



**UNIVERSIDADE FEDERAL DO CEARÁ  
DEPARTAMENTO DE COMPUTAÇÃO  
CURSO DE CIÊNCIA DA COMPUTAÇÃO**

**HENRIQUE VIANA OLIVEIRA**

**REFINAMENTO DE CONSULTAS EM LÓGICAS DE DESCRIÇÃO  
UTILIZANDO A TEORIA DOS ROUGH SETS**

**FORTALEZA, CEARÁ**

**Agosto / 2012**

**HENRIQUE VIANA OLIVEIRA**

**REFINAMENTO DE CONSULTAS EM LÓGICAS DE DESCRIÇÃO  
UTILIZANDO A TEORIA DOS ROUGH SETS**

Dissertação de mestrado apresentada à coordenação do Programa de Mestrado e Doutorado em Ciência da Computação da Universidade Federal do Ceará, como requisito para a obtenção do título de Mestre em Ciência da Computação.

Área de concentração: Lógica e Inteligência Artificial

Orientador: Prof. Dr. João Fernando Lima Alcântara

Coorientadora: Prof<sup>a</sup>. Dra. Ana Teresa de Castro Martins

**FORTALEZA, CEARÁ**

**Agosto / 2012**

A000z OLIVEIRA, H. V.  
Refinamento de Consultas em Lógicas de Descrição Utilizando a Teoria dos Rough Sets / Henrique Viana Oliveira.  
– Fortaleza, Agosto / 2012.  
111p.;il.  
Orientador: Prof. Dr. João Fernando Lima Alcântara  
Coorientadora: Prof<sup>a</sup>. Dra. Ana Teresa de Castro Martins  
Dissertação(Ciência da Computação) - Universidade Federal do Ceará, Centro de Ciências.  
1. Lógicas de Descrição 2. Rough Sets 3. Refinamento de Consultas I. Prof. Dr. João Fernando Lima Alcântara  
II. Universidade Federal do Ceará - Ciência da Computação (Mestrado) III. Mestre

CDD:000.0

**HENRIQUE VIANA OLIVEIRA**

**REFINAMENTO DE CONSULTAS EM LÓGICAS DE DESCRIÇÃO  
UTILIZANDO A TEORIA DOS ROUGH SETS**

Dissertação de mestrado apresentada à coordenação do Programa de Mestrado e Doutorado em Ciência da Computação da Universidade Federal do Ceará, como requisito para a obtenção do título de Mestre em Ciência da Computação. Área de concentração: Lógica e Inteligência Artificial

Aprovada em: \_\_/\_\_/----

**BANCA EXAMINADORA**

---

Prof. Dr. João Fernando Lima Alcântara  
Universidade Federal do Ceará - UFC  
Orientador

---

Prof<sup>a</sup>. Dra. Ana Teresa de Castro Martins  
Universidade Federal do Ceará - UFC  
Coorientadora

---

Prof. Dr. Carlos Eduardo Fish de Brito  
Universidade Federal do Ceará - UFC

---

Prof. Dr. Mário Roberto Folhadela Benevides  
Universidade Federal do Rio de Janeiro - UFRJ

*Thank you, India*  
*Thank you, terror*  
*Thank you, disillusionment*  
*Thank you, frailty*  
*Thank you, consequence*  
*Thank you,*  
*Thank you, silence*  
***Alanis Morissette***



## RESUMO

Refinamento de consulta consiste de técnicas que modificam os termos de uma consulta com o objetivo de alterar os resultados obtidos inicialmente. Para a realização de tal fim, diversas abordagens podem ser aplicadas e diferentes tipos de refinamentos podem ser considerados. Este trabalho propõe aplicar a teoria dos Rough Sets como uma nova alternativa de solução para o problema. Através das noções presentes nessa teoria, iremos desenvolver técnicas que serão aplicadas nas linguagens de Lógicas de Descrição, que são comumente utilizadas em problemas de representação de bases de conhecimento ou ontologias. Além disso, introduziremos duas extensões de Lógicas de Descrição capazes de representar as operações da teoria dos Rough Sets. Provaremos os resultados de complexidade de decisão dessas duas lógicas, assim como os resultados de complexidade das técnicas de refinamentos desenvolvidas. Por fim, mostraremos métricas de qualidade que poderão ser usadas para melhorar o resultado dos refinamentos obtidos.

Palavras-Chave: Lógicas de Descrição, Rough Sets, Refinamento de Consultas.

## **ABSTRACT**

Query Refinement consists of methods that modify the terms of a consult aiming the change of its result obtained previously. Refinements can be done of several ways and different approaches can be applied to it. This work proposes to apply methods of Query Refinement based on Rough Set theory, using it as an alternative for the refinement problem. The proposed methods will be grounded in the languages of Description Logics, which are commonly used on problems involving knowledge bases or ontologies representation. Two extensions of Description Logics with the Rough Set theory are introduced in this dissertation. We will prove the complexity of satisfiability of these logics, as well as the complexities of the query refinement methods applied to these logics. Finally, we will show quality measures which will aid to choose the results of the refinements obtained.

Keywords: Description Logic, Rough Sets, Query Refinement.

## LISTA DE FIGURAS

Figura 2.1	Representação das regiões geradas pelas aproximações de conjuntos	.....	25
------------	---	-------	----

## LISTA DE TABELAS

Tabela 1.1	Testes de segurança com carros	14
Tabela 2.1	Tabela de pessoas com problemas de visão	20
Tabela 2.2	Tabela de decisão de pessoas com problemas de visão	21
Tabela 2.3	Tabela de decisão derivada da Tabela 2.2	23
Tabela 2.4	Matriz de discernibilidade do sistema de informação da Tabela 2.3	28
Tabela 5.1	Tabelas verdade para $\wedge, \vee$ e $\neg$	47
Tabela 5.2	Tabelas verdade para $\mapsto, \supset$ e $\rightarrow$	47
Tabela 5.3	Tabelas verdade para $\wedge, \vee, \leftrightarrow$ e $\neg$	52
Tabela 5.4	Um exemplo de uma tabela de dados incompleta	55
Tabela 6.1	Tabela de Decisão $D = (\{x_1, \dots, x_7\}, \{BL, P, L\}, M, \{\mathbf{t}, \mathbf{f}, \mathbf{i}, \mathbf{u}\}, \{\mathbf{t}, \mathbf{f}, \mathbf{i}, \mathbf{u}\})$	82
Tabela 7.1	Interpretação dos conceitos $A_1, A_2, A_3$ e $C$	92
Tabela 7.2	Interpretação dos conceitos $A_1, A_2, A_3$ e $C$	96
Tabela 7.3	Tabela de decisão dos atributos $B_1, B_2$ e $D$	97

## CONTEÚDO

<b>1</b>	<b>INTRODUÇÃO</b> .....	12
<b>1.1</b>	<b>Motivação</b> .....	12
<b>1.2</b>	<b>Formulação do Problema</b> .....	15
<b>1.3</b>	<b>Contribuições</b> .....	16
<b>1.4</b>	<b>Estrutura da Dissertação</b> .....	17
<b>2</b>	<b>ROUGH SETS</b> .....	19
<b>2.1</b>	<b>Sistemas de Informação e Tabelas de Decisão</b> .....	19
<b>2.2</b>	<b>Relação de Indiscernibilidade</b> .....	21
<b>2.3</b>	<b>Aproximações de Conjuntos e <i>Rough Sets</i></b> .....	24
<b>2.4</b>	<b>Redução de Atributos</b> .....	26
<b>2.5</b>	<b>Medidas de Qualidade</b> .....	28
<b>3</b>	<b>LÓGICAS DE DESCRIÇÃO</b> .....	31
<b>3.1</b>	<b>Lógica de Descrição <math>\mathcal{AL}</math></b> .....	31
3.1.1	Sintaxe .....	31
3.1.2	Semântica .....	33
3.1.3	A Família de Linguagens de $\mathcal{AL}$ .....	34
3.1.4	Inferência .....	35
<b>3.2</b>	<b>Lógica de Descrição <i>Rough</i> <math>\mathcal{ALC}</math></b> .....	38
3.2.1	Sintaxe .....	38
3.2.2	Semântica .....	39
3.2.3	Aproximação Contextual .....	40
<b>4</b>	<b>REFINAMENTO DE CONSULTAS EM LÓGICAS DE DESCRIÇÃO</b> .....	42
<b>4.1</b>	<b>Refinamento em ABoxes - Lógica de Descrição Possibilística</b> .....	42
<b>4.2</b>	<b>Refinamento em TBoxes - Subsunção Aproximada</b> .....	44
<b>5</b>	<b>LÓGICAS DE DESCRIÇÃO <i>ROUGH</i> PARACONSISTENTE</b> .....	46
<b>5.1</b>	<b>Lógica de Belnap</b> .....	46
<b>5.2</b>	<b>Lógica de Descrição <i>Rough</i> <math>\mathcal{ALC}_4</math></b> .....	48
<b>5.3</b>	<b><i>Rough Sets</i> Paraconsistentes</b> .....	50

5.3.1	Semântica Tetravalorada .....	51
5.3.2	Conjuntos Tetravalorados .....	52
<b>5.4</b>	<b>Lógica de Descrição <i>Rough</i> Paraconsistente <math>\mathcal{ALC}</math></b> .....	<b>55</b>
5.4.1	Aproximação Contextual .....	57
5.4.2	Refinamento de Consultas .....	60
5.4.3	Complexidade de Inferência .....	65
<b>6</b>	<b>OBTENDO CONTEXTOS PARA REFINAMENTO DE CONSULTAS</b> .....	<b>76</b>
<b>6.1</b>	<b>Usando as Aproximações</b> .....	<b>76</b>
<b>6.2</b>	<b>Contextos para Aproximações Inferiores</b> .....	<b>79</b>
<b>6.3</b>	<b>Contextos para Aproximações Superiores</b> .....	<b>84</b>
<b>6.4</b>	<b>Algoritmos de Refinamento de Consultas</b> .....	<b>86</b>
<b>7</b>	<b>MEDIDAS DE QUALIDADE PARA SELEÇÃO DE CONTEXTOS</b> .....	<b>91</b>
<b>7.1</b>	<b>Medidas de Precisão e Imprecisão</b> .....	<b>92</b>
<b>7.2</b>	<b>Medidas de Granularidade</b> .....	<b>96</b>
<b>7.3</b>	<b>Medidas de Qualidade Compostas</b> .....	<b>102</b>
<b>8</b>	<b>CONCLUSÕES E TRABALHOS FUTUROS</b> .....	<b>104</b>
<b>8.1</b>	<b>Conclusões</b> .....	<b>104</b>
<b>8.2</b>	<b>Trabalhos Futuros</b> .....	<b>105</b>
	<b>BIBLIOGRAFIA</b> .....	<b>107</b>

# 1 INTRODUÇÃO

Imagine o seguinte cenário na Internet: um usuário através de uma ferramenta de busca de sítios faz uma consulta através de uma ou mais palavras-chave. Muitas vezes os resultados para suas buscas nem sempre são satisfatórios; pode ser que a quantidade de resultados relevantes seja pequena ou em contraposição, que seja grande. As causas para um número pequeno de resultados são as mais variadas: o usuário pode ter digitado alguma palavra errada ou raramente utilizada, ou pode até mesmo acontecer o caso de não existirem páginas com o determinado assunto procurado. Seria desejável que, ao ocorrer algum desses casos, a ferramenta de busca conseguisse realizar um novo tipo de consulta, modificando um pouco a consulta original quer retirando ou substituindo algumas palavras-chave para obter novos resultados. Uma intuição similar poderia ser aplicada quando o número de respostas a uma busca é relativamente grande: uma nova consulta poderia ser feita com o objetivo de limitar a quantidade de resultados, mostrando apenas os mais relevantes.

Existem várias abordagens na literatura para tratar desse problema, comumente chamado de Refinamento de Consultas, que trabalha com a ideia de relaxamentos e restrições de consultas. Esta dissertação tem como objetivo aplicar a teoria matemática dos *Rough Sets* (PAWLAK, 1982, 1991; PAWLAK; MAREK, 1984) como uma ferramenta de refinamento de consultas na linguagem das Lógicas de Descrição (BAADER, 2003), uma família de lógicas que são fragmentos decidíveis da Lógica de Primeira Ordem e que são bastante utilizadas em aplicações que envolvem base de dados e ontologias. As Lógicas de Descrição são utilizadas como uma das linguagens de representação padrão para aplicações da Web Semântica, que corresponde a um grande esforço de várias organizações para padronizar os conteúdos presentes na Web e torná-lo mais compreensivo.

## 1.1 Motivação

Pretendemos nesta dissertação investigar a teoria dos *rough sets* como uma nova abordagem para o problema de refinamentos de consultas. Dois aspectos foram fundamentais nessa escolha como solução para o problema: (i) a teoria dos *rough sets* possui uma fundamentação matemática de fácil entendimento e aplicação e (ii) trata-se de uma abordagem semântica para refinamentos em oposição à grande parte das abordagens presentes, cujo cariz é puramente sintático. Um grande problema que algumas abordagens sintáticas podem acarretar é que, ao serem aplicadas às operações de refinamento, o resultado alcançado pode não ter relação alguma com a consulta original. Quando olhamos as operações de aproximação dos *rough sets* como um método de refinamento, garantimos, através de sua semântica, que o resultado gerado tem algum grau de proximidade com a consulta original.

Uma característica precípua da teoria dos *rough sets* é a sua capacidade de representar e raciocinar com incerteza via operações de aproximação. É fácil reconhecer que a incerteza possui um importante papel nas tarefas realizadas pelos seres humanos. Por exemplo, as decisões que tomamos diariamente são cercadas por incertezas. E realizar decisões conscientes

em todas as suas variedades é talvez a capacidade mais fundamental dos seres humanos. Para entender esta capacidade de lidar com a incerteza, nós precisamos entender primeiramente o que é incerteza.

A noção de incerteza começou a ser explorada nas ciências no fim do século XIX com métodos estatísticos aplicados na Física e na Química (KLIR, 2005). Posteriormente, no começo do século XX, esses métodos estatísticos (a teoria das probabilidades) foram aceitos pela comunidade científica como uma legítima área da ciência. O estudo da incerteza tornou-se reconhecido como útil ou até mesmo essencial em certas questões científicas. Entretanto, era acreditado que a incerteza poderia sempre ser tratada com a teoria das probabilidades. Foi levado mais que metade de um século para reconhecer que o conceito de incerteza é muito vasto para ser capturado apenas pela teoria das probabilidades, e conseqüentemente começar os estudos de suas várias manifestações não-probabilísticas.

Existem várias manifestações de informação incerta: ela pode ser incompleta, imprecisa, aleatória, duvidosa ou contraditória. Em geral, essas deficiências de informação determinam o tipo de incerteza associada e o tipo de teoria matemática a ser aplicada. Raciocinar com incertezas é uma tarefa muito importante para diversos campos da Inteligência Artificial (IA), especialmente para sistemas especialistas, robótica e redes neurais (RUSSELL; NORVIG, 2010). Quando lidamos com sistemas clássicos de representação de conhecimento em IA, trabalhamos com a ideia de que algo é verdadeiro ou falso. O raciocínio com incerteza aplicado a esses sistemas vai além dessa suposição, podendo falar de diversas noções de verdade. Podemos ter casos em que algo pode ter valor verdadeiro ou falso, mas não sabemos qual é esse valor, então calculamos a probabilidade de ele ser verdadeiro ou falso. Alternativamente, pode existir outra situação em que algo será classificado como parcialmente verdadeiro. Essas situações citadas são comuns em diversas aplicações como em diagnósticos de doenças, em que um conjunto de sintomas pode definir a probabilidade de uma pessoa possuir determinada doença. Um outro evento suscetível a acontecer é quando uma propriedade não é definida precisamente, como no caso de determinar se uma pessoa está realmente saudável.

Apesar de serem comuns e de fundamental importância, esses e outros casos de incerteza não podem ser representados em abordagens clássicas de sistemas de representação de conhecimento. Para resolver esse problema, diversas formas de raciocínio com incerteza são estudadas e aplicadas a esses sistemas. As abordagens mais bem sucedidas são a teoria das probabilidades (KLIR, 2005), a teoria dos *fuzzy sets* (ZADEH, 1965) e a teoria das possibilidades (KLIR, 2005). Cada uma dessas abordagens trata de um tipo específico de incerteza: a aleatoriedade, a imprecisão e a possibilidade, respectivamente.

- **Aleatoriedade:** É a incerteza presente em um caso com clara definição, mas que nem sempre ocorre, pois podem haver instabilidades no que tange a um elemento pertencer a um certo conceito ou de um evento acontecer. É uma incerteza estatística que indica a probabilidade de um evento acontecer;
- **Imprecisão:** É a incerteza incluída em um mundo com definições incompletas. Essa imprecisão é medida por uma função de pertinência *fuzzy*, presente na teoria dos *fuzzy sets* (ZADEH, 1965), que lidam com a similaridade (pertinência) de um elemento em um

Carro	Estação	Seguro
$a$	$s_1$	sim
$a$	$s_2$	não
$b$	$s_2$	não
$c$	$s_1$	sim
$c$	$s_2$	sim
$d$	$s_1$	sim

Tabela 1.1: Testes de segurança com carros

determinado conceito. É uma incerteza que representa fatos que não podem ser precisamente descritos, como por exemplo, descrições em termos das noções de alto, baixo, grande, pequeno, rápido, lento, claro, escuro, cheio, etc.;

- **Possibilidade:** Trata da capacidade de um determinado evento ser verdadeiro ou de ele ocorrer. A noção parece bem similar à ideia de aleatoriedade, mas a diferença é que quando falamos de aleatoriedade, existe um modelo matemático (probabilístico) utilizado na representação e interpretação dos eventos. Já com relação à possibilidade, esse modelo matemático não está presente. A possibilidade pode então ser vista como uma crença ou opinião sobre a veracidade de um evento.

A teoria dos *rough sets* (PAWLAK, 1982) surge como uma abordagem específica da teoria das possibilidades e que trata também da noção de crença. A ideia é que a partir de um conjunto de informações ou eventos disponíveis, seja possível deduzir qualitativamente informações (incertas) que não estão explicitamente apresentadas, mas que existe a possibilidade de acontecerem.

Para ilustrar a motivação dessa teoria, suponha que temos o exemplo da Tabela 1.1, em que um teste foi realizado com alguns carros. Todos os carros foram testados independentemente em duas estações diferentes, as quais definiram se eles estavam seguros ou não para serem dirigidos. Podemos ver que o carro  $b$  foi testado apenas pela estação  $s_2$  e diagnosticado como não sendo seguro. Já o carro  $c$  foi testado pelas estações  $s_1$  e  $s_2$  e diagnosticado como sendo seguro em ambas estações. Porém, temos que o carro  $a$  foi determinado como sendo seguro pela estação  $s_1$  e como não sendo seguro pela estação  $s_2$ . Temos, então, aqui um tipo de informação incerta. A teoria dos *rough sets* propõe-se a identificar e raciocinar sobre essas incertezas, que não estão explicitamente definidas em uma representação de dados. Para isso, são feitas aproximações nesses dados com o objetivo de identificar os objetos a apresentar algum tipo de informação incerta. Considerando o exemplo anterior, poderíamos descobrir através das aproximações, por exemplo, quais carros podem ser considerados necessariamente seguros ou também quais carros são considerados possivelmente seguros.

## 1.2 Formulação do Problema

Uma das formas mais comuns de buscar informação na Web ou em uma base de dados é através de buscas baseadas em palavras-chave: o usuário digita uma consulta consistindo de uma ou mais palavras-chave e o sistema de busca retorna uma lista de páginas da Web ordenadas de acordo com alguma medida de relevância. O mesmo procedimento é geralmente usado em aplicações de comércio eletrônico que tentam relacionar as consultas dos usuários aos produtos do catálogo de algumas companhias.

Em uma abordagem básica, as páginas não são retornadas como resultados de busca se elas não contêm exatamente as palavras-chave da consulta. Existem várias razões pela qual uma consulta pode falhar. Um problema encontrado é a incompatibilidade de palavras: o usuário sabe o que ele está procurando e é capaz de descrevê-lo, mas os termos da consulta utilizados não correspondem aos termos utilizados nos documentos contendo a informação desejada, porque existem diferenças de vocabulários. Esse problema é ainda mais significativo no contexto da Web, pois há uma enorme quantidade de fontes heterogêneas de informação, que se expressam em diferentes jargões ou até mesmo em diferentes idiomas. Além das diferenças no vocabulário, é também comum um usuário não ser capaz de descrever precisamente o que ele está procurando; o conhecido fenômeno “Eu vou saber quando eu encontrar”. Além disso, muitos termos na linguagem são ambíguos. Por exemplo, um usuário consultando a palavra “java” pode tanto estar querendo informações sobre a linguagem de programação, um tipo de café ou uma ilha da Indonésia. Para satisfazer os usuários que esperam que as ferramentas de busca produzam resultados que digam “o que eles querem e o que eles não querem”, técnicas mais sofisticadas precisam ser usadas ao invés de apenas retornar os documentos que contêm os termos dados pelos usuários. Uma das opções a ser usada é o refinamento de consultas (COCK; CORNELIS, 2005).

A teoria dos *rough sets* vem sendo exitosamente empregada no processo de refinamento de consultas (COCK; CORNELIS; KERRE, 2004; COCK; CORNELIS, 2005). De fato, uma consulta pode ser tratada como um conjunto, no qual podem ser aplicadas as operações de aproximação. *Rough sets* podem ser vistos como mais uma ferramenta utilizada na área de descoberta de conhecimento. Alguns resultados interessantes da relação entre *rough sets* e outras abordagens para raciocínio sob incerteza foram reportados na literatura (KOMOROWSKI et al., 1998).

Um desses resultados mostra que essa teoria tem provado ser útil no campo de mineração de dados e descoberta de conhecimento (FAYYAD et al., 1996), onde ela oferece meios matemáticos para a descoberta de padrões ocultos nos dados, identificação de dependências parciais ou totais em bases de dados, eliminação de dados redundantes, realização de aproximações com valores nulos, dados perdidos, dados dinâmicos, entre outros aspectos. Além disso, a teoria dos *rough sets* proporciona métodos de classificação de dados (ZHANG; YAO, 2004; YAO, 2010) em consonância com diversos critérios, que auxiliam na escolha das respostas obtidas durante a análise dos dados.

Um fator positivo com relação a essa teoria é que ela vem sendo aplicada jun-

tamente com outros tipos de abordagens, gerando assim diversas extensões para *rough sets*. Uma destas extensões trata de representar e de aproximar dados inconsistentes e desconhecidos (VITÓRIA; SZALAS; MAŁUSZYŃSKI, 2008). Inconsistências surgem naturalmente na descrição do mundo real e isso ocorre em vários contextos. Uma forma de lidar com elas em sistemas de representação de conhecimento é através de lógicas paraconsistentes (COSTA, 1993; JAŚKOWSKI, 1948; BELNAP, 1977). Também chamadas de lógicas tolerantes a inconsistências, as lógicas paraconsistentes diferem da lógica clássica pois não seguem o princípio de que ao se chegar a uma contradição, qualquer informação pode ser inferida. Essa é uma característica importante, pois a partir dela, é possível tratar informação inconsistente de uma forma controlada e discriminatória.

Como dito previamente, a teoria dos *rough sets* permite a integração com outros tipos de abordagens, assim como a sua representação em diversas linguagens de representação de conhecimento. No caso das Lógicas de Descrição, que formalizam a descrição de ontologias ou bases de conhecimento com o intuito de modelar conhecimento do mundo real, a teoria dos *rough sets* nos permite inferir informação que não estão explicitamente definidas nas ontologias ou bases de conhecimento, assim como representar um novo tipo de informação.

### 1.3 Contribuições

Seguem abaixo as principais contribuições desta dissertação.

- Definição de extensões de Lógicas de Descrição nas quais serão aplicadas os refinamentos de consultas. Essas extensões permitem a representação de conhecimento incompleto e contraditório, seguindo os princípios de lógicas paraconsistentes. Além disso, essa nova lógica permite a representação das operações de aproximação dos *rough sets*. Consequentemente, é permitido o raciocínio sobre um determinado tipo de informação incerta assim como realizar aproximações mesmo com a presença de informação incompleta e contraditória;
- Prova de um limite superior de complexidade para o problema de satisfação dessas lógica criadas;
- Utilização da teoria dos *rough sets* como uma ferramenta de refinamento de consultas;
- Desenvolvimento de um método de refinamento de consultas para Lógicas de Descrição. Destacamos que este é o primeiro método de refinamento em Lógicas de Descrição envolvendo *rough sets*; ele consiste em um conjunto de algoritmos que buscam possíveis refinamentos para as fórmulas de Lógica de Descrição consultadas;
- Demonstração dos resultados de complexidade dos algoritmos de refinamento;
- Definição de métricas nos possíveis resultados dos algoritmos que visam selecionar os melhores refinamentos para uma consulta.

Como publicações desenvolvidas a partir das contribuições desta Dissertação de Mestrado, temos os artigos **Paraconsistent Rough Description Logic** (VIANA; ALCÂNTARA; MARTINS, 2011), publicado no 24th International Workshop on Description Logics, DL 2011 e **Searching Contexts in a Paraconsistent Rough Description Logic** (VIANA; ALCÂNTARA; MARTINS, 2012), submetido ao 25th International Workshop on Description Logics, DL 2012. Além disso, foram desenvolvidos trabalhos paralelos durante o mestrado, que envolvem alguns assuntos presentes aqui e que contribuíram no desenvolvimento desta dissertação, mas que não possuem relação direta com o seu tema principal. São eles

- **A Tractable Paraconsistent Fuzzy Description Logic** (VIANA et al., 2010) (publicado no 6th International Workshop on Uncertainty Reasoning for the Semantic Web, URSW 2010). Nesse artigo, propomos uma Lógica de Descrição paraconsistente e *fuzzy* com complexidade de decisão polinomial;
- **A Seven-Valued Extension of Rough Sets** (VIANA; ALCÂNTARA; MARTINS, 2011) (submetido ao 6th Workshop on Logical and Semantic Frameworks, with Applications, LSFA 2011). Nesse artigo, propomos uma extensão da teoria dos *rough sets* para uma lógica multi-valorada com sete valores verdade que permitem a representação de vários níveis de incerteza;
- **A Description Logic for InferenceNet.Br** (FRANCO et al., 2012) (publicado no International Conference on Computational Processing of Portuguese, PROPOR 2012). Nesse artigo, utilizamos uma Lógica de Descrição que é capaz de representar as informações de uma base de conhecimento de senso comum, chamada InferenceNet.Br, e que é utilizada em aplicações envolvendo Processamento de Linguagem Natural. Para isso, mostramos uma tradução das fórmulas de Lógica de Descrição para as bases de conhecimento da InferenceNet.Br e vice-versa.

#### 1.4 Estrutura da Dissertação

O presente trabalho está organizado da seguinte maneira:

- No Capítulo 2, apresentaremos definições básicas da teoria dos *rough sets*;
- No Capítulo 3, apresentaremos uma introdução às Lógicas de Descrição e Lógicas de Descrição *Rough*;
- No Capítulo 4, mostraremos abordagens presentes na literatura sobre refinamento de consultas em Lógicas de Descrição;
- No Capítulo 5, começaremos a apresentar as contribuições desta dissertação, introduzindo duas Lógicas de Descrição *Rough* Paraconsistentes, chamadas de *Rough  $\mathcal{ALC}_4$*  e *PR $\mathcal{ALC}$* , com seus respectivos resultados de complexidade de satisfação;

- No Capítulo 6, definiremos um método automático de refinamento de consultas que será aplicado à *Rough  $\mathcal{ALC}_4$*  e  *$PR_{\mathcal{ALC}}$* . Apresentaremos dois algoritmos de refinamento com seus resultados de complexidade;
- No Capítulo 7, definiremos medidas de qualidade para refinamentos, que são baseadas em vários critérios distintos e que possuem como objetivos escolher os melhores resultados dentre os obtidos pelo método automático de refinamento de consultas;
- No Capítulo 8, finalizaremos esta dissertação com conclusões e possíveis trabalhos futuros.

## 2 ROUGH SETS

Este capítulo apresenta um breve resumo da teoria dos *rough sets* com suas principais características. A teoria dos *rough sets* foi introduzida por Zdzislaw Pawlak (PAWLAK, 1982, 1991; PAWLAK; MAREK, 1984) em 1982 como uma metodologia para tratar qualitativamente incertezas em conjuntos de dados. A sua ideia fundamental é que diversos objetos podem parecer similares quando analisados a partir de um determinado ponto de vista devido às limitações do nosso conhecimento. Logo, é mais simples distinguir classes de objetos ao invés de objetos individualmente. Consequentemente, descrições aproximadas de conjuntos de objetos similares podem ser construídas.

Na Seção 2.1, apresentaremos as ideias de sistemas de informação e tabelas de decisão, que serão usadas na representação dos dados. Na Seção 2.2, definiremos a noção de indiscernibilidade, que também pode ser definida como uma relação de equivalência entre os objetos nas tabelas. Na Seção 2.3, apresentaremos a noção de aproximação de conceitos, e formalizaremos a noção de *rough set*. Na Seção 2.4, mostraremos medidas de qualidade associadas com as aproximações de conceitos.

### 2.1 Sistemas de Informação e Tabelas de Decisão

Em muitos problemas um conjunto de dados pode ser representado como relações em uma base de dados ou como uma tabela, em que cada linha representa um caso, um evento, um paciente, ou simplesmente um objeto. Cada coluna dessa tabela, por sua vez, representa um atributo (uma variável, uma observação, uma relação, uma propriedade, etc.) que pode ser medida para cada objeto. Para esse tipo de representação, daremos o nome de *sistema de informação*.

**Definição 1 (Sistema de Informação)** *Um sistema de informação é uma tupla  $S = (U, A, V_A)$ , em que  $U$ , chamado de universo, é um conjunto finito não-vazio de objetos,  $A$  é um conjunto finito não-vazio de atributos, tal que cada atributo  $a \in A$  pode ser visto como uma função total  $a : U \rightarrow V_A$  e  $V_A$  é o conjunto de possíveis valores dos atributos em  $A$ .*

*Exemplo 1.* Na Tabela 2.1, o conjunto universo é dado por  $U = \{x_1, x_2, \dots, x_{14}\}$ , que representa um grupo de pessoas. O conjunto de atributos é dado por  $A = \{Ast, Idade, ProdL, \acute{O}culos\}$ , em que cada atributo tem o seguinte significado: *Ast* representa astigmatismo e o seu conjunto de possíveis valores são 0 (sem astigmatismo) ou 1 (com astigmatismo); *Idade* pode receber os valores 0 (pessoas com até 20 anos de idade), 1 (pessoas entre 20 e 50 anos de idade), e 2 (pessoas acima de 50 anos de idade); *ProdL* representa produção lacrimal e pode ter os valores 1 (produção lacrimal reduzida) ou 2 (produção lacrimal normal); e *Óculos* pode assumir os valores 0 (não utilizando óculos) e 1 (utilizando óculos).

Note que, para os objetos  $x_1$  e  $x_6$ , assim como para  $x_5$  e  $x_{10}$ , os valores de seus atributos são exatamente os mesmos. Nesse caso dizemos que tais objetos são indiscerníveis

	Ast	Idade	ProdL	Óculos
$x_1$	0	0	1	0
$x_2$	1	1	1	0
$x_3$	1	0	2	1
$x_4$	1	1	1	0
$x_5$	1	0	1	0
$x_6$	0	0	1	0
$x_7$	0	2	1	1
$x_8$	0	2	1	1
$x_9$	0	0	2	1
$x_{10}$	1	0	1	0
$x_{11}$	1	0	2	1
$x_{12}$	1	2	2	1
$x_{13}$	0	0	2	1
$x_{14}$	1	0	2	1

Tabela 2.1: Tabela de pessoas com problemas de visão

no que tange aos atributos avaliados, i.e.,  $x_1$  é indiscernível de  $x_6$ , e  $x_5$  é indiscernível de  $x_{10}$ . Isso quer dizer que analisando os objetos através do conjunto de atributos dado, não podemos distinguir  $x_1$  de  $x_6$ , nem  $x_5$  de  $x_{10}$ .

Em muitas aplicações, é possível classificar os objetos com uma determinada propriedade a partir de um conjunto de atributos. Trata-se de um conhecimento posterior que é expresso por um atributo distinto chamado de *atributo de decisão*. Por exemplo, em um teste realizado para identificar se pessoas possuem determinada doença; as pessoas correspondem os objetos da tabela, um conjunto de sintomas que essas pessoas podem possuir correspondem os atributos, e o resultado do teste confirmando se a pessoa tem ou não a doença é representado pelo atributo de decisão. *Sistemas de informação* acrescidos de um *atributo de decisão* são chamados de *tabela de decisão*:

**Definição 2 (Tabela de Decisão)** *Uma tabela de decisão é uma tupla  $D = (U, A, d, V_A, V_d)$ , em que  $U$  é um conjunto finito não-vazio de objetos chamado de universo;  $A$  é um conjunto finito não-vazio de atributos, tal que para cada  $a \in A$ ,  $a : U \rightarrow V_A$ , em que  $a$  é uma função total e  $V_A$  é o conjunto de possíveis valores dos atributos em  $A$ ;  $d$  é um atributo de decisão, tal que  $d$  é uma função total, ou seja,  $d : U \rightarrow V_d$ , em que  $V_d$  é o conjunto de possíveis valores de  $d$ .*

*Exemplo 2.* Na tabela de decisão  $D = (U, A, Lentes, \{0, 1, 2\}, \{0, 1\})$  apresentada na Tabela 2.2, o atributo de decisão Lentes representa lentes de contato, o qual pode receber os valores 0 (representando as pessoas que começaram a utilizar lentes, mas pararam após um tempo) ou 1 (as pessoas que continuaram a utilizar lentes). Podemos então dividir o conjunto de pessoas em dois grupos: o grupo daquelas que abandonaram o uso de lentes de contato e o grupo daquelas que não abandonaram. Ou seja, cada objeto  $x \in U$  pode ser classificado como  $Lentes(x) = 0$  ou  $Lentes(x) = 1$ .

	Ast	Idade	ProdL	Óculos	Lentes
$x_1$	0	0	1	0	0
$x_2$	1	1	1	0	0
$x_3$	1	0	2	1	1
$x_4$	1	1	1	0	0
$x_5$	1	0	1	0	0
$x_6$	0	0	1	0	0
$x_7$	0	2	1	1	1
$x_8$	0	2	1	1	1
$x_9$	0	0	2	1	0
$x_{10}$	1	0	1	0	0
$x_{11}$	1	0	2	1	1
$x_{12}$	1	2	2	1	1
$x_{13}$	0	0	2	1	1
$x_{14}$	1	0	2	1	0

Tabela 2.2: Tabela de decisão de pessoas com problemas de visão

Uma importante pergunta a ser feita é: quando podemos recomendar a uma pessoa o uso de lentes de contato baseado nos valores dos atributos Ast, Idade, ProdL e Óculos? Para responder essa pergunta, precisamos construir uma descrição do conceito Lentes em termos das classes de indiscernibilidade (classes de equivalência) que serão detalhadas na próxima seção.

## 2.2 Relação de Indiscernibilidade

Uma tabela de decisão expressa o conhecimento disponível sobre um determinado conjunto de dados. Porém essa tabela pode ser desnecessariamente grande, pois pode conter informação redundante em pelo menos dois aspectos: o primeiro aspecto é que os objetos indiscerníveis podem ser representados diversas vezes, e o segundo aspecto é que alguns de seus atributos podem ser supérfluos, i.e., determinados atributos podem não ser necessários para distinguir os objetos com relação ao atributo de decisão. A seguir, apresentaremos o primeiro aspecto de redução, que está diretamente ligado com as noções de relação de indiscernibilidade, aproximação de conjuntos e *rough sets*, a começar pela relação de indiscernibilidade.

Inicialmente, iremos relembrar a noção de relação de equivalência. Uma relação binária  $R \subseteq U \times U$  é uma relação de equivalência se ela é reflexiva ( $\forall x \in U, R(x, x)$ ), simétrica ( $\forall x, y \in U$ , se  $R(x, y)$  então  $R(y, x)$ ) e transitiva ( $\forall x, y, z \in U$ , se  $R(x, y)$  e  $R(y, z)$ , então  $R(x, z)$ ). Através de uma relação de equivalência, podemos particionar o domínio em classes de equivalência. A classe de equivalência de um elemento  $x \in U$  com relação a  $R$ , denotada por  $[x]_R$ , consiste de todos os objetos  $y \in U$  tal que  $R(x, y)$ . Na sequência, temos a definição formal de relação de indiscernibilidade.

**Definição 3 (Relação de Indiscernibilidade)** *Seja  $D = (U, A, d, V_A, V_d)$  uma tabela de decisão.*

Para qualquer  $B \subseteq A$ , define-se a relação de indiscernibilidade  $IND(B)$  como sendo a relação de equivalência

$$IND(B) = \{(x, y) \in U \times U \mid \forall a \in B, a(x) = a(y)\}.$$

Se  $(x, y) \in IND(B)$ , então temos que os objetos  $x$  e  $y$  são indiscerníveis em cada atributo de  $B$ . A classe de equivalência de um objeto  $x$  com relação a um conjunto de atributos  $B$  é definida por  $[x]_B = \{y \in U \mid (x, y) \in IND(B)\}$ . Se considerarmos, por exemplo, os atributos  $\{Ast, Idade, ProdL, \acute{O}culos\}$  na Tabela 2.2, a relação  $IND$  induz diversas partições do universo; algumas delas são

- $IND(\{Idade\})$  induz três partições do universo:  $\{x_1, x_3, x_5, x_6, x_9, x_{10}, x_{11}, x_{13}, x_{14}\}$ ,  $\{x_2, x_4\}$  e  $\{x_7, x_8, x_{12}\}$ ;
- $IND(\{Ast, Idade\})$  induz cinco partições do universo:  $\{x_1, x_6, x_9, x_{13}\}$ ,  $\{x_2, x_4\}$ ,  $\{x_3, x_5, x_{10}, x_{11}, x_{14}\}$ ,  $\{x_7, x_8\}$  e  $\{x_{12}\}$ ;
- $IND(\{Ast, Idade, ProdL\})$  induz sete partições do universo:  $\{x_1, x_6\}$ ,  $\{x_2, x_4\}$ ,  $\{x_3, x_{11}, x_{14}\}$ ,  $\{x_5, x_{10}\}$ ,  $\{x_7, x_8\}$ ,  $\{x_9, x_{13}\}$  e  $\{x_{12}\}$ .

Analisando  $IND$  com relação ao atributo  $\{Idade\}$ , temos que os objetos  $x_7$ ,  $x_8$  e  $x_{12}$  pertencem à mesma classe de equivalência e portanto são indiscerníveis. Podemos então criar uma nova tabela, em que os objetos são representados pelas suas respectivas classes de equivalência. Consequentemente, existe a possibilidade de reduzir o tamanho de uma tabela dado que se temos pelo menos dois objetos indiscerníveis, a tabela com a descrição de classes de objetos utilizará apenas uma linha para representá-los.

*Exemplo 3.* A tabela de decisão  $D' = (U', A, Lentes', \{0, 1, 2\}, \{\{0\}, \{1\}, \{0, 1\}\})$ , apresentada na Tabela 2.3, é derivada da Tabela 2.2. A coluna Classe designa uma classe de equivalência dos objetos da Tabela 2.2. Logo o universo  $U'$  é composto por classes de equivalência. O conjunto de atributos permanece como anteriormente, ou seja,  $A = \{Ast, Idade, ProdL, \acute{O}culos\}$ . Os valores para o atributo de decisão  $Lentes'$  são agora subconjuntos não-vazios de  $\{0, 1\}$ . Se uma classe de equivalência  $C$  possui apenas objetos em que o resultado para o atributo  $Lentes$  é 0 (1), então o atributo de decisão  $Lentes'$  terá o valor  $\{0\}$  ( $\{1\}$ ). Entretanto, se alguns objetos na classe de equivalência  $C$  possuem resultado 0 para  $Lentes$  enquanto outros objetos pertencentes a  $C$  possuem resultado 1, então o atributo de decisão  $Lentes'$  terá o valor  $\{0, 1\}$ .

Cotejando a Tabela 2.3 com a Tabela 2.2, identificamos 7 classes de equivalência levando em consideração  $IND(\{Ast, Idade, ProdL, \acute{O}culos\})$ :

$$\begin{aligned} C_1 &= \{x_1, x_6\}, \\ C_2 &= \{x_2, x_4\}, \\ C_3 &= \{x_3, x_{11}, x_{14}\}, \\ C_4 &= \{x_5, x_{10}\}, \\ C_5 &= \{x_7, x_8\}, \\ C_6 &= \{x_9, x_{13}\}, \\ C_7 &= \{x_{12}\}. \end{aligned}$$

Classe	Ast	Idade	ProdL	Óculos	Lentes'
$C_1$	0	0	1	0	{0}
$C_2$	1	1	1	0	{0}
$C_3$	1	0	2	1	{0, 1}
$C_4$	1	0	1	0	{0}
$C_5$	0	2	1	1	{1}
$C_6$	0	0	2	1	{0, 1}
$C_7$	1	2	2	1	{1}

Tabela 2.3: Tabela de decisão derivada da Tabela 2.2

Note que na terceira linha da Tabela 2.3, podemos ver que a classe de equivalência  $C_3$  contém alguns objetos cujo resultado para o atributo de decisão Lentes é 0 ( $x_{14}$ ) enquanto que outros objetos em  $C_3$  possuem resultado 1 ( $x_3$  e  $x_{11}$ ).

*Exemplo 4.* Considere novamente a Tabela 2.2. O atributo de decisão dessa tabela representa dois conceitos: aqueles indivíduos que continuaram a usar lentes de contato (Lentes) e aqueles que pararam de usar lentes de contato ( $\neg$ Lentes, ou simplesmente o complemento de Lentes). O primeiro conceito pode ser descrito pelo seguinte conjunto

$$\text{Lentes} = \{x_3, x_7, x_8, x_{11}, x_{12}, x_{13}\},$$

enquanto o segundo conceito pode ser descrito pelo conjunto

$$\neg\text{Lentes} = \{x_1, x_2, x_4, x_5, x_6, x_9, x_{10}, x_{14}\}.$$

Voltando à pergunta feita no final da Seção 2.1:

- Quando podemos recomendar a uma pessoa o uso de lentes de contato, baseado nos atributos Ast, Idade, ProdL e Óculos e em seus valores mostrados na tabela?

Através da descrição obtida em termos das classes de equivalência feita acima, temos, por exemplo, que  $C_5 \subseteq \text{Lentes}$ . Logo uma condição suficiente para satisfazer a questão apontada é que as pessoas possuam as seguintes características:

$$\text{Ast} = 0 \text{ e Idade} = 2 \text{ e ProdL} = 1 \text{ e Óculos} = 1.$$

Da mesma forma, temos que  $C_7 \subseteq \text{Lentes}$ . Logo, uma outra resposta que satisfaz a questão é

$$\text{Ast} = 1 \text{ e Idade} = 2 \text{ e ProdL} = 2 \text{ e Óculos} = 1.$$

Contudo, as classes de equivalência  $C_3$  e  $C_6$  parecem apresentar um problema quando analisadas nesse contexto, pois  $C_3 \not\subseteq \text{Lentes}$  e  $C_6 \not\subseteq \text{Lentes}$ , assim como também  $C_3 \not\subseteq \neg\text{Lentes}$

e  $C_6 \not\subseteq \neg\text{Lentes}$ . Isso indica que os conceitos Lentes e  $\neg\text{Lentes}$  não podem ser descritos precisamente usando as informações disponíveis, pois existem elementos que estão na mesma classe de equivalência, mas possuem diferentes pertinências para esses conceitos.

A partir do exemplo acima, podemos concluir que por conta das informações contraditórias, nem sempre será possível obter descrições precisas de um conjunto. É nesse cenário que a noção de *rough sets* irá emergir. Para tal, vamos introduzir a ideia de aproximação de conjuntos.

### 2.3 Aproximações de Conjuntos e *Rough Sets*

Como visto na seção anterior, um conceito pode não estar definido de maneira precisa. É aqui que a noção de *rough sets* surge. Apesar de não podermos defini-lo precisamente, é possível delinear os objetos que certamente pertencem a esse conceito, os objetos que certamente não pertencem a esse conceito e, finalmente, os objetos que estão na fronteira entre pertencer ou não a esse conceito. Essas noções serão formalizadas na sequência. Primeiramente, apresentamos a ideia de aproximações de conjuntos:

**Definição 4 (Aproximação Superior/Inferior)** *Sejam  $D = (U, A, d, V_A, V_d)$  uma tabela de decisão,  $B \subseteq A$  e  $X \subseteq U$ . A teoria dos rough sets introduz dois tipos de aproximação para o conjunto  $X$ , chamadas de aproximações superior e inferior de  $X$  em relação a  $B$ , denotadas por  $\overline{X}^B$  e  $\underline{X}_B$ , respectivamente, em que*

$$\overline{X}^B = \{x \in U \mid [x]_B \cap X \neq \emptyset\} \text{ e } \underline{X}_B = \{x \in U \mid [x]_B \subseteq X\}.$$

Os objetos em  $\underline{X}_B$  são classificados como os membros que necessariamente pertencem a  $X$  em relação aos atributos de  $B$ , enquanto os objetos em  $\overline{X}^B$  são classificados como os possíveis membros de  $X$  em relação aos atributos de  $B$ . Quando todos os atributos presentes na tabela de decisão (com exceção do atributo de decisão) são utilizados no cálculo das aproximações inferior e superior de  $X$ , usaremos as notações  $\underline{X}$  e  $\overline{X}$ , em vez de  $\underline{X}_A$  e  $\overline{X}^A$ .

*Exemplo 5.* Considere novamente a Tabela 2.2 e os seguintes conjuntos

$$\text{Lentes} = \{x_3, x_7, x_8, x_{11}, x_{12}, x_{13}\} \text{ e } \neg\text{Lentes} = \{x_1, x_2, x_4, x_5, x_6, x_9, x_{10}, x_{14}\}.$$

As aproximações inferior e superior para esses conjuntos são dadas por

- $\underline{\text{Lentes}} = \{x \in U \mid [x]_A \subseteq \text{Lentes}\} = \{x \in U \mid [x]_A \subseteq \{x_3, x_7, x_8, x_{11}, x_{12}, x_{13}\}\} = \{x_7, x_8, x_{12}\};$
- $\overline{\text{Lentes}} = \{x \in U \mid [x]_A \cap \text{Lentes} \neq \emptyset\} = \{x \in U \mid [x]_A \cap \{x_3, x_7, x_8, x_{11}, x_{12}, x_{13}\} \neq \emptyset\} = \{x_3, x_7, x_8, x_9, x_{11}, x_{12}, x_{13}, x_{14}\};$
- $\underline{\neg\text{Lentes}} = \{x \in U \mid [x]_A \subseteq \neg\text{Lentes}\} = \{x \in U \mid [x]_A \subseteq \{x_1, x_2, x_4, x_5, x_6, x_9, x_{10}, x_{14}\}\} = \{x_1, x_2, x_4, x_5, x_6, x_{10}\};$

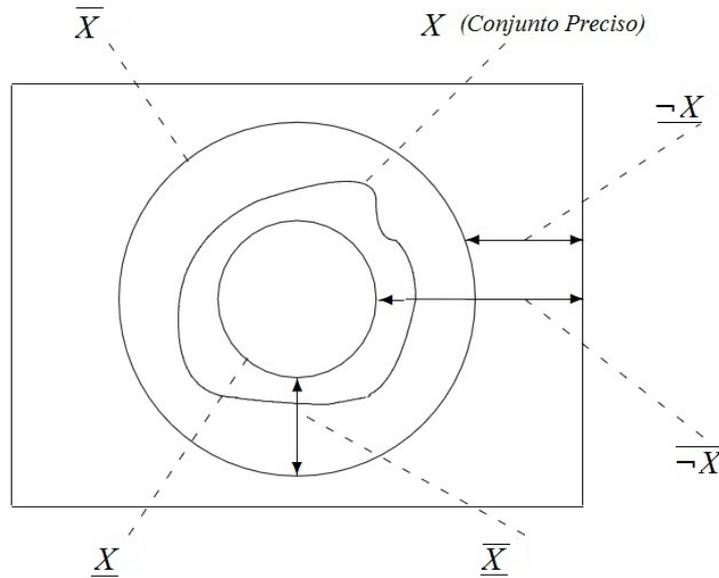


Figura 2.1: Representação das regiões geradas pelas aproximações de conjuntos

- $\overline{\neg\text{Lentes}} = \{x \in U \mid [x]_A \cap \neg\text{Lentes} \neq \emptyset\} = \{x \in U \mid [x]_A \cap \{x_1, x_2, x_4, x_5, x_6, x_9, x_{10}, x_{14}\} \neq \emptyset\} = \{x_1, x_2, x_3, x_4, x_5, x_6, x_9, x_{10}, x_{11}, x_{13}, x_{14}\}$ .

A aproximação superior  $\overline{\text{Lentes}}$  descreve as pessoas que possivelmente não terão problemas em utilizar lentes de contato, enquanto a aproximação inferior  $\underline{\text{Lentes}}$  descreve aquelas pessoas que necessariamente terão problemas devido ao uso de lentes de contato. Note que os objetos  $x_3, x_{11}$  e  $x_{14}$  pertencem tanto a  $\overline{\text{Lentes}}$  como a  $\underline{\text{Lentes}}$ . Esse fato indica que existem evidências contraditórias e, portanto, não é possível declarar precisamente se para tais casos, às pessoas com

$$\text{Ast} = 1 \text{ e Idade} = 0 \text{ e ProdL} = 2 \text{ e Óculos} = 1,$$

deveria ser recomendado o uso de lentes de contato. Note que algumas dúvidas podem existir nessas descrições, pois podem surgir objetos que pertencem a  $\overline{X}$  (i.e. objetos que estão classificados como possíveis membros de  $X$ ) mas não pertencem a  $\underline{X}$  (objetos que estão classificados como certamente membros de  $X$ ). Denotamos por *fronteira* o conjunto  $\overline{X} - \underline{X}$  e que corresponde a esses casos conflitantes.

As regiões geradas pelas aproximações de conjuntos podem ser visualizadas na Figura 2.1, as quais temos que  $\underline{X}$  refere-se a região positiva de  $X$ ;  $\overline{X}$  refere-se à região negativa de  $X$ ;  $\overline{\overline{X}}$  refere-se à região positiva-fronteira de  $X$ ;  $\underline{\underline{X}}$  refere-se à região negativa-fronteira de  $X$ ;  $\overline{\underline{X}}$  refere-se à região negativa-fronteira de  $X$ . Todo conjunto que pode ser definido precisamente, i.e., a sua região da fronteira é vazia, é chamado de conjunto *crisp* (preciso, exato). Caso contrário, ele é chamado de *rough* (vago, impreciso, inexato). Um conjunto que não pode ser precisamente definido pode então ser representado por um *rough set*.

**Definição 5 (Rough Set)** *Sejam  $D = (U, A, d, V_A, V_d)$  uma tabela de decisão,  $B \subseteq A$  e  $X \subseteq U$ . Um rough set  $X$  é definido como um par  $X = (\underline{X}_B, \overline{X}^B)$ .*

*Exemplo 6.* Considere novamente o *Exemplo 5*. Como o conjunto Lentes não pode ser precisamente definido em termos de classes de equivalência, ele é então representado pelo *rough set*  $Lentes = (\underline{Lentes}, \overline{Lentes})$ , em que  $\underline{Lentes} = \{x_7, x_8, x_{12}\}$  e  $\overline{Lentes} = \{x_3, x_7, x_8, x_9, x_{11}, x_{12}, x_{13}, x_{14}\}$ .

## 2.4 Redução de Atributos

Na seção anterior, apresentamos um dos aspectos de redução de informação, que leva em consideração identificar classes de equivalência. Com objetivo de eliminar objetos redundantes, apenas um elemento da classe de equivalência é necessário para representar a classe inteira. O outro aspecto de redução de dados é manter apenas aqueles atributos que preservam a relação de indiscernibilidade, e conseqüentemente a aproximação dos conjuntos. Os atributos rejeitados são redundantes, pois sua remoção não pode piorar a classificação dos objetos. Para expressar mais precisamente essas ideias, introduziremos algumas novas noções.

**Definição 6 (Reduto)** *Sejam  $S = (U, A, V_A)$  um sistema de informação,  $B \subseteq A$  e  $a \in B$ . Dizemos que  $a$  é dispensável em  $B$  se  $IND(B) = IND(B - \{a\})$ . Caso contrário,  $a$  é dito indispensável em  $B$ . Um conjunto  $B$  é chamado independente se todos seus atributos são indispensáveis. Um subconjunto  $B'$  de  $B$  é chamado de reduto de  $B$  se  $B'$  é independente e  $IND(B') = IND(B)$ .*

Um reduto é um subconjunto de atributos que preserva partições (indiscernibilidade). Isso significa que um reduto é um subconjunto minimal de atributos que permite a mesma classificação de elementos do universo que o conjunto completo de atributos. Em outras palavras, atributos que não pertencem a um reduto são supérfluos com relação à classificação de elementos do universo.

Computar redutos não é uma tarefa trivial. Para um sistema de informação com  $n$  atributos presentes, a quantidade de subconjuntos possíveis é  $2^n$ . A computação das classes de equivalências é feita diretamente, mas encontrar um reduto minimal entre o conjunto de atributos de um sistema de informação é um problema NP-difícil (SKOWRON; RAUSZER, 1992). A fim de computar os redutos, podemos usar a noção de matriz de discernibilidade (SKOWRON; RAUSZER, 1992):

**Definição 7 (Matriz de Discernibilidade)** *Seja  $S = (U, A, V_A)$  um sistema de informação com  $n$  objetos. A matriz de discernibilidade de  $S$  é uma matriz simétrica  $n \times n$  com entradas  $S_{ij}$ , em que*

$$S_{ij} = \{a \in A \mid a(x_i) \neq a(x_j)\}, \text{ para } i, j = 1, \dots, n.$$

Cada entrada consiste do conjunto de atributos em que os objetos  $x_i$  e  $x_j$  diferem. Como a matriz de discernibilidade é simétrica e  $S_{ii} = \emptyset$  para  $i = 1, \dots, n$ , essa matriz pode ser representada usando apenas os elementos na sua parte triangular inferior. Para cada matriz de discernibilidade podemos associar uma função booleana  $f$ , chamada de função de discernibilidade:

**Definição 8 (Função de Discernibilidade)** *A função de discernibilidade de um sistema de informação  $S = (U, A, V_A)$  com  $n$  objetos é definida por:*

$$f(S) = \bigwedge (\bigvee S_{ij}), \text{ para todo } i, j = 1, \dots, n, S_{ij} \neq \emptyset.$$

A expressão  $(\bigvee S_{ij})$  é a disjunção de todos os atributos da matriz de discernibilidade  $S_{ij}$ , indicando que um par de objetos  $(x, y)$  pode ser distinguido por algum atributo em  $S_{ij}$ . A expressão  $\bigwedge (\bigvee S_{ij})$  é a conjunção de todos  $(\bigvee S_{ij})$ , indicando que a família de pares de objetos discerníveis podem ser distinguidos por um conjunto de atributos satisfazendo  $\bigwedge (\bigvee S_{ij})$ .

*Exemplo 7.* Usando as definições acima, obtemos a matriz de discernibilidade (Tabela 2.4) do sistema de informação referente à Tabela 2.3, em que  $A, I, P, O$  correspondem aos atributos Ast, Idade, ProdL, Óculos, respectivamente. A função de discernibilidade para essa tabela é

$$\begin{aligned} f(S) = & (A \vee I) \wedge (A \vee P \vee O) \wedge (I \vee P \vee O) \wedge A \wedge I \wedge (P \vee O) \wedge \\ & (I \vee O) \wedge (A \vee I \vee O) \wedge (A \vee I \vee P) \wedge (A \vee I \vee O) \wedge (P \vee O) \wedge \\ & (A \vee I \vee P \vee O) \wedge A \wedge (A \vee P \vee O) \wedge (I \vee P) \wedge (A \vee I \vee P \vee O) \wedge \\ & (I \vee P \vee O) \wedge I \wedge (I \vee P \vee O) \wedge (A \vee P) \wedge (A \vee I). \end{aligned}$$

A função de discernibilidade pode ser usada para estabelecer um resultado importante a respeito do conjunto de redutos de um sistema de informação, como mostra o seguinte teorema de Skowron e Rauszer:

**Teorema 1** (SKOWRON; RAUSZER, 1992) *O problema de encontrar um reduto de um conjunto é equivalente ao problema de transformar a função de discernibilidade para sua forma normal disjuntiva simplificada. Cada conjunção da forma normal disjuntiva simplificada é chamada de um implicante primo. Sejam  $S = (U, A, V_A)$  e  $B \subseteq A$ . Um conjunto de atributos  $B = \{a_1, \dots, a_m\}$  é um reduto se e somente se a conjunção de todos os atributos em  $B$ , denotada por  $a_1 \wedge \dots \wedge a_m$ , é um primo implicante de  $f(S)$ .*

O conjunto de todos implicantes primos de  $f$  determina o conjunto de todos os redutos de  $A$ . O processo de descoberta de implicantes primos pode ser feito através da simplificação da fórmula booleana encontrada por  $f(S)$ .

O processo de simplificação consiste em transformar a fórmula  $f(S)$  em forma normal conjuntiva, para uma fórmula normal disjuntiva reduzida. Esse procedimento pode ser feito através das leis de absorção (i.e.,  $a \wedge (a \vee b) \equiv a$  e  $a \vee (a \wedge b) \equiv a$ ) e distributividade (i.e.,  $a \wedge (b \vee c) \equiv (a \wedge b) \vee (a \wedge c)$  e  $a \vee (b \wedge c) \equiv (a \vee b) \wedge (a \vee c)$ ) presentes na lógica proposicional, e é realizado em tempo exponencial (BOROS; ELBASSIONI; MAKINO, 2008; MILTERSEN; RADHAKRISHNAN; WEGENER, 2005). Simplificando  $f(S)$  temos que

$$\begin{aligned} f(S) &= A \wedge I \wedge (P \vee O) \\ &= (A \wedge I \wedge P) \vee (A \wedge I \wedge O). \end{aligned}$$

	$S_1$	$S_2$	$S_3$	$S_4$	$S_5$	$S_6$	$S_7$
$S_1$	$\emptyset$						
$S_2$	$A, I$	$\emptyset$					
$S_3$	$A, P, O$	$I, P, O$	$\emptyset$				
$S_4$	$A$	$I$	$P, O$	$\emptyset$			
$S_5$	$I, O$	$A, I, O$	$A, I, P$	$A, I, O$	$\emptyset$		
$S_6$	$P, O$	$A, I, O, P$	$A$	$A, P, O$	$I, P$	$\emptyset$	
$S_7$	$A, I, P, O$	$I, P, O$	$I$	$I, P, O$	$A, P$	$A, I$	$\emptyset$

Tabela 2.4: Matriz de discernibilidade do sistema de informação da Tabela 2.3

Nesse caso, os implicantes primos  $(A \wedge I \wedge P)$  ou  $(A \wedge I \wedge O)$  são os redutos da Tabela 2.4. Em outras palavras apenas com os atributos  $Ast, Idade, ProdL$  ou  $Ast, Idade, \acute{U}culos$ , ja é possível distinguir o conjunto de todos os elementos do sistema de informação da Tabela 2.3.

## 2.5 Medidas de Qualidade

Vimos que utilizando *rough sets* podemos reduzir as informações de uma tabela através de seus objetos ou de seus atributos. A redução de atributos é obtida com a ideia de reduto, que pode ser calculado através da função de discernibilidade. A função de discernibilidade pode apresentar como resultado uma série de conjuntos minimais, que serão os redutos. Todos os redutos possuem a característica essencial de preservar a indiscernibilidade do conjunto original de atributos, mas eles podem apresentar outras características que os diferenciam entre si. Para conseguir realizar essa distinção entre os redutos obtidos e conseqüentemente buscar um reduto mais representativo, são aplicadas medidas de qualidade de aproximações de conjuntos.

Na literatura, Pawlak sugere duas medidas numéricas para caracterizar a imprecisão das aproximações em *rough sets* (PAWLAK, 1982), com o objetivo de estimar a qualidade das aproximações dos conjuntos obtidos, i.e. a quantidade de incerteza que um *rough set* apresenta:

**Definição 9 (Medida de Precisão)** *Sejam  $X \subseteq U$ ,  $X \neq \emptyset$  e  $B \subseteq A$ . A medida de precisão  $\alpha_B(X)$  é definida por:*

$$\alpha_B(X) = \frac{|X_B|}{|\bar{X}^B|},$$

em que  $|X|$  denota a cardinalidade de  $X$ .

A precisão  $\alpha_B(X)$  é a razão do número de objetos que necessariamente pertencem a  $X$  com o número de objetos que possivelmente pertencem a  $X$ . Isso mostra uma medida de quão próximo o *rough set* está do conjunto  $X$ . Obviamente  $0 \leq \alpha_B(X) \leq 1$ . Se  $\alpha_B(X) = 1$  então  $X$  é *crisp* (ou preciso) com respeito a  $B$ , e caso contrário, se  $\alpha_B(X) < 1$ ,  $X$  é *rough* (ou impreciso) com respeito a  $B$ . Baseado na medida de precisão, a medida de imprecisão é definida por

$$\rho_B(X) = 1 - \alpha_B(X).$$

A medida de imprecisão pode ser vista como o complemento da medida de precisão e está relacionada com a noção da região de fronteira. Outra forma de defini-la é

$$\rho_B(X) = \frac{|\overline{X}^B - \underline{X}_B|}{|\overline{X}^B|}.$$

*Exemplo 8.* A medida de precisão do conjunto Lentes com relação ao conjunto de atributos  $A$  é dada por  $\alpha_A(\text{Lentes}) = \frac{|\underline{\text{Lentes}}|}{|\text{Lentes}|} = \frac{3}{8} = 0.375$ . Enquanto que a medida de imprecisão de Lentes com relação a  $A$  é dada por  $\rho_A(\text{Lentes}) = 1 - \alpha_A(\text{Lentes}) = 1 - 0.375 = 0.625$ .

Outro exemplo de uma medida numérica presente na teoria dos *rough sets* é a da dependência de atributos (YAO, 2010).

**Definição 10 (Medida de Dependência)** *Sejam  $B, D \subseteq A$ . A medida de dependência  $\gamma_B(D)$  denota a dependência de  $D$  com relação a  $B$  e é definida por:*

$$\gamma_B(D) = \frac{|POS_B(D)|}{|U|},$$

$$\text{em que } POS_B(D) = \bigcup_{X \in U/D} \underline{X}_B.$$

A função  $POS_B(D)$ , chamada da região positiva da partição  $U/D$  com respeito a  $B$ , é o conjunto de todos elementos de  $U$  que podem ser unicamente classificados em blocos de partição  $U/D$  por meio de  $B$ . Intuitivamente, um conjunto de atributos  $D$  depende de um conjunto de atributos  $B$ , se todos os valores de atributos de  $D$  são unicamente determinados pelos valores de atributos de  $B$ . Alternativamente, podemos definir a medida de dependência como

$$\gamma_B(D) = \sum_{X \in U/D} \frac{|\underline{X}_B|}{|U|}.$$

A dependência  $\gamma_B(D)$  pode ser interpretada como a proporção de objetos que determinam a partir dos valores dos atributos em  $B$ , os valores de atributos em  $D$ . Em outras palavras, o coeficiente  $\gamma_B(D)$  expressa a razão de todos os elementos do universo que podem ser corretamente classificados nas partições de  $U/D$  por meio de  $B$ . Se  $\gamma_B(D) = 1$ , dizemos que  $D$  depende totalmente de  $B$ . Se  $0 < \gamma_B(D) < 1$ , dizemos que  $D$  depende parcialmente de  $B$ . E quando  $\gamma_B(D) = 0$ , dizemos que  $D$  não depende de  $B$ .

*Exemplo 9.* O grau de dependência do conjunto Lentes com relação ao conjunto de atributos  $A$  é dado por  $\gamma_A(\text{Lentes}) = \frac{|POS_A(\text{Lentes})|}{|U|} = \frac{|\underline{\text{Lentes}} \cup \overline{\text{Lentes}}|}{|U|} = \frac{9}{14} = 0.642$ .

Esses são alguns conceitos fundamentais da teoria dos *rough sets* que envolvem tarefas como representação de informações contraditórias ou vagas através de aproximações de

conjuntos, redução dos dados representados nas tabelas a partir de classificações de objetos ou atributos, e medidas de qualidades nas classificações. É válido ressaltar que na ideia original dos *rough sets*, todas as informações são tidas como completas (i.e., todos os atributos são especificados); contudo existem extensões da teoria dos *rough sets* em que as informações fornecidas podem ser incompletas. Nesta dissertação, as informações incompletas são fundamentais para refinamento de consultas. Em Lógicas de Descrição, por exemplo, a informação incompleta pode ser expressa de diversas formas (CALVANESE et al., 2011). Ao utilizarmos essas lógicas como base para a representação de informações, será necessário definir extensões para a teoria dos *rough sets* que permitam lidar com aproximações. Nos próximos capítulos, serão abordadas com mais detalhes essas extensões. Primeiramente, começaremos com uma introdução às Lógicas de Descrição.

### 3 LÓGICAS DE DESCRIÇÃO

Este capítulo apresenta uma breve introdução sobre Lógicas de Descrição (DLs) (BAADER, 2003). É sabido que essas lógicas são uma família de linguagens de representação de conhecimento baseadas nas teorias de redes semânticas (SOWA, 1992) e frames (MINSKY, 1981), que descrevem o conhecimento de um domínio de uma maneira estruturada e de fácil entendimento. Sistemas escritos em DLs provêm serviços de inferência que deduzem conhecimento implícito a partir do conhecimento explicitamente representado. Eles podem ser empregados em vários domínios de aplicação tais como Web Semântica (HORROCKS, 2002), Ontologias (BAADER; HORROCKS; SATTLER, 2004), Banco de Dados (BORGIDA; LENZINI; ROSATI, 2003) e Engenharia de Software (BERARDI; CALVANESE; GIACOMO, 2005).

As DLs formalizam terminologias de um domínio de aplicação ao definir noções importantes como objetos, classes de objetos e relacionamento entre objetos. É possível também deduzir conhecimento a partir dessas definições como relacionamentos de subclasses ou relacionamentos entre instâncias. Além disso, essas lógicas são mais expressivas que a lógica proposicional e apresentam resultados de decisão mais eficientes que lógica de primeira ordem (BAADER, 2003). Em geral, possuem uma boa relação entre expressividade e complexidade.

A expressividade de uma Lógica de Descrição é medida pela família de linguagens de descrição que pode representar suas fórmulas. Assim, a família com menor expressividade é chamada  $\mathcal{AL}$  (Attributive Language) (BAADER, 2003), e suporta descrições que usam conceitos atômicos, negação atômica, interseção de conceitos, quantificação universal e quantificação existencial limitada. Essa família pode ser estendida com a adição de outros construtores como a união de conceitos, quantificador existencial completo, restrições numéricas, ou negação de conceitos (e não apenas conceitos atômicos), ou qualquer combinação de  $\mathcal{AL}$  com um conjunto formado pelas extensões descritas, dependendo do tipo de conhecimento que se pretende representar.

Com relação ao nosso trabalho, o processo de refinamento de consultas será alicerçado a partir da extensão de Lógicas de Descrição com os operadores de *rough sets* (Rough DLs) e consiste de aproximar axiomas com as aproximações inferiores e superiores. Na Seção 3.1 apresentaremos uma breve introdução sobre Lógicas de Descrição. Na Seção 3.2 apresentaremos as Lógicas de Descrição *Rough*.

#### 3.1 Lógica de Descrição $\mathcal{AL}$

##### 3.1.1 Sintaxe

A  $\mathcal{AL}$  assume que existem três tipos de informação distintas a serem representadas. São elas os indivíduos, os conceitos e os papéis.

**Definição 11 (Alfabeto)** *O alfabeto em  $\mathcal{AL}$  é composto dos construtores  $\neg, \sqcap, \exists, \forall, \top, \perp$  e três*

conjuntos disjuntos entre si: o conjunto de indivíduos  $N_I$ , o conjunto de nomes de conceitos  $N_C$  e o conjunto de nomes de papéis  $N_R$ .

Um indivíduo representa um elemento de um universo de discurso (domínio). Nomes de conceitos representam predicados unários relacionados com indivíduos. Por exemplo, sejam *joão* um indivíduo e *Pessoa* um nome de conceito; para representar que *joão* é uma instância de *Pessoa*, escrevemos *Pessoa(joão)*. Nomes de papéis representam predicados binários entre indivíduos. Sejam *josé* e *maria* indivíduos e *temFilho* um nome de papel; para descrever que o indivíduo *josé* está relacionado com o indivíduo *maria* por *temFilho*, escrevemos *temFilho(josé,maria)*. Descrições elementares como *Pessoa(joão)* e *temFilho(josé,maria)* são chamadas, respectivamente, de *conceitos atômicos* e *papéis atômicos*. Descrições gerais podem ser indutivamente geradas a partir de construtores de conceitos.

**Definição 12 (Conceitos)** *Conceitos em  $\mathcal{AL}$  são definidos pela regras de sintaxe abaixo, em que  $C$  e  $D$  são conceitos,  $A$  é um conceito atômico e  $R$  é um papel atômico:*

$$\begin{aligned}
C, D \longrightarrow & A \mid (\text{conceito atômico}) \\
& \top \mid (\text{conceito top}) \\
& \perp \mid (\text{conceito bottom}) \\
& \neg A \mid (\text{negação atômica}) \\
& C \sqcap D \mid (\text{interseção}) \\
& \exists R. \top \mid (\text{quantificação existencial não qualificada}) \\
& \forall R. C \mid (\text{restrição de valor}).
\end{aligned}$$

Um conceito atômico  $A$  representa uma propriedade associada a um conjunto de indivíduos. Os conceitos *top* e *bottom* correspondem ao conjunto de todos os indivíduos e ao conjunto composto por nenhum indivíduo, respectivamente. A negação  $\neg A$  corresponde ao complemento de  $A$ , ou seja, todos os indivíduos do domínio que não têm a propriedade  $A$ .  $C \sqcap D$  representa todos indivíduos que possuem as propriedades  $C$  e  $D$ .  $\exists R. \top$  representa todos os indivíduos que se relacionam com pelo menos um indivíduo pelo papel  $R$ . Já  $\forall R. C$  representa os indivíduos que se relacionam apenas com indivíduos que possuem a propriedade  $C$  através do papel  $R$ .

Para dar alguns exemplos do que pode ser expresso em  $\mathcal{AL}$ , suponha que *Pessoa* e *Feminino* são conceitos atômicos. Então *Pessoa*  $\sqcap$  *Feminino* e *Pessoa*  $\sqcap$   $\neg$ *Feminino* são conceitos descrevendo, intuitivamente, mulheres e homens, respectivamente. Além disso, suponha que *temFilho* é um papel atômico. Podemos formar os conceitos *Pessoa*  $\sqcap$   $\exists$ *temFilho*. $\top$  e *Pessoa*  $\sqcap$   $\forall$ *temFilho*.*Feminino*, denotando respectivamente aquelas pessoas que tem pelo menos um filho, e aquelas pessoas em que todos seus filhos são do sexo feminino. Usando o conceito *bottom*, podemos descrever aquelas pessoas sem filhos pelo conceito *Pessoa*  $\sqcap$   $\forall$ *temFilho*. $\perp$ .

A representação de conhecimento em  $\mathcal{AL}$  pode ser dividida em dois componentes: TBox e ABox. TBox introduz a terminologia, i.e., o vocabulário de um alfabeto, enquanto

o ABox contém asserções sobre indivíduos em termos desse vocabulário. O TBox pode ser usado para assinalar nomes para descrições de conceitos complexos. Sintaticamente, ele é representado por um conjunto finito de *axiomas terminológicos* com a forma

$$C \sqsubseteq D \text{ ou } C \equiv D,$$

em que  $C, D$  são conceitos. Axiomas do primeiro tipo são chamados inclusões, enquanto axiomas do segundo tipo são chamados igualdades. O axioma de inclusão  $C \sqsubseteq D$  significa que cada indivíduo de  $C$  é um indivíduo de  $D$ . Por exemplo, podemos definir  $Mãe \sqsubseteq Mulher$ . Enquanto que o axioma de igualdade  $C \equiv D$  significa que cada indivíduo de  $C$  é um indivíduo de  $D$  e cada indivíduo de  $D$  é um indivíduo de  $C$ . Por exemplo, o axioma  $Mãe \equiv Mulher \sqcap \exists temFilho$ .  $\top$  define o conceito  $Mãe$  como sendo exatamente qualquer mulher que possui pelo menos um filho. O ABox consiste de um conjunto finito de axiomas de asserções que possuem a forma

$$C(a) \text{ ou } R(a, b),$$

em que  $a, b$  são indivíduos,  $C$  é um conceito e  $R$  é um papel atômico. O axioma do primeiro tipo, chamado de asserção de conceito, denota que o indivíduo  $a$  pertence ao conceito  $C$ . Já o axioma do segundo tipo, chamado asserção de papel, denota que o indivíduo  $a$  está relacionado com o indivíduo  $b$  por  $R$ . Por exemplo, sejam *pedro*, *paulo* e *maria* indivíduos; então  $Pai(pedro)$  significa que Pedro é um pai e  $temFilho(maria, paulo)$  significa que Paulo é um filho de Maria.

**Definição 13 (Base de conhecimento)** *Uma base de conhecimento ou ontologia  $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$  em  $\mathcal{AL}$  consiste de um TBox  $\mathcal{T}$  e um ABox  $\mathcal{A}$ .*

### 3.1.2 Semântica

De um ponto de vista semântico, indivíduos são interpretados como elementos de um domínio, conceitos são interpretados como subconjuntos desse domínio, enquanto papéis são interpretados como relações binárias sobre tal domínio. Mais precisamente, uma interpretação  $I = (\Delta^I, \cdot^I)$  consiste de um domínio de interpretação  $\Delta^I$  e uma função de interpretação  $\cdot^I$  mapeando todo conceito atômico  $A$  em um subconjunto de  $\Delta^I$  e todo papel atômico  $R$  para um subconjunto de  $\Delta^I \times \Delta^I$ . A função de interpretação  $\cdot^I$  é estendida para conceitos complexos de  $\mathcal{AL}$  como segue:

- $a^I \in \Delta^I$ ;
- $A^I \subseteq \Delta^I$ ;
- $R^I \subseteq \Delta^I \times \Delta^I$ ;
- $\top^I = \Delta^I$ ;
- $\perp^I = \emptyset$ ;

- $(\neg A)^I = \Delta^I \setminus A^I$ ;
- $(C \sqcap D)^I = C^I \cap D^I$ ;
- $(\exists R. \top)^I = \{a \in \Delta^I \mid \exists b \in \Delta^I, (a, b) \in R^I\}$ ;
- $(\forall R. C)^I = \{a \in \Delta^I \mid \forall b \in \Delta^I, (a, b) \in R^I \rightarrow b \in C^I\}$ .

**Definição 14 (Satisfatibilidade)** A noção de satisfatibilidade de um axioma  $\alpha$  por uma interpretação  $I = (\Delta^I, \cdot^I)$ , denotada por  $I \models \alpha$ , é definida a seguir:

- $I \models C \sqsubseteq D$  se e somente se  $C^I \subseteq D^I$ ;
- $I \models C \equiv D$  se e somente se  $C^I = D^I$ ;
- $I \models C(a)$  se e somente se  $a^I \in C^I$ ;
- $I \models R(a, b)$  se e somente se  $(a^I, b^I) \in R^I$ .

Se  $I \models \alpha$  dizemos que  $I$  é um modelo de  $\alpha$ . Um TBox  $\mathcal{T}$  é satisfeito por uma interpretação  $I$ , ou seja,  $I \models \mathcal{T}$  se e somente se para todo  $C \sqsubseteq D$  e  $C \equiv D$  em  $\mathcal{T}$ ,  $I \models C \sqsubseteq D$  e  $I \models C \equiv D$ . Um ABox  $\mathcal{A}$  é satisfeito por uma interpretação  $I$ , ou seja,  $I \models \mathcal{A}$  se e somente se para toda asserção da forma  $C(a)$  e  $R(a, b)$  em  $\mathcal{A}$ ,  $I \models C(a)$  e  $I \models R(a, b)$ .  $I$  satisfaz (é um modelo de) uma base de conhecimento  $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$ , denotado por  $I \models \mathcal{K}$ , se e somente se  $I \models \mathcal{T}$  e  $I \models \mathcal{A}$ .

### 3.1.3 A Família de Linguagens de $\mathcal{AL}$

Nós obtemos mais linguagens expressivas se nós adicionarmos outros construtores para  $\mathcal{AL}$ . Acrescentando a união de conceitos  $C \sqcup D$ , temos a DL  $\mathcal{ALU}$ . O conceito  $C \sqcup D$  é interpretado como

$$(C \sqcup D)^I = C^I \cup D^I$$

e representa todos os indivíduos que pertencem ao conceito  $C$  ou ao conceito  $D$ . Adicionando a quantificação existencial completa  $\exists R.C$  a  $\mathcal{AL}$ , criamos a DL  $\mathcal{ALE}$ . O conceito  $\exists R.C$  representa todos os indivíduos que se relacionam via papel  $R$  com pelo menos um indivíduo em  $C$ . A quantificação existencial completa é interpretada como

$$(\exists R.C)^I = \{a \in \Delta^I \mid \exists b \in \Delta^I, (a, b) \in R^I \wedge b \in C^I\}.$$

Note que  $\exists R.C$  difere de  $\exists R. \top$  no fato que conceitos arbitrários são permitidos ocorrer no escopo de quantificador existencial. Restrições numéricas podem ser adicionadas a  $\mathcal{AL}$ , gerando a DL  $\mathcal{ALN}$ . Os conceitos que representam as restrições numéricas são  $\geq nR$  (restrição de no mínimo) e  $\leq nR$  (restrição de no máximo). Elas são interpretadas como

$$(\geq nR)^I = \{a \in \Delta^I \mid |\{b \mid (a,b) \in R^I\}| \geq n\} \text{ e } (\leq nR)^I = \{a \in \Delta^I \mid |\{b \mid (a,b) \in R^I\}| \leq n\},$$

em que  $|\cdot|$  denota a cardinalidade de um conjunto. Com esses conceitos, podemos falar, por exemplo, dos indivíduos que possuem no máximo três filhos, através de  $\leq 3\text{temFilho}$ , ou de indivíduos que possuem exatamente um filho por  $(\leq 1\text{temFilho} \sqcap \geq 1\text{temFilho})$ .

Por fim, ao acrescentar à  $\mathcal{AL}$  a negação de conceitos arbitrários  $\neg C$ , temos a DL  $\mathcal{ALC}$ . Esse conceito é interpretado como

$$(\neg C)^I = \Delta^I \setminus C^I.$$

Estendendo a  $\mathcal{AL}$  com qualquer subconjunto dos construtores acima produz uma linguagem  $\mathcal{AL}$  particular. Nós nomeamos cada linguagem  $\mathcal{AL}$  por uma string da forma

$$\mathcal{AL}[\mathcal{U}][\mathcal{E}][\mathcal{N}][C],$$

em que uma letra no nome significa a presença do construtor correspondente. Por exemplo,  $\mathcal{ALEN}$  é a extensão de  $\mathcal{AL}$  com a quantificação existencial completa e restrições numéricas. Uma exceção a regra é a DL  $\mathcal{ALC}$ , que na literatura representa na verdade a DL  $\mathcal{ALUC}$ , ou seja, ao falarmos de  $\mathcal{ALC}$ , estamos-nos referindo à extensão de  $\mathcal{AL}$  com união de conceitos e negação arbitrária. A DL  $\mathcal{ALC}$  (Attributive Concept Language with Complements) foi introduzida por Manfred Schmidt-Schauß e Gert Smolka em 1991 (SCHMIDT-SCHAUSS; SMOLKA, 1991) e é tratada como a base de muitas DLs mais expressivas. Além disso, ela é a “menor” DL que é proposicionalmente fechada aos operadores booleanos. Mais precisamente, a  $\mathcal{ALC}$  contém os construtores: negação ( $\neg$ ), conjunção ( $\sqcap$ ), disjunção ( $\sqcup$ ), restrição existencial ( $\exists$ ) e restrição universal ( $\forall$ ).

É claro que esses não são os únicos construtores de conceitos que podem ser utilizados para estender uma Lógica de Descrição. Ademais, as extensões não se limitam apenas aos conceitos; é possível também criar construtores para papéis (BAADER, 2003). Porém a adição de novos construtores a uma linguagem de descrição pode influenciar na sua complexidade de decisão. É isso que analisaremos na sequência.

### 3.1.4 Inferência

Um sistema em DL não apenas armazena terminologias e asserções, mas também oferece serviços de raciocínio sobre eles. Determinar quando uma descrição de um axioma terminológico é satisfatível, ou quando uma descrição é mais geral que outra são típicas tarefas de raciocínio para um TBox. Um problema importante para um ABox é descobrir quando seus conjuntos de asserções são consistentes (i.e., quando eles têm um modelo). Tanto a checagem de satisfatibilidade de axiomas terminológicos como a checagem de consistência de axiomas de asserções são úteis par determinar quando uma base de conhecimento é significativa de fato.

Os diferentes tipos de raciocínio realizados por uma DL são definidos como inferências lógicas. Na sequência, apresentaremos essas inferências para TBox e ABox. Averiguar a satisfatibilidade de conceitos é visto como um processo chave de inferência, pois como mostraremos a seguir, os outros tipos de inferência podem ser reduzidos à (in)satisfatibilidade. Começaremos com as tarefas de raciocínio para um TBox:

- **Satisfatibilidade:** Um conceito  $C$  é satisfável com respeito ao TBox  $\mathcal{T}$  se existe um modelo  $I$  de  $\mathcal{T}$  tal que  $C^I$  não é vazio;
- **Subsunção:** Um conceito  $C$  é subsumido por um conceito  $D$  com respeito a  $\mathcal{T}$  se  $C^I \subseteq D^I$  para todo modelo  $I$  de  $\mathcal{T}$ . Nesse caso, escrevemos  $\mathcal{T} \models C \sqsubseteq D$ ;
- **Equivalência:** Dois conceitos  $C$  e  $D$  são equivalentes com respeito a  $\mathcal{T}$  se  $C^I = D^I$  para todo modelo  $I$  de  $\mathcal{T}$ . Nesse caso, escrevemos  $\mathcal{T} \models C \equiv D$ ;
- **Disjunção:** Dois conceitos  $C$  e  $D$  são disjuntos com respeito a  $\mathcal{T}$  se  $C^I \cap D^I = \emptyset$  para todo modelo  $I$  de  $\mathcal{T}$ .

*Exemplo 10.* De acordo com o TBox abaixo, *Pessoa* subsume *Mulher*, *Mulher* e *Genitor* subsumem *Mãe*, e *Mãe* subsume *Avó*. Além disso, *Mulher* e *Homem*, e *Pai* e *Mãe* são disjuntos.

$$\begin{aligned}
 \textit{Mulher} &\equiv \textit{Pessoa} \sqcap \textit{Feminino} \\
 \textit{Homem} &\equiv \textit{Pessoa} \sqcap \neg \textit{Feminino} \\
 \textit{Mãe} &\equiv \textit{Mulher} \sqcap \exists \textit{temFilho.Pessoa} \\
 \textit{Pai} &\equiv \textit{Homem} \sqcap \exists \textit{temFilho.Pessoa} \\
 \textit{Genitor} &\equiv \textit{Pai} \sqcup \textit{Mãe} \\
 \textit{Avó} &\equiv \textit{Mãe} \sqcap \exists \textit{temFilho.Genitor} \\
 \textit{MãeSemFilha} &\equiv \textit{Mãe} \sqcap \forall \textit{temFilho.}\neg \textit{Mulher} \\
 \textit{Esposa} &\equiv \textit{Mulher} \sqcap \exists \textit{temMarido.Homem}
 \end{aligned}$$

**Proposição 1** (BAADER, 2003) *Considere os conceitos  $C$  e  $D$ . As tarefas de inferência abaixo podem ser reduzidas à insatisfatibilidade:*

- $C$  é subsumido por  $D \Leftrightarrow C \sqcap \neg D$  é insatisfável;
- $C$  e  $D$  são equivalentes  $\Leftrightarrow C \sqcap \neg D$  e  $\neg C \sqcap D$  são insatisfáveis;
- $C$  e  $D$  são disjuntos  $\Leftrightarrow C \sqcap D$  é insatisfável.

Em suma, independente do tipo de inferência utilizada, ela poderá ser reduzida à insatisfatibilidade, e conseqüentemente à satisfatibilidade. Para o ABox podemos citar as seguintes tarefas de raciocínio:

- **Checagem de Consistência de um ABox com respeito a um TBox:** Um ABox  $\mathcal{A}$  é consistente com respeito a um TBox  $\mathcal{T}$  se existe uma interpretação  $I$  que é modelo para  $\mathcal{T}$  e para  $\mathcal{A}$ ;
- **Checagem de Instância:** Checa quando uma asserção  $C(a)$  ou  $R(a,b)$  é acarretada por um ABox  $\mathcal{A}$ , i.e., toda interpretação  $I$  que é modelo de  $\mathcal{A}$  também é um modelo de  $C(a)$  ou  $R(a,b)$ . Nesse caso, escrevemos  $\mathcal{A} \models C(a)$  ou  $\mathcal{A} \models R(a,b)$ .

Se um axioma de asserção é da forma  $R(a,b)$ , a checagem de instância é fácil, pois a linguagem de descrição não contém construtores para formar papéis complexos. Se o axioma de asserção é da forma  $C(a)$ , podemos reduzir checagem de instância para o problema de consistência para ABox porque existe a seguinte conexão:

**Proposição 2** (BAADER, 2003) *Sejam  $\mathcal{A}$  um ABox,  $a$  um indivíduo e  $C$  um conceito. A checagem de instância pode ser reduzida à checagem de consistência:*

- $\mathcal{A} \models C(a) \Leftrightarrow \mathcal{A} \cup \{\neg C(a)\}$  é inconsistente.

*Exemplo 11.* Seja  $\mathcal{A}$  o ABox abaixo.

$M\tilde{a}eSemFilhas(maria)$

$Pai(pedro)$

$temFilho(maria, pedro)$

$temFilho(maria, paulo)$

$temFilho(pedro, marcos)$

Fazendo a interpretação  $I = (\Delta^I, \cdot^I)$ , em que  $\Delta^I = \{maria, pedro, paulo, marcos\}$ ,  $M\tilde{a}eSemFilhas^I = \{maria\}$ ,  $Pai^I = \{pedro\}$ ,  $temFilho^I = \{(maria, pedro), (maria, paulo), (pedro, marcos)\}$ ,  $maria^I = maria$ ,  $pedro^I = pedro$ ,  $paulo^I = paulo$  e  $marcos^I = marcos$ , temos que  $\mathcal{A}$  é consistente, pois  $I$  é um modelo para cada asserção em  $\mathcal{A}$ . Se considerarmos a asserção  $Pai(pedro)$ , temos que  $\mathcal{A} \models Pai(pedro)$ . Se por acaso, quisermos checar a asserção  $Pai(paulo)$ , temos que  $\mathcal{A} \not\models Pai(paulo)$ , pois existe uma interpretação que satisfaz  $\mathcal{A}$ , mas não satisfaz  $Pai(paulo)$  (considere a interpretação como sendo verdadeiras todas as asserções abaixo e a asserção  $Pai(paulo)$  sendo falsa).

Em relação às questões de complexidades relacionadas com a checagem de satisfatibilidade de um TBox e checagem de consistência de um ABox, temos os seguintes resultados para a  $\mathcal{ALC}$ , quer será a lógica explorada na dissertação.

**Teorema 2** (SCHILD, 1991) *A checagem de satisfatibilidade de conceitos de um TBox em  $\mathcal{ALC}$  com apenas axiomas do tipo  $C \equiv D$  tem complexidade PSPACE-completo. A checagem de satisfatibilidade de conceitos de um TBox em  $\mathcal{ALC}$  com axiomas do tipo  $C \equiv D$  e  $C \sqsubseteq D$  tem complexidade EXPTIME-completo.*

Por outro lado, quando tratamos de uma base de conhecimento com um TBox vazio e um ABox não-vazio, temos que

**Teorema 3** (BAADER et al., 2005) *A checagem de consistência de um ABox em  $\mathcal{ALC}$  tem complexidade PSPACE-completo.*

As tarefas de inferências para TBox e ABox possuem os mesmos resultados de complexidade quando os dois são analisados separadamente. Se considerarmos agora uma base de conhecimento composta de TBox e ABox não-vazios, teremos que

**Teorema 4** (TOBIES, 2001) *A checagem de consistência de um ABox e TBox em  $\mathcal{ALC}$  com apenas axiomas do tipo  $C \equiv D$  tem complexidade PSPACE-completo. A checagem de consistência de um ABox e TBox em  $\mathcal{ALC}$  com axiomas do tipo  $C \equiv D$  e  $C \sqsubseteq D$  tem complexidade EXPTIME-completo.*

### 3.2 Lógica de Descrição Rough $\mathcal{ALC}$

Para finalizar com as apresentações de noções preliminares de DLs, nesta seção abordaremos as extensões de DLs com a teoria dos *rough sets*, chamadas de Lógicas de Descrição Rough (RDLs). As RDLs (FANIZZI et al., 2008; JIANG et al., 2009; KEET, 2010; SCHLOBACH; KLEIN, 2007) introduziram um mecanismo que permite modelar raciocínio incerto por meio de aproximações de conceitos. Elas estendem DLs com duas operações: as aproximações inferiores e superiores. Ambas aproximações são responsáveis por capturar incerteza a partir de uma relação de indiscernibilidade. Podemos formalmente definir a aproximação superior de um conceito como o conjunto de indivíduos que são indiscerníveis de pelo menos um indivíduo que é conhecido pertencer ao conceito. Similarmente, podemos definir a aproximação inferior de um conceito  $C$  como o conjunto de indivíduos  $a$  tal que todos os indivíduos indiscerníveis de  $a$  pertencem ao conceito. A seguir, apresentaremos com mais detalhes as características básicas da RDL  $\mathcal{ALC}$ , iniciando pela sintaxe, a semântica e por fim, formas alternativas de representação de aproximações.

#### 3.2.1 Sintaxe

Como mencionado acima, a ideia básica por de trás das RDLs é bem direta: pode-se aproximar um conceito incerto  $C$  através de limites inferiores e superiores. A aproximação superior de  $C$ , denotada por  $\overline{C}$ , é o conjunto de todos indivíduos que possivelmente pertencem a  $C$ . Ortogonalmente, a aproximação inferior de  $C$ , denotada por  $\underline{C}$ , é o conjunto de todos indivíduos que definitivamente pertencem a  $C$ .

**Definição 15 (Conceitos)** *Os conceitos em Rough  $\mathcal{ALC}$  são definidos pela regras de sintaxe abaixo, em que  $C$  e  $D$  são conceitos,  $A$  é um conceito atômico e  $R$  é um papel atômico:*

$$C, D \longrightarrow A \mid \top \mid \perp \mid \neg C \mid C \sqcap D \mid C \sqcup D \mid \exists R.C \mid \forall R.C \mid \overline{C} \mid \underline{C}.$$

RDLs não são restritas para DLs particulares, e podem ser definidas para uma DL qualquer. A linguagem de uma RDL baseada em uma DL arbitrária é construída com a adição das aproximações superiores e inferiores como construtores unários de conceitos, ou seja, se  $C$  é um conceito em uma RDL, então  $\overline{C}$  e  $\underline{C}$  também são conceitos em RDL. As noções de *rough* TBox e ABox, assim como uma base de conhecimento *rough* estendem as noções usuais.

*Exemplo 12.* (Exemplo da Sépsis) (KEET, 2010; SCHLOBACH; KLEIN, 2007). Sépsis é uma infecção que pode se desenvolver em um paciente a partir de qualquer infecção grave. Devido a essa infecção o paciente torna-se severamente doente, no qual facilmente resulta em falência dos órgãos e eventualmente a morte. A causa desta infecção não é clara, o que impede uma precisa caracterização do paciente séptico. Portanto, uma definição consensual de sépsis foi estabelecida em 1992 para definir diversos estágios de sépsis. Essa definição não é precisa, mas provê os critérios em que houve um consenso no tocante aos sintomas verificados em pacientes com sépsis severa. Esses critérios são chamados de critérios do osso. Pacientes que satisfazem o critério do osso podem ter sépsis, contudo, esse não é necessariamente o caso. Referimo-nos a esses pacientes como sendo possivelmente sépticos. Por outro lado, podemos definir como sépticos com certeza um grupo de pacientes que preencheram os critérios do osso e têm falência múltipla de órgãos. Referimo-nos a esses pacientes como sendo definitivamente pacientes sépticos. Os seguintes exemplos são conceitos em *rough*  $\mathcal{ALC}$ :

$$\overline{\text{Septico}}, \underline{\text{Septico}}, \exists \text{diag.} \overline{\text{Septico}}, \exists \text{diag.} \underline{\text{Septico}}, \forall \text{diag.} (\text{Septico} \sqcup \neg \text{Septico}).$$

### 3.2.2 Semântica

Uma interpretação *rough* é uma tripla  $I = (\Delta^I, \cdot^I, R^\sim)$ , em que  $\Delta^I$  é um domínio de indivíduos,  $\cdot^I$  é uma função de interpretação e  $R^\sim$  é uma relação de equivalência sobre  $\Delta^I$ . A função  $\cdot^I$  mapeia conceitos RDL para subconjuntos de  $\Delta^I$  e nome de papéis para relações binárias sobre o domínio  $\Delta^I$ . A interpretação para conceitos assim como para papéis permanece a mesma vista em  $\mathcal{ALC}$ ; a diferença surge da interpretação dos dois novos conceitos:

- $(\overline{C})^I = \{a \in \Delta^I \mid \exists b((a, b) \in R^\sim \wedge b \in C^I)\}$ ,
- $(\underline{C})^I = \{a \in \Delta^I \mid \forall b((a, b) \in R^\sim \rightarrow b \in C^I)\}$ .

Uma das vantagens desta forma de modelar incertezas em conceitos é que a complexidade de inferência não aumenta. De fato, raciocínio com RDL pode ser reduzido ao raciocínio com DL, traduzindo conceitos *rough* em conceitos DL usuais com uma nova relação que é reflexiva, simétrica e transitiva. A principal desvantagem é que nessa abordagem não há um ganho de expressividade nas RDLs em relação às DLs, ou seja, utilizando apenas DL é possível simular as aproximações presentes em RDL. Uma função de tradução de conceitos  $\cdot^t : RDL \rightarrow DL$  é definida a seguir mostrando essa equivalência de expressividade (introduzindo um novo papel  $R$  para a relação de indiscernibilidade):

- $A^t = A$ , para todos conceitos atômicos  $A$  em RDL,

- $(\overline{C})^t = \exists R.C^t$  e  $(\underline{C})^t = \forall R.C^t$ , para todos os outros conceitos em RDL.

A função de tradução é recursivamente aplicada em subconceitos para todos os outros construtores e para os axiomas da base de conhecimento. Essa definição é também estendida para axiomas de inclusões e asserções. Para qualquer linguagem DL com quantificação universal e existencial, RDLs podem ser simulados em DLs: um conceito  $C$  em RDL é satisfatível em uma interpretação *rough* relativa a um TBox  $\mathcal{T}$  se e somente se o conceito  $C^t$  em DL é satisfatível em uma interpretação relativa a um TBox  $\mathcal{T}^t$  (KEET, 2010; SCHLOBACH; KLEIN, 2007).

Uma outra forma de representar as equivalências entre indivíduos foi proposta em (FANIZZI et al., 2008), no qual foi introduzida uma maneira alternativa de aproximação mais expressiva em DL. Nessa abordagem, uma aproximação depende de um conjunto específico de conceitos para determinar a indiscernibilidade dos indivíduos (e não de uma relação de indiscernibilidade já explícita). Detalharemos melhor essa ideia a seguir.

### 3.2.3 Aproximação Contextual

Em (FANIZZI et al., 2008), foi introduzida a noção de relação de indiscernibilidade contextual em RDLs como uma maneira de definir uma relação de equivalência baseada em critérios de indiscernibilidade. Em particular, a noção de contexto é introduzida, permitindo a definição de relacionamentos específicos de equivalência para serem usados em aproximações. A principal vantagem dessa abordagem é que o raciocínio com classes de equivalência é otimizado, pois as relações de equivalências são descobertas no processo de inferência, diferentemente das tradicionais RDLs, em que a relação de equivalência deve ser explicitamente definida.

Nesta seção, serão apresentadas as noções de aproximação contextual em DL, iniciando pela noção de um contexto via uma coleção de conceitos. Então será mostrada a definição de uma relação de equivalência relacionada a um contexto e por fim, definidas as aproximações superiores e inferiores de conceitos usando essa relação de equivalência. Primeiramente, relembremos a noção de uma função de projeção em DL (FANIZZI; D'AMATO; ESPOSITO, 2007):

**Definição 16 (Projeção)** *Sejam  $I = (\Delta^I, \cdot^I)$  uma interpretação e  $C$  um conceito. A função de projeção  $\pi_C^I : \Delta^I \rightarrow \{0, *, 1\}$  é definida como*

$$\forall a \in \Delta^I : \pi_C^I(a) = \begin{cases} 1, & I \models C(a); \\ 0, & I \models \neg C(a); \\ *, & \text{caso contrário.} \end{cases}$$

Um contexto pode ser definido como um conjunto de características relevantes na forma de um conjunto de conceitos em DL, que podem codificar um tipo de similaridade a ser medida (GOLDSTONE; MEDIN; HALBERSTADT, 1997).

**Definição 17 (Contexto)** Um contexto  $\Sigma = \{C_1, \dots, C_n\}$  é um conjunto finito não-vazio de conceitos em DL.

*Exemplo 13.* Um possível contexto  $\Sigma$  para a campanha de uma propaganda de um novo produto é dado a seguir:

$$\Sigma = \{\text{Salário-Milionário}, \text{Proprietário-de-Imóvel}, \text{Empresário}\},$$

em que *Salário-Milionário*, *Proprietário-de-Imóvel* e *Empresário* são conceitos DL.

Dois indivíduos  $a$  e  $b$  são indiscerníveis com respeito ao contexto  $\Sigma = \{C_1, \dots, C_n\}$  e uma interpretação  $I$  se e somente se para todo  $C_i$ , em que  $i \in \{1, \dots, n\}$ ,  $\pi_{C_i}^I(a) = \pi_{C_i}^I(b)$ . Isso facilmente induz a uma relação de equivalência:

**Definição 18 (Relação de Indiscernibilidade Contextual)** Sejam  $\Sigma = \{C_1, \dots, C_n\}$  um contexto e  $I$  uma interpretação. A relação de indiscernibilidade  $R_\Sigma$  induzida por  $\Sigma$  é definida como:

$$R_\Sigma = \{(a, b) \in \Delta^I \times \Delta^I \mid \text{para todo } C_i \text{ em que } i \in \{1, \dots, n\}, \pi_{C_i}^I(a) = \pi_{C_i}^I(b)\}.$$

Por fim, as aproximações contextuais são definidas como

**Definição 19 (Aproximação Contextual Inferior/Superior)** Sejam  $\Sigma = \{C_1, \dots, C_n\}$  um contexto,  $C$  um conceito e  $I$  uma interpretação. Então, as aproximações contextuais superior e inferior de  $C$  em relação a  $\Sigma$ , denotadas por  $\overline{C}^\Sigma$  e  $\underline{C}_\Sigma$ , respectivamente, são definidas como

- $(\overline{C}^\Sigma)^I = \{a \in \Delta^I \mid \exists b((a, b) \in R_\Sigma \wedge b \in C^I)\},$
- $(\underline{C}_\Sigma)^I = \{a \in \Delta^I \mid \forall b((a, b) \in R_\Sigma \rightarrow b \in C^I)\}.$

As aproximações contextuais desempenharão um papel central no processo de refinamento de consultas. Em nosso trabalho, refinar um conceito significa aplicar a aproximação contextual inferior (no caso de uma restrição) ou a aproximação contextual superior (no caso de um relaxamento). Para isso, é preciso identificar primeiramente um contexto que possa ser aplicado nas aproximações. Esse processo de descoberta de contextos e mais algumas outras questões envolvendo DLs serão abordados nos capítulos seguintes. Antes de começarmos a apresentar as contribuições da dissertação, iremos mostrar trabalhos da literatura envolvendo refinamento de consultas em DLs.

## 4 REFINAMENTO DE CONSULTAS EM LÓGICAS DE DESCRIÇÃO

Neste capítulo, mostraremos as abordagens de refinamento de consultas em DLs presentes na literatura. Por refinamento de consultas em DLs, compreendemos determinar a satisfatibilidade de um conceito através de outros meios que não sejam computar a satisfatibilidade do próprio conceito. Por exemplo, uma possível maneira de aproximar um conceito seria traduzi-lo em um novo conceito tal que os dois conceitos possuam algum tipo de relação. Podemos dizer que uma aproximação de um conceito é feita com objetivo de deduzir informações que não se apresentam explícitas em uma base de conhecimento. Sendo assim, apresentaremos a seguir duas abordagens propostas na literatura para tratar do problema de refinamento em DLs: a primeira abordagem trata do refinamento apenas em axiomas de ABoxes, enquanto que a segunda abordagem aplica o refinamento apenas em axiomas de TBoxes.

### 4.1 Refinamento em ABoxes - Lógica de Descrição Possibilística

No trabalho apresentado por (LIAU; YAO, 2001), foi aplicada uma abordagem possibilística à DL  $\mathcal{ALC}$ , introduzindo a lógica  $\mathcal{P}\mathcal{ALC}$ , a fim de resolver problemas de relaxamento e restrição de consultas. Os relaxamentos estão relacionados com o operador de possibilidade, enquanto que as restrições estão relacionadas com o operador de necessidade. O raciocínio possibilístico é feito a partir de relações entre os indivíduos que medem o grau de proximidade entre eles, de acordo com um conjunto de propriedades. Essas relações de proximidades entre indivíduos tratam de um conhecimento que se tem acerca do problema e da aplicação e que é utilizada para a resolução de um tipo específico de problema.

A sintaxe de  $\mathcal{P}\mathcal{ALC}$  é baseada na sintaxe de  $\mathcal{ALC}$  com a adição de operadores de necessidade e possibilidade. Sejam  $a$  um indivíduo,  $A$  um conceito atômico,  $R$  um papel atômico e  $C$  e  $D$  conceitos complexos. As regras de formação para os demais conceitos são as seguintes:

$$C, D \longrightarrow A \mid \top \mid \perp \mid a \mid \neg C \mid C \sqcap D \mid C \sqcup D \mid \exists R.C \mid \forall R.C \mid [\alpha]C \mid [\alpha]^+C \mid \langle \alpha \rangle C \mid \langle \alpha \rangle^+ C,$$

em que  $\alpha \in [0, 1]$ . Note que um nome de indivíduo é também um conceito, o qual é chamado de nominal. O propósito da representação desse conceito é tratá-lo como um conjunto unitário. Os novos operadores  $[\alpha]$  e  $[\alpha]^+$  são utilizados para a quantificação da medida de necessidade e  $\langle \alpha \rangle$  e  $\langle \alpha \rangle^+$  para a quantificação da medida de possibilidade. Por exemplo, um indivíduo  $a$  está no conceito  $\langle \alpha \rangle C$  se e somente se  $a$  é similar a outro indivíduo que está em  $C$  com pelo menos um grau numérico  $\alpha$ . Os axiomas presentes em  $\mathcal{P}\mathcal{ALC}$  são os mesmos encontrados em  $\mathcal{ALC}$ , i.e.,  $C \equiv D$  e  $C \sqsubseteq D$  para os axiomas terminológicos e  $C(a)$  e  $R(a, b)$  para axiomas de asserção.

A semântica de  $\mathcal{P}\mathcal{ALC}$  é dada por uma tripla  $I = (\Delta^I, \cdot^I, E)$ , em que  $E$  é uma relação de similaridade na qual  $E : \Delta^I \times \Delta^I \rightarrow [0, 1]$  e  $(\Delta^I, \cdot^I)$  é uma interpretação em  $\mathcal{ALC}$ . Os conceitos  $[\alpha]C$ ,  $[\alpha]^+C$ ,  $\langle \alpha \rangle C$  e  $\langle \alpha \rangle^+ C$  são interpretados como

- $([\alpha]C)^I = \{a \in \Delta^I \mid \forall b((a, b) \in E_\alpha \rightarrow b \in C^I)\}$ ;

- $([\alpha]C^+)^I = \{a \in \Delta^I \mid \forall b((a,b) \in E_\alpha^+ \rightarrow b \in C^I)\};$
- $(\langle \alpha \rangle C)^I = \{a \in \Delta^I \mid \exists b((a,b) \in E_\alpha \wedge b \in C^I)\};$
- $(\langle \alpha \rangle^+ C)^I = \{a \in \Delta^I \mid \exists b((a,b) \in E_\alpha^+ \wedge b \in C^I)\}.$

em que  $E_\alpha = \{(u,v) \mid E(u,v) \geq \alpha\}$  e  $E_\alpha^+ = \{(u,v) \mid E(u,v) > \alpha\}$  denotando as medidas de similaridade entre os indivíduos. Essas medidas podem significar, por exemplo, a distância entre dois pontos quando eles estiverem representados em um mapa ou também o grau de semelhança/coincidência entre dois indivíduos com relação a um conjunto de conceitos.

Os conceitos  $[\alpha]C$  e  $[\alpha]C^+$  representam os indivíduos que necessariamente estão em  $C$  de acordo com um grau de similaridade respectivamente maior ou igual e maior a  $\alpha$ . Já  $\langle \alpha \rangle C$  e  $\langle \alpha \rangle C^+$  representam os indivíduos que possivelmente estão em  $C$  de acordo com um grau de similaridade respectivamente maior ou igual e maior a  $\alpha$ . As definições de satisfação e consequência lógica são as mesmas encontradas em  $\mathcal{P} \mathcal{ALC}$ . Para aplicações de  $\mathcal{P} \mathcal{ALC}$  em refinamento de consultas, vamos considerar o seguinte exemplo.

*Exemplo 14.* (Relaxamento de consulta) (LIAU; YAO, 2001) Sejam  $\mathcal{K}$  uma base de conhecimento em  $\mathcal{P} \mathcal{ALC}$  e  $C$  um conceito. Então para uma consulta de  $C$  em  $\mathcal{K}$ , o problema é encontrar um indivíduo  $a$  tal que  $\mathcal{K} \models C(a)$ . Contudo, às vezes,  $C$  é muito restritivo, o que pode acarretar em um resultado que não é satisfatório para a consulta. Nesse caso, pode-se relaxar a consulta usando o conceito  $\langle \alpha \rangle C$  para algum  $\alpha < 1$ . Por exemplo, pode-se imaginar um sistema de reservas de hotéis, em que um usuário pode inserir a seguinte consulta  $C$

*Perto-estação-trem*  $\sqcap$   $\neg$  *Caro*  $\sqcap$  *Permitido-animais*  $\sqcap$   $\exists$ *temQuarto.Duplo*,

e o sistema consequentemente encontrar um hotel satisfazendo as condições. Contudo, o hotel desejado não está disponível durante o período escolhido (o hotel  $a$  não está disponível). Assim o usuário poderia relaxar a consulta para  $\langle 0.8 \rangle C$  ( $\alpha$  indicando uma medida de similaridade baseada na distância entre os hotéis) para encontrar um hotel próximo satisfazendo as condições, i.e., se  $\langle 0.8 \rangle C$  for satisfeito, significa que existe algum hotel próximo ao hotel  $a$  com um grau de proximidade 0.8 que satisfaz as condições da consulta  $C$ .

Note que no escopo das aplicações, é necessário existir uma coleção de relações de similaridade, ou seja, um conhecimento prévio sobre o problema a ser resolvido. Contudo, em muitos casos essas relações de similaridade não podem ser geradas automaticamente a partir da base de conhecimento das aplicações. Em outras palavras, nem sempre teremos um outro tipo de conhecimento sobre um problema além do que já está representado em uma base de conhecimento. Motivados por esse obstáculo, no Capítulo 5 nós propomos uma abordagem de aproximação em DL, em que um tipo específico de relação de similaridade é gerado automaticamente a partir da própria base de conhecimento.

## 4.2 Refinamento em TBoxes - Subsunção Aproximada

A ideia do raciocínio aproximado em terminologias é baseada na interpretação da ontologia de uma maneira diferente da convencional. Em particular, a ideia é buscar subconjuntos e superconjuntos de interpretações chamados de aproximações inferior e superior de uma interpretação. Essa abordagem é apresentada em (SCHLOBACH et al., 2007), seguindo as ideias de (SCHAERF; CADOLI, 1995). Outros trabalhos nessa linha de pesquisa são propostos em (AKAHANI; HIRAMATSU; SATOH, 2003; STUCKENSCHMIDT, 2007b, 2007a). Um problema presente nesses trabalhos é que se tratam de abordagens sintáticas, (diferente da abordagem semântica da lógica  $\mathcal{P} \mathcal{ALC}$ ) e portanto não existe a garantia de que os resultados de um refinamento tenham relação com a consulta original.

Seguindo o trabalho de (SCHAERF; CADOLI, 1995), para serem realizadas as aproximações em (SCHLOBACH et al., 2007) é utilizada uma nova noção de interpretação, em que o vocabulário que é levado em conta na definição dessa interpretação é restrito para um determinado conjunto (um subconjunto do vocabulário utilizado pela interpretação usual), denotado por  $S$  e chamado de conjunto de aproximação. As definições das  $S$ -aproximações inferiores e superiores são as seguintes:

**Definição 20 (Aproximação Inferior/Superior)** (SCHLOBACH et al., 2007) *A aproximação inferior  $\cdot^{I\bar{S}}$  e a aproximação superior  $\cdot^{I^+S}$  são definidas como:*

- Para o conceito atômico  $A$ ,
  - $A^{I\bar{S}} = A^{I^+S} = A^I$ , se  $A \in S$ ;
  - $A^{I\bar{S}} = \emptyset$  e  $A^{I^+S} = \Delta^I$ , se  $A \notin S$ .
- Para os demais conceitos,
  - $(\neg C)^{I\bar{S}} = \Delta^I \setminus C^{I^+S}$  e  $(\neg C)^{I^+S} = \Delta^I \setminus C^{I\bar{S}}$ ;
  - $(C \sqcap D)^{I\bar{S}} = C^{I\bar{S}} \cap D^{I\bar{S}}$  e  $(C \sqcap D)^{I^+S} = C^{I^+S} \cap D^{I^+S}$ ;
  - $(C \sqcup D)^{I\bar{S}} = C^{I\bar{S}} \cup D^{I\bar{S}}$  e  $(C \sqcup D)^{I^+S} = C^{I^+S} \cup D^{I^+S}$ ;
  - $(\exists R.C)^{I\bar{S}} = \{a \in \Delta^I \mid \exists b((a, b) \in R^I \wedge b \in C^{I\bar{S}})\}$ ;
  - $(\exists R.C)^{I^+S} = \{a \in \Delta^I \mid \exists b((a, b) \in R^I \wedge b \in C^{I^+S})\}$ ;
  - $(\forall R.C)^{I\bar{S}} = \{a \in \Delta^I \mid \forall b((a, b) \in R^I \rightarrow b \in C^{I\bar{S}})\}$ ;
  - $(\forall R.C)^{I^+S} = \{a \in \Delta^I \mid \forall b((a, b) \in R^I \rightarrow b \in C^{I^+S})\}$ .

A intuição dessas interpretações é que a aproximação superior pode se tratar de uma interpretação mais ampla que a interpretação clássica, o que vai depender do conjunto  $S$  associado a interpretação superior, pois os conceitos atômicos que não pertencem a  $S$  serão interpretados como o conjunto universo  $\Delta^I$ . Por outro lado, a aproximação inferior tem o objetivo de criar uma interpretação mais restrita que a interpretação clássica, já que os conceitos atômicos que não pertencem a  $S$  serão interpretados como o conjunto vazio.

Baseado nessas aproximações para uma interpretação, podemos definir a noção de um modelo aproximado para uma terminologia. A intuição básica é que gradativamente a terminologia é relaxada ao forçar que o lado esquerdo dos axiomas terminológicos ( $C \sqsubseteq D$ ) seja mais específico, enquanto que o lado direito dos axiomas seja mais gerai. Isso pode ser alcançado utilizando as duas noções apresentadas acima, pois elas possuem as seguintes propriedades: dada uma interpretação  $I$ , a aproximação inferior  $I_S^-$  é um subconjunto de  $I$  e a aproximação superior  $I_S^+$  é um superconjunto de  $I$ . Conseqüentemente, a aplicação dessas operações em axiomas terminológicos pode acarretar em resultados que antes não eram gerados.

**Definição 21 (Modelo S-Aproximado)** (SCHLOBACH et al., 2007) *Sejam  $C, D$  conceitos e  $\mathcal{T}$  um TBox. Uma interpretação  $I$  é um modelo S-aproximado de um axioma  $C \sqsubseteq D$  sse  $C^{I_S^-} \subseteq D^{I_S^+}$ . Uma interpretação  $I$  é um modelo S-aproximado de um TBox  $\mathcal{T}$ , se e somente se ele é um modelo S-aproximado para todos axiomas  $\alpha \in \mathcal{T}$ . Neste caso, escrevemos que  $I \models_S \mathcal{T}$ .*

Aproximações preservam modelos, i.e., se uma interpretação é um modelo para um axioma ou um TBox, ela é também um modelo S-aproximado para qualquer subconjunto  $S$  de um vocabulário. A subsunção aproximada pode ser definida de uma maneira clássica como uma relação de subconjunto entre conceitos interpretados, mas agora apenas levando em consideração os modelos aproximados de um TBox.

**Definição 22 (Subsunção aproximada)** (SCHLOBACH et al., 2007) *Seja  $\mathcal{T}$  um TBox. Dizemos que um conceito  $D$  S-aproximadamente subsume um conceito  $C$  com relação ao TBox  $\mathcal{T}$  ( $\mathcal{T} \models_S C \sqsubseteq D$ ) sse  $C^I \subseteq D^I$  para todos modelos S-aproximados  $I$  de  $\mathcal{T}$ .*

A partir das ideias propostas neste capítulo, aplicaremos intuição semelhante na aproximação de axiomas de asserções; desta vez utilizando a abordagem de *rough sets* como método de aproximação. Os detalhes da nossa abordagem serão apresentados ao longo dos Capítulos 5-7. Destacamos que focaremos apenas em aproximações dos axiomas de ABox e que aproximações em axiomas de TBox serão investigadas apenas em trabalhos futuros.

A partir do próximo capítulo, começaremos a apresentar as contribuições desta dissertação. A primeira delas trata-se da criação de uma extensão da DL *Rough  $\mathcal{ALC}$* , capaz de representar e raciocinar sobre dois diferentes tipos de incerteza. Logo após, desenvolveremos um método para o refinamento de consultas nessa nova lógica. E finalizando o trabalho, sugeriremos medidas de qualidade para os resultados obtidos por esse método de refinamento.

## 5 LÓGICAS DE DESCRIÇÃO *ROUGH* PARACONSISTENTE

A primeira contribuição a ser apresentada nesta dissertação trata-se da representação de uma linguagem em DL capaz de representar informações incertas e aproximações de conceitos. Em especial, nos focaremos em DLs capazes de representar raciocínio paraconsistente, um tipo específico de incerteza. Na literatura de DL já existe uma versão com raciocínio paraconsistente, proposta em (MA; HITZLER; LIN, 2006), sendo uma extensão tetravalorada da  $\mathcal{ALC}$ . A semântica dessa lógica tetravalorada segue a da bem conhecida lógica paraconsistente de Belnap (BELNAP, 1977). Com relação a trabalhos com *rough sets* e DLs, apenas abordagens envolvendo DLs bivaloradas e DLs *fuzzy* (BOBILLO; STRACCIA, 2009) foram propostas na literatura. Entretanto, no trabalho apresentado em (VITÓRIA; SZAŁAS; MAŁUSZYŃSKI, 2008), foi proposta uma extensão tetravalorada paraconsistente de *rough sets*, mas que segue uma semântica diferenciada daquela de Belnap.

Ao longo deste capítulo, iremos propor duas extensões paraconsistentes e *rough* de DLs: a primeira trata-se de uma extensão da lógica apresentada em (MA; HITZLER; LIN, 2006), acrescentando os operadores de aproximação superior e inferior; e a segunda trata-se da criação de uma DL que segue a semântica dos *rough sets* paraconsistentes de (VITÓRIA; SZAŁAS; MAŁUSZYŃSKI, 2008). Além disso, mostraremos seus resultados de complexidade obtidos até agora. Consequentemente, no próximo capítulo, utilizaremos os axiomas dessas duas lógicas para a fundamentação do processo de refinamento de consultas.

Na Seção 5.1, nós apresentaremos uma breve explicação sobre a Lógica de Belnap. Na Seção 5.2, nós apresentaremos uma extensão paraconsistente *rough* de  $\mathcal{ALC}$ , chamada de *Rough  $\mathcal{ALC}_4$* , que é baseada na semântica de Belnap. Na seção 5.3, nós mostraremos algumas definições que serão utilizadas na segunda DL proposta: noções de operações, conjuntos e aproximações tetravaloradas definidas em (VITÓRIA; SZAŁAS; MAŁUSZYŃSKI, 2008). Por fim, durante a Seção 5.4, nós introduziremos a lógica  $PR_{\mathcal{ALC}}$ . Na Subseção 5.4.1, aplicaremos a aproximação contextual à  $PR_{\mathcal{ALC}}$ . Na Seção 5.4.2, nós apresentaremos algumas tarefas de consultas que podem ser usadas em  $PR_{\mathcal{ALC}}$ ; em particular, as noções de aproximações *loose* e *tight*. Finalizando, na Seção 5.4.3 provaremos a complexidade de satisfação de  $PR_{\mathcal{ALC}}$ , através de uma tradução para a DL  $\mathcal{ALC}$ .

### 5.1 Lógica de Belnap

A Lógica de Belnap (BELNAP, 1977) possui quatro valores verdade no lugar dos dois valores verdade clássicos. Esses valores são o *verdadeiro*, *falso*, *desconhecido* e *inconsistente*. Usaremos os símbolos **t**, **f**, **u**, **i**, respectivamente, para esses valores. O valor verdade **i** representa a informação contraditória. O valor **u** significa nem verdadeiro nem falso, i.e. a falta de qualquer informação sobre veracidade ou falsidade.

Sintaticamente, Lógica de Belnap é muito similar à lógica clássica. Contudo, ela apresenta outras noções de implicação. De fato, iremos apresentar três noções de implicação presentes na literatura. Os conectivos que serão utilizados aqui são a negação ( $\neg$ ), a disjunção

$\wedge$	f	u	i	t	$\vee$	f	u	i	t	$\neg$	f	t
f	f	f	f	f	f	f	u	i	t	f	t	
u	f	u	f	u	u	u	t	t		u	u	
i	f	f	i	i	i	i	t	i	t	i	i	
t	f	u	i	t	t	t	t	t		t	f	

Tabela 5.1: Tabelas verdade para  $\wedge, \vee$  e  $\neg$ 

$\mapsto$	f	u	i	t	$\supset$	f	u	i	t	$\rightarrow$	f	u	i	t
f	t	t	t	t	f	t	t	t	t	f	t	t	t	t
u	u	u	t	t	u	t	t	t	t	u	u	t	u	t
i	i	t	i	t	i	f	u	i	t	i	f	u	i	t
t	f	u	i	t	t	f	u	i	t	t	f	u	f	t

Tabela 5.2: Tabelas verdade para  $\mapsto, \supset$  e  $\rightarrow$ 

( $\vee$ ), a conjunção ( $\wedge$ ), a implicação material ( $\mapsto$ ), a implicação interna ( $\supset$ ) e a implicação forte ( $\rightarrow$ ). Discutiremos essas implicações com mais detalhes na sequência.

A implicação material pode ser definida por meio da negação e disjunção como na lógica clássica. Contudo, ela não satisfaz Modus Ponens ou o teorema da dedução. A implicação interna satisfaz Modus Ponens e teorema da dedução, mas não pode ser definida por meios dos outros conectivos. Além disso, a implicação interna não satisfaz a propriedade de contraposição (i.e.  $x \supset y$  não implica  $\neg x \supset \neg y$ ). A implicação forte é descrita a partir da implicação interna, no qual ela adicionalmente satisfaz a contraposição. De fato, uma maneira alternativa de visualizar as tabelas verdade para os conectivos de implicação é:

- $x \mapsto y$  é definível como  $\neg x \vee y$ ;
- $x \supset y$  é avaliado para  $y$  se  $x \in \{t, i\}$ ;  $t$  se  $x \in \{f, u\}$ ;
- $x \rightarrow y$  é definível como  $(x \supset y) \wedge (\neg y \supset \neg x)$ .

A interpretação das fórmulas são mapeamentos do conjunto de fórmulas para o conjunto dos quatro possíveis valores verdade, respeitando a tabela verdade dos conectivos como detalhado nas Tabela 5.1 e 5.2. Modelos são definidos a seguir, em que  $\{t, i\}$  são os valores designados (i.e. os valores considerados verdade em uma consequência).

**Definição 23 (Modelo)** *Sejam  $I$  uma interpretação tetravalorada,  $\Gamma$  um conjunto de fórmulas e  $\varphi$  uma fórmula na Lógica de Belnap. Dizemos que  $I$  é um modelo tetravalorado de  $\varphi$  se e somente se  $\varphi^I \in \{t, i\}$ .  $I$  é um modelo tetravalorado de  $\Gamma$  se e somente se  $I$  é um modelo tetravalorado de cada fórmula em  $\Gamma$ .  $\Gamma$  acarreta  $\varphi$ , escrito  $\Gamma \models \varphi$ , se e somente se todo modelo tetravalorado de  $\Gamma$  é um modelo tetravalorado de  $\varphi$ .*

## 5.2 Lógica de Descrição *Rough* $\mathcal{ALC}_4$

Nós iremos descrever a sintaxe e semântica da Lógica de Descrição *Rough*  $\mathcal{ALC}_4$ . Essa abordagem é uma adaptação da DL  $\mathcal{ALC}_4$  introduzida em (MA; HITZLER; LIN, 2006) com a adição dos operadores de aproximação de conjuntos. Como mostrado no Capítulo 3, a aproximação de conjuntos não aumenta a expressividade e complexidade de uma DL em particular; essa adaptação preserva todas as propriedades de  $\mathcal{ALC}_4$ .

Sintaticamente, *Rough*  $\mathcal{ALC}_4$  quase não difere de *Rough*  $\mathcal{ALC}$ . Conceitos complexos e asserções são definidas da mesma maneira. Para inclusões de conceitos, a  $\mathcal{ALC}_4$  (assim como a *Rough*  $\mathcal{ALC}_4$ ) possui três tipos de axiomas de inclusão de conceitos, referentes às três implicações mostradas anteriormente:

$$C \mapsto D \text{ (axioma de inclusão material),}$$

$$C \sqsubset D \text{ (axioma de inclusão interna),}$$

$$C \rightarrow D \text{ (axioma de inclusão forte).}$$

Como usual, semanticamente, interpretações mapeiam indivíduos para elementos do domínio da interpretação. Para conceitos, entretanto, é preciso fazer algumas modificações na noção de interpretação a fim de permitir o raciocínio com inconsistências.

Intuitivamente, em uma lógica tetravalorada precisamos considerar quatro situações que podem ocorrer em termos de pertinência de um indivíduo a um conceito: (1) sabemos que ele está contido no conceito, (2) sabemos que ele não está contido, (3) não temos conhecimento se ele está contido ou não, (4) temos informação contraditória, afirmando que o indivíduo está contido e não está contido no conceito. Existem várias maneiras equivalentes de formalizar essa noção; uma delas é descrita na sequência.

Para um dado domínio  $\Delta^I$  e um conceito  $C$ , uma interpretação sobre  $\Delta^I$  assinala para  $C$  um par  $\langle P, N \rangle$  de subconjuntos (não necessariamente disjuntos) de  $\Delta^I$ . Intuitivamente,  $P$  é o conjunto de elementos que sabemos que pertencem a  $C$ , enquanto que  $N$  é o conjunto de elementos que sabemos que não pertencem a  $C$ . Por simplicidade de notação, as funções  $\text{proj}^+(\cdot)$  e  $\text{proj}^-(\cdot)$  são definidas por

$$\text{proj}^+ \langle P, N \rangle = P \text{ e } \text{proj}^- \langle P, N \rangle = N.$$

Formalmente, uma interpretação tetravalorada é uma tripla  $I = (\Delta^I, \cdot^I, R^\sim)$  com  $\Delta^I$  como domínio,  $R^\sim$  é uma relação de equivalência sobre  $\Delta^I$  e em que  $\cdot^I$  é uma função de assinalamento de elementos de  $\Delta^I$  para indivíduos e subconjuntos de  $\Delta^I \times \Delta^I$  para conceitos tal que as condições mostradas abaixo são satisfeitas:

- Para conceitos atômicos  $A$ ,  $A^I = \langle P, N \rangle$ , em que  $P, N \subseteq \Delta^I$ ;
- Para papéis atômicos  $R$ ,  $R^I = \langle P_1 \times P_2, N_1 \times N_2 \rangle$ , em que  $P_1 \times P_2, N_1 \times N_2 \subseteq \Delta^I \times \Delta^I$ ;

- $\top^I = \langle \Delta^I, \emptyset \rangle$ ;
- $\perp^I = \langle \emptyset, \Delta^I \rangle$ ;
- $(\neg C)^I = \langle N, P \rangle$  se  $C^I = \langle P, N \rangle$ ;
- $(C \cap D)^I = \langle P_1 \cap P_2, N_1 \cup N_2 \rangle$ , se  $C^I = \langle P_1, N_1 \rangle$  e  $D^I = \langle P_2, N_2 \rangle$ ;
- $(C \cup D)^I = \langle P_1 \cup P_2, N_1 \cap N_2 \rangle$ , se  $C^I = \langle P_1, N_1 \rangle$  e  $D^I = \langle P_2, N_2 \rangle$ ;
- $(\exists R.C)^I = \langle \{x \mid \exists y, (x, y) \in \text{proj}^+(R^I) \text{ e } y \in \text{proj}^+(C^I)\}, \{x \mid \forall y, (x, y) \in \text{proj}^+(R^I) \text{ implica } y \in \text{proj}^-(C^I)\} \rangle$ ;
- $(\forall R.C)^I = \langle \{x \mid \forall y, (x, y) \in \text{proj}^+(R^I) \text{ implica } y \in \text{proj}^+(C^I)\}, \{x \mid \exists y, (x, y) \in \text{proj}^+(R^I) \text{ e } y \in \text{proj}^-(C^I)\} \rangle$ ;
- $(\overline{C})^I = \langle \{x \mid \exists y, (x, y) \in R^\sim \text{ e } y \in \text{proj}^+(C^I)\}, \{x \mid \forall y, (x, y) \in R^\sim \text{ implica } y \in \text{proj}^-(C^I)\} \rangle$ ;
- $(\underline{C})^I = \langle \{x \mid \forall y, (x, y) \in R^\sim \text{ implica } y \in \text{proj}^+(C^I)\}, \{x \mid \exists y, (x, y) \in R^\sim \text{ e } y \in \text{proj}^-(C^I)\} \rangle$ .

Note que as condições acima para as restrições de papéis são descritas de uma maneira que as equivalências lógicas  $\neg(\forall R.C) = \exists R.(\neg C)$  e  $\neg(\exists R.C) = \forall R.(\neg C)$  são preservadas. Essa foi uma maneira conveniente encontrada em (MA; HITZLER; LIN, 2006) para lidar com restrições de papéis e que permite uma tradução direta de  $\mathcal{ALC}_4$  para a  $\mathcal{ALC}$ . Note também que nessa linguagem, apenas a parte positiva da interpretação dos papéis é requerida, pois são apenas utilizados papéis atômicos.

Obviamente, sob as restrições de que  $P \cap N = \emptyset$  e  $P \cup N = \Delta^I$ , as interpretações tetravaloradas tornam-se bivaloradas padrão. A correspondência entre os valores verdade e a pertinência de conceitos e papéis são descritas da seguinte maneira: sejam  $a, b \in \Delta^I$ ,  $C$  um nome de conceito e  $R$  um nome de papel. Temos que

- $C^I(a) = \mathbf{t}$ , sse  $a^I \in \text{proj}^+(C^I)$  e  $a^I \notin \text{proj}^-(C^I)$ ;
- $C^I(a) = \mathbf{f}$ , sse  $a^I \notin \text{proj}^+(C^I)$  e  $a^I \in \text{proj}^-(C^I)$ ;
- $C^I(a) = \mathbf{i}$ , sse  $a^I \in \text{proj}^+(C^I)$  e  $a^I \in \text{proj}^-(C^I)$ ;
- $C^I(a) = \mathbf{u}$ , sse  $a^I \notin \text{proj}^+(C^I)$  e  $a^I \notin \text{proj}^-(C^I)$ ;
- $R^I(a, b) = \mathbf{t}$ , sse  $(a^I, b^I) \in \text{proj}^+(R^I)$  e  $(a^I, b^I) \notin \text{proj}^-(R^I)$ ;
- $R^I(a, b) = \mathbf{f}$ , sse  $(a^I, b^I) \notin \text{proj}^+(R^I)$  e  $(a^I, b^I) \in \text{proj}^-(R^I)$ ;
- $R^I(a, b) = \mathbf{i}$ , sse  $(a^I, b^I) \in \text{proj}^+(R^I)$  e  $(a^I, b^I) \in \text{proj}^-(R^I)$ ;

- $R^I(a, b) = \mathbf{u}$ , sse  $(a^I, b^I) \notin \text{proj}^+(R^I)$  e  $(a^I, b^I) \notin \text{proj}^-(R^I)$ .

Por fim, seguem as semânticas dos diferentes tipos de axiomas.

- $I \models C \mapsto D$  sse  $\Delta^I \setminus \text{proj}^-(C^I) \subseteq \text{proj}^+(D^I)$ ;
- $I \models C \sqsubset D$  sse  $\text{proj}^+(C^I) \subseteq \text{proj}^+(D^I)$ ;
- $I \models C \rightarrow D$  sse  $\text{proj}^+(C^I) \subseteq \text{proj}^+(D^I)$  e  $\text{proj}^-(D^I) \subseteq \text{proj}^-(C^I)$ ;
- $I \models C(a)$  sse  $a^I \in \text{proj}^+(C^I)$ ;
- $I \models R(a, b)$  sse  $(a^I, b^I) \in \text{proj}^+(R^I)$ .

Dizemos que uma interpretação tetravalorada  $I$  satisfaz uma base de conhecimento tetravalorada  $\mathcal{K}$  (i.e.,  $I$  é um modelo de  $\mathcal{K}$ ) se e somente se ela satisfaz cada axioma de asserção e inclusão em  $\mathcal{K}$ . Uma base de conhecimento  $\mathcal{K}$  é satisfatível (insatisfatível) se e somente se existe (não existe) tal modelo.

Com relação a complexidade de decisão de *Rough*  $\mathcal{ALC}_4$ , foi mostrado em (MA; HITZLER; LIN, 2006) que a complexidade de  $\mathcal{ALC}_4$  é equivalente a complexidade de  $\mathcal{ALC}$ . Esse resultado mostra que esse raciocínio paraconsistente não é mais expressivo que o raciocínio bivalorado e que pode ser representado em uma DL bivalorada sem o acréscimo de complexidade. Para provar tal resultado, foi mostrada uma tradução polinomial de uma base de conhecimento de  $\mathcal{ALC}_4$  para uma base de conhecimento de  $\mathcal{ALC}$ , que preserva todos os seus resultados de inferência. De acordo com esse resultado, podemos mostrar facilmente que a DL *Rough*  $\mathcal{ALC}_4$  tem a mesma complexidade de  $\mathcal{ALC}_4$  através de uma tradução de *Rough*  $\mathcal{ALC}_4$  para  $\mathcal{ALC}_4$ , que é similar à tradução de *Rough*  $\mathcal{ALC}$  para  $\mathcal{ALC}$  (KEET, 2010; SCHLOBACH; KLEIN, 2007).

### 5.3 *Rough Sets* Paraconsistentes

Em (VIANA; ALCÂNTARA; MARTINS, 2011) (publicado no 24th International Workshop on Description Logics, DL 2011), introduzimos uma diferente extensão tetravalorada da DL *Rough*  $\mathcal{ALC}$ , chamada de  $PR_{\mathcal{ALC}}$ , que permite a representação de conceitos contraditórios e incompletos assim como suas aproximações inferiores e superiores. Além disso, explorando a representação de informação incompleta e contraditória, apresentamos algumas formas de gerar similaridades entre os conceitos que serão usadas posteriormente nas aproximações.

Esse trabalho é baseado em (VITÓRIA; SZALAS; MAŁUSZYŃSKI, 2008), o qual introduziu a noção de *rough sets* paraconsistentes, que estende as ideias da teoria dos *rough sets*, com intuito de representar informação incompleta e contraditória. Em vez de aplicar aproximações a conjuntos clássicos, foram levados em consideração conjuntos tetra-valorados.

Além disso, uma relação de similaridade (i.e., uma versão mais fraca da relação de indiscernibilidade que é pelo menos reflexiva) é utilizada em vez de uma relação de equivalência com o objetivo de modelar diferentes níveis de incerteza.

Nesse trabalho, nós adaptamos as noções de *rough sets* paraconsistentes para DL, introduzindo uma RDL paraconsistente (PRDL), chamada  $PR_{\mathcal{ALC}}$ . Nós introduzimos duas relações de similaridade para lidar com aproximações de conceitos e também aplicamos a noção de aproximação contextual (FANIZZI et al., 2008) em nossa abordagem como uma forma alternativa de aproximação de conceitos. Finalmente, nós apresentamos algumas tarefas de raciocínio relacionadas à refinamento de consultas que podem ser aplicadas com essas operações introduzidas.

A fim de representar informação incompleta e contraditória, foi introduzida uma abordagem de *rough set* levando em consideração conjuntos tetravalorados em vez de conjuntos elementares (VITÓRIA; SZAŁAS; MAŁUSZYŃSKI, 2008). Nesses conjuntos tetravalorados, um elemento pode pertencer a um dado conjunto ou ele pode não pertencer ao conjunto ou sua pertinência ao conjunto pode ser desconhecida devido à informação incompleta ou até mesmo ser inconsistente devido a evidências contraditórias. Sob esse ponto de vista, funções de pertinência e operações sobre conjuntos são também tetravaloradas, em que os valores lógicos são **t** (*verdadeiro*), **f** (*falso*), **i** (*inconsistente*) e **u** (*desconhecido*). Além disso, como o conhecimento disponível é incompleto, em vez de relações de indiscernibilidade, os autores recorreram à relações de similaridades para agrupar elementos que estão próximos uns aos outros. Como existe a presença de informação incompleta, nem sempre será possível distinguir precisamente os indivíduos através de um conjunto de atributos. Consequentemente, as noções de aproximação superior e inferior serão extensões das definições usuais de aproximações da teoria.

### 5.3.1 Semântica Tetravalorada

A linguagem utilizada em (VITÓRIA; SZAŁAS; MAŁUSZYŃSKI, 2008) para representar os quatro valores verdade foi adaptada da Lógica de Belnap (BELNAP, 1977), que é fundamentada na noção de birreticulados (GINSBERG, 1988). Pela sua definição, os birreticulados possuem duas ordens de valores verdade: a ordem de verdade ( $\leq_t$ ) e a ordem de conhecimento ( $\leq_k$ ). Entretanto, em (VITÓRIA; SZAŁAS; MAŁUSZYŃSKI, 2008), a construção da semântica é levemente diferente da construção empregada por Belnap. A mudança é motivada por alguns resultados apontados em (MAŁUSZYŃSKI; SZAŁAS; VITÓRIA, 2007). Para dar maiores detalhes, vamos considerar o seguinte exemplo envolvendo um teste com carros:

*Exemplo 15.* (MAŁUSZYŃSKI; SZAŁAS; VITÓRIA, 2007) Suponha que nós temos os carros  $a$  e  $b$  pertencentes ao mesmo homem, em que  $Seguro(a) = \mathbf{u}$  e  $Seguro(b) = \mathbf{i}$ , denotando que a segurança do carro  $a$  é desconhecida e a segurança do carro  $b$  é inconsistente. Queremos saber se todos os seus carros são seguros, i.e., temos a consulta  $Seguro(a) \wedge Seguro(b)$ , em que  $\wedge$  denota a operação de conjunção com respeito à ordem  $\leq_t$  na Lógica de Belnap. Temos que  $(\mathbf{u} \wedge \mathbf{i}) = \mathbf{f}$ . Entretanto, nesse caso não sabemos se ambos os carros são seguros pois não temos informação alguma sobre a segurança do carro  $a$ . Em contraste às respostas obtidas na Lógica

$\wedge$	<b>f</b>	<b>u</b>	<b>i</b>	<b>t</b>	$\vee$	<b>f</b>	<b>u</b>	<b>i</b>	<b>t</b>	$\leftrightarrow$	<b>f</b>	<b>u</b>	<b>i</b>	<b>t</b>	$\neg$	
<b>f</b>	<b>u</b>	<b>i</b>	<b>t</b>	<b>f</b>	<b>t</b>	<b>t</b>	<b>t</b>	<b>t</b>	<b>f</b>	<b>t</b>						
<b>u</b>	<b>f</b>	<b>u</b>	<b>u</b>	<b>u</b>	<b>u</b>	<b>u</b>	<b>u</b>	<b>i</b>	<b>t</b>	<b>u</b>	<b>u</b>	<b>u</b>	<b>i</b>	<b>t</b>	<b>u</b>	<b>u</b>
<b>i</b>	<b>f</b>	<b>u</b>	<b>i</b>	<b>i</b>	<b>i</b>	<b>i</b>	<b>i</b>	<b>i</b>	<b>t</b>	<b>i</b>	<b>i</b>	<b>i</b>	<b>i</b>	<b>t</b>	<b>i</b>	<b>i</b>
<b>t</b>	<b>f</b>	<b>u</b>	<b>i</b>	<b>t</b>	<b>f</b>	<b>u</b>	<b>i</b>	<b>t</b>	<b>t</b>	<b>f</b>						

Tabela 5.3: Tabelas verdade para  $\wedge, \vee, \leftrightarrow$  e  $\neg$ 

de Belnap, **u** parece ser a resposta mais intuitiva para esse caso. Similarmente, se nós queremos averiguar se o homem tem um carro seguro, i.e.,  $\text{Seguro}(a) \vee \text{Seguro}(b)$ , em que  $\vee$  denota a operação de disjunção com respeito à ordem  $\leq_t$ , a Lógica de Belnap resultaria no valor verdade **t**. Isso difere da nossa intuição, pois nós sabemos que a segurança do carro  $b$  é contraditória e nós não sabemos nada sobre a segurança do carro  $a$ . Nesse caso, **i** parece ser a resposta mais intuitiva.

Para contornar esses resultados inesperados, a ordem  $\leq_t$  foi redefinida como a menor relação reflexiva e transitiva satisfazendo  $\mathbf{f} < \mathbf{u} < \mathbf{i} < \mathbf{t}$ . Consequentemente, as operações de conjunção e disjunção na ordem de verdade são definidas como o maior limite inferior (GLB) e como o menor limite superior (LUB) nessa nova ordem, respectivamente, i.e.,  $(x \wedge y) = \text{GLB}\{x, y\}$  e  $(x \vee y) = \text{LUB}\{x, y\}$ . As tabelas-verdade para as operações de conjunção e disjunção assim como as operações de negação e implicação podem ser vistas na Tabela 5.3.

A implicação  $\leftrightarrow$  presente na linguagem naturalmente estende a implicação da lógica clássica, que é definida como  $(\neg x \vee y)$ . Consequentemente, a implicação tem a seguinte propriedade:  $(x \leftrightarrow y) \equiv (\neg y \leftrightarrow \neg x)$ , mas ela não satisfaz a propriedade de Modus Ponens se nós assumirmos que  $\{\mathbf{t}, \mathbf{i}\}$  é o conjunto de valores designados. Por exemplo,  $(\mathbf{i} \leftrightarrow \mathbf{f}) \wedge \mathbf{i}$  não resulta em **f**. Uma explicação mais detalhada da motivação para  $\leftrightarrow$  pode ser encontrada em (VITÓRIA; SZALAS; MAŁUSZYŃSKI, 2008). Finalmente, a semântica de  $\forall$  e  $\exists$  é dada por

$$\forall x P(x) = \text{GLB}_{x \in U} \{P(x)\} \text{ e } \exists x P(x) = \text{LUB}_{x \in U} \{P(x)\},$$

em que  $U$  é o conjunto universo e  $P(x)$  denota que  $x$  tem a propriedade  $P$ , que é determinada por um dos quatro valores verdade.

### 5.3.2 Conjuntos Tetravalorados

Apresentamos agora a noção de um conjunto tetravalorado. Dados os conjunto disjuntos  $U$  e  $\neg U$ , em que  $\neg U = \{\neg x \mid x \in U\}$ , um conjunto tetravalorado  $A$  sobre  $U$  é definido como qualquer subconjunto de  $U \cup \neg U$  (a negação utilizada para a representação de um conjunto tetravalorado consiste de apenas um abuso de notação e não trata-se realmente do conectivo da negação anteriormente apresentado). Intuitivamente  $x \in A$  representa que existe uma evidência que  $x$  está em  $A$ , e  $(\neg x) \in A$  representa que existe uma evidência que  $x$  não está em  $A$ . Nós assumimos que  $(\neg(\neg x))$  é equivalente a  $x$ . Nessa abordagem, a pertinência de um elemento em um conjunto é tetravalorada e estende a usual pertinência bivalorada.

**Definição 24 (Pertinência de Conjunto)** A pertinência de conjunto, denotada por  $\bar{\epsilon} : U \times 2^{U \cup \neg U} \rightarrow \{t, f, i, u\}$ , é definida como

$$x\bar{\epsilon}A = \begin{cases} t, & \text{se } x \in A, (\neg x) \notin A, \\ f, & \text{se } x \notin A, (\neg x) \in A, \\ i, & \text{se } x \in A, (\neg x) \in A, \\ u, & \text{se } x \notin A, (\neg x) \notin A. \end{cases}$$

**Definição 25 (Operações de Conjuntos)** O complemento  $\neg A$  de um conjunto tetravalorado  $A$  é definido por  $\neg A = \{\neg x \mid x \in A\}$ . A inclusão de conjuntos tetravalorados é definida por  $X \subseteq Y = \forall x \in U (x\bar{\epsilon}X \leftrightarrow x\bar{\epsilon}Y)$ . A interseção e união de conjuntos tetravalorados são definidas como  $x\bar{\epsilon}(X \cap Y) = (x\bar{\epsilon}X) \wedge (x\bar{\epsilon}Y)$  e  $x\bar{\epsilon}(X \cup Y) = (x\bar{\epsilon}X) \vee (x\bar{\epsilon}Y)$ , respectivamente.

*Exemplo 16.* Sejam  $U = \{a, b, c\}$  e  $Vermelho = \{a, \neg a, b\}$ . Nesse caso,  $b\bar{\epsilon}Vermelho = t$  e  $a\bar{\epsilon}Vermelho = i$ , pois  $a \in Vermelho$  e  $\neg a \in Vermelho$ . Além disso,  $\neg Vermelho = \{a, \neg a, \neg b\}$ .

Uma extensão tetravalorada de *rough sets* é, então, definida por aproximações de conjuntos tetravalorados (VITÓRIA; SZALAS; MAŁUSZYŃSKI, 2008). Primeiro, a relação de equivalência é trocada por uma relação de similaridade, que também é tetravalorada. Uma relação de similaridade estende a relação de indiscernibilidade coletando na mesma classe objetos que não são necessariamente indiscerníveis, mas suficientemente próximos ou similares. Em outras palavras, uma relação de similaridade é construída a partir de uma relação de indiscernibilidade e retirando algumas condições necessárias para indiscernibilidade e possivelmente acrescentando novas condições. Essa construção pode ser realizada de várias maneiras, dando muitas possibilidades para similaridades.

**Definição 26 (Relação de Similaridade Tetravalorada)** Uma relação de similaridade tetravalorada  $\sigma$  é qualquer relação binária tetravalorada em um conjunto universo  $U$  satisfazendo pelo menos a condição de reflexividade, i.e., para qualquer elemento  $x \in U$ ,  $(x, x)\bar{\epsilon}\sigma = t$ . A vizinhança de um elemento  $x$  com relação a  $\sigma$  é dada pelo conjunto tetravalorado

$$\sigma(x) = (x, y)\bar{\epsilon}\sigma.$$

A partir das definições acima podemos construir as aproximações inferior e superior de conjuntos tetravalorados.

**Definição 27 (Aproximação Inferior/Superior)** Seja  $A$  um conjunto tetravalorado. Então, as aproximações inferior e superior de  $A$  com relação a  $\sigma$ , denotadas por  $\underline{A}_\sigma$  e  $\overline{A}^\sigma$ , respectivamente, são definidas como

$$x\bar{\epsilon}\underline{A}_\sigma = \sigma(x) \subseteq A \text{ e } x\bar{\epsilon}\overline{A}^\sigma = \exists y \in U [y\bar{\epsilon}(\sigma(x) \cap A)].$$

A pertinência de um elemento  $x$  na aproximação inferior de um conjunto tetraavaliado  $A$  é determinada pela pertinência da inclusão de sua vizinhança  $\sigma(x)$  em  $A$ . A pertinência de  $x$  na aproximação superior de um conjunto tetraavaliado  $A$  é determinada pela maior pertinência que um vizinho de  $x$  pode possuir na interseção de  $\sigma(x)$  e  $A$ .

*Exemplo 17.* (VITÓRIA; SZALAS; MAŁUSZYŃSKI, 2008) O universo  $U = \{a, b, c, d\}$  (Tabela 5.2) consiste de objetos que serão classificados como sendo perigosos ou não. O conjunto tetraavaliado Perigo representa essa propriedade. Note que essa classificação pode ser incompleta em alguns casos e inconsistente em outros. Por exemplo, para o objeto  $d$ , não existe informação sobre sua periculosidade ( $d \in \text{Perigo} = \mathbf{u}$ ) enquanto que para o objeto  $c$  existem evidências contraditórias sobre se ele é perigoso ( $c \in \text{Perigo} = \mathbf{i}$ ). Suponha que temos quatro relações de similaridades,  $\sigma_1$  à  $\sigma_4$ , modelando similaridades entre os indivíduos. Cada relação de similaridade pode ser vista como uma avaliação obtida de uma fonte distinta de informação. A pertinência de  $a$  na aproximação inferior de Perigo para a relação  $\sigma_1$  é obtida por

$$\begin{aligned}
a \in \underline{\text{Perigo}}_{\sigma_1} &= \sigma_1(a) \in \text{Perigo} \\
&= \forall y \in U (y \in \sigma_1(a) \leftrightarrow y \in \text{Perigo}) \\
&= \text{GLB}_{y \in U} \{(a, y) \in \sigma_1 \leftrightarrow y \in \text{Perigo}\} \\
&= \text{GLB}\{\mathbf{t} \leftrightarrow \mathbf{t}, \mathbf{i} \leftrightarrow \mathbf{f}, \mathbf{u} \leftrightarrow \mathbf{i}, \mathbf{f} \leftrightarrow \mathbf{u}\} \\
&= \text{GLB}\{\mathbf{t}, \mathbf{i}, \mathbf{i}, \mathbf{t}\} = \mathbf{i}.
\end{aligned}$$

O indivíduo  $a$  possui pertinência verdadeira com relação ao conjunto Perigo, mas sua pertinência na aproximação inferior de Perigo com relação a  $\sigma_1$  é inconsistente devido a algumas incertezas presentes em  $\sigma_1$ . A pertinência de  $a$  na aproximação superior de Perigo para a relação  $\sigma_1$  é obtida por

$$\begin{aligned}
a \in \overline{\text{Perigo}}^{\sigma_1} &= \exists y \in U [y \in (\sigma_1(a) \cap \text{Perigo})] \\
&= \text{LUB}_{y \in U} \{(a, y) \in \sigma_1 \wedge y \in \text{Perigo}\} \\
&= \text{LUB}\{\mathbf{t} \wedge \mathbf{t}, \mathbf{i} \wedge \mathbf{f}, \mathbf{u} \wedge \mathbf{i}, \mathbf{f} \wedge \mathbf{u}\} \\
&= \text{LUB}\{\mathbf{t}, \mathbf{f}, \mathbf{u}, \mathbf{f}\} = \mathbf{t}.
\end{aligned}$$

A pertinência de  $a$  na aproximação superior de Perigo para a relação  $\sigma_1$  é verdadeira. De fato, como a pertinência de  $a$  em Perigo é verdadeira, a pertinência de sua aproximação superior também será, pois as relações de similaridade são definidas como sendo pelo menos reflexivas. Logo, ao aplicarmos a aproximação superior,  $a$  será similar a pelo menos um indivíduo (ele mesmo) que tem pertinência verdadeira ao conjunto Perigo.

	$\sigma_1(a)$	$\sigma_2(a)$	$\sigma_3(a)$	$\sigma_4(a)$	Perigo
$a$	<b>t</b>	<b>t</b>	<b>t</b>	<b>t</b>	<b>t</b>
$b$	<b>i</b>	<b>f</b>	<b>t</b>	<b>u</b>	<b>f</b>
$c$	<b>u</b>	<b>f</b>	<b>u</b>	<b>f</b>	<b>i</b>
$d$	<b>f</b>	<b>f</b>	<b>t</b>	<b>t</b>	<b>u</b>

Tabela 5.4: Um exemplo de uma tabela de dados incompleta

#### 5.4 Lógica de Descrição *Rough* Paraconsistente $\mathcal{ALC}$

Como uma contribuição dessa dissertação, nesta seção nós estendemos os formalismos das RDLs para *rough sets* paraconsistentes, introduzindo uma RDL tetravalorada geral o suficiente para abranger dois tipos de relações de similaridades: aquelas explicitamente definidas e aquelas definidas em termos de um contexto. Um aspecto distintivo do nosso formalismo é que ele permite sucessivos refinamentos de uma consulta. Na sequência, nós introduzimos a sintaxe, semântica e tarefas de raciocínio para essa lógica, chamada de  $PR_{\mathcal{ALC}}$ .

**Definição 28 (Alfabeto)** *O alfabeto de  $PR_{\mathcal{ALC}}$  é composto pelos símbolos  $\neg, \sqcap, \sqcup, \exists, \forall, \top, \perp, \underline{\cdot}, \overline{\cdot}$  e quatro conjuntos disjuntos: o conjunto de indivíduos  $N_I$ , o conjunto de nomes de conceitos  $N_C$ , o conjunto de nomes de papéis  $N_R$  e o conjunto de relações de similaridades  $N_S$ .*

A construção dos conceitos em  $PR_{\mathcal{ALC}}$  é semelhante à construção dos conceitos na *Rough  $\mathcal{ALC}_4$* .

**Definição 29 (Conceitos)** *Conceitos em  $PR_{\mathcal{ALC}}$  são definidos pelas regras de sintaxe abaixo, em que  $C$  e  $D$  são conceitos,  $A$  é um conceito atômico,  $R$  é um papel atômico e  $S$  é uma relação de similaridade:*

$$C, D \longrightarrow A \mid \top \mid \perp \mid \neg A \mid C \sqcap D \mid C \sqcup D \mid \exists R.C \mid \forall R.C \mid \overline{C}^S \mid \underline{C}_S.$$

A diferença fica por conta dos conceitos  $\overline{C}^S$  e  $\underline{C}_S$ , que agora são aproximados por uma relação de similaridade  $S$  (substituindo a relação de indiscernibilidade). A semântica dos conceitos em  $PR_{\mathcal{ALC}}$  também difere da semântica da *Rough  $\mathcal{ALC}$* . Uma interpretação é definida de acordo com conjuntos tetravalorados e suas semânticas tetravaloradas.

**Definição 30 (Semântica)** *A semântica de indivíduos, conceitos atômicos, papéis atômicos e relações de similaridade em  $PR_{\mathcal{ALC}}$  é dada por uma interpretação  $I = (\Delta^I, \cdot^I)$  em que o domínio  $\Delta^I$  é um conjunto não-vazio de elementos e  $\cdot^I$  é uma função de mapeamento definida como segue: cada indivíduo  $a \in N_I$  é mapeado para  $a^I \in \Delta^I$ ; cada conceito atômico  $A \in N_C$  é mapeado para  $A^I : \Delta^I \rightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{i}, \mathbf{u}\}$ ; cada papel atômico  $R \in N_R$  é mapeado para  $R^I : \Delta^I \times \Delta^I \rightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{i}, \mathbf{u}\}$ ; cada relação de similaridade  $S \in N_S$  é mapeada para  $S^I : \Delta^I \times \Delta^I \rightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{i}, \mathbf{u}\}$  satisfazendo pelo menos a condição de reflexividade, i.e.,  $S^I(x, x) = \mathbf{t}$  para qualquer  $x \in \Delta^I$ . Para os demais conceitos, em que  $x \in \Delta^I$ :*

- $\top^I(x) = \mathbf{t}$ ;
- $\perp^I(x) = \mathbf{f}$ ;
- $(\neg A)^I(x) = \neg(A^I(x))$ ;
- $(C \sqcap D)^I(x) = (C^I(x) \wedge D^I(x))$ ;
- $(C \sqcup D)^I(x) = (C^I(x) \vee D^I(x))$ ;
- $(\exists R.C)^I(x) = \text{LUB}_{y \in \Delta^I}(R^I(x, y) \wedge C^I(y))$ ;
- $(\forall R.C)^I(x) = \text{GLB}_{y \in \Delta^I}(R^I(x, y) \leftrightarrow C^I(y))$ ;
- $(\overline{C}^S)^I(x) = \text{LUB}_{y \in \Delta^I}(S^I(x, y) \wedge C^I(y))$ ;
- $(\underline{C}_S)^I(x) = \text{GLB}_{y \in \Delta^I}(S^I(x, y) \leftrightarrow C^I(y))$ .

Um fato interessante é que a semântica dos conceitos é baseada na semântica de DLs *Fuzzy* (BOBILLO; STRACCIA, 2009) e de *fuzzy rough sets* (COCK; CORNELIS; KERRE, 2004). Podemos ver as operações de conjunção ( $\wedge$ ), disjunção ( $\vee$ ), implicação ( $\leftrightarrow$ ) e negação ( $\neg$ ) como as funções de t-normas, t-conormas, funções de implicação e negação das abordagens de *fuzzy sets*. A diferença é que em nossa abordagem, a semântica dessas operações são fundamentadas nas operações tetravaloradas de (VITÓRIA; SZALAS; MAŁUSZYŃSKI, 2008). Agora definiremos as noções de TBox, ABox e base de conhecimento em  $PR_{\mathcal{ALC}}$ .

**Definição 31 (TBox)** *Um TBox é um conjunto finito de expressões das formas  $C \sqsubseteq D$  ou  $C \equiv D$ , chamados de axiomas de inclusão e equivalência de conceitos, respectivamente, em que  $C$  e  $D$  são conceitos em  $PR_{\mathcal{ALC}}$ .*

**Definição 32 (ABox)** *Um ABox consiste de um conjunto finito de axiomas de asserção da forma  $C(a)$ ,  $R(a, b)$  e  $S(a, b)$ , em que  $C$  é um conceito,  $R$  é um papel,  $S$  é uma relação de similaridade e  $a$  e  $b$  são indivíduos em  $PR_{\mathcal{ALC}}$ .*

Quando queremos referenciar os axiomas de inclusão e asserção indistintamente, nós os chamaremos apenas de axiomas. A semântica dos axiomas é dada por:

**Definição 33 (Satisfatibilidade)** *A noção de satisfatibilidade de um axioma  $\alpha$  por uma interpretação  $I = (\Delta^I, \cdot^I)$ , denotada por  $I \models \alpha$ , é definida a seguir:*

- $I \models C \sqsubseteq D$  sse  $i \leq \text{GLB}_{x \in \Delta^I}(C^I(x) \leftrightarrow D^I(x))$ ;
- $I \models C \equiv D$  sse  $i \leq \text{GLB}_{x \in \Delta^I}(C^I(x) \leftrightarrow D^I(x))$  e  $i \leq \text{GLB}_{x \in \Delta^I}(D^I(x) \leftrightarrow C^I(x))$ ;

- $I \models C(a)$  sse  $i \leq C^I(a)$ ;
- $I \models R(a, b)$  sse  $i \leq R^I(a, b)$ ;
- $I \models S(a, b)$  sse  $i \leq S^I(a, b)$ .

Note que a semântica da inclusão de conceitos  $C \sqsubseteq D$  é derivada das operações tetravaloradas apresentadas anteriormente. Nós assumimos que  $\{\mathbf{i}, \mathbf{t}\}$  é o conjunto de valores designados, portanto um axioma  $\alpha$  é satisfeito por uma interpretação  $I$  se e somente se  $I$  avalia  $\alpha$  para os valores verdades  $\mathbf{t}$  ou  $\mathbf{i}$  (os valores contidos no conjunto de valores designados).

**Definição 34 (Base de Conhecimento)** *Uma base de conhecimento ou ontologia  $\mathcal{K}$  é um conjunto composto por um TBox  $\mathcal{T}$  e um ABox  $\mathcal{A}$ .*

Para um axioma  $\alpha$ , se  $I \models \alpha$  dizemos que  $I$  é um modelo de  $\alpha$ . Um TBox  $\mathcal{T}$  é satisfeito por uma interpretação  $I$ , ou seja,  $I \models \mathcal{T}$  se e somente se para todo  $C \sqsubseteq D$  e  $C \equiv D$  em  $\mathcal{T}$ ,  $I \models C \sqsubseteq D$  e  $I \models C \equiv D$ . Um ABox  $\mathcal{A}$  é satisfeito por uma interpretação  $I$ , ou seja,  $I \models \mathcal{A}$  se e somente se para toda asserção da forma  $C(a)$ ,  $R(a, b)$  e  $S(a, b)$  em  $\mathcal{A}$ ,  $I \models C(a)$ ,  $I \models R(a, b)$  e  $I \models S(a, b)$ .  $I$  satisfaz (é um modelo de) uma base de conhecimento  $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$ , denotado por  $I \models \mathcal{K}$ , se e somente se  $I \models \mathcal{T}$  e  $I \models \mathcal{A}$ .

**Definição 35 (Consequência Lógica)** *Um axioma  $\alpha$  é uma consequência lógica de uma base de conhecimento  $\mathcal{K}$ , denotada por  $\mathcal{K} \models \alpha$ , se e somente se todo modelo de  $\mathcal{K}$  satisfaz  $\alpha$ .*

### 5.4.1 Aproximação Contextual

Nesta seção, aplicaremos as noções de aproximação contextual em  $PR_{\mathcal{ALC}}$ , estendendo as definições de aproximações inferior e superior de um conceito. Introduziremos duas relações de similaridades específicas baseadas em contextos, e finalmente definiremos as aproximações inferior e superior usando essas novas similaridades. Destacamos que todos as definições apresentadas nesta seção podem ser também aplicadas à *Rough  $\mathcal{ALC}_4$* . A primeira mudança será na definição de contexto apresentado no Capítulo 3.

**Definição 36 (Contexto)** *Um contexto é um conjunto finito não vazio de conceitos atômicos em  $DL \Sigma = \{A_1, \dots, A_n\}$ .*

Por questões de complexidade, utilizaremos apenas conceitos atômicos em nosso método de refinamento de consultas em  $PR_{\mathcal{ALC}}$ . Como trabalho futuro, pretendemos analisar a utilização de conceitos complexos dentro de um contexto. A segunda mudança será na definição de projeção, também apresentada no Capítulo 3.

**Definição 37 (Projeção)** *Sejam  $\mathcal{K}$  uma base de conhecimento e  $A$  um conceito atômico, a função de projeção  $\pi_A^{\mathcal{K}} : N_C \rightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{i}, \mathbf{u}\}$  é definida como*

$$\forall c \in N_C : \pi_A^{\mathcal{K}}(c) = \begin{cases} \mathbf{t}, & \mathcal{K} \models A(c) \text{ e } \mathcal{K} \not\models \neg A(c), \\ \mathbf{f}, & \mathcal{K} \not\models A(c) \text{ e } \mathcal{K} \models \neg A(c), \\ \mathbf{i}, & \mathcal{K} \models A(c) \text{ e } \mathcal{K} \models \neg A(c), \\ \mathbf{u}, & \text{caso contrário.} \end{cases}$$

A motivação para essa mudança segue do princípio que iremos trabalhar mais adiante com a noção de consequência lógica durante o processo de refinamento de consultas ( $\mathcal{K} \models \alpha$ ). Assim, a relação de indiscernibilidade também será redefinida com essa nova projeção.

**Definição 38 (Relação de Indiscernibilidade)** *Seja  $\Sigma = \{A_1, \dots, A_n\}$  um contexto. A relação de indiscernibilidade  $R_\Sigma$  induzida por  $\Sigma$  é definida como segue:*

$$R_\Sigma = \{(a, b) \in N_C \times N_C \mid \text{para todo } A_i \text{ em que } i \in \{1, \dots, n\}, \pi_{A_i}^{\mathcal{K}}(a) = \pi_{A_i}^{\mathcal{K}}(b)\}.$$

Como também estamos lidando com informação incompleta, uma relação de similaridade deveria ser mais adequada para modelar relacionamentos entre indivíduos, pois ela permite agrupar indivíduos que estão próximos, mas não necessariamente que sejam indiscerníveis. Por exemplo, suponha que temos dois indivíduos  $a$  e  $b$ , e um conceito  $A$ . Um problema que poderia surgir com apenas a presença da relação de indiscernibilidade é que se a projeção de  $a$  de acordo com  $A$  é verdadeira e a projeção de  $b$  de acordo com  $A$  é desconhecida, teríamos que os dois indivíduos são discerníveis. Mas isso não é possível afirmar, pois não temos informação suficiente sobre  $b$ . Nós introduziremos a seguinte relação de similaridade, que é baseada no trabalho apresentado em (GRZYMALA-BUSSE, 2006) e que de certa forma afrouxa a condição original da indiscernibilidade:

**Definição 39 (Relação de Similaridade - Conceitos Desconhecidos)** *Seja  $\Sigma = \{A_1, \dots, A_n\}$  um contexto. A relação de similaridade  $S_\Sigma$  induzida por  $\Sigma$  é definida como segue:*

$$S_\Sigma = \{(a, b) \in N_C \times N_C \mid \text{para todo } A_i \text{ em que } i \in \{1, \dots, n\}, \pi_{A_i}^{\mathcal{K}}(a) = \pi_{A_i}^{\mathcal{K}}(b) \text{ ou } \pi_{A_i}^{\mathcal{K}}(a) = \mathbf{u} \text{ ou } \pi_{A_i}^{\mathcal{K}}(b) = \mathbf{u}\}.$$

O propósito da relação de similaridade  $S_\Sigma$  é aproximar informação incompleta considerando o valor verdade  $\mathbf{u}$  como sendo similar a qualquer outro valor verdade e vice-versa. Em  $S_\Sigma$ , a informação relevante é apenas aquela que é afirmada, i.e., um indivíduo avaliado para  $\mathbf{u}$  em um conceito é considerado como não-relevante. Portanto, um elemento do domínio  $a$  é similar a  $b$  em um contexto  $\Sigma$  se para todos conceitos em  $\Sigma$ , a interpretação de  $a$  e  $b$  são iguais, ou a interpretação de algum deles é igual a  $\mathbf{u}$ .

Com o objetivo de aproximar ambas informações contraditória e incompleta, introduziremos uma segunda relação de similaridade, denotada por  $P_\Sigma$ . Nessa nova relação de similaridade, os valores verdades  $\mathbf{t}$  e  $\mathbf{f}$  são considerados similares a  $\mathbf{i}$ .

**Definição 40 (Relação de Similaridade - Conceitos Desconhecidos e Inconsistentes)** *Seja  $\Sigma = \{A_1, \dots, A_n\}$  um contexto. A relação de similaridade  $P_\Sigma$  induzida por  $\Sigma$  é definida como segue:*

$$P_\Sigma = \{(a, b) \in N_C \times N_C \mid \text{para todo } A_i \text{ em que } i \in \{1, \dots, n\}, \pi_{A_i}^{\mathcal{K}}(a) = \pi_{A_i}^{\mathcal{K}}(b) \text{ ou}$$

$$\pi_{A_i}^{\mathcal{K}}(a) = \mathbf{u} \text{ ou } \pi_{A_i}^{\mathcal{K}}(b) = \mathbf{u}$$

$$\text{ou se } \pi_{A_i}^{\mathcal{K}}(a) = \mathbf{t} \text{ então } \pi_{A_i}^{\mathcal{K}}(b) = \mathbf{i}$$

$$\text{ou se } \pi_{A_i}^{\mathcal{K}}(a) = \mathbf{f} \text{ então } \pi_{A_i}^{\mathcal{K}}(b) = \mathbf{i}\}.$$

Em  $P_\Sigma$ , cuja definição foi baseada em uma relação de similaridade apresentada em (GRZYMALA-BUSSE, 2006), é assumido que a informação pode ser parcialmente descrita por causa de nosso conhecimento incompleto ou contraditório. Desse ponto de vista, um elemento do domínio  $a$  pode ser considerado similar a  $b$  se a informação obtida em  $a$  é também obtida em  $b$ . Logo, para um conceito  $A$ , em que  $\pi_A^{\mathcal{K}}(a) = \mathbf{t}$  e  $\pi_A^{\mathcal{K}}(b) = \mathbf{i}$ , o indivíduo  $a$  é similar a  $b$  pois o valor verdade  $\mathbf{t}$  está contido em  $\mathbf{i}$ . Note que o contrário não é verdade:  $b$  não é similar a  $a$  de acordo com  $P_\Sigma$ , pois nem toda informação de  $b$  está contida em  $a$ . Destacamos que as relações de similaridades introduzidas neste trabalho são bivaloradas, mas nada impede que elas sejam tetravaloradas.

As aproximações contextuais inferior e superior de um conceito  $C$  com relação ao contexto  $\Sigma$  são denotadas por  $\underline{C}_{\text{Sim}_\Sigma}$  e  $\overline{C}^{\text{Sim}_\Sigma}$ , respectivamente, em que  $\text{Sim}$  é uma R, S ou P. É importante enfatizar que para aproximações com essas relações mostradas anteriormente, nós temos os seguintes resultados:

**Proposição 3** *Dados um conceito  $C$ , um contexto  $\Sigma$ , um indivíduo  $x$  e uma interpretação  $I$ , é válido que:*

$$(\underline{C}_{P_\Sigma})^I(x) \leq (\underline{C}_{S_\Sigma})^I(x) \leq (\underline{C}_{R_\Sigma})^I(x) \leq C^I(x) \leq (\overline{C}^{R_\Sigma})^I(x) \leq (\overline{C}^{S_\Sigma})^I(x) \leq (\overline{C}^{P_\Sigma})^I(x).$$

*Prova.* Vamos mostrar primeiramente que  $(\underline{C}_{S_\Sigma})^I(x) \leq (\underline{C}_{R_\Sigma})^I(x)$ . Os outros casos para a aproximação inferior são similares. Temos que  $R_\Sigma(x) \subseteq S_\Sigma(x)$ , em que  $R_\Sigma(x) = \{y \in \Delta^I \mid (x, y) \in R_\Sigma\}$  e  $S_\Sigma(x) = \{y \in \Delta^I \mid (x, y) \in S_\Sigma\}$ , pois  $S_\Sigma$  é mais geral que  $R_\Sigma$ . Como consequência,  $R_\Sigma^I(x, y) \leq S_\Sigma^I(x, y)$ . Como  $R_\Sigma$  e  $S_\Sigma$  são duas relações bivaloradas (suas interpretações mapeiam apenas para  $\mathbf{t}$  ou  $\mathbf{f}$ ), Temos os seguintes casos a analisar:

- $S_\Sigma^I(x, y) = \mathbf{t}$  e  $R_\Sigma^I(x, y) = \mathbf{t}$ . Temos então que  $R_\Sigma^I(x, y) \leftrightarrow C^I(y) = S_\Sigma^I(x, y) \leftrightarrow C^I(y)$ .
- $S_\Sigma^I(x, y) = \mathbf{t}$  e  $R_\Sigma^I(x, y) = \mathbf{f}$ . Temos então que  $R_\Sigma^I(x, y) \leftrightarrow C^I(y) \geq S_\Sigma^I(x, y) \leftrightarrow C^I(y)$ .

$$\text{Logo, } (\underline{C}_{R_\Sigma})^I(x) = \text{GLB}_{y \in \Delta^I}(R_\Sigma^I(x, y) \leftrightarrow C^I(y)) \geq \text{GLB}_{y \in \Delta^I}(S_\Sigma^I(x, y) \leftrightarrow C^I(y)) = (\underline{C}_{S_\Sigma})^I(x).$$

Para os casos relativos à aproximação superior, iremos mostrar apenas a prova de  $(\overline{C}^{R_\Sigma})^I(x) \leq (\overline{C}^{S_\Sigma})^I(x)$ . Os demais casos também são similares. Temos os seguintes casos a analisar:

- $S_{\Sigma}^I(x, y) = \mathbf{t}$  e  $R_{\Sigma}^I(x, y) = \mathbf{t}$ . Temos então que  $R_{\Sigma}^I(x, y) \wedge C^I(y) = S_{\Sigma}^I(x, y) \wedge C^I(y)$ .
- $S_{\Sigma}^I(x, y) = \mathbf{t}$  e  $R_{\Sigma}^I(x, y) = \mathbf{f}$ . Temos então que  $R_{\Sigma}^I(x, y) \wedge C^I(y) \leq S_{\Sigma}^I(x, y) \wedge C^I(y)$ .

Logo,  $(\overline{C}^{\mathbf{R}\Sigma})^I(x) = \text{LUB}_{y \in \Delta^I}(R_{\Sigma}^I(x, y) \wedge C^I(y)) \leq \text{LUB}_{y \in \Delta^I}(S_{\Sigma}^I(x, y) \wedge C^I(y)) = (\overline{C}^{\mathbf{S}\Sigma})^I(x)$ .  $\square$

A proposição mostra que se aplicarmos, por exemplo, uma aproximação superior com respeito à relação de indiscernibilidade e ela não for satisfeita, podemos tentar aplicar a aproximação com respeito às relações de similaridades, pois elas são mais gerais que a indiscernibilidade e consequentemente sua satisfação poderá ser diferente.

#### 5.4.2 Refinamento de Consultas

Como ressaltado anteriormente, a teoria dos *rough sets* é uma ferramenta candidata a ser aplicada em refinamento de consultas, pois suas aproximações dão a ideia de respostas possíveis ou necessárias. A motivação é que podemos imaginar uma situação em que uma consulta resulta em uma resposta vazia ou pequena; nesse caso, sua aproximação superior pode ser aplicada para possivelmente produzir algum elemento do domínio relacionado à consulta. Por outro lado, uma resposta a uma consulta pode resultar em muitos indivíduos; então a aproximação inferior poderia ser aplicada para possivelmente restringir o número de indivíduos relacionados à consulta, resultando apenas naqueles mais relevantes.

*Exemplo 18.* Considere o conjunto  $\Delta^I = \{x_1, x_2, x_3, x_4\}$ , o conceito  $C$  e suas interpretações  $C^I(x_1) = \mathbf{t}$ ,  $C^I(x_2) = \mathbf{u}$ ,  $C^I(x_3) = \mathbf{t}$ ,  $C^I(x_4) = \mathbf{u}$  e a relação de similaridade  $S$  em  $\Delta^I$  definida por

$$S^I = \{(x_1, x_1), (x_1, x_2), (x_1, x_4), (x_2, x_2), (x_2, x_3), (x_2, x_4), (x_3, x_1), (x_3, x_3), (x_4, x_2), (x_4, x_4)\}.$$

O indivíduo  $x_1$  está possivelmente, mas não necessariamente presente no conceito  $C$  com relação a  $S$ , isto é,  $(\underline{C}_S)^I(x_1) = \mathbf{u}$  e  $(\overline{C}^S)^I(x_1) = \mathbf{t}$ . Nesse caso, a aproximação inferior parece muito restrita (o resultado da aproximação não corresponde a um dos valores designados). Podemos expandir essa consulta com uma aproximação superior e descobrir que  $(\overline{C}_S^S)^I(x_1) = \mathbf{t}$ , ou seja,  $x_1$  possui ainda um certo grau de necessidade com o conceito  $C$ . Da mesma forma, podemos restringir a aproximação superior e descobrir que ela possui um grau de relevância, pois  $(\underline{C}_S^S)^I(x_1) = \mathbf{t}$ .

Nessa seção, nós combinamos a flexibilidade da aproximação superior com o rigor da aproximação inferior aplicando-as alternativamente ou sucessivamente (COCK; CORNELIS, 2005; COCK; CORNELIS; KERRE, 2004). Essas estratégias serão definidas em  $PR_{\mathcal{A}LC}$  na sequência:

**Definição 41 (Aproximações Inferior/Superior tight e loose)** *As aproximações tight e loose das aproximações inferior e superior de um conceito  $C$  com respeito à relação de similaridade  $\text{Sim}$  são definidas como*

- $(\overline{C}_{\text{Sim}}^{\text{Sim}})^I(x) = \text{GLB}_{y \in \Delta^I}(\text{Sim}^I(x, y) \leftrightarrow \text{LUB}_{z \in \Delta^I}(\text{Sim}^I(y, z) \wedge C^I(z)));$
- $(\overline{C}^{\text{Sim}})^I(x) = \text{LUB}_{y \in \Delta^I}(\text{Sim}^I(x, y) \wedge \text{LUB}_{z \in \Delta^I}(\text{Sim}^I(y, z) \wedge C^I(z)));$
- $(\underline{C}_{\text{Sim}}^{\text{Sim}})^I(x) = \text{GLB}_{y \in \Delta^I}(\text{Sim}^I(x, y) \leftrightarrow \text{GLB}_{z \in \Delta^I}(\text{Sim}^I(y, z) \leftrightarrow C^I(z)));$
- $(\underline{C}_{\text{Sim}}^{\text{Sim}})^I(x) = \text{LUB}_{y \in \Delta^I}(\text{Sim}^I(x, y) \wedge \text{GLB}_{z \in \Delta^I}(\text{Sim}^I(y, z) \leftrightarrow C^I(z))).$

Podemos generalizar as definições de aplicações sucessivas de aproximação inferior e superior de conceitos, utilizando a noção de *relação-passo-k* (WU; ZHANG, 2002).

**Definição 42 (Relação-passo-k)** *Sejam  $\Delta^I$  o conjunto universo não-vazio,  $S$  uma relação binária em  $\Delta^I$  e  $k$  um número natural. A relação-passo-k de  $S$ , denotada por  $S_k$ , é definida como:*

- $S_1 = S;$
- $S_{k+1} = S_k \cup \{(x, y) \in \Delta^I \times \Delta^I \mid \text{existe } y_1, y_2, \dots, y_k \in \Delta^I, \text{ tal que } xSy_1, y_1Sy_2, \dots, y_kSy\},$   
para  $k \geq 1.$

*Exemplo 19.* Sejam  $U = \{1, 2, 3\}$  e  $S = \{(1, 2), (2, 3), (3, 1)\}$ . Podemos checar que:

$$S_2 = \{(1, 2), (1, 3), (2, 1), (2, 3), (3, 1), (3, 2)\},$$

$$S_3 = \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (3, 3)\}.$$

É fácil ver que  $S_k \subseteq S_{k+1}$ . De fato,  $S_n$  é fecho transitivo de grau  $n$  da relação  $S$ . Nós chamamos uma relação de passo maior de  $S_k$  qualquer relação  $S_j$ , em que  $j > k$ . A semântica de aplicações sucessivas das aproximações pode estar então relacionada com uma *relação-passo-k*:

**Definição 43 (Aproximações loose e tight Generalizadas)** *A aproximação superior loose e a aproximação inferior tight de um conceito  $C$  com respeito à relação de similaridade  $\text{Sim}$  em  $n$  passos são denotadas por  $\overline{C}^{(\text{Sim}, n)}$  e  $\underline{C}_{(\text{Sim}, n)}$ , respectivamente, e definidas como*

- $(\overline{C}^{(\text{Sim}, n)})^I(x) = \text{LUB}_{y \in \Delta^I}(\text{Sim}_n^I(x, y) \wedge C^I(y));$
- $(\underline{C}_{(\text{Sim}, n)})^I(x) = \text{GLB}_{y \in \Delta^I}(\text{Sim}_n^I(x, y) \leftrightarrow C^I(y)).$

**Proposição 4** *Dados um conceito  $C$ , uma relação de similaridade  $\text{Sim}_\Sigma$ , um indivíduo  $x$  e uma interpretação  $I$ , é válido que*

$$\bullet (\underline{C}_{\text{Sim}_\Sigma \text{Sim}_\Sigma})^I(x) \leq (\underline{C}_{\text{Sim}_\Sigma})^I(x) \leq C^I(x) \leq (\overline{C}^{\text{Sim}_\Sigma})^I(x) \leq (\overline{C}^{\text{Sim}_\Sigma \text{Sim}_\Sigma})^I(x);$$

- $(\underline{C}_{\text{Sim}_\Sigma})^I(x) \leq (\overline{\underline{C}_{\text{Sim}_\Sigma}}^{\text{Sim}_\Sigma})^I(x) \leq (\overline{C}^{\text{Sim}_\Sigma})^I(x)$ ;
- $(\underline{C}_{\text{Sim}_\Sigma})^I(x) \leq (\overline{\underline{C}_{\text{Sim}_\Sigma}}^{\text{Sim}_\Sigma})^I(x) \leq (\overline{C}^{\text{Sim}_\Sigma})^I(x)$ .

*Prova.* 1.  $(\underline{C}_{\text{Sim}_\Sigma \text{Sim}_\Sigma})^I(x) \leq (\underline{C}_{\text{Sim}_\Sigma})^I(x) \leq C^I(x) \leq (\overline{C}^{\text{Sim}_\Sigma})^I(x) \leq (\overline{\underline{C}_{\text{Sim}_\Sigma}}^{\text{Sim}_\Sigma})^I(x)$ . Vamos mostrar o caso  $(\underline{C}_{\text{Sim}_\Sigma \text{Sim}_\Sigma})^I(x) \leq (\underline{C}_{\text{Sim}_\Sigma})^I(x)$ . Os outros casos são similares. Como  $\text{Sim}_\Sigma$  é uma relação bivalorada, temos os seguintes casos a analisar:

- $(\underline{C}_{\text{Sim}_\Sigma})^I(x) = \mathbf{t}$ . Trivialmente podemos ver que  $(\underline{C}_{\text{Sim}_\Sigma \text{Sim}_\Sigma})^I(x) \leq \mathbf{t}$ .
- $(\underline{C}_{\text{Sim}_\Sigma})^I(x) = \mathbf{i}$ . Por definição,  $\text{GLB}_{y \in \Delta^I}(\text{Sim}_\Sigma^I(x, y) \leftrightarrow C^I(y)) = \mathbf{i}$ . Pela semântica de  $\leftrightarrow$  e GLB, podemos dizer que existe um  $w \in \Delta^I$  no qual  $\text{Sim}_\Sigma^I(x, w) = \mathbf{t}$  e  $C^I(w) = \mathbf{i}$ , que é o menor valor para  $\text{GLB}_{y \in \Delta^I}(\text{Sim}_\Sigma^I(x, y) \leftrightarrow C^I(y))$ . Consequentemente,  $\text{GLB}_{z \in \Delta^I}(\text{Sim}_\Sigma^I(w, z) \leftrightarrow C^I(z)) \leq \mathbf{i}$ , pois para  $z = w$ ,  $\text{Sim}_\Sigma^I(w, w) = \mathbf{t}$  e  $C^I(w) = \mathbf{i}$ . Logo,  $\text{GLB}_{y \in \Delta^I}(\text{Sim}_\Sigma^I(x, y) \leftrightarrow \text{GLB}_{z \in \Delta^I}(\text{Sim}_\Sigma^I(y, z) \leftrightarrow C^I(z))) \leq \mathbf{i}$  (fazendo  $y = z = w$ ). Portanto,  $(\underline{C}_{\text{Sim}_\Sigma \text{Sim}_\Sigma})^I(x) \leq \mathbf{i}$ .
- $(\underline{C}_{\text{Sim}_\Sigma})^I(x) = \mathbf{u}$  e  $(\underline{C}_{\text{Sim}_\Sigma})^I(x) = \mathbf{f}$ . São similares ao caso  $(\underline{C}_{\text{Sim}_\Sigma})^I(x) = \mathbf{i}$ .

2.  $(\underline{C}_{\text{Sim}_\Sigma})^I(x) \leq (\overline{\underline{C}_{\text{Sim}_\Sigma}}^{\text{Sim}_\Sigma})^I(x) \leq (\overline{C}^{\text{Sim}_\Sigma})^I(x)$ . O caso  $(\overline{\underline{C}_{\text{Sim}_\Sigma}}^{\text{Sim}_\Sigma})^I(x) \leq (\overline{C}^{\text{Sim}_\Sigma})^I(x)$  é similar à prova do caso 1. Vamos mostrar o caso  $(\underline{C}_{\text{Sim}_\Sigma})^I(x) \leq (\overline{\underline{C}_{\text{Sim}_\Sigma}}^{\text{Sim}_\Sigma})^I(x)$ . Para isso, temos os seguintes casos a analisar:

- $(\overline{\underline{C}_{\text{Sim}_\Sigma}}^{\text{Sim}_\Sigma})^I(x) = \mathbf{t}$ . Trivialmente podemos ver que  $(\underline{C}_{\text{Sim}_\Sigma})^I(x) \leq \mathbf{t}$ .
- $(\overline{\underline{C}_{\text{Sim}_\Sigma}}^{\text{Sim}_\Sigma})^I(x) = \mathbf{i}$ . Por definição,  $\text{LUB}_{y \in \Delta^I}(\text{Sim}_\Sigma^I(x, y) \wedge \text{GLB}_{z \in \Delta^I}(\text{Sim}_\Sigma^I(y, z) \leftrightarrow C^I(z))) = \mathbf{i}$ . Pela semântica da interseção  $\wedge$  e de LUB, podemos dizer que existe um  $w \in \Delta^I$  no qual  $\text{Sim}_\Sigma^I(x, w) = \mathbf{t}$  e  $\text{GLB}_{z \in \Delta^I}(\text{Sim}_\Sigma^I(w, z) \leftrightarrow C^I(z)) = \mathbf{i}$ , que garante o maior valor para  $\text{LUB}_{y \in \Delta^I}(\text{Sim}_\Sigma^I(x, y) \wedge \text{GLB}_{z \in \Delta^I}(\text{Sim}_\Sigma^I(y, z) \leftrightarrow C^I(z)))$ . Pela semântica de  $\leftrightarrow$ , existe um  $v \in \Delta^I$  em que  $\text{Sim}_\Sigma^I(w, v) = \mathbf{t}$  e  $C^I(v) = \mathbf{i}$ . Como  $\text{Sim}_\Sigma^I(x, w) = \mathbf{t}$  e  $\text{Sim}_\Sigma^I(w, v) = \mathbf{t}$ , por transitividade  $\text{Sim}_\Sigma^I(x, v) = \mathbf{t}$ . Consequentemente,  $\text{GLB}_{y \in \Delta^I}(\text{Sim}_\Sigma^I(x, y) \leftrightarrow C^I(y)) \leq \mathbf{i}$ , pois para  $y = v$ ,  $\text{Sim}_\Sigma^I(x, v) = \mathbf{t}$  e  $C^I(v) = \mathbf{i}$ . Portanto,  $(\underline{C}_{\text{Sim}_\Sigma})^I(x) \leq \mathbf{i}$ .
- $(\overline{\underline{C}_{\text{Sim}_\Sigma}}^{\text{Sim}_\Sigma})^I(x) = \mathbf{u}$  e  $(\overline{\underline{C}_{\text{Sim}_\Sigma}}^{\text{Sim}_\Sigma})^I(x) = \mathbf{f}$ . São similares ao caso  $(\overline{\underline{C}_{\text{Sim}_\Sigma}}^{\text{Sim}_\Sigma})^I(x) = \mathbf{i}$ .

3.  $(\underline{C}_{\text{Sim}_\Sigma})^I(x) \leq (\overline{\underline{C}_{\text{Sim}_\Sigma}}^{\text{Sim}_\Sigma})^I(x) \leq (\overline{C}^{\text{Sim}_\Sigma})^I(x)$ . Similar à prova do Caso 2.  $\square$

**Proposição 5** *Dados um conceito C, a relação de indiscernibilidade  $R_\Sigma$ , um indivíduo x e uma interpretação I, é válido que*

- $(\underline{C}_{R_\Sigma R_\Sigma})^I(x) = (\underline{C}_{R_\Sigma})^I(x) \leq C^I(x) \leq (\overline{C}^{R_\Sigma})^I(x) = (\overline{\overline{C}^{R_\Sigma}})^I(x)$ ;
- $(\underline{C}_{R_\Sigma})^I(x) = (\overline{\underline{C}_{R_\Sigma}})^I(x) \leq (\overline{C}^{R_\Sigma})^I(x)$ ;
- $(\underline{C}_{R_\Sigma})^I(x) \leq (\overline{\underline{C}_{R_\Sigma}})^I(x) = (\overline{C}^{R_\Sigma})^I(x)$ .

*Prova.* 1.  $(\underline{C}_{R_\Sigma R_\Sigma})^I(x) = (\underline{C}_{R_\Sigma})^I(x) \leq C^I(x) \leq (\overline{C}^{R_\Sigma})^I(x) = (\overline{\overline{C}^{R_\Sigma}})^I(x)$ . Vamos demonstrar o caso  $(\underline{C}_{R_\Sigma R_\Sigma})^I(x) = (\underline{C}_{R_\Sigma})^I(x)$ . O caso em que  $(\overline{C}^{R_\Sigma})^I(x) = (\overline{\overline{C}^{R_\Sigma}})^I(x)$  segue ideia similar. Temos os seguintes casos a analisar:

- $(\underline{C}_{R_\Sigma})^I(x) = \mathbf{t}$ . Por definição,  $\text{GLB}_{y \in \Delta^I}(\mathbf{R}_\Sigma^I(x, y) \leftrightarrow C^I(y)) = \mathbf{t}$ . Pela semântica da implicação  $\leftrightarrow$  e do GLB, para todo  $w \in \Delta^I$ , temos que ou  $\mathbf{R}_\Sigma^I(x, w) = \mathbf{t}$  e  $C^I(w) = \mathbf{t}$ , ou  $\mathbf{R}_\Sigma^I(x, w) = \mathbf{f}$ . Para mostrar que  $\text{GLB}_{y \in \Delta^I}(\mathbf{R}_\Sigma^I(x, y) \leftrightarrow \text{GLB}_{z \in \Delta^I}(\mathbf{R}_\Sigma^I(y, z) \leftrightarrow C^I(z))) = \mathbf{t}$ , teremos os seguintes casos considerando qualquer  $y \in \Delta^I$ :

- $\mathbf{R}_\Sigma^I(x, y) = \mathbf{f}$ . Pela semântica a implicação  $\leftrightarrow$ , temos  $\mathbf{R}_\Sigma^I(x, y) \leftrightarrow \text{GLB}_{z \in \Delta^I}(\mathbf{R}_\Sigma^I(y, z) \leftrightarrow C^I(z)) = \mathbf{t}$ .
- $\mathbf{R}_\Sigma^I(x, y) = \mathbf{t}$ . Como  $\mathbf{R}_\Sigma$  é uma relação de equivalência, para todo  $w \in \Delta^I$ , se  $\mathbf{R}_\Sigma^I(x, w) = \mathbf{t}$  e  $C^I(w) = \mathbf{t}$ , então  $\mathbf{R}_\Sigma^I(y, w) = \mathbf{t}$  e  $C^I(w) = \mathbf{t}$ . Se  $\mathbf{R}_\Sigma^I(x, w) = \mathbf{f}$ , então  $\mathbf{R}_\Sigma^I(y, w) = \mathbf{f}$ . Logo, pela semântica a implicação  $\leftrightarrow$ , temos  $\mathbf{R}_\Sigma^I(x, y) \leftrightarrow \text{GLB}_{z \in \Delta^I}(\mathbf{R}_\Sigma^I(y, z) \leftrightarrow C^I(z)) = \mathbf{t}$ .

Logo,  $\text{GLB}_{y \in \Delta^I}(\mathbf{R}_\Sigma^I(x, y) \leftrightarrow \text{GLB}_{z \in \Delta^I}(\mathbf{R}_\Sigma^I(y, z) \leftrightarrow C^I(z))) = \mathbf{t}$ . Portanto,  $(\underline{C}_{R_\Sigma R_\Sigma})^I(x) = \mathbf{t}$ .

- $(\underline{C}_{R_\Sigma})^I(x) = \mathbf{i}$ ,  $(\underline{C}_{R_\Sigma})^I(x) = \mathbf{u}$  e  $(\underline{C}_{R_\Sigma})^I(x) = \mathbf{t}$ . São similares ao caso  $(\underline{C}_{R_\Sigma})^I(x) = \mathbf{t}$ .

2.  $(\underline{C}_{R_\Sigma})^I(x) = (\overline{\underline{C}_{R_\Sigma}})^I(x) \leq (\overline{C}^{R_\Sigma})^I(x)$ . Similar à prova do caso 1.

3.  $(\underline{C}_{R_\Sigma})^I(x) \leq (\overline{\underline{C}_{R_\Sigma}})^I(x) = (\overline{C}^{R_\Sigma})^I(x)$ . Similar à prova do caso 1.  $\square$

Note que a aplicação das aproximações *tight* e *loose* com respeito à relação de indiscernibilidade não resulta em novos resultados para uma consulta. Por sua vez, a proposição acima garante que se nós recorrermos às relações de similaridades, sucessivas aplicações de aproximações podem resultar em diferentes respostas. Por conseguinte, relações de similaridade podem ser empregadas no refinamento/restricção de consultas em DL. Para uma aplicação de refinamento de consultas em  $PR_{\mathcal{A}LC}$ , vamos considerar o exemplo abaixo com informações incompletas e contraditórias.

*Exemplo 20.* (Relaxamento/Restricção de Consultas) Sejam  $\{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$  um conjunto de indivíduos representando casas; *BoaLocalização*, *Porão*, *Lareira*, *Caro*, *Barato* e *Médio* conceitos;  $\Sigma = \{\text{BoaLocalização}, \text{Porão}, \text{Lareira}\}$  um contexto e  $I = (\Delta^I, \cdot^I)$  uma interpretação em  $PR_{\mathcal{A}LC}$  tal que

- $\Delta^I = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$ ;
- $BL^I(x_1) = \mathbf{t}; BL^I(x_2) = \mathbf{f}; BL^I(x_3) = \mathbf{t}; BL^I(x_4) = \mathbf{f}; BL^I(x_5) = \mathbf{u}; BL^I(x_6) = \mathbf{t}; BL^I(x_7) = \mathbf{f}$ ;
- $P^I(x_1) = \mathbf{t}; P^I(x_2) = \mathbf{i}; P^I(x_3) = \mathbf{f}; P^I(x_4) = \mathbf{t}; P^I(x_5) = \mathbf{u}; P^I(x_6) = \mathbf{i}; P^I(x_7) = \mathbf{t}$ ;
- $L^I(x_1) = \mathbf{t}; L^I(x_2) = \mathbf{f}; L^I(x_3) = \mathbf{u}; L^I(x_4) = \mathbf{f}; L^I(x_5) = \mathbf{t}; L^I(x_6) = \mathbf{t}; L^I(x_7) = \mathbf{t}$ ;
- $M^I(x_1) = \mathbf{f}; M^I(x_2) = \mathbf{f}; M^I(x_3) = \mathbf{t}; M^I(x_4) = \mathbf{t}; M^I(x_5) = \mathbf{t}; M^I(x_6) = \mathbf{f}; M^I(x_7) = \mathbf{t}$ ;
- $C^I(x_1) = \mathbf{t}; C^I(x_2) = \mathbf{f}; C^I(x_3) = \mathbf{f}; C^I(x_4) = \mathbf{f}; C^I(x_5) = \mathbf{f}; C^I(x_6) = \mathbf{t}; C^I(x_7) = \mathbf{f}$ ;
- $B^I(x_1) = \mathbf{f}; B^I(x_2) = \mathbf{t}; B^I(x_3) = \mathbf{f}; B^I(x_4) = \mathbf{f}; B^I(x_5) = \mathbf{f}; B^I(x_6) = \mathbf{f}; B^I(x_7) = \mathbf{f}$ .

Primeiramente iremos considerar um exemplo utilizando relaxamento de consulta. Suponha que queremos saber quais casas são caras. Temos que

$$I \models \text{Caro}(x_1), I \not\models \text{Caro}(x_2), I \not\models \text{Caro}(x_3), I \not\models \text{Caro}(x_4) \text{ e } I \not\models \text{Caro}(x_5).$$

Isso quer dizer que  $x_1$  é a única casa cara. Mas suponha que queremos saber que casas são possivelmente caras (casas que não são caras, mas possuem características de casas caras) de acordo com o contexto  $\Sigma$ . Relaxando essa consulta (utilizaremos a relação de similaridade  $S_\Sigma$  para esse exemplo), nós temos que

$$I \models \overline{\text{Caro}}^{S_\Sigma}(x_1) \text{ e } I \models \overline{\text{Caro}}^{S_\Sigma}(x_5).$$

Logo  $x_1$  e  $x_5$  são casas possivelmente caras. Observe que  $x_5$  é possivelmente cara por ser similar a  $x_1$ , que é avaliada como sendo cara. Se usarmos o relaxamento de consulta novamente teremos que

$$I \models \overline{\overline{\text{Caro}}^{S_\Sigma}}^{S_\Sigma}(x_1), I \models \overline{\overline{\text{Caro}}^{S_\Sigma}}^{S_\Sigma}(x_3) \text{ e } I \models \overline{\overline{\text{Caro}}^{S_\Sigma}}^{S_\Sigma}(x_5).$$

Nós temos agora que  $x_3$  é possivelmente uma casa possivelmente cara, pois  $x_3$  é similar a  $x_5$  de acordo com a relação  $S_\Sigma$  (ou seja,  $x_3$  tem menos características de casas caras do que  $x_5$ ). Iremos mostrar agora outro exemplo relacionado com refinamento de consultas, mas usando restrição de consulta: suponha que queremos saber quais casas possuem valor médio. Temos então que

$$I \not\models \text{Médio}(x_1), I \not\models \text{Médio}(x_2), I \models \text{Médio}(x_3), I \models \text{Médio}(x_4) \text{ e } I \models \text{Médio}(x_5).$$

Ou seja,  $x_1, x_3, x_4$  e  $x_5$  são casas com valor médio. Usando restrição de consulta através do contexto  $\Sigma$ , concluímos que

$$I \models \underline{\text{Médio}}_{S_\Sigma}(x_3), \text{ mas } I \not\models \underline{\text{Médio}}_{S_\Sigma}(x_4) \text{ e } I \not\models \underline{\text{Médio}}_{S_\Sigma}(x_5).$$

Os indivíduos  $x_4$  e  $x_5$  não possuem necessariamente um valor médio. Se usarmos restrição de consulta novamente, concluiremos que

$$I \not\models \underline{Médio}_{S\Sigma S\Sigma}(x_3).$$

Logo teremos que  $x_3$  necessariamente não deve ter valor médio, i.e.,  $x_3$  é similar a alguma casa que necessariamente não tem valor médio. Focando um pouco na relação de similaridade para informações inconsistentes e desconhecidas, temos que

$$I \not\models \overline{Barato}(x_4) \text{ e } I \not\models \overline{Barato}^{S\Sigma}(x_4), \text{ porém } I \models \overline{Barato}^{P\Sigma}(x_4).$$

Isso significa que  $P\Sigma$  pode ser usada para descobrir indivíduos que possuem relações com informações contraditórias dentro de um contexto. Sabendo que  $I \not\models \overline{Barato}^{S\Sigma}(x_4)$  e  $I \models \overline{Barato}^{P\Sigma}(x_4)$ , podemos inferir que ao aceitarmos a presença de similaridades com contradições em  $\Sigma$ ,  $x_4$  pode ser visto como um objeto possivelmente barato. Uma intuição parecida pode ser usada para a aproximação inferior em busca de encontrar aqueles indivíduos que certamente possuem determinada propriedade ao serem evidenciadas relações com contradições. Por exemplo, como

$$I \models \underline{Médio}_{S\Sigma}(x_3) \text{ e } I \not\models \underline{Médio}_{P\Sigma}(x_3),$$

podemos concluir que se permitirmos relações de similaridade com contradições,  $x_3$  não será considerada uma casa de valor médio, mas se analisarmos apenas as similaridades com valores desconhecidos, ela será considerada uma casa de valor médio. Em consideração ao indivíduo  $x_7$ , o resultado é que

$$I \models \underline{Médio}_{S\Sigma}(x_7) \text{ e } I \models \underline{Médio}_{P\Sigma}(x_7).$$

Isso mostra que independente de se analisar as similaridades com valores contraditórios ou não, o resultado será o mesmo. Em outras palavras, para  $x_7$  não existe a presença de similaridades com informações contraditórias explícitas. Outros tipos de consultas podem ser feitas, como por exemplo, aplicações alternadas entre aproximações superiores e inferiores. Como podemos ver, em  $PR_{\mathcal{ALC}}$ , podemos representar refinamentos de consultas bem elaborados.

### 5.4.3 Complexidade de Inferência

Para mostrar a complexidade do problema de satisfação de  $PR_{\mathcal{ALC}}$ , iremos traduzi-la na DL  $\mathcal{ALC}$ , tal que a partir disso, o seu raciocínio paraconsistente pode ser simulado usando os algoritmos de inferência de  $\mathcal{ALC}$ . Essa tradução apresentada abaixo é baseada no trabalho de (MA; HITZLER; LIN, 2006). Primeiramente, começaremos com a tradução dos conceitos de  $PR_{\mathcal{ALC}}$  para conceitos de  $\mathcal{ALC}$ .

**Definição 44 (Tradução  $\dagger$  - Conceitos)** Para cada conceito  $C_l$  em que  $l \in \{t, f, i, u\}$ , a sua tradução  $C_l^\dagger$  é o conceito obtido de  $C_l$  pelas seguintes transformações definidas indutivamente:

- $A_t^\dagger = A_t$ ;  $A_f^\dagger = A_f$ ;  $A_i^\dagger = A_i$ ;  $A_u^\dagger = A_u$ ;
- $\top_t^\dagger = \top$ ;  $\top_f^\dagger = \perp$ ;  $\top_i^\dagger = \perp$ ;  $\top_u^\dagger = \perp$ ;
- $\perp_t^\dagger = \perp$ ;  $\perp_f^\dagger = \top$ ;  $\perp_i^\dagger = \perp$ ;  $\perp_u^\dagger = \perp$ ;
- $(\neg C)_t^\dagger = C_f$ ;  $(\neg C)_f^\dagger = C_t$ ;  $(\neg C)_i^\dagger = C_i$ ;  $(\neg C)_u^\dagger = C_u$ ;
- $(C \cap D)_t^\dagger = C_t^\dagger \cap D_t^\dagger$ ;
- $(C \cap D)_f^\dagger = C_f^\dagger \sqcup D_f^\dagger$ ;
- $(C \cap D)_i^\dagger = (C_i^\dagger \cap (D_i^\dagger \sqcup D_t^\dagger)) \sqcup (C_t^\dagger \cap D_i^\dagger)$ ;
- $(C \cap D)_u^\dagger = (C_u^\dagger \cap \neg(D_f^\dagger)) \sqcup (\neg(C_f^\dagger) \cap D_u^\dagger)$ ;
- $(C \sqcup D)_t^\dagger = C_t^\dagger \sqcup D_t^\dagger$ ;
- $(C \sqcup D)_f^\dagger = C_f^\dagger \cap D_f^\dagger$ ;
- $(C \sqcup D)_i^\dagger = (C_i^\dagger \cap \neg(D_t^\dagger)) \sqcup (\neg(C_t^\dagger) \cap D_i^\dagger)$ ;
- $(C \sqcup D)_u^\dagger = (C_u^\dagger \cap (D_f^\dagger \sqcup D_u^\dagger)) \sqcup (C_f^\dagger \cap D_u^\dagger)$ ;
- $(\exists R.C)_t^\dagger = \exists R_t.C_t^\dagger$ ;
- $(\exists R.C)_f^\dagger = \forall R_u.C_f^\dagger \cap \forall R_i.C_f^\dagger \cap \forall R_t.C_f^\dagger$ ;
- $(\exists R.C)_i^\dagger = (\exists R_i.(C_i^\dagger \sqcup C_t^\dagger) \sqcup \exists R_t.C_i^\dagger) \cap \forall R_t.\neg(C_t^\dagger)$ ;
- $(\exists R.C)_u^\dagger = (\exists R_u.\neg(C_f^\dagger) \sqcup \exists R_i.C_u^\dagger \sqcup \exists R_t.C_u^\dagger) \cap \forall R_i.(C_f^\dagger \sqcup C_u^\dagger) \cap \forall R_t.(C_f^\dagger \sqcup C_u^\dagger)$ ;
- $(\forall R.C)_t^\dagger = \forall R_u.C_t^\dagger \sqcup \forall R_i.C_t^\dagger \sqcup \forall R_t.C_t^\dagger$ ;
- $(\forall R.C)_f^\dagger = \exists R_t.C_f^\dagger$ ;
- $(\forall R.C)_i^\dagger = (\exists R_u.C_i^\dagger \sqcup \exists R_i.\neg(C_t^\dagger) \sqcup \exists R_t.C_i^\dagger) \cap \forall R_t.(C_i^\dagger \sqcup C_t^\dagger) \cap \forall R_u.(C_i^\dagger \sqcup C_t^\dagger)$ ;
- $(\forall R.C)_u^\dagger = (\exists R_u.(C_f^\dagger \sqcup C_u^\dagger) \sqcup \exists R_t.C_u^\dagger) \cap \forall R_t.\neg(C_f^\dagger)$ ;
- $(\underline{C}_S)_t^\dagger = (\forall S.C)_t^\dagger$ ;  $(\underline{C}_S)_f^\dagger = (\forall S.C)_f^\dagger$ ;  $(\underline{C}_S)_i^\dagger = (\forall S.C)_i^\dagger$ ;  $(\underline{C}_S)_u^\dagger = (\forall S.C)_u^\dagger$ ;
- $(\overline{C}^S)_t^\dagger = (\exists S.C)_t^\dagger$ ;  $(\overline{C}^S)_f^\dagger = (\exists S.C)_f^\dagger$ ;  $(\overline{C}^S)_i^\dagger = (\exists S.C)_i^\dagger$ ;  $(\overline{C}^S)_u^\dagger = (\exists S.C)_u^\dagger$ ;

Essa tradução tem como objetivo preservar a noção de satisfatibilidade dos conceitos: a interpretação de um conceito  $C$  em  $PR_{\mathcal{ALC}}$  é satisfável (**t** ou **i**) se e somente se a interpretação da tradução de  $C$  é satisfável (**t**) em  $\mathcal{ALC}$ . Diferente da DL *Rough*  $\mathcal{ALC}_4$ , a DL  $PR_{\mathcal{ALC}}$  não apresenta a propriedade de De Morgan. i.e.  $\neg(C \sqcap D) \not\equiv \neg C \sqcup \neg D$  e  $\neg(C \sqcup D) \not\equiv \neg C \sqcap \neg D$ . Por exemplo, se  $C^I(x) = \mathbf{i}$  e  $D^I(x) = \mathbf{u}$ , temos que  $(\neg(C \sqcap D))^I(x) = \mathbf{u}$ , mas  $(\neg C \sqcup \neg D)^I = \mathbf{i}$ . Devido a esse problema, não é possível fazer uma tradução recursiva direta dos conceitos de  $PR_{\mathcal{ALC}}$  para os conceitos de  $\mathcal{ALC}$ .

A ideia geral da tradução é que cada conceito  $C$  em  $PR_{\mathcal{ALC}}$  será transformado em quatro novos conceitos de  $\mathcal{ALC}$ :  $C_t, C_i, C_f$  e  $C_u$ . A fim de preservar a noção de satisfatibilidade usamos a seguinte ideia: Se a interpretação  $I$  em  $PR_{\mathcal{ALC}}$  é  $C^I(x) = \mathbf{t}$ , então a interpretação  $I$  em  $\mathcal{ALC}$  é  $x \in C_t^I$ . A mesma intuição é aplicada para os outros valores verdade.

Sintaticamente, as traduções dos conceitos são construídas obedecendo tabela verdade de seus operadores. Para ficar mais clara a construção da tradução de conceitos, considere a regra

$$(C \sqcap D)_t^\dagger = C_t^\dagger \sqcap D_t^\dagger.$$

A semântica do operador  $\sqcap$  é obtida através do conectivo da interseção ( $\wedge$ ), que determina que  $(C \sqcap D)^I(x) = \mathbf{t}$  quando  $C^I(x) = \mathbf{t}$  e  $D^I(x) = \mathbf{t}$ . Isso implica que  $x$  deve estar contido nas interpretações dos conceitos  $C_t$  e  $D_t$ . Em outras palavras, sempre que  $(C \sqcap D)^I(x) = \mathbf{t}$ , o conceito  $C_t^\dagger \sqcap D_t^\dagger$  é satisfável em  $\mathcal{ALC}$ . Já para a regra

$$(C \sqcap D)_i^\dagger = (C_i^\dagger \sqcap (D_i^\dagger \sqcup D_f^\dagger)) \sqcup (C_f^\dagger \sqcap D_i^\dagger),$$

$(C \sqcap D)^I(x) = \mathbf{i}$  quando os seguintes casos ocorrem: (1)  $C^I(x) = \mathbf{i}$  e  $D^I(x) = \mathbf{i}$ , (2)  $C^I(x) = \mathbf{i}$  e  $D^I(x) = \mathbf{t}$  ou (3)  $C^I(x) = \mathbf{t}$  e  $D^I(x) = \mathbf{i}$ . Todos esses casos são capturados pelo conceito  $(C_i^\dagger \sqcap (D_i^\dagger \sqcup D_f^\dagger)) \sqcup (C_f^\dagger \sqcap D_i^\dagger)$ . Isso implica que sempre que  $(C \sqcap D)^I(x) = \mathbf{i}$ , o conceito  $(C_i^\dagger \sqcap (D_i^\dagger \sqcup D_f^\dagger)) \sqcup (C_f^\dagger \sqcap D_i^\dagger)$  deve ser satisfável em  $\mathcal{ALC}$ .

A tradução dos axiomas presentes nos TBox e ABox também seguirá a partir dessa ideia, respeitando a definição de suas semânticas em  $PR_{\mathcal{ALC}}$ .

**Definição 45 (Tradução  $\tau$  - Axiomas)** Para cada axioma  $\alpha$ , a sua tradução  $\tau(\alpha)$  é obtida através das seguintes regras:

- $\tau(C \sqsubseteq D) = \{(C_u^\dagger \sqcup C_t^\dagger) \sqsubseteq (D_i^\dagger \sqcup D_f^\dagger), C_i^\dagger \sqsubseteq \top\}$ ;
- $\tau(C \equiv D) = \{\tau(C \sqsubseteq D), \tau(D \sqsubseteq C)\}$ ;
- $\tau(C(a)) = (C_t^\dagger \sqcup C_i^\dagger)(a)$ ;
- $\tau(R(a, b)) = R_T(a, b)$ ;
- $\tau(S(a, b)) = S_T(a, b)$ .

Note que a função de transformação cresce exponencialmente no tamanho da base de conhecimento, pois ela é aplicada recursivamente em cada subconceito e cada transformação de um subconceito aumenta o tamanho do conceito resultante. O teorema abaixo mostra que o raciocínio paraconsistente pode, de fato, ser simulado através do raciocínio usual por meios da tradução dada. Isso implica que esse raciocínio paraconsistente da nossa linguagem não é mais expressivo que o raciocínio de  $\mathcal{ALC}$ , porém possui uma complexidade maior. Assumimos que para qualquer base de conhecimento  $\mathcal{K}$ , a função  $\tau(\mathcal{K})$  é definida como o conjunto  $\{\tau(\alpha) \mid \alpha \text{ é um axioma de } \mathcal{K}\}$ .

**Teorema 5** *Para qualquer base de conhecimento  $\mathcal{K}$  de  $PR_{\mathcal{ALC}}$  e  $\alpha$  um axioma de  $PR_{\mathcal{ALC}}$ , nós temos que  $\mathcal{K} \models \alpha$  se e somente se  $\tau(\mathcal{K}) \models_2 \tau(\alpha)$ , em que  $\models_2$  é a relação de consequência lógica em  $\mathcal{ALC}$ .*

No resto desta seção provaremos o Teorema 5. Para isso, primeiramente precisamos de algumas notações e definições.

**Definição 46 (Decomposição)** *A semântica tetravalorada de  $PR_{\mathcal{ALC}}$  pode ser decomposta na semântica bivalorada de  $\mathcal{ALC}$  se e somente se para qualquer conceito  $C$ , papel  $R$  e relação de similaridade  $S$  em uma base de conhecimento  $\mathcal{K}$  de  $PR_{\mathcal{ALC}}$ , existem quatro conceitos  $C_t, C_f, C_i, C_u$  e dez papéis  $R_t, R_f, R_i, R_u, R_T, S_t, S_f, S_i, S_u$  e  $S_T$  em uma base de conhecimento  $\tau(\mathcal{K})$  de  $\mathcal{ALC}$ , tal que para qualquer interpretação tetravalorada  $I$  de  $\mathcal{K}$ , existe uma interpretação bivalorada  $I$  de  $\tau(\mathcal{K})$ , em que para todo  $x, y \in \Delta^I$ :*

- $C^I(x) = \mathbf{t}$  sse  $x \in (C_t^\dagger)^I$ ;  $C^I(x) = \mathbf{f}$  sse  $x \in (C_f^\dagger)^I$ ;  
 $C^I(x) = \mathbf{i}$  sse  $x \in (C_i^\dagger)^I$ ;  $C^I(x) = \mathbf{u}$  sse  $x \in (C_u^\dagger)^I$ ;
- $R^I(x, y) = \mathbf{t}$  sse  $(x, y) \in R_t^I$ ;  $R^I(x, y) = \mathbf{f}$  sse  $(x, y) \in R_f^I$ ;  
 $R^I(x, y) = \mathbf{i}$  sse  $(x, y) \in R_i^I$ ;  $R^I(x, y) = \mathbf{u}$  sse  $(x, y) \in R_u^I$ ;
- $S^I(x, y) = \mathbf{t}$  sse  $(x, y) \in S_t^I$ ;  $S^I(x, y) = \mathbf{f}$  sse  $(x, y) \in S_f^I$ ;  
 $S^I(x, y) = \mathbf{i}$  sse  $(x, y) \in S_i^I$ ;  $S^I(x, y) = \mathbf{u}$  sse  $(x, y) \in S_u^I$ .

A decomposição de  $PR_{\mathcal{ALC}}$  significa que a semântica tetravalorada do conceito  $C$ , do papel  $R$  e da relação de similaridade  $S$  pode ser transformada na semântica bivalorada dos conceitos  $C_t, C_f, C_i, C_u$  e dos papéis  $R_t, R_f, R_i, R_u, S_t, S_f, S_i$  e  $S_u$  de  $\mathcal{ALC}$ .

Para eliminar a confusão de símbolos entre a linguagem original e a linguagem transformada, denotamos a linguagem original como  $\mathcal{L} = \{C, R, S, a\}$ , em que  $C$  representa um conceito,  $R$  um papel atômico,  $S$  uma relação de similaridade e  $a$  um indivíduo. A linguagem transformada  $\tau(\mathcal{L})$  é definida como  $\tau(\mathcal{L}) = \{C_t, C_f, C_i, C_u, R_t, R_f, R_i, R_u, S_t, R_T, S_f, S_i, S_u, S_T, a\}$ , em que  $C_l, l \in \{t, f, i, u\}$  representam as transformações do conceito  $C$  e  $R_l, S_l$  são novos papéis para  $R, S$ .  $R_T$  e  $S_T$  são definidos como  $R_T = R_t \cup R_i$  e  $S_T = S_t \cup S_i$ , respectivamente.

**Definição 47 (Interpretação Clássica Induzida)** *Seja  $I = (\Delta^I, \cdot^I)$  uma interpretação de  $PR_{\mathcal{ALC}}$ . A interpretação clássica induzida  $I = (\Delta^I, \cdot^I)$  de  $I$  é definida como*

- $I$  e  $I$  têm o mesmo domínio, i.e.  $\Delta^I = \Delta^I$ ;
- $I$  e  $I$  interpretam nomes de instâncias do mesmo jeito, i.e.  $a^I = a^I$ ;
- Para qualquer conceito atômico  $A$ ,
  - se  $A^I(x) = \mathbf{t}$  então  $x \in A_t^I$ ; se  $A^I(x) = \mathbf{f}$  então  $x \in A_f^I$ ;
  - se  $A^I(x) = \mathbf{i}$  então  $x \in A_i^I$ ; se  $A^I(x) = \mathbf{u}$  então  $x \in A_u^I$ ;
- Para os conceitos  $\top$  e  $\perp$ ,
  - Como  $\top^I(x) = \mathbf{t}$ , logo  $x \in \top_t^I, x \notin \top_f^I, x \notin \top_i^I$  e  $x \notin \top_u^I$ ;
  - Como  $\perp^I(x) = \mathbf{f}$ , logo  $x \in \perp_f^I, x \notin \perp_t^I, x \notin \perp_i^I$  e  $x \notin \perp_u^I$ ;
- Para qualquer papel  $R$ ,
  - se  $R^I(x, y) = \mathbf{t}$  então  $(x, y) \in R_t^I$ ; se  $R^I(x, y) = \mathbf{f}$  então  $(x, y) \in R_f^I$ ;
  - se  $R^I(x, y) = \mathbf{i}$  então  $(x, y) \in R_i^I$ ; se  $R^I(x, y) = \mathbf{u}$  então  $(x, y) \in R_u^I$ ;
- Para qualquer relação de similaridade  $S$ ,
  - se  $S^I(x, y) = \mathbf{t}$  então  $(x, y) \in S_t^I$ ; se  $S^I(x, y) = \mathbf{f}$  então  $(x, y) \in S_f^I$ ;
  - se  $S^I(x, y) = \mathbf{i}$  então  $(x, y) \in S_i^I$ ; se  $S^I(x, y) = \mathbf{u}$  então  $(x, y) \in S_u^I$ ;
- A interpretação clássica induzida dos conceitos complexos é obtida de acordo com a interpretação dos conceitos em  $\mathcal{ALC}$ .

Como mencionado anteriormente, as interpretações são construídas de modo que a noção de satisfação seja preservada em ambas as linguagens. A interpretação clássica induzida tem como objetivo preservar a satisfação de um conceito de  $PR_{\mathcal{ALC}}$  para um conceito de  $\mathcal{ALC}$ . Já a interpretação tetravalorada induzida preserva a satisfação de um conceito de  $\mathcal{ALC}$  para um conceito de  $PR_{\mathcal{ALC}}$ . Para evitar confusões com relação à notação, utilizaremos  $I$  para referir-se a interpretação tetravalorada de  $PR_{\mathcal{ALC}}$  e  $I$  para referir-se a interpretação bivalorada de  $\mathcal{ALC}$ .

**Definição 48 (Interpretação Tetravalorada Induzida)** *Seja  $I = (\Delta^I, \cdot^I)$  uma interpretação de  $\mathcal{ALC}$ . A interpretação tetravalorada induzida  $I = (\Delta^I, \cdot^I)$  de  $I$  é definida como:*

- $I$  e  $I$  têm o mesmo domínio, i.e.  $\Delta^I = \Delta^I$ ;
- $I$  e  $I$  interpretam nomes de instâncias do mesmo jeito, i.e.  $a^I = a^I$ ;
- Para qualquer conceito atômico  $A$ , para todo  $x \in \Delta^I$ ,

- se  $x \in A_t^I$  então  $A^I(x) = \mathbf{t}$  e  $(\neg A)^I(x) = \mathbf{f}$ ; se  $x \in A_f^I$  então  $A^I(x) = \mathbf{f}$  e  $(\neg A)^I(x) = \mathbf{t}$ ;  
se  $x \in A_i^I$  então  $A^I(x) = \mathbf{i}$  e  $(\neg A)^I(x) = \mathbf{i}$ ; se  $x \in A_u^I$  então  $A^I(x) = \mathbf{u}$  e  $(\neg A)^I(x) = \mathbf{u}$ ;
- Para os conceitos  $\top$  e  $\perp$ ,
  - Temos que  $x \in \top_t^I, x \notin \top_f^I, x \notin \top_i^I$  e  $x \notin \top_u^I$ , logo  $\top^I(x) = \mathbf{t}$ ;
  - Temos que  $x \in \perp_f^I, x \notin \perp_t^I, x \notin \perp_i^I$  e  $x \notin \perp_u^I$ , logo  $\perp^I(x) = \mathbf{f}$ ;
- Para qualquer papel  $R$ , para todo  $x, y \in \Delta^I$ ,
  - se  $(x, y) \in R_t^I$  então  $R^I(x, y) = \mathbf{t}$ ; se  $(x, y) \in R_f^I$  então  $R^I(x, y) = \mathbf{f}$ ;  
se  $(x, y) \in R_i^I$  então  $R^I(x, y) = \mathbf{i}$ ; se  $(x, y) \in R_u^I$  então  $R^I(x, y) = \mathbf{u}$ ;
- Para qualquer relação de similaridade  $S$ , para todo  $x, y \in \Delta^I$ ,
  - se  $(x, y) \in S_t^I$  então  $S^I(x, y) = \mathbf{t}$ ; se  $(x, y) \in S_f^I$  então  $S^I(x, y) = \mathbf{f}$ ;  
se  $(x, y) \in S_i^I$  então  $S^I(x, y) = \mathbf{i}$ ; se  $(x, y) \in S_u^I$  então  $S^I(x, y) = \mathbf{u}$ ;
- A interpretação tetravalorada induzida dos conceitos complexos é obtida de acordo com a interpretação dos conceitos em  $PR_{\mathcal{A}LC}$ .

Agora provaremos a decomposição de  $PR_{\mathcal{A}LC}$ .

**Lema 1** *A DL  $PR_{\mathcal{A}LC}$  pode ser decomposta para a semântica bivalorada da DL  $\mathcal{A}LC$ .*

*Prova.* Sejam  $C$  um conceito e  $x$  um indivíduo em  $PR_{\mathcal{A}LC}$ . Para qualquer interpretação tetravalorada  $I$ , nós mostraremos por indução na estrutura do conceito  $C$  que

- se  $C^I(x) = \mathbf{t}$  então  $x \in (C_t^\dagger)^I$ ; se  $C^I(x) = \mathbf{f}$  então  $x \in (C_f^\dagger)^I$ ;  
se  $C^I(x) = \mathbf{i}$  então  $x \in (C_i^\dagger)^I$ ; se  $C^I(x) = \mathbf{u}$  então  $x \in (C_u^\dagger)^I$ ;

em que  $I$  é a interpretação clássica induzida de  $I$ .

- Casos  $C = A$ ,  $C = \top$  e  $C = \perp$ . Seguem das Definições 47 e 48.
- Hipótese de Indução: Para quaisquer conceitos  $D$  e  $E$ ,
  - se  $D^I(x) = \mathbf{t}$  então  $x \in (D_t^\dagger)^I$ ; se  $D^I(x) = \mathbf{f}$  então  $x \in (D_f^\dagger)^I$ ;  
se  $D^I(x) = \mathbf{i}$  então  $x \in (D_i^\dagger)^I$ ; se  $D^I(x) = \mathbf{u}$  então  $x \in (D_u^\dagger)^I$ ;
  - se  $E^I(x) = \mathbf{t}$  então  $x \in (E_t^\dagger)^I$ ; se  $E^I(x) = \mathbf{f}$  então  $x \in (E_f^\dagger)^I$ ;  
se  $E^I(x) = \mathbf{i}$  então  $x \in (E_i^\dagger)^I$ ; se  $E^I(x) = \mathbf{u}$  então  $x \in (E_u^\dagger)^I$ .
- Caso  $C = \neg D$ .

- Suponha que  $C^I(x) = \mathbf{t}$ . Então, pela semântica da negação,  $D^I(x) = \mathbf{f}$ . Por Hipótese de Indução (HI),  $x \in (D_f^\dagger)^I$ . Como  $C_t^\dagger = (\neg D)_t^\dagger = D_f^\dagger$ , temos que  $x \in (C_t^\dagger)^I$ .
  - Suponha que  $C^I(x) = \mathbf{f}$ . Então, pela semântica da negação,  $D^I(x) = \mathbf{t}$ . Por Hipótese de Indução (HI),  $x \in (D_t^\dagger)^I$ . Como  $C_f^\dagger = (\neg D)_f^\dagger = D_t^\dagger$ , temos que  $x \in (C_f^\dagger)^I$ .
  - Suponha que  $C^I(x) = \mathbf{i}$ . Então, pela semântica da negação,  $D^I(x) = \mathbf{i}$ . Por Hipótese de Indução (HI),  $x \in (D_i^\dagger)^I$ . Como  $C_i^\dagger = (\neg D)_i^\dagger = D_i^\dagger$ , temos que  $x \in (C_i^\dagger)^I$ .
  - Suponha que  $C^I(x) = \mathbf{u}$ . Então, pela semântica da negação,  $D^I(x) = \mathbf{u}$ . Por Hipótese de Indução (HI),  $x \in (D_u^\dagger)^I$ . Como  $C_u^\dagger = (\neg D)_u^\dagger = D_u^\dagger$ , temos que  $x \in (C_u^\dagger)^I$ .
- Caso  $C = D \sqcap E$ .
- Suponha que  $C^I(x) = \mathbf{t}$ . Então, pela semântica da interseção,  $D^I(x) = \mathbf{t}$  e  $E^I(x) = \mathbf{t}$ . Por Hipótese de Indução (HI),  $x \in (D_t^\dagger)^I$  e  $x \in (E_t^\dagger)^I$ . Logo,  $x \in (D_t^\dagger \sqcap E_t^\dagger)^I$ . Como  $C_t^\dagger = (D \sqcap E)_t^\dagger = D_t^\dagger \sqcap E_t^\dagger$ , temos que  $x \in (C_t^\dagger)^I$ .
  - Suponha que  $C^I(x) = \mathbf{f}$ . Então, pela semântica da interseção,  $D^I(x) = \mathbf{f}$  ou  $E^I(x) = \mathbf{f}$ . Temos dois casos a analisar:
    - \* Suponha que  $D^I(x) = \mathbf{f}$ . Por HI,  $x \in (D_f^\dagger)^I$ . Logo,  $x \in (D_f^\dagger \sqcup E_f^\dagger)^I$ . Como  $C_f^\dagger = (D \sqcap E)_f^\dagger = D_f^\dagger \sqcup E_f^\dagger$ , temos que  $x \in (C_f^\dagger)^I$ .
    - \* Suponha que  $E^I(x) = \mathbf{f}$ . Por HI,  $x \in (E_f^\dagger)^I$ . Logo,  $x \in (D_f^\dagger \sqcup E_f^\dagger)^I$ . Como  $C_f^\dagger = (D \sqcap E)_f^\dagger = D_f^\dagger \sqcup E_f^\dagger$ , temos que  $x \in (C_f^\dagger)^I$ .
  - Suponha que  $C^I(x) = \mathbf{i}$ . Então, pela semântica da interseção, temos três casos a considerar:
    - \*  $D^I(x) = \mathbf{i}$  e  $E^I(x) = \mathbf{i}$ . Por HI,  $x \in (D_i^\dagger)^I$  e  $x \in (E_i^\dagger)^I$ . Logo,  $x \in (D_i^\dagger \sqcap E_i^\dagger)^I$ . Consequentemente,  $x \in (D_i^\dagger \sqcap (E_i^\dagger \sqcup E_t^\dagger))^I$  e  $x \in ((D_i^\dagger \sqcap (E_i^\dagger \sqcup E_t^\dagger)) \sqcup (D_t^\dagger \sqcap E_i^\dagger))^I$ . Como  $C_i^\dagger = (D \sqcap E)_i^\dagger = (D_i^\dagger \sqcap (E_i^\dagger \sqcup E_t^\dagger)) \sqcup (D_t^\dagger \sqcap E_i^\dagger)$ , temos que  $x \in (C_i^\dagger)^I$ .
    - \*  $D^I(x) = \mathbf{i}$  e  $E^I(x) = \mathbf{t}$ . Por HI,  $x \in (D_i^\dagger)^I$  e  $x \in (E_t^\dagger)^I$ . Logo,  $x \in (D_i^\dagger \sqcap E_t^\dagger)^I$ . Consequentemente,  $x \in (D_i^\dagger \sqcap (E_i^\dagger \sqcup E_t^\dagger))^I$  e  $x \in ((D_i^\dagger \sqcap (E_i^\dagger \sqcup E_t^\dagger)) \sqcup (D_t^\dagger \sqcap E_i^\dagger))^I$ . Como  $C_i^\dagger = (D \sqcap E)_i^\dagger = (D_i^\dagger \sqcap (E_i^\dagger \sqcup E_t^\dagger)) \sqcup (D_t^\dagger \sqcap E_i^\dagger)$ , temos que  $x \in (C_i^\dagger)^I$ .
    - \*  $D^I(x) = \mathbf{t}$  e  $E^I(x) = \mathbf{i}$ . Por HI,  $x \in (D_t^\dagger)^I$  e  $x \in (E_i^\dagger)^I$ . Logo,  $x \in (D_t^\dagger \sqcap E_i^\dagger)^I$ . Consequentemente,  $x \in ((D_i^\dagger \sqcap (E_i^\dagger \sqcup E_t^\dagger)) \sqcup (D_t^\dagger \sqcap E_i^\dagger))^I$ . Como  $C_i^\dagger = (D \sqcap E)_i^\dagger = (D_i^\dagger \sqcap (E_i^\dagger \sqcup E_t^\dagger)) \sqcup (D_t^\dagger \sqcap E_i^\dagger)$ , temos que  $x \in (C_i^\dagger)^I$ .
  - Suponha que  $C^I(x) = \mathbf{u}$ . Então, pela semântica da interseção, temos os seguintes casos a considerar:
    - \*  $D^I(x) = \mathbf{u}$  e  $E^I(x) = \mathbf{u}$ . Por HI,  $x \in (D_u^\dagger)^I$  e  $x \in (E_u^\dagger)^I$ . Logo,  $x \in (D_u^\dagger \sqcap E_u^\dagger)^I$ . Consequentemente,  $x \in (D_u^\dagger \sqcap \neg(E_f^\dagger))^I$ , pois  $\neg E_f = (E_t \sqcup E_i \sqcup E_u)$ , e então  $x \in ((D_u^\dagger \sqcap \neg(E_f^\dagger)) \sqcup (\neg(D_f^\dagger) \sqcap E_u^\dagger))^I$ . Como  $C_u^\dagger = (D \sqcap E)_u^\dagger = (D_u^\dagger \sqcap \neg(E_f^\dagger)) \sqcup (\neg(D_f^\dagger) \sqcap E_u^\dagger)$ , temos que  $x \in (C_u^\dagger)^I$ .
    - \*  $D^I(x) = \mathbf{u}$  e  $E^I(x) = \mathbf{i}$ . Por HI,  $x \in (D_u^\dagger)^I$  e  $x \in (E_i^\dagger)^I$ . Logo,  $x \in (D_u^\dagger \sqcap E_i^\dagger)^I$ . Consequentemente,  $x \in (D_u^\dagger \sqcap \neg(E_f^\dagger))^I$ , pois  $\neg E_f = (E_t \sqcup E_i \sqcup E_u)$ , e então  $x \in ((D_u^\dagger \sqcap \neg(E_f^\dagger)) \sqcup (\neg(D_f^\dagger) \sqcap E_u^\dagger))^I$ . Como  $C_u^\dagger = (D \sqcap E)_u^\dagger = (D_u^\dagger \sqcap \neg(E_f^\dagger)) \sqcup (\neg(D_f^\dagger) \sqcap E_u^\dagger)$ , temos que  $x \in (C_u^\dagger)^I$ .

- \*  $D^I(x) = \mathbf{u}$  e  $E^I(x) = \mathbf{t}$ . Por HI,  $x \in (D_u^\dagger)^I$  e  $x \in (E_t^\dagger)^I$ . Logo,  $x \in (D_u^\dagger \cap E_t^\dagger)^I$ . Consequentemente,  $x \in (D_u^\dagger \cap \neg(E_f^\dagger))^I$ , pois  $\neg E_f = (E_t \sqcup E_i \sqcup E_u)$ , e então  $x \in ((D_u^\dagger \cap \neg(E_f^\dagger)) \sqcup (\neg(D_f^\dagger) \cap E_u^\dagger))^I$ . Como  $C_u^\dagger = (D \cap E)_u^\dagger = (D_u^\dagger \cap \neg(E_f^\dagger)) \sqcup (\neg(D_f^\dagger) \cap E_u^\dagger)$ , temos que  $x \in (C_u^\dagger)^I$ .
- \*  $D^I(x) = \mathbf{i}$  e  $E^I(x) = \mathbf{u}$ . Por HI,  $x \in (D_i^\dagger)^I$  e  $x \in (E_u^\dagger)^I$ . Logo,  $x \in (D_i^\dagger \cap E_u^\dagger)^I$ . Consequentemente,  $x \in (\neg(D_f^\dagger) \cap E_u^\dagger)^I$ , pois  $\neg D_f = (D_t \sqcup D_i \sqcup D_u)$ , e então  $x \in ((D_u^\dagger \cap \neg(E_f^\dagger)) \sqcup (\neg(D_f^\dagger) \cap E_u^\dagger))^I$ . Como  $C_u^\dagger = (D \cap E)_u^\dagger = (D_u^\dagger \cap \neg(E_f^\dagger)) \sqcup (\neg(D_f^\dagger) \cap E_u^\dagger)$ , temos que  $x \in (C_u^\dagger)^I$ .
- \*  $D^I(x) = \mathbf{t}$  e  $E^I(x) = \mathbf{u}$ . Por HI,  $x \in (D_t^\dagger)^I$  e  $x \in (E_u^\dagger)^I$ . Logo,  $x \in (D_t^\dagger \cap E_u^\dagger)^I$ . Consequentemente,  $x \in (\neg(D_f^\dagger) \cap E_u^\dagger)^I$ , pois  $\neg D_f = (D_t \sqcup D_i \sqcup D_u)$ , e então  $x \in ((D_u^\dagger \cap \neg(E_f^\dagger)) \sqcup (\neg(D_f^\dagger) \cap E_u^\dagger))^I$ . Como  $C_u^\dagger = (D \cap E)_u^\dagger = (D_u^\dagger \cap \neg(E_f^\dagger)) \sqcup (\neg(D_f^\dagger) \cap E_u^\dagger)$ , temos que  $x \in (C_u^\dagger)^I$ .

- Caso  $C = D \sqcup E$ . A prova desse caso segue uma intuição similar ao caso  $C = D \cap E$ .
- Caso  $C = \exists R.D$ .

- Suponha que  $C^I(x) = \mathbf{t}$ , então pela semântica do conceito existencial, existe  $y \in \Delta^I$  tal que  $R^I(x, y) = \mathbf{t}$  e  $D^I(y) = \mathbf{t}$ . Pelas Definições 47 e 48 temos que  $(x, y) \in R_t^I$ . Por HI,  $y \in (D_t^\dagger)^I$ . Logo,  $x \in (\exists R_t.D_t^\dagger)^I$ . Como  $C_t^\dagger = (\exists R.D)_t^\dagger = \exists R_t.D_t^\dagger$ , temos que  $x \in (C_t^\dagger)^I$ .
- Suponha que  $C^I(x) = \mathbf{f}$ , então pela semântica do conceito existencial, para todo  $y \in \Delta^I$ , se  $R^I(x, y) = \mathbf{u}$  então  $D^I(y) = \mathbf{f}$  e se  $R^I(x, y) = \mathbf{i}$  então  $D^I(y) = \mathbf{f}$  e se  $R^I(x, y) = \mathbf{t}$  então  $D^I(y) = \mathbf{f}$ . Pelas Definições 47 e 48 temos que  $(x, y) \in R_u^I$ ,  $(x, y) \in R_i^I$ ,  $(x, y) \in R_t^I$ . Por HI,  $y \in (D_f^\dagger)^I$ . Logo,  $x \in (\forall R_u.D_f^\dagger \cap \forall R_i.D_f^\dagger \cap \forall R_t.D_f^\dagger)^I$ . Como  $C_f^\dagger = (\exists R.D)_f^\dagger = \forall R_u.D_f^\dagger \cap \forall R_i.D_f^\dagger \cap \forall R_t.D_f^\dagger$ , temos que  $x \in (C_f^\dagger)^I$ .
- Suponha que  $C^I(x) = \mathbf{i}$ . Pela semântica do conceito existencial, existe  $y \in \Delta^I$  tal que um dos seguintes casos podem ocorrer:
  - \*  $R^I(x, y) = \mathbf{i}$  e  $D^I(y) = \mathbf{i}$ , e para todo  $z \in \Delta^I$  que  $R^I(x, z) = \mathbf{t}$  então  $D^I(z) \neq \mathbf{t}$ . Pelas Definições 47 e 48,  $(x, y) \in R_i^I$  e  $(x, z) \in R_t^I$  (apenas para os casos em que  $R^I(x, z) = \mathbf{t}$ ). Por HI,  $y \in (D_i^\dagger)^I$  e  $z \in (\neg(D_t^\dagger))^I$  (todo  $z$  em que  $R^I(x, z) = \mathbf{t}$ ). Logo,  $x \in (\exists R_i.D_i^\dagger \cap \forall R_t.\neg(D_t^\dagger))^I$ . Consequentemente,  $x \in ((\exists R_i.(D_i^\dagger \sqcup D_t^\dagger)) \sqcup \exists R_t.D_i^\dagger \cap \forall R_t.\neg(D_t^\dagger))^I$ . Como  $C_i^\dagger = (\exists R.D)_i^\dagger = (\exists R_i.(D_i^\dagger \sqcup D_t^\dagger)) \sqcup \exists R_t.D_i^\dagger \cap \forall R_t.\neg(D_t^\dagger)$ , temos que  $x \in (C_i^\dagger)^I$ .
  - \*  $R^I(x, y) = \mathbf{i}$  e  $D^I(x) = \mathbf{t}$ , e para todo  $z \in \Delta^I$  que  $R^I(x, z) = \mathbf{t}$ , então  $D^I(z) \neq \mathbf{t}$ . Similar ao sub-item anterior.
  - \*  $R^I(x, y) = \mathbf{t}$  e  $D^I(x) = \mathbf{i}$ , e para todo  $z \in \Delta^I$  que  $R^I(x, z) = \mathbf{t}$ , então  $D^I(z) \neq \mathbf{t}$ . Similar ao sub-item anterior.
- Suponha  $C^I(x) = \mathbf{u}$ . Pela semântica do conceito existencial, existe  $y \in \Delta^I$  tal que um dos seguintes casos ocorrem:
  - \*  $R^I(x, y) = \mathbf{u}$ ,  $D^I(y) \neq \mathbf{f}$  e para todo  $z \in \Delta^I$  se  $R^I(x, z) = \mathbf{i}$  então  $D^I(z) = \mathbf{f}$  ou  $D^I(z) = \mathbf{u}$ , e se  $R^I(x, z) = \mathbf{t}$  então  $D^I(z) = \mathbf{f}$  ou  $D^I(z) = \mathbf{u}$ . Pelas Definições

47 e 48 temos que  $(x, y) \in R_u^\dagger$ ,  $(x, z) \in R_i^\dagger$  (caso  $R^I(x, z) = \mathbf{i}$ ) e  $(x, z) \in R_t^\dagger$  (caso  $R^I(x, z) = \mathbf{t}$ ). Por HI,  $y \in (\neg(D_f^\dagger))^I$  e  $z \in (D_f^\dagger \sqcup D_u^\dagger)^I$ . Logo,

$x \in (\exists R_u. \neg(D_f^\dagger) \sqcap \forall R_i. (D_f^\dagger \sqcup D_u^\dagger) \sqcap R_t. (D_f^\dagger \sqcup D_u^\dagger))^I$ . Consequentemente,

$x \in ((\exists R_u. \neg(D_f^\dagger) \sqcup \exists R_i. D_u^\dagger \sqcup \exists R_t. D_u^\dagger) \sqcap \forall R_i. (D_f^\dagger \sqcup D_u^\dagger) \sqcap R_t. (D_f^\dagger \sqcup D_u^\dagger))^I$ . Como  $(\exists R_u. \neg(D_f^\dagger) \sqcup \exists R_i. D_u^\dagger \sqcup \exists R_t. D_u^\dagger) \sqcap \forall R_i. (D_f^\dagger \sqcup D_u^\dagger) \sqcap R_t. (D_f^\dagger \sqcup D_u^\dagger) = (\exists R. D)_u^\dagger = C_u^\dagger$ , temos que  $x \in (C_u^\dagger)^I$ .

- \*  $R^I(x, y) = \mathbf{i}$ ,  $D^I(y) = \mathbf{u}$  e para todo  $z \in \Delta^I$  se  $R^I(x, z) = \mathbf{i}$  então  $D^I(z) = \mathbf{f}$  ou  $D^I(z) = \mathbf{u}$ , e se  $R^I(x, z) = \mathbf{t}$  então  $D^I(z) = \mathbf{f}$  ou  $D^I(z) = \mathbf{u}$ . Similar ao sub-item anterior.
- \*  $R^I(x, y) = \mathbf{t}$ ,  $D^I(y) = \mathbf{u}$  e para todo  $z \in \Delta^I$  se  $R^I(x, z) = \mathbf{i}$  então  $D^I(z) = \mathbf{f}$  ou  $D^I(z) = \mathbf{u}$ , e se  $R^I(x, z) = \mathbf{t}$  então  $D^I(z) = \mathbf{f}$  ou  $D^I(z) = \mathbf{u}$ . Similar ao sub-item anterior.

- Casos  $C = \forall R. D$ ,  $C = \underline{D}_S$  e  $C = \overline{D}^S$ . As provas desses casos seguem uma intuição similar ao caso  $C = \exists R. D$ . □

Agora provaremos o Teorema 5.

*Prova.* ( $\Rightarrow$ ) Para qualquer interpretação  $I$  de  $\mathcal{K}$ , seja  $I$  a interpretação clássica induzida de  $I$ . Iremos mostrar que para quaisquer interpretação  $I$  e axioma  $\alpha$  em  $PR_{\mathcal{ALC}}$ , se  $I \models \alpha$  então  $I \models_2 \tau(\alpha)$ , em que  $I$  é uma interpretação em  $\mathcal{ALC}$  e  $\tau(\alpha)$  um axioma de  $\mathcal{ALC}$ .

De acordo com a função de tradução  $\tau$ , para qualquer inclusão da forma  $C \sqsubseteq D \in \mathcal{K}$ , temos que  $(C_u^\dagger \sqcup C_t^\dagger) \sqsubseteq (D_i^\dagger \sqcup D_t^\dagger)$ ,  $C_i^\dagger \sqsubseteq \top \in \tau(\mathcal{K})$ . Suponha que  $I \models C \sqsubseteq D$ . Pela semântica da inclusão  $\sqsubseteq$ ,  $\mathbf{i} \leq \text{GLB}_{x \in \Delta^I}(C^I(x) \leftrightarrow D^I(x))$ .

- Caso  $\text{GLB}_{x \in \Delta^I}(C^I(x) \leftrightarrow D^I(x)) = \mathbf{i}$ . Pela definição da semântica da implicação  $\leftrightarrow$ , temos os seguintes casos a analisar:
  - para todo  $x \in \Delta^I$ ,  $C^I(x) = \mathbf{i}$  e  $D^I(x) \neq \mathbf{t}$ . Pelo Lema 1,  $x \in (C_i^\dagger)^I$  e  $x \in (D_f^\dagger \sqcup D_i^\dagger \sqcup D_u^\dagger)^I$ . Logo  $I \models_2 C_i^\dagger \sqsubseteq D_f^\dagger \sqcup D_i^\dagger \sqcup D_u^\dagger$ . Consequentemente,  $I \models_2 C_i^\dagger \sqsubseteq D_f^\dagger \sqcup D_i^\dagger \sqcup D_u^\dagger \sqcup D_t^\dagger$  e  $I \models_2 C_i^\dagger \sqsubseteq \top$ . Além disso, por vacuidade  $I \models_2 (C_u^\dagger \sqcup C_t^\dagger) \sqsubseteq (D_i^\dagger \sqcup D_t^\dagger)$ .
  - para todo  $x \in \Delta^I$ ,  $D^I(x) = \mathbf{i}$  e
    - \*  $C^I(x) = \mathbf{t}$ . Pelo Lema 1,  $x \in (C_t^\dagger)^I$  e  $x \in (D_i^\dagger)^I$ . Logo,  $I \models_2 C_t^\dagger \sqsubseteq D_i^\dagger$ . Consequentemente,  $I \models_2 (C_t^\dagger \sqcup C_u^\dagger) \sqsubseteq (D_i^\dagger \sqcup D_t^\dagger)$  e  $I \models_2 C_i^\dagger \sqsubseteq \top$ .
    - \*  $C^I(x) = \mathbf{u}$ . Similar ao caso em que  $C^I(x) = \mathbf{t}$ .
    - \*  $C^I(x) = \mathbf{i}$ . Similar ao caso em que  $C^I(x) = \mathbf{i}$  e  $D^I(x) \neq \mathbf{t}$ .
- Caso  $\text{GLB}_{x \in \Delta^I}(C^I(x) \leftrightarrow D^I(x)) = \mathbf{t}$ . Pela definição da semântica da inclusão, temos os seguintes casos a analisar:
  - para todo  $x \in \Delta^I$ ,  $C^I(x) = \mathbf{f}$ . Pelo Lema 1,  $x \in (C_f^\dagger)^I$ . Por vacuidade,  $I \models_2 (C_t^\dagger \sqcup C_u^\dagger) \sqsubseteq (D_i^\dagger \sqcup D_t^\dagger)$  e  $I \models_2 C_i^\dagger \sqsubseteq \top$ , para qualquer interpretação de  $D$ .

- para todo  $x \in \Delta^I$ ,  $D^I(x) = \mathbf{t}$ . Pelo Lema 1,  $x \in (D_i^\dagger)^I$ . Pela semântica da inclusão, temos que  $I \models_2 (C_i^\dagger \sqcup C_u^\dagger) \sqsubseteq (D_i^\dagger \sqcup D_i^\dagger)$  e  $I \models_2 C_i^\dagger \sqsubseteq \top$  para qualquer interpretação de  $C$ .

Para os axiomas da forma  $C \equiv D$ , podemos traduzi-los para os axiomas  $C \sqsubseteq D$  e  $D \sqsubseteq C$  e utilizarmos os mesmos argumentos apresentados acima. Para o axioma de asserção da forma  $C(a)$ , se  $I \models C(a)$ , temos os casos

- $C^I(a) = \mathbf{i}$ . Pelo Lema 1,  $a \in (C_i^\dagger)^I$  e, portanto,  $I \models_2 (C_i^\dagger \sqcup C_i^\dagger)(a)$ .
- $C^I(a) = \mathbf{t}$ . Pelo Lema 1,  $a \in (C_i^\dagger)^I$  e, portanto,  $I \models_2 (C_i^\dagger \sqcup C_i^\dagger)(a)$ .

Para os axiomas de asserção da forma  $R(a, b)$  e  $S(a, b)$ , suponha que  $I \models R(a, b)$  e  $I \models S(a, b)$ . Temos então:

- $R^I(a, b) = \mathbf{i}$ . Pelo Lema 1,  $(a, b) \in (R_i^\dagger)^I$ . Consequentemente,  $(a, b) \in (R_T^\dagger)$ . Portanto  $I \models_2 (R_T^\dagger)(a, b)$ .
- $R^I(a, b) = \mathbf{t}$ . Pelo Lema 1,  $(a, b) \in (R_i^\dagger)^I$ . Consequentemente,  $(a, b) \in (R_T^\dagger)$ . Portanto  $I \models_2 (R_T^\dagger)(a, b)$ .
- $S^I(a, b) = \mathbf{i}$ . Pelo Lema 1,  $(a, b) \in (S_i^\dagger)^I$ . Consequentemente,  $(a, b) \in (S_T^\dagger)$ . Portanto  $I \models_2 (S_T^\dagger)(a, b)$ .
- $S^I(a, b) = \mathbf{t}$ . Pelo Lema 1,  $(a, b) \in (S_i^\dagger)^I$ . Consequentemente,  $(a, b) \in (S_T^\dagger)$ . Portanto  $I \models_2 (S_T^\dagger)(a, b)$ .

( $\Leftarrow$ ) Para qualquer interpretação  $I = (\Delta^I, \cdot^I)$  de  $\tau(\mathcal{K})$ , seja  $I$  a interpretação tetravalorada induzida de  $I$ . A volta consiste em mostrar que para quaisquer interpretação  $I$  e axioma  $\tau(\alpha)$  em  $\mathcal{ALC}$ , se  $I \models_2 \tau(\alpha)$  então  $I \models \alpha$ , em que  $I$  é uma interpretação em  $PR_{\mathcal{ALC}}$  e  $\alpha$  um axioma de  $PR_{\mathcal{ALC}}$ . A prova pode ser feita similarmente à demonstração da ida ( $\Rightarrow$ ).  $\square$

Como resultado, temos que a  $PR_{\mathcal{ALC}}$  pode ser simulada através da  $\mathcal{ALC}$ , porém a complexidade de sua tradução cresce exponencialmente, e portanto a  $PR_{\mathcal{ALC}}$  tem um limite superior de complexidade maior que  $\mathcal{ALC}$ . Mais precisamente, temos os seguintes teoremas.

**Teorema 6** *A checagem de satisfatibilidade de conceitos de um TBox em  $PR_{\mathcal{ALC}}$  com apenas axiomas do tipo  $C \equiv D$  tem complexidade EXPSpace. A checagem de satisfatibilidade de conceitos de um TBox em  $PR_{\mathcal{ALC}}$  com axiomas do tipo  $C \equiv D$  e  $C \sqsubseteq D$  tem complexidade 2-EXPTIME.*

*Prova.* Seja uma base de conhecimento  $\mathcal{K} = (\mathcal{T}, \emptyset)$ . A tradução  $\tau(\mathcal{K}) = (\tau(\mathcal{T}), \emptyset)$ , consiste da tradução dos axiomas de inclusão e equivalência de conceitos de  $\mathcal{T}$ . De acordo com a construção das Definições 44 e 45, a tradução desses axiomas é feita utilizando espaço exponencial. Como mostrado no Teorema 2, a complexidade de satisfatibilidade de conceitos de um

TBox em  $\mathcal{ALC}$  com apenas axiomas do tipo  $C \equiv D$  é PSPACE-completo. Temos que o TBox  $\tau(\mathcal{T})$  de  $\mathcal{ALC}$  cresce exponencialmente em espaço com relação ao TBox original  $\mathcal{T}$  de  $PR_{\mathcal{ALC}}$ . Desse modo, a complexidade de satisfação de conceitos de  $\tau(\mathcal{T})$  com somente axiomas do tipo  $C \equiv D$  é EXPSPACE.

Além disso, de acordo com o Teorema 2, a checagem de satisfatibilidade de conceitos de um TBox em  $PR_{\mathcal{ALC}}$  com axiomas do tipo  $C \equiv D$  e  $C \sqsubseteq D$  tem complexidade EXPTIME-completo. Como o TBox  $\tau(\mathcal{T})$  de  $\mathcal{ALC}$  também cresce exponencialmente em tempo com relação ao TBox original  $\mathcal{T}$  de  $PR_{\mathcal{ALC}}$ , a complexidade de satisfação de conceitos de  $\tau(\mathcal{T})$  com axiomas do tipo  $C \equiv D$  e  $C \sqsubseteq D$  é 2-EXPTIME.  $\square$

**Teorema 7** *A checagem de consistência de um ABox em  $PR_{\mathcal{ALC}}$  tem complexidade EXPSPACE.*

*Prova.* Seja uma base de conhecimento  $\mathcal{K} = (\emptyset, \mathcal{A})$ . A tradução  $\tau(\mathcal{K}) = (\emptyset, \tau(\mathcal{A}))$ , consiste da tradução dos axiomas de assertão de  $\mathcal{A}$ . De acordo com a construção das Definições 44 e 45, a tradução desses axiomas é feita utilizando espaço exponencial. Como mostrado no Teorema 3, a checagem de consistência de conceitos de um ABox em  $\mathcal{ALC}$  é PSPACE-completo. Temos que o ABox  $\tau(\mathcal{A})$  de  $\mathcal{ALC}$  cresce exponencialmente em espaço com relação ao ABox original  $\mathcal{A}$  de  $PR_{\mathcal{ALC}}$ . Desse modo, a complexidade da checagem de consistência de conceitos de  $\tau(\mathcal{A})$  é EXPSPACE.  $\square$

**Teorema 8** *A checagem de consistência de um ABox e TBox em  $PR_{\mathcal{ALC}}$  com apenas axiomas do tipo  $C \equiv D$  tem complexidade EXPSPACE. A checagem de consistência de um ABox e TBox em  $PR_{\mathcal{ALC}}$  com axiomas do tipo  $C \equiv D$  e  $C \sqsubseteq D$  tem complexidade 2-EXPTIME-completo.*

*Prova.* Seja uma base de conhecimento  $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ . A tradução  $\tau(\mathcal{K}) = (\tau(\mathcal{T}), \tau(\mathcal{A}))$ , consiste da tradução dos axiomas de  $\mathcal{T}$  e  $\mathcal{A}$ . De acordo com a construção das Definições 44 e 45, a tradução desses axiomas é feita utilizando tempo e espaço exponenciais. Como mostrado no Teorema 4, a checagem de consistência de conceitos de um ABox e TBox em  $\mathcal{ALC}$  é PSPACE-completo, utilizando apenas axiomas do tipo  $C \equiv D$  no TBox; e EXPTIME-completo, com axiomas do tipo  $C \equiv D$  e  $C \sqsubseteq D$  no TBox. Temos que os ABox  $\tau(\mathcal{A})$  e TBox  $\tau(\mathcal{T})$  de  $\mathcal{ALC}$  crescem exponencialmente em tempo e espaço com relação aos ABox  $\mathcal{A}$  e TBox  $\mathcal{T}$  de  $PR_{\mathcal{ALC}}$ . Desse modo, a complexidade da checagem de consistência de conceitos de  $\tau(\mathcal{A})$  e  $\tau(\mathcal{T})$  é EXPSPACE, utilizando apenas axiomas do tipo  $C \equiv D$  no TBox; e 2-EXPTIME, com axiomas do tipo  $C \equiv D$  e  $C \sqsubseteq D$  no TBox.  $\square$

Destacamos que esse resultado mostra apenas um limite superior de  $PR_{\mathcal{ALC}}$ . Uma questão em aberto é se existe um limite superior menor para a checagem de consistência em  $PR_{\mathcal{ALC}}$ . Conjecturamos que existe a possibilidade desse limite superior ser EXPTIME, a mesma complexidade de  $\mathcal{ALC}$ , e que seja possível mostrá-lo através de um método de prova (tableaux) similar ao de  $\mathcal{ALC}$  (BAADER, 2003).

## 6 OBTENDO CONTEXTOS PARA REFINAMENTO DE CONSULTAS

Neste capítulo apresentaremos alguns resultados sobre aproximações em *Rough*  $\mathcal{ALC}_4$  e  $PR_{\mathcal{ALC}}$  e discutiremos um método de obtenção de consultas refinadas de axiomas dessas lógicas. Trata-se de um método automático de seleção de contextos que satisfaz critérios fundamentais dos relaxamentos e das restrições de consultas.

Aproximações contextuais diferem das aproximações usuais por utilizarem a noção de contexto, no qual as relações (de equivalência ou similaridade) entre indivíduos são descobertas a partir de um determinado contexto. Como consequência, diferentes contextos podem gerar diferentes relações entre indivíduos. Quando pensamos em um refinamento de um axioma em DL usando *rough sets*, estamos falando em encontrar um contexto adequado para ser utilizado na aproximação. Por exemplo, se estamos procurando relaxar um conceito  $C$  usando a relação de equivalência  $R$ , procuramos buscar um contexto  $\Sigma$ , tal que  $\overline{C}^{R_\Sigma}$  seja consequência lógica da base de conhecimento. Contudo, nem sempre é possível encontrar um contexto para um refinamento, ou possivelmente pode acontecer de haver diversos contextos adequados para um refinamento. Seria viável ter um mecanismo responsável por retornar os possíveis resultados de um refinamento (caso eles existam).

Se pensarmos sobre um mecanismo automático de refinamento de consultas, as possibilidades de gerar todos os contextos é de  $2^{|N_C|}$ , em que  $N_C$  é o conjunto de todos os conceitos atômicos presentes na base de conhecimento. Além do mais, grande parte desses contextos podem ser redundantes. Dizemos que um contexto é redundante quando um dos dois seguintes casos acontecem: (1) no problema de restrição de consulta, um contexto  $\Sigma$  é dito redundante, quando existe um contexto  $\Sigma_1$  (não-vazio), tal que  $\Sigma_1 \subset \Sigma$  e  $\Sigma_1$  é adequado para uma restrição; (2) no problema de relaxamento de consulta, um contexto  $\Sigma$  é dito redundante, quando existe um contexto  $\Sigma_1$ , tal que  $\Sigma \subset \Sigma_1$  e  $\Sigma_1$  é adequado para um relaxamento. Essa questão será abordada detalhadamente adiante durante este capítulo. De fato, o que pretendemos durante a busca de contextos é encontrar conjuntos de contextos minimais (quando trabalhamos com restrições) e maximais (quando trabalhamos com relaxamentos). A partir dos contextos minimais, sabemos que todos seus superconjuntos também são adequados para a restrição de consulta. A partir dos contextos maximais, sabemos que todos seus subconjuntos também são adequados para o relaxamento de consulta. A fim de resolver essa questão, neste capítulo apresentaremos um método baseado nas noções de matrizes de discernibilidade (YAO; ZHAO, 2009) para computar contextos para aproximações inferiores e superiores.

### 6.1 Usando as Aproximações

O principal problema encontrado em refinamento de consultas com *rough sets* e aproximação contextual é determinar um conjunto de conceitos (contexto) que faça uma restrição (aproximação inferior) ou um relaxamento (aproximação superior) de um conceito ser consequência lógica de uma base de conhecimento. Chamaremos esses contextos de contextos adequados. Os seguintes resultados nos ajudarão a descobrir esses conjuntos adequados.

**Proposição 6** *Dados dois contextos  $\Sigma_1$  e  $\Sigma_2$  tal que  $\Sigma_1 \subseteq \Sigma_2$ , as seguintes equações valem para qualquer conceito  $C$ , indivíduo  $x$ , interpretação  $I$  e relação de similaridade  $\text{Sim} \in \{\mathbf{R}, \mathbf{S}, \mathbf{P}\}$ :*

$$(\underline{C}_{\text{Sim}_{\Sigma_1}})^I(x) \leq (\underline{C}_{\text{Sim}_{\Sigma_2}})^I(x) \text{ e } (\overline{C}^{\text{Sim}_{\Sigma_2}})^I(x) \leq (\overline{C}^{\text{Sim}_{\Sigma_1}})^I(x).$$

*Prova.* Como  $\Sigma_1 \subseteq \Sigma_2$ , temos que  $\text{Sim}_{\Sigma_2} \subseteq \text{Sim}_{\Sigma_1}$ . Consequentemente,  $\text{Sim}_{\Sigma_2}(x) \subseteq \text{Sim}_{\Sigma_1}(x)$ , em que  $\text{Sim}_{\Sigma}(x) = \{y \in \Delta^I \mid (x, y) \in \text{Sim}_{\Sigma}\}$  e  $\text{Sim}_{\Sigma_2}^I(x, y) \leq \text{Sim}_{\Sigma_1}^I(x, y)$ . Pela definição da interpretação da aproximação inferior,  $(\underline{C}_{\text{Sim}_{\Sigma_1}})^I(x) = \text{GLB}_{y \in \Delta^I}(\text{Sim}_{\Sigma_1}^I(x, y) \leftrightarrow C^I(y))$  e  $(\underline{C}_{\text{Sim}_{\Sigma_2}})^I(x) = \text{GLB}_{y \in \Delta^I}(\text{Sim}_{\Sigma_2}^I(x, y) \leftrightarrow C^I(y))$ . Como  $\text{Sim}_{\Sigma}$  é uma relação bivalorada (é mapeada para  $\mathbf{t}$  ou  $\mathbf{f}$ ).

Temos os seguintes casos, para qualquer  $y \in \Delta^I$ :

- Suponha  $\text{Sim}_{\Sigma_1}^I(x, y) = \mathbf{t}$ . Como  $\text{Sim}_{\Sigma_2}^I(x, y) \leq \text{Sim}_{\Sigma_1}^I(x, y)$ , temos que se
  - $\text{Sim}_{\Sigma_2}^I(x, y) = \mathbf{t}$ , claramente  $\text{Sim}_{\Sigma_1}^I(x, y) \leftrightarrow C^I(y) = \text{Sim}_{\Sigma_2}^I(x, y) \leftrightarrow C^I(y)$ .
  - $\text{Sim}_{\Sigma_2}^I(x, y) = \mathbf{f}$ , pela semântica de  $\leftrightarrow$ ,  $\text{Sim}_{\Sigma_1}^I(x, y) \leftrightarrow C^I(y) \leq \text{Sim}_{\Sigma_2}^I(x, y) \leftrightarrow C^I(y)$ .
- Suponha  $\text{Sim}_{\Sigma_1}^I(x, y) = \mathbf{f}$ . Como  $\text{Sim}_{\Sigma_2}^I(x, y) \leq \text{Sim}_{\Sigma_1}^I(x, y)$ , temos que
  - $\text{Sim}_{\Sigma_2}^I(x, y) = \mathbf{f}$ , e claramente  $\text{Sim}_{\Sigma_1}^I(x, y) \leftrightarrow C^I(y) = \text{Sim}_{\Sigma_2}^I(x, y) \leftrightarrow C^I(y)$ .

Logo,  $(\underline{C}_{\text{Sim}_{\Sigma_1}})^I(x) = \text{GLB}_{y \in \Delta^I}(\text{Sim}_{\Sigma_1}^I(x, y) \leftrightarrow C^I(y)) \leq \text{GLB}_{y \in \Delta^I}(\text{Sim}_{\Sigma_2}^I(x, y) \leftrightarrow C^I(y)) = (\underline{C}_{\text{Sim}_{\Sigma_2}})^I(x)$ .

Para provar que  $(\overline{C}^{\text{Sim}_{\Sigma_2}})^I(x) \leq (\overline{C}^{\text{Sim}_{\Sigma_1}})^I(x)$ , podemos usar argumento similar, só que utilizando as semânticas de  $\wedge$  e LUB, em vez de  $\leftrightarrow$  e GLB.  $\square$

Intuitivamente, a Proposição 6 afirma que aumentando o tamanho do contexto, o tamanho da interpretação de um conceito também cresce para a aproximação inferior e decresce para a aproximação superior. Portanto, em busca de encontrar um contexto para satisfazer a aproximação inferior de um conceito qualquer, apenas aqueles contextos minimais são suficientes, pois todos seus superconjuntos também preservam a propriedade de satisfação. Analogamente, em relação aos contextos para a aproximação superior, apenas os maximais satisfazendo a aproximação bastarão. Finalmente, para aproximações *loose* e *tight*, as seguintes sentenças valem:

**Proposição 7** *Dados um contexto  $\Sigma$ , um conceito  $C$ , um indivíduo  $x$ , uma interpretação  $I$  e um número natural  $n$ , é válido que*

- $(\overline{C}^{(\mathbf{R}_{\Sigma}, n)})^I(x) = (\overline{C}^{(\mathbf{R}_{\Sigma}, n+1)})^I(x)$  e  $(\underline{C}_{(\mathbf{R}_{\Sigma}, n)})^I(x) = (\underline{C}_{(\mathbf{R}_{\Sigma}, n+1)})^I(x)$ ;
- $(\overline{C}^{(\mathbf{S}_{\Sigma}, n)})^I(x) \leq (\overline{C}^{(\mathbf{S}_{\Sigma}, n+1)})^I(x)$  e  $(\underline{C}_{(\mathbf{S}_{\Sigma}, n+1)})^I(x) \leq (\underline{C}_{(\mathbf{S}_{\Sigma}, n)})^I(x)$ ;
- $(\overline{C}^{(\mathbf{P}_{\Sigma}, n)})^I(x) \leq (\overline{C}^{(\mathbf{P}_{\Sigma}, n+1)})^I(x)$  e  $(\underline{C}_{(\mathbf{P}_{\Sigma}, n+1)})^I(x) \leq (\underline{C}_{(\mathbf{P}_{\Sigma}, n)})^I(x)$ .

*Prova.* Como  $R_\Sigma$  é uma relação de equivalência (reflexiva, simétrica e transitiva), temos que  $R_{\Sigma_n}^I(x, y) = R_{\Sigma_{n+1}}^I(x, y)$ , para quaisquer  $x, y \in \Delta^I$ . Logo, por definição

$$\begin{aligned} (\overline{C}^{(R_\Sigma, n)})^I(x) &= \text{LUB}_{y \in \Delta^I} (R_{\Sigma_n}^I(x, y) \wedge C^I(y)) = \text{LUB}_{y \in \Delta^I} (R_{\Sigma_{n+1}}^I(x, y) \wedge C^I(y)) = (\overline{C}^{(R_\Sigma, n+1)})^I(x) \text{ e} \\ (\underline{C}_{(R_\Sigma, n)})^I(x) &= \text{GLB}_{y \in \Delta^I} (R_{\Sigma_n}^I(x, y) \leftrightarrow C^I(y)) = \text{GLB}_{y \in \Delta^I} (R_{\Sigma_{n+1}}^I(x, y) \leftrightarrow C^I(y)) = (\underline{C}_{(R_\Sigma, n+1)})^I(x). \end{aligned}$$

Para as relações de similaridade  $S_\Sigma$  e  $P_\Sigma$ , de acordo com a definição da *relação-passo-k* (Definição 42), teremos que  $S_{\Sigma_n}^I(x, y) \leq S_{\Sigma_{n+1}}^I(x, y)$  e  $P_{\Sigma_n}^I(x, y) \leq P_{\Sigma_{n+1}}^I(x, y)$ , pois  $S_{\Sigma_n} \subseteq S_{\Sigma_{n+1}}$  e  $P_{\Sigma_n} \subseteq P_{\Sigma_{n+1}}$ . Segue-se então que

$$\begin{aligned} (\overline{C}^{(S_\Sigma, n)})^I(x) &= \text{LUB}_{y \in \Delta^I} (S_{\Sigma_n}^I(x, y) \wedge C^I(y)) \leq \text{LUB}_{y \in \Delta^I} (S_{\Sigma_{n+1}}^I(x, y) \wedge C^I(y)) = (\overline{C}^{(S_\Sigma, n+1)})^I(x) \\ &\quad \text{(pelas semânticas de } \wedge \text{ e LUB),} \\ (\underline{C}_{(S_\Sigma, n)})^I(x) &= \text{GLB}_{y \in \Delta^I} (S_{\Sigma_n}^I(x, y) \leftrightarrow C^I(y)) \geq \text{GLB}_{y \in \Delta^I} (S_{\Sigma_{n+1}}^I(x, y) \leftrightarrow C^I(y)) = (\underline{C}_{(S_\Sigma, n+1)})^I(x) \\ &\quad \text{(pelas semânticas de } \leftrightarrow \text{ e GLB),} \\ (\overline{C}^{(P_\Sigma, n)})^I(x) &= \text{LUB}_{y \in \Delta^I} (P_{\Sigma_n}^I(x, y) \wedge C^I(y)) \leq \text{LUB}_{y \in \Delta^I} (P_{\Sigma_{n+1}}^I(x, y) \wedge C^I(y)) = (\overline{C}^{(P_\Sigma, n+1)})^I(x) \\ &\quad \text{(pelas semânticas de } \wedge \text{ e LUB),} \\ (\underline{C}_{(P_\Sigma, n)})^I(x) &= \text{GLB}_{y \in \Delta^I} (P_{\Sigma_n}^I(x, y) \leftrightarrow C^I(y)) \geq \text{GLB}_{y \in \Delta^I} (P_{\Sigma_{n+1}}^I(x, y) \leftrightarrow C^I(y)) = (\underline{C}_{(P_\Sigma, n+1)})^I(x) \\ &\quad \text{(pelas semânticas de } \leftrightarrow \text{ e GLB).} \end{aligned}$$

□

Uma aplicação para as aproximações superiores *loose* surge quando não existem contextos que satisfaçam a aproximação superior de um conceito. Em outras palavras, uma relação de similaridade de um passo maior pode ser usada na busca de um contexto. Similarmemente, a aproximação *tight* pode ser aplicada para descobrir um conjunto de conceitos que reforçam a aproximação inferior, i.e., contextos que preservam a aproximação inferior em uma relação de similaridade de um passo maior. Note que o resultado para a aproximação superior *loose* não muda para a relação de indiscernibilidade (ou não diminui quando a aproximação inferior *tight* é considerada), pois ela é transitiva e não aumenta o tamanho da interpretação quando ela é aplicada sucessivamente.

*Exemplo 21.* Retornando ao *Exemplo 20*, no qual temos:

- $\Delta^I = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$ ;
- $BL^I(x_1) = \mathbf{t}; BL^I(x_2) = \mathbf{f}; BL^I(x_3) = \mathbf{t}; BL^I(x_4) = \mathbf{f}; BL^I(x_5) = \mathbf{u}; BL^I(x_6) = \mathbf{t}; BL^I(x_7) = \mathbf{f}$ ;
- $P^I(x_1) = \mathbf{t}; P^I(x_2) = \mathbf{i}; P^I(x_3) = \mathbf{f}; P^I(x_4) = \mathbf{t}; P^I(x_5) = \mathbf{u}; P^I(x_6) = \mathbf{i}; P^I(x_7) = \mathbf{t}$ ;
- $L^I(x_1) = \mathbf{t}; L^I(x_2) = \mathbf{f}; L^I(x_3) = \mathbf{u}; L^I(x_4) = \mathbf{f}; L^I(x_5) = \mathbf{t}; L^I(x_6) = \mathbf{t}; L^I(x_7) = \mathbf{t}$ ;
- $M^I(x_1) = \mathbf{f}; M^I(x_2) = \mathbf{f}; M^I(x_3) = \mathbf{t}; M^I(x_4) = \mathbf{t}; M^I(x_5) = \mathbf{t}; M^I(x_6) = \mathbf{f}; M^I(x_7) = \mathbf{t}$ ;

- $C^I(x_1) = \mathbf{t}; C^I(x_2) = \mathbf{f}; C^I(x_3) = \mathbf{f}; C^I(x_4) = \mathbf{f}; C^I(x_5) = \mathbf{f}; C^I(x_6) = \mathbf{t}; C^I(x_7) = \mathbf{f};$
- $B^I(x_1) = \mathbf{f}; B^I(x_2) = \mathbf{t}; B^I(x_3) = \mathbf{f}; B^I(x_4) = \mathbf{f}; B^I(x_5) = \mathbf{f}; B^I(x_6) = \mathbf{f}; B^I(x_7) = \mathbf{f}.$

A consulta perguntando que se  $x_3$  é uma casa cara de acordo com a interpretação  $I$  não é satisfeita, i.e.,

$$I \not\models \text{Caro}(x_3).$$

Ao consultarmos agora se  $x_3$  é uma casa possivelmente cara com relação ao contexto  $\Sigma = \{\text{BoaLocalização}, \text{Porão}, \text{Lareira}\}$  e a similaridade  $S$ , novamente também não será satisfeita, ou seja,

$$I \not\models \overline{\text{Caro}}^{(S_\Sigma, 1)}(x_3).$$

Mas se aplicarmos novamente a aproximação superior para o mesmo contexto, chegaremos ao resultado que

$$I \models \overline{\text{Caro}}^{(S_\Sigma, 2)}(x_3).$$

O principal problema a ser considerado é descobrir quais contextos servem para serem utilizados durante as aproximações. No exemplo acima, o contexto  $\Sigma$  não foi suficiente para satisfazer o conceito  $\overline{\text{Caro}}^{(S_\Sigma, 1)}(x_3)$ . A pergunta que surge é a seguinte: Não existiria um contexto  $\Sigma'$  que conseguisse satisfazer  $\overline{\text{Caro}}^{(S_{\Sigma'}, 1)}(x_3)$ ? A resposta é sim, pois o contexto  $\Sigma' = \{\text{Porão}, \text{Lareira}\}$  satisfaz  $\overline{\text{Caro}}^{(S_{\Sigma'}, 1)}(x_3)$ .

Motivados por esse problema de encontrar contextos adequados para cada tipo de consulta, a seguir iremos propor um método de busca de contextos para as aproximações em *Rough  $\mathcal{ALC}_4$*  e *PR $_{\mathcal{ALC}}$* .

## 6.2 Contextos para Aproximações Inferiores

Primeiro iremos considerar o problema de procurar um contexto adequado para restrição de consultas. Como dito anteriormente, neste trabalho iremos considerar apenas refinamentos em ABoxes. Maiores detalhes sobre a questão de se considerar refinamentos também em TBoxes serão apresentados no capítulo de conclusão da dissertação. Nosso problema de restrição de consulta pode ser formulado da seguinte maneira:

- **Entrada:** O universo  $\Delta^I$ , o conjunto de nomes de conceitos  $N_C$ , um ABox  $\mathcal{A}$ , uma relação de similaridade  $\text{Sim}$  e um conceito  $C$ .
- **Saída:** Contextos  $\Sigma \subseteq N_C - \text{atom}(C)$  e um número natural  $n$  tal que  $\mathcal{A} \models \underline{C}_{(\text{Sim}_\Sigma, n)}$ .

A função  $atom(C)$  retorna o conjunto de nomes de conceitos atômicos presentes no conceito  $C$ . Fazendo uma analogia com uma tabela de decisão em *rough sets*, o conceito  $C$  pode ser visto como o atributo de decisão e o contexto como os atributos condicionais. Temos que na literatura dos *rough sets*, o atributo de decisão não possui relação direta com os atributos condicionais, portanto não permitiremos aqui que os conceitos presentes em um contexto façam parte do conceito  $C$ . Uma análise detalhada dessa questão será abordada na conclusão da dissertação.

O problema de selecionar um subconjunto de conceitos para ser utilizado como um contexto se assemelha ao problema de encontrar redutos em *rough sets*, que são subconjuntos de atributos que preservam a informação do conjunto original. Para uma base de conhecimento em DL, os redutos são determinados por conjuntos minimais de conceitos que preservam discernibilidade entre todos indivíduos. Um reduto é, portanto, um conjunto minimal de conceitos que possui a mesma relação de indiscernibilidade do conjunto de todos os conceitos.

Na teoria dos *rough sets*, a computação de todos os tipos de redutos é baseada em matrizes de discernibilidade (SKOWRON; RAUSZER, 1992) como mostrado no Capítulo 2. Tais matrizes são construídas a partir da relação de discernibilidade. Neste trabalho, consideraremos dissimilaridade (STEPANIUK, 1998), pois estamos trabalhando com noções de similaridades. Realçamos que uma relação de dissimilaridade pode ser vista como o complemento de uma relação de similaridade.

Outra consideração importante é que, para resolver o nosso problema não precisamos utilizar todos os elementos de uma matriz de dissimilaridade, pois como estamos preocupados com contextos para uma asserção de conceito (e.g.  $C(a)$ ), precisamos considerar apenas uma linha da matriz, i.e. a linha das dissimilaridades do indivíduo da asserção. Para capturar essa ideia, definiremos a função de dissimilaridade  $DIS(\Sigma, x, y, Sim)$ .

**Definição 49 (Função de Dissimilaridade - DIS)** *Sejam  $\Delta^I$  o conjunto universo,  $x, y \in \Delta^I$ ,  $\Sigma$  um contexto e  $Sim$  uma relação de similaridade. A função de dissimilaridade  $DIS(\Sigma, x, y, Sim)$  é definida como*

$$DIS(\Sigma, x, y, Sim) = \{A_i \in \Sigma \mid y \notin Sim_{\{A_i\}}(x)\}, \text{ tal que } Sim_{\{A_i\}}(x) = \{y \in \Delta^I \mid (x, y) \in Sim_{\{A_i\}}\}.$$

Intuitivamente,  $DIS(\Sigma, x, y, Sim)$  descreve o conjunto de todos os conceitos em  $\Sigma$  em que cada indivíduo  $x$  não é similar a  $y$  com relação a  $Sim$ .

*Exemplo 22.* Considerando as seguintes afirmações do *Exemplo 20*:

- $\Delta^I = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$ ;
- $BL^I(x_1) = \mathbf{t}; BL^I(x_2) = \mathbf{f}; BL^I(x_3) = \mathbf{t}; BL^I(x_4) = \mathbf{f}; BL^I(x_5) = \mathbf{u}; BL^I(x_6) = \mathbf{t}; BL^I(x_7) = \mathbf{f}$ ;
- $P^I(x_1) = \mathbf{t}; P^I(x_2) = \mathbf{i}; P^I(x_3) = \mathbf{f}; P^I(x_4) = \mathbf{t}; P^I(x_5) = \mathbf{u}; P^I(x_6) = \mathbf{i}; P^I(x_7) = \mathbf{t}$ ;
- $L^I(x_1) = \mathbf{t}; L^I(x_2) = \mathbf{f}; L^I(x_3) = \mathbf{u}; L^I(x_4) = \mathbf{f}; L^I(x_5) = \mathbf{t}; L^I(x_6) = \mathbf{t}; L^I(x_7) = \mathbf{t}$ ;
- $M^I(x_1) = \mathbf{f}; M^I(x_2) = \mathbf{f}; M^I(x_3) = \mathbf{t}; M^I(x_4) = \mathbf{t}; M^I(x_5) = \mathbf{t}; M^I(x_6) = \mathbf{f}; M^I(x_7) = \mathbf{t}$ ;

Seja o contexto  $\Sigma = \{BL, P, L, M\}$ . A função  $\text{DIS}(\Sigma, x_1, x_3, R) = \{P, L, M\}$  diz que  $x_1$  e  $x_3$  não são equivalentes (relação R) com respeito aos conceitos  $P$ ,  $L$  e  $M$ . Enquanto que  $\text{DIS}(\Sigma, x_1, x_3, S) = \{P, M\}$  diz que  $x_1$  e  $x_3$  não são similares (relação S) com respeito aos conceitos  $P$  e  $M$ .

É importante destacar aqui que não iremos trabalhar diretamente com a noção de reduto. Em DL, um reduto seria calculado levando em consideração todas as informações presentes em uma base de conhecimento, i.e. todas as relações entre os indivíduos da base de conhecimento. Nosso problema consiste em encontrar contextos para asserções de conceitos (e.g.  $C(a)$ ), logo não é necessário ter uma noção geral de todas as informações da base de conhecimento. Precisamos apenas dos relacionamentos entre o indivíduo mapeado da asserção (e.g.  $a^I = x$ ) com os outros indivíduos (e.g. as similaridades  $\text{Sim}(x, y)$  para todo  $y \in \Delta^I$ ).

Para representar essa ideia, utilizaremos um tipo de reduto que leva em consideração um conceito e um indivíduo específico, chamado de reduto relativo (STEPANIUK, 1998). Um reduto relativo é calculado através de uma outra função de dissimilaridade obtida através de algumas mudanças na função de discernibilidade (vide Definição 7).

**Definição 50 (Função de Dissimilaridade -  $f$ )** (STEPANIUK, 1998) *A função de dissimilaridade de um objeto  $x \in U$  e uma relação de similaridade  $\text{Sim}$  em uma tabela de decisão  $D = (U, A, d, V_A, V_d)$  com  $n$  objetos é definida por:*

$$f(D, \text{Sim}, x) = \bigwedge_{y \in U, d(x) \neq d(y)} (\bigvee D_{xy}),$$

em que  $D_{xy} = \{a \in A \mid y \notin \text{Sim}_{\{a\}}(x)\}$ , tal que  $\text{Sim}_{\{a\}}(x) = \{y \in U \mid (x, y) \in \text{Sim}_{\{a\}}\}$ .

A primeira diferença com respeito à função de discernibilidade é o cálculo da matriz de dissimilaridade  $D_{xy}$ . A matriz de dissimilaridade segue o mesmo cálculo da matriz de discernibilidade, mas em vez de utilizar uma relação de discernibilidade, é utilizada uma relação de dissimilaridade (vide Definição 7). Outra diferença é que a matriz de dissimilaridade de um reduto relativo não precisa ser calculada completamente; apenas as linhas correspondentes ao objeto  $x$  são necessárias. Além disso, apenas aqueles objetos que apresentam valor de decisão diferente de  $x$  precisam ser levados em consideração. De acordo com (STEPANIUK, 1998), o resultado da função de dissimilaridade de  $x$  corresponde ao conjunto de redutos relativos de  $x$  em uma tabela de decisão. O reduto é chamado de relativo, pois é preciso olhar somente uma parte dos relacionamentos entre indivíduos (os relacionamentos de um indivíduo  $x$ ).

*Exemplo 23.* Adaptando as informações do *Exemplo 22*, temos a tabela de decisão  $D = (\{x_1, \dots, x_7\}, \{BL, P, L\}, M, \{\mathbf{t}, \mathbf{f}, \mathbf{i}, \mathbf{u}\}, \{\mathbf{t}, \mathbf{f}, \mathbf{i}, \mathbf{u}\})$  (Tabela 6.1). Se considerarmos a função de dissimilaridade  $f$  para a relação de similaridade S e o indivíduo  $x_3$  teremos o resultado

$$f(D, \text{Sim}, x_3) = \bigwedge_{y \in U, M(x_3) \neq M(y)} (\bigvee D_{x_3y})$$

	<i>BL</i>	<i>P</i>	<i>L</i>	<i>M</i>
$x_1$	<b>t</b>	<b>t</b>	<b>t</b>	<b>f</b>
$x_2$	<b>f</b>	<b>i</b>	<b>f</b>	<b>f</b>
$x_3$	<b>t</b>	<b>f</b>	<b>u</b>	<b>t</b>
$x_4$	<b>f</b>	<b>t</b>	<b>f</b>	<b>t</b>
$x_5$	<b>u</b>	<b>u</b>	<b>t</b>	<b>t</b>
$x_6$	<b>t</b>	<b>i</b>	<b>t</b>	<b>f</b>
$x_7$	<b>f</b>	<b>t</b>	<b>t</b>	<b>t</b>

Tabela 6.1: Tabela de Decisão  $D = (\{x_1, \dots, x_7\}, \{BL, P, L\}, M, \{t, f, i, u\}, \{t, f, i, u\})$

$$\begin{aligned}
&= (\bigvee D_{x_3x_1}) \wedge (\bigvee D_{x_3x_2}) \wedge (\bigvee D_{x_3x_6}) \\
&= P \wedge (BL \vee P) \wedge P \\
&= P.
\end{aligned}$$

Desta forma, um reduto relativo dos atributos  $BL, P$  e  $L$  para o atributo de decisão  $M$  e o indivíduo  $x_3$  é  $P$ . Isto quer dizer que se olharmos de acordo com o resultado de  $M$  e o indivíduo  $x_3$ , o atributo  $P$  já é suficiente para distinguir  $x_3$  dos demais indivíduos.

Iremos adaptar a definição da função de dissimilaridade  $f$  para *Rough  $\mathcal{ALC}_4$*  e  $PR_{\mathcal{ALC}}$ , a fim de calcular o conjunto de redutos relativos de um indivíduo  $x$  com relação a um conceito  $C$ , os quais serão os contextos para restrições de consulta.

**Definição 51 (Função de Dissimilaridade - DL)** *A função de dissimilaridade de um indivíduo  $x$  com respeito a um contexto  $\Sigma$ , um conceito  $C$  e uma relação de similaridade  $\text{Sim}$  é definida por*

$$f(\Sigma, C, x, \text{Sim}) = \bigwedge_{y \in \Delta^I, \mathcal{A} \models C(x) \Leftrightarrow \mathcal{A} \not\models C(y)} (\bigvee \text{DIS}(\Sigma, x, y, \text{Sim})).$$

Pela Proposição 6, iremos calcular a interseção de todas as dissimilaridades do indivíduo  $x$ , porque estamos interessados apenas em contextos minimais.

*Exemplo 24.* Considere o ABox  $\mathcal{A}$  abaixo criado a partir do *Exemplo 20*, em que  $\Delta^I = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$ .

*BoaLocalização*( $x_1$ );  $\neg$ *BoaLocalização*( $x_2$ ); *BoaLocalização*( $x_3$ );  $\neg$ *BoaLocalização*( $x_4$ );

*BoaLocalização*( $x_6$ );  $\neg$ *BoaLocalização*( $x_7$ );

*Porão*( $x_1$ ); *Porão*( $x_2$ );  $\neg$ *Porão*( $x_2$ );  $\neg$ *Porão*( $x_3$ ); *Porão*( $x_4$ ); *Porão*( $x_6$ );  $\neg$ *Porão*( $x_6$ ); *Porão*( $x_7$ );

*Lareira*( $x_1$ );  $\neg$ *Lareira*( $x_2$ );  $\neg$ *Lareira*( $x_4$ ); *Lareira*( $x_5$ ); *Lareira*( $x_6$ ); *Lareira*( $x_7$ );

$\neg$ *Médio*( $x_1$ );  $\neg$ *Médio*( $x_2$ ); *Médio*( $x_3$ ); *Médio*( $x_4$ ); *Médio*( $x_5$ );  $\neg$ *Médio*( $x_6$ ); *Médio*( $x_7$ ).

Um resultado obtido é que  $\mathcal{A} \models \text{Médio}(x_3)$ . Para saber se  $x_3$  necessariamente tem a propriedade de valor médio, podemos aplicar  $f(\Sigma, \text{Médio}, x_3, S)$ , em que  $\Sigma = \{\text{BoaLocalização}, \text{Porão}, \text{Lareira}\}$  e  $S$  a relação de similaridade com informação desconhecida:

$$\begin{aligned}
f(\Sigma, \text{Médio}, x_3, S) &= \bigwedge_{y \in \Delta^I, \mathcal{A} \models \text{Médio}(x) \Leftrightarrow \mathcal{A} \not\models \text{Médio}(y)} (\bigvee \text{DIS}(\Sigma, x_3, y, S)). \\
&= (\bigvee \text{DIS}(\Sigma, x_3, x_1, S)) \wedge (\bigvee \text{DIS}(\Sigma, x_3, x_2, S)) \wedge (\bigvee \text{DIS}(\Sigma, x_3, x_6, S)). \\
&= \text{Porão} \wedge (\text{BoaLocalização} \vee \text{Porão}) \wedge \text{Porão}. \\
&= \text{Porão}.
\end{aligned}$$

Segue-se que  $\mathcal{A} \models \underline{\text{Médio}}_{(\Sigma_1, 1)}(x_3)$ , em que  $\Sigma_1 = \{\text{Porão}\}$ . Esse resultado diz que o contexto *Porão* já é suficiente para ser usado na aproximação inferior de *Médio*. De fato, podemos ter que  $\mathcal{A} \models \underline{\text{Médio}}_{(\Sigma, 1)}(x_3)$  é também satisfeito, pois como mencionado na Proposição 6,  $\Sigma_1 \subseteq \Sigma$  implica que  $\underline{\text{Médio}}_{\Sigma_1}(x_3) \sqsubseteq \underline{\text{Médio}}_{\Sigma}(x_3)$ . Se consultarmos a função de dissimilaridade com a relação R, teremos como resultado

$$\begin{aligned}
f(\Sigma, \text{Médio}, x_3, R) &= \bigwedge_{y \in \Delta^I, \mathcal{A} \models \text{Médio}(x) \Leftrightarrow \mathcal{A} \not\models \text{Médio}(y)} (\bigvee \text{DIS}(\Sigma, x_3, y, R)). \\
&= (\bigvee \text{DIS}(\Sigma, x_3, x_1, R)) \wedge (\bigvee \text{DIS}(\Sigma, x_3, x_2, R)) \wedge (\bigvee \text{DIS}(\Sigma, x_3, x_6, R)). \\
&= (\text{Porão} \vee \text{Lareira}) \wedge (\text{BoaLocalização} \vee \text{Porão} \vee \text{Lareira}) \wedge \\
&\quad (\text{Porão} \vee \text{Lareira}). \\
&= (\text{Porão} \vee \text{Lareira}).
\end{aligned}$$

O contexto  $\Sigma_2 = \{\text{Porão}, \text{Lareira}\}$  satisfaz  $\mathcal{A} \models \underline{\text{Médio}}_{(\Sigma_2, 1)}(x_3)$ . Comparando com a relação de similaridade S podemos perceber que o contexto  $\Sigma_2$  apresenta similaridades com informação desconhecida, já que ele não aparece em  $f(\Sigma, \text{Médio}, x_3, S)$ , mas é encontrado em  $f(\Sigma, \text{Médio}, x_3, R)$ . Ao aplicarmos a função de dissimilaridade com a relação P teremos que

$$\begin{aligned}
f(\Sigma, \text{Médio}, x_3, P) &= \bigwedge_{y \in \Delta^I, \mathcal{A} \models \text{Médio}(x) \Leftrightarrow \mathcal{A} \not\models \text{Médio}(y)} (\bigvee \text{DIS}(\Sigma, x_3, y, P)). \\
&= (\bigvee \text{DIS}(\Sigma, x_3, x_1, P)) \wedge (\bigvee \text{DIS}(\Sigma, x_3, x_2, P)) \wedge (\bigvee \text{DIS}(\Sigma, x_3, x_6, P)). \\
&= \text{Porão} \wedge \text{BoaLocalização} \wedge \emptyset \\
&= \emptyset.
\end{aligned}$$

O resultado vazio mostra que não existe contexto que satisfaça a aproximação inferior de *Médio* para  $x_3$  e a similaridade P. Intuitivamente, podemos dizer que ao permitirmos similaridades do valores de  $x_3$  com informações contraditórias presentes em P, haverá um indivíduo similar a  $x_3$  que não satisfaz o conceito *Médio*; nesse caso o indivíduo  $x_6$ .

### 6.3 Contextos para Aproximações Superiores

O segundo problema pode ser formulado como segue:

- **Entrada.** O conjunto universo  $\Delta^I$ , o conjunto de nomes de conceitos  $N_C$ , um ABox  $\mathcal{A}$ , uma relação de similaridade  $\text{Sim}$  e um conceito  $C$ .
- **Saída.** Contextos  $\Sigma \subseteq N_C - \text{atom}(C)$  e um número natural  $n$  tal que  $\mathcal{A} \models \bar{C}^{(\text{Sim}_\Sigma, n)}$ .

Diferente da aproximação inferior, utilizaremos agora a ideia de função de similaridade em vez de função de dissimilaridade. A motivação por trás da busca por um contexto para aproximação superior vem da seguinte ideia: quando um axioma de asserção não é satisfeito por uma base de conhecimento, nós precisamos encontrar indivíduos que satisfaçam a asserção do conceito consultado e que apresentem algumas similaridades com o indivíduo da consulta original. Essas similaridades irão caracterizar o contexto. Por isso iremos considerar agora função de similaridade.

**Definição 52 (Função de Similaridade - SIM)** *Sejam  $\Delta^I$  o conjunto universo,  $\Sigma$  um contexto e  $\text{Sim}$  uma relação de similaridade. A função de similaridade  $\text{SIM}(\Sigma, x, y, \text{Sim})$  é definida como*

$$\text{SIM}(\Sigma, x, y, \text{Sim}) = \{A_i \in \Sigma \mid y \in \text{Sim}_{\{A_i\}}(x)\}.$$

A função  $\text{SIM}(\Sigma, x, y, \text{Sim})$  descreve o conjunto de todos os conceitos atômicos em  $\Sigma$ , em que o indivíduo  $x$  é similar a  $y$  com relação a  $\text{Sim}$ . A partir da função de similaridade podemos construir a função de similaridade, que será responsável por obter os contextos para os relaxamentos de consulta.

**Definição 53 (Função de Similaridade - g - DL)** *A função de similaridade de um indivíduo  $x$  com respeito a um contexto  $\Sigma$ , um conceito  $C$  e uma relação de similaridade  $\text{Sim}$  é definida por*

$$g(\Sigma, C, x, \text{Sim}) = \bigvee_{y \in \Delta^I, \mathcal{A} \not\models C(x) \Leftrightarrow \mathcal{A} \models C(y)} (\bigwedge \text{SIM}(\Sigma, x, y, \text{Sim})).$$

Ao contrário da aproximação inferior, não estamos interessados em encontrar o conjunto minimal de conceitos. Pela Proposição 6, queremos encontrar os conjuntos maximais de conceitos, por isso é realizada a disjunção de todas as similaridades de  $x$ . De um outro ponto de vista, podemos ver um relaxamento de consulta como uma operação dual à restrição de consulta, logo a função  $g(\Sigma, C, x, \text{Sim})$  é uma espécie de função dual de  $f(\Sigma, C, x, \text{Sim})$ .

*Exemplo 25.* Considere agora o ABox  $\mathcal{A}^I$  abaixo criado a partir do *Exemplo 20*, em que  $\Delta^I = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$ .

$$\begin{aligned}
& \text{BoaLocalização}(x_1); \neg \text{BoaLocalização}(x_2); \text{BoaLocalização}(x_3); \neg \text{BoaLocalização}(x_4); \\
& \quad \text{BoaLocalização}(x_6); \neg \text{BoaLocalização}(x_7); \\
& \text{Porão}(x_1); \text{Porão}(x_2); \neg \text{Porão}(x_2); \neg \text{Porão}(x_3); \text{Porão}(x_4); \text{Porão}(x_6); \neg \text{Porão}(x_6); \text{Porão}(x_7); \\
& \quad \text{Lareira}(x_1); \neg \text{Lareira}(x_2); \neg \text{Lareira}(x_4); \text{Lareira}(x_5); \text{Lareira}(x_6); \text{Lareira}(x_7); \\
& \quad \neg \text{Caro}(x_1); \neg \text{Caro}(x_2); \neg \text{Caro}(x_3); \neg \text{Caro}(x_4); \neg \text{Caro}(x_5); \text{Caro}(x_6); \neg \text{Caro}(x_7).
\end{aligned}$$

Temos que  $\mathcal{A}' \not\models \text{Caro}(x_7)$ . Nós podemos aplicar a função  $g(\Sigma, \text{Caro}, x_7, \text{S})$ , em que  $\Sigma = N_C - \{\text{Caro}\}$ , para descobrir se existe um contexto satisfazendo a aproximação superior de *Caro* de acordo com  $x_7$ :

$$\begin{aligned}
g(\Sigma, \text{Caro}, x_7, \text{S}) &= \bigvee_{y \in \Delta^I, \mathcal{A}' \not\models \text{Caro}(x) \Leftrightarrow \mathcal{A}' \models \text{Caro}(y)} (\bigwedge \text{SIM}(\Sigma, x_7, y, \text{S})). \\
&= (\bigwedge \text{SIM}(\Sigma, x_7, x_1, \text{S})) \vee (\bigwedge \text{SIM}(\Sigma, x_7, x_6, \text{S})). \\
&= (\text{Porão} \wedge \text{Lareira}) \vee \text{Lareira}. \\
&= (\text{Porão} \wedge \text{Lareira}).
\end{aligned}$$

Portanto, temos  $\mathcal{A}' \models \overline{\text{Caro}}^{(\Sigma_1, 1)}(x_7)$ , em que  $\Sigma_1 = \{\text{Porão}, \text{Lareira}\}$ . Nós escolhemos  $\Sigma_1$  como o resultado da simplificação, pois pelo Lema 6, todos subconjuntos não-vazios  $\Sigma_2$  de  $\Sigma_1$  também satisfazem a aproximação superior de *Caro*( $x_7$ ) (por exemplo *Lareira*), isto é  $\mathcal{A}' \models \overline{\text{Caro}}^{(\Sigma_2, 1)}(x_7)$ . Consideramos  $\Sigma_1$  um contexto otimizado devido ele cobrir uma quantidade maior de conceitos para a aproximação superior. Para a função de similaridade com a relação R, temos o resultado

$$\begin{aligned}
g(\Sigma, \text{Caro}, x_7, \text{R}) &= \bigvee_{y \in \Delta^I, \mathcal{A}' \not\models \text{Caro}(x) \Leftrightarrow \mathcal{A}' \models \text{Caro}(y)} (\bigwedge \text{SIM}(\Sigma, x_7, y, \text{R})). \\
&= (\bigwedge \text{SIM}(\Sigma, x_7, x_1, \text{R})) \vee (\bigwedge \text{SIM}(\Sigma, x_7, x_6, \text{R})). \\
&= (\text{Porão} \wedge \text{Lareira}) \vee \text{Lareira}. \\
&= (\text{Porão} \wedge \text{Lareira}).
\end{aligned}$$

Nesse exemplo, obtemos os mesmos resultados para as relações R e S. Nesse caso, podemos concluir que não existe a presença de informação desconhecida relacionada com o indivíduo  $x_7$ . Para a função de similaridade com a relação P temos

$$\begin{aligned}
g(\Sigma, \text{Caro}, x_7, \text{P}) &= \bigvee_{y \in \Delta^I, \mathcal{A}' \not\models \text{Caro}(x) \Leftrightarrow \mathcal{A}' \models \text{Caro}(y)} (\bigwedge \text{SIM}(\Sigma, x_7, y, \text{P})). \\
&= (\bigwedge \text{SIM}(\Sigma, x_7, x_1, \text{P})) \vee (\bigwedge \text{SIM}(\Sigma, x_7, x_6, \text{P})).
\end{aligned}$$

$$\begin{aligned}
&= (Por\tilde{a}o \wedge Lareira) \vee (Por\tilde{a}o \wedge Lareira). \\
&= (Por\tilde{a}o \wedge Lareira).
\end{aligned}$$

O resultado é o mesmo obtido com as relações R e S. Porém, diferente das outras duas relações, o resultado de  $(\wedge \text{SIM}(\Sigma, x_7, x_6, P)) = (Por\tilde{a}o \wedge Lareira)$ . Isso indica que existem evidências de similaridades com contradições presentes em P, mas nesse caso essas evidências são redundantes, pois não alteraram o valor do resultado original.

#### 6.4 Algoritmos de Refinamento de Consultas

Na sequência apresentaremos os algoritmos responsáveis por encontrar contextos para refinamento de consultas. Eles buscam por conjuntos minimais de conceitos que satisfaçam a aproximação inferior de uma asserção (para restrições de consulta) ou conjuntos maximais de conceitos que satisfaçam a aproximação superior de uma asserção (para relaxamentos de consulta) de acordo com uma relação de similaridade. Se nenhum resultado for encontrado, então o processo é repetido levando em consideração uma relação de similaridade de um passo maior. Os algoritmos terminam quando algum contexto é encontrado (o qual será a solução do problema) ou quando eles procuram em todo o fecho transitivo da relação de similaridade e nenhum resultado é retornado (neste caso, um conjunto vazio será a solução do problema). A seguir, explicaremos detalhadamente os algoritmos assim como seus resultados de complexidade.

---

##### Algoritmo 1 SimplificaFND( $S, n$ )

---

```

1:  $i \leftarrow 1$ 
2: while  $i \leq n$  do
3:    $j \leftarrow i + 1$ 
4:   while  $j \leq n$  do
5:     if  $S_i \subseteq S_j$  then
6:        $S \leftarrow \text{Extrair}(S, S_j)$ 
7:        $j \leftarrow j + 1$ 
8:     else
9:       if  $S_j \subseteq S_i$  then
10:         $S \leftarrow \text{Extrair}(S, S_i)$ 
11:         $i \leftarrow i + 1$ 
12:         $j \leftarrow i + 1$ 
13:      else
14:         $j \leftarrow j + 1$ 
15:      end if
16:    end if
17:  end while
18:   $i \leftarrow i + 1$ 
19: end while
20: retorna  $S$ 

```

---

Assumimos no problema que uma fórmula está representada na Forma Normal Conjuntiva (FNC) (se ela é entrada do **Algoritmo 2**) ou na Forma Normal Disjuntiva (FND) (se ela é entrada do **Algoritmo 1**). Por exemplo,  $S = \{\{A_1, A_2\}, \{A_2, A_3\}\}$  pode ser tratado como  $S = (A_1 \wedge A_2) \vee (A_2 \wedge A_3)$  (se ela é entrada do **Algoritmo 1**) ou como  $S = (A_1 \vee A_2) \wedge (A_2 \vee A_3)$

(se ela é entrada do **Algoritmo 2**). **Algoritmo 1** simplifica uma fórmula  $S$  na FND removendo cláusulas redundantes. Esse procedimento é realizado empregando a lei de absorção, i.e.,  $(a \wedge b) \vee a \equiv a$ , que é aplicada na função *Extrair* (linhas 6 e 10). Apenas essa lei é necessária para a simplificação, porque a entrada do algoritmo contém apenas conceitos atômicos (aqueles em  $N_C$ ). A função *Extrair* consiste simplesmente em retirar uma cláusula dentro de um conjunto de cláusulas.

A função *SimplificaFND2*( $S, n$ ) (**Algoritmo 2**) segue a mesma ideia do **Algoritmo 1**, mas em vez da lei de absorção, é aplicada a regra  $(a \wedge b) \vee a \equiv (a \wedge b)$  para representar contextos maximais. A função *SimplificaFND2*( $S, n$ ) pode ser obtida pela troca das linhas 5 e 9 do **Algoritmo 1**.

---

**Algoritmo 2** *SimplificaFND2*( $S, n$ )

---

```

1:  $i \leftarrow 1$ 
2: while  $i \leq n$  do
3:    $j \leftarrow i + 1$ 
4:   while  $j \leq n$  do
5:     if  $S_j \subseteq S_i$  then
6:        $S \leftarrow \text{Extrair}(S, S_j)$ 
7:        $j \leftarrow j + 1$ 
8:     else
9:       if  $S_i \subseteq S_j$  then
10:         $S \leftarrow \text{Extrair}(S, S_i)$ 
11:         $i \leftarrow i + 1$ 
12:         $j \leftarrow i + 1$ 
13:      else
14:         $j \leftarrow j + 1$ 
15:      end if
16:    end if
17:  end while
18:   $i \leftarrow i + 1$ 
19: end while
20: retorna  $S$ 

```

---

**Teorema 9** *Os algoritmos *SimplificaFND*( $S, n$ ) e *SimplificaFND2*( $S, n$ ) têm complexidades de tempo  $O(|N_C|.n^2)$ .*

*Prova.* Os algoritmos são compostos por dois laços aninhados do tipo **while**, em que cada laço é limitado superiormente por  $n$  (logo,  $O(n^2)$ ), em que  $n$  é o número de cláusulas da fórmula na forma normal disjuntiva recebida como entrada. A complexidade da função *Extrair* tem limite superior linear no tamanho da fórmula, ou seja,  $O(|N_C|)$ , em que  $N_C$  é o conjunto de todos os conceitos atômicos. Logo, a complexidade dos algoritmos *SimplificaFND*( $S, n$ ) e *SimplificaFND2*( $S, n$ ) é de  $O(|N_C|.n^2)$ .  $\square$

**Teorema 10** *O algoritmo *SimplificaFNC*( $S, n$ ) tem complexidade de tempo  $O(2^{|N_C|})$ .*

*Prova.* **Algoritmo 3** simplifica uma fórmula  $S$  na FNC removendo cláusulas redundantes, similarmente ao **Algoritmo 1**. Isso é alcançado pela tradução de uma fórmula em FNC para uma

fórmula em FND via  $\text{Distribuir}(S)$ . Essa tradução é necessária pois o conjunto de todos os re-  
 dutos é obtido através de um fórmula na FND reduzida (vide Seção 2.4). A complexidade do  
 algoritmo que traduz FNC para FND é de  $O(2^n)$  (MILTERSEN; RADHAKRISHNAN; WE-  
 GENER, 2005), em que  $n$  é o número de diferentes variáveis da fórmula em FNC. O algoritmo  
 procede então com a simplificação da fórmula em FND através da função  $\text{SimplificaFND}(S, n)$ .  
 A complexidade da linha 1 do **Algoritmo 3** é  $O(2^{|N_C|})$ , em que  $|N_C|$  é a quantidade de conceitos  
 atômicos e o limite de variáveis que podem aparecer na fórmula em FNC. A complexidade da  
 linha 2 é de  $O(|N_C|.n^2)$  como mostrado anteriormente. Portanto, a complexidade do **Algoritmo**  
**3** é de  $O(2^{|N_C|})$ .  $\square$

---

**Algoritmo 3**  $\text{SimplificaFNC}(S, n)$ 


---

1:  $A \leftarrow \text{Distribuir}(S)$   
 2: **retorna**  $\text{SimplificaFND}(S, n)$

---

**Algoritmo 4** encontra conjuntos de contextos para aproximação inferior de um axi-  
 oma de asserção. Em outras palavras, ele implementa a função de dissimilaridade  $f(\Sigma, C, x, \text{Sim})$ .  
 Primeiro, é calculada a matriz de dissimilaridade de um indivíduo específico com relação a uma  
 relação de similaridade  $\text{Sim}$  (linhas 8-18). Logo após, é calculada a função *booleana* para  
 aproximação inferior (a função de dissimilaridade) através da função  $\text{SimplificaFNC}$  (linha 19),  
 que tem o objetivo de procurar por conjuntos minimais de conceitos.

---

**Algoritmo 4**  $f(N_C, C, x, \text{Sim})$ 


---

1:  $j \leftarrow 1$   
 2:  $\Sigma_1 \leftarrow \emptyset$   
 3:  $\Sigma \leftarrow N_C - \{C\}$   
 4:  $\text{Sim}_{0\Sigma} \leftarrow \emptyset$   
 5: **while**  $\Sigma_1 = \emptyset$  e  $\text{Sim}_{j-1\Sigma} \neq \text{Sim}_{j\Sigma}$  **do**  
 6:      $S \leftarrow \emptyset$   
 7:      $i \leftarrow 1$   
 8:     **for all**  $y \in \Delta^I$  **do**  
 9:         **if**  $\mathcal{A} \models C(x) \Leftrightarrow \mathcal{A} \not\models C(y)$  **then**  
 10:              $S_i \leftarrow \text{DIS}(N_C - \{C\}, x, y, \text{Sim}_j)$   
 11:             **if**  $S_i = \emptyset$  **then**  
 12:                 **retorna**  $\emptyset$   
 13:             **else**  
 14:                  $S \leftarrow S \cup \{S_i\}$   
 15:                  $i \leftarrow i + 1$   
 16:             **end if**  
 17:     **end if**  
 18:     **end for**  
 19:      $\Sigma_1 \leftarrow \text{SimplificaFNC}(S, i - 1)$   
 20:      $j \leftarrow j + 1$   
 21: **end while**  
 22: **retorna**  $\Sigma_1, j - 1$

---

Se o resultado é não-vazio, então ele será a solução do problema. Caso contrário, o  
 procedimento é reiniciado com uma relação de um passo maior. Ao encontrar uma relação de  
 similaridade de passo  $k$  que retorna um conjunto não-vazio de conceitos, esse conjunto será o  
 resultado do problema. Se por acaso, todas as relações de similaridade de passo  $k$  retornarem  
 conjuntos vazios, o resultado do problema será então vazio.

**Teorema 11** *O algoritmo  $f(N_C, C, x, \text{Sim})$  tem complexidade EXPTIME em Rough  $\mathcal{ALC}_4$  e complexidade EXPSPACE em  $PR_{\mathcal{ALC}}$ .*

*Prova.* A complexidade do laço **for all** (linhas 8-18) leva em consideração a complexidade da consequência lógica de uma asserção de conceito em *Rough  $\mathcal{ALC}_4$*  ou  $PR_{\mathcal{ALC}}$  ( $\mathcal{A} \models C(x)$ ) e o cálculo da matriz de dissimilaridade DIS. Por definição, DIS também depende do cálculo da consequência lógica de asserções de conceitos, ou seja, é PSPACE para *Rough  $\mathcal{ALC}_4$*  (vide Seções 3.1.4 e 5.2) e EXPSPACE para a  $PR_{\mathcal{ALC}}$  (vide Seção 5.4.2). Logo, a complexidade do laço das linhas 8-18 é PSPACE para *Rough  $\mathcal{ALC}_4$*  e EXPSPACE para  $PR_{\mathcal{ALC}}$ . Como apontado na Proposição 5, as aproximações *tight* e *loose* são monotônicas, e como estamos lidando com um conjunto finito de indivíduos, existe um número finito de *relações-passo-k* para esse problema, que é limitado por  $|\Delta^I|$ . Portanto, a quantidade de comparações realizadas através do **while** (linha 5) é de  $O(|\Delta^I|)$ . Por fim, temos que a  $\text{SimplificaFNC}(S, n)$  tem complexidade de tempo  $O(2^{|N_C|})$ .  $\square$

**Algoritmo 5** segue os princípios do **Algoritmo 4**, mas ele constrói uma matriz de similaridade. Após isso, o algoritmo cria uma função *booleana* para aproximação superior (a função de similaridade) através de  $\text{SimplificaFND2}$ , que resulta em conjuntos maximais de conceitos. Assim como o **Algoritmo 4**, o **Algoritmo 5** termina quando uma *relação-passo-k* retorna um conjunto não-vazio de conceitos.

---

**Algoritmo 5**  $g(N_C, C, x, \text{Sim})$

---

```

1:  $j \leftarrow 1$ 
2:  $\Sigma_1 \leftarrow \emptyset$ 
3:  $\Sigma \leftarrow N_C - \{C\}$ 
4:  $\text{Sim}_{0\Sigma} \leftarrow \emptyset$ 
5: while  $\Sigma_1 = \emptyset$  e  $\text{Sim}_{j-1\Sigma} \neq \text{Sim}_{j\Sigma}$  do
6:    $S \leftarrow \emptyset$ 
7:    $i \leftarrow 1$ 
8:   for all  $y \in \Delta^I$  do
9:     if  $\mathcal{A} \not\models C(x) \Leftrightarrow \mathcal{A} \models C(y)$  then
10:        $S_i \leftarrow \text{SIM}(N_C - \{C\}, x, y, \text{Sim}_j)$ 
11:       if  $S_i \neq \emptyset$  then
12:          $S \leftarrow S \cup \{S_i\}$ 
13:          $i \leftarrow i + 1$ 
14:       end if
15:     end if
16:   end for
17:    $\Sigma_1 \leftarrow \text{SimplificaFND2}(S, i - 1)$ 
18:    $j \leftarrow j + 1$ 
19: end while
20: retorna  $\Sigma_1, j - 1$ 

```

---

**Teorema 12** *O algoritmo  $g(N_C, C, x, \text{Sim})$  tem complexidade PSPACE para Rough  $\mathcal{ALC}_4$  e complexidade EXPSPACE para  $PR_{\mathcal{ALC}}$ .*

*Prova.* Assim como no **Algoritmo 4**, a construção da matriz de similaridade SIM leva em consideração a complexidade da consequência lógica de uma asserção de conceito (linhas 8-14) e tem complexidade PSPACE para *Rough  $\mathcal{ALC}_4$*  e EXPSPACE para  $PR_{\mathcal{ALC}}$ . Da mesma

forma, o número de *relações-passo- $k$*  é limitado por  $O(|\Delta^I|)$  (linha 5). A complexidade de  $\text{SimplificaFND2}(S, n)$  é de ordem polinomial, porém a complexidade geral do **Algoritmo 5** é PSPACE para *Rough  $\mathcal{ALC}_4$*  e EXPSPACE para *PR $\mathcal{ALC}$* .  $\square$

Note que esses resultados foram obtidos devidos às complexidades das lógicas utilizadas e que nesse caso o problema de refinamento não é o responsável pelo aumento de complexidade. De fato, os **Algoritmos 1, 2 e 3** mostram que o problema de encontrar refinamentos em tabelas de decisão utilizando *rough sets* possui tempo exponencial em relação a entrada para a restrição de consulta e possui tempo polinomial em relação a entrada para o relaxamento de consulta.

## 7 MEDIDAS DE QUALIDADE PARA SELEÇÃO DE CONTEXTOS

Dado que os métodos propostos no Capítulo 6 retornam conjuntos maximais e minimais, pode ocorrer que um problema retorne mais de um contexto como resposta. Neste capítulo apresentaremos algumas medidas de qualidade que podem ser aplicadas às aproximações de conceitos em *Rough  $\mathcal{ALC}_4$*  e *PR $_{\mathcal{ALC}}$* . Essas medidas têm como intuito determinar os melhores contextos (relativos a critérios específicos) para serem usados nas aproximações.

O exemplo abaixo ilustra um caso em que dois contextos minimais são obtidos como resposta para uma restrição de uma consulta.

*Exemplo 26.* Sejam o universo  $\Delta^I = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$  e o ABox  $\mathcal{A}$ :

$$\begin{aligned} &A_1(x_1); \neg A_1(x_2); \neg A_1(x_3); \neg A_1(x_4); A_1(x_5); A_1(x_6); \neg A_1(x_7); \\ &\neg A_2(x_1); A_2(x_2); A_2(x_3); \neg A_2(x_4); \neg A_2(x_5); \neg A_2(x_6); A_2(x_7); \\ &A_3(x_1); A_3(x_2); A_3(x_3); \neg A_3(x_4); A_3(x_5); A_3(x_6); \neg A_3(x_7); \\ &C(x_1); C(x_2); C(x_3); \neg C(x_4); \neg C(x_5); \neg C(x_6); \neg C(x_7). \end{aligned}$$

Para facilitar o entendimento, podemos traduzir essas informações na Tabela 7.1. Suponha que queiramos obter um contexto adequado para a aproximação inferior do objeto  $x_2$  em relação ao conceito  $C$ . Aplicando a função de dissimilaridade  $f(\{A_1, A_2, A_3\}, C, x_2, \mathcal{R})$ , temos o seguinte resultado:

$$(A_2 \vee A_3) \wedge (A_1 \vee A_2) \wedge (A_1 \vee A_2) \wedge A_3 = (A_1 \vee A_2) \wedge A_3 = (A_1 \wedge A_3) \vee (A_2 \wedge A_3).$$

Em outras palavras,  $\Sigma_1 = \{A_1, A_3\}$  e  $\Sigma_2 = \{A_2, A_3\}$  são dois contextos válidos para o problema. Aparentemente não está claro qual o melhor contexto a ser utilizado. As medidas de qualidade surgiriam para auxiliar nessa escolha.

Como apontado no Capítulo 2, com a finalidade de medir numericamente a precisão de uma aproximação, Pawlak introduziu duas medidas quantitativas de precisão e imprecisão em (PAWLAK, 1991). A precisão de um subconjunto é definida como a razão das cardinalidades das aproximações inferior e superior do subconjunto, que é limitada por 0 e 1; a imprecisão de um subconjunto é calculada simplesmente como  $1 - x$ , em que  $x$  é a medida de precisão desse subconjunto. Em suma, imprecisão é vista como o complemento da precisão da aproximação. Uma medida de imprecisão de um subconjunto igual a 1 mostra que não temos conhecimento algum sobre ele; em contrapartida, se tal medida for igual a 0, o conhecimento sobre esse subconjunto será total. Desse ponto de vista, a medida de imprecisão é um importante indicador de incerteza e precisão associada a dado subconjunto.

Outra medida bastante utilizada na literatura dos *rough sets* para medir a qualidade das aproximações é a medida de granularidade (WEI-HUA; XIAO-YAN; WEN-XIU, 2009). Neste trabalho, as funções de entropia e co-entropia são utilizadas como medidas de granu-

	$A_1$	$A_2$	$A_3$	$C$
$x_1$	<b>t</b>	<b>f</b>	<b>t</b>	<b>t</b>
$x_2$	<b>f</b>	<b>t</b>	<b>t</b>	<b>t</b>
$x_3$	<b>f</b>	<b>t</b>	<b>t</b>	<b>t</b>
$x_4$	<b>f</b>	<b>f</b>	<b>f</b>	<b>f</b>
$x_5$	<b>t</b>	<b>f</b>	<b>t</b>	<b>f</b>
$x_6$	<b>t</b>	<b>f</b>	<b>t</b>	<b>f</b>
$x_7$	<b>f</b>	<b>t</b>	<b>f</b>	<b>f</b>

Tabela 7.1: Interpretação dos conceitos  $A_1, A_2, A_3$  e  $C$

laridade, e elas representam a quantidade de medida de incerteza associada aos elementos de conjuntos de partições.

A medida de precisão pode ser vista como a medida de qualidade da informação de um subconjunto de elementos enquanto que a medida de granularidade pode ser vista como a qualidade da informação que uma relação de similaridade apresenta ao particionar vários subconjuntos de elementos.

Ao longo desta seção, adaptaremos as medidas de precisão e granularidade da teoria dos *rough sets* para serem aplicadas em *Rough  $\mathcal{ALC}$*  e *PR $_{\mathcal{ALC}}$* . Por fim, utilizaremos uma medida composta de qualidade que tem como objetivo escolher o contexto mais representativo baseado nesses critérios de precisão e granularidade. Destacamos aqui que todas as medidas apresentadas durante este capítulo são aplicadas apenas aos contextos resultantes dos métodos propostos no Capítulo 6, e que todos os resultados apresentados são válidos exclusivamente com relação a esses contextos.

## 7.1 Medidas de Precisão e Imprecisão

Como estará justificado mais adiante, utilizaremos como base a medida de precisão  $\gamma_B(X)$  apresentada em (YAO, 2010), que é diferente da medida de precisão  $\gamma_B(X) = \frac{|POS_B(X)|}{|U|} = \frac{|X_B|}{|U|}$  de Pawlak (PAWLAK, 1982).

**Definição 54 (Medida de Precisão/Imprecisão)** (YAO, 2010) *Sejam um sistema de informação  $S = (U, A, V_A)$ ,  $X \subseteq U$  e  $B \subseteq A$ . As medidas de precisão e imprecisão de aproximações são dadas por  $\gamma_B(X)$  e  $\beta_B(X)$ , em que*

$$\gamma_B(X) = \frac{|POS_B(X)| + |NEG_B(X)|}{|U|} = \frac{|X_B| + |\neg X_B|}{|U|} \text{ e } \beta_B(X) = 1 - \gamma_B(X).$$

A diferença é que na medida de precisão de Pawlak, não se tem uma ideia geral de como se comporta a aproximação, pois apenas o conjunto  $X$  é considerado. Em relação à medida de precisão vista em Yao, são analisados tanto elementos em  $X$  como os que não estão em  $X$ . Em nosso trabalho, essa visão de diferentes conjuntos será mais adequada, pois estamos lidando com quatro partições distintas (conjuntos tetravalorados), sendo necessário

saber como os elementos de uma aproximação se comportam em relação a cada um desses conjuntos. Maiores detalhes serão dados ao longo da seção.

Assim como na medida de precisão de Pawlak, temos que  $\gamma_B(X) = 1$  se e somente se  $\underline{X}_B = \overline{X}^B$ . Já  $\gamma_B(X) = 0$  se e somente se  $\underline{X}_B = \emptyset$  e  $\overline{X}^B = U$ . Também temos os seguintes resultados de monotonicidade (respectivamente antimonicidade) para a medida de precisão (imprecisão) em relação ao tamanho do conjunto de atributos.

**Proposição 8** *Sejam  $S = (U, A, V_A)$ ,  $X \subseteq U$  e  $B_1, B_2 \subseteq A$ :*

- Se  $B_1 \subseteq B_2$ , então  $\gamma_{B_1}(X) \leq \gamma_{B_2}(X)$ ;
- Se  $B_1 \subseteq B_2$ , então  $\beta_{B_1}(X) \geq \beta_{B_2}(X)$ .

*Prova.* Iremos primeiro mostrar que se  $B_1 \subseteq B_2$ , então  $IND(B_2) \subseteq IND(B_1)$ . Por contradição, assumamos que  $B_1 \subseteq B_2$  e  $IND(B_2) \not\subseteq IND(B_1)$ . Por definição,  $IND(B_2) = \{(x, y) \in U \times U \mid \forall a \in B_2, a(x) = a(y)\} \supset \{(x, y) \in U \times U \mid \forall a \in B_1, a(x) = a(y)\} = IND(B_1)$ . Logo, existem  $x_1, x_2 \in U$  tal que  $(x_1, x_2) \in IND(B_2)$  e  $(x_1, x_2) \notin IND(B_1)$ . Ou seja,  $\forall a \in B_2, a(x_1) = a(x_2)$ . Como  $B_1 \subseteq B_2$ , temos que  $\forall a \in B_1, a(x_1) = a(x_2)$ . Portanto,  $(x_1, x_2) \in IND(B_1)$ . Contradição com  $(x_1, x_2) \notin IND(B_1)$ .

Pela definição de classe da equivalência  $[x]_B = \{y \in U \mid (x, y) \in IND(B)\}$ , temos facilmente que se  $B_1 \subseteq B_2$ , então  $[x]_{B_2} \subseteq [x]_{B_1}$ . Pela definição de  $\gamma_{B_1}(X)$ , temos

$$\gamma_{B_1}(X) = \frac{|POS_{B_1}(X)| + |NEG_{B_1}(X)|}{|U|} = \frac{|\underline{X}_{B_1}| + |\overline{X}_{B_1}|}{|U|}.$$

Como  $B_1 \subseteq B_2$ ,  $\underline{X}_{B_1} = \{x \in U \mid [x]_{B_1} \subseteq X\} \subseteq \{x \in U \mid [x]_{B_2} \subseteq X\} = \underline{X}_{B_2}$ . O mesmo é válido para o conjunto  $\neg X$ , i.e.  $\overline{X}_{B_1} \subseteq \overline{X}_{B_2}$ . Logo,  $|\underline{X}_{B_1}| \leq |\underline{X}_{B_2}|$ ,  $|\overline{X}_{B_1}| \leq |\overline{X}_{B_2}|$  e

$$\gamma_{B_1}(X) = \frac{|\underline{X}_{B_1}| + |\overline{X}_{B_1}|}{|U|} \leq \frac{|\underline{X}_{B_2}| + |\overline{X}_{B_2}|}{|U|} = \gamma_{B_2}(X).$$

Consequentemente, Se  $B_1 \subseteq B_2$ , então  $1 - \gamma_{B_1}(X) \geq 1 - \gamma_{B_2}(X)$ , i.e.  $\beta_{B_1}(X) \geq \beta_{B_2}(X)$ .  $\square$

Com relação a conjuntos elementares, temos que cada um deles pode ser dividido em duas partições: a dos elementos que pertencem ao conjunto e a dos elementos que não pertencem. Para utilizar as medidas de precisão e imprecisão nos conceitos de *Rough  $\mathcal{ALC}$*  e *PR $\mathcal{ALC}$* , nós precisaremos dividir cada conceito em quatro partições distintas, que serão correspondentes aos valores verdade **t, f, i** e **u**.

**Definição 55 (Transformação de um Conceito Tetravalorado)** *Sejam  $C$  um conceito tetravalorado,  $I$  uma interpretação e  $x$  um indivíduo.*

- Se  $C$  é um conceito de  $\mathcal{ALC}_4$ , então

- $x \in C_t^I$  sse  $x \in \text{proj}^+(C^I)$  e  $x \notin \text{proj}^-(C^I)$ ;
- $x \in C_f^I$  sse  $x \notin \text{proj}^+(C^I)$  e  $x \in \text{proj}^-(C^I)$ ;
- $x \in C_i^I$  sse  $x \in \text{proj}^+(C^I)$  e  $x \in \text{proj}^-(C^I)$ ;
- $x \in C_u^I$  sse  $x \notin \text{proj}^+(C^I)$  e  $x \notin \text{proj}^-(C^I)$ .

• Se  $C$  é um conceito de  $PR_{\mathcal{ALC}}$ , então

- $x \in C_t^I$  sse  $C^I(x) = t$ ;
- $x \in C_f^I$  sse  $C^I(x) = f$ ;
- $x \in C_i^I$  sse  $C^I(x) = i$ ;
- $x \in C_u^I$  sse  $C^I(x) = u$ .

*Exemplo 27.* Dados o domínio  $\Delta^I = \{x_1, x_2, x_3, x_4\}$  e a interpretação do conceito  $C^I = \langle \{x_1, x_3\}, \{x_2, x_3\} \rangle$  em  $\mathcal{ALC}_4$ , podemos transformar  $C$  nos seguintes conceitos:  $C_t = \{x_1\}$ ,  $C_f = \{x_2\}$ ,  $C_i = \{x_3\}$  e  $C_u = \{x_4\}$ .

Na sequência, introduziremos para *Rough*  $\mathcal{ALC}$  e  $PR_{\mathcal{ALC}}$  as medidas de precisão e imprecisão que serão extensões de  $\gamma_B(X)$  e  $\beta_B(X)$  considerando agora as quatro partições geradas por um conceito  $C$ .

**Definição 56 (Medida de Precisão/Imprecisão - DL)** *Sejam  $\Sigma$  um contexto,  $C$  um conceito,  $\text{Sim}$  uma relação de similaridade e  $I$  uma interpretação. As medidas de precisão e imprecisão em  $\mathcal{ALC}_4$  e  $PR_{\mathcal{ALC}}$  são denotadas por  $\gamma(\Sigma, C, \text{Sim}, I)$  e  $\beta(\Sigma, C, \text{Sim}, I)$ , em que:*

- $\gamma(\Sigma, C, \text{Sim}, I) = \frac{|(C_t)_{\text{Sim}_\Sigma}^I| + |(C_f)_{\text{Sim}_\Sigma}^I| + |(C_i)_{\text{Sim}_\Sigma}^I| + |(C_u)_{\text{Sim}_\Sigma}^I|}{|\Delta^I|}$ ;
- $\beta(\Sigma, C, \text{Sim}, I) = 1 - \gamma(\Sigma, C, \text{Sim}, I)$ .

Note que ao transformar um conceito tetravalorado, estamos dividindo-o em quatro conceitos de  $\mathcal{ALC}$ , e que a aproximação inferior de um conceito  $C$  em  $\mathcal{ALC}$  com relação a um contexto  $\Sigma$  e uma relação de similaridade  $\text{Sim}$  é dada por  $(C_{\text{Sim}_\Sigma})^I = \{x \in \Delta^I \mid \forall y((x, y) \in \text{Sim}_\Sigma \rightarrow y \in C^I)\}$ . Claramente  $0 \leq \gamma(\Sigma, C, \text{Sim}, I)$  e  $\beta(\Sigma, C, \text{Sim}, I) \leq 1$ . Os resultados de monotonicidade e antimonotonicidade vistos em  $\gamma_B(X)$  e  $\beta_B(X)$  também continuam valendo para as funções  $\gamma(\Sigma, C, \text{Sim}, I)$  e  $\beta(\Sigma, C, \text{Sim}, I)$ :

**Proposição 9** *Considerando os contextos  $\Sigma_1$  e  $\Sigma_2$  e  $\text{Sim} \in \{R, S, P\}$ , nós temos que*

$$\text{se } \Sigma_1 \subseteq \Sigma_2, \text{ então } \text{Sim}_{\Sigma_2} \subseteq \text{Sim}_{\Sigma_1}.$$

*Prova.* A afirmação de que se  $\Sigma_1 \subseteq \Sigma_2$  então  $R_{\Sigma_2} \subseteq R_{\Sigma_1}$  é similar à prova de que se  $B_1 \subseteq B_2$  então  $IND(B_2) \subseteq IND(B_1)$  (Proposição 8). Para os casos das relações de similaridade S e P, dois indivíduos são similares em relação a um contexto quando eles são similares em cada conceito pertencente ao contexto:

$$S_{\Sigma} = \{(a, b) \in \Delta^I \times \Delta^I \mid \text{para todo } C_i \text{ em que } i \in \{1, \dots, n\}, \pi_{C_i}^{\mathcal{K}}(a) = \pi_{C_i}^{\mathcal{K}}(b) \text{ ou} \\ \pi_{C_i}^{\mathcal{K}}(a) = \mathbf{u} \text{ ou } \pi_{C_i}^{\mathcal{K}}(b) = \mathbf{u}\}, \text{ e}$$

$$P_{\Sigma} = \{(a, b) \in \Delta^I \times \Delta^I \mid \text{para todo } C_i \text{ em que } i \in \{1, \dots, n\}, \pi_{C_i}^{\mathcal{K}}(a) = \pi_{C_i}^{\mathcal{K}}(b) \text{ ou} \\ \pi_{C_i}^{\mathcal{K}}(a) = \mathbf{u} \text{ ou } \pi_{C_i}^{\mathcal{K}}(b) = \mathbf{u} \\ \text{ou se } \pi_{C_i}^{\mathcal{K}}(a) = \mathbf{t} \text{ então } \pi_{C_i}^{\mathcal{K}}(b) = \mathbf{i} \\ \text{ou se } \pi_{C_i}^{\mathcal{K}}(a) = \mathbf{f} \text{ então } \pi_{C_i}^{\mathcal{K}}(b) = \mathbf{i}\}.$$

Considerado a relação S, por contradição, suponha que  $\Sigma_1 \subseteq \Sigma_2$  e  $S_{\Sigma_2} \not\subseteq S_{\Sigma_1}$ . Logo, existem  $x_1, x_2 \in \Delta^I$  tal que  $(x_1, x_2) \in S_{\Sigma_2}$  e  $(x_1, x_2) \notin S_{\Sigma_1}$ . Ou seja, para todo  $C_i \in \Sigma_2$ ,  $\pi_{C_i}^{\mathcal{K}}(x_1) = \pi_{C_i}^{\mathcal{K}}(x_2)$  ou  $\pi_{C_i}^{\mathcal{K}}(x_1) = \mathbf{u}$  ou  $\pi_{C_i}^{\mathcal{K}}(x_2) = \mathbf{u}$ . Como  $\Sigma_1 \subseteq \Sigma_2$ , temos que para todo  $C_i \in \Sigma_1$ ,  $\pi_{C_i}^{\mathcal{K}}(x_1) = \pi_{C_i}^{\mathcal{K}}(x_2)$  ou  $\pi_{C_i}^{\mathcal{K}}(x_1) = \mathbf{u}$  ou  $\pi_{C_i}^{\mathcal{K}}(x_2) = \mathbf{u}$ . Consequentemente,  $(x_1, x_2) \in S_{\Sigma_1}$ . Contradição com  $(x_1, x_2) \notin S_{\Sigma_1}$ .

Considerado a relação P, por contradição, suponha que  $\Sigma_1 \subseteq \Sigma_2$  e  $P_{\Sigma_2} \not\subseteq P_{\Sigma_1}$ . Logo, existem  $x_1, x_2 \in \Delta^I$  tal que  $(x_1, x_2) \in P_{\Sigma_2}$  e  $(x_1, x_2) \notin P_{\Sigma_1}$ . Ou seja, para todo  $C_i \in \Sigma_2$ ,  $\pi_{C_i}^{\mathcal{K}}(x_1) = \pi_{C_i}^{\mathcal{K}}(x_2)$  ou  $\pi_{C_i}^{\mathcal{K}}(x_1) = \mathbf{u}$  ou  $\pi_{C_i}^{\mathcal{K}}(x_2) = \mathbf{u}$ , ou se  $\pi_{C_i}^{\mathcal{K}}(x_1) = \mathbf{t}$  então  $\pi_{C_i}^{\mathcal{K}}(x_2) = \mathbf{i}$ , ou se  $\pi_{C_i}^{\mathcal{K}}(x_1) = \mathbf{f}$  então  $\pi_{C_i}^{\mathcal{K}}(x_2) = \mathbf{i}$ . Como  $\Sigma_1 \subseteq \Sigma_2$ , temos que para todo  $C_i \in \Sigma_1$ ,  $\pi_{C_i}^{\mathcal{K}}(x_1) = \pi_{C_i}^{\mathcal{K}}(x_2)$  ou  $\pi_{C_i}^{\mathcal{K}}(x_1) = \mathbf{u}$  ou  $\pi_{C_i}^{\mathcal{K}}(x_2) = \mathbf{u}$ , ou se  $\pi_{C_i}^{\mathcal{K}}(x_1) = \mathbf{t}$  então  $\pi_{C_i}^{\mathcal{K}}(x_2) = \mathbf{i}$ , ou se  $\pi_{C_i}^{\mathcal{K}}(x_1) = \mathbf{f}$  então  $\pi_{C_i}^{\mathcal{K}}(x_2) = \mathbf{i}$ . Consequentemente,  $(x_1, x_2) \in P_{\Sigma_1}$ . Contradição com  $(x_1, x_2) \notin P_{\Sigma_1}$ .  $\square$

**Proposição 10** Considerando os contextos  $\Sigma_1$  e  $\Sigma_2$ , temos que

- Se  $\Sigma_1 \subseteq \Sigma_2$ , então  $\gamma(\Sigma_1, C, \text{Sim}, I) \leq \gamma(\Sigma_2, C, \text{Sim}, I)$ ;
- Se  $\Sigma_1 \subseteq \Sigma_2$ , então  $\beta(\Sigma_1, C, \text{Sim}, I) \geq \beta(\Sigma_2, C, \text{Sim}, I)$ .

*Prova.* Suponha que  $\Sigma_1 \subseteq \Sigma_2$ . De acordo com a Proposição 9, se  $\Sigma_1 \subseteq \Sigma_2$ , então  $\text{Sim}_{\Sigma_2} \subseteq \text{Sim}_{\Sigma_1}$ . Consequentemente, pela semântica da aproximação inferior de um conceito, para qualquer conceito  $C$ ,  $(\underline{C}_{\text{Sim}_{\Sigma_1}})^I = \{x \in \Delta^I \mid \forall y((x, y) \in \text{Sim}_{\Sigma_1} \rightarrow y \in C^I)\} \subseteq \{x \in \Delta^I \mid \forall y((x, y) \in \text{Sim}_{\Sigma_2} \rightarrow y \in C^I)\} = (\underline{C}_{\text{Sim}_{\Sigma_2}})^I$ . Logo,  $|(\underline{C}_{\text{Sim}_{\Sigma_1}})^I| \leq |(\underline{C}_{\text{Sim}_{\Sigma_2}})^I|$ , para todo  $l \in \{t, f, i, u\}$ . Portanto, temos

$$\gamma(\Sigma_1, C, \text{Sim}, I) = \frac{|(\underline{C}_{\text{Sim}_{\Sigma_1}})^I| + |(\underline{C}_{f_{\text{Sim}_{\Sigma_1}}})^I| + |(\underline{C}_{i_{\text{Sim}_{\Sigma_1}}})^I| + |(\underline{C}_{u_{\text{Sim}_{\Sigma_1}}})^I|}{|\Delta^I|} \leq \\ \frac{|(\underline{C}_{\text{Sim}_{\Sigma_2}})^I| + |(\underline{C}_{f_{\text{Sim}_{\Sigma_2}}})^I| + |(\underline{C}_{i_{\text{Sim}_{\Sigma_2}}})^I| + |(\underline{C}_{u_{\text{Sim}_{\Sigma_2}}})^I|}{|\Delta^I|} = \gamma(\Sigma_2, C, \text{Sim}, I).$$

Consequentemente, Se  $\Sigma_1 \subseteq \Sigma_2$ , então  $1 - \gamma(\Sigma_1, C, \text{Sim}, I) \geq 1 - \gamma(\Sigma_2, C, \text{Sim}, I)$ , i.e.  $\beta(\Sigma_1, C, \text{Sim}, I) \geq \beta(\Sigma_2, C, \text{Sim}, I)$ .  $\square$

	$A_1$	$A_2$	$A_3$	$C$
$x_1$	<b>t</b>	<b>f</b>	<b>t</b>	<b>t</b>
$x_2$	<b>f</b>	<b>t</b>	<b>t</b>	<b>f</b>
$x_3$	<b>t</b>	<b>t</b>	<b>f</b>	<b>i</b>
$x_4$	<b>t</b>	<b>u</b>	<b>t</b>	<b>t</b>

Tabela 7.2: Interpretação dos conceitos  $A_1, A_2, A_3$  e  $C$ 

Com os seguintes resultados temos que ao diminuir o tamanho de um contexto, a precisão de sua aproximação tende a diminuir, enquanto que a imprecisão da aproximação tende a aumentar. Se estivermos interessados em buscar um contexto mais representativo para a aproximação inferior, um critério a ser utilizado é retornar o contexto que apresenta o maior grau de precisão. Logo, uma boa estratégia a ser utilizada é buscar contextos com maiores cardinalidades. Se o problema for encontrar um contexto para a aproximação superior, um critério a ser utilizado é retornar o contexto que apresenta o maior grau de imprecisão. Portanto, uma boa estratégia nesse caso é buscar os contextos de menores cardinalidades.

*Exemplo 28.* Considere a Tabela 7.2 com as informações sobre o universo  $\Delta^I = \{x_1, x_2, x_3, x_4\}$ . Ao procurarmos um contexto para a aproximação inferior de  $x_1$  com relação ao conceito  $C$  e a similaridade  $S$ , temos  $f(\{A_1, A_2, A_3\}, C, x_1, S) = A_1 \vee A_2$ . Nesse caso temos dois contextos minimais como resposta:  $\{A_1\}$  e  $\{A_2\}$ . Aplicando a função de precisão nos dois contextos, temos os seguintes resultados:

$$\gamma(\{A_1\}, C, S, I) = \frac{|(C_{tS_{\{A_1\}}})^I| + |(C_{fS_{\{A_1\}}})^I| + |(C_{iS_{\{A_1\}}})^I| + |(C_{uS_{\{A_1\}}})^I|}{|\Delta^I|} = \frac{0+1+0+0}{4} = \frac{1}{4}.$$

$$\gamma(\{A_2\}, C, S, I) = \frac{|(C_{tS_{\{A_2\}}})^I| + |(C_{fS_{\{A_2\}}})^I| + |(C_{iS_{\{A_2\}}})^I| + |(C_{uS_{\{A_2\}}})^I|}{|\Delta^I|} = \frac{1+0+0+0}{4} = \frac{1}{4}.$$

De acordo com o critério de precisão, o contexto  $\{A_1\}$  apresenta o mesmo valor de precisão do contexto  $\{A_2\}$ . Porém, temos que os contextos  $\{A_1, A_2\}$ ,  $\{A_1, A_3\}$ ,  $\{A_2, A_3\}$  e  $\{A_1, A_2, A_3\}$  são escolhas mais representativas que  $\{A_1\}$  e  $\{A_2\}$ , pois seus graus de precisão são maiores que  $\frac{1}{4}$ :

$$\begin{aligned} \gamma(\{A_1, A_2\}, C, S, I) &= \frac{2+1+1+0}{4} = 1, \\ \gamma(\{A_1, A_3\}, C, S, I) &= \frac{2+1+1+0}{4} = 1, \\ \gamma(\{A_2, A_3\}, C, S, I) &= \frac{2+1+1+0}{4} = 1, \\ \gamma(\{A_1, A_2, A_3\}, C, S, I) &= \frac{2+1+1+0}{4} = 1. \end{aligned}$$

## 7.2 Medidas de Granularidade

A medida de precisão de aproximações (ou completude do conhecimento) é baseada nas aproximações de *rough sets* como subconjuntos do universo. Contudo, a informação que diz

	$B_1$	$B_2$	$D$
$x_1$	<b>t</b>	<b>t</b>	<b>t</b>
$x_2$	<b>t</b>	<b>f</b>	<b>t</b>
$x_3$	<b>f</b>	<b>t</b>	<b>t</b>
$x_4$	<b>f</b>	<b>f</b>	<b>f</b>
$x_5$	<b>f</b>	<b>f</b>	<b>f</b>
$x_6$	<b>t</b>	<b>t</b>	<b>t</b>

Tabela 7.3: Tabela de decisão dos atributos  $B_1, B_2$  e  $D$ 

respeito à composição de partições de uma relação de indiscernibilidade (ou de similaridade) nas aproximações não é considerada. Em outras palavras, uma medida de precisão apenas reflete um aspecto da qualidade ou uma característica da aproximação. Isso tem levado a críticas e modificações nas medidas de precisão. Nesse sentido, em adição à medida de precisão, temos a medida de entropia.

A entropia é um conceito que foi derivado no século XIX durante o estudo de sistemas termodinâmicos. A entropia possui definições matemáticas e estatísticas precisas, mas pode ser aproximadamente definida como o grau de desordem ou incerteza em um sistema.

Em adição à entropia da termodinâmica, equações estatísticas para entropia foram derivadas por Boltzmann (SERWAY, 1999) na física e por Shannon (SHANNON, 1948) no campo da teoria da informação. Em relação aos *rough sets*, a medida de entropia é derivada da medida de Shannon.

É importante reconhecer entropia como um indicador do estado (ou nível) de um sistema. Ela pode ser vista como uma maneira importante de definir e medir sustentabilidade. Se o grau de desordem é muito grande (entropia é alta), então o sistema carece de sustentabilidade. Se a entropia é baixa, a sustentabilidade é mais fácil. Fazendo um paralelo com a teoria dos *rough sets*, sua entropia pode ser vista como o grau de desordem (granularidade) da definição dos indivíduos pertencentes a uma determinada classe.

*Exemplo 29.* De acordo com a Tabela 7.3, temos que  $D = \{x_1, x_2, x_3, x_6\}$  e  $\neg D = \{x_4, x_5\}$ . Se analisarmos esses conjunto por meios de classes de equivalências entre os indivíduos com relação aos atributos  $B_1$  e  $B_2$ , teremos que  $U/IND(D) = \{\{x_1, x_6\}, \{x_2\}, \{x_3\}\}$ , i.e.,  $x_1, x_2, x_3$  e  $x_6$  pertencem a  $D$ , mas apenas  $x_1$  e  $x_6$  são indiscerníveis entre si, e  $U/IND(\neg D) = \{\{x_4, x_5\}\}$ , i.e.,  $x_4$  e  $x_5$  são indiscerníveis entre si e pertencem ao complemento de  $D$ .

Comparando o nível de desordem (granularidade) entre os conjuntos  $D$  e  $\neg D$ , podemos ver que  $\neg D$  tem apenas uma partição de indivíduos ( $\{x_4, x_5\}$ ). Isso quer dizer que a definição de qualquer indivíduo  $x$  que está em  $\neg D$  pode ser feita pelas condições  $B_1(x) = \mathbf{f}$  e  $B_2(x) = \mathbf{f}$ . Em outras palavras, se  $B_1(x) = \mathbf{f}$  e  $B_2(x) = \mathbf{f}$  então  $D(x) = \mathbf{f}$ . O nível de desordem (granularidade) na descrição de  $\neg D$  é o menor possível, pois existe apenas uma maneira de descrever os indivíduos que não estão em  $D$ . Com o conjunto  $D$ , a situação é diferente, pois ele é composto por três partições:  $\{x_1, x_6\}, \{x_2\}, \{x_3\}$ . Isso implica que existem três maneiras distintas de definir um elemento que está em  $D$ . A primeira é se  $B_1(x) = \mathbf{t}$  e  $B_2(x) = \mathbf{t}$  então  $D(x) = \mathbf{t}$  (referentes a  $x_1$  e  $x_6$ ). A segunda é se  $B_1(x) = \mathbf{t}$  e  $B_2(x) = \mathbf{f}$  então  $D(x) = \mathbf{t}$  (referente a

$x_2$ ), e a terceira se  $B_1(x) = \mathbf{f}$  e  $B_2(x) = \mathbf{t}$  então  $D(x) = \mathbf{t}$  (referente a  $x_3$ ). Nesse caso, a entropia da descrição de  $D$  é alta, pois existem diferentes maneiras de caracterizar esse conjunto.

Um outro exemplo que aponta uma distinção entre precisão e granularidade das aproximações pode ser visto a seguir:

*Exemplo 30.* (BEAUBOUF; PETRY; ARORA, 1998) Seja  $S = (U, A, V_A)$  um sistema de informação, em que  $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9\}$  e  $B_1, B_2, B_3 \subseteq A$ . Suponha que temos as seguintes relações de indiscernibilidade  $IND(B_1)$ ,  $IND(B_2)$  e  $IND(B_3)$  com suas respectivas classes de equivalência:

$$\begin{aligned} U/IND(B_1) &= \{\{x_1, x_2, x_3, x_4\}, \{x_5, x_6, x_7\}, \{x_8, x_9\}\}, \\ U/IND(B_2) &= \{\{x_1, x_2\}, \{x_3, x_4\}, \{x_5, x_6, x_7\}, \{x_8, x_9\}\}, \\ U/IND(B_3) &= \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_5, x_6, x_7\}, \{x_8, x_9\}\}. \end{aligned}$$

Considere também as seguintes definições de aproximações:

- As noções de aproximações superior e inferior usuais como mostradas no Capítulo 2,  $\overline{X}^B = \{x \mid [x]_B \cap X \neq \emptyset\}$  e  $\underline{X}_B = \{x \mid [x]_B \subseteq X\}$ ;
- Uma noção alternativa das aproximações superior e inferior, que retornam classes de equivalências no lugar de indivíduos e são definidas como  $\overline{\overline{X}}^B = \{[x]_B \mid [x]_B \cap X \neq \emptyset\}$  e  $\underline{\underline{X}}_B = \{[x]_B \mid [x]_B \subseteq X\}$ , respectivamente.

Seja  $X = \{x_1, x_2, x_3, x_4, x_5, x_8\}$ . As aproximações superiores e inferiores (usuais) de  $X$  com os atributos  $B_1, B_2$  e  $B_3$  são

- $\underline{X}_{B_1} = \underline{X}_{B_2} = \underline{X}_{B_3} = \{x_1, x_2, x_3, x_4\}$ ;
- $\overline{X}_{B_1} = \overline{X}_{B_2} = \overline{X}_{B_3} = \emptyset$ ;
- $\overline{\overline{X}}^{B_1} = \overline{\overline{X}}^{B_2} = \overline{\overline{X}}^{B_3} = U$ .

A informação sobre as partições das classes de equivalências não está explicitamente mostrada. De acordo com a medida de precisão  $\gamma_B(X)$ , teríamos os mesmos valores para os três pares de aproximações, i.e.,

$$\gamma_{B_1}(X) = \gamma_{B_2}(X) = \gamma_{B_3}(X) = \frac{4+0}{9} = \frac{4}{9}.$$

Contudo, a granularidade (entropia ou co-entropia) das partições não é refletida. Se considerarmos a noção alternativa das operações de aproximações, teremos três diferentes pares de aproximações

- $\underline{\underline{X}}_{B_1} = \{\{x_1, x_2, x_3, x_4\}\}$ ;

- $\underline{X}_{B_2} = \{\{x_1, x_2\}, \{x_3, x_4\}\}$ ;
- $\underline{X}_{B_3} = \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}\}$ ;
- $\overline{\overline{X}}^{B_1} = \{\{x_1, x_2, x_3, x_4\}, \{x_5, x_6, x_7\}, \{x_8, x_9\}\}$ ;
- $\overline{\overline{X}}^{B_2} = \{\{x_1, x_2\}, \{x_3, x_4\}, \{x_5, x_6, x_7\}, \{x_8, x_9\}\}$ ;
- $\overline{\overline{X}}^{B_3} = \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_5, x_6, x_7\}, \{x_8, x_9\}\}$ .

A granularidade da informação das diferentes partições agora se torna explícita nas aproximações. O exemplo mostra que a segunda definição de aproximação permite-nos expressar mais um tipo de informação: a granularidade das aproximações.

Para medir o grau de granularidade das partições em *rough sets*, é utilizada a medida de entropia de Shannon (SHANNON, 1948). Um estudo bem detalhado sobre medidas de granularidades de partições pode ser encontrado em (BEAUBOUEF; PETRY; ARORA, 1998; ZIARKO, 2008; ZHU, 2011). Antes de mostrarmos a definição de entropia, precisamos definir algumas noções.

**Definição 57 (Partições de uma relação de Similaridade)** *Sejam  $S = (U, A, V_A)$  um sistema de informação,  $B \subseteq A$  e  $\text{Sim}$  uma relação de similaridade. O conjunto  $\pi = U/\text{Sim}(B) = \{X_1, \dots, X_m\}$  compreende o conjunto de todas as partições induzidas por  $\text{Sim}$ , i.e., cada  $X_i \in \pi$ ,  $1 \leq i \leq m$ , corresponde a  $\text{Sim}_B(x) = \{y \in U \mid (x, y) \in \text{Sim}(B)\}$ , para algum  $x \in U$ .*

Para dois conjuntos de partições  $\pi_1 = \{X_1, \dots, X_m\}$  e  $\pi_2 = \{Y_1, \dots, Y_n\}$  sob o mesmo domínio  $U$ , dizemos que  $\pi_1 \preceq \pi_2$  se e somente se para qualquer  $X_i \in \pi_1$ , existe  $Y_j \in \pi_2$  tal que  $X_i \subseteq Y_j$ , para  $1 \leq i \leq m, 1 \leq j \leq n$ . Dizemos que  $\pi_1$  é mais *fino* que  $\pi_2$  quando  $\pi_1 \preceq \pi_2$ ; caso contrário dizemos que  $\pi_1$  é mais *grosso* que  $\pi_2$ . A relação dos conjuntos de partições com os contextos em DL segue abaixo.

**Proposição 11** *Considerando os contextos  $\Sigma_1$  e  $\Sigma_2$  e os conjuntos de partições  $\pi_1 = \Delta^I/\text{Sim}_{\Sigma_1}$  e  $\pi_2 = \Delta^I/\text{Sim}_{\Sigma_2}$ , temos que*

$$\text{se } \Sigma_1 \subseteq \Sigma_2, \text{ então } \pi_1 \succeq \pi_2.$$

*Prova.* Suponha que  $\Sigma_1 \subseteq \Sigma_2$ . De acordo com a Proposição 9, se  $\Sigma_1 \subseteq \Sigma_2$ , então  $\text{Sim}_{\Sigma_2} \subseteq \text{Sim}_{\Sigma_1}$ . Seja  $\text{Sim}_{\Sigma}(x) = \{y \in \Delta^I \mid (x, y) \in \text{Sim}_{\Sigma}\}$  o conjunto de indivíduos similares a  $x$  com respeito a relação  $\text{Sim}$ . Como  $\text{Sim}_{\Sigma_2} \subseteq \text{Sim}_{\Sigma_1}$ , para qualquer  $x$ ,  $\text{Sim}_{\Sigma_2}(x) \subseteq \text{Sim}_{\Sigma_1}(x)$ . Sejam também os conjuntos de partições  $\pi_1 = \Delta^I/\text{Sim}_{\Sigma_1} = \{X_1, \dots, X_n\}$  e  $\pi_2 = \Delta^I/\text{Sim}_{\Sigma_2} = \{Y_1, \dots, Y_m\}$ , em que cada  $X_i \in \pi_1$  e  $Y_j \in \pi_2$ ,  $1 \leq i \leq n$  e  $1 \leq j \leq m$ , correspondem a  $\text{Sim}_{\Sigma_1}(x)$  e  $\text{Sim}_{\Sigma_2}(x)$  de algum  $x \in \Delta^I$ , respectivamente. Como  $\text{Sim}_{\Sigma_2}(x) \subseteq \text{Sim}_{\Sigma_1}(x)$  para qualquer  $x$ , temos para todo  $Y_j$ , existe um  $X_i$ , tal que  $Y_j \subseteq X_i$ . Portanto, por definição,  $\pi_2 \preceq \pi_1$ .  $\square$

**Definição 58 (Medida de Entropia)** (YAO, 2010) *Sejam  $S = (U, A, V_A)$  um sistema de informação,  $B \subseteq A$  e  $\text{Sim}$  uma relação de similaridade. A medida de entropia da partição  $\pi = U/\text{Sim}(B) = \{X_1, X_2, \dots, X_m\}$  é dada por*

$$H(\pi) = - \sum_{i=1}^m \frac{|X_i|}{|U|} \log \frac{|X_i|}{|U|}.$$

A entropia chega ao seu valor máximo  $\log|U|$  para o conjunto de partições mais fino, consistindo dos conjuntos unitários de  $U$  e chega ao seu valor mínimo 0 para o conjunto de partições mais grosso  $\{U\}$ . Complementarmente à medida de entropia, temos a medida de co-entropia.

**Definição 59 (Medida de Co-Entropia)** (YAO, 2010) *Sejam  $S = (U, A, V_A)$  um sistema de informação,  $B \subseteq A$  e  $\text{Sim}$  uma relação de similaridade. A medida de co-entropia da partição  $\pi = U/\text{Sim}(B) = \{X_1, X_2, \dots, X_m\}$  é dada por*

$$G(\pi) = \frac{1}{|U|} \sum_{i=1}^m |X_i| \log |X_i|.$$

O conjunto de partições mais grosso  $\{U\}$  tem o valor máximo de co-entropia  $\log|U|$ , e o conjunto de partições mais fino  $\{\{x\} \mid x \in U\}$  tem o valor mínimo de co-entropia que é 0. Intuitivamente a medida de entropia (co-entropia) mede o grau de desordem (respectivamente ordem) entre as partições geradas por uma relação. Quanto mais próximo esse valor é de  $\log|U|$ , maior é o grau de desordem (respectivamente ordem) entre as partições. Com relação às medidas de granularidade, temos os seguintes resultados de monotonicidade e antimonotonicidade:

**Proposição 12** (YAO, 2010) *Dadas duas partições  $\pi_1$  e  $\pi_2$ :*

- Se  $\pi_1 \preceq \pi_2$ , então  $H(\pi_1) \geq H(\pi_2)$ ;
- Se  $\pi_1 \preceq \pi_2$ , então  $G(\pi_1) \leq G(\pi_2)$ .

Para adaptar as medidas de entropia e co-entropia em *Rough  $\mathcal{ALC}$*  e *PR $_{\mathcal{ALC}}$* , precisaremos primeiramente definir uma noção de aproximação relacionada às partições de uma relação de similaridade:

**Definição 60 (Partições de Aproximação Inferior/Superior)** *Os conjuntos das partições das aproximações inferior e superior de um conceito  $C$ , com relação a um contexto  $\Sigma$  e uma relação de similaridade  $\text{Sim}$  são dados por:*

$$\begin{aligned} (\underline{C}_{\text{Sim}_\Sigma})^I &= \{\text{Sim}_\Sigma(x) \in \Delta^I/\text{Sim}_\Sigma \mid \text{Sim}_\Sigma(x) \subseteq ((C_t \sqcup C_i)_{\text{Sim}_\Sigma})^I\} \text{ e} \\ (\overline{C}^{\text{Sim}_\Sigma})^I &= \{\text{Sim}_\Sigma(x) \in \Delta^I/\text{Sim}_\Sigma \mid \text{Sim}_\Sigma(x) \subseteq ((\overline{C_t \sqcup C_i})^{\text{Sim}_\Sigma})^I\}, \end{aligned}$$

em que  $\text{Sim}_\Sigma(x) = \{y \in \Delta^I \mid (x, y) \in \text{Sim}_\Sigma\}$  e  $C_t$  e  $C_i$  são os conjuntos transformados de  $C$ .

Essas partições de aproximações inferiores/superiores retornam as partições dos indivíduos que satisfazem as aproximações inferiores/superiores em vez de retornar apenas os indivíduos. Sendo assim, as medidas de granularidade em *Rough  $\mathcal{ALC}$*  e *PR $\mathcal{ALC}$*  ficarão responsáveis por medir a ordem ou a desordem nas partições obtidas pelas aproximações inferior e superior.

**Definição 61 (Medidas de Entropia/Co-Entropia - DL)** *Sejam  $\Sigma$  um contexto,  $C$  um conceito,  $\text{Sim}$  uma relação de similaridade e  $I$  uma interpretação. As medidas de entropia e co-entropia em  $\mathcal{ALC}_4$  e  $PR_{\mathcal{ALC}}$  são definidas como*

- $H_1(\Sigma, C, \text{Sim}, I) = H((\underline{C}_{\text{Sim}_\Sigma})^I)$  e  $H_2(\Sigma, C, \text{Sim}, I) = H((\overline{C}^{\text{Sim}_\Sigma})^I)$ ;
- $G_1(\Sigma, C, \text{Sim}, I) = G((\underline{C}_{\text{Sim}_\Sigma})^I)$  e  $G_2(\Sigma, C, \text{Sim}, I) = G((\overline{C}^{\text{Sim}_\Sigma})^I)$ .

Da mesma forma que as funções  $H(\pi)$  e  $G(\pi)$ ,  $H_i(\Sigma, C, \text{Sim}, I)$  e  $G_i(\Sigma, C, \text{Sim}, I)$  têm valor mínimo de 0 e valor máximo de  $\log|\Delta^I|$ . A partir das Proposições 6 e 12 chegamos ao seguinte resultado:

**Corolário 1** *Considerando os contextos  $\Sigma_1$  e  $\Sigma_2$ , temos que:*

- Se  $\Sigma_1 \subseteq \Sigma_2$  então  $H_1(\Sigma_1, C, \text{Sim}, I) \geq H_1(\Sigma_2, C, \text{Sim}, I)$  e  $H_2(\Sigma_2, C, \text{Sim}, I) \geq H_2(\Sigma_1, C, \text{Sim}, I)$ ;
- Se  $\Sigma_1 \subseteq \Sigma_2$  então  $G_1(\Sigma_1, C, \text{Sim}, I) \leq G_1(\Sigma_2, C, \text{Sim}, I)$  e  $G_2(\Sigma_2, C, \text{Sim}, I) \leq G_2(\Sigma_1, C, \text{Sim}, I)$ .

*Exemplo 31.* Considere novamente a Tabela 7.2. A medida de co-entropia para os conjuntos minimais  $\{A_1\}$  e  $\{A_2\}$  é

- $G_1(\{A_1\}, C, \text{Sim}, I) = G((\underline{C}_{\text{Sim}_{\{A_1\}}})^I) = G(\{\{x_1, x_3, x_4\}\}) = \frac{3 \log 3}{4} = 0.357$ ;
- $G_1(\{A_2\}, C, \text{Sim}, I) = G((\underline{C}_{\text{Sim}_{\{A_2\}}})^I) = G(\{\{x_1\}\}) = \frac{1 \log 1}{4} = 0$ .

Nesse caso, o contexto  $\{A_1\}$  é o que apresenta o maior grau de co-entropia. Isso significa que  $\{A_1\}$  apresenta uma descrição mais uniforme que  $\{A_2\}$  com relação às informações presentes na Tabela 7.2.

### 7.3 Medidas de Qualidade Compostas

Cada medida apresentada anteriormente captura um diferente aspecto de qualidade das aproximações. Neste trabalho, utilizaremos uma medida combinada como uma medida de adequação de uma aproximação que pretende refletir o quão boa uma aproximação é de acordo com os critérios de precisão e granularidade.

**Definição 62 (Medidas de Qualidade - DL)** *Sejam  $\Sigma$  um contexto,  $C$  um conceito,  $\text{Sim}$  uma relação de similaridade,  $I$  uma interpretação e  $x$  um indivíduo. As medidas de qualidade compostas em  $\mathcal{ALC}_4$  e  $PR_{\mathcal{ALC}}$  são definidas como*

$$L(\Sigma, C, \text{Sim}, I) = \gamma(\Sigma, C, \text{Sim}, I) * G_1(\Sigma, C, \text{Sim}, I) \text{ e}$$

$$U(\Sigma, C, \text{Sim}, I) = \beta(\Sigma, C, \text{Sim}, I) * G_2(\Sigma, C, \text{Sim}, I).$$

As medidas compostas  $L(\Sigma, C, \text{Sim}, I)$  e  $U(\Sigma, C, \text{Sim}, I)$  são aplicadas para contextos de aproximações inferiores e superiores, respectivamente. A função  $L(\Sigma, C, \text{Sim}, I)$  reflete um balanço entre precisão e co-entropia. Ela tem seu valor mínimo de 0 e máximo de  $\log|\Delta^I|$ . Quanto mais próximo esse valor é de  $\log|\Delta^I|$ , melhor será a qualidade da aproximação inferior em termos gerais. Já a função  $U(\Sigma, C, \text{Sim}, I)$  captura um balanço entre imprecisão e co-entropia. Ela também tem valor mínimo de 0 e valor máximo de  $\log|\Delta^I|$  e quanto mais próximo de  $\log|\Delta^I|$ , melhor será a qualidade geral da aproximação superior. Deve ser percebido que uma medida composta não é sempre significativa, embora ela tenha a vantagem de ser um único número. Na prática, é melhor trabalhar com apenas um número do que com uma tupla de números fornecidos pelas medidas de qualidade.

**Proposição 13** *Considerando os contextos  $\Sigma_1$  e  $\Sigma_2$ , um conceito  $C$ , uma relação de similaridade  $\text{Sim}$  e um indivíduo  $x$ , temos que*

- Se  $\Sigma_1 \subseteq \Sigma_2$  então  $L(\Sigma_1, C, \text{Sim}, I) \leq L(\Sigma_2, C, \text{Sim}, I)$ ;
- Se  $\Sigma_1 \subseteq \Sigma_2$  então  $U(\Sigma_1, C, \text{Sim}, I) \geq U(\Sigma_2, C, \text{Sim}, I)$ .

*Prova.* Suponha que  $\Sigma_1 \subseteq \Sigma_2$ . Pela Proposição 10, temos que  $\gamma(\Sigma_1, C, \text{Sim}, I) \leq \gamma(\Sigma_2, C, \text{Sim}, I)$  e  $\beta(\Sigma_1, C, \text{Sim}, I) \geq \beta(\Sigma_2, C, \text{Sim}, I)$ . Pelo Corolário 1, nós temos que  $G_1(\Sigma_1, C, \text{Sim}, I) \leq G_1(\Sigma_2, C, \text{Sim}, I)$  e  $G_2(\Sigma_1, C, \text{Sim}, I) \geq G_2(\Sigma_2, C, \text{Sim}, I)$ . Logo,

$$\begin{aligned} L(\Sigma_1, C, \text{Sim}, I) &= \gamma(\Sigma_1, C, \text{Sim}, I) * G_1(\Sigma_1, C, \text{Sim}, I) \leq \gamma(\Sigma_2, C, \text{Sim}, I) * G_1(\Sigma_2, C, \text{Sim}, I) = \\ &L(\Sigma_2, C, \text{Sim}, I) \text{ e} \\ U(\Sigma_1, C, \text{Sim}, I) &= \beta(\Sigma_1, C, \text{Sim}, I) * G_2(\Sigma_1, C, \text{Sim}, I) \geq \beta(\Sigma_2, C, \text{Sim}, I) * G_2(\Sigma_2, C, \text{Sim}, I) = \\ &U(\Sigma_2, C, \text{Sim}, I). \end{aligned}$$

□

Com relação à medida composta  $L(\Sigma, C, \text{Sim}, I)$ , esse resultado mostra que ao aumentarmos um contexto, as medidas de precisão e co-entropia de sua aproximação inferior tendem a aumentar. Se trabalharmos com contextos para restrições de consulta, uma alternativa é buscar contextos com maiores cardinalidades. Com relação à medida  $U(\Sigma, C, \text{Sim}, I)$ , esse resultado mostra que ao diminuirmos um contexto, as medidas de imprecisão e co-entropia de sua aproximação superior tendem a aumentar. Se trabalharmos com contextos para relaxamentos de consulta, uma alternativa é buscar contextos com menores cardinalidades.

*Exemplo 34.* A medida de qualidade composta dos contextos minimais  $\{A_1\}$  e  $\{A_2\}$  da Tabela 7.2, com relação ao conceito  $D$ , a relação de similaridade  $S$  e a interpretação  $I$  é

$$L(\{A_1\}, D, S, I) = 0.25 * 0.357 = 0.08925,$$

$$L(\{A_2\}, D, S, I) = 0.5 * 0 = 0.$$

Nesse caso, o contexto  $\{A_1\}$  apresenta uma medida de qualidade maior que  $\{A_2\}$  e portanto, se trata de uma escolha geral melhor que  $\{A_2\}$ .

## 8 CONCLUSÕES E TRABALHOS FUTUROS

### 8.1 Conclusões

O objetivo principal desta dissertação foi apresentar a teoria dos *rough sets* como uma ferramenta a ser utilizada no processo de Refinamento de Consultas. Uma consulta pode ser refinada por diversos motivos, mas uma característica que está sempre presente nos refinamentos é ausência de certeza. Escolhemos então tratar de dois tipos de incerteza: a que é resultante de contradição e a que resulta de desconhecimento. Para isso, utilizamos como fundamentação teórica duas abordagens baseadas em lógicas paraconsistentes tetravaloradas (BELNAP, 1977; VITÓRIA; SZALAS; MAŁUSZYŃSKI, 2008). Além do mais, consideramos adequar essas definições na linguagem das Lógicas de Descrição, que são bastante utilizadas em aplicações que envolvem base de dados e ontologias.

Apresentamos duas Lógicas de Descrição paraconsistentes e *rough* capazes de representação e aproximação de informação contraditória e desconhecida. A aproximação dos axiomas dessas lógicas caracterizam os Refinamentos de Consultas. Lidamos neste trabalho com dois tipos de refinamentos: as restrições e os relaxamentos. Esses refinamentos são gerados através de um método que consiste em buscar contextos para as operações de aproximações. Por fim, propomos também métricas que nos ajudam a medir a qualidade dessas aproximações a fim de buscar o melhor contexto para um refinamento.

Um ponto chave desta dissertação é a representação de incertezas utilizando *rough sets*. Para serem realizadas as aproximações, é necessário analisarmos o grau de proximidade dos objetos de um domínio através dos valores de suas propriedades. Porém, como trabalhamos com informação incerta e contraditória, é possível descrever diversas formas de proximidade entre objetos. Mostramos durante o trabalho três modos distintos de relações de proximidade considerando informação completa, contraditória e desconhecida. Consequentemente, o desenvolvimento do método de busca de refinamentos e das medidas de qualidade levam em consideração essas relações de proximidade entre objetos.

Alguns problemas foram encontrados durante o desenvolvimento deste trabalho. O primeiro problema trata da relação de expressividade das *Rough DLs* com a teoria dos *rough sets*. As *Rough DLs* são, de fato, mais expressivas que *rough sets* clássicos, pois permitem a aproximação de conceitos complexos. Fazendo um paralelo da teoria dos *rough sets* com DLs, podemos dizer que ela permite apenas aproximações de conceitos atômicos. Além disso, as *Rough DLs* permitem axiomas de inclusão de conceitos que conseguem definir relação de dependência entre os conceitos. Na teoria dos *rough sets*, entretanto, as informações dos atributos de uma tabela de decisão são independentemente dispostas.

Na prática, podem surgir alguns problemas durante o processo de refinamento de axiomas em DL. Ao tentarmos refinar um conceito, buscamos por um contexto formado por outros conceitos que não estão diretamente ligados a ele. Em *Rough DL*, ao aproximarmos um conceito  $C$  com um contexto  $\Sigma$  e  $C \in \Sigma$ , temos que  $C \equiv \overline{C}^{\Sigma} \equiv \underline{C}_{\Sigma}$ , o que torna trivial e irrelevante a aproximação. Esse caso não se limita apenas quando o próprio conceito está contido no

contexto. Dependendo do conceito utilizado, se algum de seus sub-conceitos estiver presente no contexto, o resultado da aproximação também se torna trivial. Por esse motivo é que no Capítulo 6, optamos por retirar todos os conceitos relacionados com o conceito a ser aproximado durante o processo de busca de contextos. Por esse mesmo motivo é que optamos também por não utilizarmos TBox no processo de refinamento, pois os axiomas de inclusão de conceitos apresentam dependências entre os conceitos. Além de excluirmos todos os sub-conceitos de conceito durante o processo, deveríamos analisar se através dos axiomas de inclusões, outros conceitos não desempenham papel similar aos sub-conceitos retirados.

Outro problema apresentado foi com relação a aproximação de axiomas terminológicos. Decidimos não utilizá-los nesse trabalho por dois motivos: aparentemente são mais complicados de tratar que os axiomas de asserção e podem ser aplicados de várias maneiras. Enquanto que os axiomas de asserção tratam da aproximação de apenas um indivíduo em uma base de conhecimento, os axiomas terminológicos necessitam uma análise de todos os indivíduos da base de conhecimento.

Por fim, um refinamento em um axioma terminológico pode ser pensado de várias maneiras. Como estamos trabalhando com axiomas do tipo  $C \equiv D$  ou  $C \sqsubseteq D$ , podemos pensar, por exemplo, em um relaxamento desses axiomas como a restrição do conceito  $C$  e o relaxamento do conceito  $D$  ou nessas operações realizadas separadamente. Podemos, inclusive, desenvolver restrições e relaxamentos dos conceitos contidos em  $C$  e  $D$ . Essas e outras questões precisam ser analisadas mais detalhadamente.

## 8.2 Trabalhos Futuros

Algumas questões ficaram abertas durante a dissertação e pretendemos investigá-las em trabalhos futuros. São elas

- Mostrar um limite superior menor de complexidade para  $PR_{\mathcal{ALC}}$ .

Como mostrado no Capítulo 5, conseguimos um limite superior de complexidade de 2-EXPTIME para  $PR_{\mathcal{ALC}}$  através de uma tradução para  $\mathcal{ALC}$ . Supomos que seja possível conseguir um limite menor equivalente à complexidade de  $\mathcal{ALC}$ : EXPTIME. Como trabalho futuro, pretendemos desenvolver um método de prova como os tableaux de DLs para se ter uma ideia mais clara de sua complexidade.

- Aplicar Refinamento de Consultas em Axiomas de Asserção considerando TBoxes não-vazios.

Como dito na seção anterior, realizar uma busca de contextos para um conceito considerando ABox e TBox é aparentemente mais complexo que considerando apenas o ABox. Pretendemos descobrir exatamente qual a função dos axiomas terminológicos no processo de refinamento, desenvolver um método de busca de contextos para refinamento considerando TBoxes e mostrar sua complexidade.

- Aplicar Refinamento de Consultas em Axiomas Terminológicos.

Consequentemente, a partir dos resultados obtidos do item anterior, poderemos entender o processo de refinamento também para os axiomas terminológicos.

- Extensões Multi-valoradas.

Nos nossos estudos iniciais, nós focamos em extensões de *rough sets* com lógicas multi-valoradas como as lógicas paraconsistentes (MAŁUSZYŃSKI; SZĄŁAS; VITÓRIA, 2007, 2008; VITÓRIA, 2005; VITÓRIA; MAŁUSZYŃSKI, 2002) e lógica *fuzzy* (BOBILLO; STRACCIA, 2009; THIELE, 1998; YAO, 2004). Um artigo foi escrito (**A Seven-Valued Extension of Rough Sets** (VIANA; ALCÂNTARA; MARTINS, 2011)) estendendo a ideia de *rough sets* para uma lógica heptavalorada, com intuito de representar e aproximar informação incompleta levando em conta valores perdidos (GRZYMALA-BUSSE, 1997).

O objetivo desse artigo foi aplicar as ideias dos *rough sets* para diversos graus de informações incompletas e contraditórias. Uma das possíveis linhas de estudo segue em adaptar raciocínio com essa abordagem multi-valorada para Lógicas de Descrição.

- Extensões Multi-valoradas de Relações de Similaridade.

No Capítulo 5, trabalhamos com três tipos de relações de similaridade. Porém, todas as relações são bivaloradas a despeito da linguagem permitir relações tetravaloradas como apresentado em (VITÓRIA; SZĄŁAS; MAŁUSZYŃSKI, 2008). Pretendemos investigar como introduzir e aplicar novos tipos de relações de similaridade.

## BIBLIOGRAFIA

- AKAHANI, J.; HIRAMATSU, K.; SATOH, T. Approximate Query Reformulation Based on Hierarchical Ontology Mapping. *In Proc. of International Workshop on Semantic Web Foundations and Application Technologies (SWFAT)*, p. 43–46, 2003.
- BAADER, F. *The Description Logic Handbook: Theory, Implementation, and Applications*. New York, NY, USA: Cambridge University Press, 2003.
- BAADER, F.; HORROCKS, I.; SATTLER, U. Description Logics. In: STAAB, S.; STUDER, R. (Ed.). *Handbook on Ontologies*. Berlin, Heidelberg, New York: Springer, 2004, (International Handbooks on Information Systems). p. 3–28. ISBN 3-540-40834-7.
- BAADER, F. et al. *Integrating Description Logics and Action Formalisms for Reasoning about Web Services*. Germany, 2005. See <http://lat.inf.tu-dresden.de/research/reports.html>.
- BEAUBOUEF, T.; PETRY, F. E.; ARORA, G. Information-Theoretic Measures of Uncertainty for Rough Sets and Rough Relational Databases. *Inf. Sci.*, Elsevier Science Inc., New York, NY, USA, v. 109, n. 1-4, p. 185–195, 1998. ISSN 0020-0255.
- BELNAP, N. D. A Useful Four-Valued Logic. *Modern Uses of Multiple-Valued Logic*. D. Reidel. Ed: J. Michael Dunn and G. Epstein, p. 8–37, 1977.
- BERARDI, D.; CALVANESE, D.; GIACOMO, G. D. Reasoning on uml class diagrams. *Artif. Intell.*, v. 168, n. 1-2, p. 70–118, 2005.
- BOBILLO, F.; STRACCIA, U. Supporting Fuzzy Rough Sets in Fuzzy Description Logics. *In European Conferences on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU)*, p. 676–687, 2009.
- BORGIDA, A.; LENZERINI, M.; ROSATI, R. Description Logics for Databases. In: BAADER, F. et al. (Ed.). *Description Logic Handbook*. Cambridge: Cambridge University Press, 2003. p. 462–484. ISBN 0-521-78176-0.
- BOROS, E.; ELBASSIONI, K.; MAKINO, K. On Berge Multiplication for Monotone Boolean Dualization. In: *Proceedings of the 35th international colloquium on Automata, Languages and Programming, Part I*. Berlin, Heidelberg: Springer-Verlag, 2008. (ICALP'08), p. 48–59. ISBN 978-3-540-70574-1.
- CALVANESE, D. et al. View-based Query Answering in Description logics: Semantics and Complexity. *J. of Computer and System Sciences*, 2011.
- COCK, M. D.; CORNELIS, C. Fuzzy Rough Set Based Web Query Expansion. *Proceedings of Rough Sets and Soft Computing in Intelligent Agent and Web Technology, International Workshop at WIAT2005 (2005 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology)*, p. 9–16, 2005.
- COCK, M. D.; CORNELIS, C.; KERRE, E. E. Fuzzy Rough Sets: Beyond the Obvious. *Proceedings of FUZZ-IEEE2004. Vol. 1*, p. 103–108, 2004.

- COSTA, N. C. A. *Sistemas formais inconsistentes (PhD Thesis 1963)*. BR: Edidora UFPR, 1993.
- FANIZZI, N.; D'AMATO, C.; ESPOSITO, F. Randomized Metric Induction and Evolutionary Conceptual Clustering for Semantic Knowledge Bases. In: SILVA, M. J. et al. (Ed.). *CIKM*. ACM Press: ACM, 2007. p. 51–60. ISBN 978-1-59593-803-9.
- FANIZZI, N. et al. Representing Uncertain Concepts in Rough Description Logics via Contextual Indiscernibility Relations. *URSW*, 2008.
- FAYYAD, U. M. et al. (Ed.). *Advances in Knowledge Discovery and Data Mining*. Menlo Park, CA, USA: American Association for Artificial Intelligence, 1996. ISBN 0-262-56097-6.
- FRANCO, W. et al. A Description Logic for InferenceNet.Br. In: CASELI, H. de M. et al. (Ed.). *PROPOR*. Published on: Springer, 2012. (Lecture Notes in Computer Science, v. 7243), p. 298–305. ISBN 978-3-642-28884-5.
- GINSBERG, M. Multivalued Logics: A Uniform Approach to Reasoning in Artificial Intelligence. *Computational Intelligence*, v. 4, p. 265–316, 1988.
- GOLDSTONE, R. L.; MEDIN, D. L.; HALBERSTADT, J. Similarity in Context. *Memory and Cognition*, p. 237–255, 1997.
- GRZYMALA-BUSSE, J. Modified Algorithms LEM1 and LEM2 for Rule Induction from Data with Missing Attribute Value. *Proc. of the Fifth International Workshop on Rough Sets and Sof Computing (RSSC'97) at the Third Joint Conference on Information Sciences (JCSI'97), Research Triangle Park, NC, March 2-5*, p. 69–72, 1997.
- GRZYMALA-BUSSE, J. Rough Set Strategies to Data with Missing Attribute Values. In: *Foundations and Novel Approaches in Data Mining*. Berlin, Heidelberg, New York: Springer, 2006. p. 197–212.
- HORROCKS, I. DAML+OIL: A Description Logic for the Semantic Web. *IEEE Data Eng. Bull.*, v. 25, n. 1, p. 4–9, 2002.
- JAŚKOWSKI, S. A propositional Calculus for Inconsistent Deductive Systems. (*reprinted in: Studia Logica*, 24 1969, pp 143-157 and in: *Logic and Logical Philosophy* 7, 1999 pp. 35-56), 1948.
- JIANG, Y. et al. Reasoning with Rough Description Logics: An Approximate Concepts Approach. *Inf. Sci.*, v. 179, n. 5, p. 600–612, 2009.
- KEET, C. M. On the Feasibility of Description Logic Knowledge Bases with Rough Concepts and Vague Instances. *Description Logics (DL2010)*, 2010.
- KLIR, G. J. *Uncertainty and Information: Foundations of Generalized Information Theory*. USA: Wiley-IEEE Press; 1 edition, 2005.
- KOMOROWSKI, J. et al. *Rough Sets: A tutorial*. 1998.
- LIAU, C.-J.; YAO, Y. Y. Information Retrieval by Possibilistic Reasoning. In: MAYR, H. C. et al. (Ed.). *Database and Expert Systems Applications, 12th International Conference, DEXA 2001 Munich, Germany, September 3-5, 2001, Proceedings*. Berlin, Heidelberg, New York: Springer, 2001. (Lecture Notes in Computer Science, v. 2113), p. 52–61. ISBN 3-540-42527-6.

- MA, Y.; HITZLER, P.; LIN, Z. *Paraconsistent Reasoning with OWL — Algorithms and the ParOWL Reasoner*. Germany, 2006.
- MAŁUSZYŃSKI, J.; SZAŁAS, A.; VITÓRIA, A. A Four-valued Logic for Rough Set-like Approximate Reasoning. *T. Rough Sets*, v. 6, p. 176–190, 2007.
- MAŁUSZYŃSKI, J.; SZAŁAS, A.; VITÓRIA, A. Paraconsistent Logic Programs with Four-Valued Rough Sets. p. 41–51, 2008.
- MILTERSEN, P.; RADHAKRISHNAN, J.; WEGENER, I. On Converting CNF to DNF. *Theor. Comput. Sci.*, Elsevier Science Publishers Ltd., Essex, UK, v. 347, n. 1-2, p. 325–335, 2005. ISSN 0304-3975.
- MINSKY, M. A Framework for Representing Knowledge. In: HAUGELAND, J. (Ed.). *Mind Design: Philosophy, Psychology, Artificial Intelligence*. Cambridge, MA: MIT Press, 1981. p. 95–128.
- PAWLAK, Z. Rough Sets. *International Journal of Information and Computer Science*, v. 11, p. 341–356, 1982.
- PAWLAK, Z. *Rough Sets. Theoretical Aspects of Reasoning About Data*. Dordrecht: Kluwer Academic Publishers, 1991.
- PAWLAK, Z.; MAREK, W. Rough Sets and Information Systems. *Fundamenta Informaticae*, v. 17, p. 105–115, 1984.
- RUSSELL, S. J.; NORVIG, P. *Artificial Intelligence - A Modern Approach (3. internat. ed.)*. UK: Pearson Education, 2010. I-XVIII, 1-1132 p. ISBN 978-0-13-207148-2.
- SCHAERF, M.; CADOLI, M. Tractable Reasoning via Approximation. *Artificial Intelligence*, v. 74, p. 249–310, 1995.
- SCHILD, K. A correspondence theory for terminological logics: preliminary report. In: *Proceedings of the 12th international joint conference on Artificial intelligence - Volume 1*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1991. (IJCAI'91), p. 466–471. ISBN 1-55860-160-0. Disponível em: <<http://dl.acm.org/citation.cfm?id=1631171.1631241>>.
- SCHLOBACH, S. et al. *Anytime Classification by Ontology Approximation*. 2007.
- SCHLOBACH, S.; KLEIN, M. Description Logics with Approximate Definitions: Precise Modeling of Vague Concepts. *Proceedings of the 20th International Joint Conference on Artificial Intelligence, IJCAI 07*, 2007.
- SCHMIDT-SCHAUSS, M.; SMOLKA, G. Attributive Concept Descriptions with Complements. *Artif. Intell.*, v. 48, n. 1, p. 1–26, 1991.
- SERWAY, R. A. *Physics for Scientists and Engineers*. Pacific Grove: Brooks/Cole Publishing Company, 1999. Hardcover. ISBN 0030317169.
- SHANNON, C. E. A mathematical theory of communication. *Bell system technical journal*, v. 27, 1948.

- SKOWRON, A.; RAUSZER, C. The Discernibility Matrices and Functions in Information Systems. In: *Slowiński, R.(Ed.), Handbook of Applications and Advances of the Rough Sets Theory*. Kluwer, Dordrecht, 1992.
- SOWA, J. F. Semantic Networks. In: SHAPIRO, S. C. (Ed.). *Encyclopedia of Artificial Intelligence*. 2. ed. Wiley, 1992. Disponível em: <<http://www.jfsowa.com/pubs/semnet.htm>>.
- STEPANIUK, J. Approximation Spaces, Reducts and Representatives. (eds.) L. Polkowski, A. Skowron, *Rough Sets in Knowledge Discovery, Part I and II*, Physica-Verlag, 1998.
- STUCKENSCHMIDT, H. *Approximate Subsumption in  $\mathcal{ALCQ}$* . 2007.
- STUCKENSCHMIDT, H. Partial Matchmaking using Approximate Subsumption. In *AAAI'07*, p. 1459–1464, 2007.
- THIELE, H. Fuzzy Rough Sets versus Rough Fuzzy Sets - An Interpretation and a Comparative Study Using Concepts of Modal Logic. *Technical Report, ISSN 1433-3325, University of Dortmund*, 1998.
- TOBIES, S. Complexity Results and Practical Algorithms for Logics in Knowledge Representation. *CoRR*, cs.LO/0106031, 2001.
- VIANA, H.; ALCÂNTARA, J.; MARTINS, A. T. Paraconsistent Rough Description Logic. *Proceedings of the 24th International Workshop on Description Logics (DL 2011)*, p. 587–597, 2011.
- VIANA, H.; ALCÂNTARA, J.; MARTINS, A. T. A Seven-Valued Extension of Rough Sets. *Submetido ao LSFA 2011*, 2011.
- VIANA, H.; ALCÂNTARA, J.; MARTINS, A. T. Searching Contexts in a Paraconsistent Rough Description Logic. *A submeter*, 2012.
- VIANA, H. et al. A Tractable Paraconsistent Fuzzy Description Logic. In: BOBILLO, F. et al. (Ed.). *URSW*. Published on: CEUR-WS.org, 2010. (CEUR Workshop Proceedings, v. 654), p. 101–104.
- VITÓRIA, A. A Framework for Reasoning with Rough Sets. *T. Rough Sets*, p. 178–276, 2005.
- VITÓRIA, A.; MAŁUSZYŃSKI, J. A Logic Programming Framework for Rough Sets. In: ALPIGINI, J. et al. (Ed.). *Proc. of the 3rd International Conference on Rough Sets and Current Trends in Computing, RSCTC'02*. Berlin, Heidelberg, New York: Springer-Verlag, 2002. (LNCS/LNAI, 2475), p. 205–212.
- VITÓRIA, A.; SZAŁAS, A.; MAŁUSZYŃSKI, J. Four-Valued Extension of Rough Sets. *Proceedings of Third International Conference on Rough Sets and Knowledge Technology, Chengdu, China (17th–19th May 2008)*. (LNCS, v. 5009), p. 106–114, 2008.
- WEI-HUA, X.; XIAO-YAN, Z.; WEN-XIU, Z. Knowledge Granulation, Knowledge Entropy and Knowledge Uncertainty Measure in Ordered Information Systems. *Appl. Soft Comput.*, Elsevier Science Publishers B. V., Amsterdam, The Netherlands, The Netherlands, v. 9, n. 4, p. 1244–1251, 2009. ISSN 1568-4946.

WU, W.; ZHANG, W. Neighborhood Operator Systems and Approximations. *Inf. Sci.*, v. 144, n. 1-4, p. 201–217, 2002.

YAO, Y. Semantics of Fuzzy Sets in Rough Set Theory. *T. Rough Sets*, p. 297–318, 2004.

YAO, Y. Notes on Rough Set Approximations and Associated Measures. *Proceedings of Second International Workshop on Rough Sets Theory (RST'10)*, Zhoushan, Zhejiang, China, October 19-21, Journal of Zhejiang Ocean University (Natural Science), Vol. 29, No. 5, p. 399–410, 2010.

YAO, Y.; ZHAO, Y. Discernibility Matrix Simplification for Constructing Attribute Reducts. *Inf. Sci.*, v. 179, n. 7, p. 867–882, 2009.

ZADEH, L. A. Fuzzy Sets. *Information Control*, v. 8, p. 338–353, 1965.

ZHANG, M.; YAO, J. T. A Rough Sets Based Approach to Feature Selection. *Proceedings of The 23rd International Conference of NAFIPS, Banff, Canada*. 2004. 434-439 p.

ZHU, P. An Axiomatic Approach to the Roughness Measure of Rough Sets. *Fundam. Inform.*, p. 463–480, 2011.

ZIARKO, W. Probabilistic Approach to Rough Sets. *Int. J. Approx. Reasoning*, Elsevier Science Inc., New York, NY, USA, v. 49, n. 2, p. 272–284, 2008. ISSN 0888-613X.