

**UNIVERSIDADE FEDERAL DE SANTA CATARINA  
DEPARTAMENTO DE INFORMÁTICA E ESTATÍSTICA**

MunIQUE MittELMANN

**MODELO DE INFERÊNCIA EM REDES  
FUZZY-BAYESIANAS DINÂMICAS PARA  
TRATAMENTO DE INCERTEZAS**

Florianópolis

2019



MunIQUE MittelmANN

**MODELO DE INFERÊNCIA EM REDES  
FUZZY-BAYESIANAS DINÂMICAS PARA  
TRATAMENTO DE INCERTEZAS**

Dissertação submetida ao Programa  
de Pós Graduação em Ciência da Com-  
putação para a obtenção do Grau de  
Mestre em Ciência da Computação.  
Orientador: Prof. Aldo von Wange-  
nheim, Dr.  
Coorientadora: Prof<sup>a</sup>. Jerusa Marchi,  
Dr<sup>a</sup>.

Florianópolis

2019

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Mittelman, Munyque

Modelo de Inferência em Redes Fuzzy-Bayesianas Dinâmicas para Tratamento de Incertezas / Munyque Mittelman ; orientador, Aldo von Wangenheim, coorientadora, Jerusa Marchi, 2019.

111 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós Graduação em Ciência da Computação, Florianópolis, 2019.

Inclui referências.

1. Ciência da Computação. 2. Tratamento de Incerteza. 3. Rede Bayesiana Dinâmica. 4. Teoria Fuzzy. 5. Inferência Fuzzy-Bayesiana. I. von Wangenheim, Aldo. II. Marchi, Jerusa. III. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Ciência da Computação. IV. Título.

Munye Mittelmann

**MODELO DE INFERÊNCIA EM REDES  
FUZZY-BAYESIANAS DINÂMICAS PARA  
TRATAMENTO DE INCERTEZAS**

Esta Dissertação foi julgada aprovada para a obtenção do Título de “Mestre em Ciência da Computação”, e aprovada em sua forma final pelo Programa de Pós Graduação em Ciência da Computação.

Florianópolis, 24 de janeiro 2019.

---

Prof. José Luís Almada Güntzel, Dr.  
Coordenador do Curso

---

Prof. Aldo von Wangenheim, Dr.  
Orientador

**Banca Examinadora:**

---

Prof<sup>a</sup>. Jerusa Marchi, Dr<sup>a</sup>.  
Universidade Federal de Santa Catarina  
Presidente

---

Prof. Rafael de Santiago, Dr.  
Universidade Federal de Santa Catarina



---

Prof. Ricardo Azambuja Silveira, Dr.  
Universidade Federal de Santa Catarina

---

Prof. Eros Comunello, Dr.  
Universidade do Vale do Itajaí

---

Prof. Juliano Tonizetti Brignoli, Dr.  
Instituto Federal Catarinense



## AGRADECIMENTOS

Agradeço aos meus pais, Mauristela Aparecida Odorizzi Kossmann e Euzébio Mittelman, pelo carinho e apoio.

Ao meu noivo Tércio Lippel, pela compreensão, paciência e companheirismo.

Ao meu orientador, professor Aldo von Wangenheim, pelas oportunidades e pelas contribuições durante a elaboração deste trabalho.

À minha coorientadora, professora Jerusa Marchi, pela amizade, dedicação e motivação, que foram imprescindíveis para o desenvolvimento desta dissertação.

Por fim, agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior pelo apoio financeiro durante o mestrado.



*“I don't see why it's impossible to express  
everything that's on one's mind.”*

(Ivan Turguêniev)



## RESUMO

Diversos problemas computacionais requerem a realização de inferência sobre informações multissensoriais incertas. Redes Bayesianas (BN) são uma ferramenta para representação e inferência de variáveis que possuem incerteza devido a aleatoriedade. Redes Bayesianas Dinâmicas (DBN) estendem este conceito introduzindo dependências temporais entre variáveis, de modo a incorporar o reconhecimento de padrões temporais quando os dados possuem características sequenciais. Para modelar um ambiente com variáveis contínuas de forma eficiente com BN, é necessário discretizar os dados em categorias. Quando a discretização é especificada por um usuário especialista no domínio, os conceitos linguísticos utilizados para descrever as classes frequentemente possuem imprecisão por vagueza, isto é, não podem ser definidas adequadamente por limiares numéricos. A Teoria Fuzzy provê ferramentas para modelar a imprecisão por vagueza, permitindo a definição de classes com transições gradativas. Embora existam abordagens híbridas que integrem a Teoria Fuzzy no processo de inferência de BN, ainda existem limitações na utilização de observações *fuzzy* como evidências de uma DBN. Os trabalhos correlatos restringem a modelagem das evidências em particionamentos restritos a apenas duas funções de pertinência *fuzzy*. Para sobrepor esta limitação, este trabalho propõe um método de inferência Fuzzy-Bayesiana Dinâmica com evidências não-dicotômicas. O método modifica a inferência por filtragem de DBN, introduzindo qualificadores *fuzzy* para tratar a imprecisão por vagueza nas evidências da rede. É demonstrado que a inferência proposta respeita a propriedade da Soma das Probabilidades. Para verificar a proposta, o modelo foi aplicado como um classificador em bases de dados disponíveis da literatura e comparado com outras abordagens existentes. Os resultados obtidos mostram uma melhora na precisão e na acurácia da classificação, em relação às demais abordagens.

**Palavras-chave:** Tratamento de Incerteza. Rede Bayesiana Dinâmica. Teoria Fuzzy. Inferência Fuzzy-Bayesiana.



## ABSTRACT

Many computational problems require inference over incomplete or uncertain sensory data. Bayesian Networks (BN) are a tool for representing and inferring over variables with aleatory uncertainty. Dynamic Bayesian networks (DBN) extend this concept to introduce temporal dependencies that catch dynamic behaviors from the domain variables. An effective and efficient modeling through BN demands data discretization on categories. One of the main methods for discretizing continuous data for use in BN is manual specification by an expert user. However, the linguistic labels used by humans frequently have vagueness uncertainty, once they use labels that haven't exact numerical thresholds. Fuzzy Theory provides a framework for modeling vagueness uncertainty. Although there are hybrid theories that integrate Fuzzy Theory to the inference process in BN, there are still limitations on using fuzzy evidence on a DBN. The related works restrict the evidence modeling in partitioning with only two fuzzy membership functions. Thereby, this work proposes a method for Dynamic Fuzzy-Bayesian inference over not dichotomous variables. The proposed method modifies the filtering inference of DBN, introducing fuzzy qualifiers for dealing with the vagueness in the network evidence. It is demonstrated that the new inference respects the probability property. To evaluate the proposal, the model is applied as a classifier on open databases and compared with other approaches. The obtained results show an improvement in the accuracy of the classification, in comparison to the other approaches.

**Keywords:** Dealing with Uncertainty. Dynamic Bayesian Network. Fuzzy Theory. Fuzzy-Bayesian Inference.



## LISTA DE FIGURAS

Figura 1	Classificação da Incerteza, construída a partir de (THUNNISSEN, 2003; LI; CHEN; FENG, 2013) .....	32
Figura 2	Exemplo de Rede Bayesiana .....	36
Figura 3	Exemplo de Rede Bayesiana Dinâmica .....	38
Figura 4	<i>Core, support e boundaries</i> de uma função de pertinência <i>fuzzy</i> (ROSS, 2010) .....	41
Figura 5	Particionamento <i>fuzzy</i> , adaptado de (TOMASO; BALDWIN, 2008) .....	51
Figura 6	Rede Bayesiana para o problema <i>Wall-Following</i> .....	54
Figura 7	Modelo de Inferência Fuzzy-Bayesiana Dinâmica .....	58
Figura 8	Rede Bayesiana Dinâmica para o problema <i>Wall-Following</i> .....	62
Figura 9	Distribuição dos valores da Variável <i>Direction</i> no <i>Wall-Following Dataset</i> .....	65
Figura 10	Disposição dos Dados com $T$ fatias temporais .....	66
Figura 11	Matriz de Confusão para Classificação Binária (THARWAT, 2018) .....	68
Figura 12	Curva ROC, Adaptado de (THARWAT, 2018) .....	69
Figura 13	Modelagem Fuzzy para o <i>Occupancy Detection Dataset</i> .....	72
Figura 14	Topologia da DBN com 4 <i>time slice</i> para o <i>Occupancy Detection Dataset</i> .....	76
Figura 15	Curva ROC para DBN, Fuzzy-BN e Fuzzy-DBN para a classificação positiva de ocupação .....	77
Figura 16	Modelagem Fuzzy para o <i>Wall-Following Dataset</i> .....	77
Figura 17	Topologia da DBN com 7 <i>time slice</i> para o <i>Wall-Following Dataset</i> .....	78



## LISTA DE TABELAS

Tabela 1	Matriz de Confusão para a Variável Binária <i>Occupancy</i>	72
Tabela 2	Comparação da Classificação com Fuzzy-BN, DBN e Fuzzy-DBN ( <i>Occupancy Detection Dataset</i> )	73
Tabela 3	Matriz de Confusão para a Variável <i>Direction</i> (Fuzzy-BN)	74
Tabela 4	Comparação da Classificação com BN e Fuzzy-BN ( <i>Wall-Following Dataset</i> )	75
Tabela 5	Matriz de Confusão para a Variável <i>Direction</i> (Fuzzy-DBN com 3 <i>time slice</i> )	79
Tabela 6	Comparação da Classificação com DBN e Fuzzy-DBN e 3 <i>time slice</i> ( <i>Wall-Following Dataset</i> )	79
Tabela 7	Matriz de Confusão para a Variável <i>Direction</i> (Fuzzy-DBN com 5 <i>time slice</i> )	80
Tabela 8	Comparação da Classificação com DBN e Fuzzy-DBN e 5 <i>time slice</i> ( <i>Wall-Following Dataset</i> )	80
Tabela 9	Matriz de Confusão para a Variável <i>Direction</i> (Fuzzy-DBN com 7 <i>time slice</i> )	81
Tabela 10	Comparação da Classificação com DBN e Fuzzy-DBN e 7 <i>time slice</i> ( <i>Wall-Following Dataset</i> )	81
Tabela 11	Quantidade de Artigos recuperados em cada Estágio de Seleção	101
Tabela 12	Distribuição <i>a priori</i> da variável não-observável <i>Occupancy</i> no tempo inicial	105
Tabela 13	Modelo de Transição ( <i>Occupancy Detection Dataset</i> )	105
Tabela 14	Modelo de Sensores ( <i>Occupancy Detection Dataset</i> )	106
Tabela 15	Distribuição <i>a priori</i> da variável não-observável <i>Direction</i> no tempo inicial	107
Tabela 16	Modelo de Transição ( <i>Wall-Following Dataset</i> )	108
Tabela 17	Modelo de Sensores ( <i>Wall-Following Dataset</i> )	109



## LISTA DE QUADROS

Figura 1	Trabalhos Relacionados - Fuzzy-DBN.....	47
Figura 2	Trabalhos relacionados - DBN com Entradas Fuzzy....	50



## LISTA DE ABREVIATURAS E SIGLAS

BN	Rede Bayesiana ( <i>Bayesian Network</i> ) . . . . .	27
DBN	Rede Bayesiana Dinâmica ( <i>Dynamic Bayesian Network</i> ) . . . . .	27
CPT	Tabela de Probabilidade Condicional ( <i>Conditional Probability Table</i> ) . . . . .	36
<i>argmax</i>	Argumento de Máxima . . . . .	38
Fuzzy-DBN	Rede Fuzzy-Bayesiana Dinâmica . . . . .	45
UCI	UC Irvine Machine Learning Repository . . . . .	63
TP	Verdadeiro Positivo ( <i>True Positive</i> ) . . . . .	67
FP	Falso Positivo ( <i>False Positive</i> ) . . . . .	67
TN	Verdadeiro Negativo ( <i>True Negative</i> ) . . . . .	67
FN	Falso Negativo ( <i>False Negative</i> ) . . . . .	67
ROC	<i>Receiver Operating Characteristics</i> . . . . .	69
Fuzzy-BN	Inferência Fuzzy-Bayesiana . . . . .	71



## LISTA DE SÍMBOLOS

$\Omega$	Espaço Amostral .....	32
$P$	Função do valor de probabilidade de um evento ou conjunto de eventos .....	32
$\emptyset$	Conjunto vazio .....	33
$\bar{A}$	Negação de $A$ .....	33
$P(A)$	Probabilidade a priori do evento $A$ .....	33
$P(A B)$	Probabilidade condicional do evento $A$ dado a ocorrência do evento $B$ .....	33
$X^t$	Variável de estado não-observável no tempo $t$ .....	37
$E^t$	Variável de evidência observável no tempo $t$ .....	37
$X^{a:b}$	Conjunto de variáveis de $X^a$ até $X^b$ .....	37
$P(X^t X^{t-1})$	Modelo de transição .....	37
$P(E^t X^t)$	Modelo de sensores .....	37
$\mu_A$	Função de pertinência à classe $A$ .....	40
$\mu_A(x)$	Grau de pertinência de $x$ à classe $A$ .....	40
$\mathcal{P}$	Probabilidade sobre evidências imprecisas .....	42
$\alpha$	Constante de normalização .....	55
$\tilde{P}$	Possibilidade condicional de evidências <i>fuzzy</i> .....	56
$\forall$	Para todo .....	95



## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	27
1.1 DEFINIÇÃO DO PROBLEMA .....	28
1.2 OBJETIVOS .....	29
1.2.1 Objetivos Específicos .....	29
1.3 ESTRUTURA DO TRABALHO .....	30
<b>2 RACIOCÍNIO SOBRE INCERTEZA</b> .....	31
2.1 RACIOCÍNIO PROBABILÍSTICO .....	32
2.1.1 Teoria da Probabilidade .....	32
2.1.1.1 Teorema de Bayes .....	34
2.1.2 Redes Bayesianas .....	35
2.1.3 Redes Bayesianas Dinâmicas .....	37
2.1.3.1 Filtragem .....	39
2.2 RACIOCÍNIO FUZZY .....	39
2.2.1 Conjuntos Fuzzy .....	40
2.2.2 Inferência Fuzzy-Bayesiana .....	42
<b>3 TRABALHOS RELACIONADOS</b> .....	45
3.1 INFERÊNCIA POR MEIO DE REGRAS FUZZY .....	46
3.2 INFERÊNCIA POR FILTRAGEM DICOTÔMICA .....	48
3.3 PROBABILIDADE CONJUNTA DAS VARIÁVEIS NÃO-OBSERVÁVEIS .....	48
3.4 VISÃO GERAL DOS TRABALHOS RELACIONADOS ....	49
<b>4 INFERÊNCIA FUZZY-BAYESIANA DINÂMICA</b> ...	53
4.1 INDUÇÃO À INFERÊNCIA FUZZY-BAYESIANA DINÂMICA .....	54
4.2 MODELO DE INFERÊNCIA FUZZY-BAYESIANA DINÂMICA .....	57
4.3 PROVA DA PROPRIEDADE DA SOMA DAS PROBABILIDADES NA INFERÊNCIA FUZZY-BAYESIANA DINÂMICA .....	59
<b>5 METODOLOGIA DE AVALIAÇÃO</b> .....	63
5.1 BASES DE DADOS .....	63
5.2 TRATAMENTO DOS DADOS .....	64
5.3 TREINAMENTO E AVALIAÇÃO DA DBN .....	66
5.4 MÉTODOS PARA ANÁLISE DA CLASSIFICAÇÃO .....	67
<b>6 AVALIAÇÃO</b> .....	71
6.1 <i>OCCUPANCY DETECTION DATASET</i> .....	71
6.2 <i>WALL-FOLLOWING DATASET</i> .....	74

<b>7 CONSIDERAÇÕES FINAIS</b> .....	83
7.1 TRABALHOS FUTUROS .....	84
<b>REFERÊNCIAS</b> .....	85
<b>APÊNDICE A – Prova da Propriedade da Soma das Probabilidades na Inferência Fuzzy-Bayesiana</b> .....	95
<b>APÊNDICE B – Protocolo da Revisão Sistemática da Literatura</b> .....	99
<b>APÊNDICE C – Distribuições de Probabilidade para a DBN</b> .....	105

# 1 INTRODUÇÃO

Diversos problemas computacionais requerem a realização de inferência sobre informações multissensoriais incompletas ou incertas. Modelos probabilísticos, como as Redes Bayesianas (*Bayesian Network* - BN), são uma ferramenta para representação e inferência de variáveis que possuem incerteza devido a aleatoriedade inerente ao mundo real (TOMASO; BALDWIN, 2008). A partir do conhecimento obtido por observações, BN permitem inferir, por exemplo, a probabilidade de ocorrer precipitação forte, dadas as evidências de que o clima está quente e a umidade está alta.

Entretanto, BN não realizam o reconhecimento de padrões temporais. Esta tarefa é necessária quando os dados possuem características sequenciais, como os dados provenientes de séries temporais ou gerados por sistemas dinâmicos. Para abranger esta situação, as BN podem ser estendidas em Redes Bayesianas Dinâmicas (*Dynamic Bayesian Network* - DBN), que introduzem dependências temporais para capturar os comportamentos dinâmicos das variáveis do domínio (NADERPOUR; LU; ZHANG, 2013a).

Embora medições sensoriais usualmente assumam valores contínuos (*e.g.* números reais), existe uma limitação prática na implementação das BN, pois a maioria dos softwares e algoritmos não permite acomodar variáveis contínuas (NOJAVAN; QIAN; STOW, 2017). Para utilizar as BN de forma eficiente e obter eficácia na modelagem do ambiente, é necessário discretizar os dados em categorias (BEUZEN; MARSHALL; SPLINTER, 2018).

A discretização pode ser especificada manualmente por um usuário especialista no domínio (BEUZEN; MARSHALL; SPLINTER, 2018). Entretanto, os conceitos linguísticos utilizados por humanos frequentemente possuem imprecisão por vagueza, isto é, são classes que não podem ser definidas precisamente por limiares numéricos (LI; CHEN; FENG, 2013). Por exemplo, as classes frio, morno e quente são conceitos linguísticos vagos. A imprecisão por vagueza é frequentemente modelada pela Teoria Fuzzy (ZADEH, 1965).

Com o objetivo de permitir a inferência de BN sobre classes com vagueza, a partir dos estudos de Viertl (1987, 1989, 1995, 2008), diversos trabalhos investigaram modelos híbridos de inferência Fuzzy-Bayesiana. Dentre os quais, Brignoli (2001) apresenta um modelo híbrido Fuzzy-Bayesiano que realiza inferência por meio de uma Rede Bayesiana com variáveis difusas, cujos estados são organizados de forma

dicotômica<sup>1</sup>. Em Tibiriçá (2005), o modelo de Brignoli (2001) é ampliado para considerar a representação de uma Rede Bayesiana com variáveis de múltiplos estados. Entretanto, o processo de inferência continua sendo dicotômico, pois realiza a estimativa considerando apenas um estado da evidência e seu complemento.

Como extensão dos modelos apresentados em Brignoli (2001) e Tibiriçá (2005), em Brignoli (2013), Brignoli et al. (2015) há a inclusão de variáveis difusas não-dicotômicas durante o processo de inferência. O modelo de Brignoli (2013), Brignoli et al. (2015) para inferência Bayesiana com entradas difusas é baseado no princípio de superposição de estados.

As abordagens Fuzzy-Bayesianas citadas não permitem dependências temporais entre as variáveis do modelo. Assim, a integração da Teoria Fuzzy com DBN visa permitir a inferência sobre variáveis temporais em cenários com características *fuzzy*.

## 1.1 DEFINIÇÃO DO PROBLEMA

As pesquisas em Redes Fuzzy-Bayesianas Dinâmicas podem ser divididas em relação ao momento no qual ocorre a inclusão da Teoria Fuzzy às DBN. Este processo pode ocorrer antes (pré-processamento), após (pós-processamento) e/ou durante a inferência da DBN.

A integração da Teoria Fuzzy à DBN na etapa de pré-processamento é apresentada em Yao et al. (2015), Li e Gao (2016), Feng et al. (2017), He et al. (2018), Zheng e Chien (2018). Nesta abordagem, a Teoria Fuzzy é empregada com o objetivo de prover a parametrização de DBN. Por sua vez, as pesquisas que utilizam a Teoria Fuzzy no pós-processamento da DBN, aplicam Sistemas de Inferência Fuzzy sobre a distribuição de probabilidade estimada pela DBN (SEKKAS; ANAGNOSTOPOULOS; HADJIEFTHYMIADES, 2007; EL-SAWAH et al., 2007; SYKES, 2010; KHALID; WANG; NURUDEEN, 2013).

Ao calcular as probabilidades condicionais e *a posteriori*, as abordagens de pré e pós-processamento não abordam a vagueza presente em variáveis de evidências definidas por estados *fuzzy*. Deste modo, os estudos apresentados em Naderpour, Lu e Zhang (2013b, 2014b, 2014a), Tomaso e Baldwin (2008) apresentam modelos que integram Fuzzy e DBN durante o processo de inferência, o que permite que as variáveis de evidência da rede tenham estados *fuzzy*. Entretanto, estas pesquisas

---

<sup>1</sup>Neste texto, considera-se como variável dicotômica aquela que é dividida logicamente em apenas dois conceitos complementares.

restringem a modelagem das evidências em particionamentos *fuzzy*, os quais podem ter apenas duas funções de pertinência. Deste modo, não é possível que haja variáveis de evidência com estados não-dicotômicos.

Tendo em vista as limitações ao integrar observações *fuzzy* a DBN, este trabalho busca propor a inferência Fuzzy-Bayesiana Dinâmica com evidências não-dicotômicas. Para atingir este objetivo, a inferência Fuzzy-Bayesiana apresentada em Brignoli (2013) é integrada ao processo de filtragem em DBN.

Para avaliar a proposta, o modelo é aplicado como um classificador nas bases de dados abertas *Wall-Following Dataset* (FREIRE et al., 2009) e *Occupancy Dataset* (CANDANEDO; FELDHHEIM, 2016). Caso não haja adição de fatias temporais, o modelo proposto é equivalente à inferência Fuzzy-Bayesiana. Assim, investiga-se a hipótese de que, em domínios descritos por fenômenos dinâmicos, a introdução de dependências temporais na inferência Fuzzy-Bayesiana melhora sua capacidade de classificação. Em contrapartida, a proposta estende o processo de filtragem em DBN para incluir observações *fuzzy*. Caso as variáveis de evidência observáveis possuam pertinência total e exclusiva a apenas uma função de pertinência *fuzzy*, então a proposta é equivalente a DBN tradicional. Deste modo, também é investigada a hipótese de que a inclusão de variáveis de evidência *fuzzy* melhora a classificação da filtragem em DBN.

De acordo com a Teoria da Probabilidade (KOLMOGOROV, 1956), a distribuição de probabilidade de um espaço amostral deve somar um. Para garantir a validade desta propriedade pela inferência Fuzzy-Bayesiana Dinâmica proposta, é apresentada a prova de que a soma das probabilidades da variável de saída é igual a um.

## 1.2 OBJETIVOS

O objetivo geral deste trabalho é propor um modelo de inferência para filtragem em Redes Fuzzy-Bayesianas Dinâmicas com evidências não-dicotômicas, de modo a aumentar a acurácia em relação às abordagens existentes.

### 1.2.1 Objetivos Específicos

Este trabalho tem como objetivos específicos:

- (a) Definir o processo de inferência Fuzzy-Bayesiana Dinâmica para

filtragem;

- (b) Permitir a inclusão de evidências com estados não-dicotômicos na rede Fuzzy-Bayesiana Dinâmica;
- (c) Provar que a inferência Fuzzy-Bayesiana Dinâmica proposta respeita a Propriedade da Soma das Probabilidades;
- (d) Avaliar a filtragem proposta, aplicando-a como um classificador nas bases de dados abertas *Wall-Following Dataset* (FREIRE et al., 2009) e *Occupancy Dataset* (CANDANEDO; FELDHEIM, 2016).

### 1.3 ESTRUTURA DO TRABALHO

O trabalho está estruturado como segue: no Capítulo 2 é apresentada a fundamentação teórica da pesquisa acerca de raciocínio sobre incerteza, no Capítulo 3 são apresentados os trabalhos relacionados, no Capítulo 4 é apresentada a proposta de inferência Fuzzy-Bayesiana Dinâmica, no Capítulo 5 é descrita a metodologia empregada para a avaliação da proposta, no Capítulo 6 são descritos os experimentos e os resultados obtidos. Por fim, no Capítulo 7 são apresentadas as considerações finais desta pesquisa e as sugestões para trabalhos futuros.

## 2 RACIOCÍNIO SOBRE INCERTEZA

A incerteza representa a falta de conhecimento, que pode ser decorrente de diversas fontes, como a imprecisão, dependência do acaso, falta de confiança, informações conflitantes, dentre outros (THUNNISSEN, 2003; DUTT; KURIAN, 2013).

Diversas áreas, tais como a economia e a engenharia, precisam analisar ou manipular dados e/ou informações que contêm incerteza. Por este motivo, existem numerosas classificações de incerteza divergentes na literatura, conforme elucidado por Thunnissen (2003). Neste trabalho, considera-se a divisão da incerteza de acordo com sua origem: aleatória ou epistêmica, pela definição apresentada por Li, Chen e Feng (2013).

A incerteza por aleatoriedade representa a variabilidade natural do mundo físico, ou seja, reflete a aleatoriedade inerente na natureza, que existe independente do conhecimento humano sobre o domínio (LI; CHEN; FENG, 2013).

Por sua vez, a incerteza epistêmica surge da falta de conhecimento humano do mundo físico, incompletude de informações ou da falta de habilidade para medir e modelar o mundo físico (LI; CHEN; FENG, 2013). A classificação de Li, Chen e Feng (2013) não especifica claramente os tipos de incerteza epistêmica. Assim, a subdivisão da incerteza epistêmica empregada neste texto é baseada em Thunnissen (2003), que descreve três tipos de incerteza quando há incompletude de informações na modelagem computacional:

- (i) Vagueza, na qual a informação é definida imprecisamente em classes com limites mal-definidos;
- (ii) Não-especificidade, que ocorre quando há uma variedade de alternativas em determinada situação e todas são possíveis;
- (iii) Dissonância, que refere-se a existência de evidências conflitantes (parcial ou totalmente).

Este trabalho busca tratar a incerteza presente em fenômenos aleatórios, cujas variáveis são descritas por classes vagas. A classificação de incerteza utilizada nesta pesquisa é ilustrada pela Figura 1, que destaca os tipos de incerteza investigados.



Figura 1 – Classificação da Incerteza, construída a partir de (THUNNISSEN, 2003; LI; CHEN; FENG, 2013)

Uma das formas de tratar incerteza por aleatoriedade é através da Teoria da Probabilidade, que será apresentada na Seção 2.1. Por sua vez, a incerteza por vagueza é tratada por meio da Teoria dos Conjuntos Fuzzy, que será apresentada na Seção 2.2.

## 2.1 RACIOCÍNIO PROBABILÍSTICO

A incerteza por aleatoriedade ocorre quando é impossível determinar com certeza a incidência de um evento, mesmo com o conhecimento de casos anteriores (*e.g.* lançamento de um dado). A Teoria da Probabilidade mede a crença da ocorrência de um evento em termos de probabilidade e obtêm o valor probabilístico com base na estatística e na análise de experimentos aleatórios (LI; CHEN; FENG, 2013).

A Subseção 2.1.1 especifica os fundamentos da Teoria da Probabilidade e as Subseções 2.1.2 e 2.1.3 apresentam os modelos gráficos probabilísticos Redes Bayesianas e Redes Bayesianas Dinâmicas, que permitem inferência sobre variáveis aleatórias.

### 2.1.1 Teoria da Probabilidade

Seja  $\Omega$  um espaço amostral. Cada subconjunto de  $\Omega$  é chamado de evento. Uma função  $P$  representa o valor de probabilidade de um evento ou conjunto de eventos. Assumindo A e B como eventos, então uma função real  $P$  sobre  $\Omega$  deve satisfazer os axiomas básicos da

probabilidade, ou Axiomas de Kolmogorov (KOLMOGOROV, 1956):

1. Não-negatividade:  $P(A) \geq 0$ ;
2. Normalização:  $P(\Omega) = 1$ ;
3. Aditividade:  $P(A \cup B) = P(A) + P(B)$  quando  $A \cap B = \emptyset$ , onde  $\emptyset$  é o conjunto vazio.

Como  $A + \bar{A} = \Omega$  e dados os Axiomas 2 e 3, segue a seguinte propriedade (KOLMOGOROV, 1956):

$$P(A) + P(\bar{A}) = 1 \text{ ou } P(\bar{A}) = 1 - P(A) \quad (2.1)$$

Considerando o Axioma 2 e a existência de  $n$  eventos no espaço amostral  $\Omega$ , ou seja,  $\Omega = A_1 + \dots + A_n$ , então  $P(\Omega) = P(A_1) + \dots + P(A_n) = 1$ . Ou seja, a somatória das probabilidades de um espaço amostral deve ser 1 (NORVIG; RUSSELL, 2014):

$$\sum_{i=1}^n P(A_i) = 1 \quad (2.2)$$

Por conveniência, este texto utilizará o termo “Propriedade da Soma das Probabilidades” para referir-se à propriedade apresentada na Equação 2.2.

A probabilidade de um evento  $A$  na ausência de quaisquer outras informações é denominada probabilidade *a priori* ou probabilidade incondicional e é denotada por  $P(A)$ .

Segundo Kolmogorov (1956), a probabilidade condicional, denotada por  $P(A|B)$ , representa a probabilidade da ocorrência do evento  $A$ , dada a ocorrência do evento  $B$ . Se  $P(B) > 0$ , então:

$$P(A|B) = \frac{P(A \cup B)}{P(B)} \quad (2.3)$$

Pela Regra do Produto, pode-se definir a Equação 2.3 da seguinte forma:

$$P(A \cup B) = P(A|B)P(B) \quad (2.4)$$

Ou:

$$P(A \cup B) = P(B|A)P(A) \quad (2.5)$$

De acordo com Giambene (2005), dois eventos  $A$  e  $B$  são independentes se:

$$P(A \cup B) = P(A)P(B) \quad (2.6)$$

Se os eventos  $A$  e  $B$  são independentes então:

$$P(A|B) = P(A) \text{ e } P(B|A) = P(B) \quad (2.7)$$

Um conjunto de eventos  $B_1, B_2, \dots, B_n$  são mutuamente independentes sob o evento  $A$  se (DAWID, 1979):

$$P(B_1 \cap B_2 \cap \dots \cap B_n | A) = P(B_1 | A)P(B_2 | A) \dots P(B_n | A) \quad (2.8)$$

A distribuição de probabilidade de  $A$ , denotada por  $P(A)$  é a declaração das probabilidades de todos os valores possíveis que a variável aleatória  $A$  pode assumir. Uma distribuição de probabilidade conjunta denota as probabilidades de todas as combinações de valores de um conjunto de variáveis aleatórias. Quando uma distribuição de probabilidade conjunta abrange o conjunto completo de todas as variáveis utilizadas para descrever o mundo, é chamada de distribuição de probabilidade total (NORVIG; RUSSELL, 2014).

### 2.1.1.1 Teorema de Bayes

Com base nas Equações 2.4 e 2.5, é possível obter o Teorema de Bayes, que permite calcular a probabilidade de um evento de hipótese  $A$  dado o evento de evidência  $B$  (KOLMOGOROV, 1956):

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (2.9)$$

Se o espaço amostral  $\Omega$  pode ser dividido em uma quantidade finita de eventos mutuamente exclusivos  $A_1, A_2, \dots, A_n$  e se  $B$  é um evento com  $P(B) > 0$ , para cada  $A_i \in \Omega$ , então o Teorema da Probabilidade Total afirma que (KOLMOGOROV, 1956):

$$P(B) = \sum_{j=1}^n P(B|A_j)P(A_j) \quad (2.10)$$

Neste caso, a fórmula generalizada do Teorema de Bayes é (KOLMOGOROV, 1956):

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^n P(B|A_j)P(A_j)} \quad (2.11)$$

Nota-se que assim como na propriedade apresentada na Equação 2.2, a distribuição de probabilidade  $A$  dado  $B$  deve somar um, ou seja:

$$\sum_{j=1}^n P(A_j|B) = 1 \quad (2.12)$$

Na Equação 2.9, a expressão  $\frac{1}{P(B)}$  é constante independente do valor que  $A$  assumir. Esta expressão é chamada de constante de normalização  $\alpha$  para a distribuição, pois assegura que as entradas de  $P(A|B)$  tenham soma igual a 1.

Considerando a existência de  $k$  evidências combinadas, representadas por  $B_1, \dots, B_k$ , o Teorema de Bayes pode ser expresso como:

$$P(A_i|B_1 \cap \dots \cap B_k) = \frac{P(B_1 \cap \dots \cap B_k|A_i)P(A_i)}{P(B_1 \cap \dots \cap B_k)} \quad (2.13)$$

Assumindo que as evidências  $B_1, \dots, B_k$  são independentes entre si sob a hipótese  $A_i$ , então, pela Equação 2.8 (BERRAR, 2018):

$$P(A_i|B_1 \cap \dots \cap B_k) = \frac{P(B_1|A_i)\dots P(B_k|A_i)P(A_i)}{P(B_1 \cap \dots \cap B_k)} \quad (2.14)$$

Ou:

$$P(A_i|B_1 \cap \dots \cap B_k) = \frac{P(A_i) \prod_{l=1}^k P(B_l|A_i)}{P(B_1 \cap \dots \cap B_k)} \quad (2.15)$$

A Equação 2.15 é conhecida como Modelo Naive Bayes, sendo a forma geral do Teorema de Bayes com uma combinação de  $k$  evidências com independência condicional sob a hipótese  $A_i$ .

### 2.1.2 Redes Bayesianas

Uma Rede Bayesiana (BN) é um grafo orientado acíclico, cujos nós são um conjunto de variáveis aleatórias e cujas arestas conectam pares de nós, representando dependências entre variáveis. Se houver uma aresta do nó  $X$  até o nó  $Y$ ,  $X$  será denominado *pai* de  $Y$ . Cada nó  $X_i$  tem uma distribuição de probabilidade condicional  $P(X_i|Pais(X_i))$  que quantifica o efeito dos pais sobre o nó (PEARL, 1988). Os relacionamentos probabilísticos entre as variáveis são expressos usando Tabelas

de Probabilidade Condicional (*Conditional Probability Tables - CPT*) (NOJAVAN; QIAN; STOW, 2017).

A topologia da rede especifica o conjunto de nós e seus vínculos, representando os relacionamentos de independência condicional válidos no domínio. Uma aresta de X até Y representa que X tem uma influência direta sobre Y. Por exemplo, na BN ilustrada pela Figura 2, a variável aleatória Precipitação influencia nas variáveis Umidade e Temperatura.



Figura 2 – Exemplo de Rede Bayesiana

A BN representa a distribuição de probabilidade conjunta total. Uma entrada genérica na distribuição de probabilidade conjunta é a probabilidade de uma conjução de atribuições específicas a cada variável e pode ser calculada pela seguinte fórmula:

$$P(x_1, \dots, x_n) = \prod_{i=1}^n P(x_i | \text{pais}(X_i)) \quad (2.16)$$

Na qual  $\text{pais}(X_i)$  denota os valores específicos das variáveis em  $\text{Pais}(X_i)$ .

Quando uma BN implementa o Modelo Naïve Bayes<sup>1</sup> (e.g. Figura 2), seus nós folhas são denominados variáveis de evidência, enquanto que uma variável que seja pai de uma variável de evidência é denominada variável de hipótese ou de saída.

As BN possuem diversas características que as tornam úteis para modelagem de problemas, como o baixo custo computacional, capacidade de lidar com falta de dados, dados de diferentes fontes e inclusão explícita de incertezas (BEUZEN; MARSHALL; SPLINTER, 2018).

---

<sup>1</sup>Ver Equação 2.15.

### 2.1.3 Redes Bayesianas Dinâmicas

Uma Rede Bayesianas Dinâmica (DBN) é uma forma de estender Redes Bayesianas para modelar distribuições de probabilidade sobre coleções semi-infinitas de variáveis aleatórias (DEAN; KANAZAWA, 1989; MURPHY, 2002).

Cada fatia de tempo de uma DBN pode ter qualquer número de variáveis de estados não-observáveis  $X^t$  no tempo  $t$  e variáveis de evidência observáveis  $E^t$ . A notação  $X^{a:b}$  denota o conjunto correspondente de variáveis desde  $X^a$  até  $X^b$ . Para construir uma DBN, deve-se especificar a distribuição anterior sobre as variáveis de estados no tempo inicial  $P(X^0)$ ; o modelo de transição  $P(X^t|X^{t-1})$  e o modelo de sensores ou observações  $P(E^t|X^t)$  (MIHAJLOVIC; PETKOVIC, 2001).

O modelo de transição  $P(X^t|X^{t-1})$  descreve como as variáveis não-observáveis  $X^t$  são afetadas pelos estados das variáveis não-observáveis nos tempos anteriores, isto é, nos estados  $X^{0:t-1}$ . Seguindo esta definição, cada nodo de variáveis  $X^t$  poderia envolver um número ilimitado de nodos pais. O problema de manipular uma quantidade potencialmente infinita de pais, é resolvido considerando a hipótese de Markov. A hipótese de Markov afirma que o estado atual depende apenas de um histórico finito de estados anteriores. Em um processo de Markov de primeira ordem, o estado atual depende apenas do estado anterior, ou seja (WANG et al., 2017):

$$P(X^t|X^{0:t-1}) = P(X^t|X^{t-1}) \quad (2.17)$$

O modelo de sensores  $P(E^t|X^t)$  descreve como as variáveis de evidência observáveis (sensores) são afetadas pelo estado real do mundo. Para especificar os modelos de transição e de sensores, deve-se especificar a topologia das conexões entre fatias sucessivas e entre as variáveis de estados não-observáveis e de evidências.

Assumindo a hipótese de que as mudanças no estado do mundo são causadas por um processo estacionário, os modelos de transição e de sensores são os mesmos para todo  $t$  (WANG et al., 2017).

Considerando-se, que a DBN descreve um processo de Markov de primeira ordem, a distribuição conjunta completa das  $T$  variáveis de estados não-observáveis  $X = \{X^0, \dots, X^{T-1}\}$  e das  $T$  variáveis observáveis  $E = \{E^0, \dots, E^{T-1}\}$ , onde  $T$  é o limite de tempo para o evento que está sendo investigado, é dada por (MIHAJLOVIC; PETKOVIC, 2001):

$$P(X, E) = P(X^0) \prod_{i=1}^{T-1} P(X^i | X^{i-1}) \prod_{i=0}^{T-1} P(E^i | X^i) \quad (2.18)$$

A Figura 3 exemplifica uma DBN para uma situação na qual um agente possui apenas sensores de umidade e de temperatura e precise determinar o nível de precipitação. Esta DBN assume a hipótese de que o processo que descreve as relações entre a variável não-observável Precipitação e as variáveis observáveis Umidade e Temperatura é estacionário e, também, um processo de Markov de primeira ordem.

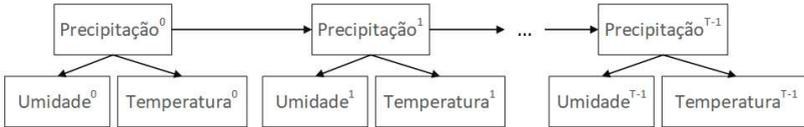


Figura 3 – Exemplo de Rede Bayesiana Dinâmica

Os principais tipos (ou tarefas) de inferência em DBN são a filtragem, a previsão, a suavização e a explicação mais provável (MIHAJLOVIC; PETKOVIC, 2001; MURPHY, 2002; TOMASO; BALDWIN, 2008):

- A filtragem consiste em calcular a distribuição posterior sobre o estado atual, dada toda a evidência até o momento, ou seja  $P(X^t | e^{0:t})$ ;
- A previsão corresponde a calcular a distribuição posterior sobre o estado futuro, dada toda a evidência até o momento, ou seja  $P(X^{t+k} | e^{0:t})$  para algum  $k > 0$ ;
- A suavização calcula a distribuição posterior sobre um estado passado, dada toda a evidência até o presente, ou seja,  $P(X^k | e^{0:t})$ , para algum  $k$  tal que  $0 \leq k < t$ ;
- Dada uma sequência de observações, a explicação mais provável calcula  $\operatorname{argmax}_{x_{0:t}} P(x_{0:t} | e^{0:t})$ , ou seja, a sequência de estados que mais provavelmente gerou estas observações<sup>2</sup>.

Dada uma série de observações, esta pesquisa tem o objetivo de permitir a inferência sobre a situação atual do ambiente. Ou seja,

<sup>2</sup>O argumento de máxima (*argmax*) é o ponto do domínio de alguma função no qual o valor da função é maximizado.

realizar a tarefa de inferência filtragem. Por este motivo, esta tarefa de inferência é especificada na Subseção 2.1.3.1.

### 2.1.3.1 Filtragem

A filtragem pode ser estimada por meio de avaliação recursiva: dado o resultado da filtragem até o tempo  $t - 1$  e com uma nova evidência  $e^t$ , é possível estimar o resultado em  $t$ .

Assumindo a hipótese de Markov de primeira ordem sobre os estados não-observáveis e estados de evidência, a estimação de  $P(X^t|e^{0:t})$  é definida por (MURPHY, 2002):

$$P(X^t|e^{0:t}) = \alpha P(e^t|X^t) \sum_{x^{t-1}} P(X^t|X^{t-1})P(x^{t-1}|e^{0:t-1}) \quad (2.19)$$

na qual  $P(e^{0:t}|X^t)$  pode ser obtida diretamente a partir do modelo de sensores. No somatório,  $P(X^t|X^{t-1})$  é o modelo de transição e  $P(x^{t-1}|e^{0:t-1})$  é a distribuição de estados atual, ou seja, o passo recursivo.

É possível visualizar o processo de filtragem em duas partes (MIHAJLOVIC; PETKOVIC, 2001; MURPHY, 2002):

- (i) na etapa conhecida como *forward*, a resolução recursiva de  $\sum_{x^{t-1}} P(X^t|X^{t-1})P(x^{t-1}|e^{0:t-1})$  projeta a distribuição de estados atuais de  $t - 1$  para  $t$ .
- (ii) na etapa de atualização, a resolução da Equação 2.19 atualiza a distribuição de estados por meio da utilização da nova evidência  $e^t$ .

Embora a Teoria da Probabilidade e DBN forneçam suporte para inferir sobre fenômenos dinâmicos com incerteza por aleatoriedade, elas não lidam com a imprecisão por vagueza, que é apresentada na Seção 2.2.

## 2.2 RACIOCÍNIO FUZZY

A incerteza por vagueza representa a imprecisão contida nos conceitos linguísticos usados por humanos (*e.g.* alto, pequeno, distante), os quais não podem ser precisamente definidos por limiares numéricos.

A Teoria Fuzzy, proposta por Zadeh (1965), busca reduzir a dissônança cognitiva entre a conceitualização linguística e sua representação computacional (LI; CHEN; FENG, 2013).

Um conjunto *crisp* representa classes com limites precisos, cuja pertinência pode ser apenas verdadeira ou falsa. Por sua vez, um conjunto *fuzzy* expressa uma classe linguística por meio de graus de pertinência contínuos, que representam a transição gradual entre a não-pertinência e a pertinência completa ao conjunto (ZADEH, 1965; LI; CHEN; FENG, 2013). A Subseção 2.2.1 apresenta as definições básicas de conjuntos *fuzzy* e a Subseção 2.2.2 descreve uma abordagem para realizar inferência Bayesiana sobre evidências difusas.

### 2.2.1 Conjuntos Fuzzy

Seja  $X$  um espaço de pontos (ou objetos), com um elemento genérico de  $X$  denotado por  $x$ . Um conjunto fuzzy (ou classe)  $A$  em  $X$  é caracterizado por uma função de pertinência  $\mu_A : X \rightarrow [0, 1]$ , que associa cada ponto em  $X$  com um número real no intervalo  $[0, 1]$ , com o valor de  $\mu_A(x)$  representando o “grau de pertinência” de  $x$  em  $A$ . Assim, quanto mais próximo  $\mu_A(x)$  estiver de 1, maior o grau de pertinência de  $x$  em  $A$ . Quando  $A$  é um conjunto *crisp*, a função de pertinência pode assumir apenas os valores 0 ou 1, com  $\mu_A(x) = 1$  ou  $\mu_A(x) = 0$  representando, respectivamente, a pertinência ou a não-pertinência de  $x$  ao conjunto  $A$  (ZADEH, 1965).

As principais operações sobre conjuntos *fuzzy* são apresentadas a seguir:

- Um conjunto *fuzzy* está vazio se e somente se sua função de pertinência é identicamente zero em  $X$ .
- Dois conjuntos *fuzzy*  $A$  e  $B$  são iguais, denotado por  $A = B$ , se e somente se  $\mu_A(x) = \mu_B(x) \forall x \in X$ .
- O complemento de um conjunto *fuzzy*  $A$  é denotado por  $A'$  e é definido por  $\mu_{A'}(x) = 1 - \mu_A(x)$ .
- A união de dois conjuntos *fuzzy*  $A$  e  $B$ , com as respectivas funções de pertinência  $\mu_A(x)$  e  $\mu_B(x)$  é um conjunto *fuzzy*  $C$ , denotado por  $C = A \cup B$ , cuja função de pertinência é definida por  $\mu_C(x) = \text{Mínimo}[\mu_A(x), \mu_B(x)]$ ,  $x \in X$ .
- A intersecção de dois conjuntos *fuzzy*  $A$  e  $B$ , com as respectivas funções de pertinência  $\mu_A(x)$  e  $\mu_B(x)$  é um conjunto *fuzzy*  $C$ ,

denotado por  $C = A \cap B$ , cuja função de pertinência é definida por  $\mu_C(x) = \text{Máximo}[\mu_A(x), \mu_B(x)]$ ,  $x \in X$ .

Dado um conjunto *fuzzy*  $A$ , Ross (2010) apresenta as seguintes definições referentes a uma função de pertinência  $\mu_A$  *fuzzy*:

- O núcleo (*core*) de  $\mu_A$  é definido como a região do universo  $X$  que é caracterizada pela completa pertinência ao conjunto  $A$ . Ou seja, o núcleo de  $\mu_A$  corresponde aos elementos  $x \in X$  tal que  $\mu_A(x) = 1$ .
- O suporte (*support*) de  $\mu_A$  é definido como a região do universo  $X$  que é caracterizada por terem grau de pertinência não-zero ao conjunto  $A$ . Ou seja, o suporte de  $\mu_A$  corresponde aos elementos  $x \in X$  tal que  $\mu_A(x) > 0$ .
- Os limites (*boundaries*) de  $\mu_A$  são definidos como as regiões do universo  $X$  que contêm elementos sem pertinência completa ou não-pertinência ao conjunto  $A$ . Ou seja, os limites de  $\mu_A$  correspondem aos elementos  $x \in X$  tal que  $0 < \mu_A(x) < 1$ .

As definições de núcleo, suporte e limites de uma função de pertinência *fuzzy* são ilustradas na Figura 4.

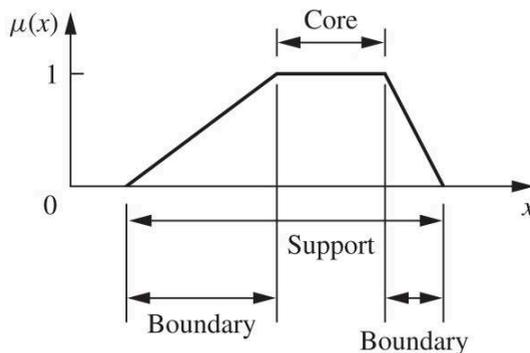


Figura 4 – *Core*, *support* e *boundaries* de uma função de pertinência *fuzzy* (ROSS, 2010)

### 2.2.2 Inferência Fuzzy-Bayesiana

Variáveis difusas usualmente não são descritos pela inferência Bayesiana padrão e, assim, é necessário modelá-los de modo a incorporar seu aspecto *fuzzy* antes de analisá-los por métodos estatísticos (VIERTL, 1987, 2008). Os estudos de Viertl (1987, 1989, 1995, 2008), investigam a generalização da estatística Bayesiana para variáveis difusas. Tais abordagens híbridas são conhecidas pelos termos inferência Fuzzy-Bayesiana ou modelo Difuso-Probabilístico.

Em Brignoli (2001) é apresentado um modelo híbrido Fuzzy-Bayesiano que infere por meio de uma Rede Bayesiana com variáveis difusas, cujos estados são organizados de forma dicotômica. Em Tibiriçá (2005), o modelo de Brignoli (2001) é ampliado de modo a considerar a representação de uma Rede Bayesiana com variáveis de múltiplos estados, entretanto a inferência continua sendo dicotômica.

Como extensão dos modelos apresentados por Brignoli (2001) e Tibiriçá (2005), em Brignoli (2013), Brignoli et al. (2015) há a inclusão de variáveis difusas não-dicotômicas. O modelo de Brignoli (2013) para inferência Bayesiana com entradas difusas é baseado no princípio de superposição de estados. A probabilidade  $\mathcal{P}$  sobre evidências imprecisas é definida pela seguinte equação:

$$\mathcal{P}(H_i|E_j) = \frac{P(H_i) \prod_{j=1}^m \sum_{k=1}^u (P_{x_k}(E_j|H_i) \cdot \mu_{x_k}(E_j))}{\sum_{l=1}^n P(H_l) \prod_{j=1}^m \sum_{k=1}^u (P_{x_k}(E_j|H_l) \cdot \mu_{x_k}(E_j))} \quad (2.20)$$

na qual,  $n$  é a quantidade de hipóteses;  $m$  é a quantidade de evidências;  $u$  é a quantidade de estados de uma evidência;  $H_i$  é o vetor de hipóteses, com  $1 \leq i \leq n$ ;  $E_j$  é o vetor de evidências, com  $1 \leq j \leq m$ ;  $P_{x_k}(E_j|H_i)$  é a probabilidade condicional do estado  $x_k$  da evidência  $E_j$  dado  $H_i$ , com  $1 \leq k \leq u$ ;  $P(H_i)$  é a probabilidade *a priori* sobre  $H_i$ ; e  $\mu_{x_k}(E_j)$  representa o grau de pertinência do estado  $x_k$  da evidência  $E_j$  em relação a função de pertinência difusa.

Assim,  $\mathcal{P}(H_i|E_j)$  é a probabilidade condicional de  $H_i$  ajustada pela imprecisão sobre as variáveis de evidência  $E_j$ .

Apesar de realizar simulações nas quais é mantida a Propriedade da Soma das Probabilidades<sup>3</sup>, em Brignoli (2013), Brignoli et al. (2015) não é demonstrado que inferência Fuzzy-Bayesiana proposta respeita esta propriedade em qualquer caso. Desta forma, no Apêndice A apresenta-se a prova direta de que a inferência Fuzzy-Bayesiana

---

<sup>3</sup>Ver Equação 2.2.

respeita a Propriedade da Soma das Probabilidades.

As abordagens híbridas Fuzzy-Bayesianas não possuem dependências temporais entre as variáveis da rede, assim, não é possível capturar comportamentos dinâmicos como ao utilizar DBN. Deste modo, o Capítulo 3 apresenta abordagens que integram Teoria Fuzzy e DBN.



### 3 TRABALHOS RELACIONADOS

Este Capítulo brevemente revisa os trabalhos relacionados que integram Lógica Fuzzy a Redes Bayesianas Dinâmicas (*Dynamic Bayesian Networks* - DBN), de modo a identificar as abordagens usadas e propor, em contrapartida, a inferência sobre Redes Fuzzy-Bayesianas Dinâmicas (Fuzzy-DBN).

Por meio de uma revisão sistemática elaborada de acordo com a metodologia de Kitchenham (2004)<sup>1</sup>, foram identificadas três abordagens gerais para incorporar Fuzzy à DBN: (i) explorando parâmetros Fuzzy da rede (pré-processamento); (ii) usando um Sistema de Inferência Fuzzy na saída da rede (pós-processamento); e (iii) utilizando Entradas Fuzzy como as evidências observadas pela rede (durante a inferência da DBN).

Na abordagem de pré-processamento, a Lógica Fuzzy é utilizada de diferentes modos com o objetivo de prover parâmetros para uma DBN. Em Kao e Chen (2014), a estimação *fuzzy* é empregada para prover os parâmetros temporais da DBN, de modo a identificar relações temporais estacionárias, durante a aprendizagem *data-driven* de parâmetros da DBN. Por sua vez, Feng et al. (2017), Zhang et al. (2018) optam por fuzzificar as opiniões de especialistas do domínio. Como números *fuzzy* não podem ser diretamente aplicados na rede, em Feng et al. (2017), as opiniões dos especialistas são convertidos em valores crisp antes de serem usados para obter os parâmetros. Já em Zhang et al. (2018), a probabilidade condicional é definida pelo limite superior mínimo do conjunto de opiniões *fuzzy* dos especialistas.

Outra abordagem de pré-processamento, apresentada por Zaiyi (2018), compara o grau de pertinência *fuzzy* como um *threshold* para que os valores observados sejam considerados como pertencentes a determinado estado de um nodo de evidência. Em Yao et al. (2015), Chen e Gao (2009), Yang, Gao e Chen (2010), Li e Gao (2016), o grau de pertinência *fuzzy* é utilizado como probabilidade *a priori* para descrever a distribuição de probabilidade condicional da rede. De forma similar, em Zheng e Chien (2018), He et al. (2018) são utilizados números triangulares *fuzzy* ao definir a probabilidade da ocorrência de eventos que sejam definidos por expressões linguísticas. Contudo, nota-se que ao calcular as probabilidades condicionais e *a posteriori*, as entradas são consideradas como evidências confirmadas, ao invés de Entradas Fuzzy.

A abordagem de pós-processamento usa Sistemas de Inferência

---

<sup>1</sup>O protocolo da Revisão Sistemática é apresentado no Apêndice B.

Fuzzy e é empregada em Anagnostopoulos, Sekkas e Hadjiefthymiades (2007), Sekkas, Anagnostopoulos e Hadjiefthymiades (2007), Naderpour, Lu e Zhang (2013b, 2014b, 2014a), El-Sawah et al. (2007). Inicialmente, a DBN estima as probabilidades condicionais. Em seguida, conjuntos *fuzzy* são usados para caracterizar cada valor de probabilidade por meio de um conjunto de termos linguísticos. Por fim, regras de inferência *fuzzy* são usadas para inferir o resultado. De modo similar, em Sykes (2010), Khalid, Wang e Nurudeen (2013) é proposto o uso de uma Rede Neuro-Fuzzy para o Sistema de Inferência Fuzzy.

A terceira abordagem para incorporar Fuzzy a DBN é permitir evidências *fuzzy*, isto é, valores *fuzzy* como as entradas da rede e utilizá-las durante a inferência da DBN. Assim, os estados de um nodo de evidência não são divididos por valores de *threshold*, mas definidos por funções de pertinência *fuzzy*. Desta maneira, é possível tratar a vagueza existente em variáveis linguísticas que possivelmente descrevem as evidências observadas (*e.g.* evidência de dor de cabeça alta ou baixa).

A aplicação de Fuzzy-DBN ocorre em uma ampla variedade de campos de aplicação, como por exemplo: no rastreamento de mãos e reconhecimento de gestos (EL-SAWAH et al., 2007), na avaliação de situações de risco (NADERPOUR; LU; ZHANG, 2013b, 2014b, 2014a), dentre outros. O Quadro 1 sumariza os trabalhos relacionados sobre Fuzzy-DBN, com o objetivo de comparar e situar a proposta desta pesquisa, mostrando a abordagem para incorporar Fuzzy e DBN. Neste trabalho, busca-se propor um modelo de inferência Fuzzy-Bayesiano Dinâmico com evidências *fuzzy*. Assim, a próxima subseção apresenta características dos estudos que utilizam esta abordagem, com o objetivo de distingui-los da proposta.

### 3.1 INFERÊNCIA POR MEIO DE REGRAS FUZZY

Na abordagem de Teixeira e Zaverucha (2003, 2004a, 2004b, 2005) para previsão em DBN, a probabilidade de uma variável *fuzzy* é igual a expectativa da função de pertinência, que é estimada a partir da amostra de dados. As probabilidades *fuzzy* são estimadas para cada saída (nodos de hipótese em um dado momento) e cada atributo de entrada *fuzzy* (nodos de evidência).

A abordagem faz um paralelo com regras *fuzzy*, nas quais se diferentes exemplos geram regras com as mesmas precondições, é mantida apenas a regra com maior grau. Assim, é considerada uma “fórmula

Quadro 1 – Trabalhos Relacionados - Fuzzy-DBN

Referência	Abordagem para incorporar Fuzzy a DBN		
	(i) Pré-processamento	(ii) Pós-processamento	(iii) Durante a inferência
Teixeira e Zaverucha (2003, 2004a, 2004b, 2005)			X
Hua, Rui e Jizhou (2006), Zhang e Li (2006)			X
Anagnostopoulos, Sekkas e Hadjiefthymiades (2007), Sekkas, Anagnostopoulos e Hadjiefthymiades (2007)		X	
El-Sawah et al. (2007)		X	
Tomaso e Baldwin (2008)			X
Chen e Gao (2009), Yang, Gao e Chen (2010)	X		
Wang, Gao e Thompson (2009)			X
Sykes (2010)		X	
Wu et al. (2012)			X
Khalid, Wang e Nurudeen (2013)		X	
Naderpour, Lu e Zhang (2013b, 2014b, 2014a)		X	X
Kao e Chen (2014)	X		
Yao et al. (2015)	X		
Li e Gao (2016)	X		
Feng et al. (2017)	X		
Jinyong, Keke e Wenjing (2017)			X
Zaiyi (2018)	X		
He et al. (2018)	X		
Zhang et al. (2018)	X		
Zheng e Chien (2018)	X		
Esta proposta			X

de filtragem”, similar a uma regra *fuzzy*. Isto significa que ao computar as probabilidades *fuzzy*, cada variável *fuzzy* assume o grau de pertinência para uma função ativa (assim como ocorre em uma regra *fuzzy* como “*if X is Baixo*”). Deste modo, apenas um estado está sendo considerado durante a inferência, o que possivelmente leva a perda de informação/precisão pois o grau de pertinência de cada estado não é considerado simultaneamente.

Em Zhang e Li (2006), Hua, Rui e Jizhou (2006), esta estratégia é aplicada com o objetivo de construir um modelo de personalidade-emoção que simula emoções e comunicações humanas.

### 3.2 INFERÊNCIA POR FILTRAGEM DICOTÔMICA

Em Naderpour, Lu e Zhang (2013b, 2014b, 2014a) é apresentada uma abordagem de filtragem para avaliação de situação baseada em Fuzzy-DBN para auxiliar operadores no processo de decisão de operadores em situações arriscadas. As observações sensoriais são discretizadas por um método quantizador *fuzzy* proposto por Chai e Wang (2011). O método quantizador *fuzzy* considera apenas variáveis *fuzzy* com dois estados possíveis. A CPT é construída estimando a pertinência *fuzzy* do primeiro estado e seu complemento é considerado a pertinência *fuzzy* do segundo estado. Ou seja, o processo de inferência é dicotômico, pois só considera dois estados, para cada evidência, durante o processo de inferência.

De forma similar, na filtragem apresentada por Tomaso e Baldwin (2008), são definidas *soft evidence* como particionamentos *fuzzy* que sejam igualmente espaçados, com conjuntos *fuzzy* simétricos e apenas duas funções de pertinência. Os valores de pertinência a estas funções definem a distribuição de probabilidade sobre os elementos da partição.

### 3.3 PROBABILIDADE CONJUNTA DAS VARIÁVEIS NÃO-OBSERVÁVEIS

Em Wang (2009), Wu et al. (2012), Jinyong, Keke e Wenjing (2017)<sup>2</sup> conjuntos *fuzzy*, cuja modelagem não é discutida pelos autores, são estabelecidos de acordo com o estado discreto das variáveis da rede. Os valores contínuos observados são classificados por funções *fuzzy* para obter o grau de pertinência de cada observação para cada conjunto *fuzzy*.

Uma Rede Bayesiana Estática Discreta com nodos não-observáveis e nodos de observação (evidências) é aplicada com as propriedades de independência condicional das Redes Bayesianas. Para obter a fórmula de inferência da Fuzzy-DBN, após fuzzificar as observações contínuas, a distribuição de probabilidade conjunta a *posteriori* das variáveis não-observáveis é estimada. Então, aplica-se o princípio da probabilidade a esta distribuição e, por fim, o valor resultante é ponderado pelo produto dos graus de pertinência *fuzzy* de cada estado. Os autores não consideram a situação de estados Não-Dicotômicos, mas a abordagem suporta múltiplos estados e considera todos na estimação das probabi-

---

<sup>2</sup>As pesquisas mencionadas são baseadas em Shi, Gao e Li (2006), que pertence à base China Academic Journals (CNKI) e não possui acesso livre.

lidades.

Diferentemente dos trabalhos citados, que multiplicam a probabilidade das variáveis escondidas sendo confirmadas as observações pelo produtório dos graus de pertinência das observações, a proposta apresentada neste trabalho pondera cada probabilidade condicional dos estados de evidências ocorrerem sendo confirmada uma hipótese com seu respectivo grau de pertinência *fuzzy* e usa estes valores para estimar a probabilidade das hipóteses dadas as observações. Assim, a probabilidade condicional de cada estado é avaliada em relação à sua pertinência ao conjunto.

### 3.4 VISÃO GERAL DOS TRABALHOS RELACIONADOS

Como introduzido na Seção 2.1.3 a hipótese de Markov, que afirma que o estado atual depende apenas de uma quantidade finita de estados anteriores, pode ser utilizada em DBN para solucionar o problema de manipular uma quantidade potencialmente infinita de nodos pais. Por este motivo, esta premissa é assumida por Tomaso e Baldwin (2008), Teixeira e Zaverucha (2003, 2004a, 2004b, 2005) e Zhang e Li (2006), Hua, Rui e Jizhou (2006).

Além disto, Tomaso e Baldwin (2008) e Teixeira e Zaverucha (2003, 2004a, 2005) assumem a hipótese de que as mudanças no estado do mundo são causadas por um processo estacionário, ou seja, são invariantes com o tempo.

As tarefas de inferência realizadas pelos trabalhos relacionados são: (i) a previsão, por Teixeira e Zaverucha (2003, 2004a, 2004b, 2005) e Zhang e Li (2006), Hua, Rui e Jizhou (2006); (ii) a filtragem, por Tomaso e Baldwin (2008) e Naderpour, Lu e Zhang (2013b); e (iii) a explicação mais provável, por Tomaso e Baldwin (2008) e Naderpour, Lu e Zhang (2014a). Ao invés de definirem tarefas de inferência, os trabalhos apresentados por Jinyong, Keke e Wenjing (2017), Wang (2009), Wu et al. (2012), apresentam a probabilidade conjunta das variáveis não-observáveis. O Quadro 2 compara este e outros aspectos dos trabalhos relacionados com a proposta deste trabalho.

Para permitir múltiplos estados nas variáveis de evidência *fuzzy*, os trabalhos de Teixeira e Zaverucha (2003, 2004a, 2004b, 2005), Zhang e Li (2006), Hua, Rui e Jizhou (2006), Tomaso e Baldwin (2008) e Naderpour, Lu e Zhang (2013b, 2014b, 2014a) baseiam-se em particionamentos do universo de discurso *fuzzy* de cada evidência em regiões ou partições.

Quadro 2 – Trabalhos relacionados - DBN com Entradas Fuzzy

Referência	Inferência Não-Dicotômica	Modelagem Não-Dicotômica de Variáveis	Tarefa de Inferência ou Probabilidade Estimada
Teixeira e Zaverucha (2003, 2004a, 2004b, 2005)	Não	Não	Previsão
Hua, Rui e Jizhou (2006), Zhang e Li (2006)	Não	Não	Previsão
Tomaso e Baldwin (2008)	Não	Não	Filtragem e Explicação mais provável
Naderpour, Lu e Zhang (2013b, 2014b, 2014a)	Não	Não	Filtragem e Explicação mais provável
Jinyong, Keke e Wenjing (2017)	Não	Indefinido	Probabilidade conjunta das variáveis não-observáveis
Wang, Gao e Thompson (2009)	Não	Indefinido	Probabilidade conjunta das variáveis não-observáveis
Wu et al. (2012)	Não	Indefinido	Probabilidade conjunta das variáveis não-observáveis
Esta proposta	Sim	Sim	Filtragem

Neste sentido, uma partição *fuzzy*, ilustrada na Figura 5, no universo  $\Omega$ , é o conjunto de conjuntos *fuzzy*  $f_1, \dots, f_p$  tal que  $\forall x \in \Omega$ ,

$$\sum_{i=1}^p \mu_{f_i}(x) = 1 \quad (3.1)$$

Na qual  $\mu_{f_i}$  é a função de pertinência de  $f_i$  (TOMASO; BALDWIN, 2008). Logo, a utilização de particionamentos *fuzzy* restringe a modelagem das variáveis a intervalos compostos apenas por funções de pertinência complementares, cuja soma deve ser igual a 1.

Além desta restrição na modelagem das evidências, os trabalhos relacionados definem o processo de inferência de forma dicotômica. Como citado na Seção 3.1, as pesquisas Teixeira e Zaverucha (2003, 2004a, 2004b, 2005) e Zhang e Li (2006), Hua, Rui e Jizhou (2006)

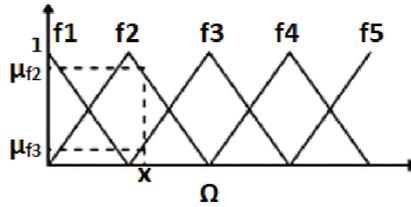


Figura 5 Particionamento *fuzzy*, adaptado de (TOMASO; BALDWIN, 2008)

utilizam apenas o maior grau de pertinência durante o processo de inferência. A inferência de Naderpour, Lu e Zhang (2013b, 2013a, 2014a) considera o grau de pertinência *fuzzy* de um estado de evidência e seu respectivo complemento, desconsiderando a existência de quaisquer outros estados. De modo semelhante, Tomaso e Baldwin (2008) utilizam o grau de pertinência das duas funções da partição *fuzzy* para definir a distribuição de probabilidade de uma evidência.

Entretanto, em um cenário real, existem fenômenos que não podem ser modelados por termos booleanos ou dicotômicos. Inclusive, alguns fenômenos podem ter sobreposição semântica entre termos. Por exemplo, se um paciente tem magreza “extrema”, “grave” ou “elevada”, a vagueza dos termos permite a sua sobreposição semântica (BRIGNOLI et al., 2015). Os trabalhos relacionados realizam tratamento dicotômico às variáveis associadas aos estados de conceitos sobrepostos. Deste modo, uma Fuzzy-DBN deve ser capaz de suportar nodos de evidência com termos sobrepostos, isto é, com estados Não-Dicotômicos. Este trabalho busca resolver esta questão, permitindo nodos de evidência mais genéricos, com múltiplos estados. Para cada evidência, todos os graus de pertinência são considerados, propagando toda a informação *fuzzy* para a estimação probabilística.

A proposta deste trabalho é um modelo de inferência não-dicotômica para a tarefa de inferência filtragem, na qual a modelagem das variáveis de evidência não é restrita em particionamentos *fuzzy*. A proposta de inferência é apresentada no Capítulo 4.



## 4 INFERÊNCIA FUZZY-BAYESIANA DINÂMICA

A abordagem desenvolvida nesta pesquisa visa apresentar uma alternativa para inferir sobre informações que contêm incerteza. Como apresentado na Seção 2.2.1, Conjuntos Fuzzy permitem expressar a incerteza por vagueza, enquanto que a Teoria da Probabilidade, apresentada na Seção 2.1.1, estuda eventos com incerteza por aleatoriedade. Em razão de existirem domínios nos quais ambos os tipos de incerteza estão presentes, as abordagens híbridas Fuzzy-Bayesianas, introduzidas na Seção 2.2.2, estendem o raciocínio Bayesiano para possibilitar a inferência sobre variáveis de entrada *fuzzy*.

Neste trabalho, investiga-se a hipótese de que, em domínios descritos por fenômenos dinâmicos, a introdução de dependências temporais na inferência Fuzzy-Bayesiana melhora seu desempenho caso utilizada como um classificador. Para atingir este objetivo, este trabalho estende o processo de filtragem em DBN para incluir observações *fuzzy*.

Nos trabalhos relacionados, conforme descrito na Seção 3.4, as DBN realizam inferência dicotômica e não permitem a sobreposição de mais de dois estados nas variáveis de evidência. Por este motivo, a proposta é baseada na estratégia da inferência Fuzzy-Bayesiana não-dicotômica de Brignoli (2013).

Para ilustrar a necessidade de inclusão de dependências temporais, é introduzido o problema de navegação robótica de sequência de parede (*Wall-Following*). Neste problema, dadas as medições sensoriais da distância à frente e à esquerda de um robô, o mesmo deve decidir qual direção deve tomar para seguir a parede. A metodologia, a modelagem e os experimentos no problema *Wall-Following*, utilizando dados reais, são descritos posteriormente nas Seções 5 e 6.2. A Figura 6 mostra a topologia utilizada na inferência Fuzzy-Bayesiana para o problema *Wall-Following*, composta por duas variáveis de evidência observáveis e uma variável não-observável.

Para exemplificar uma deficiência da inferência Fuzzy-Bayesiana neste problema, selecionou-se uma instância na qual a medição da variável de evidência *sensor\_reading\_front* é igual a 1,721, da variável de evidência *sensor\_reading\_left* é igual a 0,417 e cuja saída esperada da hipótese é *Slight\_Right\_Turn*.

Com a modelagem *fuzzy* descrita na Seção 6.2, ao aplicar a inferência Fuzzy-Bayesiana nesta instância, o vetor de probabilidades dos estados da hipótese, ou seja  $\{Move\_Forward; Slight\_Right\_Turn; Sharp\_Right\_Turn; Slight\_Left\_Turn\}$ , assume, respectivamente,

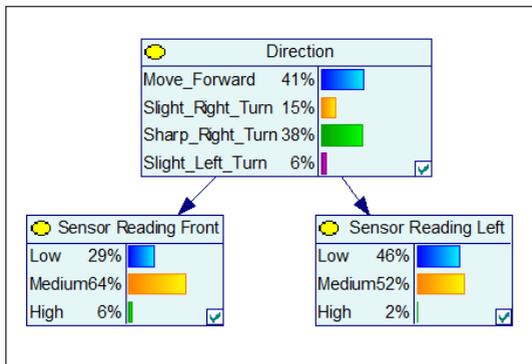


Figura 6 – Rede Bayesiana para o problema *Wall-Following*

os seguintes valores:  $\{0,4931; 0,2900; 0,1650; 0,0518\}$ . Assim, utilizando a abordagem Fuzzy-Bayesiana como um classificador, ao invés da classificação correta como *Slight\_Right\_Turn*, a instância é um falso positivo da classe *Move\_Forward*. Na Seção 6.2, é visto que este não é um caso isolado, mas que todas as instâncias de *Slight\_Right\_Turn* são erroneamente classificadas como *Move\_Forward* pela inferência Fuzzy-Bayesiana.

Deste modo, retoma-se a hipótese de que a adição de dependências temporais na inferência Fuzzy-Bayesiana pode melhorar sua classificação ao ser aplicada em domínios dinâmicos, como é o caso do problema *Wall-Following*.

#### 4.1 INDUÇÃO À INFERÊNCIA FUZZY-BAYESIANA DINÂMICA

Este trabalho tem o objetivo de propor a inferência de DBN sobre variáveis de observação cujos estados são descritos por classes *fuzzy*. Com o objetivo de permitir a sobreposição de estados, a estratégia adotada é baseada na inferência Fuzzy-Bayesiana de Brignoli (2013).

A tarefa de inferência que é de interesse desta pesquisa é a filtragem. Por este motivo, será apresentada a indução à equação de inferência para filtragem ajustada por evidências *fuzzy*, ou seja  $\mathcal{P}(x_i^t | e^{0:t})$ .

Retomando a Equação 2.19, que assume a hipótese de Markov de primeira ordem sobre os estados não-observáveis e estados de evidência (observáveis), tem-se que a estimativa filtrada é definida por:

$$P(x_i^t | e^{0:t}) = \alpha P(e^t | x_i^t) \sum_{l=1}^n P(x_i^t | x_l^{t-1}) P(x_l^{t-1} | e^{0:t-1}) \quad (4.1)$$

onde o índice  $i$  representa o  $i$ -ésimo estado da variável não-observável tal que  $x_i^t \in P(\Omega^t)$  e  $n$  é a quantidade total de estados da variável não-observável. Explicitando a constante de normalização  $\alpha = \frac{1}{P(e^{0:t})}$ , tem-se que:

$$P(x_i^t | e^{0:t}) = \frac{P(e^t | x_i^t) \sum_{l=1}^n P(x_i^t | x_l^{t-1}) P(x_l^{t-1} | e^{0:t-1})}{P(e^{0:t})} \quad (4.2)$$

Se o espaço amostral  $\Omega^t$  pode ser dividido em uma quantidade finita  $n$  de eventos mutuamente exclusivos  $x_j^t$ , onde  $\sum_{j=1}^n x_j^t = \Omega^t$  e se  $P(e^{0:t}) > 0$ , então, conforme o Teorema da Probabilidade Total (Equação 2.10), é possível definir  $P(e^{0:t})$  como:

$$P(e^{0:t}) = \sum_{j=1}^n (P(e^t | x_j^t) P(x_j^t)) \quad (4.3)$$

Considerando que na filtragem em DBN a distribuição de probabilidade dos estados  $x_j$  de  $t-1$  para  $t$  é projetada pela etapa de *forward*<sup>1</sup>, ou seja, por  $\sum_{l=1}^n P(x_i^t | x_l^{t-1}) P(x_l^{t-1} | e^{0:t-1})$ , então  $P(e^{0:t})$  pode ser definido como:

$$P(e^{0:t}) = \sum_{j=1}^n (P(e^t | x_j^t) \sum_{l=1}^n P(x_i^t | x_l^{t-1}) P(x_l^{t-1} | e^{0:t-1})) \quad (4.4)$$

Deste modo, aplicando a Equação 4.4 na inferência por filtragem da Equação 4.2, obtêm-se:

$$P(x_i^t | e^{0:t}) = \frac{P(e^t | x_i^t) \sum_{l=1}^n P(x_i^t | x_l^{t-1}) P(x_l^{t-1} | e^{0:t-1})}{\sum_{j=1}^n P(e^t | x_j^t) \sum_{l=1}^n P(x_i^t | x_l^{t-1}) P(x_l^{t-1} | e^{0:t-1})} \quad (4.5)$$

Considerando a existência de  $k$  evidências combinadas, representadas por  $e_1^{0:t}, \dots, e_k^{0:t}$  denotadas por  $E^{0:t}$  e assumindo que as  $k$  evidências são independentes entre si sob a hipótese  $x_i^t$ , é possível adotar a

---

<sup>1</sup>A etapa de *forward* na filtragem de DBN é apresentada na Seção 2.1.3.1.

propriedade de independência condicional<sup>2</sup>, da mesma forma que nas Equações 2.13 e 2.14. Então, a filtragem é definida por:

$$P(x_i^t | E^{0:t}) = \frac{\sum_{l=1}^n P(x_i^t | x_l^{t-1}) P(x_l^{t-1} | E^{0:t-1}) \prod_{m=1}^k P(e_m^t | x_i^t)}{\sum_{j=1}^n (\sum_{l=1}^n P(x_i^t | x_l^{t-1}) P(x_l^{t-1} | E^{0:t-1}) \prod_{m=1}^k P(e_m^t | x_j^t))} \quad (4.6)$$

O modelo de inferência posposto visa considerar a representação de uma Rede Fuzzy-Bayesiana dinâmica com  $k$  variáveis de evidência ( $E^t$ ) e uma variável não-observável (hipótese  $x^t$ ). Cada variável  $e_j^t$  poderá conter múltiplos estados, discretos ou contínuos. Os estados contínuos de  $e_j^t$  são classificados por meio de um qualificador *fuzzy*, de modo que  $e_j^t s_w$  é a evidência  $e_j^t$  observada no estado *fuzzy*  $s_w$  e  $P_{s_w}(e_j^t | x_i^t)$  representa a probabilidade condicional da evidência  $e_j^t$  no estado  $s_w$  dado  $x_i^t$ . Por fim,  $\mu_{s_w}(e_j^t)$  representa o grau de pertinência de  $e_j^t$  ao estado  $s_w$ .

Baseando-se na inferência apresentada por (BRIGNOLI, 2013; BRIGNOLI et al., 2015), é definida a possibilidade condicional  $\tilde{P}$  de  $e_j^t$  dado  $x_i^t$  por:

$$\tilde{P}(e_j^t | x_i^t) = \sum_{w=1}^u P_{s_w}(e_j^t | x_i^t) \cdot (\mu_{s_w}(e_j^t)) \quad (4.7)$$

onde  $u$  é a quantidade de estados *fuzzy* da variável  $e_j^t$ .

Utilizando a estratégia da Equação (4.7) para considerar os estados *fuzzy*  $s_w$  da variável  $e_j^t$  na Equação (4.6), tem-se que a probabilidade  $\mathcal{P}$  de  $x_i^t$  dadas as evidências imprecisas  $E^{0:t}$  é definida por:

$$\mathcal{P}(x_i^t | E^{0:t}) = \frac{\sum_{l=1}^n P(x_i^t | x_l^{t-1}) \mathcal{P}(x_l^{t-1} | E^{0:t-1}) \prod_{m=1}^k \tilde{P}(e_m^t | x_i^t)}{\sum_{j=1}^n (\sum_{l=1}^n P(x_j^t | x_l^{t-1}) \mathcal{P}(x_l^{t-1} | E^{0:t-1}) \prod_{m=1}^k \tilde{P}(e_m^t | x_j^t))} \quad (4.8)$$

Ou:

---

<sup>2</sup>A Equação 2.8 define independência condicional e a Equação 2.15 apresenta sua aplicação em um Modelo Naïve Bayes.

$$\mathcal{P}(x_i^t | E^{0:t}) = \frac{\sum_{l=1}^n P(x_i^t | x_l^{t-1}) \mathcal{P}(x_l^{t-1} | E^{0:t-1}) \prod_{m=1}^k \sum_{w=1}^u P_{s_w}(e_i^t | x^t) (\mu_{s_w}(e_j^t))}{\sum_{j=1}^n (\sum_{l=1}^n P(x_j^t | x_l^{t-1}) \mathcal{P}(x_l^{t-1} | E^{0:t-1}) \prod_{m=1}^k \sum_{w=1}^u P_{s_w}(e_j^t | x^t) (\mu_{s_w}(e_j^t)))} \quad (4.9)$$

A Equação 4.9 define o núcleo da contribuição desta pesquisa e é denominada Inferência por Filtragem Fuzzy-Bayesiana Dinâmica para variáveis não-dicotômicas.

## 4.2 MODELO DE INFERÊNCIA FUZZY-BAYESIANA DINÂMICA

A Seção 4.1 apresentou a indução matemática de um modelo de inferência em Redes Fuzzy-Bayesianas Dinâmicas, que suporta múltiplas evidências e múltiplos estados, os quais podem conter sobreposição. Para aplicar o modelo, é necessário que sejam especificadas as seguintes informações *a priori*:

- (i) Modelagem Fuzzy das Evidências: descrição por meio de Conjuntos Fuzzy dos estados das variáveis observáveis;
- (ii) Topologia da DBN: especificação das variáveis, das suas relações e da quantidade de fatias temporais. O modelo supõe que o processo sendo descrito pela DBN seja estacionário, descreva um processo de Markov de primeira ordem e que as variáveis de evidência sejam independentes entre si sob a variável não-observável;
- (iii) Parametrização da DBN: parametrização da distribuição inicial da variável não-observável, do modelo de sensores e do modelo de transição. Os parâmetros da rede podem ser especificados manualmente por especialistas do domínio ou por algoritmos de aprendizagem.

Em cada momento, a entrada do modelo é o conjunto de medidas observadas no tempo atual  $t$ . As medidas são estimadas em graus de pertinência por meio da Modelagem Fuzzy e incorporadas na inferência conforme a Equação 4.7. Como definido na Seção 2.1.3.1, o processo de inferência por filtragem pode ser dividido em duas partes: a etapa recursiva de *forward* e a etapa de atualização. Na inferência proposta, a etapa de *forward* é o seguinte trecho da Equação 4.9:

$$\sum_{l=1}^n P(x_l^t | x_l^{t-1}) \mathcal{P}(x_l^{t-1} | E^{0:t-1}) \quad (4.10)$$

Por ser uma definição recursiva, no tempo inicial ( $t=0$ ), o passo base da etapa de *forward* é a distribuição inicial da variável não-observável. A etapa de atualização refere-se a aplicação da Equação 4.9, tendo sido resolvida a recursão da etapa *forward*. A saída do modelo é a probabilidade das variáveis não observáveis em  $t$  dado o conjunto de evidências imprecisas  $E^{0:t}$ . A Figura 7 resume o funcionamento do modelo.

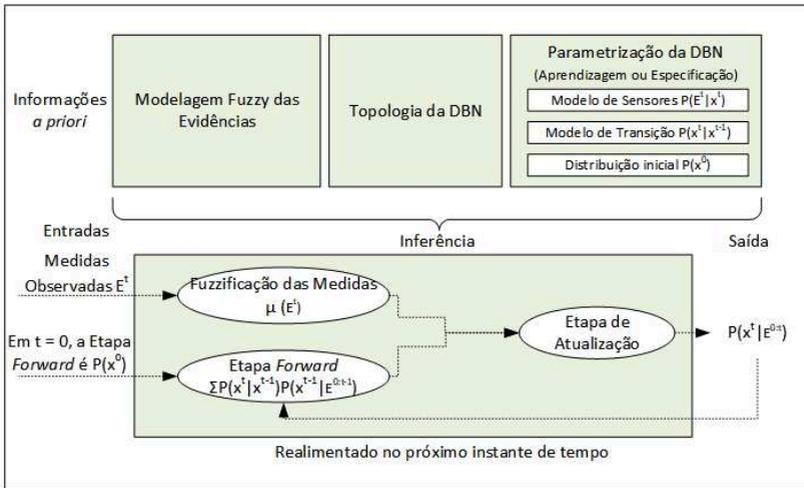


Figura 7 – Modelo de Inferência Fuzzy-Bayesiana Dinâmica

Retomando o exemplo do início deste Capítulo, ao utilizar 3 *time slices* com dependências temporais na variável não-observável do problema *Wall-Following*, tem-se a topologia da DBN ilustrada pela Figura 8. Para a mesma instância exemplificada anteriormente, ao adicionar as medições sensoriais de dois momentos anteriores e utilizar a inferência Fuzzy-Bayesiana Dinâmica, o vetor de probabilidades dos estados da variável não-observável, ou seja  $\{Move\_Forward; Slight\_Right\_Turn; Sharp\_Right\_Turn; Slight\_Left\_Turn\}$ , assume, respectivamente, os seguintes valores:  $\{0,3655; 0,5205; 0,0925; 0,0214\}$ . Assim, utilizando o modelo proposto como um classificador, neste caso, a classificação seria a classe esperada, isto é, *Slight\\_Right\\_Turn*.

A Seção 4.3 apresenta a prova de que a inferência por filtragem

proposta na Equação 4.8 mantém a Propriedade da Soma das Probabilidades<sup>3</sup> sobre a variável não-observável.

#### 4.3 PROVA DA PROPRIEDADE DA SOMA DAS PROBABILIDADES NA INFERÊNCIA FUZZY-BAYESIANA DINÂMICA

A inferência Fuzzy-Bayesiana Dinâmica proposta neste trabalho deve manter a Propriedade das Soma das Probabilidades<sup>4</sup>, que afirma que a somatória das probabilidades de um espaço amostral deve ser 1. A prova é análoga a apresentada no Apêndice A e é definida a seguir por motivos elucidativos.

Retomando a Equação 4.8, tem-se que a probabilidade  $\mathcal{P}$  de  $x_i^t$  dadas as evidências imprecisas  $E^{0:t}$  é definida por:

$$\mathcal{P}(x_i^t | E^{0:t}) = \frac{\sum_{l=1}^n P(x_i^t | x_l^{t-1}) \mathcal{P}(x_l^{t-1} | E^{0:t-1}) \prod_{m=1}^k \tilde{P}(e_m^{0:t} | x_i^t)}{\sum_{j=1}^n (\sum_{l=1}^n P(x_j^t | x_l^{t-1}) \mathcal{P}(x_l^{t-1} | E^{0:t-1}) \prod_{m=1}^k \tilde{P}(e_m^{0:t} | x_j^t))} \quad (4.11)$$

Considerando que o espaço amostral seja composto pelo conjunto de variáveis de estados não-observáveis  $\mathcal{X}$ , deseja-se provar que  $\sum_{i=1}^n \mathcal{P}(x_i^t | E^{0:t}) = 1$ , onde  $n = |\mathcal{X}|$  é a cardinalidade de  $\mathcal{H}$ .

Utilizando a Equação 4.11, a soma de  $\mathcal{P}(x_i^t | E^{0:t}) \forall x_i^t \in \mathcal{X}$ , é definida por:

$$\sum_{i=1}^n \mathcal{P}(x_i^t | E^{0:t}) = \sum_{i=1}^n \frac{\sum_{l=1}^n P(x_i^t | x_l^{t-1}) \mathcal{P}(x_l^{t-1} | E^{0:t-1}) \prod_{m=1}^k \tilde{P}(e_m^{0:t} | x_i^t)}{\sum_{j=1}^n (\sum_{l=1}^n P(x_j^t | x_l^{t-1}) \mathcal{P}(x_l^{t-1} | E^{0:t-1}) \prod_{m=1}^k \tilde{P}(e_m^{0:t} | x_j^t))} \quad (4.12)$$

Considera-se que a expressão  $c_{x_i^t}$  representa a “possibilidade” condicional das evidências observadas com estados contínuos  $e_m^{0:t}$  dada a hipótese  $x_i^t$ , tal que:

<sup>3</sup>A Propriedade da Soma das Probabilidades está definida na Equação 2.2.

<sup>4</sup>Ver Equação 2.2.

$$c_{x_i^t} = \prod_{m=1}^k \tilde{P}(e_m^{0:t} | x_i^t) \quad (4.13)$$

para algum  $x_i^t \in \mathcal{X}$ . Assim, é possível reescrever a Equação 4.12 da seguinte forma:

$$\begin{aligned} \sum_{i=1}^n \mathcal{P}(x_i^t | E^{0:t}) &= \\ &= \sum_{i=1}^n \frac{\sum_{l=1}^n P(x_i^t | x_l^{t-1}) \mathcal{P}(x_l^{t-1} | E^{0:t-1}) c_{x_i^t}}{\sum_{j=1}^n (\sum_{l=1}^n P(x_j^t | x_l^{t-1}) \mathcal{P}(x_l^{t-1} | E^{0:t-1}) c_{x_j^t})} \quad (4.14) \end{aligned}$$

Considerando que o denominador é constante durante o somatório, tem-se que:

$$\begin{aligned} \sum_{i=1}^n \mathcal{P}(x_i^t | E^{0:t}) &= \\ &= \frac{1}{\sum_{j=1}^n (\sum_{l=1}^n P(x_j^t | x_l^{t-1}) \mathcal{P}(x_l^{t-1} | E^{0:t-1}) c_{x_j^t})} \cdot \\ &= \sum_{i=1}^n (\sum_{l=1}^n P(x_i^t | x_l^{t-1}) \mathcal{P}(x_l^{t-1} | E^{0:t-1}) c_{x_i^t}) \quad (4.15) \end{aligned}$$

Ou seja:

$$\begin{aligned} \sum_{i=1}^n \mathcal{P}(x_i^t | E^{0:t}) &= \\ &= \frac{\sum_{i=1}^n (\sum_{l=1}^n P(x_i^t | x_l^{t-1}) \mathcal{P}(x_l^{t-1} | E^{0:t-1}) c_{x_i^t})}{\sum_{j=1}^n (\sum_{l=1}^n P(x_j^t | x_l^{t-1}) \mathcal{P}(x_l^{t-1} | E^{0:t-1}) c_{x_j^t})} \quad (4.16) \end{aligned}$$

Considerando-se que o limite inicial e superior de ambos os somatórios são idênticos, então o denominador e o numerador da fração na Equação (4.16) são iguais, de modo que:

$$\sum_{i=1}^n \mathcal{P}(x_i^t | E^{0:t}) = \frac{\sum_{i=1}^n (\sum_{l=1}^n P(x_i^t | x_l^{t-1}) \mathcal{P}(x_l^{t-1} | E^{0:t-1}) c_{x_i^t})}{\sum_{j=1}^n (\sum_{l=1}^n P(x_j^t | x_l^{t-1}) \mathcal{P}(x_l^{t-1} | E^{0:t-1}) c_{x_j^t})} = \frac{1}{1} = 1 \quad (4.17)$$

Pela a Equação 4.17, observa-se que  $\sum_{i=1}^n \mathcal{P}(x_i^t | E^{0:t}) = 1$ , ou seja, a Propriedade da Soma das Probabilidades é mantida pela inferência Fuzzy-Bayesiana Dinâmica proposta para a variável não-observável  $x_i^t$ .

As Seções 5 e 6 descrevem a metodologia, modelagem e os resultados obtidos ao aplicar a proposta em duas bases de dados, uma das quais é o exemplo utilizado neste capítulo.

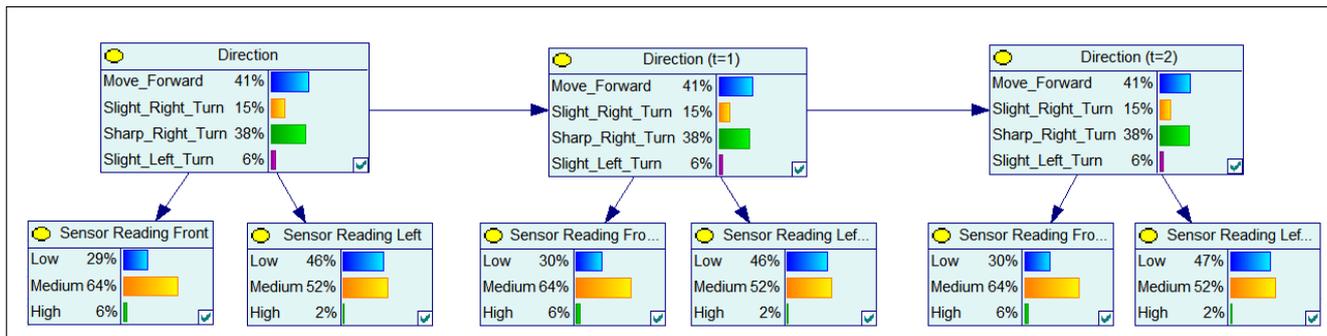


Figura 8 – Rede Bayesiana Dinâmica para o problema *Wall-Following*

## 5 METODOLOGIA DE AVALIAÇÃO

Para avaliar a inferência por filtragem proposta, o modelo foi aplicado como um classificador em bases de dados abertas, em problemas onde a modelagem *fuzzy* possui sobreposição de estados.

Os trabalhos relacionados que definem filtragem ou a previsão em Fuzzy-DBNs, além de realizarem a inferência de forma dicotômica, restringem a modelagem das evidências cenário de aplicação em particionamentos *fuzzy*, os quais são limitados a duas classes de pertinência complementares<sup>1</sup>. Em função desta restrição, não é possível aplicá-los às situações de interesse desta pesquisa.

Deste modo, esta proposta é comparada com a inferência por filtragem tradicional em DBN (MURPHY, 2002) e com a inferência Fuzzy-Bayesiana de Brignoli (2013). Nota-se que, unicamente no caso de não haverem adições de fatias temporais (ou seja, se  $t = 0$ ), o modelo proposto é equivalente à inferência Fuzzy-Bayesiana. Em contrapartida, caso as variáveis de evidência observáveis possuam pertinência total e exclusiva a uma função de pertinência *fuzzy* (ou seja, se as evidências são *crisp*), então o modelo é equivalente a DBN tradicional. Assim, as comparações entre as três abordagens tem o objetivo de mostrar que (i) há uma melhora na classificação quando são inseridas fatias temporais na inferência Fuzzy-Bayesiana; e (ii) há uma melhora na classificação quando as variáveis observáveis assumem valores *fuzzy* (ou seja, valores que não são *crisp*).

As próximas Seções descrevem as bases de dados utilizadas, a preparação dos dados, o treinamento e avaliação da rede e os índices empregados para análise da qualidade da classificação.

### 5.1 BASES DE DADOS

Para avaliação do modelo, foram utilizadas duas bases de dados públicas do UC Irvine Machine Learning Repository (UCI) (DHEERU; TANISKIDOU, 2017). A primeira base é o *Occupancy Detection Dataset* (CANDANEDO; FELDHEIM, 2016), que é composta por dados experimentais para classificação binária da ocupação de um quarto a partir de dados de temperatura, umidade e CO<sub>2</sub>. Esta base de dados possui três conjuntos de dados sequenciais. O conjunto utilizado nesta pesquisa tem 8143 instâncias, das quais 21.23% (1729 instâncias) são da

---

<sup>1</sup>Os trabalhos relacionados são apresentados na Seção 3.4.

classe ocupado e as 78.77% restantes (6414 instâncias) são da classe desocupado. Como variáveis de observação, foram utilizados os seguintes atributos: *Temperature* (temperatura em graus Celsius), *Humidity* (umidade relativa em porcentagem), *Light* (luz em Lux) e CO<sub>2</sub> (em ppm ou partes por milhão). A variável não observável do modelo é *Occupancy* e assume o valor *Yes* ou *No* para representar o estado de ocupação do quarto.

A segunda base de dados é o Conjunto de Dados Simplificados para Sequência de Parede ou *Wall-Following Dataset* (FREIRE et al., 2009), que é composta por dados coletados enquanto um robô navega através de uma sala, enquanto segue a parede no sentido horário. O conjunto de dados possui as distâncias da frente e da esquerda do robô, obtidas pela leitura de sensores ultra-som, e o nome da classe correspondente à direção que o robô deve seguir para continuar seguindo a parede. O conjunto de dados tem 5456 instâncias e quatro classes referentes à direção. Existem 2205 instâncias da classe *Move\_Forward* (mover para frente), 2097 instâncias da classe *Sharp\_Right\_Turn* ou (curva acentuada à direita), 328 instâncias da classe *Slight\_Right\_Turn* ou (curva leve à direita) e 328 instâncias da classe *Slight\_Left\_Turn* ou (curva leve à esquerda). A Figura 9 apresenta a distribuição destas classes no conjunto de dados. Como variáveis observáveis, foram utilizadas *sensor\_reading\_front*, representando a leitura mínima do sensor em um arco de 60 graus localizado na frente do robô (número real) e *sensor\_reading\_left*, representando a leitura mínima do sensor em um arco de 60 graus localizado à esquerda do robô (número real). A variável não observável é a classe *Direction*, que representa a direção que o robô deve seguir e assume um dos seguintes valores *Move\_Forward*, *Slight\_Right\_Turn*, *Sharp\_Right\_Turn* ou *Slight\_Left\_Turn*.

## 5.2 TRATAMENTO DOS DADOS

As variáveis observáveis de ambos os conjuntos de dados assumem valores contínuos. Por este motivo, foram discretizadas em categorias. As categorias utilizadas são conceitos linguísticos que informam se a medição de cada variável está alta, média ou baixa e foram modeladas utilizando Conjuntos Fuzzy<sup>2</sup>. A modelagem foi realizada a partir da observação do intervalo e da distribuição dos valores de cada variável. Os parâmetros da modelagem *fuzzy* foram refinados após a realização

---

<sup>2</sup>A modelagem *fuzzy* foi realizada através da biblioteca jFuzzyLogic, disponível em <http://jfuzzylogic.sourceforge.net/html/index.html>.

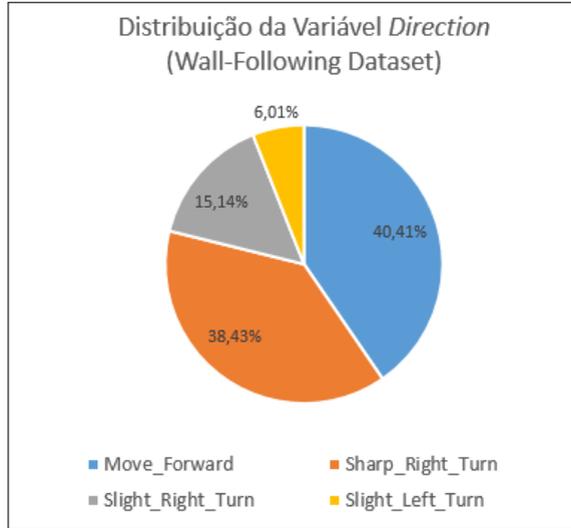


Figura 9 – Distribuição dos valores da Variável *Direction* no *Wall-Following Dataset*

de alguns testes preliminares. Para otimizar e reduzir a arbitrariedade do processo de parametrização das classes *fuzzy*, é necessário investigar métodos de treinamento de parâmetros a partir de bases de dados. Entretanto, nota-se que este aspecto disto do escopo central desta pesquisa. Para reduzir os efeitos da arbitrariedade na parametrização *fuzzy*, tanto a proposta quanto os trabalhos escolhidos para comparação foram implementados utilizando exatamente a mesma modelagem *fuzzy*.

Para possibilitar a aprendizagem dos parâmetros das BN e a classificação por meio da DBN padrão, as variáveis observáveis foram pré-classificadas por meio da modelagem *fuzzy*. Assim, cada valor observado foi categorizado na classe *fuzzy* com maior pertinência. Entretanto, o valor real observado também foi mantido, para que pudesse ser utilizado pelas abordagens híbridas Fuzzy-Bayesianas.

Em seguida, os dados foram organizados em uma planilha de modo que cada linha representa uma instância em um tempo  $m$  com o conjunto das variáveis observáveis  $E^t$  e não observáveis  $x^t$  no intervalo de tempo  $t = [m - T, m]$ , onde  $T$  representa a quantidade de fatias temporais que estão sendo recuperadas. A Figura 10 exemplifica a forma descrita de organização dos dados. Considerando que o tempo

inicial é 0, então para  $m = 0$  até  $m = T - 1$ , não é possível obter todo o conjunto de variáveis  $E^t$  e  $x^t$ , pois o intervalo  $t = [m - T, m]$  possuiria valores menores que o tempo inicial, para os quais não foram capturadas observações. Assim, as primeiras  $T - 1$  observações do conjunto de dados são descartadas.

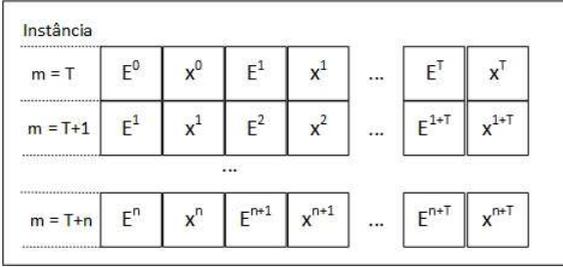


Figura 10 – Disposição dos Dados com  $T$  fatias temporais

Após a organização dos dados, estes são divididos em dois subconjuntos: subconjunto de treinamento e subconjunto de avaliação. 70% do conjunto total de dados disponíveis foi aleatoriamente escolhido para o subconjunto de treinamento. Por sua vez, os dados restantes (30%) integraram o subconjunto de avaliação.

### 5.3 TREINAMENTO E AVALIAÇÃO DA DBN

O subconjunto de treinamento foi utilizado na aprendizagem dos parâmetros da BN. Para obter a distribuição *a priori* sobre as variáveis não-observáveis no tempo inicial  $P(X^0)$ , o modelo de transição  $P(X^t|X^{t-1})$  e o modelo de sensores ou observações  $P(E^t|X^t)$  foi utilizado o Algoritmo Expectation Maximization (DEMPSTER; LAIRD; RUBIN, 1977), por meio da ferramenta GeNIe Modeler<sup>3</sup>.

Após estas etapas, a proposta é aplicada no subconjunto de avaliação. Primeiramente, os valores de evidência observados são fuzzificados, de modo a obter o grau de pertinência de cada evidência  $e_j^t$  a cada estado  $s_w$ , ou seja:  $\mu_{s_w}(e_j^t)$ . Como definido na Seção 4, para cada fatia temporal, são realizadas duas etapas para estimar a inferência a etapa de *forward*, que projeta a distribuição do estado anterior para o atual e a etapa de atualização, na qual a adição de novas evidências atualiza

<sup>3</sup>GeNIe Modeler é uma ferramenta para modelagem e aprendizagem de BN, disponível em <https://www.bayesfusion.com/genie/>.

a distribuição da variável não-observável.

O modelo proposto foi utilizado como um classificador, no qual o estado da variável não-observável de maior probabilidade é considerado a classe de saída. As entradas da rede são as evidências observadas  $E^0$  até  $E^T$ , sendo  $T$  a quantidade de fatias temporais. Após a aplicação das etapas *forward* e atualização, o modelo gera a probabilidade da variável não-observável  $x_i^T$  para todo  $1 \leq i \leq n$ , com  $n$  representando a quantidade de estados da variável  $x^T$ . Como classificador, a saída será o estado  $x_i^T$  com o maior valor de probabilidade.

## 5.4 MÉTODOS PARA ANÁLISE DA CLASSIFICAÇÃO

O desempenho da classificação é representado por valores escalares em diferentes métricas, como precisão, sensibilidade e especificidade (THARWAT, 2018).

Uma matriz de confusão apresenta informação sobre quão bem um determinado comportamento é detectado corretamente e com que frequência é classificado como outro comportamento (RUUSKA et al., 2018).

A Figura 11 mostra uma Matriz de Confusão para um problema de classificação binário, isto é, com duas classes. A diagonal em verde representa as predições corretas do classificador, enquanto que a diagonal rosa representa as incorretas. Supondo que existam as classes positivo e negativo, as saídas do classificador binário são contabilizadas como(THARWAT, 2018):

- Verdadeiro Positivo (*True Positive* - TP), quando a classe real da amostra é positiva e a classe prevista é positiva;
- Falso Positivo (*False Positive* - FP), quando a classe real da amostra é positiva e a classe prevista é negativa;
- Verdadeiro Negativo (*True Negative* - TN), quando a classe real da amostra é negativa e a classe prevista é negativa;
- Falso Negativo (*False Negative* - FN), quando a classe real da amostra é positiva e a classe prevista é positiva;

A Acurácia é definida como a relação entre as amostras corretamente classificadas e a quantidade total de amostras, ou seja (SOKOLOVA; JAPKOWICZ; SZPAKOWICZ, 2006):

		True/Actual Class	
		Positive (P)	Negative (N)
Predicted Class	True (T)	True Positive (TP)	False Positive (FP)
	False (F)	False Negative (FN)	True Negative (TN)
		$P = TP + FN$	$N = FP + TN$

Figura 11 – Matriz de Confusão para Classificação Binária (THARWAT, 2018)

$$\text{Acurácia} = \frac{TP + TN}{TP + TN + FP + FN} \quad (5.1)$$

A Sensitividade, Recall ou Taxa de Verdadeiros Positivos representa os positivos corretamente classificados na quantidade total de amostras positivas (SOKOLOVA; JAPKOWICZ; SZPAKOWICZ, 2006):

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5.2)$$

A Especificidade, Recall Inverso ou Taxa de Verdadeiros Negativos representa a relação entre as amostras corretamente classificadas como negativas e a quantidade total de amostras negativas (SOKOLOVA; JAPKOWICZ; SZPAKOWICZ, 2006):

$$\text{Recall Inverso} = \frac{TN}{FP + TN} \quad (5.3)$$

O Valor Preditivo Positivo ou Precisão representa a proporção de amostras positivas que foram corretamente classificadas em relação à quantidade total de amostras previstas como positivas (SOKOLOVA; JAPKOWICZ; SZPAKOWICZ, 2006):

$$\text{Precisão} = \frac{TP}{FP + TP} \quad (5.4)$$

O Valor Preditivo Negativo ou Precisão Inversa mede a proporção de amostras negativas que foram corretamente classificadas em relação à quantidade total de amostras previstas como negativas (SOKOLOVA; JAPKOWICZ; SZPAKOWICZ, 2006):

$$\text{Precisão Inversa} = \frac{TN}{FN + TN} \quad (5.5)$$

Por fim, o  $F_1$ -score representa a média harmônica entre a Precisão e o Recall (SOKOLOVA; JAPKOWICZ; SZPAKOWICZ, 2006):

$$F_1\text{-score} = \frac{2 \cdot \text{Precisão} \cdot \text{Recall}}{\text{Precisão} + \text{Recall}} \quad (5.6)$$

Todas as métricas apresentadas variam de zero a um e maiores valores indicam melhor desempenho de classificação. Além destas métricas, existem métodos de avaliação gráfica. Dentre os quais a Curva *Receiver Operating Characteristics* (ROC) é um gráfico bidimensional, na qual o Eixo Y é representado pelo Recall e o Eixo X, por  $1 - \text{Recall Inverso}$  (Taxa de Falsos Positivos) (THARWAT, 2018).

Uma Curva ROC descreve as compensações relativas entre benefícios (Verdadeiros Positivos) e custos (Falsos Positivos). Classificadores a partir de Redes Bayesianas podem ser usados com um valor de *threshold* para produzir uma saída binária. Cada valor de *threshold* produz um diferente ponto no espaço ROC. Com a variação do *threshold*, é possível obter a Curva ROC. Na Figura 12, o Ponto A (0, 0) representa a estratégia de nunca emitir uma classificação positiva. A estratégia oposta, de emitir classificações positivas incondicionalmente, é representada pelo Ponto C (1, 1). O Ponto D (1, 0) representa um classificador no qual todas as amostras positivas e negativas foram classificadas erroneamente. Por sua vez, o Ponto B (0, 1) representa uma classificação perfeita (FAWCETT, 2005).

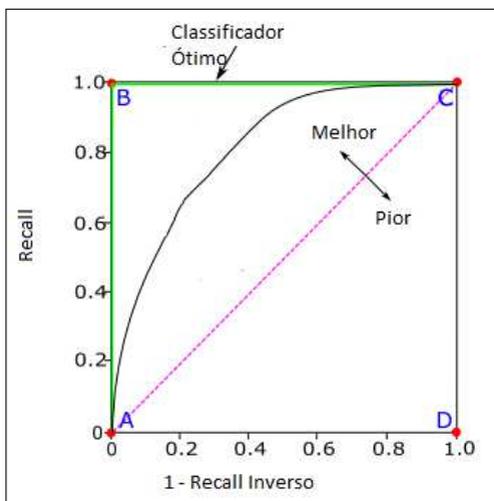


Figura 12 – Curva ROC, Adaptado de (THARWAT, 2018)

As métricas descritas foram utilizadas para avaliar o desempenho da proposta como um classificador. Os resultados obtidos após a aplicação da metodologia são detalhados no Capítulo 6.

## 6 AVALIAÇÃO

Este capítulo descreve a aplicação da metodologia de avaliação descrita no Capítulo 5. Tendo em vista a comparação das abordagens, a partir desta Seção, o termo Fuzzy-DBN será utilizado para referir-se à proposta, o termo Fuzzy-BN indicará a inferência Fuzzy-Bayesiana de Brignoli et al. (2015) e, por fim, o termo DBN se referirá a inferência por filtragem tradicional em Redes Bayesianas Dinâmicas.

De acordo com o modelo apresentado na Seção 4.2, as informações fornecidas antes do processo de inferência correspondem à modelagem *fuzzy* das evidências, à descrição da topologia da rede e sua parametrização. Nota-se que é utilizada a mesma modelagem *fuzzy* para as abordagens Fuzzy-DBN e Fuzzy-BN. Do mesmo modo, é mantida a parametrização e topologia da DBN na Fuzzy-DBN. A topologia da Fuzzy-BN é equivalente à topologia da DBN no tempo inicial, sem a adição de fatias temporais.

As próximas Seções apresentam a modelagem e resultados obtidos com a aplicação da proposta nas bases de dados *Occupancy Detection* e *Wall-Following*, conforme Seção 5.1.

### 6.1 OCCUPANCY DETECTION DATASET

Para o *Occupancy Detection Dataset* (CANDANEDO; FELDHEIM, 2016), foram modeladas quatro variáveis de evidência (ou observação) e uma variável não-observável, que indica o estado de ocupação de um quarto. A Figura 13 ilustra a modelagem *fuzzy* utilizada para cada uma das variáveis de observação: *Light* (a),  $\text{CO}_2$  (b), *Humidity* (c) e *Temperature* (d). As variáveis de observação possuem sobreposição de estados e os valores de pertinência não são necessariamente complementares. Por exemplo, uma medição de 19,5 graus Celsius possui o grau de pertinência à classe *low* igual a 0,755, à classe *medium* igual a 0,325 e à classe *high* igual a 0,186.

Conforme ilustrado na Figura 14, durante os experimentos foram utilizadas quatro fatias temporais na topologia da Fuzzy-DBN. As distribuições de probabilidade obtidas durante a aprendizagem de parâmetros são especificadas no Apêndice C.1.

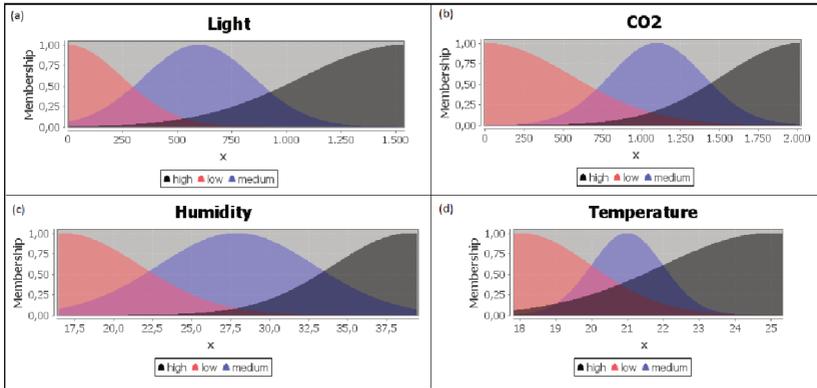


Figura 13 Modelagem Fuzzy para o *Occupancy Detection Dataset*

A Matriz de Confusão que descreve os erros e acertos da classificação pela Fuzzy-DBN é apresentada na Tabela 1. Com base nestes dados, foram estimadas as métricas de erro, que são listadas e comparadas com a classificação por meio de Fuzzy-BN e DBN na Tabela 2.

Tabela 1 Matriz de Confusão para a Variável Binária *Occupancy*

		Valor Real	
		Yes	No
Valor Estimado	Yes	<b>474</b>	33
	No	39	<b>1896</b>

Em relação à Fuzzy-BN, houve um aumento de 1,37% na acurácia da classificação. Ou seja, ao adicionar fatias temporais na Fuzzy-BN, a proposta obteve mais acertos durante a classificação. A melhoria no Recall indica um aumento no acerto da classificação positiva de ocupação, quando o quarto realmente estava ocupado. Apesar de haver uma queda na Precisão, o  $F_1$ -score indica que o aumento no Recall foi mais significativo para a qualidade da classificação do que a diminuição na Precisão.

A inclusão de evidências *fuzzy* na DBN resultou em um aumento de 0,34% na acurácia da classificação. Fuzzy-DBN obteve melhores resultados no Recall, na Precisão Inversa e no  $F_1$ -score. Isto indica que a proposta melhora a capacidade da rede de acertar quando o quarto está ocupado, como prejuízo, a diminuição da Precisão mostra que há

Tabela 2 – Comparação da Classificação com Fuzzy-BN, DBN e Fuzzy-DBN (*Occupancy Detection Dataset*)

	Fuzzy-BN	DBN	Fuzzy-DBN
Acurácia	0,957412	0,96724	0,970516
Recall	0,849903	0,879159	0,923977
Recall Inverso	0,986003	0,994121	0,982893
Precisão	0,941685	0,978558	0,934911
Precisão Inversa	0,961091	0,96423	0,979845
$F_1$ -score	0,893443	0,926199	0,929412

um aumento na quantidade de falsos positivos. Apesar disto, o  $F_1$ -score mostra que o aumento nos acertos supera o aumento na quantidade de falsos positivos durante a classificação da Fuzzy-DBN, o que sugere a efetividade da proposta.

A Figura 15 apresenta a Curva ROC para a classificação positiva de ocupação. A Fuzzy-BN e Fuzzy-DBN obtiveram resultados similares. Entretanto, a Curva ROC da DBN tradicional mostra que a Taxa de Falsos Positivos teve um aumento mais acentuado, conforme a variação do *threshold*, do que as outras duas abordagens.

Tanto ao incluir fatias temporais na Fuzzy-BN quanto ao incluir evidências *fuzzy* na DBN, houve uma melhora na classificação ao utilizar a Fuzzy-DBN, principalmente em relação à acurácia. A próxima Seção compara o desempenho das três abordagens com outra base de dados, ao mesmo tempo que explora como o aumento na quantidade de fatias temporais afeta a classificação.

## 6.2 WALL-FOLLOWING DATASET

Para o *Wall-Following Dataset* (FREIRE et al., 2009), foram utilizadas duas variáveis de evidência (ou observação) e uma variável não-observável, que indica a direção que o robô deve seguir. A Figura 16 ilustra a modelagem *fuzzy* utilizada para cada uma das variáveis de observação: *Sensor\_Reading\_Front* (a) e *Sensor\_Reading\_Left* (b), que representam a leitura frontal e à esquerda do robô, respectivamente. As variáveis de observação possuem sobreposição de estados e os valores de pertinência não são necessariamente complementares.

Como estão sendo utilizadas apenas duas informações sensoriais, tem-se o objetivo de investigar a hipótese de que adição de fatias temporais aumenta a qualidade da classificação. Deste modo, a quantidade de fatias temporais variou entre 1 (quando a DBN é equivalente a uma BN) e 7, cuja topologia é ilustrada na Figura 17. As distribuições de probabilidade obtidas durante a aprendizagem de parâmetros são especificadas no Apêndice C.2.

Diferentemente da detecção de ocupação tratada na Seção 6.1, o *Wall-Following Dataset* não é um problema de classificação binária. A variável não observável *direction* possui quatro estados possíveis: (i) *Move\_Forward*, (ii) *Slight\_Right\_Turn*, (iii) *Sharp\_Right\_Turn* e (iv) *Slight\_Left\_Turn*.

Sem a adição de fatias temporais, a DBN é equivalente a uma BN de estrutura Naïve Bayes. Enquanto que a Fuzzy-DBN torna-se equivalente à Fuzzy-BN. A Matriz de Confusão para a variável não-observável *Direction*, utilizando a Fuzzy-BN, é apresentada na Tabela 3. Das 265 instâncias da classe *Slight\_Right\_Turn*, a Fuzzy-BN não produziu nenhuma classificação positiva. A maior parcela das instâncias de *Slight\_Right\_Turn* foi classificada erroneamente na classe *Move\_Forward*. Este problema também ocorreu com a classe *Slight\_Left\_Turn*, cuja totalidade dos falsos negativos foi classificada na classe *Move\_Forward*.

Tabela 3 – Matriz de Confusão para a Variável *Direction* (Fuzzy-BN)

		Valor Real			
		(i)	(ii)	(iii)	(iv)
Valor Estimado	(i) <i>Move_Forward</i>	<b>639</b>	256	153	99
	(ii) <i>Slight_Right_Turn</i>	0	<b>0</b>	0	0
	(iii) <i>Sharp_Right_Turn</i>	0	9	<b>470</b>	0
	(iv) <i>Slight_Left_Turn</i>	0	0	6	<b>3</b>

Estes problemas na classificação da Fuzzy-BN levaram à uma queda de 7,70% na Acurácia Geral, em comparação com a BN (Tabela 4). A quantidade alta de falsos positivos na classe *Move\_Forward* acarretou na diminuição do Recall Inverso em 41,20% e da Precisão em 24,01%. Além disso, devido à baixa quantidade de verdadeiros positivos, o Recall das classes *Slight\_Right\_Turn* e *Slight\_Left\_Turn* foi igual ou próximo a zero.

Tabela 4 – Comparação da Classificação com BN e Fuzzy-BN (*Wall-Following Dataset*)

	BN				Fuzzy-BN			
Acurácia Geral	0,746789				0,680122			
Classe	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
Acurácia	0,7865	0,8404	0,9156	0,9512	0,6893	0,8379	0,8972	0,9358
Recall	0,7136	0,9472	0,7806	0,2255	1,0000	0,0000	0,7472	0,0294
Recall Inverso	0,8333	0,8197	1,0000	0,9993	0,4900	1,0000	0,9911	0,9961
Precisão	0,7331	0,5040	1,0000	0,9583	0,5571	Div/0 <sup>1</sup>	0,9812	0,3333
Precisão Inversa	0,8193	0,9877	0,8794	0,9510	1,0000	0,8379	0,8625	0,9391
$F_1$ -score	0,7232	0,6579	0,8768	0,3651	0,7156	Div/0 <sup>1</sup>	0,8484	0,0541

Em geral, a BN obteve melhores desempenhos na classificação. Entretanto, o Recall e o  $F_1$ -score da classe *Slight\_Left\_Turn* mostram que a BN também falhou em classificar corretamente suas instâncias positivas. Deste modo, a utilização da DBN e da Fuzzy-DBN, investiga se a inclusão de dependências temporais aperfeiçoa a capacidade de classificação da BN e da Fuzzy-BN.

Utilizando 3 *time slice*, a Fuzzy-DBN, cuja Matriz de Confusão é apresentada an Tabela 5, obteve progressos na classificação de *Slight\_Right\_Turn*. Enquanto que a Fuzzy-BN apresentou 265 falsos negativos, a Fuzzy-DBN reduziu este número para menos da metade (128 falsos negativos).

---

<sup>1</sup>Métricas indefinidas devido à divisão por zero.

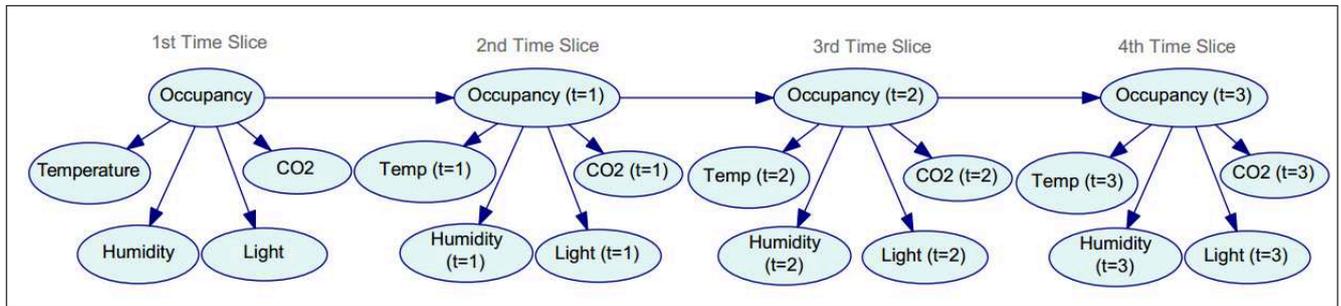


Figura 14 – Topologia da DBN com 4 *time slice* para o *Occupancy Detection Dataset*

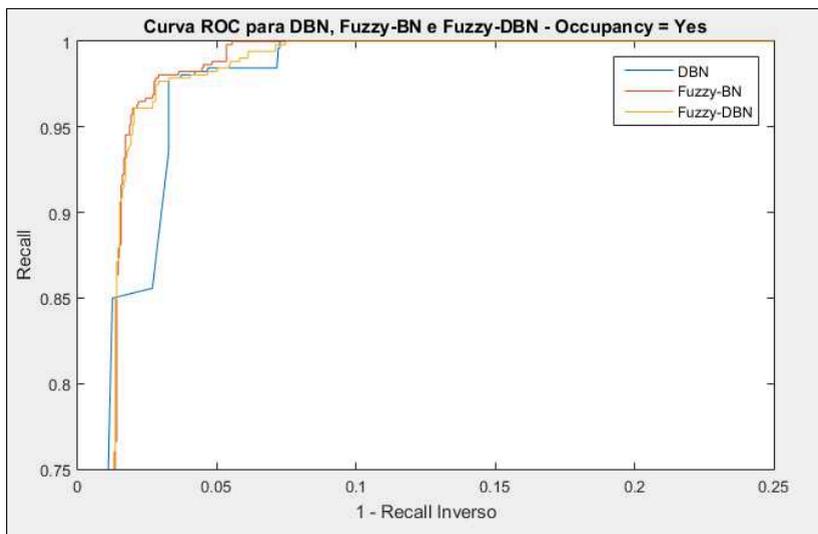


Figura 15 – Curva ROC para DBN, Fuzzy-BN e Fuzzy-DBN para a classificação positiva de ocupação

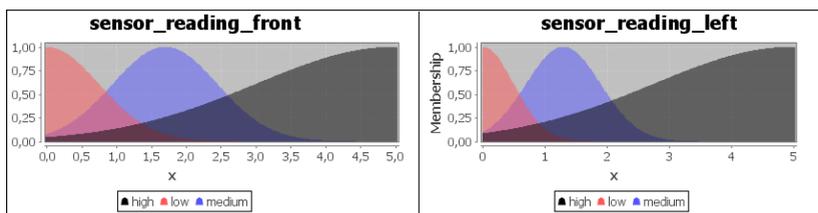


Figura 16 – Modelagem Fuzzy para o *Wall-Following Dataset*

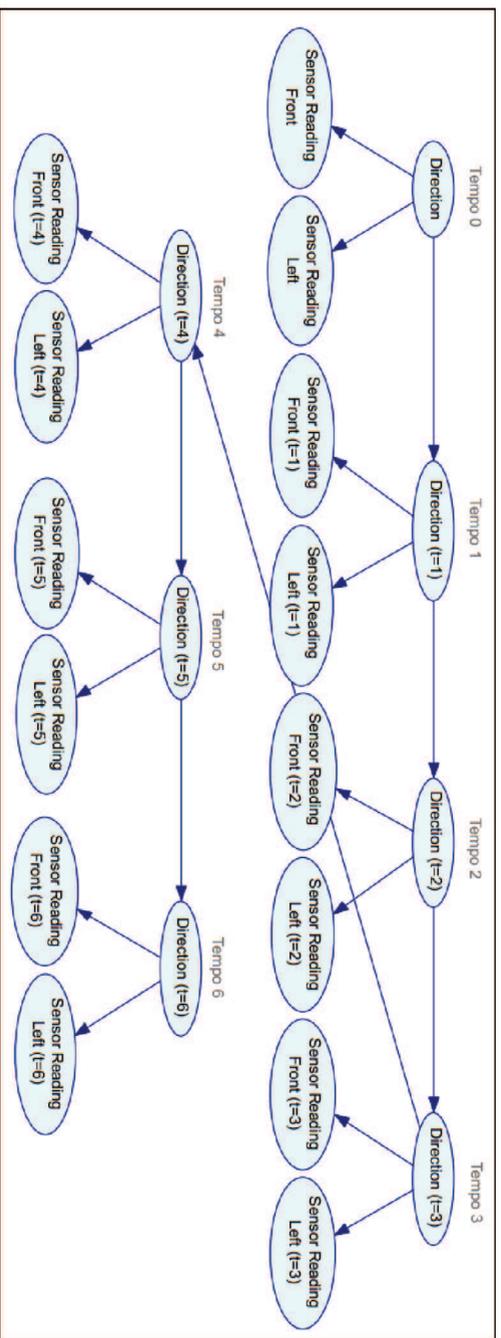


Figura 17 – Topologia da DBN com 7 time slice para o *Wall-Following Dataset*

Tabela 5 – Matriz de Confusão para a Variável *Direction* (Fuzzy-DBN com 3 *time slice*)

		Valor Real			
		(i)	(ii)	(iii)	(iv)
Valor Estimado	(i) Move_Forward	<b>623</b>	119	145	95
	(ii) Slight_Right_Turn	1	<b>137</b>	1	0
	(iii) Sharp_Right_Turn	15	7	<b>472</b>	0
	(iv) Slight_Left_Turn	0	2	11	<b>7</b>

Em relação a Fuzzy-BN, houve uma melhora de 11,42% na Acurácia Geral da classificação ao utilizar a Fuzzy-DBN com 3 *time slice*. As principais melhoras são percebidas na Acurácia das classes *Move\_Forward*, *Slight\_Right\_Turn* e *Sharp\_Right\_Turn*. Além disso, a diminuição dos falsos positivos da classe *Move\_Forward* impactou positivamente no Recall e na Precisão da Fuzzy-DBN (Tabela 6).

Tabela 6 – Comparação da Classificação com DBN e Fuzzy-DBN e 3 *time slice* (*Wall-Following Dataset*)

	DBN				Fuzzy-DBN			
Acurácia Geral	0,751682				0,757798			
Classe	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
Acurácia	0,7945	0,8495	0,9089	0,9505	0,7706	0,9205	0,8905	0,9339
Recall	0,7011	0,9472	0,7981	0,2745	0,9750	0,5170	0,7504	0,0686
Recall Inverso	0,8544	0,8307	0,9781	0,9954	0,6396	0,9985	0,9781	0,9915
Precisão	0,7555	0,5197	0,9580	0,8000	0,6344	0,9856	0,9555	0,3500
Precisão Inversa	0,8167	0,9878	0,8857	0,9538	0,9755	0,9144	0,8624	0,9412
$F_1$ -score	0,7273	0,6711	0,8708	0,4088	0,7687	0,6782	0,8406	0,1148

Com uma leve redução na quantidade de falsos negativos e um aumento de falsos positivos, a classe *Slight\_Left\_Turn*, obteve um desempenho melhor que o atingido pela Fuzzy-BN. Entretanto, em relação à DBN, o Recall e a Precisão mostram que a quantidade de falsos negativos ainda é significativa. Ainda assim, a Acurácia Geral da Fuzzy-DBN é 0,81% superior à DBN, ambas com 3 *time slice*. Estes resultados são apresentados na Tabela 6, na qual a melhora na classificação pode ser percebida principalmente no Recall e  $F_1$ -score de *Move\_Forward* e na Precisão e no  $F_1$ -score de *Slight\_Right\_Turn*.

Ao utilizar 5 *time slice*, é possível perceber que a Fuzzy-DBN, cuja Matriz de Confusão é apresentada na Tabela 7, continua apresentando uma melhora na classificação positiva das classes com alta quantidade de falsos negativos na Fuzzy-BN inicial. A classe *Slight\_Right*

*\_Turn* obteve apenas 22 casos de falsos negativos, ou seja, apenas 8,3% das instâncias de *Slight\_Right\_Turn* não foram classificadas corretamente. Esta melhoria é significativa, uma vez que a Fuzzy-BN havia obtido 100% de falsos negativos nesta classe. Isto mostra o impacto positivo da adição de dependências temporais entre variáveis da rede no problema do *Wall-Following Dataset*.

Tabela 7 – Matriz de Confusão para a Variável *Direction* (Fuzzy-DBN com 5 *time slice*)

		Valor Real			
		(i)	(ii)	(iii)	(iv)
Valor Estimado	(i) Move_Forward	<b>576</b>	18	134	74
	(ii) Slight_Right_Turn	53	<b>243</b>	20	0
	(iii) Sharp_Right_Turn	8	2	<b>462</b>	0
	(iv) Slight_Left_Turn	2	2	13	<b>28</b>

A Fuzzy-DBN com 5 *time slice* obteve um aumento na Acurácia Geral de 5,65% em relação à Fuzzy-DBN com 3 *time slice* e de 5,74% em relação à DBN com 5 *time slice*. A comparação entre a Fuzzy-DBN e a DBN, ambas com 5 *time slice* é apresentada na Tabela 8. É possível notar uma melhoria na classificação, principalmente em relação à Acurácia e ao  $F_1$ -score das classes *Move\_Forward* e *Slight\_Right\_Turn*. Embora o Recall e  $F_1$ -score de *Slight\_Right\_Turn* ainda sejam inferiores na Fuzzy-DBN, a diferença não é tão significativa como havia sido durante a comparação entre Fuzzy-BN e BN e durante a comparação entre Fuzzy-DBN e DBN com 3 *time slice*.

Tabela 8 – Comparação da Classificação com DBN e Fuzzy-DBN e 5 *time slice* (*Wall-Following Dataset*)

	DBN				Fuzzy-DBN			
	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
Acurácia Geral	0,757187				0,800612			
Classe	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
Acurácia	0,8000	0,8544	0,9089	0,9511	0,8232	0,9419	0,8917	0,9443
Recall	0,7136	0,9358	0,7981	0,3137	0,9014	0,9170	0,7345	0,2745
Recall Inverso	0,8554	0,8387	0,9781	0,9935	0,7731	0,9467	0,9901	0,9889
Precisão	0,7600	0,5288	0,9580	0,7619	0,7182	0,7690	0,9788	0,6222
Precisão Inversa	0,8232	0,9854	0,8857	0,9561	0,9244	0,9833	0,8564	0,9535
$F_1$ -score	0,7361	0,6757	0,8708	0,4444	0,7994	0,8365	0,8392	0,3810

Por fim, a Fuzzy-DBN com 7 *time slice* apresenta uma Matriz de Confusão (Tabela 9) com uma redução na quantidade de falsos negativos da classe *Slight\_Right\_Turn* e um aumento nos fal-

os negativos das classes *Move\_Forward* e *Sharp\_Right\_Turn*. Os falsos negativos destas classes foram erroneamente classificados como *Slight\_Right\_Turn*.

Tabela 9 – Matriz de Confusão para a Variável *Direction* (Fuzzy-DBN com 7 *time slice*)

		Valor Real			
		(i)	(ii)	(iii)	(iv)
Valor Estimado	(i) Move_Forward	<b>551</b>	17	144	46
	(ii) Slight_Right_Turn	75	<b>242</b>	60	0
	(iii) Sharp_Right_Turn	6	3	<b>411</b>	0
	(iv) Slight_Left_Turn	7	3	14	<b>56</b>

A Acurácia Geral da Fuzzy-DBN com 7 *time slice* teve uma melhora de 1,77% em relação à DBN com 7 *time slice*. Entretanto, o aumento na quantidade de falsos positivos da classe *Slight\_Right\_Turn* reduziu a Acurácia Geral em relação à Fuzzy-DBN com 5 *time slice*. Em contrapartida, a Fuzzy-DBN com 7 *time slice* foi a rede com melhor  $F_1$ -score e Recall para a classe *Slight\_Left\_Turn*, que tanto nas DBN quanto nas Fuzzy-DBN apresentadas anteriormente, havia apresentado problemas na sua classificação positiva. Além disto, as classes *Move\_Forward* e *Slight\_Right\_Turn* novamente obtiveram melhores Acurácia e  $F_1$ -score, em relação a DBN.

Tabela 10 – Comparação da Classificação com DBN e Fuzzy-DBN e 7 *time slice* (*Wall-Following Dataset*)

	DBN				Fuzzy-DBN			
Acurácia Geral	0,757187				0,770642			
Classe	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
Acurácia	0,8006	0,8538	0,9089	0,9511	0,8196	0,9034	0,8612	0,9572
Recall	0,7105	0,9321	0,7965	0,3529	0,8623	0,9132	0,6534	0,5490
Recall Inverso	0,8584	0,8387	0,9791	0,9909	0,7922	0,9015	0,9911	0,9843
Precisão	0,7630	0,5278	0,9598	0,7200	0,7269	0,6419	0,9786	0,7000
Precisão Inversa	0,8221	0,9846	0,8850	0,9584	0,8997	0,9817	0,8206	0,9704
$F_1$ -score	0,7358	0,6739	0,8705	0,4739	0,7888	0,7539	0,7836	0,6154

Estes resultados mostram que (i) a inclusão de fatias temporais reduziu as falhas de classificação da Fuzzy-BN, o que é percebido principalmente na classe *Slight\_Right\_Turn*; (ii) a inclusão de evidências *fuzzy*, na Fuzzy-DBN, melhorou a Acurácia Geral com 3, 5 e 7 *time slice*, em relação a DBN; (iii) a definição da quantidade de fatias temporais mais adequada depende não só do domínio, como dos seus objetivos,

pois apesar da rede que obteve maior Acurácia Geral ter 5 *time slice*, em um contexto no qual a classe *Slight\_Left\_Turn* tivesse identificação prioritária, a utilização de 7 *time slice* seria mais adequada, pois obteve maior especificidade (Recall) na classificação desta classe.

## 7 CONSIDERAÇÕES FINAIS

As abordagens existentes para inferir em Redes Bayesianas Dinâmicas sobre dados com imprecisão por vagueza restringem a modelagem e a sobreposição de estados de evidência para apenas dois estados *fuzzy*. Deste modo, com o objetivo de permitir a sobreposição de múltiplos estados e de considerá-los durante o processo de inferência, este trabalho propôs um modelo para inferência Fuzzy-Bayesiana Dinâmica sobre variáveis não-dicotômicas.

O modelo proposto é inspirado em um modelo Fuzzy-BN, onde o grau de pertinência *fuzzy* é utilizado para definir a possibilidade condicional das evidências *fuzzy*. É apresentado um modelo genérico, no sentido de que permite múltiplas evidências, múltiplos estados *fuzzy* e múltiplas fatias temporais.

O modelo foi aplicado como um classificador em duas bases de dados. Na primeira, *Detection Occupancy Dataset*, a proposta obteve 97% de Acurácia, 92% de  $F_1$ -score e 92% de Recall. A inferência Fuzzy-BN é equivalente a Fuzzy-DBN proposta se e somente se não são consideradas fatias temporais. Assim, os resultados mostraram que a Fuzzy-DBN obteve melhores resultados de classificação ao adicionar fatias temporais a inferência Fuzzy-BN. Em contrapartida, a comparação da proposta com DBN, sugere que o uso de evidências *fuzzy* aumenta a qualidade da previsão.

Com a segunda base de dados, *Wall-Following Dataset*, foi investigado um problema de classificação não binário para o qual a Fuzzy-BN apresentou diversos problemas de classificação. Os experimentos novamente mostraram que a inclusão de fatias temporais aumentou a Acurácia Geral em relação a Fuzzy-BN, além de reduzir suas falhas de classificação. Em comparação com a DBN, pode-se notar que a inclusão de evidências *fuzzy* também melhorou a Acurácia Geral. Nesta base de dados, a utilização de cinco fatias temporais resultou na obtenção de 80% de Acurácia Geral, que representa um aumento de 5,65% em relação a DBN com a mesma quantidade de fatias temporais e de 17,72% em relação à Fuzzy-BN.

A avaliação da filtragem proposta produziu resultados favoráveis às hipóteses da pesquisa. Isto é, por meio dos experimentos, foi observado que tanto a introdução de fatias temporais em Fuzzy-BN quanto a inclusão de evidências *fuzzy* em DBN melhoraram a qualidade da classificação. Além disto, por meio da prova de que o somatório da distribuição de probabilidade da variável não-observável é igual a um,

demonstrou-se que a inferência proposta manteve a Propriedade da Soma das Probabilidades.

Como produto desta pesquisa, o artigo intitulado “Fusão de Dados Através de Redes Fuzzy-Bayesianas para a Construção de Crenças de Percepção em Agentes Cognitivos” foi aceito para publicação pela Revista de Informática Teórica e Aplicada e encontra-se na etapa de edição.

Os objetivos deste trabalho, que incluem a proposta de um modelo de inferência por filtragem Fuzzy-Bayesiano Dinâmico com evidências *fuzzy* não-dicotômicas, sua avaliação e a prova de que respeita a Propriedade da Soma das Probabilidades, foram alcançados com êxito.

## 7.1 TRABALHOS FUTUROS

Como trabalhos futuros, é possível estender a proposta para as outras tarefas de inferência, como previsão, suavização e explicação mais provável.

Outra possibilidade de continuidade a este trabalho é o desenvolvimento de um *framework* para modelagem e construção de Redes Fuzzy-Bayesianas Dinâmicas, que implemente a inferência proposta e, assim, possibilite maior automatização do processo de avaliação utilizado.

Além disto, um problema em potencial em Fuzzy-DBN diz respeito ao custo computacional no armazenamento de informações da rede e no processamento da inferência. Por este motivo, é necessário investigar métodos de inferência aproximada em DBN e adaptá-los para Fuzzy-DBN.

Por fim, considerando cenários nos quais a inferência seja *online*, a estrutura da rede pode crescer sem limite à medida que são adicionadas novas evidências observadas. Uma alternativa para reduzir o espaço de armazenamento é limitar dinamicamente a quantidade de fatias temporais mantidas na memória. Neste caso é necessário analisar a relação entre o aumento da acurácia da rede e a quantidade de fatias temporais empregadas.

## REFERÊNCIAS

- ANAGNOSTOPOULOS, C.; SEKKAS, O.; HADJIEFTHYMIADIS, S. Context fusion: Dealing with sensor reliability. *2007 IEEE International Conference on Mobile Adhoc and Sensor Systems, MASS*, n. 0, p. 1–6, 2007.
- BERRAR, D. Bayes' Theorem and Naive Bayes Classifier. *Reference Module in Life Sciences*, v. 1, 2018. <<http://linkinghub.elsevier.com/retrieve/pii/B9780128096338204731>>.
- BEUZEN, T.; MARSHALL, L.; SPLINTER, K. D. A comparison of methods for discretizing continuous variables in Bayesian Networks. *Environmental Modelling & Software*, Elsevier, v. 108, n. December 2017, p. 61–66, 2018. ISSN 13648152. <<https://linkinghub.elsevier.com/retrieve/pii/S1364815217313269>>.
- BRIGNOLI, J. T. *Modelo Híbrido Difuso-Probabilístico: uma alternativa para sistemas especialistas*. [S.l.]: Dissertação (Mestrado) - Programa de Pós-Graduação em Ciência da Computação. Universidade Federal de Santa Catarina. Florianópolis, 2001. 118 p.
- BRIGNOLI, J. T. *Um Modelo para Suporte ao Raciocínio Diagnóstico Diante da Dinâmica do Conhecimento sobre Incertezas*. 260 p. Tese (Doutorado) — Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento. Universidade Federal de Santa Catarina. Florianópolis, 2013.
- BRIGNOLI, J. T. et al. A fuzzy-Bayesian model based on the superposition of states applied to the clinical reasoning support. *IntelliSys 2015 - Proceedings of 2015 SAI Intelligent Systems Conference*, p. 210–219, 2015.
- CANDANEDO, L. M.; FELDHEIM, V. Accurate occupancy detection of an office room from light, temperature, humidity and CO<sub>2</sub> measurements using statistical learning models. *Energy and Buildings*, Elsevier B.V., v. 112, p. 28–39, 2016. ISSN 03787788. <<http://dx.doi.org/10.1016/j.enbuild.2015.11.071>>.
- CHAI, H.; WANG, B. A Hierarchical Situation Assessment Model Based on Fuzzy Bayesian Network. In: *International Conference on Artificial Intelligence and Computational Intelligence*. [S.l.]: Springer, 2011. p. 444–454.

CHEN, H.; GAO, X. Ship recognition based on improved forwards-backwards algorithm. *6th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2009*, v. 5, p. 509–513, 2009.

DAWID, A. P. Conditional Independence in Statistical Theory. *Journal of the Royal Statistical Society. Series B (Methodological)*, [Royal Statistical Society, Wiley], v. 41, n. 1, p. 1–31, 1979. ISSN 00359246. <<http://www.jstor.org/stable/2984718>>.

DEAN, T.; KANAZAWA, K. A model for reasoning about persistence and causation. *Computational intelligence*, Wiley Online Library, v. 5, n. 2, p. 142–150, 1989.

DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society. Series B (methodological)*, JSTOR, p. 1–38, 1977.

DHEERU, D.; TANISKIDOU, E. K. *(UCI) Machine Learning Repository*. 2017. <<http://archive.ics.uci.edu/ml>>.

DUTT, L. S.; KURIAN, M. Handling of Uncertainty – A Survey. *International Journal of Scientific and Research Publications*, v. 3, n. 1, p. 1–4, 2013.

EL-SAWAH, A. et al. A Framework for 3D Hand Tracking and Gesture Recognition using Elements of Genetic Programming. *Fourth Canadian Conference on Computer and Robot Vision (CRV '07)*, n. Figure 2, p. 495–502, 2007.

FAWCETT, T. An introduction to ROC analysis. *Irbm*, v. 35, n. 6, p. 299–309, 2005. ISSN 18760988.

FENG, Y. et al. Engagement Evaluation for Autism Intervention by Robots Based on Dynamic Bayesian Network and Expert Elicitation. *IEEE Access*, v. 5, p. 19494–19504, 2017. ISSN 21693536.

FREIRE, A. L. et al. Short-term memory mechanisms in neural network learning of robot navigation tasks: A case study. *2009 6th Latin American Robotics Symposium, LARS 2009*, n. 4, 2009.

GIAMBENE, G. Survey on Probability Theory. In: *Queueing Theory and Telecommunications: Networks and Applications*. Boston, MA: Springer US, 2005. p. 237–304. ISBN 978-0-387-24066-4.

- HE, R. et al. A quantitative risk analysis model considering uncertain information. *Process Safety and Environmental Protection*, Institution of Chemical Engineers, v. 118, p. 361–370, 2018. ISSN 09575820. <<https://doi.org/10.1016/j.psep.2018.06.029>>.
- HUA, Z.; RUI, L.; JIZHOU, S. An emotional model for nonverbal communication based on fuzzy dynamic bayesian network. *Electrical and Computer Engineering, ...*, n. May, 2006.
- JINYONG, Y.; KEKE, L.; WENJING, W. Ship-Aircraft Joint Situation Assessment by Using Fuzzy Dynamic Bayesian network. In: *2017 IEEE International Conference on Unmanned Systems (ICUS)*. [S.l.: s.n.], 2017. p. 220 – 224. ISBN 9780136042594.
- KAO, H. Y.; CHEN, B. S. Efficiency classification by hybrid Bayesian networks - The dynamic multidimensional models. *Applied Soft Computing Journal*, Elsevier B.V., v. 24, p. 842–850, 2014. ISSN 15684946. <<http://dx.doi.org/10.1016/j.asoc.2014.08.061>>.
- KHALID, A. J.; WANG, J.; NURUDEEN, M. A new fault classification model for prognosis and diagnosis in CNC machine. *2013 25th Chinese Control and Decision Conference, CCDC 2013*, p. 3538–3543, 2013.
- KITCHENHAM, B. *Procedures for performing systematic reviews*. [S.l.: s.n.], 2004. 28 p. ISSN 13537776. ISBN 1353-7776.
- KOLMOGOROV, A. N. *Foundations of the Theory of Probability*. [S.l.: s.n.], 1956. 47 p. ISSN 0971-8044. ISBN 0828400237.
- LI, J.; GAO, X. Fuzzy variable structure dynamic Bayesian network applying target recognition. *2016 First IEEE International Conference on Computer Communication and the Internet (ICCCI)*, p. 434–438, 2016. <<http://ieeexplore.ieee.org/document/7778959/>>.
- LI, Y.; CHEN, J.; FENG, L. Dealing with uncertainty: A survey of theories and practices. *IEEE Transactions on Knowledge and Data Engineering*, v. 25, n. 11, p. 2463–2482, 2013. ISSN 10414347.
- MIHAJLOVIC, V.; PETKOVIC, M. Dynamic bayesian networks: A state of the art. *CTIT technical reports series*, v. 34, p. 1–37, 2001. ISSN 13813625.
- MURPHY, K. P. *Dynamic Bayesian Networks: Representation, Inference and Learning*. 225 p. Tese (Doutorado) — University Of California, Berkeley, 2002.

NADERPOUR, M.; LU, J.; ZHANG, G. A fuzzy dynamic bayesian network-based situation assessment approach. *IEEE International Conference on Fuzzy Systems*, 2013. ISSN 10987584.

NADERPOUR, M.; LU, J.; ZHANG, G. A fuzzy dynamic bayesian network-based situation assessment approach. In: *2013 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. [S.l.: s.n.], 2013. p. 1–8. ISSN 1098-7584.

NADERPOUR, M.; LU, J.; ZHANG, G. An intelligent situation awareness support system for safety-critical environments. *Decision Support Systems*, Elsevier B.V., v. 59, n. 1, p. 325–340, 2014. ISSN 01679236. <<http://dx.doi.org/10.1016/j.dss.2014.01.004>>.

NADERPOUR, M.; LU, J.; ZHANG, G. The explosion at institute: Modeling and analyzing the situation awareness factor. *Accident Analysis and Prevention*, Elsevier Ltd, v. 73, p. 209–224, 2014. ISSN 00014575. <<http://dx.doi.org/10.1016/j.aap.2014.09.008>>.

NOJAVAN, F. A.; QIAN, S. S.; STOW, C. A. Comparative analysis of discretization methods in Bayesian networks. *Environmental Modelling and Software*, Elsevier Ltd, v. 87, p. 64–71, 2017. ISSN 13648152. <<http://dx.doi.org/10.1016/j.envsoft.2016.10.007>>.

NORVIG, P.; RUSSELL, S. *Inteligência Artificial*. [S.l.]: Elsevier Brasil, 2014.

PEARL, J. Chapter 3 - MARKOV AND BAYESIAN NETWORKS: Two Graphical Representations of Probabilistic Knowledge. In: PEARL, J. (Ed.). *Probabilistic Reasoning in Intelligent Systems*. San Francisco (CA): Morgan Kaufmann, 1988. p. 77–141. ISBN 978-0-08-051489-5.

ROSS, T. J. Properties of Membership Functions, Fuzzification, and Defuzzification. In: *Fuzzy Logic with Engineering Applications*. 3rd ed. ed. Chichester, UK: John Wiley & Sons, Ltd, 2010. p. 89–116. <<http://doi.wiley.com/10.1002/9781119994374.ch4>>.

RUUSKA, S. et al. Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle. *Behavioural Processes*, Elsevier, v. 148, n. January, p. 56–62, 2018. ISSN 18728308. <<https://doi.org/10.1016/j.beproc.2018.01.004>>.

SEKKAS, O.; ANAGNOSTOPOULOS, C. B.; HADJIEFTHYMI-  
ADES, S. Context fusion through imprecise reasoning. *2007 IEEE  
International Conference on Pervasive Services, ICPS*, p. 88–91, 2007.

SHI, J.-g.; GAO, X.-g.; LI, X.-m. Modeling air combat situation  
assessment by using fuzzy dynamic Bayesian network. *Journal of  
System Simulation*, v. 18, n. 5, p. 1093–1100, 2006.

SOKOLOVA, M.; JAPKOWICZ, N.; SZPAKOWICZ, S. Beyond  
Accuracy, F-Score and ROC: A Family of Discriminant Measures  
for Performance Evaluation BT - AI 2006: Advances in Artificial  
Intelligence. In: SATTAR, A.; KANG, B.-h. (Ed.). *Journal of the  
America Pomological Society*. Berlin, Heidelberg: Springer Berlin  
Heidelberg, 2006. v. 59, n. 1, p. 1015–1021. ISBN 978-3-540-49788-2.  
ISSN 15273741.

SYKES, E. R. Preliminary findings of visualization of the interruptible  
moment. *Lecture Notes in Computer Science (including subseries  
Lecture Notes in Artificial Intelligence and Lecture Notes in  
Bioinformatics)*, v. 5976 LNCS, p. 215–229, 2010. ISSN 03029743.

TEIXEIRA, M.; ZAVERUCHA, G. Fuzzy Markov predictor in  
multi-step electric load forecasting. *Proceedings of the International  
Joint Conference on Neural Networks, 2003.*, v. 4, p. 3065–3070, 2003.

TEIXEIRA, M.; ZAVERUCHA, G. Fuzzy hidden Markov predictor in  
electric load forecasting. *2004 IEEE International Joint Conference  
on Neural Networks (IEEE Cat. No.04CH37541)*, p. 315–320, 2004.  
ISSN 10987576.

TEIXEIRA, M. A.; ZAVERUCHA, G. A Partitioning Method for  
Fuzzy Probabilistic Predictors. In: *Neural Information Processing*.  
[S.l.: s.n.], 2004. p. 929–934.

TEIXEIRA, M. A.; ZAVERUCHA, G. Fuzzy multi-hidden Markov  
predictor in electric load forecasting. *Proceedings of the International  
Joint Conference on Neural Networks*, v. 3, p. 1758–1763, 2005. ISSN  
10987576.

THARWAT, A. Classification assessment methods. *Applied  
Computing and Informatics*, The Author, 2018. ISSN 22108327.  
<<https://linkinghub.elsevier.com/retrieve/pii/S2210832718301546>>.

THUNNISSEN, D. P. Uncertainty Classification for the Design and  
Development of Complex Systems. 2003.

TIBIRIÇÁ, C. A. G. *Uma abordagem híbrida Fuzzy-bayesiana para modelagem de incertezas*. [S.l.]: Dissertação (Mestrado) - Programa de Pós-Graduação em Ciência da Computação. Universidade Federal de Santa Catarina. Florianópolis, 2005.

TOMASO, E. D.; BALDWIN, J. F. An approach to hybrid probabilistic models. *International Journal of Approximate Reasoning*, v. 47, n. 2, p. 202–218, 2 2008. ISSN 0888-613X. <<https://www.sciencedirect.com/science/article/pii/S0888613X07000485>>.

VIERTL, R. Is it Necessary to Develop a Fuzzy Bayesian Inference ? In: VIERTL, R. (Ed.). *Probability and Bayesian Statistics*. Boston, MA: Springer US, 1987. p. 471–475. ISBN 978-1-4613-1885-9.

VIERTL, R. Modeling of Fuzzy Measurements in Reliability Estimation. 1989.

VIERTL, R. Statistics with Fuzzy Data. In: RICCIA, G. D.; KRUSE, R.; VIERTL, R. (Ed.). *Proceedings of the ISSEK94 Workshop on Mathematical and Statistical Methods in Artificial Intelligence*. Vienna: Springer Vienna, 1995. p. 33–49. ISBN 978-3-7091-2690-5.

VIERTL, R. Fuzzy Bayesian Inference. *Smps*, p. 10–15, 2008.

WANG, H.; GAO, X.; THOMPSON, C. P. Aircraft target recognition based on recursive inference of fuzzy discrete DBNs. *Proceedings - International Conference on Advanced Computer Control, ICACC 2009*, p. 183–187, 2009.

WANG, P. Analogy in a general-purpose reasoning system. *Cognitive Systems Research*, v. 10, n. 3, p. 286–296, 9 2009. ISSN 1389-0417. <<file://www.sciencedirect.com/science/article/pii/S1389041708000764>>.

WANG, Y. F. et al. Fire probability prediction of offshore platform based on Dynamic Bayesian Network. *Ocean Engineering*, Elsevier Ltd, v. 145, n. April, p. 112–123, 2017. ISSN 0029-8018. <<https://doi.org/10.1016/j.oceaneng.2017.08.035>>.

WU, H. et al. Modeling and simulation of pilot behavior interfaced with avionics system. *Communications in Computer and Information Science*, v. 327 CCIS, n. PART 2, p. 390–398, 2012. ISSN 18650929.

YANG, S.; GAO, X.; CHEN, H. Ground Thread Identification of the Reconnaissance and Strike Integrated VAV Based on Improved Direct Inference Algorithm. In: *2010 3rd International Conference on*

*Advanced Computer Theory and Engineering(1CACTE) Ground*. [S.l.: s.n.], 2010. p. 577–580. ISBN 9781424465422.

YAO, J. Y. et al. Modeling System Based on Fuzzy Dynamic Bayesian Network for Fault Diagnosis and Reliability Prediction. 2015.

ZADEH, L. Fuzzy sets. *Information and Control*, v. 8, n. 3, p. 338–353, 1965. ISSN 00199958.

ZAIYI, P. Network security situation analysis based on a dynamic Bayesian network and phase space reconstruction. *The Journal of Supercomputing*, Springer US, 2018. ISSN 0920-8542. <<http://link.springer.com/10.1007/s11227-018-2575-3>>.

ZHANG, H.; LI, R. Emotional nonverbal communication based on fuzzy dynamic Bayesian network. *9th International Conference on Control, Automation, Robotics and Vision, 2006, ICARCV '06*, 2006.

ZHANG, Q. et al. A Fuzzy Probability Bayesian Network Approach for Dynamic Cybersecurity Risk Assessment in Industrial Control Systems. *IEEE Transactions on Industrial Informatics*, IEEE, v. 14, n. 6, p. 2497–2506, 2018. ISSN 15513203.

ZHENG, W.; CHIEN, S. Risk Analysis for the Wireless Communication of the High-Speed Maglev under the Cognitive Uncertainties. *2018 IEEE Intelligent Vehicles Symposium (IV)*, IEEE Intelligent Vehicles Symposium, n. Iv, p. 1063–1070, 2018.



**APÊNDICE A – Prova da Propriedade da Soma das  
Probabilidades na Inferência Fuzzy-Bayesiana**



A inferência Fuzzy-Bayesiana (BRIGNOLI, 2013) deve manter a Propriedade das Soma das Probabilidades<sup>1</sup>, que afirma que a somatória das probabilidades de um espaço amostral deve ser 1. Ou seja, considerando que o espaço amostral seja composto pelo conjunto de hipóteses  $\mathcal{H}$ , então  $\sum_{i=1}^n \mathcal{P}(\mathbf{H}_i | \mathbf{E}_j \mathbf{x}_k) = \mathbf{1}$ , onde  $n = |\mathcal{H}|$  é a cardinalidade de  $\mathcal{H}$ . A seguir, deseja-se provar que esta propriedade é mantida na Inferência Fuzzy-Bayesiana.

Retomando a Equação 2.20, tem-se que a probabilidade condicional  $\mathcal{P}$  sobre evidências imprecisas é dada por (BRIGNOLI, 2013):

$$\mathcal{P}(\mathbf{H}_i | \mathbf{E}_j) = \frac{P(\mathbf{H}_i) \prod_{j=1}^m \sum_{k=1}^u (P_{\mathbf{x}_k}(\mathbf{E}_j | \mathbf{H}_i) \cdot \mu_{\mathbf{x}_k}(\mathbf{E}_j))}{\sum_{l=1}^n P(\mathbf{H}_l) \prod_{j=1}^m \sum_{k=1}^u (P_{\mathbf{x}_k}(\mathbf{E}_j | \mathbf{H}_l) \cdot \mu_{\mathbf{x}_k}(\mathbf{E}_j))} \quad (\text{A.1})$$

Pela Equação A.1, tem-se que a soma de  $\mathcal{P}(\mathbf{H}_i | \mathbf{E}_j) \forall \mathbf{H}_i \in \mathcal{H}$ , é definida por:

$$\sum_{i=1}^n \mathcal{P}(\mathbf{H}_i | \mathbf{E}_j) = \sum_{i=1}^n \frac{P(\mathbf{H}_i) \prod_{j=1}^m \sum_{k=1}^u (P_{\mathbf{x}_k}(\mathbf{E}_j | \mathbf{H}_i) \cdot \mu_{\mathbf{x}_k}(\mathbf{E}_j))}{\sum_{l=1}^n P(\mathbf{H}_l) \prod_{j=1}^m \sum_{k=1}^u (P_{\mathbf{x}_k}(\mathbf{E}_j | \mathbf{H}_l) \cdot \mu_{\mathbf{x}_k}(\mathbf{E}_j))} \quad (\text{A.2})$$

A Propriedade da Soma das Probabilidades não se aplica sobre as evidências apresentadas à Rede Fuzzy-Bayesiana, pois os qualificadores *fuzzy* na formulação matemática do modelo de inferência transformam a conotação das variáveis  $\mathbf{E}_j$  após a distribuição das probabilidades conjuntas, de modo que seu aspecto altere de probabilístico para possibilístico. Assim, uma evidência  $\mathbf{E}_j$  passa a ser um agregado de conceitos que confirmam uma medida de possibilidade para múltiplos estados superpostos, ou seja uma possibilidade aferida (BRIGNOLI, 2013).

Considera-se que a expressão  $c_{\mathbf{H}_i}$  representa a “possibilidade” condicional das evidências observadas  $\mathbf{E}_j$  dada a hipótese  $\mathbf{H}_i$ , tal que:

$$c_{\mathbf{H}_i} = \prod_{j=1}^m \sum_{k=1}^u (P_{\mathbf{x}_k}(\mathbf{E}_j | \mathbf{H}_i) \cdot \mu_{\mathbf{x}_k}(\mathbf{E}_j)) \quad (\text{A.3})$$

para algum  $\mathbf{H}_i \in \mathcal{H}$ . Nota-se que, considerando um espaço amostral

---

<sup>1</sup>Ver Equação 2.2.

que possua uma instanciação para  $\mathbf{u}$  estados de  $\mathbf{m}$  evidências,  $\mathbf{c}_{H_i}$  possui um valor único para cada  $H_i \in \mathcal{H}$ .

Assim, é possível reescrever a Equação A.2 da seguinte forma:

$$\sum_{i=1}^n \mathcal{P}(H_i|E_j) = \sum_{i=1}^n \frac{P(H_i) \cdot \mathbf{c}_{H_i}}{\sum_{l=1}^n P(H_l) \cdot \mathbf{c}_{H_l}} \quad (\text{A.4})$$

Considerando que o denominador é constante durante o somatório, tem-se que:

$$\sum_{i=1}^n \mathcal{P}(H_i|E_j) = \frac{1}{\sum_{l=1}^n P(H_l) \cdot \mathbf{c}_{H_l}} \cdot \sum_{i=1}^n P(H_i) \cdot \mathbf{c}_{H_i} \quad (\text{A.5})$$

Ou seja:

$$\sum_{i=1}^n \mathcal{P}(H_i|E_j) = \frac{\sum_{i=1}^n P(H_i) \cdot \mathbf{c}_{H_i}}{\sum_{l=1}^n P(H_l) \cdot \mathbf{c}_{H_l}} \quad (\text{A.6})$$

Considerando-se que o limite inicial e superior de ambos os somatórios são idênticos, então o denominador e o numerador da fração na Equação (A.6) são iguais, de modo que:

$$\sum_{i=1}^n \mathcal{P}(H_i|E_j) = \frac{\sum_{i=1}^n P(H_i) \cdot \mathbf{c}_{H_i}}{\sum_{l=1}^n P(H_l) \cdot \mathbf{c}_{H_l}} = \frac{1}{1} = 1 \quad (\text{A.7})$$

Assim, para qualquer cardinalidade do conjunto  $\mathcal{H}$  maior que zero<sup>2</sup>, é mantida a Propriedade da Soma das Probabilidades, isto é, a somatória das probabilidades de um espaço amostral é igual a 1. Nota-se que o axioma é mantido independentemente do valor de  $\mathbf{c}_{H_i}$ . Deste modo, a quantidade  $\mathbf{m}$  de evidências, a quantidade  $\mathbf{t}$  de estados das evidências e a possível sobreposição destes estados não podem tornar  $\sum_{i=1}^n \mathcal{P}(H_i|E_j) \neq 1$ , ou seja, respeitam o axioma da Teoria das Probabilidades para  $\mathcal{P}(H_i|E_j)$ .

---

<sup>2</sup>A Equação (2.20) prevê que haja pelo menos uma hipótese, uma evidência e um estado de evidência, pois os limites iniciais dos produtórios e somatórios são iguais a 1.

**APÊNDICE B - Protocolo da Revisão Sistemática da  
Literatura**



## B.1 PROTOCOLO DA REVISÃO

O objetivo desta revisão sistemática é realizar um levantamento da literatura relevante acerca de Redes Fuzzy-Bayesianas Dinâmicas, para identificar e distinguir abordagens híbridas que apliquem conjuntos Fuzzy a Redes Bayesianas Dinâmicas. Com o intuito de cumprir os objetivos da revisão, foi elaborado um protocolo de busca, de modo a levantar o máximo da literatura sobre o assunto. A revisão sistemática foi realizada de acordo com a metodologia de Kitchenham (KITCHENHAM, 2004). A seguir será descrito o protocolo de busca elaborado.

### B.1.1 Pergunta de Pesquisa

1. Pergunta de Pesquisa: Como integrar conjuntos Fuzzy e Redes Bayesianas Dinâmicas?
2. Pergunta de Pesquisa Adicional: Quais enfoques de Redes Fuzzy-Bayesianas Dinâmicas existem?
3. População: Pesquisas realizadas na área de Redes Bayesianas Dinâmicas.
4. Intervenção: Para a pergunta principal: Estado da arte em Redes Fuzzy-Bayesianas Dinâmicas.  
Para a pergunta secundária: Abordagens utilizadas na área de Redes Fuzzy-Bayesianas Dinâmicas.
5. Resultados: Comparativo dos enfoques utilizados para integrar conjuntos Fuzzy e Redes Bayesianas Dinâmicas.
6. Contexto: Artigos publicados nas bases de dados eletrônicas: (i) ACM Digital Library , (ii) IEEEExplore , (iii) ScienceDirect e (iv) Springer.

## B.2 BASES DE DADOS E STRINGS DE BUSCA

A seguir são descritas as strings de busca e os filtros utilizados em cada uma das bases de dados que foram consultadas nesta revisão sistemática. As strings de busca foram adaptadas para cada fonte de pesquisa, dado que cada base possui um mecanismo de busca próprio.

## (i) ACM Digital Library

String: (+fuzzy +bayesian +dynamic)

Filtro: Nenhum.

## (ii) IEEEExplore

String: (("Document Title": fuzzy AND "Document Title": bayesian AND "Document Title": dynamic) OR ("Abstract": fuzzy AND "Abstract": bayesian AND "Abstract": dynamic))

Filtro: "Conferences AND Journals &amp; Magazines".

## (iii) ScienceDirect

String: tak(fuzzy bayesian dynamic).

Filtro: "Research articles".

## (iv) Springer

String: fuzzy AND dynamic AND bayesian AND "dynamic bayesian network"

Filtro: "Discipline: Computer Science".

### B.3 CRITÉRIOS E PROCEDIMENTOS DE SELEÇÃO PARA TRABALHOS

Após a construção e uso das strings de busca em cada fonte de pesquisa, foram encontrados diversos trabalhos. Por esta razão, os trabalhos foram filtrados a partir dos seguintes Estágios de Seleção:

(i) Pela análise do título e resumo do trabalho;

(ii) Pela análise do resumo, introdução e conclusão do trabalho.

Para que fossem selecionados trabalhos com maior relação com o tema da pesquisa, foram definidos dois critérios de inclusão e cinco critérios de exclusão. Os critérios de inclusão definidos foram:

(i) Artigos escritos na língua inglesa;

(ii) Estudos primários que descrevem em detalhes a abordagem Fuzzy-Bayesina Dinâmica empregada.

Por sua vez, os Critérios de Exclusão determinados foram:

(i) Artigos curtos, *e.g.* resumos estendidos;

- (ii) Artigos de revisão;
- (iii) Artigos ainda em processo de impressão, ou seja, cuja versão final não foi publicada;
- (iv) Estudos que não foquem na integração de conjuntos Fuzzy a Redes Bayesianas Dinâmicas;
- (v) Estudos sobre Fuzzy C-Means Clustering ou Processos Analíticos Hierárquicos Fuzzy.

#### B.4 EXECUÇÃO DO PROTOCOLO DE REVISÃO SISTEMÁTICA DA LITERATURA

A revisão sistemática foi iniciada em março de 2018 e atualizada continuamente com o objetivo de recuperar as publicações mais recentes. A última atualização da revisão foi realizada em 28/12/2018. Depois dos testes iniciais e dos refinamentos nas strings de busca, foi iniciada a coleta e análise dos trabalhos. Para cada base de dados, a Tabela 11 exibe a data da busca, a quantidade inicial de artigos recuperados e a quantidade restante de artigos em cada um dos Estágios de Seleção descritos na Seção B.3. Inicialmente as buscas resultaram um total de 267 artigos, os quais foram reduzidos a 24 trabalhos, após a aplicação dos procedimentos de seleção.

Tabela 11 – Quantidade de Artigos recuperados em cada Estágio de Seleção

Base de Dados	Quantidade de artigos recuperados por estágio		
	Total	(i)	(ii)
(i) ACM Digital Library	9	0	0
(ii) IEEEXplore	58	21	19
(iii) ScienceDirect	43	11	5
(iv) Springer	160	4	4
-Total recuperado	270	36	28

Os trabalhos selecionados foram lidos, analisados e comparados com o objetivo de responder as perguntas de pesquisa, conforme apresentado na Seção 3.



## APÊNDICE C - Distribuições de Probabilidade para a DBN



## C.1 OCCUPANCY DETECTION DATASET

Esta seção apresenta integralmente a distribuição *a priori* sobre as variáveis não-observáveis no tempo inicial (Tabela 12), o modelo de transição (Tabela 13) e o modelo de sensores ou observações (Tabela 14) para o *Occupancy Detection Dataset*.

Tabela 12 – Distribuição *a priori* da variável não-observável *Occupancy* no tempo inicial

P(Occupancy)	
Yes	0,212406
No	0,787594

Tabela 13 – Modelo de Transição (*Occupancy Detection Dataset*)

P(Occupancy_1  Occupancy)	Yes	No
Yes	0,987201	0,003230
No	0,012799	0,996770
P(Occupancy_2  Occupancy_1)	Yes	No
Yes	0,986364	0,003007
No	0,013636	0,996993
P(Occupancy_3  Occupancy_2)	Yes	No
Yes	0,987158	0,003672
No	0,012842	0,996328

Tabela 14 – Modelo de Sensores (*Occupancy Detection Dataset*)

	Tempo 0		Tempo 1		Tempo 2		Tempo 3	
P(Temperature   Occupancy)	Yes	No	Yes	No	Yes	No	Yes	No
Low	0,011836	0,387911	0,011846	0,388048	0,011047	0,388011	0,009382	0,387652
Medium	0,583265	0,508428	0,584573	0,507869	0,586026	0,508643	0,587196	0,508979
High	0,404900	0,103661	0,403581	0,104083	0,402927	0,103346	0,403422	0,103369
P(Light   Occupancy)	Yes	No	Yes	No	Yes	No	Yes	No
Low	0,004404	0,964432	0,004408	0,963326	0,004419	0,962238	0,001932	0,962452
Medium	0,995321	0,035494	0,995317	0,036377	0,995305	0,037243	0,997792	0,037029
High	0,000275	0,000074	0,000275	0,000297	0,000276	0,000519	0,000276	0,000519
P(Humidity   Occupancy)	Yes	No	Yes	No	Yes	No	Yes	No
Low	0,331407	0,345140	0,331680	0,345286	0,333333	0,344610	0,333057	0,344242
Medium	0,451142	0,616470	0,450689	0,616555	0,446838	0,617034	0,446468	0,617617
High	0,217451	0,038390	0,217631	0,038159	0,219829	0,038356	0,220475	0,038142
P(CO2   Occupancy)	Yes	No	Yes	No	Yes	No	Yes	No
Low	0,161299	0,938813	0,163912	0,938604	0,163491	0,937978	0,163355	0,937964
Medium	0,707129	0,055320	0,704408	0,055531	0,703673	0,056161	0,704746	0,056174
High	0,131572	0,005866	0,131680	0,005865	0,132836	0,005861	0,131898	0,005862

## C.2 WALL-FOLLOWING DATASET

Esta seção apresenta integralmente a distribuição *a priori* sobre as variáveis não-observáveis no tempo inicial (Tabela 15), o modelo de transição (Tabela 16) e o modelo de sensores ou observações (Tabela 17) para o *Wall-Following Dataset*. Nas tabelas, os estados da variável *Direction* são representados pela seguinte numeração: (i) *Move\_Forward*, (ii) *Slight\_Right\_Turn*, (iii) *Sharp\_Right\_Turn* e (iv) *Slight\_Left\_Turn*.

Tabela 15 – Distribuição *a priori* da variável não-observável *Direction* no tempo inicial

P(Direction)	
(i) <i>Move_Forward</i>	0,408453
(ii) <i>Slight_Right_Turn</i>	0,146592
(iii) <i>Sharp_Right_Turn</i>	0,383552
(iv) <i>Slight_Left_Turn</i>	0,061402

Tabela 16 – Modelo de Transição (*Wall-Following Dataset*)

P(Direction_1   Direction)	(i)	(ii)	(iii)	(iv)
(i) Move_Forward	0,930244	0,059375	0,041837	0,047872
(ii) Slight_Right_Turn	0,012989	0,907589	0,019296	0,005319
(iii) Sharp_Right_Turn	0,047627	0,029018	0,938012	0,018085
(iv) Slight_Left_Turn	0,009140	0,004018	0,000854	0,928723
P(Direction_2   Direction_1)	(i)	(ii)	(iii)	(iv)
(i) Move_Forward	0,928824	0,050627	0,041042	0,051907
(ii) Slight_Right_Turn	0,014942	0,923387	0,021969	0,001059
(iii) Sharp_Right_Turn	0,047719	0,025538	0,936819	0,009534
(iv) Slight_Left_Turn	0,008515	0,000448	0,000170	0,937500
P(Direction_3   Direction_2)	(i)	(ii)	(iii)	(iv)
(i) Move_Forward	0,921895	0,063485	0,041098	0,060638
(ii) Slight_Right_Turn	0,013745	0,907618	0,019270	0,005319
(iii) Sharp_Right_Turn	0,055789	0,024956	0,938779	0,018085
(iv) Slight_Left_Turn	0,008571	0,003940	0,000853	0,915957
P(Direction_4   Direction_3)	(i)	(ii)	(iii)	(iv)
(i) Move_Forward	0,926595	0,072496	0,044733	0,052802
(ii) Slight_Right_Turn	0,013835	0,903779	0,019750	0,005388
(iii) Sharp_Right_Turn	0,052246	0,023286	0,935348	0,014009
(iv) Slight_Left_Turn	0,007324	0,000439	0,000169	0,927802
P(Direction_5   Direction_4)	(i)	(ii)	(iii)	(iv)
(i) Move_Forward	0,927576	0,074647	0,042004	0,058370
(ii) Slight_Right_Turn	0,014420	0,899735	0,021086	0,001101
(iii) Sharp_Right_Turn	0,049417	0,023410	0,936066	0,014317
(iv) Slight_Left_Turn	0,008587	0,002208	0,000843	0,926211
P(Direction_6   Direction_5)	(i)	(ii)	(iii)	(iv)
(i) Move_Forward	0,929793	0,071492	0,048818	0,063053
(ii) Slight_Right_Turn	0,010491	0,895648	0,023818	0,001106
(iii) Sharp_Right_Turn	0,050516	0,030639	0,926520	0,005531
(iv) Slight_Left_Turn	0,009199	0,002220	0,000845	0,930310

Tabela 17 – Modelo de Sensores (*Wall-Following Dataset*)

	Tempo 0				Tempo 1				Tempo 2			
P(sensor <sub>front</sub>   Direction)	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
Low	0,0002	0,0006	0,7666	0,0014	0,0002	0,0006	0,7679	0,0014	0,0002	0,0006	0,7710	0,0014
Medium	0,8758	0,9435	0,2284	0,9972	0,8749	0,9397	0,2271	0,9972	0,8760	0,9340	0,2246	0,9972
High	0,1240	0,0560	0,0050	0,0014	0,1249	0,0597	0,0050	0,0014	0,1238	0,0654	0,0043	0,0014
P(sensor <sub>left</sub>   Direction)	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
Low	0,3152	0,9988	0,4873	0,0014	0,3126	0,9988	0,4914	0,0014	0,3191	0,9988	0,4886	0,0014
Medium	0,6846	0,0006	0,5064	0,7631	0,6872	0,0006	0,5023	0,7768	0,6807	0,0006	0,5057	0,7631
High	0,0002	0,0006	0,0064	0,2355	0,0002	0,0006	0,0064	0,2218	0,0002	0,0006	0,0057	0,2355
	Tempo 3				Tempo 4				Tempo 5			
P(sensor <sub>front</sub>   Direction)	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
Low	0,0002	0,0006	0,7646	0,0014	0,0002	0,0006	0,7701	0,0015	0,0002	0,0006	0,7705	0,0015
Medium	0,8765	0,9303	0,2298	0,9971	0,8771	0,9317	0,2256	0,9971	0,8769	0,9349	0,2246	0,9971
High	0,1233	0,0691	0,0056	0,0014	0,1227	0,0677	0,0043	0,0015	0,1229	0,0645	0,0050	0,0015
P(sensor <sub>left</sub>   Direction)	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
Low	0,3101	0,9988	0,4904	0,0014	0,3158	0,9988	0,4942	0,0015	0,3159	0,9988	0,4914	0,0015
Medium	0,6897	0,0006	0,5039	0,7601	0,6839	0,0006	0,4989	0,7548	0,6839	0,0006	0,5016	0,7670
High	0,0002	0,0006	0,0056	0,2385	0,0002	0,0006	0,0070	0,2438	0,0002	0,0006	0,0070	0,2316
	Tempo 6											
P(sensor <sub>front</sub>   Direction)	(i)	(ii)	(iii)	(iv)								
Low	0,0002	0,0006	0,7741	0,0015								
Medium	0,8790	0,9359	0,2223	0,9971								
High	0,1208	0,0635	0,0036	0,0015								
P(sensor <sub>left</sub>   Direction)	(i)	(ii)	(iii)	(iv)								
Low	0,3161	0,9988	0,4921	0,0015								
Medium	0,6837	0,0006	0,5009	0,7592								
High	0,0002	0,0006	0,0070	0,2394								