{aligned} B(y) &= \sup_{x \in R^n} A(x) \ \wedge \ R_M(x,y) \\ &= \sup_{x \in R^n} A(x) \ \wedge \ A_i(x) \ \wedge \ B_i(y) \\ &= \sup_{x \in R^n} (A \cap A_i)(x) \ \wedge \ B_i(y) \end{aligned} \tag{6}$$

para todo $y \in R^m$.

Dado uma t-norma t contínua à esquerda, o método de inferência residual de Gödel consiste em considerar a base de regras sup-t e a composição relacional $\otimes = \bullet_t$ na Equação (5), isto é, a composição inf-i_t onde i_t é a implicação residual com respeito à t. Neste caso obtemos o seguinte método de inferência $I : \mathcal{F}(R^n) \to \mathcal{F}(R^m)$:

$$B(y) \ = \ \inf_{x \in R^n} A(x) \ i \ R_t(x,y). \tag{7}$$

para todo $y \in R^m$. .

O módulo de defuzzificação associa a saída $B = I(A)$ do método de inferência a um elemento do R^m. Um dos defuzzificadores mais utilizados é o método do centro de gravidade (centróide) explicitado a seguir [9]. Seja B um subconjunto fuzzy de $X \subseteq R$, o defuzzificador D dado pelo centro de gravidade de B é dado por

$$D(B) = \int_X xB(x)dx \int_X B(x)dx \tag{8}$$

claramente, aqui estamos assumindo que a integral acima existe.

4 Incompletude

Em um sistema de base de regra clássico $\{A_i \to B_i \ / \ i = 1, ..., n\}$, onde $A_i \subseteq X$ e $B_i \subseteq Y$, uma forma de analisar a completude desse conjunto de regras consiste em verificar a seguinte igualdade:

$$\cup_{i=1}^{n} A_i = X, \forall \ x \in X$$

onde $\to$ denota a implicação Booleana. Assim, a base de regras é dita incompleta se:

$$\cup_{i=1}^{n} A_i \neq X, \forall \ x \in X.$$

Em um sistema de base de regras fuzzy [3], a incompletude é identificada ao analisar se a união dos suportes dos conjuntos associados aos antecedentes cobrem todo o domínio. Mais precisamente, dizemos que a base de regras fuzzy do tipo "Se x **is** A_i então y **is** B_i", para $i = 1, ..., n$, é incompleto se existe $x \in X$ tal que $max_{i=1,...,n} A_i(x) = 0$

Geralmente, essa anomalia não é de grande interesse na prática, pois, a maior parte das técnicas de particionamentos já se preocupam com a controle dos suportes dos antecedentes. Por exemplo, os métodos que determinam regras fuzzy automaticamente a partir de dados, geralmente, levam em consideração a propriedade de completude na determinação dos antecedentes. Assim, na prática, a incompletude além de ser facilmente evitável também é fácilmente identificado, bastando analisar os suportes dos antecedentes da base de regra fuzzy.

5 Redundância

Outro tipo de anomalia que pode encontrada, é a redundancia. Segundo [4], para um sistema de base de regras implicativas, dada uma base de regras R composta por n regras, $R_1, ..., R_n$, uma regra R' será redundante com respeito a R se, e somente se

$$\mu_R(x, y) = min_{i=1,...,n} \mu_{R_i}(x, y) \leq \mu_{R'}(x, y), \ \forall x, y \in X \times Y, \tag{9}$$

onde $\mu_{R_i}, \mu'_R \in \mathcal{F}(U \times V)$ são as funções de pertinência de R_i e R' dadas em termos de uma implicação fuzzy.

Do mesmo modo, para um sistema de bases de regras conjuntivas temos R' será redundante com respeito a R se

$$\mu_R(x,y) = max_{i=1,\ldots,n}\mu_{R_i}(x,y) \geq \mu_{R'}(x,y),\ \forall x,y \in X \times Y. \tag{10}$$

onde $\mu_{R_i}, \mu'_R \in \mathcal{F}(U \times V)$ são as funções de pertinência de R_i e R' dadas em termos de uma t-norma.

No entanto, a redundância não representa necessariamente um problema, pois, não altera a saída do sistema baseado em regras fuzzy. De fato, a redundância não altera os resultados da inferência, apenas aumenta o custo computacional de processamento dos resultados. Em geral, esse aumento de processamento não menos relevantes dependendo do tambanho da base de regras. Por outro lado, possíveis contradições podem gerar erros de inferência acarretando em resultados inconsistentes e, por isso, será o foco principal dessa monografia.

6 Medida de Similaridade

A verificação da existência de anomalias em uma base de regras fuzzy é essencial para que o modelo matemático baseado nessas regras expresse corretamente o fenômino estudado. Neste trabalho estudaremos o método proposto em [11] para identificação de pares de regras possivelmente contraditórias, que por sua vez podem produzir saídas inconsistentes se ambas regras forem consideradas nos cálculos do SFBR em questão.

O modelo de Scarpelli *et al.* pode ser aplicável para analisar tanto base de regras implicativas quanto conjuntivas, por considerar apenas medidas de similaridade entre os antecedentes e consequentes sem levar em conta o papel da condicional "Se-Então" na regra. Mais especificamente, este método compara as regras fuzzy duas-a-duas analisando a consistência de tais regras através do *índice de consistência* que é definido em termos de uma dada medida de similaridade.

Uma medida de similaridade em $\mathcal{F}(X)$ é uma $S : \mathcal{F}(X) \times \mathcal{F}(X) \to [0,1]$ que satisfaz as seguintes propriedades [3]:

- $S(A,B) = S(B,A)$ para todo $A, B \in \mathcal{F}(X)$;

- $S(X, \emptyset) = 0$;

- $S(A, A) = 1$ para todo $A \in \mathcal{F}(X)$;

- Se $A \subseteq B \subseteq C$, então, $S(A, C) \leq S(A, B)$ e $S(A, C) \leq S(B, C)$.

Um exemplo bem conhecido de medida de similaridade é a função

$$S(A,B) = \begin{cases} \frac{|A \cap B|}{|A \cup B|} & , \text{ se } |A \cup B| > 0 \\ 1 & , \text{ se } |A \cup B| = 0 \end{cases} \tag{11}$$

para todo $A, B \in \mathcal{F}(X)$, onde $|C|$ denota a cardinalidade do conjunto fuzzy $C \in \mathcal{F}(X)$ e é definida como se segue:

$$|C| = \begin{cases} \int_X C(x)dx & , \text{ se } X \text{ é não enumerável} \\ \sum_{x_i \in X} C(x_i) & , \text{ se } X \text{ é enumerável} \end{cases} , \tag{12}$$

Evidentemente, na definição de cardinalidade de conjuntos fuzzy acima, estamos assumindo que as integrais ou as séries existem e são finitas nos casos respectivamente onde X é não enumerável ou X é enumerável e infinito.

7 Introduzindo a Noção de Consistência de Bases de Regras Fuzzy

Para anomalias do tipo contraditórias, abordaremos a noção de consitência de uma base de regras. Esta irá nos auxiliar a mensurar possíveis contradições em bases de regras fuzzy.

É natural esperar que se uma dada entrada ativa ao mesmo tempo mais de uma regra, então as saídas devem ser semelhantes [4]. Para o caso clássico, essa consistência é definida da seguinte foma. Um conjunto de regras $\{A_i \rightarrow B_i \ / \ i = 1, ..., n\}$ é dita consistente se $\forall \ I \subseteq \{1, ..., n\}$ temos que

$$\cap_{i \in I} A_i \neq \emptyset \Rightarrow \cap_{i \in I} B_i \neq \emptyset. \tag{13}$$

Como ilustração, utilizaremos o exemplo a seguir, que foi abordado por Dubois *et al.* em [4], que trata de um sistema de base de regras para modelar o controle de uma rota de carro.

- Dados de entrada:
  1. O_d: Obstáculo à diteita;
  2. O_e: Obstáculo à esquerda;
  3. O_f: Obstáculo à frente;
- Saídas:
  1. B_d: Vire à diteita;
  2. B_e: Vire à esquerda;
  3. B_f: Siga em frente;
- Base de regras:
  1. Se um objeto encontra-se à frente ou à esquerda, vire à direita:
  $$(O_e \cup O_f) \rightarrow B_d$$
  2. Se um objeto encontra-se à frente ou à direita, vire a esquerda:
  $$(O_d \cup O_f) \rightarrow B_e$$

Para esse sistema de base de regras com apenas duas regras, se o método de defuzzificação utilizado for o do centróide (veja Eq. (12)), então, a saída obtida será: Siga em frente (B_f). O que causaria uma colisão. Essa incoerência ocorre poque os antecedante-cedentes das regras se sobrepõem mas os consequentes não. A medida de consitência detecta isso:

$$(O_e \cup O_f) \cap (O_d \cup O_f) = O_f \neq \emptyset \text{ mas } B_d \cap B_e = \emptyset$$

No entanto, o uso de um defuzzificador mais adequado poderia contornar esse problema e devolver uma saída *vire à direita* ou *vire à esquerda*. Por exemplo, em termos fuzzy, poderíamos utilizar o defuzzificador "Mínimo do Máximo" ao invés do centroid. Vale salientar que normalmente a defuzificação utilizada é a do centróide, porém, essa nem sempre é a mais adequada dependendo do problema. Assim, uma troca no defuzificador pode ser suficiente para sanar alguma inconsistência obtida, no entanto, não nos aprofundaremos nesse tipo de abordagem.

8 Índice de Consistência de Bases de Regras Fuzzy

Tomando a equação (13) como base, Scarpelli propos uma generalização para a mesma, considerando as bases de regras fuzzy.

Seja S_X uma medida de similaridade em $\mathcal{F}(X)$ e S_Y uma medida de similaridade em $\mathcal{F}(Y)$ e duas regras r_1 :"Se x is A, então, y is B" e r_2 :"Se x is $\tilde{A}$, então, y is $\tilde{B}$", onde $A, \tilde{A} \in \mathcal{F}(X)$ e $B, \tilde{B} \in \mathcal{F}(Y)$. O índice de consistência C entre as regras r_1 e r_2 com respeito à S_X, S_Y e a implicação fuzzy i é dado por [11]:

$$C(r_1, r_2) = S_X(A_1, A_2) \; i \; S_Y(B1, B2). \tag{14}$$

De modo que o índice de consistência $C(r_1, r_2)$ é baixo quando a correspondência dos antecedentes é alta mas a correspondencia entre os respectivos consequêntes é baixa

Além disso, r_1 e r_2 são ditas potencialmente contraditórias com respeito a $\beta \in [0, 1]$ se

$$C(r_1, r_2) < \beta \tag{15}$$

onde β é um limiar de corte dado de acordo com o contexto. Cabe destacar que aqui generalizamos de maneira natural a definição original deste índice proposto por [11], onde tal conceito leva em conta escolhas particulares de medidas de similaridade.

Para n regras $r_i \in \{1, ..., n\}$, o número de comparações duas a duas realizado é dado por:

$$c = n * (n - 1)2 \tag{16}$$

Dessa forma, podemos sempre ter uma ideia do custo computacional que o problema a ser tratado acarretará.

A fim de ilustrar o método de Scarpelli *et al.* vamos considerar o seguinte exemplo de uma base de regras fuzzy que descreve o risco de uma pessoa estar com dengue baseado na sua temperatura corporal e no nível de dores das suas articulações:

- x denota a *temperatura corporal* que toma valores no domínio $X = [36, 41]$. Considere os conjuntos fuzzy $A_1 = (36; 36; 38)$, $A_2 = (37; 38; 39)$ e $A_3 = (38; 40; 41; 41)$ que modelam respectivamente os termos linguísticos *normal*, *febre moderada* e *febre alta*.

- y denota o *nível de dor nas articulações* e toma valores no domínio $Y = [0, 1]$. Considere os conjuntos fuzzy $B_1 = (0; 0; 0,5)$, $B_2 = (0; 0,5; 1)$ e $B_3 = (0,5; 1; 1)$ que modelam respectivamente os termos linguísticos *baixo*, *médio* e *alto*.

- z denota o *risco de estar com dengue* e toma valores no domínio $Z = [0, 1]$. Considere os conjuntos fuzzy $C_1 = (0; 0; 0,5)$, $C_2 = (0; 0,5; 1)$ e $C_3 = (0,5; 1; 1)$ que modelam respectivamente os termos linguísticos *baixo*, *médio* e *alto*.

- considere a seguintes regras:

  1. r_1 : **Se** (x is *febre moderada*) e (y is *médio*) **então** (z is *alto*);
  2. r_2 : **Se** (x is *febre moderada*) e (y is *alto*) **então** (z is *baixo*);
  3. r_3 : **Se** (x is *normal*) e (y is *alto*) **então** (z is *baixo*);

Considerando a medida de similaridade dada na Equação (11) e a implicação de fuzzy de Goguen, obtemos os seguintes índices de consistências:

$$C(r_1, r_2) = 0$$

$$C(r_1, r_3) = 0$$

$$C(r_2, r_3) = 1$$

Pelo método de Scarpelli *et al.*, obtemos que a regra r_1 é contraditória com as regras r_2 e r_3, indicando que uma pessoa, segundo essas regras, pode apresentar risco de ter dengue baixo e alto ao mesmo tempo.

8.1 Uma proposta de estimativa para o parâmetro β

Uma vez que β é um indíce de quais regras são potencialmentes conflitantes, o desafio é conseguir mensurar um valor que seja suficiente para detectar as regras de interesse, sem que o mesmo detecte muitas regras que não são de fato inconsistentes.

A seguir analisaremos mais de perto as regras apontadas como potencialmente contraditórias no problema tratado por SANTOS em [10], de acordo com o método baseado em uma medida de similaridade proposto por [11] para diversos valores de β. Para $\beta_1 = 0,5$, o método não apresenta nenhum par de regras potencialmete conflitantes. Para $\beta_2 = 0,7$ obtem-se 13 pares de regras potencialmete conflitantes e para $\beta_3 = 0,85$ temos 82 pares de regras potencialmente conflitantes. Obviamente, os pares de regra que apareceram para β_2 continam aparecendo para β_3.

Uma questão natural que surge nesse contexto é de como obter um limiar adequado para identificar regras possivelmente contraditórias. Motivados por essa questão, propomos determinas o valor de β atraves da integral abaixo, contudo, a aplicação e as análises dessa equação serão realizadas em trabalhos futuros:

$$\beta' = \sum_{i=1}^{n} \frac{1}{n} \int_0^1 (1-\beta) m(r_i, \beta) d\beta \tag{17}$$

onde n é o número total de regras a serem analizadas e

$$m(r_i, \beta) = |\{j \neq i \ / \ C(r_i, r_j) < \beta\}|. \tag{18}$$

Observe que encontrar as regras que geram inconsistências significa apontar os possíveis resultados incompatíveis encontrados na saída do programa, que por sua vez pode auxiliar em melhorias do SBRF considerado.

9 Conclusão

Neste projeto abordamos anomalias em sistemas de bases de regras fuzzy, analisando as inconsistencias do tipo incompletude, redundância e contradição [3], para regras conjuntivas e implicativas. Abordamos mais amplamente os conceitos de consistência de bases de regras, desde a introdução da ideia até a medida do índice de consistência, proposta por SCARPELLI *et al.*, que leva em consideração a medida de similaridade que relaciona um par de regras fuzzy por vez.

A incompletude é geralmente fácil de identificar, pois basta analisar se os suportes dos conjuntos associados aos antecedentes cobrem todo o domínio. A redundância, por sua vez, não altera a saída produzida por um sistemas baseado em regras fuzzy, mas causa aumento de processamento do sistema obtido. Por fim, resta analisar possíveis contradições entre as regras.

Na literatura há diversas abordagens nesse sentido [11],[12], [4]. No entanto, nosso trabalho se estrurou no modelo proposto por Scarpelli *et al.* que apresenta um método de identificação de possíveis contradições entre regras, baseado em medidas de similaridades entre antecedentes e consequentes. Contudo, da revisão bibliográfica que fizemos, vale destacar que nenhum método da literatura considera sistemas baseados em regras fuzzy ponderadas. Sendo assim, como pesquisa futura, pretendemos estender as ideias de Scarpelli *et al.* para lidar também com essas classes de SBRFs.

Referências

[1] Laécio Carvalho Barros, Rodney Carlos Bassanezi, and Weldon Alexander Lodwick. *First Course in Fuzzy Logic, Fuzzy Dynamical Systems, and Biomathematics*. Springer, 2016.

[2] B. Bede. *Mathematics of Fuzzy Sets and Fuzzy Logic*. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg, 2013.

[3] H. Bustince, M. Pagola, and E. Barrenechea. Construction of fuzzy indices from fuzzy DI-subsethood measures: Application to the global comparison of images. *Information Sciences*, 177(3):906 – 929, 2007.

[4] Didier Dubois, Henri Prade, and Laurent Ughetto. Checking the coherence and redundancy of fuzzy knowledge bases. *IEEE Transactions on Fuzzy Systems*, 5(3):398–417, 1997.

[5] Rene Jager. *Fuzzy logic in control*. PhD thesis, Delft University of Technology, Delft, Holanda, 1995.

[6] George Klir and Bo Yuan. *Fuzzy sets and fuzzy logic*, volume 4. Prentice hall New Jersey, 1995.

[7] Ebrahim H Mamdani and Sedrak Assilian. An experiment in linguistic synthesis with a fuzzy logic controller. *International journal of man-machine studies*, 7(1):1–13, 1975.

[8] CV Negoita and DA Ralescu. Representation theorems for fuzzy concepts. *Kybernetes*, 4(3):169–174, 1975.

[9] Witold Pedrycz and Fernando Gomide. *Fuzzy Systems Engineering: Towards Human-Centric Computing*. Wiley, IEEE Press, New York, 2007.

[10] Daniel D. C. Santos. Um estudo sobre identificação de anomalias em bases de regras fuzzy aplicado a estimação do risco de endometriose. Master's thesis, University of Campinas, Campinas, Brazil, 2018.

[11] H Scarpelli, W Pedrycz, and F Gomide. Quantification of inconsistencies in fuzzy knowledge bases. In *Proc. of the Second European Congress on Intelligent Techniques and Soft Computing, EUFIT*, volume 94, pages 1456–1460, 1994.

[12] Ronald R Yager and Henrik Legind Larsen. On discovering potential inconsistencies in validating uncertain knowledge bases by reflecting on the input. *IEEE transactions on systems, man, and cybernetics*, 21(4):790–801, 1991.

[13] Lotfi A Zadeh. Fuzzy sets. *Information and control*, 8(3):338–353, 1965.