Projeto de Iniciação Científica - Unicamp/PIBIC/CNPq

Dependência e interatividade: uma abordagem a partir da teoria de conjuntos fuzzy

Relatório Final

Aluno: Felipo Bacani

Orientador: Prof. Dr. Laécio Carvalho de Barros

Departamento de Matemática Aplicada e Computacional

Instituto de Matemática, Estatística e Computacão Científica, Unicamp

Introdução

A teoria de conjuntos fuzzy foi criada com o objetivo de representar incertezas que dizem respeito à pertinência. Para isto, criou-se uma teoria de conjuntos com fronteiras imprecisas. Neste trabalho serão investigadas questões sobre interatividade entre conjuntos dessa natureza, objetivando aplicar este conceito à inferência de dados. Isto é, a partir da pertinência de um elemento a um conjunto, pretende-se inferir a pertinência desse mesmo elemento a outro conjunto fuzzy. Através de uma leitura comparada de artigos, faz-se uma discussão detalhada do assunto, utilizando analogias com a teoria de probabilidades com o propósito de ilustrar alguns conceitos. Quanto ao modelo de inferência utilizado, este se baseia conceitualmente na teoria de inferência bayesiana.

A metodologia estudada durante o período é ilustrada aqui por meio de exemplos, e as discussões que justificam a escolha pelas abordagens e modelos utilizados no projeto foram organizadas didaticamente na forma de um relatório técnico. As teorias de probabilidades e de inferência bayesiana foram cruciais para traçar as idéias que fundamentam a metodologia utilizada para inferência fuzzy, e também para generalizar a abordagem ao problema de interatividade entre conjuntos dessa natureza, uma questão relevante desde o nascimento da própria teoria fuzzy.

Sobre a literatura

Quanto aos artigos da referência à literatura, dois deles foram essenciais. Um deles é [2], que é um dos primeiros textos sobre independência e interatividade em conjuntos fuzzy. Utilizando uma analogia com a teoria de probabilidades, o texto dialoga com os primeiros questionamentos do próprio criador da teoria fuzzy sobre o assunto[6], que surgiram na metade da década de 70. O outro artigo importante é [3], que traz a mesma questão à tona, quase 20 anos depois dos outros dois textos citados. Ele a coloca tal questão em termos de outra teoria bem consolidada que, assim como a teoria de probabilidades, parece não ter relação aparente com o assunto estudado: o problema é generalizado em termos da teoria de inferência bayesiana. Quanto aos livros [1, 5] citados no projeto inicial do estudo, foram utilizados mais com o objetivo de fornecer referências e exemplos sobre os tópicos básicos da teoria de conjuntos fuzzy. Pelo fato de o projeto envolver questões muito específicas, e que ainda não se têm um consenso sobre, um estudo minucioso dos artigos [2] e [3], citados acima, foi essencial para que o projeto pudesse ser realizado.

Objetivos do trabalho

Os objetivos do projeto são estudar aspectos teóricos e práticos sobre a questão de independência e interatividade entre conjuntos fuzzy. Do ponto de vista prático, essas questões são importantes para fornecer formas de como mensurar a influência entre diferentes fatores de um mesmo problema; por exemplo, num caso de diagnóstico médico, tal estudo pode fornecer meios de estabelecer uma relação entre alguns sintomas que são relevantes para a determinação de uma determinada patologia. Ou seja, uma medida relevante de um determinado sintoma pode influenciar outros sintomas. O presente estudo fornece maneiras de como quantificar tal influência. Além do estudo dos artigos e livros da bibliografia, ao longo do trabalho surgiu também a necessidade de também tornar o estudo realizado acessível a pesquisadores e profissionais interessados, e não apenas os da matemática e áreas correlatas. Com isso em mente, foi elaborado um relatório técnico que tornará o assunto estudado o mais didático possível, de forma a ser um texto auto contido com respeito às definições, desenvolvimento da teoria, explicitação das potencialidades práticas, e discussões relevantes e referências externas. O texto está em fase de conclusão, e trechos dele serão utilizados neste relatório para elucidar algumas questões.

Atividades desenvolvidas

O trabalho foi basicamente ler as referências e escrever sobre elas. Foi realizada a leitura dos textos da bibliografia, estabelecendo relações entre os artigos estudados, colocando os conceitos definidos nos artigos em termos de tópicos já existentes da teoria fuzzy, utilizando os livros para esclarecer alguns conceitos importantes e estudando os exemplos dos artigos. O que foi realizado está sendo organizado de forma didática no relatório técnico citado acima. Dois exemplos obtidos dos artigos [2, 3] estudados são estruturalmente simples, mas representativos de duas famílias relevantes de problemas: reconhecimento de padrões e previsão de informações (primeiro e segundo exemplo, respectivamente).

Resultados

O trabalho baseia-se no seguinte questionamento: existe alguma forma de expressar a interatividade e a independência entre conjuntos fuzzy de uma forma análoga a como esse conceito é expresso na teoria de probabilidades? Ou seja, se pudermos representar conjuntos fuzzy através de textitdistribuições, podemos utilizar essas distribuições para trabalhar de forma muito análoga com a teoria de probabilidades. Como se verá a seguir, proceder dessa forma é interessante, mas dá origem a muitos questionamentos. Tais questionamentos vêm de não ser possível trabalhar exatamente da mesma forma nas duas teorias; principalmente pelo fato dos operadores matemáticos envolvidos serem diferentes, como será explicado a seguir. As primeiras definições do trabalho, feitas abaixo, são extraídas do referido relatório técnico escrito durante o projeto.

A teoria de conjuntos fuzzy se destina a incluir no formalismo matemático objetos que envolvem certo tipos de imprecisão, inexatidão, vagueza. Tradicionalmente, esses temas têm sido formalizado matematicamente por meio da teoria de probabilidades. Porém, como veremos, a proposta da teoria de conjuntos fuzzy é totalmente diferente da de probabilidades. A teoria dos conjuntos fuzzy lida com conceitos incertos (inexatos ou vagos) cuja "fronteira" é imprecisa. Isto é, conceitos que expressam "dúvidas" a respeito da pertinência de elementos a uma certa classe.

Intuitivamente (e ingenuamente), podemos dizer que originalmente a teoria dos conjuntos fuzzy lida com incerteza quanto à *pertinência*, enquanto a teoria de probabilidades lida com a incerteza ligada à *ocorrência*. Por exemplo, no lançamento de uma moeda a "incerteza probabilística" termina logo após a realização do experimento. Porém, a dúvida no reconhecimento da face da moeda pode persistir, a depender do estado da moeda. É no reconhecimento do resultado que reside a "incerteza fuzzy". É claro que o exemplo citado é um tanto fictício, mas torna claros os conceitos envolvidos. Na medicina por exemplo, um diagnóstico correto depende fundamentalmente da identificação dos sinais do paciente [4], sinais estes que muitas vezes estão sujeitos à incertezas que podem ser descritos por conjuntos fuzzy.

Subconjuntos fuzzy são conjuntos onde os elementos de um determinado universo têm graus de pertinência. No caso da teoria clássica de conjuntos, um elemento têm uma condição bivalente com relação a qualquer conjunto: pode pertencer ou não ao subconjunto - que corresponde a graus de pertinência 1 e 0, respectivamente. No caso da teoria de conjuntos fuzzy, cada elemento pertence a conjuntos desse tipo de forma incerta ou parcial. Assim, para cada subconjunto fuzzy do conjunto universo existe uma função que estende o conceito de função característica (ou indicadora) de conjuntos clássicos.

Definição 1.1. *[Função indicadora] Seja U um conjunto clássico não-vazio, e A um subconjunto* *de U. A função característica do conjunto A, $\chi_A : U \to \{0,1\}$, é dada por:*

$$\chi_A(x) = \begin{cases} 1, & se\, x \in A \\ 0, & se\, x \notin A \end{cases}$$

A equação acima descreve totalmente quais elementos de U pertencem ao conjunto A, identificando cada elemento de U com um grau de pertinência, que pode ser 0 ou 1.

Podemos imaginar, num contexto mais geral, conjuntos *fuzzy*, ou seja, que têm fronteiras que não são bem definidas. Suponha então que os elementos de U possam pertencer *parcialmente* a conjuntos desse tipo. Sendo assim, tais elementos teriam um grau de pertinência que estaria *entre* 0 e 1. Então, se fosse definir uma função equivalente à característica para tais conjuntos, obtém-se algo análogo:

Definição 1.2. *[Função de pertinência] Seja U um conjunto clássico não-vazio. Um subconjunto fuzzy F de U é descrito pela função φ_F:*

$$\varphi_F : U \to [0,1]$$

chamada função de pertinência.

Uma forma de representar funções de pertinência é vê-las como *distribuições de possibilidades.* Esta representação têm a vantagem de explorar analogias com a teoria de probabilidades, e esta será a representação utilizada no presente estudo. Sendo assim, as três sentenças a seguir designam o mesmo conceito:

1. Um subconjunto fuzzy X de um conjunto universo U
2. Uma função de pertinência $\varphi_X : U \to [0,1]$
3. Uma distribuição de possibilidades $\Pi_X : U \to [0,1]$

Quando se têm dois conjuntos universo U_1 e U_2, a função de pertinência conjunta de dois conjuntos fuzzy X_1 de U_1 e X_2 de U_2 é representada como uma *distribuição de possibilidades conjunta* $\Pi_{(X_1,X_2)}$:$U_1 \times U_2$→[0,1], onde $\Pi_{(X_1,X_2)}(u_1,u_2)$, $u_1 \in U_1$ e $u_2 \in U_2$, representa o valor da pertinência conjunta do par (u_1,u_2) a $(X_1 \times X_2)$. Como dissemos antes, analogias com probabilidades serão exploradas. Assim, mais adiante utilizar-se-à a noção de distribuições de probabilidades condicionais para definir algo análogo e central no presente estudo: *distribuições de possibilidade condicionais.*

Utilizam-se neste estudo alguns dos principais conceitos da **teoria de probabilidades**:

Definição 1.3. *[Espaço de probabilidades] Um espaço de probabilidades é uma tripla* $(\Omega,\ \mathcal{A},\ P)$ *que consiste de:*

- *O espaço amostral Ω: conjunto arbitrário não-vazio*
- *A σ-algebra $\mathcal{A} \subseteq 2^{\Omega}$: conjunto formado por subconjuntos de Ω, denominados eventos*
- *A medida de probabilidades $\boldsymbol{P} : \mathcal{A} \to [0,1]$*

Definição 1.4. *[Variável aleatória] Uma variável aleatória X é uma aplicação cujo domínio é o espaço amostral Ω, que associa eventos a números reais.*

Definição 1.5. *[Distribuição de probabilidades] Se X é uma variável aleatória,*

$$\boldsymbol{P}_X(x) = Prob\{w \in \Omega : X(w) = x\}$$

é sua distribuição de probabilidades.

Se X_1 e X_2 forem variáveis aleatórias em Ω_1 e Ω_2 respectivamente, então a distribuição de probabilidades conjunta do vetor aleatório (X_1, X_2), onde Prob e a medida de probabilidades, é definida por

$$\mathbf{P}_{X_1,X_2}(x_1, x_2) = \mathrm{Prob}\{(w_1, w_2) \in \Omega_{1\times}\Omega_2 : (X_1(w_1), X_2(w_2)) = (x_1, x_2)\}$$

Da mesma forma que define-se distribuições condicionais em probabilidades, pode-se também definir distribuições condicionais de possibilidade. Este conceito será o centro do presente estudo, na medida que permite com que se possa conhecer o valor de pertinência de um elemento a um conjunto fuzzy *dado que* se conhece o grau de pertinência a outro conjunto fuzzy desse mesmo elemento. Agora, definamos tais distribuições condicionais de probabilidade. O caso que interessa aqui é quando os conjuntos envolvidos X_1 e X_2 são discretos. Se $\mathbf{P}(x_1) > 0$, então temos:

$$\mathbf{P}_{(X_2|X_1)}(x_2|x_1) = \frac{\mathbf{P}_{(X_1,X_2)}(x_1, x_2)}{\mathbf{P}_{X_1}(x_1)}$$

ou ainda

$$\mathbf{P}_{(X_1,X_2)}(x_1, x_2) = \mathbf{P}_{X_1}(x_1) \cdot \mathbf{P}_{(X_2|X_1)}(x_2|x_1)$$

Temos ainda as relações

$$\mathbf{P}_{X_1}(x_1) = \sum_{x_2} \mathbf{P}_{(X_1,X_2)}(x_1, x_2) = \sum_{x_2} \mathbf{P}_{X_2}(x_2) \cdot \mathbf{P}_{(X_1|X_2)}(x_1|x_2)$$

$$\sum_{x_2} \mathbf{P}_{(X_2|X_1)}(x_2|x_1) = 1$$

Uma vez que as definições básicas sobre conjuntos fuzzy foram feitas, e foi explicitado como eles serão estudados no presente trabalho (através de distribuições de possibilidade), convém agora dar uma idéia de como tais noções foram exploradas pelos artigos da bibliografia. É comum na teoria fuzzy trabalhar-se com operadores lógicos como os de mínimo e máximo, o que ocorre também em outras áreas da matemática, mas na teoria fuzzy operadores desse tipo têm uma importância muito maior que o usual. Por esse motivo, os artigos da referência trabalham com a substituição dos operadores de soma e produto nas equações da teoria de probabilidades acima por outros operadores, como o de máximo e mínimo. Então as distribuições de possibilidade são razoavelmente parecidas com as de probabilidade, só que com relação à outras operações. Um dos casos particulares mais importantes, estudado com detalhes no artigo [2], é visto a seguir.

Artigo de Ellen Hisdal

Como já foi dito anteriormente, um caso particular de substituição de operadores é o de trocar a operação de soma pela de máximo, e a de produto pelo de mínimo. Então, neste caso, a fórmula probabilística

$$\mathbf{P}_{(X_1,X_2)}(x_1,x_2) = \mathbf{P}_{X_1}(x_1)\cdot\mathbf{P}_{(X_2|X_1)}(x_2|x_1)$$

tem sua correspondente na teoria de possibilidades a fórmula

$$\Pi_{(X_1,X_2)}(u_1,u_2) = \Pi_{X_1}(u_1)\wedge\Pi_{(X_2|X_1)}(u_2|u_1), \tag{1.1}$$

trocando os índices 1 e 2, temos

$$\Pi_{(X_1,X_2)}(u_1,u_2) = \Pi_{X_2}(u_2)\wedge\Pi_{(X_1|X_2)}(u_1|u_2), \tag{1.2}$$

e de forma análoga a equação da distribuição marginal de probabilidades:

$$\mathbf{P}_{X_1}(x_1) = \sum_{x_2}\mathbf{P}_{(X_1,X_2)}(x_1,x_2),$$

tem como correspondente a fórmula

$$\Pi_{X_1}(u_1) = \vee_{u_2}\left[\Pi_{(X_1,X_2)}(u_1,u_2)\right] \tag{1.3}$$

, e trocando os índices 1 e 2, temos:

$$\Pi_{X_2}(u_2) = \vee_{u_1}\left[\Pi_{(X_1,X_2)}(u_1,u_2)\right] \tag{1.4}$$

A analogia como foi feita acima é útil no sentido que define algo inicialmente desconhecido em termos de conceitos de uma teoria que é mais familiar, a teoria de probabilidades. Mas fazer isso dá origem a alguns problemas, que surgem do fato dos operadores de máximo e mínimo não serem tão bem comportados como os de soma de produto. O que ocorre é que tais operadores podem dar origem a soluções que não são necessariamente únicas, podendo ocorrer de a solução para uma equação pode ser qualquer valor em um determinado intervalo. Por exemplo, a fórmula (1.1), se resolvida em termos da distribuição condicional de possibilidades $\Pi_{(X_2|X_1)}(u_2|u_1)$, tem como solução

$$\Pi_{(X_2|X_1)}(u_2|u_1) = \begin{cases}\alpha = \Pi_{(X_1,X_2)}(u_1,u_2) & \text{para } \Pi_{X_1}(u_1) > \Pi_{(X_1,X_2)}(u_1,u_2)\\ \alpha\in[\Pi_{(X_1,X_2)}(u_1,u_2),1] & \text{para } \Pi_{X_1}(u_1) = \Pi_{(X_1,X_2)}(u_1,u_2)\end{cases} \tag{1.5}$$

e analogamente trocando os índices 1 e 2, a equação (1.2) dá origem à solução

$$\Pi_{(X_1|X_2)}(u_1|u_2) = \begin{cases}\alpha = \Pi_{(X_1,X_2)}(u_1,u_2) & \text{para } \Pi_{X_2}(u_2) > \Pi_{(X_1,X_2)}(u_1,u_2)\\ \alpha\in[\Pi_{(X_1,X_2)}(u_1,u_2),1] & \text{para } \Pi_{X_2}(u_2) = \Pi_{(X_1,X_2)}(u_1,u_2)\end{cases} \tag{1.6}$$

Note que a solução (1.5) não é necessariamente única. De fato, quando $\Pi_{X_1}(u_1) = \Pi_{(X_1,X_2)}(u_1,u_2)$ têm-se que a solução é qualquer elemento que pertença ao intervalo fechado $[\Pi_{(X_1,X_2)}(u_1,u_2),1]$.

$u_2 \rightarrow$	u_1 $\downarrow$	$\Pi_{(X_1,X_2)}(u_1,u_2)$ 1	2	3	$\Pi(u_1)$
	1	0.1*	0.1	0	0.1
	2	0.1	0.9*	0.5	0.9
	3	0	0.5	1*	1
	$\Pi(u_2)$	0.1	0.9	1	

	u_2=1	u_2=2	u_2=3	$\Pi(u_2)$
$\Pi_{(X_2\|X_1)}(u_2\|u_1=1)$	1*	0.5*	0*	0.1
$\Pi_{(X_2\|X_1)}(u_2\|u_1=2)$	0.1	1*	0.5*	0.9
$\Pi_{(X_2\|X_1)}(u_2\|u_1=3)$	0*	0.5*	1*	1
$\Pi_{X_2}(u_2)$	0.1	0.9	1	

	$\Pi_{(X_1\|X_2)}(u_1\|u_2=1)$	$\Pi_{(X_1\|X_2)}(u_1\|u_2=2)$	$\Pi_{(X_1\|X_2)}(u_1\|u_2=3)$	$\Pi_{X_1}(u_1)$
u_1=1	1*	0.1	0*	0.1
u_1=2	0.5	1*	0.5*	0.9
u_1=3	0*	0.5*	1*	1
$\Pi_{X_2}(u_2)$	0.1	0.9	1	

Tabela 1.1: Distribuições de possibilidades para classe de letras "U ou V maiúscula". As entradas sem asterisco foram computadas a partir das entradas marcadas por asterisco.

No restante do domínio - $\Pi_{X_1}(u_1) > \Pi_{(X_1,X_2)}(u_1,u_2)$ -, têm-se que a distribuição condicional $\Pi_{(X_2|X_1)}(u_2|u_1)$ tem o mesmo valor da distribuição conjunta $\Pi_{(X_1,X_2)}(u_1,u_2)$.

A não-unicidade da solução explorada no parágrafo acima é fundamental para levar a discussão adiante; tal fato inicia um questionamento acerca da *independência e interatividade* entre conjuntos fuzzy, uma das questões centrais do presente projeto. Essa questão é de muita importancia pelo fato de relacionar-se fortemente com aplicações, que são modelos de inferência de dados descritos por conjuntos fuzzy. Ou seja, se há alguma relação entre dois conjuntos fuzzy (eles têm algum tipo interatividade e/ou dependência), saber informações sobre um deles ajuda a predizer dados sobre o outro conjunto. É nessa premissa que se baseiam sistemas que inferem informações sobre um determinado efeito, com base em dados conhecidos sobre fatores periféricos ao sistema.

Um dos resultados mais inesperados quando se segue trabalhando com essa analogia com teoria de probabilidades é que, no caso fuzzy os conceitos de interatividade e independência não são equivalentes, o que ocorre em probabilidades. Esse fato generaliza a abordagem utilizada nos primeiros questionamentos sobre o assunto, feitos pelo próprio criador da teoria fuzzy[6].

A autora do artigo [2] prossegue utilizando as equações (1-6) para ilustar um exemplo: um sistema de reconhecimento de caligrafias. Este é problema de reconhecimento de padrões relativamente simples, mas representa uma família de problemas similares. Com o objetivo de reconhecer as letras U e V maiúsculas, os caracteres serão divididos em duas metades, onde a parte de baixo da letra em reconhecimento é a parte que a divide as duas metades. Essas três letras foram escolhidas por encaixarem-se no mesmo padrão de serem "dois traços essencialmente verticais conectados na parte de baixo". A altura da metade esquerda será representada por $u_1 \in X_1$, e a da direita por $u_2 \in X_2$, onde X_1 e X_2 são subconjuntos fuzzy de U_1 e U_2, respectivamente. Para propósitos de ilustração, considera-se aqui os valores possíveis das alturas de cada uma das metades como 3, 6 e 9 milímetros. Dessa forma, $U_1 = U_2 = \{1, 2, 3\}$ (em unidades de 3mm). Cada uma das distribuições de possibilidade $\Pi_{(X_1,X_2)}(u_1,u_2)$, $\Pi_{(X_2|X_1)}(u_2|u_1)$ e $\Pi_{(X_1|X_2)}(u_1|u_2)$ representam um subconjunto fuzzy do universo $U = U_1 \times U_2$.

Trabalha-se no exemplo da seguinte forma: supôe-se alguns valores das distribuições que sejam razoáveis, e utilizam-se estes valores iniciais e as equações (1-6) para obter todos os valores restantes. Como este é um problema discreto de pequeno porte, alguns valores iniciais foram suficientes para se obter todos os outros. A tabela 1 contém todos os valores das distribuições deste exemplo.

Este exemplo é simples, mas mostra a potencialidade da abordagem utilizada. Pode-se ter que em muitos problemas de reconhecimento de padrões, abordagens como a utilizada por Ellen Hisdal podem ser interessantes, na medida em que um modelo simples permite, utilizando alguns valores iniciais definidos baseados no que se têm de conhecimento específico do problema (informações que podem ter sido fornecidas por um especialista, que não necessariamente precisa ter conhecimento sobre conjuntos fuzzy ou mesmo sobre matemática), obter conclusões que podem auxiliar em uma tomada de decisões, por exemplo.

Artigo de Stephanie Lapointe

A abordagem de Lapointe no artigo [3] vai em uma direção que de certa forma complementa o trabalho de Ellen Hisdal; novamente a teoria de probabilidades tem relevância, mas desta vez um dos seus ramos específicos mais reconhecidos é utilizado como fonte de analogias: a teoria de *Inferência Bayesiana* tem papel muito relevante nesse artigo.

Como já foi dito, a teoria fuzzy trabalha bastante com operadores que não são muito usuais, como o de mínimo e máximo. O artigo de Ellen Hisdal (explorado no tópico acima) substitui os operadores de multiplicação pelo de mínimo e o de soma pelo de máximo. Lapointe generaliza a abordagem de Hisdal da seguinte forma: substitui a multiplicação e a soma por operadores que generalizam os de mínimo e máximo. Tais operadores gerais são muito utilizados na teoria fuzzy; são as chamadas T-normas (que têm o mínimo como caso particular) e T-conormas (máximo como caso particular). Dessa forma, a equação (1.1), reproduzida abaixo:

$$\Pi_{(X_1,X_2)}(u_1,u_2) = \Pi_{X_1}(u_1) \wedge \Pi_{(X_2|X_1)}(u_2|u_1)$$

têm sua forma geral dada por:

$$\Pi_{(X_1,X_2)}(u_1,u_2) = T(\,\Pi_{X_1}(u_1),\, \Pi_{(X_2|X_1)}(u_2|u_1)\,)$$

onde T é alguma T-norma. Para o caso onde $T = min$, temos o caso da equação (1.1). De forma similar, a equação (1.3), reproduzida abaixo,

$$\Pi_{X_1}(u_1) = \vee_{u_2}\left[\Pi_{(X_1,X_2)}(u_1,u_2)\right],$$

têm sua correspondente generalizada com relação aos operadores

$$\Pi_{X_1}(u_1) = S_{u_2}\left[\Pi_{(X_1,X_2)}(u_1,u_2)\right],$$

onde S é alguma T-conorma.

Tal generalização quanto aos operadores envolvidos é útil no sentido de que agora existem mais possibilidades de modelagem. Ela também serve para contextualizar as distribuições de possibilidade num contexto mais geral: o de relações fuzzy. Uma relação fuzzy sobre os conjuntos U_1, U_2 é qualquer subconjunto fuzzy de $U_1 \times U_2$. Como as distribuições de possibilidade são na verdade funções de pertinência de subconjuntos fuzzy, pode-se interpretar tais distribuições como subcon-juntos fuzzy de um conjunto universo $U_1 \times U_2$.

Tratar distribuições condicionais como relações fuzzy quando conveniente tem a grande vantagem de procurarmos as soluções do problema que nos interessa com as mesmas ferramentas que se usa para resolver um problema de relações fuzzy, sendo que as relações fuzzy já as têm ferramentas necessárias para encontrar suas soluções, e tais ferramentas devem-se em grande parte à sua grande relevância na resolução de problemas das mais diversas áreas (como as áreas médicas e biológicas). Em suma, representar distribuições de possibilidade como relações fuzzy permite resolver um problema novo utilizando ferramentas já existentes; ao invés de criarmos uma nova forma de achar soluções, utilizamos ferramentas já conhecidas. Além disso, essa representação é feita sem que haja nenhuma perda de generalidade por parte das distribuições condicionais, o que é extremamente satisfatório.

Representar distribuições de possibilidade dessa forma têm também a vantagem de resolver o problema da não-unicidade das soluções, problema enfrentado por Ellen Hisdal (tópico anterior) quando isolou a distribuição condicional $\Pi_{(X_2|X_1)}(u_2|u_1)$ de (1.1), dando origem à solução

$$\Pi_{(X_2|X_1)}(u_2|u_1) = \begin{cases} \alpha = \Pi_{(X_1,X_2)}(u_1,u_2) & \text{para } \Pi_{X_1}(u_1) > \Pi_{(X_1,X_2)}(u_1,u_2) \\ \alpha \in [\Pi_{(X_1,X_2)}(u_1,u_2),1] & \text{para } \Pi_{X_1}(u_1) = \Pi_{(X_1,X_2)}(u_1,u_2) \end{cases}$$

O trabalho de Lapointe contorna o problema acima com a seguinte premissa: quando for obtido um intervalo como solução, sempre se escolherá o maior valor deste intervalo. Logo, a solução da equação acima na abordagem de Lapointe fica:

$$\Pi_{(X_2|X_1)}(u_2|u_1) = \begin{cases} \Pi_{(X_1,X_2)}(u_1,u_2) & \text{para } \Pi_{X_1}(u_1) > \Pi_{(X_1,X_2)}(u_1,u_2) \\ 1 & \text{para } \Pi_{X_1}(u_1) = \Pi_{(X_1,X_2)}(u_1,u_2) \end{cases} \tag{1.7}$$

pelo fato de $\sup\{x \in [\Pi_{(X_1,X_2)}(u_1,u_2),1]\} = 1$. Além de o problema citado acima ter sido resolvido, o procedimento utilizado para encontrar o valor de $\Pi_{(X_2|X_1)}(u_2|u_1)$ na equação (1.7) independe da T-norma utilizada. Isso quer dizer que para quase qualquer T-norma, é garantida a existência de solução para a equação (1.7). Logo o caso onde $T = min$ é obtido como caso particular.

Com base nessa abordagem, foi estudado um exemplo. O tipo de aplicação estudada é chamada de *Processamento de previsões*. Assuma que um tomador de decisões têm que realizar uma ação que depende de uma variável crítica X, cujo valor só será conhecido no futuro. Ele recebe uma previsão imperfeita Y da variável X, e ele têm à sua disposição um registro histórico de previsões Y e de valores correspondentes observados X. Esse registro provém duas informações: informa a habilidade de quem faz a previsão, e informa também sobre a distribuição da variável X. O objetivo do tomador de decisões é combinar a previsão Y com a distribuição do histórica de X, de forma a determinar a distribuição posteriori da variável X *condicionada* à previsão Y; a distribuição resultante desse processo será a utilizada para tomar a decisão. Processamento de previsões refere-se ao problema de modificar a distribuição de X levando em conta uma previsão Y.

De certa forma, procede-se de maneira análoga ao outro exemplo estudado (tópico sobre Ellen Hisdal), no sentido que precisa-se de algumas informações iniciais para, a partir destas, obter as distribuições condicionais de probabilidades. As informações que foram necessárias inicialmente foram: a distribuição marginal do conjunto X, $\Pi_{X_1}(u_1)$, e a distribuição condicional $\Pi_{(Y|X)}(y|x)$. Ambas as distribuições foram construídas com base no registro histórico de Y por X. Uma vantagem que tal modelagem tem com relação ao exemplo de Hisdal é que pode-se trabalhar com distribuições contínuas de forma totalmente livre, sem a menor perda de generalidade. Além disso, para um mesmo modelo pode-se obter estimativas com relação a diferentes T-normas, o que permite, à medida que o modelo for sendo testado, escolher a que melhor se encaixa no problema particular. A quantidade de problemas que poderiam ser modelados pelo exemplo geral proposto acima é muito grande, de forma que a sua relevância é muito grande para o presente estudo.

Observações sobre a continuidade do trabalho

A própria elaboração do relatório técnico citado, trabalhando de forma didática entre os conceitos novos e as ferramentas já existentes da teoria fuzzy, tem o objetivo de permitir que haja a continuidade do trabalho. Assim que este for concluído, será publicado como literatura interna do instituto (ao menos inicialmente). O discente estuda a possibilidade de continuar pesquisando nessa área em nível de mestrado.

Conclusões

O que se conclui é que as teorias de probabilidades e de inferência bayesiana foram essenciais para o presente trabalho. O modelo proposto por Stephanie Lapointe têm uma vantegem com relação à abordagem bayesiana quando lida com problemas similares pelo fato de apresentar soluções analíticas ao invés de numéricas. À medida em que o registro histórico aumenta, é fácil alterar as funções de distribuição de forma a ajustar de forma consistente os novos dados.

Conclui-se também é que, na prática, dificilmente se é capaz de dizer que dois determinados fatores de fato não inluenciam um no outro (são não interativos/independentes). O que ocorre é que, baseando-se em informações sobre tais fatores, pode-se ter indicações indiretas da relação que se estabelece entre eles.

Referências Bibliográficas

[1] L. C. Barros and C. Rodney. *Tópicos de Teoria Fuzzy e Biomatemática*, volume 5. Coleção IMECC - Textos Didáticos, 1 edition, 2006.

[2] Ellen Hisdal. Conditional possibilities independence and noninteraction. *Fuzzy Sets and Systems*, 1(4):283–297, 1978.

[3] Stéphane Lapointe and Bernard Bobée. Revision of possibility distributions: A bayesian inference pattern. *Fuzzy Sets and Systems*, 116(2):119–140, 2000.

[4] E. Massad, Neli R. S. Ortega, C. Barros, Laécio, and Cláudio José Struchiner. *Fuzzy Logic in Action: Applications in Epidemiology and Beyond*, volume 232 of *Studies in Fuzziness and Soft Computing*. Springer, 1 edition, 2008.

[5] C. Negoita, V. and D. Ralescu, A. *Applications of Fuzzy Sets to System Analysis*, volume 11 of *Interdisciplinary Systems Research*. Birkhäuser Verlag Basel, 1975.

[6] L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1(1):3–28, 1978.