

CENTRO UNIVERSITÁRIO DA FEI  
EDILSON JOSÉ RODRIGUES

**DEFINIÇÃO E IMPLEMENTAÇÃO DE UMA SEMÂNTICA DE SUPERVALORAÇÃO  
PARA POLISSEMIA EM EXPRESSÕES ESPACIAIS**

São Bernardo do Campo

2019



EDILSON JOSÉ RODRIGUES

**DEFINIÇÃO E IMPLEMENTAÇÃO DE UMA SEMÂNTICA DE SUPERVALORAÇÃO  
PARA POLISSEMIA EM EXPRESSÕES ESPACIAIS**

Tese de Doutorado apresentada ao Centro Universitário da FEI para obtenção do título de Doutor em Engenharia Elétrica. Orientado pelo Prof. Dr. Paulo Eduardo Santos.

São Bernardo do Campo

2019

Rodrigues, Edilson José.

Definição e implementação de uma semântica de superavaliação para polissemia em expressões espaciais / Edilson José Rodrigues. São Bernardo do Campo, 2019.

113 p. : il.

Tese - Centro Universitário FEI.

Orientador: Prof. Dr. Paulo Eduardo Santos.

1. Polissemia. 2. Raciocínio Espacial Qualitativo. 3. Preposições Espaciais. I. Santos, Paulo Eduardo, orient. II. Título.

**Aluno:** Edilson José Rodrigues

**Matrícula:** 515302-8

**Título do Trabalho:** Definição e implementação de uma semântica de superavaliação para polissemia em expressões espaciais.

**Área de Concentração:** Inteligência Artificial Aplicada à Automação e Robótica

**Orientador:** Prof. Dr. Paulo Eduardo Santos

**Data da realização da defesa:** 31/01/2019

**ORIGINAL ASSINADA**

Avaliação da Banca Examinadora

São Bernardo do Campo, / / .

**MEMBROS DA BANCA EXAMINADORA**

Prof. Dr. Paulo Eduardo Santos	Ass.: _____
Prof. Dr. Plinio Thomaz Aquino Junior	Ass.: _____
Prof. Dr. Raul Cesar Gouveia Fernandes	Ass.: _____
Prof. Dr. Marcos Fernando Lopes	Ass.: _____
Prof. Dr. Marcelo Finger	Ass.: _____

A Banca Examinadora acima-assinada atribuiu ao aluno o seguinte:

APROVADO

REPROVADO

**VERSÃO FINAL DA TESE**

**ENDOSSO DO ORIENTADOR APÓS A INCLUSÃO DAS  
RECOMENDAÇÕES DA BANCA EXAMINADORA**

Aprovação do Coordenador do Programa de Pós-graduação

Prof. Dr. Carlos Eduardo Thomaz



Aos meus pais, José e Maria



## **AGRADECIMENTOS**

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Agradeço a Deus por mostrar que com Ele tudo é possível.

À CAPES, pelo apoio financeiro.

Aos meus irmãos e pais, pelo incentivo.

Ao meu orientador, Prof. Dr. Paulo E. Santos, pela disposição em me orientar.

Aos demais amigos e colegas que foram importantes na elaboração desta tese.



## RESUMO

Nesta tese apresentamos um formalismo para o tratamento da polissemia em expressões espaciais. O problema da polissemia ocorre quando uma mesma palavra pode ter sentidos diferentes, podendo levar à ambiguidade de interpretações. Por exemplo, nas frases “o vaso está sobre a mesa” e “o quadro está sobre a parede” tem-se o emprego da mesma preposição (sobre) porém com sentidos diferentes. Enquanto que na primeira frase há, implicitamente, a ideia de verticalidade, o mesmo não ocorre na segunda frase. Isso acontece pois não há um número suficiente de expressões espaciais para categorizar todos os tipos de relações espaciais entre os objetos. Sendo assim, é necessário estabelecer um meio de interpretar o sentido de uma preposição dentro de uma frase, sendo esse o propósito desta tese. Para tanto, desenvolvemos um formalismo baseado em semântica de superavaliação que leva em conta o esquema imagético, o traço semântico e o eixo espacial associados à uma preposição, como também as características semânticas dos termos que estão ligados pela preposição. Por fim, propomos um sistema computacional que fornece as interpretações possíveis, denominadas *standpoints*, e as representações visuais para uma determinada sentença espacial. Foram testadas 71 sentenças contendo a preposição inglesa *in* e o sistema classificou corretamente 87% destas sentenças, de acordo com um padrão ouro estabelecido através de gramáticas e trabalhos sobre semânticas de preposições.

Palavras-chave: Polissemia. Raciocínio Espacial Qualitativo. Preposições Espaciais



## ABSTRACT

In this thesis we present a formalism for the treatment of polysemy in spatial expressions. The problem of polysemy occurs when the same word may have different meanings and they may lead to ambiguity of interpretation. For example, in the phrases “the vase is on the table” and “the table is on the wall” one has the use of the same preposition (on) but with different meanings. Whereas in the first sentence there is implicitly the idea of verticality, the same does not occur in the second sentence. This happens because there are not enough spatial expressions to categorize all kinds of spatial relationships between objects. Thus, it is necessary to establish a means of interpreting the meaning of a preposition within a sentence, which it is the purpose of this thesis. For this, we developed a formalism based on supervaluation semantics that takes into account the imagetic schema, the semantic trait and the spatial axis associated to a preposition, as well as the semantic features of the terms that are linked by the preposition. Finally, we propose a computational system that provides the possible interpretations, denominated *standpoints*, and the pictorial representations for a certain spatial sentence. We tested 71 sentences containing the English preposition *in* and the system correctly classified 87% of these sentences, according to a gold standard established through grammars and works on prepositional semantics.

Keywords: Polysemy. Qualitative Spatial Reasoning. Spatial Prepositions



## LISTA DE ILUSTRAÇÕES

Ilustração 1 – Maçãs no Cesto . . . . .	22
Ilustração 2 – Proto-cena para a preposição <i>over</i> . . . . .	31
Ilustração 3 – Relações DENTRO, P–DENTRO e FORA. . . . .	43
Ilustração 4 – Relações TOP–DENTRO e GEO–DENTRO . . . . .	44
Ilustração 5 – Relação ESP – DENTRO. . . . .	45
Ilustração 6 – Relação NTC. . . . .	46
Ilustração 7 – Cálculo de Direções Cardinais. . . . .	46
Ilustração 8 – Esquema imagético de contenção. . . . .	51
Ilustração 9 – Algoritmo 1 - Teorias de Definição de Predicados . . . . .	82
Ilustração 10 – Árvore de hiperônimos para “vase” . . . . .	83
Ilustração 11 – Algoritmo 2 - Semântica Standpoints para Polissemia . . . . .	83
Ilustração 12 – Exemplo de proto-cenas . . . . .	83
Ilustração 13 – Possíveis interpretações para “à esquerda”. . . . .	98
Ilustração 14 – Atribuições de frente intrínseca. . . . .	98
Ilustração 15 – Atribuições de frente extrínseca. . . . .	99
Ilustração 16 – Alteração na percepção de frente intrínseca. . . . .	99
Ilustração 17 – Inclusão total . . . . .	102
Ilustração 18 – Inclusão parcial . . . . .	102
Ilustração 19 – Inclusão por tolerância . . . . .	102
Ilustração 20 – Inclusão de um líquido . . . . .	103
Ilustração 21 – Trajetor sobre uma parede . . . . .	103
Ilustração 22 – Marco em forma de anel . . . . .	103
Ilustração 23 – Encapsulamento . . . . .	103
Ilustração 24 – Marco como extremidade do corpo-humano . . . . .	103
Ilustração 25 – Inclusão total em uma bolsa . . . . .	104
Ilustração 26 – Inclusão parcial em uma bolsa . . . . .	104
Ilustração 27 – Inclusão em um meio líquido . . . . .	104
Ilustração 28 – Trajetor é um objeto fixador . . . . .	104
Ilustração 29 – Marco como suporte . . . . .	105
Ilustração 30 – Inclusão em um compartimento de armazenagem . . . . .	105
Ilustração 31 – Inclusão em um veículo . . . . .	105
Ilustração 32 – Inclusão em um tubo . . . . .	105
Ilustração 33 – Marco como assento . . . . .	105
Ilustração 34 – Inclusão de uma vala . . . . .	106
Ilustração 35 – Inclusão em um contentor aberto . . . . .	106



## LISTA DE TABELAS

Tabela 1 – As relações do CCR. . . . .	41
Tabela 2 – Relações baseadas no fecho convexo. . . . .	42
Tabela 3 – Relações baseadas no CDC . . . . .	47
Tabela 4 – Relações laterais e transversais menos rígidas . . . . .	47
Tabela 5 – Relações do eixo vertical . . . . .	48
Tabela 6 – Resultados por TDP . . . . .	75
Tabela 7 – Tipos de contenção e expressões relacionadas. . . . .	95
Tabela 8 – Relações de verticalidade. . . . .	96
Tabela 9 – Relações de contato. . . . .	97
Tabela 10 – Expressões projetivas. . . . .	100



## SUMÁRIO

<b>1</b>	<b>Introdução</b>	19
<b>2</b>	<b>Revisão de Literatura</b>	27
2.1	Semântica de Supervaloração	27
2.2	Preposições Espaciais	29
<b>2.2.1</b>	<b>O modelo de polissemia baseado em princípios</b>	29
<b>2.2.2</b>	<b>Linguagem e cognição espacial</b>	34
2.3	Representações visuais	38
2.4	Considerações Finais do Capítulo	39
<b>3</b>	<b>Raciocínio Espacial Qualitativo</b>	41
3.1	Cálculo de Conexão de Regiões	41
3.2	Cálculo de Direções Cardinais	46
<b>3.2.1</b>	<b>Relações de Verticalidade</b>	47
<b>4</b>	<b>Esquemas Imagéticos</b>	49
4.1	Esquemas Imagéticos em Linguagem	53
<b>5</b>	<b>Semântica Standpoint para Polissemia</b>	59
5.1	Características Semânticas	59
<b>6</b>	<b>Implementação da Semântica Standpoint para Polissemia</b>	67
6.1	Modelo de precisificação espacial para a preposição <i>in</i>	67
<b>6.1.1</b>	<b>Algoritmo para o SSP</b>	72
<b>6.1.2</b>	<b>Testes</b>	74
<b>6.1.3</b>	<b>Representação visual</b>	76
6.2	Modelo de precisificação para a preposição “em”	77
<b>6.2.1</b>	<b>Sentenças com a preposição “em”</b>	79
6.3	Considerações finais do Capítulo	81
<b>7</b>	<b>Discussão e Considerações Finais</b>	85
7.1	Considerações Finais	89
	<b>APÊNDICE A – Expressões Espaciais</b>	91
A.1	Contenção	92
<b>A.1.1</b>	<b>Inclusões Total e Parcial</b>	92
<b>A.1.2</b>	<b>Inclusão em um meio</b>	93
<b>A.1.3</b>	<b>Inclusão trajetora vazia</b>	93
<b>A.1.4</b>	<b>Outros tipos de inclusão</b>	94
A.2	Verticalidade	95
A.3	Contato	96
A.4	Esquema Direita-Esquerda e Frente-Trás	97
<b>A.4.1</b>	<b>Quadros de referência</b>	97
	<b>APÊNDICE B – Proto-Cenas para a preposição “in”</b>	101

**REFERÊNCIAS . . . . . 107**

## 1 INTRODUÇÃO

A polissemia ocorre quando uma palavra apresenta diversos sentidos relacionados entre si, enquanto que na homonímia há a mesma grafia para significados diferentes. Por exemplo, a palavra vela é polissêmica, pois apresenta os seguintes significados, de acordo com o Dicionário Michaelis:

- 1) Peça geralmente cilíndrica, de cera ou outra substância gordurosa, com um pavio central em toda a extensão, e cuja chama serve para iluminar.
- 2) Qualquer medida de luminosidade que fornece a potência de uma fonte de iluminação.
- 3) Antiga unidade de medida da intensidade da luz.
- 4) Dispositivo que, nos motores de explosão, produz a ignição da mistura combustível nos cilindros; vela de ignição.

É importante observar que nos exemplos acima, todos os significados estão relacionados entre si através das ideias de iluminação e combustão. Além disso, a estrutura morfológica da palavra vela é a mesma em todos os exemplos, isto é, apresentam o mesmo radical (vel-) e a vogal temática “a”. Também pode-se dizer que os outros significados derivam de um significado mais antigo, que no exemplo de vela é o de “objeto cilíndrico de cera com um pavio”. Portanto, para se dizer que uma palavra é polissêmica é necessário verificar se os significados derivam de um significado mais antigo e apresentam a mesma estrutura morfológica. Agora, observe os seguintes exemplos:

- 1) A família vela os seus mortos;
- 2) A vela do barco.

Nestes casos, temos uma situação de homonímia, pois os significados da palavra vela nos dois exemplos não estão relacionados com as definições de vela do exemplo anterior. Além disso, no caso da frase “a família vela os seus mortos”, vela é um verbo e não um substantivo. Já no exemplo da vela do barco, o sentido etimológico é de “véu, cortina”. Assim, apesar de terem a mesma grafia, não pode-se dizer que nestes dois casos há uma extensão de sentido em relação à primeira definição de vela. Portanto, a homonímia ocorre quando a grafia da palavra é a mesma mas os seus sentidos não estão relacionados entre si. No âmbito desta tese, e para o formalismo apresentado, não se fará distinção entre polissemia e homonímia.

O objetivo desta tese é apresentar um formalismo para o tratamento da polissemia em expressões espaciais baseado em semântica de superavaliação. Uma semântica de superavaliação fornece os significados de uma proposição através de precisificações, que são as condições onde a proposição é considerada como verdadeira ou falsa. A principal motivação para esta tese é que uma única expressão espacial pode apresentar uma vasta gama de interpretações (OLIVEIRA, 2009). Por exemplo, pode-se facilmente, ao ouvir uma frase do tipo “as flores estão sobre a mesa”, inferir o conhecimento de que há um conjunto de flores em cima da mesa, provavelmente dentro de um vaso. É importante observar que, neste exemplo, conceitualiza-se a

mesa como uma superfície, ignorando o fato de ela estar sendo sustentada por pernas, e que as flores estão localizadas na parte superior desta superfície. Pode-se supor também que essas flores estão em um ramalhete sobre a mesa, ou ainda que há uma toalha entre a mesa e o vaso de flores. Todas essas interpretações são possíveis pois a preposição espacial “sobre” está diretamente ligada ao esquema de suporte (mais adiante nesta tese será introduzido o conceito de esquema imagético para definir os tipos de uso de uma preposição). Além disso, o conhecimento das características geométricas dos objetos envolvidos na frase ajuda a entender como as “flores” estão dispostas sobre a “mesa”.

Agora considere um outro exemplo: “o quadro está sobre a parede”. O que difere no uso da preposição no primeiro exemplo em relação ao segundo é a questão da verticalidade. Enquanto que no primeiro exemplo, há uma relação de verticalidade entre os objetos, no segundo exemplo isso não ocorre. Não se pode dizer simplesmente que o “quadro está em cima da parede”, apesar de a parede, assim como a mesa, estar suportando o peso do quadro. Repare que o esquema de suporte acontece em ambos os casos, mas o que difere o emprego da preposição é o eixo espacial. Enquanto há uma relação de verticalidade entre “flores” e “mesa”, o mesmo não ocorre entre “quadro” e “parede”. Assim, tem-se usos diferentes para o mesmo termo e para um mesmo esquema.

Mas então o que faz perceber que as flores estão acima da mesa e o quadro colocado sobre a parede? A resposta é simples: temos a ideia de que a mesa é uma superfície plana horizontal, enquanto a parede é caracterizada por ser uma superfície não horizontal.

Há também casos em que a preposição pode ser usada em esquemas diferentes. Veja os seguintes exemplos:

- 1) O anel está na caixinha;
- 2) O anel está no dedo.

No primeiro caso, o termo “caixinha” se refere a um recipiente, e pode-se entender a frase como o anel estando no interior da caixinha, remetendo a um esquema de contenção. Já no segundo exemplo, não pode-se dizer que o anel está no interior do dedo. Assim, a mesma preposição indica dois esquemas diferentes.

Esses usos diferentes para um mesmo termo acontecem, pois os idiomas não possuem palavras suficientes para expressar todas as relações espaciais entre os objetos. A essa diversidade de sentidos para um mesmo termo dá-se o nome de polissemia. A grande questão que é abordada nesta tese é como tratar a polissemia em expressões espaciais, com o intuito de estabelecer uma forma de comunicação entre agentes humanos e robóticos que seja feita de forma inequívoca. A ideia principal é trazer um formalismo que possa identificar qual o esquema imagético e o eixo espacial que estão sendo evocados em uma frase, através das características dos objetos que estão sendo ligados pela preposição. Portanto, o sistema que se propõe nesta tese terá de ser capaz de buscar um conhecimento pré-estabelecido que uma mesa é uma superfície horizontal, que uma parede é uma superfície não horizontal, que caixinha é um recipiente, que dedo não é um recipiente.

Para tanto, o conceito de esquemas imagéticos será trabalhado para se determinar os casos de uso de uma expressão espacial. Um esquema imagético é uma estrutura recorrente dentro do processo cognitivo que estabelece padrões de entendimento e raciocínio. São versões esquemáticas de imagens e derivam de representações de experiências sensório-motoras do homem com o mundo. Por exemplo, para o esquema de Contenção pode-se atribuir preposições tais como “em”, “dentro de”, “fora de”.

Além dos esquemas imagéticos e dos eixos espaciais, um outro fator que influencia na aplicação de uma expressão espacial é o traço semântico que ela representa. No exemplo das flores sobre a mesa, o esquema é de suporte, o eixo é de verticalidade e traço semântico representado por “sobre” é o de superioridade (acima de). Um outro exemplo seria o da preposição “sob”, que pode remeter à ideia de suporte, em um eixo de verticalidade, mas com um traço semântico de inferioridade (abaixo de). Portanto, o formalismo proposto nesta tese classificará uma expressão espacial de acordo com o esquema imagético, o eixo espacial e o traço semântico representado por ela.

A segunda parte desse formalismo é classificar os termos da sentença que representam os objetos envolvidos na cena de acordo com suas características semânticas (semas). Por exemplo, o objeto “mesa” possui a característica semântica de superfície enquanto que o objeto “caixinha” é caracterizado semanticamente como um contentor. O sema é a menor unidade de significado de uma palavra. Assim, o objetivo é estabelecer classes de objetos baseados nos semas das palavras que representam tais objetos. Por exemplo, para a classe de objetos contentor podemos ter os objetos caixa, copo, cesto, jarro, etc.

Por fim, é preciso se determinar a localização precisa de um objeto em relação a um outro. Considere os seguintes exemplos:

- 1) As flores no vaso;
- 2) O leite no copo;
- 3) As maçãs no cesto.

No primeiro exemplo, o trajetor “flores” muito provavelmente estará contido parcialmente dentro do vaso, ou seja, somente uma parte das flores (o caule) está de fato dentro do vaso. As pétalas estão fisicamente fora do vaso, caracterizando uma inclusão parcial do trajetor no marco, isto é, apenas uma parte do trajetor está realmente dentro do marco. No exemplo 2, pelo fato do trajetor (leite) ser um líquido, pode-se concluir que ele está inteiramente contido dentro do marco (copo). Assim, tem-se um caso de inclusão total. Agora, no exemplo 3, imagine a situação em que uma cesta está cheia de maçãs, conforme a Figura 1<sup>1</sup>:

Tem-se, nesse caso, algumas maçãs que estão totalmente dentro do cesto, enquanto que a maçã que está no topo está fisicamente fora do cesto, porém pode-se dizer que ela está no cesto. Isso ocorre pois a pilha de maçãs é vista como uma única massa, e a maçã do topo está em contato com esta massa, que sua por vez está dentro do cesto. Em outras palavras, o limite do cesto foi extrapolado e passou a incluir a pilha de maçãs. Assim, tem-se um caso de inclusão

<sup>1</sup>Disponível em: [https://pt.pngtree.com/freepng/apple-luokuang-li\\_2749469.html](https://pt.pngtree.com/freepng/apple-luokuang-li_2749469.html)

Figura 1 – Maças no Cesto



indireta, que ocorre quando um objeto está fisicamente fora mas em contato com algo contido por um outro objeto.

A esses possíveis arranjos espaciais entre trajetor e marco, tais como inclusão total, parcial ou indireta dá-se o nome de *precisificação*. Assim, a interpretação de uma sentença contendo uma descrição espacial depende do esquema imagético, do eixo espacial e dos traços semânticos da preposição, bem como das características semânticas do trajetor e marco, e também das precisificações entre trajetor e marco. Nesta tese propomos um formalismo, baseado nesses conceitos, para estabelecer o conjunto de interpretações de sentenças espaciais e conseqüentemente tratar a polissemia de preposições espaciais. Vale ressaltar que, nesta tese, abordar-se-á apenas usos não metafóricos das preposições e limitamos as sentenças espaciais que não envolvam movimento.

Ademais, o conhecimento espacial é uma habilidade essencial para que um humano possa se comunicar e inferir conhecimento sobre objetos no espaço. Uma forma de compartilhar uma informação espacial é pelo uso de preposições para descrever as relações entre entidades em um ambiente. Geralmente, as relações entre objetos são referidas através de expressões não numéricas (por exemplo: “a xícara está em cima da mesa” ou “a cadeira está na sala”) ao invés de termos numéricos. Isso exige o uso de relações qualitativas para representar e definir formalmente os termos espaciais. Segundo Ligozat (2013), o Raciocínio Espacial Qualitativo (REQ) é uma área de pesquisa em Inteligência Artificial cujo objetivo é formalizar o espaço em termos de entidades elementares e relações primitivas, com o intuito de se aproximar da compreensão humana do espaço.

Para Moratz e Ragni (2008), o Raciocínio Espacial Qualitativo abstrai detalhes métricos do mundo físico e permite que computadores façam previsões sobre as relações espaciais, mesmo quando informações quantitativas precisas estão indisponíveis. Do ponto de vista prático, o REQ é uma abstração que resume estados quantitativos similares em uma caracterização qualitativa. Uma visão complementar, a partir da perspectiva cognitiva, é que o método qualita-

tivo compara os recursos dentro do domínio do objeto em vez de medi-los em termos de alguma escala externa artificial. Esta é a razão pela qual as descrições qualitativas são bastante naturais para os seres humanos.

Em particular, será utilizado neste trabalho um formalismo de REQ denominado Cálculo de Conexão de Regiões (CCR) (RANDELL; CUI; COHN, 1992). O CCR está baseado em uma relação primitiva reflexiva e simétrica entre regiões espaciais, representada por  $C(x,y)$ , onde  $x$  e  $y$  são regiões espaciais que possuem pelo menos um ponto em comum.

Para a aplicação do formalismo de REQ, considera-se que é importante entender como é feito o processo cognitivo na percepção espacial, pois isso afeta diretamente a forma de se relacionar com objetos em um cenário. Por exemplo, como se pode determinar se um objeto está localizado acima ou abaixo de um outro objeto, ou se está dentro ou fora, perto ou longe ou ainda se um recipiente está cheio ou vazio. Assim, é destacado o conceito de esquema imagético, que é uma estrutura recorrente dentro dos processos cognitivos que estabelecem padrões de compreensão e raciocínio. Esquemas imagéticos são formados a partir de interações corporais, experiência linguística e do contexto histórico (LANGACKER, 1987).

Uma vez definidos os esquemas imagéticos que estão relacionados às percepções espaciais, o Cálculo de Conexões de Regiões é estendido de modo que seja possível modelar as relações espaciais com definições formais não ambíguas. Um problema a ser tratado é que o termo que representa a relação espacial entre os objetos pode apresentar mais de um significado (termo polissêmico) e pode levar o formalismo a uma interpretação ambígua. Por exemplo, a frase “o pinguim está na geladeira” pode ser interpretada de três formas: o pinguim pode estar dentro ou em cima da geladeira, ou ainda aderido à porta da geladeira (se “pinguim” se referir a um ímã em formato de pinguim). Para tratar este problema será utilizado um modelo de semântica de superavaliação denominado *standpoint* (ou estrutura de crença). Uma semântica de superavaliação vê uma linguagem como um conjunto de significados distintos em si, onde cada significado é denominado precisificação (FINE, 1975; SANTOS; BENNETT; SAKELLARIOU, 2005). Um *standpoint* pode ser modelado por uma estrutura que engloba três conjuntos: i) mundos possíveis compatíveis com a crença do agente; ii) precisificações aceitáveis pelo agente; iii) teorias de definição que especificam diferentes formas em que os termos vagos podem ser representados em relação a alguma combinação lógica de limiares restritivos (BENNETT, 2011). Ressalta-se que o *standpoint* foi concebido para lidar com termos vagos. Neste trabalho, é proposta uma alteração para que se possa aplicá-lo em termos polissêmicos. Inicialmente, Bennett (2011) propõe um formalismo que determina limiares numéricos à predicados vagos. Por exemplo, para o predicado *Alto*, são definidas as precisificações onde pessoas são consideradas altas ou baixas. Essas precisificações são atribuições de limiares numéricos ao predicado. No presente trabalho, modificou-se a ideia de precisificação para que cada predicado possua um conjunto de arranjos espaciais, determinados por fórmulas do CCR. Ademais, Bennett (2011) propõe uma linguagem de primeira ordem baseado em funções de medições para definir os predicados vagos, denominado Teorias de Definição de Predicados (TDP). Neste

trabalho, as TDP são modificadas para que a linguagem de primeira ordem possa definir os predicados polissêmicos através das características semânticas dos termos que estão ligados pela preposição espacial.

A habilidade de localizar objetos no mundo é uma das mais básicas competências requeridas para um organismo vivo (COVENTRY; GARROD, 2004). Igualmente, ser capaz de descrever onde objetos estão localizados usando descrições locativas simples é considerada habilidade básica para qualquer falante de uma linguagem (CARLSON, 2000). Expressões espaciais estão presentes no cotidiano e ocorrem em uma ampla variedade de contextos: localização de objeto, representação de cenário, compreensão de conceitos de colocação de objetos. A formalização inequívoca de expressões espaciais, portanto, é certamente uma difícil e ainda aberta questão. Por exemplo, de acordo com Bowerman (1996), preposições espaciais estão entre os mais difíceis termos para que um estudante de uma linguagem estrangeira possa adquirir. Isso acontece devido ao fato que as línguas diferem na maneira de mapear termos linguísticos para conceitos espaciais. Além disso, linguagens naturais codificam apenas um número limitado de relações espaciais entre objetos e elas precisam cobrir todas as possibilidades. Isso está relacionado diretamente à questão da polissemia, que é um dos mais difíceis problemas em semânticas formais (LANDAU; JACKENDOFF, 1993).

A principal contribuição desta tese é a apresentação de um formalismo denominado Semântica Standpoint para Polissemia (SSP), que trata a polissemia em expressões espaciais utilizando conceitos de semântica de superavaliação, mundos possíveis e características semânticas. A semântica de superavaliação define as condições em que uma proposição é verdade ou falso, chamadas de precisificações. No SSP, as precisificações são fórmulas do CCR e determinam a localização espacial dos objetos, representados pelos termos de uma frase, num plano Euclidiano. Os *mundos possíveis* definem o modo como uma expressão espacial pode ser utilizada e são compostos de esquemas imagéticos, características semânticas e eixos espaciais. Assim, os mundos possíveis caracterizam os diversos sentidos que uma preposição pode ter. Por fim, as características semânticas são as unidades básicas de significado de um determinado termo. De posse destes conceitos, foi estabelecida uma linguagem de primeira ordem que relaciona as características semânticas aos predicados espaciais, que por sua vez representam expressões espaciais. Cada fórmula dessa linguagem é denominada Teoria de Definição de Predicado (TDP). O intuito do SSP é estabelecer os *standpoints* para uma determinada sentença. Assim, um *standpoint* é uma combinação de mundos possíveis, precisificações e Teoria de Definição de Predicado (TDP) aceitáveis para uma frase contendo uma expressão espacial.

Neste trabalho foi implementado um sistema computacional para o SSP tendo a preposição inglesa *in* como teste. Com base nos trabalhos de Herskovits (1987) e Tyler e Evans (2003), foram definidos cinco mundos possíveis para a preposição *in*, representando seu caráter polissêmico. Além disso, foram levantadas em dicionários de uso e na *WordNet 37* características semânticas para os termos que estão ligados pela preposição (denominados trajetor e marco). Com essas características semânticas foi possível estabelecer as TDP e então definir o conjunto

de regras para a linguagem de primeira ordem para modelar a relação entre trajetor e marco utilizando a preposição *in*. Também foram definidas oito precisificações com o intuito de definir a posição no espaço do trajetor em relação ao marco. Assim, foram levantadas frases contendo a preposição *in* no corpus COCA<sup>2</sup> e testadas em um sistema computacional implementado através da biblioteca para processamento de linguagem natural NLTK na linguagem Python. Esse sistema extrai as características semânticas do trajetor e do marco e compara com cada uma das TDP, retornando o conjunto interpretação de uma frase. Para tanto, foi utilizada uma função de similaridade disponível na NLTK, que compara o nível de similaridade entre duas palavras na árvore de hiperônimos, retornando um valor entre 0 e 1, sendo que, quanto mais próximo de 1, mais similar são as palavras. Por exemplo, as palavras “Basket” e “Container” possuem um grau de similaridade de 0.5, permitindo dizer que “Basket” é um tipo de “Container”. Assim, as 37 características semânticas formam classes de palavras e as TDP são utilizadas para classificá-las de acordo com a classe. Ademais, para cada elemento no conjunto interpretação é apresentada uma representação visual (ou cena) para avaliar visualmente se o uso da preposição é possível para uma determinada frase. Além da implementação do SSP para a preposição *in* foi apresentada uma proposta de implementação para a preposição “em”, elencando os mundos possíveis, as TDP e as precisificações para esta preposição. Não foi implementado um sistema computacional para o teste da preposição “em” pois a biblioteca NLTK não tem suporte para a língua portuguesa.

A principal motivação para a realização deste trabalho foi a necessidade de comunicação entre agentes humanos e robóticos de forma inequívoca. Segundo Nerlich (2003), a polissemia não é um fator de ambiguidade na linguagem falada, porém a ambiguidade se torna latente na comunicação entre humanos e agentes robóticos. Retomando o exemplo do “pinguim na geladeira”, é importante que um agente robótico seja capaz de perceber todas as possibilidades de interpretação para tal frase. Assim, o SSP é uma contribuição para lidar com as possíveis ambiguidades de interpretação em expressões espaciais.

Esta tese está dividida da seguinte forma: no Capítulo 2 é apresentada uma revisão da literatura sobre os trabalhos que serviram de base para esta tese; o Capítulo 3 mostra uma revisão sobre os formalismos de REQ, que serão utilizados posteriormente para se estabelecer as precisificações; no Capítulo 4 o conceito de esquema imagético é apresentado; o Capítulo 5 mostra a formalização do SSP através dos modelos de precisificação espacial; no Capítulo 6 é mostra-se a implementação do SSP para a preposição “in” e também uma proposta de implementação para a preposição “em”; no Capítulo 7 são apresentadas as considerações finais da tese e os trabalhos futuros; por fim, no Apêndice 7.1 é apresentado um estudo sobre as preposições espaciais em português e no Apêndice A.4.1 são mostrados os exemplos e as protocenas para a preposição “in”.

---

<sup>2</sup>Disponível em <https://corpus.byu.edu/coca/>.



## 2 REVISÃO DE LITERATURA

### 2.1 SEMÂNTICA DE SUPERVALORAÇÃO

A formalização de semânticas de supervaloração foi inicialmente introduzida por Van Fraassen (1966) para linguagens com lacunas de valores verdade geradas por termos singulares (nomes ou descrições definidas). A ideia de supervaloração ocorreu previamente no trabalho de Mehlberg (1958), de maneira informal, para o tratamento da vagueza. Fine (1975) apresentou formalmente o conceito de semântica de supervaloração. Keefe (2008) apresenta uma visão geral da teoria de supervaloração de vagueza. De acordo com essa teoria, uma proposição é verdadeira se e somente se é verdade em todas as formas de torná-la precisa. Isso produz predicções de caso limítrofes que não são nem verdadeiras nem falsas, mas ainda assim a lógica clássica é preservada quase inteiramente. O artigo de Keefe (2008) apresenta a visão e alguns de seus méritos e a compara brevemente com outras teorias de vagueza.

O trabalho de Bennett (2001a) examina as maneiras pelas quais a interpretação dos conceitos espaciais é afetada pela vagueza e sugere mecanismos para levar isso em conta nos sistemas de informação espacial. A teoria da semântica de supervaloração é explicada e aplicada ao domínio espacial e a problemas particulares de definição de conceitos geográficos como “floresta”. Em Santos, Bennett e Sakellariou (2005) é descrita uma ontologia para recursos de águas interiores construída usando análise formal de conceito e semântica de supervaloração. O primeiro é usado para gerar uma rede completa do domínio de água, enquanto a semântica de supervaloração é usada para modelar a variabilidade dos conceitos em termos de parâmetros de um limiar. Bennett (2001b) utiliza a semântica de supervaloração para definir o conceito vago de floresta.

Cobrerros (2008) diz que as lacunas de valores de verdade ajudam a caracterizar a vagueza como um fenômeno semântico. Se Tim for um caso limítrofe do predicado “magro”, a frase “Tim é magro” não tem valor de verdade. A ideia subjacente para as lacunas do valor de verdade é que quaisquer fatos (fatos sobre o uso, provavelmente) determinam o significado do predicado “magro”, deixam em aberto a questão de se Tim é magro. A teoria supervaloracionista fornece uma maneira de entender a vagueza como um fenômeno semântico. Existem várias maneiras em que um predicado vago como “magro” pode ser feito preciso de uma forma consistente com o uso real que se faz da expressão. Para uma série de casos, qualquer maneira de tornar preciso o predicado seria verdadeira; por exemplo, “Gandhi é magro” (supondo que ele ainda estivesse vivo) é verdade em qualquer forma de tornar preciso o predicado “magro”, que é consistente com o uso real que se faz da expressão. Para uma série de casos, as sentenças correspondentes seriam falsas de qualquer maneira de tornar preciso o predicado. Mas haverá uma série de casos para que algumas dessas maneiras de tornar preciso renderiam uma declaração verdadeira e alguns uma falsa. A teoria supervaloracionista entende que não há nada no mundo ou no uso de um predicado vago (ou em qualquer outro fator relevante para o significado

do predicado) que selecione um determinado modo de tornar preciso o predicado vago sobre os outros. Assim, qualquer maneira de tornar preciso o predicado vago é tão correta quanto qualquer outra. Este é o sentido da ideia de indecisão semântica: nada no mundo, seja no uso ou em qualquer outro fator relevante para a determinação do significado de um predicado vago, decide de quais maneiras pode-se tornar preciso o predicado. Assim como “Tim é magro” é verdade para algumas maneiras de tornar “magro” preciso, mas falso para os outros, “Tim é magro” não é verdadeiro nem falso quando atende-se ao significado completo do predicado (COBREROS, 2008).

Cobrerros (2008) faz duas observações sobre a verdade e formas de precisão. O parágrafo anterior contém implicitamente a noção supervaloracionista de super-verdade: para a superavaliação, uma sentença é verdadeira (super-verdade), para o caso de ser verdade em toda a precisificação admissível. Uma precisificação é uma atribuição de valor de verdade (verdadeiro ou falso, mas não ambos) a todas as sentenças da linguagem. Intuitivamente, uma precisificação seria alcançada se tornar todas as expressões da linguagem precisas (cada sentença receberia um valor de verdade). Mas nem toda precisificação é admissível. Na teoria supervaloracionista existem algumas restrições sobre as precisificações que desempenham um papel interessante na teoria. Em primeiro lugar, uma precisificação deve respeitar os casos claros (ou os não-casos claros): qualquer precisificação atribuída à sentença “Gandhi é magro”, o valor-verdade falso não é uma precisificação admissível. Em segundo lugar, as precisificações devem respeitar as conexões semânticas que podem se sustentar mesmo em casos limites. Por exemplo, se Tim é um pouco mais magro que Tom, qualquer ajuste que conte Tom como magro deve contar Tim como magro; assim, a frase “Se Tom é magro, então Tim é magro” é super-verdade, mesmo que Tim e Tom estejam no limite de magro. A acomodação dessas conexões semânticas é frequentemente invocada como uma vantagem da semântica de superavaliação sobre as teorias de vagueza que admitem lacunas de valores de verdade.

Neste trabalho o conceito de superavaliação foi estendido para o tratamento da polissemia. Para tanto, foi utilizada a noção de precisificação para determinar os sentidos que uma expressão espacial pode ter dentro de uma frase. Sendo assim, as precisificações serão atribuições de valores verdade às expressões espaciais polissêmicas. Assim como na vagueza nem todas as precisificações são admissíveis, o mesmo ocorre com a polissemia. Por exemplo, pode-se precisificar a preposição “em” como “dentro de”, “sobre”, “embaixo de”, mas não como “fora de”. A ideia então é verificar quais os usos polissêmicos de uma expressão espacial e a partir daí estabelecer as precisificações, ou seja, as situações onde o emprego de tal preposição possa ser admissível. Na próxima seção serão apresentados os trabalhos que tratam de preposições espaciais, e que foram a base para o desenvolvimento das precisificações do formalismo de semântica de superavaliação SSP.

## 2.2 PREPOSIÇÕES ESPACIAIS

Destacam-se nesta seção alguns trabalhos que lidam com preposições espaciais. Em Retz-Schmidt (1988), os princípios que envolvem o uso intrínseco, dêitico e extrínseco de preposições espaciais são examinados a partir de abordagens linguísticas, psicológicas e de inteligência artificial. Retz-Schmidt (1988) introduz o sistema de diálogo de linguagem natural CITYTOUR, que pode lidar com o uso intrínseco, dêitico e extrínseco de preposições espaciais, e compara-o com outros sistemas de inteligência artificial.

O trabalho de Rice (1992) traz uma análise preliminar, a partir de uma perspectiva da linguística cognitiva, do significado de três preposições de alta frequência em inglês, *in*, *on* e *at*, que são argumentadas como inerentemente polissêmicas. Coventry e Garrod (2004) diz que o uso de preposições espaciais traz uma compreensão implícita das relações funcionais entre os objetos em si e a interação humana com esses objetos. Esta é a tese rigorosamente explicada pelo autor. Ele visa dar conta não apenas da compreensão teórica das relações espaciais, mas da capacidade de interceder com eficácia no mundo dos objetos espacialmente relacionados. Van der Gucht, Willems e De Cuypere (2007) fazem um estudo do “modelo de polissemia baseado em princípios” para preposições espaciais de Tyler e Evans (2003). O artigo enfoca os critérios de Tyler e Evans (2003) para postular uma lista de 15 sentidos distintos da preposição *over*. Van der Gucht, Willems e De Cuypere (2007) argumentam, em primeiro lugar, que o modelo cognitivo em geral concebe erroneamente significados proposicionais em termos de significados léxicos em vez de instrumentais, e que os alegados sentidos distintos da preposição sobre o postulado de Tyler e Evans (2003) são de fato significados de enunciação de frases e cláusulas inteiras; isso explica o número ainda elevado de sentidos distintos atribuídos ao item preposicional. Em segundo lugar, Tyler e Evans (2003) tentaram ilustrar que a principal razão pela qual uma bateria de sentidos é postulada, em primeiro lugar, deriva de um critério não linguístico chamado de “iconicidade do significado corporificado”. Este critério leva o linguista a aceitar tantos sentidos distintos quanto a experiências prototípicas de senso comum comumente associadas (ou “refletidas por”) ao uso de um item linguístico específico em várias instâncias.

### 2.2.1 O modelo de polissemia baseado em princípios

Enquanto os linguistas cognitivos concordam que as palavras de sentido múltiplo devem ser representadas pelos elementos conceituais de redes de polissemia, muitos propuseram sistemas organizacionais diferentes das redes, tanto do sentido central quanto dos sentidos ampliados dentro de uma determinada categoria. Duas teorias que explicam como a categoria lexical está estruturada precisam ser mencionadas: a abordagem de especificação completa, que é um modelo de extensão semântica proposto por G. Lakoff (1989), e o modelo de polissemia baseado em princípios proposto por Tyler e Evans (2003), que é uma estrutura processual que fornece

princípios metodologicamente restritos usados para analisar redes semânticas. Embora o trabalho apresentado por G. Lakoff (1989) tenha sido uma grande influência na formação do campo da linguística cognitiva como se conhece hoje, alguns pesquisadores desafiaram vários aspectos de sua abordagem, em particular algumas especificidades na análise da preposição inglesa *over*. Tyler e Evans (2003) apontam os problemas do modelo de G. Lakoff (1989) baseando-se na noção de prototipicidade ao decidir o sentido primário, especialmente no tratamento de partículas espaciais que representam relações e processos. Isto é, elas argumentam que o modelo de G. Lakoff (1989) permite inevitavelmente que cada pesquisador tome decisões arbitrárias por “próprias preferências ou imaginação” ao escolher um sentido primário. Tyler e Evans (2003) escrevem que essas decisões foram principalmente afirmadas em vez de serem defendidas. Assim, eles argumentam que havia uma falta de metodologia restrita para estabelecer o sentido primário e, em consequência, os pesquisadores dependiam excessivamente de sua própria intuição. Para superar essa falha, o modelo de polissemia baseado em princípios apresenta um conjunto de critérios que os pesquisadores podem aplicar para distinguir os novos sentidos de partículas espaciais para determinar o sentido primário (que elas chamam de proto-cena). Tyler e Evans (2003) fornecem cinco características a considerar antes de determinar o sentido primário:

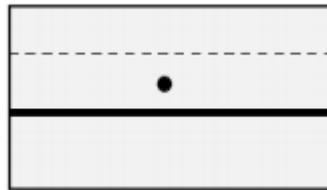
- 1) significado atestado mais antigo;
- 2) predominância na rede semântica;
- 3) relações com outras preposições;
- 4) predições gramaticais;
- 5) unidades compostas.

Tyler e Evans (2003) também enfatizam que existem estratégias importantes de inferência que um ouvinte emprega para derivar interpretações imediatas de uma dada sentença. Eles listam as três estratégias mais importantes para tal: melhor ajuste, conhecimento da dinâmica da força no mundo real e extensão topológica. Em primeiro lugar, a estratégia de inferência “melhor ajuste” explica o fato de que os falantes tentam transmitir as relações espaciais que desejam comunicar com outras pessoas escolhendo o melhor ajuste das relações conceituais a partir do conjunto limitado dos itens lexicais espaciais linguísticos que existem em uma linguagem. A segunda estratégia presume que o ouvinte e o falante supõem que todos os elementos em uma relação espacial estão sujeitos à dinâmica da força do mundo real, incluindo informações sobre gravidade e característica do marco (VANDELOISE, 1991; TALMY, 2000). A terceira extensão de noção topológica representa a capacidade de ver a relação relativística entre trajetora e marco, ao invés de noções métricas fixas como distância, tamanho, ângulo, etc (TALMY, 2000). Tyler e Evans (2003) argumentam que as estratégias de inferência acima mencionadas facilitarão uma série de interpretações imediatas derivadas de uma representação bastante abstrata de um sentido distinto. Assim, Tyler e Evans (2003) listam duas considerações para decidir se um sentido pode ser contado como um sentido distinto. A primeira é verificar se envolve um significado adicional que não é aparente em nenhum outro sentido. Em segundo lugar, verificar se

as instâncias do sentido são independentes do contexto. Para que seja independente, o sentido estendido não deve ser inferido de outro significado ou de contexto.

Outro ponto importante que Tyler e Evans (2003) postulam é que a rede de polissemia representa os sentidos distintos que estão relacionados à proto-cena em um ponto e depois diacronicamente com ela como um senso que é instanciado como uma memória semântica separada. Eles explicam que o processo de se tornar um sentido distinto ocorre através do fortalecimento pragmático, que é um processo de convencionalização iniciado de instâncias sendo extensões imediatas da proto-cena para se tornar um sentido relacionado mas distinto como resultado do uso continuado (TRAUGOTT, 1989). Esse uso continuado dá origem às barreiras definitivas dos padrões de uso que são armazenados como um sentido distinto na memória de longo prazo. Em sua reanálise da preposição *over* de G. Lakoff (1989), Tyler e Evans (2003) identificam a proto-cena, uma representação altamente abstrata de uma configuração espacial de *over*, como observado na Figura 2.

Figura 2 – Proto-cena para a preposição *over*



(a) Fonte: Tyler e Evans (2003).

O ponto é um trajetor esquemático e a linha sólida vertical é um marco esquemático. Este diagrama representa um trajetor que está mais alto que o marco, mas está ao alcance do marco (localizado abaixo da linha pontilhada) para ter potencial contato com o marco. Tyler e Evans (2003) explicam as consequências dessa configuração particular do trajetor estar localizado dentro do alcance do marco, conforme representado na proto-cena do *over* usando a noção de elemento funcional. O elemento funcional é um aspecto importante da proto-cena que Tyler e Evans (2003) argumentam ser um elemento fundamental de cada proto-cena das preposições que eles analisaram. Sua posição coincide com outros pesquisadores como Vandeloise (1991) e Herskovits (1987). No caso da preposição *over*, Tyler e Evans (2003) explicam que ela tem um elemento funcional de proximidade. Por estar fisicamente próximo ao marco, a relação entre trajetor e marco também exibe a relação funcional de controle e influência. Essa noção funcional, especialmente de proximidade na análise de *over*, se tornará um fator chave na análise dos indicadores espaciais que serão apresentados no Capítulo 6.

No geral, o modelo de polissemia baseado em princípios segue as linhas da maioria das estruturas linguísticas cognitivas, incluindo o trabalho de G. Lakoff (1989), mas ofereceu uma melhoria em relação à metodologia anterior (especialmente em comparação ao modelo de rede cognitiva de G. Lakoff (1989)) sobre como conduzir pesquisas e analisar a semântica das preposições inglesas. Especificamente, o modelo de polissemia baseado em princípios é capaz de

oferecer um método sistêmico para determinar o sentido central e uma explicação mais abrangente dos mecanismos de extensão de significado restritos envolvidos nas redes de polissemia. O modelo de polissemia baseado em princípios é similarmente significativo do ponto de vista teórico. Uma das contribuições mais importantes desta abordagem são explicações detalhadas de como uma cena espacial particular tem consequências sobre a natureza das experiências humanas envolvidas, através da correlação experiencial. O elemento funcional de cada preposição, bem como a informação geométrica, são levados em consideração. Também discutindo a noção de fortalecimento pragmático e de convencionalização de implicações, eles demonstram com sucesso a extensão do significado e o caráter sutil e complexo da mudança semântica de uma forma muito mais satisfatória.

Apesar de seu trabalho inicial sobre o modelo de polissemia baseado em princípios focular apenas as preposições inglesas, Tyler e Evans (2003) levantam a hipótese de que o modelo é aplicável a outras línguas. Com isso em mente, Shakova e Tyler (2010) testaram se o modelo poderia ser aplicado a uma linguagem diferente do inglês aplicando o modelo à preposição espacial russa *za*, que é altamente polissêmica e interage com um sistema complexo de marcação de casos. A conclusão geral de Shakova e Tyler (2010) é que o modelo de polissemia baseado em princípios pode sistematicamente explicar a complexa rede de polissemia de *za*. Além disso, Mahpeykar e Tyler (2011) analisaram a rede semântica da preposição persa *be*, altamente polissêmica, e aplicaram com sucesso o modelo de polissemia baseada em princípios. Eles identificam quinze sentidos distintos de estarem sob cinco agrupamentos individuais. A análise mostra que muitos dos significados associados a essa preposição são semelhantes à preposição inglesa, mas também há diferenças específicas de idioma. Além disso, outros idiomas foram testados com sucesso usando este modelo, incluindo alemão (LIAMKINA, 2007); polonês (KNAŚ, 2006); búlgaro (TCHIZMAROVA, 2006); e coreano (KANG, 2012).

Assim, o modelo de polissemia baseado em princípios de Tyler e Evans (2003) foi um dos trabalhos que nortearam as definições dos mundos possíveis da preposição *in*, conforme mostrado no Capítulo 6. Tyler e Evans (2003) definem o significado principal de *in* como sendo uma relação espacial em que o trajector está localizado dentro do marco, que possui três elementos estruturais salientes: um interior, um limite e um exterior. Ademais, Tyler e Evans (2003) dizem que existem relações que não envolvem marcos tridimensionais. Por exemplo, na frase “the cow munched grass in the field”, o trajector “cow” está conceitualizado como “contido” no marco “field”, mas “field” não é um marco tridimensional. Em outros casos, o marco pode ser um coletivo de indivíduos caracterizados como uma massa única, como em “the child couldn’t be seen in the crowd”. Ademais, Tyler e Evans (2003) definem outros 11 sentidos, não necessariamente espaciais, para a preposição *in*. Através desses sentidos foi possível estabelecer os cinco mundos possíveis que propomos para a preposição *in*.

Outros trabalhos também motivaram a definição dos usos das preposições espaciais, tanto em inglês quanto em português. A seguir, é apresentado um breve resumo desses trabalhos: Os trabalhos de Neves (2011), Castilho (2014) e Ilari et al. (2008) são gramáticas da língua

portuguesa, e como tal apresentam os principais tipos de uso para as preposições neste idioma. No trabalho de Oliveira (2009) é apresentado um estudo sobre a preposição “em”. Neste trabalho, Oliveira (2009) analisou 2813 ocorrências em corpus e obteve dois sentidos esquemáticos para “em”, derivados do esquema imagético de contenção: localização e especificação. Foram elencadas 22 categorias para o primeiro e 9 para o segundo. Dentre os sentidos de localização foram observadas oito categorias para localização espacial, três para o temporal e onze para outros domínios não espaciais. Para explicar a polissemia da preposição “em”, Oliveira (2009) adotou modelo teórico de linguagem baseado no uso e conceitualização de Langacker (1987); na proposta de linguagem corporificada de esquemas imagéticos de Johnson (1987); na noção de metáforas conceituais de G. Lakoff e Johnson (2008); na semântica de classes fechadas de Talmy (2000) e nas funções relacionais envolvidas na semântica de preposições espaciais de Vandeloise (1991). Outros trabalhos em língua portuguesa que foram base para a elaboração desta tese foi o de Garcia (2018) e Pontes (1992). O primeiro apresenta uma rede de preposições espaciais, denominada PrepNet, que mostra 153 usos distintos para preposições espaciais em língua portuguesa; o segundo trata-se de um livro sobre a investigação de como o espaço e tempo são codificados na língua portuguesa, através do exame de preposições, advérbios de lugar, demonstrativos e locuções adverbiais e prepositivas. Trata-se de um estudo semântico de como o espaço é pensado e como o tempo é codificado em cima do espaço.

Ademais, destacam-se outros trabalhos que foram importantes para se definir os usos das preposições *in* e “em”, mostrados no Capítulo 6. Em Cuyckens (1993) é mostrada uma análise semântica de sentidos espaciais da preposição holandesa *in*, em Hawkins (1993) é discutida a questão da universalidade no programa de pesquisa da linguística cognitiva através de um exame das estruturas semânticas no domínio das relações espaciais. Mais diretamente, há uma confrontação sobre a questão de quais estruturas semânticas características das preposições espaciais inglesas podem ser consideradas universais. Ademais, também é considerada a questão de quais fatores levam à variabilidade observada entre as linguagens dessas estruturas semânticas supostamente universais. Hottenroth (1993) discute três problemas básicos da semântica de preposições: i) preposições oferecem uma variabilidade de interpretações dependente do contexto, podendo uma única preposição expressar uma vasta gama de relações; ii) o segundo problema para uma teoria semântica de preposições surge do fato de que há uma interação semântica muito próxima entre a preposição e seu contexto linguístico, especialmente entre a preposição e seus argumentos. Não se pode pensar no significado de uma preposição espacial, por exemplo, sem ter algum lugar ou coisa em mente; iii) um terceiro problema para a semântica preposicional surge do fato de que uma mesma situação espacial pode ser vista de diferentes perspectivas, isto é, pode ser conceituada de maneiras diferentes. Por fim, Zelinsky-Wibbelt (1993) apresenta uma formalização para a tradução e interpretação de preposições. Todos esses trabalhos fornecem tipificações de uso para preposições, e que foram aproveitados para se definir as características semânticas propostas para a preposição *in* e também para a preposição portuguesa “em”. Através desses trabalhos foi possível perceber que algumas características

espaciais extrapolam a questão do idioma. Por exemplo, a ideia de contenção está presente na preposição holandesa “in”, na preposição inglesa “in” (são línguas próximas) e na preposição portuguesa “em”. Por fim, para se definir as características espaciais para as preposições *in* e “em” foi necessário relacionar a ideia de cognição espacial com os termos linguísticos. Para tanto, utilizou-se o trabalho de Herskovits (1987), que será detalhado na próxima seção.

### **2.2.2 Linguagem e cognição espacial**

Nesta seção será apresentado o trabalho de Herskovits (1987), que foi um dos pilares para a elaboração desta tese juntamente com a revisão do trabalho de Herskovits (1987) realizada por Dahl (1994). Neste livro, Herskovits (1987) propõe uma análise das expressões locativas em inglês. Seu objetivo é posicionar o estudo de expressões linguísticas (em particular, expressões espaciais) dentro de um contexto mais amplo de uso da linguagem e as convenções associadas com a comunicação de objetivos, crenças, etc. Assim, o trabalho deve ser avaliado a partir da perspectiva da ciência cognitiva como um campo interdisciplinar e não simplesmente como um tratado linguístico sobre preposições ou um modelo computacional de um subconjunto de linguagem natural. A teoria semântica subjacente que Herskovits (1987) assume difere das abordagens clássicas mais tradicionais de significado (por exemplo, (TARSKI, 1944)) e é mais semelhante às posições defendidas por Searle (1978) e Winograd (1980), e anteriormente por Wittgenstein (1953). Segundo esses autores, a visão de que o significado deriva da interpretação literal das palavras em uma expressão é inadequada, e que o que é necessário para compreender a polissemia das preposições (e de toda a linguagem) é uma teoria da informação contextual e do uso da linguagem. Herskovits (1987) incorpora esses aspectos da teoria não-clássica do significado introduzindo três noções: i) o significado ideal de uma palavra; ii) o modo como as expressões formadas a partir desta palavra podem variar devido à convenção; e iii) as influências pragmáticas da tolerância nesse significado. Herskovits (1987) classifica essas influências de acordo com uma série de tipos de uso que abstraem essas convenções e completam o mapeamento do significado lexical para o uso.

Herskovits (1987) argumenta que a teoria clássica de significados é inadequada e que é preciso enriquecer a estrutura descritiva com pares de tipos de frase, juntamente com interpretações dos casos centrais dentro de um tipo de uso. Assim, essa visão nega a interpretação padrão da tese da composicionalidade para a linguagem. Outro tema recorrente do livro de Herskovits (1987) é a distinção entre abordagens computacionais e holísticas à inteligência. A metáfora computacional afirma que um computador pode ser programado para entender completamente o inglês e incorporar inteligência em geral. A posição anti-computacional (isto é, o pensamento holístico), defendida por Dreyfus (1972) e Winograd e Flores (1986), argumenta que nenhuma especificação algorítmica pode ser dada para incorporar inteligência ou habilidades linguísticas. Essa visão afirma que as condições necessárias para entender uma frase no contexto não podem ser codificadas completa ou exaustivamente em procedimentos que uma máquina pode

executar. Além disso, Winograd e Flores (1986) sustentam que um usuário da língua não pode ter uma representação mental do mundo que ele ou ela percebe. Adotando uma posição neo-heideggeriana sobre a atividade mental, eles negam que a manipulação de símbolos desempenhe um papel em nossa compreensão do mundo e em nosso uso da linguagem. Herskovits (1987) adota essa posição até certo ponto, embora com algumas exceções importantes. Ela concorda com muitos linguistas e psicólogos cognitivos que a elaboração da teoria da construção é essencial para revelar generalizações sobre os dados e para uma melhor compreensão dos fenômenos (NELSON, 1951; CARNAP, 1967; SCHLICK, 2009). Além disso, ela adota uma forte visão representacional sobre a natureza do significado lexical (mais próxima de G. Lakoff (1989) ou Jackendoff (1983)), que está embutida em uma teoria não representativa da pragmática.

De uma perspectiva linguística, Herskovits (1987) faz as seguintes afirmações teóricas:

- 1) As preposições espaciais têm significados ideais associados a elas em suas entradas lexicais;
- 2) As maneiras pelas quais esse significado é convencionalmente explorado. Existem dois tipos de desvios do significado ideal de um termo: mudança baseada em convenção (resultando em polissemia) e processos pragmáticos de tolerância, para permitir que algo seja quase verdadeiro.

De uma perspectiva computacional, pode-se avaliar o trabalho de Herskovits (1987) de uma maneira ligeiramente diferente. Há três propostas específicas feitas no trabalho, a primeira sobre representação, as outras duas sobre análise e geração. Embora amplamente programático, o modelo computacional de Herskovits (1987) equivale a um paradigma de geração e teste para a interpretação semântica e uma mistura de abordagens de decomposição para geração. Finalmente, do ponto de vista psicológico, o trabalho pode ser visto como argumentando em favor de uma teoria prototípica de categorização (ROSCH, 1977), onde a noção de Herskovits (1987) de significado ideal é próxima de um protótipo, embora ela afirme que existem importantes diferenças entre os dois conceitos. O estudo de Herskovits (1987) é organizado em três partes. Ela primeiro apresenta um quadro descritivo para estudar preposições espaciais. Isto é seguido por uma discussão de modelos de computador para os dados apresentados na primeira seção. Finalmente, ela discute quatro estudos de caso, envolvendo as preposições “in”, “on”, “at”, e as preposições projetivas (por exemplo, “behind”, “to the right”, etc.). A visão básica pode ser resumida da seguinte maneira: existe um significado ideal de uma preposição, que é uma ideia geométrica. A partir dessa ideia geométrica, todos os usos dessa preposição são derivados por meio de operações envolvendo adaptações e mudanças nesse sentido. O próprio significado ideal é definido como uma relação entre objetos geométricos ideais, como ponto, linha, superfície, etc. Juntamente com esse significado, há um conjunto de restrições que Herskovits (1987) chama de “tipo de situação normal”. Este é o conjunto acima mencionado de casos centrais que se enquadram no tipo de uso. Embora seja um pouco vago, esses tipos de situação parecem estar associados a expressões completas e não simplesmente a itens lexicais individuais.

Existem duas maneiras para o significado de uma expressão mudar. Primeiro considere o que Herskovits (1987) chama de mudanças de sentido: suponha que a preposição “on” carrega o significado de apoio e continuidade. Isso descreve a situação em i) “the book on the table” mas não em ii) “the wrinkles on his forehead”. No entanto, a situação em ii) se assemelha ao significado ideal de alguma forma, mesmo que não haja suporte envolvido. Tal semelhança é relacionada por uma mudança de sentido. Outra maneira pela qual o significado ideal pode mudar é por tolerância. Por exemplo, considere a expressão locativa em i) novamente. Isso não seria apropriado com o significado ideal de “on” se houvesse uma toalha de mesa entre a mesa e o livro. Mas obviamente permite-se essa imprecisão na descrição das situações. Em tais casos, diz-se que o significado da expressão mudou devido à tolerância.

O significado ideal de uma preposição não se mapeia diretamente no mundo. Em vez disso, Herskovits (1987) defende um nível intermediário de conceituação geométrica, em que as funções de descrição geométrica mapeiam as descrições locativas em objetos. Em outras palavras, essas funções geométricas determinam o que a preposição contribui para o significado de uma situação particular. Muitas vezes, esse pode ser o espaço ocupado pelo objeto, seu lugar. Obviamente, a maneira como um objeto se relaciona com um espaço dependerá do objeto. Um objeto “mesa” por exemplo, identifica seu espaço de maneira diferente de um objeto “água”. Com o primeiro, o espaço é delimitado e definido, enquanto o segundo é ilimitado e indefinido. É essa função que distingue entre objetos sólidos, objetos líquidos, objetos geométricos, buracos, etc. Herskovits (1987) descreve seis outras funções geométricas descritivas (conforme Talmy (1975)), incluindo:

- 1) partes (por exemplo, bordas, bases, superfícies, partes 3-D, etc.);
- 2) idealizações (aproximações para um ponto, linha, superfície ou plano);
- 3) “*good-forms*”, (esboços, invólucros, regiões normalizadas);
- 4) volumes (interior, volume ou área associada com vértices, lâminas associadas com superfície);
- 5) eixos (eixo principal, pontos de vista associados);
- 6) projeções (num plano infinito, no solo).

Como exemplo, considere a interpretação da expressão “o pássaro na árvore”. Há uma função *good-form* que provê o fechamento *Gestalt* sobre “árvore” tal que um pássaro pode estar contido no espaço ocupado pela forma da árvore.

A principal afirmação de Herskovits (1987) é que uma simples interpretação espacial das preposições locativas é inadequada para captar a variedade de significados carregados por essas palavras. Em vez disso, precisa-se conversar em termos de como se conceitua objetos, em termos de suas extensões geométricas e espaciais. Essa visão está relacionada àquela apresentada em alguns detalhes na discussão de Hayes (1978) sobre uma física ingênua para líquidos. Também é similar em relação à interpretação de Moravcsik (1981) dos modos de explanação (*aitiae*) de Aristóteles associadas a objetos. Depois de apresentar a estrutura para a representação de preposições espaciais, Herskovits (1987) se volta para os dois problemas de decodificação e

codificação. O problema de decodificação pode ser declarado da seguinte maneira: dada uma expressão locativa usada em uma situação particular, como pode-se prever o que ela transmite? O problema da codificação é similarmente afirmado: dada uma situação com dois objetos espaciais, como pode-se descrever melhor a relação espacial obtida entre os dois objetos? Estes são simplesmente os problemas de interpretação e geração de expressões espaciais.

A maioria das abordagens computacionais para a desambiguação lexical envolve algum mecanismo de seleção entre os sentidos de palavras alternativas. Descobrir as associações entre um sentido de uma palavra e o resto da expressão ajudará na desambiguação da expressão geral. Isso pode ser feito de várias maneiras, incluindo passagem de marcadores (CHARNIAK, 1983; HIRST, 1992); passagem de mensagens (RIEGER; SMALL, 1979); semântica de preferências (WILKS, 1997) e semântica colativa (FASS, 1986). Herskovits (1987) pressupõe implicitamente que os sentidos são gerados e combinados de acordo com o contexto e as restrições situacionais. A discussão de como as expressões locativas são codificadas também é bastante programática e pouco clara. Os problemas de planejamento textual e seleção lexical no processo de geração têm sido ativamente abordados por muitos pesquisadores. Herskovits (1987) vislumbra um formalismo sistêmico para gerar descrições espaciais. Isso, de fato, parece consistente com sua discussão sobre o papel funcional das diferentes expressões e o modo como elas afetarão o ato comunicativo. Por fim Herskovits (1987) conclui seu trabalho com vários estudos de caso detalhados sobre como tantas expressões diferentes podem ser derivadas da mesma preposição, devido a adaptações e mudanças de tolerância.

Além do trabalho de Herskovits (1987), que foi a principal fonte para a categorização das preposições espaciais vistas no Capítulo 6, outros trabalhos também foram consultados, com destaque para o livro de Coventry e Garrod (2004), que traz uma estrutura funcional para explicar o significado de preposições espaciais. Coventry e Garrod (2004) dizem que, ao descrever uma cena espacial, o falante escolhe expressões que mapeiam as relações em um modelo de situação daquela cena que são consistentes com as saídas geométricas, refletindo na maneira de como se vê o mundo e as rotinas extra-geométricas dependem do conhecimento dos objetos envolvidos e da memória das relações sobre esses objetos. Assim como a abordagem de Coventry e Garrod (2004), apresentamos um formalismo que leva em conta a geometria, através de relações de REQ, e extra-geométricas tais como as características semânticas das palavras que representam os objetos.

Por fim, destaca-se também os trabalhos de Abella e Kender (1993), que apresenta um arcabouço para um sistema que descreve objetos de uma forma qualitativa, e o trabalho de Vandeloise (1991), que faz um estudo de preposições espaciais no idioma francês. A seguir, são apresentados trabalhos que relacionam descrições textuais com representações visuais, assim como se propõe nesta tese.

### 2.3 REPRESENTAÇÕES VISUAIS

Inspirado no trabalho de Ligozat, Nowak e Schmitt (2007), que utiliza a teoria de cores de Brunet (1986) para obter descrições visuais por intermédio de textos, o Semântica Standpoint para Polissemia também apresenta descrições visuais da disposição de objetos a partir de textos contendo expressões espaciais. Destacam-se os trabalhos de Chang et al. (2015), que introduz um conjunto de dados de cenas 3D anotadas com descrições de linguagem natural para realizar a aprendizagem, através desses dados, de como fundamentar descrições textuais em objetos físicos, enquanto Coyne e Sproat (2001) apresentam um sistema, denominado *WordsEyes* para converter automaticamente um texto em cenas 3D representativas. Esse sistema conta com um banco de dados de modelos 3D e representam entidades e relações. Cada modelo 3D pode ter forma, rótulos e propriedades funcionais relacionados a ele. Ademais, Coyne e Sproat (2001) descrevem uma análise linguística e as técnicas de representação utilizadas no *WordEyes*, juntamente com estratégias pelas quais conceitos mais abstratos são representados.

Johansson et al. (2005) descrevem um sistema que converte automaticamente narrativas em cenas 3D. Os textos, escritos em sueco, descrevem acidentes rodoviários. Um dos principais recursos do programa é que nele a cena gerada utiliza relações temporais entre os eventos. O sistema consiste em três módulos: interpretação de linguagem natural baseada em métodos de extração de informação; um módulo de planejamento que produz uma descrição geométrica do acidente; e um módulo de visualização que renderiza a descrição geométrica como gráficos animados. Por sua vez, Krishnaswamy e Pustejovsky (2016) desenvolveram um sistema para gerar simulações visuais tridimensionais de expressões de movimento da linguagem natural. Para tanto, eles utilizam uma *engine* para jogos denominada Unity para gerar os gráficos e processar as entradas e saídas. As entradas são sentenças simples em linguagem natural anotadas e tratadas em processadores externos, tais como o *parser* ClearNP, e referenciados aos objetos da cena. Por último, destaca-se o trabalho de Pustejovsky e Krishnaswamy (2016), que traz a especificação de uma linguagem de modelagem denominada VoxML, que codifica o conhecimento semântico de objetos do mundo real representados como modelos tridimensionais, e também de eventos e atributos relacionados a esses objetos. O objetivo do VoxML é superar as limitações existentes em linguagens de marcação visual, permitindo a codificação de uma ampla gama de conhecimentos semânticos que podem ser explorados por uma variedade de sistemas e plataformas, levando à simulações multimodais de cenários do mundo real usando objetos conceituais que representam seus valores semânticos.

Vale ressaltar que não foram adotadas nesta tese representações tridimensionais de objetos, mas sim as representações esquemáticas para a localização do trajetor em relação ao marco. Porém, esses trabalhos citados acima servirão de base para se implementar, em trabalhos futuros, um modelo de precisificação que gere imagens tridimensionais das relações espaciais polissêmicas.

## 2.4 CONSIDERAÇÕES FINAIS DO CAPÍTULO

Neste capítulo foram apresentados trabalhos relacionados com esta tese. Em primeiro lugar foram apresentados os trabalhos sobre Semântica de Supervaloração. Destacamos aqui o trabalho de Bennett (2011), que serviu de base para a elaboração do SSP. Neste trabalho, Bennett (2011) propõe uma semântica de supervaloração para o tratamento da vagueza, enquanto que nesta tese adaptou-se esse modelo para que se pudesse tratar a polissemia. Essa formalização está descrita no Capítulo 5. Para a aplicação do formalismo proposto foram escolhidas as preposições “in” e “em” e fez-se necessário entender a base cognitiva por trás dos casos de uso destas preposições. Para tanto, estudou-se o modelo de polissemia baseado em princípios de Tyler e Evans (2001) e os modelos cognitivos de Herskovits (1987) e Coventry e Garrod (2004). Além disso, para a elaboração dos mundos possíveis foram utilizadas gramáticas, dicionários de uso e outros trabalhos sobre a aplicação destas preposições, conforme visto no presente capítulo. Ademais, ao implementar o SSP para a preposição “in”, foi estabelecido um conjunto de imagens para cada interpretação de uma sentença espacial e para isso nos baseamos nos trabalhos descritos na Seção 2.3, que tratam de representações visuais a partir de descrições textuais.

Dada a amplitude do estudo da polissemia, focou-se nesta tese somente a polissemia em expressões espaciais. Para tanto, estabeleceu-se um formalismo que precisifica os significados de uma expressão espacial polissêmica através de um conjunto de características semânticas e Raciocínio Espacial Qualitativo (REQ). Assim, no próximo capítulo são introduzidos os conceitos de REQ que servirão de base para as precisificações de uma expressão espacial polissêmica.



### 3 RACIOCÍNIO ESPACIAL QUALITATIVO

O formalismo lógico apresentado neste trabalho é baseado em semântica de superavaliação juntamente com formalismos de Raciocínio Espacial Qualitativo (REQ). Neste capítulo serão abordados tais formalismos, baseados em primitivas como conectividade e fecho convexo de regiões em um plano Cartesiano. A seguir, é apresentado o Cálculo de Conexão de Regiões (CCR), baseado nos trabalhos de Randell, Cui e Cohn (1992), Cohn, Randell e Cui (1995) e Cohn et al. (1997).

#### 3.1 CÁLCULO DE CONEXÃO DE REGIÕES

O CCR tem por base uma relação diádica primitiva  $C(x,y)$ , lida como “ $x$  está conectado a  $y$ ”, onde  $x$  e  $y$  são regiões no espaço Euclidiano. Em outras palavras, as regiões  $x$  e  $y$  possuem pelo menos um ponto em comum. Ademais, a relação  $C(x,y)$  é axiomatizada como reflexiva e simétrica. Utilizando a relação  $C(x,y)$ , um conjunto de relações diádicas são definidas na Tabela 1:

Tabela 1 – As relações do CCR.

Relação	Significado
$DC(x,y) \equiv_{def} \neg C(x,y)$	$x$ está desconectado de $y$ .
$P(x,y) \equiv_{def} \forall z [C(z,x) \rightarrow C(z,y)]$	$x$ é parte de $y$ .
$PP(x,y) \equiv_{def} P(x,y) \wedge \neg P(y,x)$	$x$ é parte própria de $y$ .
$EQ(x,y) \equiv_{def} P(x,y) \wedge P(y,x)$	$x$ é igual a $y$
$O(x,y) \equiv_{def} \exists z [P(z,x) \wedge P(z,y)]$	$x$ se sobrepõe a $y$
$PO(x,y) \equiv_{def} O(x,y) \wedge \neg P(x,y) \wedge \neg P(y,x)$	$x$ se sobrepõe parcialmente a $y$ .
$DR(x,y) \equiv_{def} \neg O(x,y)$	$x$ não se sobrepõe a $y$ .
$EC(x,y) \equiv_{def} C(x,y) \wedge \neg O(x,y)$	$x$ está externamente conectado a $y$ .
$TPP(x,y) \equiv_{def} PP(x,y) \wedge \exists z [EC(z,x) \wedge EC(z,y)]$	$x$ é parte própria tangencial de $y$ .
$NTPP(x,y) \equiv_{def} PP(x,y) \wedge \neg \exists z [EC(z,x) \wedge EC(z,y)]$	$x$ é parte própria não tangencial de $y$ .

Fonte: (RANDELL; CUI; COHN, 1992)

É importante observar que as relações  $P$ ,  $PP$ ,  $TPP$  e  $NTPP$  são assimétricas e admitem inversas. As formas inversas são denotadas por  $P_i$ ,  $PP_i$ ,  $TPP_i$  e  $NTPP_i$ .

Randell, Cui e Cohn (1992) também definem um conjunto de funções Booleanas que retornam a união ( $\text{sum}(x,y)$ ), a intersecção ( $\text{prod}(x,y)$ ) e a diferença ( $\text{diff}(x,y)$ ) entre duas regiões, além do complemento ( $\text{compl}(x)$ ) de uma região em relação à uma região espacial universal  $U$ , na qual todas as regiões estão conectadas, e  $\iota$  significa a região resultante dessas funções:

$$\text{sum}(x,y) = \iota y[\forall z[C(z,y) \leftrightarrow [C(z,x) \vee C(z,y)]]] \quad (1)$$

$$\text{compl}(x) = \iota y[\forall z[[C(z,y) \leftrightarrow \neg\text{NTPP}(z,x)] \wedge [O(z,y) \leftrightarrow \neg P(z,x)]]] \quad (2)$$

$$U = \iota y[\forall z[C(z,y)]] \quad (3)$$

$$\text{prod}(x,y) = \iota z[\forall u[C(u,z) \leftrightarrow \exists v[P(v,x) \wedge P(v,y) \wedge C(u,v)]]] \quad (4)$$

$$\text{diff}(x,y) = \iota w[\forall z[C(z,w) \leftrightarrow C(z,\text{prod}(x,\text{compl}(y)))]] \quad (5)$$

Note que, devido à possibilidade de união de duas regiões quaisquer, essa união pode resultar em regiões topologicamente conectadas (uma única parte indivisível) ou desconectadas (dividida em várias partes). Assim, uma relação monádica para regiões indivisíveis é definida por:

$$\text{CON}(x) \equiv_{def} \forall yz[\text{sum}(y,z) = x \rightarrow (y,z)]. \quad (6)$$

e pode ser lida como “ $x$  é uma região com uma única parte indivisível”. Cohn, Randell e Cui (1995) descrevem uma função primitiva  $\text{conv}(x)$  (fecho convexo de  $x$ ) que retorna a menor região convexa de  $x$  e é axiomatizada como:

$$\forall x \quad P(x, \text{conv}(x)) \quad (7)$$

$$\forall x \quad P(\text{conv}(\text{conv}(x)), \text{conv}(x)) \quad (8)$$

$$\forall xyz \quad [[P(x, \text{conv}(y)) \wedge P(y, \text{conv}(z))] \rightarrow P(x, \text{conv}(z))] \quad (9)$$

$$\forall xy \quad [[P(x, \text{conv}(y)) \wedge P(y, \text{conv}(x))] \rightarrow O(x,y)] \quad (10)$$

$$\forall xy \quad [[\text{DR}(x, \text{conv}(y)) \wedge \text{DR}(y, \text{conv}(x))] \leftrightarrow \text{DR}(\text{conv}(x), \text{conv}(y))] \quad (11)$$

Ademais, o predicado  $\text{CONV}(x)$  denota que  $x$  é uma região convexa. A função  $\text{conv}(x)$  é utilizada para definir um conjunto de relações que descrevem inclusão (ou exclusão) entre regiões. Essas relações são definidas a seguir, na Tabela 2:

Tabela 2 – Relações baseadas no fecho convexo.

Relação	Significado
$\text{DENTRO}(x,y) \equiv_{def} \neg P(x,y) \wedge P(x, \text{conv}(y))$	$x$ está dentro de $y$ .
$\text{P} - \text{DENTRO}(x,y) \equiv_{def} \neg P(x,y) \wedge \text{PO}(x, \text{conv}(y)) \wedge \exists z[P(z, \text{conv}(y)) \wedge \neg P(z,y) \wedge \text{PO}(z,x)]$	$x$ está parcialmente dentro de $y$ .
$\text{FORA}(x,y) \equiv_{def} \neg P(x,y) \wedge \neg \exists z[P(z, \text{conv}(y)) \wedge \neg P(z,y) \wedge \text{PO}(z,x)]$	$x$ está fora de $y$ .

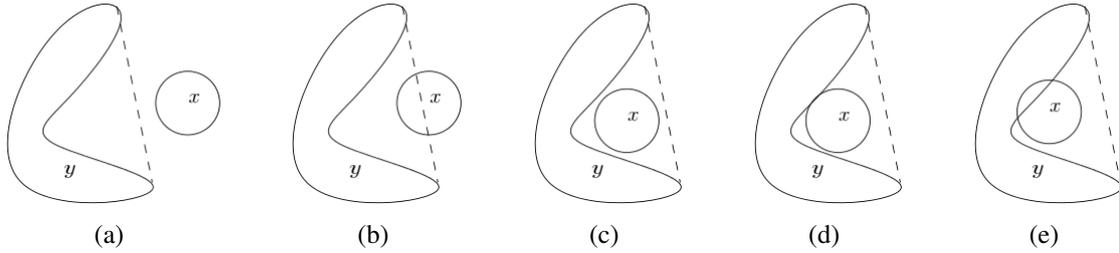
Fonte: (COHN; RANDELL; CUI, 1995)

As relações acima admitem inversas, sendo denotadas por  $\text{DENTRO}_i$ ,  $\text{P} - \text{DENTRO}_i$  e  $\text{FORA}_i$ . Cohn, Randell e Cui (1995) também definem duas funções ( $\text{dentro}(x)$ ,  $\text{fora}(x)$ ) que representam, respectivamente, as regiões interna e externa à  $x$ , e são definidas como:

$$\text{dentro}(x) = \iota y[\forall z[C(z,y) \leftrightarrow \exists w[\text{DENTRO}(w,x) \wedge C(z,w)]]] \quad (12)$$

$$\text{fora}(x) = \iota y[\forall z[C(z,y) \leftrightarrow \exists w[\text{FORA}(w,x) \wedge C(z,w)]]] \quad (13)$$

Figura 3 – Relações DENTRO, P–DENTRO e FORA.



Na Figura 3 tem-se um exemplo de aplicação das relações DENTRO, P–DENTRO e FORA. A seguir temos a descrição de cada um destes exemplos:

- (a)  $\text{FORA}(x,y) \wedge \text{FORAi}(x,y) \wedge \text{DC}(x,y)$
- (b)  $\text{P–DENTRO}(x,y) \wedge \text{FORAi}(x,y) \wedge \text{DC}(x,y)$
- (c)  $\text{DENTRO}(x,y) \wedge \text{FORAi}(x,y) \wedge \text{DC}(x,y)$
- (d)  $\text{DENTRO}(x,y) \wedge \text{FORAi}(x,y) \wedge \text{EC}(x,y)$
- (e)  $\text{DENTRO}(x,y) \wedge \text{FORAi}(x,y) \wedge \text{PO}(x,y)$

Cohn, Randell e Cui (1995) ressaltam que há duas situações em que uma região está dentro de outra: geométrica ou topologicamente. Essas relações são definidas a seguir:

$$\text{TOP–DENTRO}(x,y) \equiv_{def} \text{DENTRO}(x,y) \wedge \forall z[[\text{CON}(z) \wedge C(z,x) \wedge C(z, \text{fora}(y))] \rightarrow \text{O}(z,y)] \quad (14)$$

$$\text{GEO–DENTRO}(x,y) \equiv_{def} \text{DENTRO}(x,y) \wedge \neg \text{TOP–INSIDE}(x,y) \quad (15)$$

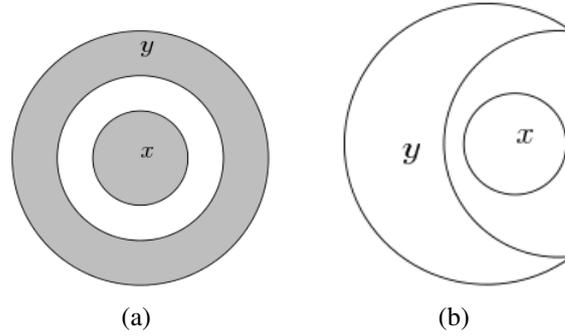
e as funções que retornam as regiões topológica ou geometricamente internas são dadas por (a Figura 4 traz, respectivamente, exemplos de inclusão topológica e geométrica):

$$\text{topdentro}(x) = \iota y[\forall yz[C(y,z) \leftrightarrow \exists w[\text{TOP–DENTRO}(w,x) \wedge C(z,w)]]] \quad (16)$$

$$\text{geodentro}(x) = \iota y[\forall yz[C(y,z) \leftrightarrow \exists w[\text{GEO–DENTRO}(w,x) \wedge C(z,w)]]] \quad (17)$$

O ponto importante de uma região estando topologicamente dentro de outra é que é preciso fazer um “corte” na região circundante para alcançar e fazer contato com a região contida. Na variante geométrica, este não é o caso.

Figura 4 – Relações TOP–DENTRO e GEO–DENTRO



Cohn, Randell e Cui (1995) argumentam que a noção de geometricamente dentro não é suficiente para representar a ideia de contenção. Assim, Cohn, Randell e Cui (1995) fazem um refinamento dessa relação para que situações de conteúdo possam ser representadas. Inicialmente eles definem o conceito de *tampa*, que é uma região que transforma parte de uma região internamente geométrica em topológica. A partir dessa definição é possível definir o conceito de *conteúdo interno* de uma região, que pode ser utilizado, por exemplo, para representar um líquido contido em um recipiente (Fórmulas 18 e 19).

$$\text{LID}(w,y,x) \equiv_{def} \text{CONV}(w) \wedge \text{P}(x, \text{geodentro}(y)) \wedge \text{P}(x, \text{topdentro}(\text{sum}(w,y))) \quad (18)$$

$$\text{CONT-DENTRO}(x,y) \equiv_{def} \text{P}(x, \text{geodentro}(y)) \wedge \exists w[\text{LID}(w,y,x)] \quad (19)$$

onde  $w$  é uma tampa,  $y$  um recipiente e  $x$  um conteúdo. Porém, essas definições não são suficientes para expressar todas as situações de inclusão geométrica, pois o lado geométrico interno pode ser dividido em dois (por exemplo, um *tubo*). Assim, Cohn, Randell e Cui (1995) definem outras duas relações onde uma região está geometricamente dentro de uma outra mas não inteiramente dentro de seu conteúdo (Fórmulas 20 e 21).

$$\text{T-DENTRO}(x,y) \equiv_{def} \text{GEO-DENTRO} \wedge \neg \text{PO}(x, \text{contdentro}(y)) \quad (20)$$

$$\text{G-DENTRO}(x,y) \equiv_{def} \text{GEO-DENTRO} \wedge \text{PO}(x, \text{contdentro}(y)) \quad (21)$$

e a função  $\text{contdentro}(x)$ , que retorna a região interna de um recipiente, é definida por:

$$\text{contdentro}(x) = \nu y[\forall yz[\text{C}(y,z) \leftrightarrow \exists w[\text{CONT-DENTRO}(w,x) \wedge \text{C}(z,w)]]] \quad (22)$$

Todas as relações vistas até o momento são aplicáveis tanto para regiões únicas quanto para regiões multi-peças (uma região multi-peças é formada por várias partes não contíguas). Porém, há casos onde é necessário descrever situações onde uma região está dentro do fecho

convexo de uma região multi-peças mas não dentro de uma peça específica. Assim, a relação ESP–DENTRO representa tal situação e é lida como “ $x$  está dentro de um região multi-peças  $y$ ”. Ademais, a relação T–DENTRO é redefinida, como segue (o predicado auxiliar  $\text{MAXP}(x,y)$  retorna verdadeiro quando  $x$  é uma sub-região maximal conectada de  $y$ ):

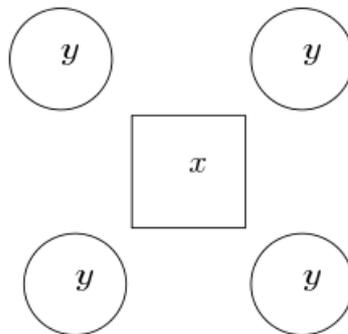
$$\text{MAXP}(x,y) \equiv_{def} \text{CON}(x) \wedge \text{P}(x,y) \wedge \neg \exists z [\text{PP}(z,y) \wedge \text{P}(z,y) \wedge \text{CON}(z)] \quad (23)$$

$$\begin{aligned} \text{ESP–DENTRO}(x,y) \equiv_{def} & \text{DENTRO}(x,y) \wedge \\ & \forall w [\text{MAXP}(w,x) \rightarrow \neg \exists z [\text{MAXP}(z,y) \wedge \text{DENTRO}(w,z)]] \end{aligned} \quad (24)$$

$$\begin{aligned} \text{T–DENTRO}(x,y) \equiv_{def} & \text{GEO–DENTRO}(x,y) \wedge \\ & \neg \text{PO}(x, \text{contdentro}(y)) \wedge \neg \text{ESP–DENTRO}(x,y) \end{aligned} \quad (25)$$

Portanto, a relação de interioridade geométrica pode ser dividida em quatro relações disjuntas par a par: CONT–DENTRO, T–DENTRO, G–DENTRO e ESP–DENTRO (veja um exemplo de ESP–DENTRO na Figura 5).

Figura 5 – Relação ESP – DENTRO.

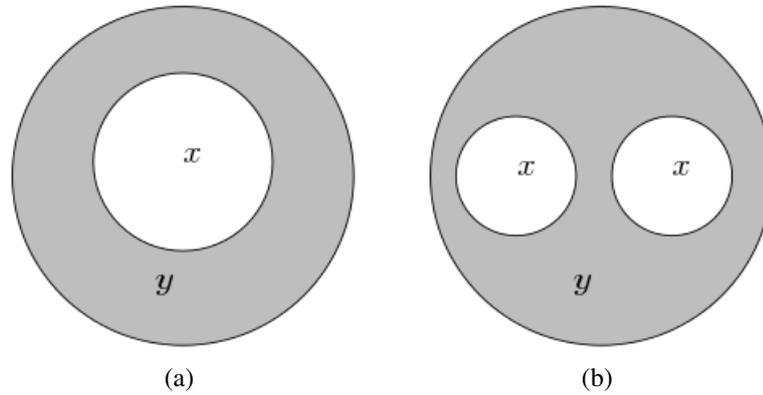


Um refinamento da relação TOP–DENTRO é dado quando uma região  $x$  está não tangencialmente cercada pela região  $y$ . O predicado não tangencialmente cercado  $\text{NTC}(x,y)$  define esse caso (a Figura 6 traz dois exemplos da relação NTC):

$$\text{NTC}(x,y) \equiv_{def} \exists z [\text{NTPP}(x,z) \wedge y = \text{prod}(z, \text{compl}(x))]. \quad (26)$$

Um refinamento para a relação de conteúdo (CONT–DENTRO) ocorre quando um recipiente está cheio. Uma dificuldade encontrada para expressar tal situação acontece quando há uma região multi-peças (por exemplo, quando há diversas concavidades agindo como recipientes). Uma solução é definir a ideia de contentor maximal em relação a uma região e definir o predicado PREENCH (toda peça que é sub-parce de  $x$  preenche completamente uma peça sub-parce de  $y$ ):

Figura 6 – Relação NTC.



$$\text{MAXCONT-DENTRO}(x,y) \equiv_{def} \text{MAXP}(x,\text{contdentro}(y)) \quad (27)$$

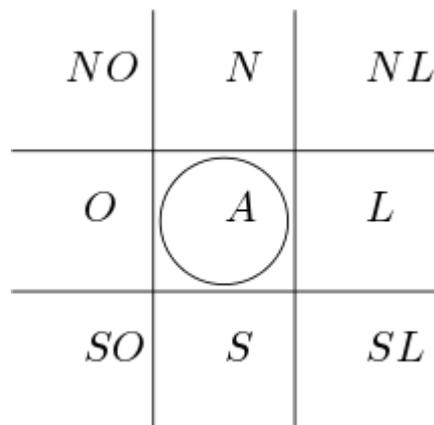
$$\text{PREENCH}(x,y) \equiv_{def} \forall z[\text{MAXP}(z,x) \rightarrow \text{MAXCONT-DENTRO}(z,y)]. \quad (28)$$

Foi visto nessa seção um conjunto de relações que tratam da mereotopologia das regiões. Essas relações servirão para estabelecer as precisificações do Semântica Standpoint para Polissemia (SSP), no Capítulo 6. Na próxima seção será mostrado um formalismo de REQ para as orientações lateral e transversal das regiões.

### 3.2 CÁLCULO DE DIREÇÕES CARDINAIS

Nesta seção será apresentado um formalismo para os eixos transversal (frente/trás) e lateral (direita/esquerda) baseado no Cálculo de Direções Cardinais (CDC), proposto por Frank (1996) e Skiadopoulos e Koubarakis (2004). O CDC é baseado na primitiva *caixa delimitadora mínima* CDM de uma região no espaço Euclidiano e consiste dividir o espaço em oito regiões adjacentes à CDM, nomeando-as com os pontos cardinais, conforme definido na Figura 7 (É importante observar que a CDM é representada pela letra *A*).

Figura 7 – Cálculo de Direções Cardinais.



Pode-se definir a função  $cdm(x)$ , que retorna a CDM da região  $x$ , e oito funções que retornam as regiões adjacentes à CDM:  $norte(x)$ ,  $sul(x)$ ,  $oeste(x)$ ,  $leste(x)$ ,  $nordeste(x)$ ,  $sudoeste(x)$  e  $sudeste(x)$ . Com base nessas funções, pode-se definir os predicados relacionados às direções transversais (frente/trás) e laterais (esquerda/direita). Esses predicados estão definidos na Tabela 3. Também é possível definir uma forma menos rígida para as relações de transversalidade e lateralidade, conforme mostrado na Tabela 4. Essas relações menos rígidas são utilizadas quando uma região  $y$  ocupa mais do que uma região adjacente à  $y$ .

Tabela 3 – Relações baseadas no CDC

Relação	Significado
$N(x,y) \equiv_{def} PP(x,norte(y))$	$x$ está atrás de $y$
$S(x,y) \equiv_{def} PP(x,sul(y))$	$x$ está à frente de $y$
$LE(x,y) \equiv_{def} PP(x,leste(y))$	$x$ está à direita de $y$
$OE(x,y) \equiv_{def} PP(x,oeste(y))$	$x$ está à esquerda de $y$
$NO(x,y) \equiv_{def} PP(x,noroeste(y))$	$x$ está atrás ou à esquerda de $y$
$NL(x,y) \equiv_{def} PP(x,nordeste(y))$	$x$ está atrás ou à direita de $y$
$SO(x,y) \equiv_{def} PP(x,sudoeste(y))$	$x$ está à frente ou à esquerda de $y$
$SL(x,y) \equiv_{def} PP(x,sudeste(y))$	$x$ está à frente ou à direita de $y$

Fonte: Autor

Tabela 4 – Relações laterais e transversais menos rígidas

Relação
$F-NORTE(x,y) \equiv_{def} P(x,norte(y)) \wedge [P(x,noroeste(y)) \vee P(x,nordeste(y))]$
$F-SUL(x,y) \equiv_{def} P(x,sul(y)) \wedge [P(x,sudoeste(y)) \vee P(x,sudeste(y))]$
$F-LESTE(x,y) \equiv_{def} P(x,leste(y)) \wedge [P(x,sudeste(y)) \vee P(x,nordeste(y))]$
$F-OESTE(x,y) \equiv_{def} P(x,oeste(y)) \wedge [P(x,noroeste(y)) \vee P(x,sudoeste(y))]$

Fonte: Autor

As relações abordadas nas Tabelas 3 e 4 serão utilizadas para as precisificações de expressões espaciais dos eixos lateral e transversal, tais como “atrás de”, “na frente de”, “à esquerda de”, “à direita de”, entre outras.

### 3.2.1 Relações de Verticalidade

Nesta seção será definida as relações para o eixo de verticalidade. Embora o foco desta tese seja em regiões num plano cartesiano, serão apresentados predicados que tratam do eixo vertical. O motivo é que algumas precisificações para as preposições são baseadas no eixo vertical. Para tanto, foram definidas as funções  $sup(x)$  e  $inf(x)$ , que retornam, respectivamente, a região que está acima e abaixo de  $x$  num plano 3D. Também foram definidos os predicados  $CIMA(x,y)$  e  $BAIXO(x,y)$ , que são lidos como “ $x$  está acima de  $y$ ” e “ $x$  está abaixo de  $y$ ” respectivamente. A Tabela 5 mostra a definição formal destes predicados.

Tabela 5 – Relações do eixo vertical

Relação	Significado
$CIMA(x,y) \equiv_{def} P(x, \sup(y))$	$x$ está acima de $y$
$BAIXO(x,y) \equiv_{def} P(x, \inf(y))$	$x$ está abaixo de $y$

Fonte: Autor

Neste capítulo foram vistas as relações de Raciocínio Espacial Qualitativo que serão utilizadas nas precisificações de expressões espaciais. No próximo capítulo será apresentado o conceito de esquema imagético, que é a base para se definir os mundos possíveis de uma expressão espacial, ou seja, os esquemas imagéticos são necessários para se entender os padrões de uso de uma expressão espacial, pois as relações espaciais de um determinado idioma são decompostas nestes esquemas. Assim, os formalismos de REQ, as características semânticas e os esquemas imagéticos são a base da Semântica Standpoint para Polissemia, abordada no Capítulo 5.

#### 4 ESQUEMAS IMAGÉTICOS

Tendo por base a noção de Linguagem Cognitiva, Johnson (1987) propôs a Tese da Corporeidade que, junto à Tese Simbólica de Langacker (1987), diz que uma linguagem é formada por representações conceituais, formas lexicais e percepções, e que com isso uma linguagem é dita corporificada. Pelo princípio objetivista, as representações são imagens reais de entidades. Assim, G. Lakoff e Johnson (1999) dizem que o raciocínio utiliza a experiência sensório-motora e é inseparável das características de corpo e cérebro. Essa abordagem remete a um refinamento da noção *Kantiana* de esquema.

Um esquema representa uma categoria ou um conceito e é oriundo da capacidade de abstração ou esquematização que permite formar padrões referentes ao uso da língua. Os esquemas são uma estrutura abstrata, uma vez que não possuem características de objetos específicos. Esses esquemas são organizados em *Gestalts* e sua capacidade de estruturação é vital para a linguagem.

*Gestalt* é uma teoria que estuda como os seres humanos percebem as coisas. A psicologia da *Gestalt* enfoca as leis mentais - os princípios que determinam a maneira como o ser-humano percebe as coisas e possui como princípios básicos (DESOLNEUX; MOISAN; MOREL, 2003):

- 1) Similaridade;
- 2) Proximidade;
- 3) Continuidade;
- 4) Alinhamento;
- 5) Experiências anteriores.

Estruturas mentais originadas na experiência sensório-motora constituem um caso específico de esquema, denominado esquema imagético. Johnson (1987) diz que os esquemas imagéticos são considerados como primitivas conceituais, onde as relações espaciais em um determinado idioma são decompostas nestes esquemas tornando-se importantes para o estudo da polissemia, entre outros casos, pois as expressões espaciais estão ligadas ao conceito de esquemas imagéticos. Por exemplo, a preposição “em” pode estar ligada ao esquema de contenção e também o de suporte. Há um outro conceito, denominado sentido incorporado, que está ligado diretamente aos esquemas imagéticos. Assim, a estrutura semântica é originada da estrutura conceitual humana e esta na incorporação do sentido. A estrutura conceitual humana é um conceito mais amplo, incluindo representações conceituais abstratas e esquemas imagéticos.

Para Johnson (1987), um esquema possui um nível básico de especificidade e não se refere a nenhuma instância de uso ou à capacidade linguística. Para Oliveira (2009), um esquema é um padrão, forma ou regularidade decorrente das atividades contínuas de ordenação, a saber: ações, percepções e concepções. Baseado nas ideias apresentadas por Croft e Cruse (2004), Evans (2006), Clausner e Croft (1999), Johnson (1987), Hampe e Grady (2005) e G. Lakoff e

Johnson (1999), os esquemas imagéticos são agrupados de acordo com a base de experiência de cada esquema:

- 1) Espaço: Em cima, Em baixo, Frente-Verso, Esquerda-Direita, Próximo-Longe, Centro-Periferia, Contato, Verticalidade, Reta;
- 2) Escala: Percurso;
- 3) contendor: Contenção, Dentro-Fora, Superfície, Cheio-Vazio, Conteúdo;
- 4) Força: Contra Força, Compulsão, Restrição, Desbloqueio, Bloqueio, Desvio, Atração, Resistência, Equilíbrio;
- 5) Unidade/Multiplicidade: Fusão, Coleção, Separação, Reiteração, Parte-Todo, Incontável-Contável, Ligação;
- 6) Identidade: Combinação, Sobreposição;
- 7) Existência: Remoção, Espaço Delimitado, Ciclo, Objeto, Processo;
- 8) Locomoção: Impulso, Origem-Percurso-Objetivo;
- 9) Equilíbrio: Eixo, Duplo Equilíbrio, Ponto de Equilíbrio, Equilíbrio.

Os esquemas imagéticos são representações análogas de experiências sensorio-motoras e são representados pelo uso de dois construtos denominados *trajetor* e *marco*, sendo o primeiro o elemento de maior proeminência e o segundo a referência. Em geral, o trajetor é mais móvel, menor e de difícil localização, ao passo que o marco é mais fixo, de dimensão maior e de localização mais fácil (ARAÚJO, 2008).

Para Johnson (1987), um esquema imagético é uma estrutura recorrente dentro dos processos cognitivos que estabelecem padrões de compreensão e raciocínio. Esses esquemas imagéticos são formados a partir de interações corporais, da experiência linguística e do contexto histórico. Hampe e Grady (2005) apresentam uma caracterização resumida da definição original de Johnson para esquemas imagéticos:

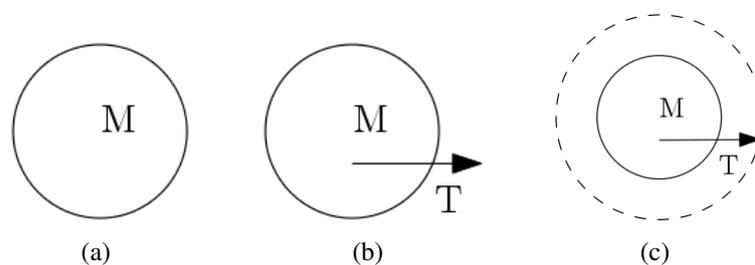
- 1) São providos de sentidos, devido ao sentido incorporado, além de serem estruturas pré-conceituais e não proposicionais, decorrentes da interação com o mundo externo e a manipulação de objetos nesse mundo;
- 2) São *Gestalts*, e como tal, são internamente estruturados (compostos de pequenas e poucas partes relacionadas) e flexíveis (esquemas imagéticos podem passar por transformações).

Evans (2006) elenca as seguintes características dos esquemas imagéticos:

- 1) São pré-conceituais na origem: Evans (2006) argumenta que eles surgem de experiências sensoriais nos primeiros estágios do desenvolvimento humano que precedem a formação de conceitos, e certamente surgem antes do aprendizado da linguagem;
- 2) Podem dar origem a conceitos mais específicos: na Figura 8 tem-se três exemplos para o esquema imagético de “Contenção”. Na Figura 8(a) o marco é o círculo representado pela letra “M” e possui dois elementos: interior (área dentro da borda

do círculo) e a borda. O exterior é a área fora do marco. As Figuras 8(b) e 8(c) representam variações do esquema de Contenção e são lexicalizados pela palavra “fora”. Na Figura 8(b), o trajetor (representado pela letra “T”) move-se de uma posição dentro do marco para ocupar um lugar fora do marco. Uma outra interpretação para a palavra “fora” pode ser vista na Figura 8(c). Suponha-se que na Figura 8(c) tem-se a presença de um líquido. Devido às características físicas do líquido, este pode ser, simultaneamente, marco e trajetor. O marco é dado pela área ocupada inicialmente pelo líquido (o círculo) e o trajetor pela área fora da borda (linha tracejada) onde o líquido transborda. Portanto, tem-se aqui um esquema genérico (Contenção) dando origem a esquemas mais específicos;

Figura 8 – Esquema imagético de contenção.



- 3) Derivam de experiências sensório-motoras e da observação do mundo: para ilustrar essa propriedade, considere o esquema imagético de Força. Esse esquema surge da experiência de atuação sobre uma entidade ou entidades atuando sobre outras entidades, resultando na transferência de energia de movimento. Evans (2006) diz que o esquema de Força só ocorre quando há interação ou potencial interação entre entidades, e portanto é derivado da atuação de um ente sobre outro, como por exemplo ao entrar em um quarto escuro e chocar-se contra uma cadeira, ou ainda a pressão que o alimento exerce no estômago ao comer acima do limite;
- 4) São inerentemente providos de sentido: devido ao motivo de que os esquemas imagéticos derivam da interação com o mundo, eles são inerentemente significativos. Ilustra-se essa propriedade com o seguinte exemplo: imagine uma xícara de café nas mãos. Se mover a xícara lentamente para cima e para baixo, ou de lado a lado, espera-se que o café mova-se com a xícara. Isso é uma consequência da contenção, uma vez que é definida pelas bordas e restringe a localização de qualquer entidade dentro destes limites. Em outras palavras, a xícara exerce um controle de força dinâmica sobre o café. Esse tipo de conhecimento é adquirido como consequência da interação com o ambiente físico. As propriedades de força dinâmica, vistas acima para o esquema de Contenção, também aparecem no significado linguístico, como ilustra o significado da preposição “em” nos exemplos em Vandeloise (1994):

**Exemplo 1.** A lâmpada está no soquete.

**Exemplo 2.** A garrafa está na tampa.

Vandeloise (1994) ressalta que as relações espaciais entre o trajetor e marco nessas duas frases são idênticas e, no entanto, enquanto o Exemplo 1 é uma sentença perfeitamente aceitável, o Exemplo 2 é semanticamente inaceitável. Vandeloise (1994) sugere ainda que não é a relação espacial entre marco e trajetor que representa a aceitabilidade ou não da frase, mas sim a dinâmica: enquanto o soquete exerce uma força na lâmpada e determina sua posição, o contrário não ocorre com a tampa e a garrafa. Em outras palavras, além da lâmpada estar contida pelo soquete, é este que a impede de sucumbir à força da gravidade. Já a posição da garrafa não depende da tampa. Assim, o conhecimento das consequências associadas a um esquema imagético (Contenção, por exemplo) afeta a aceitabilidade contextual de uma preposição;

- 5) São representações análogas: esquemas imagéticos são representações decorrentes da experiência. Nesse contexto, o termo “análogas” significa que esquemas imagéticos assumem uma forma no sistema conceitual que espelha a experiência sensorial que estão sendo representadas. Em outras palavras, embora se possa tentar descrever esquemas imagéticos usando palavras e imagens, eles não são representados na mente nesses tipos de formas simbólicas. Em vez disso, conceitos esquemáticos estão representados na mente em termos de experiências sensoriais holísticas;
- 6) São internamente complexos: esquemas imagéticos são muitas vezes compostos por aspectos mais complexos que podem ser analisados separadamente. Por exemplo, o esquema Contenção é um conceito que consiste de três elementos: interior, limite e exterior. Outro exemplo de um esquema imagético complexo é o esquema Origem-Caminho-Objetivo (ou simplesmente esquema Caminho). Um caminho é um meio de se mover de um local para outro e baseia-se em um ponto de partida ou Origem, um destino ou Objetivo e uma série de locais contíguos entre os quais se relacionam a origem e o objetivo. Como todos os esquemas imagéticos complexos, o esquema Caminho constitui uma *Gestalt* experiencial: tem estrutura interna e é coerente como um todo;
- 7) São diferentes de imagens mentais: as imagens mentais são detalhadas e resultam de um processo cognitivo parcialmente consciente que envolve a recordação na memória visual. Esquemas imagéticos são esquemáticos e, portanto, mais abstratos na natureza, emergindo da experiência corporificada em curso. Isso significa que não se pode fechar os olhos e “pensar” sobre um esquema imagético da mesma maneira que se pode “pensar” na visão do rosto de alguém ou a sensação de um determinado objeto contido nas mãos;
- 8) São multi-modais: uma das razões pelas quais não se pode fechar os olhos e “pensar” um esquema imagético é porque esquemas imagéticos derivam de experi-

ências em diferentes modalidades (diferentes tipos de experiências sensoriais) e, portanto, não são específicos de um sentido particular. Em outras palavras, os esquemas imagéticos são localizados mais profundamente dentro do sistema cognitivo, sendo padrões abstratos decorrentes de uma vasta gama de experiências perceptuais. Por exemplo, pessoas cegas têm acesso a esquemas imagéticos para contentores, caminhos e assim por diante, de forma precisa, pois os tipos de experiências que dão origem a esses esquemas imagéticos dependem de uma série de experiências sensório-perceptivas além da visão, incluindo audição, tato e da experiência de movimento e equilíbrio, entre outros;

- 9) Estão sujeitos a transformações: G. Lakoff (1989) apresentou um exemplo para ilustrar essa propriedade: suponha um rebanho de vacas suficientemente perto, de modo que consegue-se apontar e escolher cada um dos elementos do rebanho. Agora, suponha que houve um afastamento gradual desse rebanho até chegar ao ponto que não é mais possível escolher uma vaca, passando a enxergar o rebanho como uma massa única. Nesse caso, houve uma transformação do esquema Contagem no esquema Massa. Os esquemas Contagem e Massa são refletidos nos usos dos substantivos e são determinados por artigos ou numerais. Por fim, é possível transformar o esquema de Contagem em Massa e vice-versa através da conversão de substantivos contáveis em não contáveis e vice-versa (no exemplo temos vaca como contável e rebanho como não contável).
- 10) Podem ocorrer em redes ou em grupos de esquemas imagéticos: tome como exemplo o esquema de Força, que consiste em diversos outros esquemas relacionados: Compulsão, Bloqueio, Contra-Força, Espalhamento, Remoção, Restrição, Ativação e Atração. Todos esses esquemas envolvem a aplicação de força.

Como visto nas propriedades dos esquemas imagéticos acima, a aquisição, por humanos, do conceito de esquema imagético deve ocorrer primeiro do que o aprendizado linguístico e pode influenciar na escolha das palavras e na semântica. Assim, entender como uma palavra se relaciona com um esquema imagético é importante para definir os mundos possíveis de uma determinada expressão espacial. Por estar diretamente relacionado à linguagem, o conceito de esquema imagético será abordado na próxima seção com ênfase em linguagem.

#### 4.1 ESQUEMAS IMAGÉTICOS EM LINGUAGEM

Segundo Dodge e Lakoff (2005), os esquemas imagéticos estruturam a experiência independentemente da linguagem. Dodge e Lakoff (2005) citam como exemplo o experimento de vários objetos como contentores: caixas, copos, cestos, boca, salas.

Os esquemas imagéticos definem classes de experimentos que são caracterizados por uma mesma palavra. De acordo com Dodge e Lakoff (2005), esse fato levanta dois problemas: como é possível diferentes experiências terem a mesma estrutura de esquema imagético,

e também como os esquemas imagéticos são expressos pela linguagem. É razoável dizer que a forma como os esquemas imagéticos são expressos em linguagem é uma característica central da estrutura linguística. Eles podem ser expressos por preposições, verbos, casos, morfemas (LANGACKER; MUNRO, 1975). Ademais, a forma com que os esquemas imagéticos são expressos em uma linguagem é uma característica topológica da linguagem (DODGE; LAKOFF, 2005). A seguir, são apresentados alguns exemplos para discussão.

Dodge e Lakoff (2005) questionam quais tipos de esquemas imagéticos são expressados quando se referir ao movimento de uma pessoa. Considere as frases dos Exemplos 3 e 4:

**Exemplo 3.** Ele caminhou **até a** cozinha.

**Exemplo 4.** Ele caminhou **dentro da** cozinha.

Em ambas as sentenças, a pessoa que se movimenta (“ele”) altera sua localização em relação ao marco (cozinha). Porém, a mudança é somente esquematicamente especificada. Do Exemplo 3 pode-se inferir que a pessoa não estava inicialmente na cozinha, e depois de movimentar-se estava dentro da cozinha. A partir dessas descrições pode-se observar que o esquema Origem-Caminho-Destino está expresso na sentença do Exemplo 3. A frase do Exemplo 4 utiliza a locução prepositiva “dentro de”, indicando que tanto origem quanto objetivo estão dentro da cozinha. Nesse caso, elementos da estrutura esquemática do marco são utilizados para indicar a localização da pessoa, isto é, a cozinha é entendida como um contentor. Literalmente as paredes dividem a cozinha em duas partes: interior (espaço fechado pelas paredes) e exterior (o espaço em torno da cozinha). Há também nas paredes uma abertura que permite um agente se mover do exterior para o interior. Portanto, os seguintes elementos esquemáticos são relevantes:

- 1) Limite;
- 2) Interior;
- 3) Exterior;
- 4) Abertura.

De acordo com Dodge e Lakoff (2005), esses elementos em conjunto se referem, geralmente, ao esquema de contentor. Esses mesmos elementos estruturais são inconscientes e automaticamente utilizados para definir uma variedade de objetos de vários tamanhos: copos, caixas, sacos, quartos, tendas, edifícios, vales, etc. Não surpreende então que, além de descrição de pessoas que se deslocam em torno do mundo, esse esquema também pode ser usado em descrições de outros tipos de experiências, como colocar comida na boca ou maçãs em uma cesta (DODGE; LAKOFF, 2005).

**Exemplo 5.** Harry caminhou **para fora** do quarto.

**Exemplo 6.** Harry caminhou **para dentro** do quarto.

No Exemplo 5, Harry está inicialmente dentro do quarto e, portanto, move-se para o exterior. O oposto ocorre no Exemplo 6. Ainda de acordo com Dodge e Lakoff (2005), es-

As mudanças de localização podem ser analisadas em termos do esquema Origem-Caminho-Destino: “fora” significa que o exterior é o objetivo, enquanto “dentro” mostra que o objetivo é o interior. Cada uma dessas sentenças pode então ser expressa pelos esquemas de contentor e Origem-Caminho-Destino.

A forma com que os exemplos dados são estruturados pode ocorrer de diversas maneiras, mas dada uma determinada sentença, somente uma dessas maneiras pode ser utilizada. Por exemplo, um quarto é referenciado como contentor, mas também pode ser dito como tendo diferentes lados, como no Exemplo 7:

**Exemplo 7.** Harry caminhou **pelo** quarto.

O uso da preposição “através” no Exemplo 7 indica que Harry está se movendo de um lado do marco (quarto) para o lado oposto. Esses lados podem ser determinados baseados na geometria do marco, com ênfase em sua estrutura de eixos. Nesse exemplo, o tamanho das paredes do quarto pode servir para definir um eixo principal que efetivamente divide o quarto em dois lados. Como Harry se movimenta de um lado ao outro do quarto, ele cruza o eixo principal.

Como visto nos exemplos, os esquemas imagéticos estão ligados às interpretações de uma expressão espacial. Por isso, nesta tese, optou-se por categorizar os usos de uma preposição (denominados mundos possíveis) através dos esquemas imagéticos que estão associados a ela.

A estruturação dos esquemas imagéticos de experimentos significa que um esquema imagético primitivo está disponível para ser, de alguma maneira, expressado em uma dada língua. Porém, como mostram os estudos de Talmy (1983) e Langacker (1987) para linguística cruzada, há uma diversidade significativa de formas como as linguagens descrevem espaço e localização. Observe alguns exemplos adaptados de Dodge e Lakoff (2005):

- 1) Importância da forma e/ou orientação do marco: para diversos termos que designam relações espaciais, a forma e a orientação do marco não são fundamentais. Por exemplo, o termo “em cima” pode ser aplicado de diferentes formas, como em “o pássaro está em cima da árvore/casinha/mesa”;
- 2) Presença ou ausência de contato: alguns termos de relações espaciais são distintos quanto à presença ou ausência de contato. Por exemplo, os termos “sobre” e “acima” diferem nas relações verticais do objeto marco.

No entanto, apesar dessa diversidade de linguística cruzada, o número de esquemas imagéticos primitivos usados por termos espaciais parece ser bastante limitado. Talmy (1983) realizou uma extensa análise de linguística cruzada das formas gramaticais usadas na descrição linguística do espaço. Com base em sua análise, ele supõe que há um número limitado de distinções espaciais básicas que as linguagens fazem em seus sistemas de classe fechada. Esse grupo de primitivas básicas inclui:

- 1) Distinções de foco em uma cena: figura (objeto focalizado) e fundo (foco secundário, servindo como um objeto de referência para a localização da figura);

- 2) Geometria da figura e do fundo, orientações relativas;
- 3) Presença/ausência de contato da figura e do fundo;
- 4) Força-dinâmica: reflete as modalidades não visuais e é, em grande parte, independente de outras distinções espaciais.

Talmy (2000) propõe que os termos das relações espaciais utilizados na linguagem são conceitos complexos compostos de primitivas selecionadas nesse conjunto. Assim, as relações espaciais podem ser codificadas como um complexo de esquemas ao invés de apenas um esquema primitivo.

Baseando-se em um amplo estudo de linguagens, Levinson e Meira (2003) e Talmy (2000) concluíram que estas linguagens possuem três quadros de referências:

- 1) Quadro de referência intrínseco: coordenadas espaciais são determinadas utilizando características próprias do marco. Por exemplo, em “ele correu para dentro do quarto”, a relação espacial “para dentro” faz uso da estrutura do esquema imagético característico do quarto para especificar a localização do agente que está se movimentando;
- 2) Quadro de referência relativo: a determinação das coordenadas espaciais é feita em relação a um ponto de vista particular. Nesse quadro, a estrutura esquemática não é inerente ao marco. Por exemplo em “ele correu em frente à árvore”, o termo “em frente” faz uso de elementos frente/trás de um esquema. Porém, ao invés de ter componentes da estrutura esquemática própria da árvore, esses elementos estão associados com um agente que visualiza a cena. Portanto, enquanto um objeto marco está presente, o visualizador serve como um esquema âncora para esse quadro de referência;
- 3) Quadro de referência absoluto: tome como exemplo a frase: “ele correu para o norte”. Essa frase não possui referência para um marco espacial. Pelo contrário, a localização está especificada em relação a elementos fixos.

Para Dodge e Lakoff (2005), as linguagens diferem em como as distinções espaciais básicas são feitas, as combinações das distinções e os termos para relações espaciais e a classe gramatical para esses termos. Dada a diversidade linguística, se cada termo das relações espaciais representasse um esquema imagético, então poderia se dizer que as linguagens diferem fortemente em seu rol de esquemas imagéticos espaciais. Entretanto, Dodge e Lakoff (2005) dizem que as expressões de relações espaciais podem ser analisadas como uma combinação complexa de esquemas imagéticos primitivos. Consequentemente, se a expressão “esquema imagético” for restrita a apenas esquemas primitivos, a análise linguística cruzada indica que há um número limitado de esquemas imagéticos usados pelas linguagens de todo o mundo (TALMY, 2000).

O conceito de *Gestalt* afirma que na cognição espacial humana a percepção espacial é feita através da assimetria figura/fundo. Herskovits (1997) sugeriu que a construção de uma relação espacial entre entidades geométricas é um processo seletivo e eliminativo. Seletivo pois

somente uma parte da realidade é percebida pelos olhos e eliminativo pois todos os objetos carregam mais propriedades do que aquelas usadas para descrevê-los. Herskovits (1997) descreve essa construção em um processo de três passos:

- 1) Abstração: informações do cenário são deixadas de lado para se focar naquilo que está sendo descrito. Por exemplo, quando se diz “o gato está embaixo da mesa”, o fato de o gato estar sentado, parado, correndo, dormindo, etc. é irrelevante;
- 2) Idealização: a simplificação geométrica dos objetos restantes. A idealização das características de objetos geometricamente complexos se dá através de linhas, pontos, planos;
- 3) Seleção: partes ou aspectos dos objetos que são ignorados. No exemplo “o gato está embaixo da mesa” não exclui o fato de a mesa ter pés, porém isso não é mencionado pois a mesa foi selecionada como sendo uma superfície plana.

Talmy (2000) então conclui que esses passos são um processo de esquematização da realidade. As ideias apresentadas neste capítulo serão úteis para o desenvolvimento do SSP. Para tanto, neste trabalho serão listados os esquemas imagéticos e seus respectivos elementos esquemáticos e em seguida os conceitos de linguística cruzada, e o quadro de referências será utilizado para, assim como Dodge e Lakoff (2005), modelar as inferências do formalismo utilizando classes de características visuais, tais como orientação e topologia.

Ademais, através dos esquemas imagéticos será possível classificar os diversos sentidos de uma expressão espacial em mundos possíveis, pois uma mesma palavra pode remeter a diversos esquemas imagéticos. Considere, por exemplo, a preposição “em”. Quando utilizada na frase “as maçãs estão no cesto”, evoca-se o esquema de Contenção, enquanto que na frase “as maçãs estão na mesa” o esquema imagético observado é o de Suporte e Verticalidade. Na primeira frase, entende-se que as maçãs estão dentro do cesto enquanto que na segunda as maçãs estão em cima da mesa. Portanto, a caracterização dos esquemas imagéticos é importante para poder identificar quais são os sentidos que uma palavra pode ter. Assim, para tratar a polissemia de uma expressão espacial, optou-se inicialmente por identificar quais os esquemas imagéticos que estão relacionados a ela. Para tanto, é necessário extrair quais as características semânticas do trajetor e do marco. Por exemplo, na frase “as maçãs estão no cesto” é possível inferir que se trata de um esquema de contenção pois o termo “cesto” remete à ideia de um contentor, enquanto que em “as maçãs estão na mesa”, o termo “mesa” está relacionado a uma superfície plana.

Não obstante, pode-se especificar ainda mais o sentido de uma sentença espacial. Imagine no exemplo em que “as maçãs estão na mesa”: elas poderiam estar em contato diretamente com a mesa, ou poderia haver uma toalha entre a mesa e as maçãs, ou ainda as maçãs poderiam estar em um cesto e este em cima da mesa. Todas essas situações são expressas por uma mesma sentença. Com isso, além de identificar os esquemas imagéticos e as características semânticas, é necessário precisar quais as possíveis localizações do trajetor em relação ao marco. Portanto, as precisificações são aplicações dos formalismos de Raciocínio Espacial Qualitativo

a uma determinada sentença. Assim, a interpretação de uma expressão espacial dependerá do esquema imagético, das características semânticas do trajecto e marco e das precisificações. No próximo capítulo, será apresentado o formalismo SSP que utiliza esses conceitos para o tratamento da polissemia em expressões espaciais.

## 5 SEMÂNTICA STANDPOINT PARA POLISSEMIA

Neste capítulo será apresentado um formalismo denominado Semântica Standpoint para Polissemia em expressões espaciais baseado no trabalho de Bennett (2011). Inicialmente Bennett (2011) propôs um arcabouço para o tratamento da vagueza utilizando uma semântica de superavaliação. Assim, uma das contribuições do presente trabalho é adaptar o arcabouço de Bennett (2011) para a polissemia. O conceito fundamental do formalismo desenvolvido nesse trabalho consiste numa tripla ⟨esquema imagético, característica semântica, eixo espacial⟩, denominada *característica espacial* de uma expressão. A ideia do formalismo é, para cada expressão, estabelecer o conjunto de características espaciais, denominado *mundos possíveis*. Os esquemas imagéticos são obtidos conforme discutido no Capítulo 4. Já as características semânticas são adquiridas de dicionários de uso, dicionários de sinônimos e *WordNet*. Os eixos espaciais são divididos em vertical (cima/baixo), lateral (esquerda/direita), frontal (frente/trás) e conteúdo (dentro/fora) (MARQUES, 2009).

O conceito de mundos possíveis citado acima baseia-se na classificação de expressões espaciais feita por Garcia (2018), que por sua vez utilizou a proposta apresentada por Ilari et al. (2008) para dividir o significado das preposições em esquemas imagéticos, características semânticas e eixos espaciais.

### 5.1 CARACTERÍSTICAS SEMÂNTICAS

Segundo Fromkin, Rodman e Hyams (2018), as características semânticas representam os componentes conceituais básicos do significado para qualquer item léxico. Uma característica semântica individual constitui um componente de uma *intensão* de uma palavra, que é o sentido inerente ou conceito evocado por uma palavra (O'GRADY; DOBROVOLSKY; KATAMBA, 2010). Propõe-se o significado linguístico de uma palavra a partir de contrastes e diferenças significativas com outras palavras. As características semânticas permitem explicar como as palavras que compartilham determinadas características podem fazer parte do mesmo domínio semântico. De forma correspondente, o contraste nos significados das palavras é explicado por características semânticas divergentes. Nida (1979) dá o seguinte exemplo: as palavras “pai” e “filho” compartilham os componentes comuns de “humano”, “parentesco”, “masculino” e, portanto fazem parte de um domínio semântico das relações familiares masculinas. Eles diferem em termos de “geração”, que é o que dá a cada um seus significados individuais. As características semânticas serão utilizadas no SSP para descrever semanticamente cada expressão espacial. Além disso, será estabelecido um conjunto de características semânticas para os termos que acompanham a expressão espacial.

As características semânticas são denotadas entre colchetes e um sinal de mais ou menos indicando a existência ou a inexistência de tal propriedade (LIPKA, 1990). O exemplo a seguir mostra as características semânticas para os termos “homem”, “mulher”, “menino” e “menina”:

- 1) Homem: [+humano], [+masculino], [+adulto];
- 2) Mulher: [+humano], [-masculino], [+adulto];
- 3) Menino: [+humano], [+masculino], [-adulto];
- 4) Menina: [+humano], [-masculino], [-adulto];

**Definição 1.** Uma *estrutura espacial* denota os usos espaciais de uma expressão em relação aos esquemas imagéticos, às características semânticas e aos eixos espaciais. Formalmente, uma estrutura espacial é uma tupla  $\langle D, \mathcal{S}, \Sigma, S, X, w \rangle$  onde

- 1)  $D$  é o domínio das entidades (os termos que representam objetos do cenário);
- 2)  $\mathcal{S}$  é o conjunto de esquemas imagéticos;
- 3)  $\Sigma$  é o conjunto de características semânticas relacionadas às expressões espaciais;
- 4)  $S$  é o conjunto de características semânticas dos termos que acompanham a expressão espacial;
- 5)  $X$  é o conjunto de eixos espaciais;
- 6)  $w : \mathcal{S} \times \Sigma \times X$  é uma tripla esquema imagético, característica semântica, eixo espacial denominada *mundo possível*.

Portanto, o conjunto de mundos possíveis é dado por  $\text{Mundos}(D, \mathcal{S}, \Sigma, X) = \{w(\mathcal{S}, \Sigma, X)\}$ .

**Definição 2.** Um *quadro espacial* é uma tupla que especifica todos os mundos possíveis determinados por uma estrutura espacial  $\langle D, \mathcal{S}, \Sigma, S, X, W \rangle$  onde  $W = \text{Mundos}(D, \mathcal{S}, \Sigma, X)$ .

A principal diferença em relação ao trabalho de Bennett (2011) é que ele define os mundos possíveis como valorações arbitrárias de funções de medição sobre um domínio  $D$  a um número racional  $\mathbb{Q}$ . Uma vez que o propósito de Bennett (2011) é tratar a vagueza, essas funções são utilizadas para mensurar observáveis físicos, tais como tamanho, massa, volume, velocidade, etc. Neste trabalho, modificou-se o conceito de mundo possível para que se possa lidar com os sentidos que uma expressão espacial pode ter. Ao invés de funções de medição, os mundos possíveis são caracterizados por triplas de  $\langle$ esquema imagético, característica semântica, eixo espacial $\rangle$ .

Uma vez definidos os mundos possíveis, pode-se estabelecer uma linguagem de primeira ordem para as expressões espaciais baseada nas características semânticas dos termos que acompanham essas expressões.

**Definição 3.** *Trajedor* (tr) é a entidade cuja localização é relevante e pode ser estática ou dinâmica, uma pessoa ou um objeto. Neste trabalho o trajedor será limitado a um objeto estático;

**Definição 4.** *Marco* (lm) é a entidade de referência em relação à localização ou trajetória do trajedor;

**Definição 5.** *Indicadores espaciais* (si) são os termos que definem as implicações sobre as propriedades espaciais, tais como as relações entre o trajedor e o marco. Normalmente representados por preposições ou locuções prepositivas.

Seja  $\mathcal{L}_1(S,V)$  o conjunto de fórmulas de uma linguagem de primeira ordem cujos símbolos não lógicos consistem de um conjunto de funções Booleanas que representam as características semânticas dos termos que acompanham a expressão espacial  $S = \{S_1, \dots, S_i\}$  e um conjunto de símbolos variáveis relacionados aos termos da sentença  $V = \{tr, lm\}$ .

As fórmulas atômicas para a linguagem  $\mathcal{L}_1$  são definidas como  $S_i(x)$  onde  $x \in V$  e  $S_i \in S$ . As fórmulas complexas são definidas por fórmulas atômicas, conectivos lógicos e quantificadores universais. Cada uma das fórmulas da linguagem  $\mathcal{L}_1$  representa uma característica semântica do trajetor ou do marco. Essas características semânticas são importantes para definir o uso de um termo em uma sentença. Por exemplo, na sentença “a maçã está no cesto”, o trajetor é representado pela palavra “maçã” e o marco por “cesto”. Assim, a fórmula da linguagem  $\mathcal{L}_1$  que representa essa sentença pode ser dada por  $Fruta(tr) \wedge Container(lm)$ . Isso expressa o comportamento que uma palavra possui em relação a uma sentença. Com isso, pode-se estender a linguagem  $\mathcal{L}_1$  com a definição de novos predicados pelos significados de fórmulas no formato:

$$\forall x_1, \dots, x_k [R(x_1, \dots, x_k) \Leftrightarrow \Phi(S_1, \dots, S_j, x_1, \dots, x_m)] \quad (29)$$

onde  $\Phi(S_1, \dots, S_j, x_1, \dots, x_m)$  é qualquer fórmula da linguagem  $\mathcal{L}_1(S,V)$ . Esses predicados representam os indicadores espaciais de uma sentença. Por exemplo, na frase “as maçãs estão no cesto”, o indicador espacial é dado pela preposição “em”. Portanto, os predicados em  $R$  estão associados com preposições espaciais. Por exemplo, pela Equação 29, o predicado EM pode ser definido como:

$$\forall tr, lm [EM(tr, lm) \Leftrightarrow Container(lm)].$$

A fórmula acima é um exemplo típico de um predicado definido dessa maneira. Aqui,  $Container(lm)$  é função que representa uma característica semântica. Uma interpretação informal desta fórmula é que uma entidade está dentro de uma outra caso essa segunda entidade for um contentor.

Seja  $\mathcal{L}_2(S,V,R)$  a linguagem obtida pela expansão da linguagem  $\mathcal{L}_1$  incluindo novos predicados  $R$ , então:

**Definição 6.** Uma Teoria de Definição de Predicado (TDP) define cada predicado em  $R$  em relação a uma fórmula da linguagem  $\mathcal{L}_2$ , fixando os casos de uso de uma expressão espacial em termos de características semânticas. Assim, uma TDP é um conjunto de fórmulas contendo exatamente uma fórmula para cada predicado  $R_i \in R$  no formato da Equação 29.

**Definição 7.** O conjunto  $\Theta = \{\theta_1, \dots, \theta_i\}$  é o conjunto de todas as Teorias de Definição de Predicado para a linguagem  $\mathcal{L}_2$ .

**Definição 8.** Um *modelo de precisificação espacial* é uma estrutura

$$\mathfrak{M} = \langle \mathcal{F}, T, R, V, \Theta, K, \xi, P \rangle \quad (30)$$

onde:

- 1)  $\mathcal{F} = \langle D, \mathcal{I}, \Sigma, S, X, W \rangle$  é um quadro espacial;
- 2)  $T$  é o conjunto de predicados de REQ;
- 3)  $R$  é o conjunto de predicados que definem os usos de uma expressão espacial;
- 4)  $V = \{tr, si, lm\}$  é o conjunto de termos de uma sentença;
- 5)  $\Theta = \{\dots, \theta_i, \dots\}$  onde cada  $\theta_i$  é uma Teoria de Definição de Predicado;
- 6)  $K(d) = \{\kappa \mid \kappa : D \rightarrow S\}$  é o conjunto de predicados de características semânticas de um termo  $d \in D$ ;
- 7)  $\xi : V \rightarrow D$  mapeia um termo de uma sentença a um objeto;
- 8)  $P = \{p \mid p : R \rightarrow T\}$  é o conjunto de mapeamentos de expressões espaciais para relações de REQ.  $P$  é chamado de o conjunto de *precisificações*.

O intuito de um modelo de precisificação é, dada uma sentença contendo uma expressão espacial, estabelecer as precisificações possíveis para tal expressão. Bennett (2011) define um “modelo de precisificação parametrizado” onde as precisificações são atribuições de valores racionais a limiares previamente estabelecidos. Esses limiares são casos limite de funções de medição. Por exemplo, para o predicado Alto, é definido uma função de medição  $Altura(x)$ , um limiar T–alto e uma precisificação 1,81m. Assim, pessoas com altura acima de 1,81m são consideradas altas. No modelo desenvolvido nesta tese, as precisificações são relações de REQ que determinam as possíveis localizações espaciais do trajector em relação ao marco. Por exemplo, na sentença “a pera está no cesto”, as precisificações possíveis para a preposição “em” são:

- 1)  $p_1 : \text{GEO-DENTRO}(pera, cesto)$ : a pera está geometricamente dentro do cesto;
- 2)  $p_2 : \text{P-DENTRO}(pera, cesto)$ : a pera está parcialmente contida no cesto;
- 3)  $p_3 : \text{FORA}(pera, cesto) \wedge \exists z [[\text{GEO-DENTRO}(z, cesto) \vee \text{P-DENTRO}(z, cesto)] \wedge \text{EC}(pera, z)]$ : a pera está fora do cesto porém em contato com um objeto  $z$  localizado dentro do cesto.

Na precisificação  $p_1$  a pera está completamente contida no cesto, enquanto que nas precisificações  $p_2$  e  $p_3$  a pera está em contato com outros elementos e parcialmente contida ( $p_2$ ) ou fora ( $p_3$ ) do cesto. Assim, para a mesma sentença tem-se três possibilidades de localização espacial. Portanto, pode-se observar que a polissemia ocorre não somente em nível da palavra (a preposição “em” tem o mesmo sentido nas três situações) mas também em localização espacial.

É importante ressaltar que a localização de um objeto deve ser precisa. Imagine um sistema robótico que opere com localização espacial a partir de uma frase. Ele deve ser capaz de analisar todas as possibilidades de configuração espacial. Portanto, o modelo de precisificação apresentado neste trabalho é capaz de prover tais informações. Assim, dado um conjunto de

Teoria de Definição de Predicado, um conjunto de mundos possíveis e um conjunto de precisificações, pode-se definir uma função de interpretação semântica, conforme descrita na Definição 9 abaixo:

**Definição 9.** A função de interpretação semântica  $\llbracket \phi \rrbracket_{\mathfrak{M}}^{w,p,\theta}$  fornece a denotação de qualquer sentença espacial  $\phi$  em relação a um modelo de precisificação  $\mathfrak{M}$ , um mundo possível  $w \in W$ , uma Teoria de Definição de Predicado  $\theta \in \Theta$  e uma precisificação  $p \in P$  (BENNETT, 2011).

A função de interpretação semântica (Definição 9) pode ser especificada, onde os conectivos lógicos possuem suas interpretações padrões:

$$\llbracket \neg \phi \rrbracket_{\mathfrak{M}}^{w,p,\theta} = \mathbf{v} \text{ se } \llbracket \phi \rrbracket_{\mathfrak{M}}^{w,p,\theta} = \mathbf{f}, \text{ caso contrário } = \mathbf{f} \quad (31)$$

$$\llbracket \phi \wedge \psi \rrbracket_{\mathfrak{M}}^{w,p,\theta} = \mathbf{v} \text{ se } \llbracket \phi \rrbracket_{\mathfrak{M}}^{w,p,\theta} = \mathbf{v} \text{ e } \llbracket \psi \rrbracket_{\mathfrak{M}}^{w,p,\theta} = \mathbf{v}, \text{ caso contrário } = \mathbf{f}. \quad (32)$$

**Definição 10.** A relação de satisfação semântica é definida por:

$$\mathfrak{M}, \langle w, p, \theta \rangle \Vdash \phi \text{ sse } \llbracket \phi \rrbracket_{\mathfrak{M}}^{w,p,\theta} = \mathbf{v}. \quad (33)$$

**Definição 11.** O conjunto interpretação de uma sentença em relação a um modelo  $\mathfrak{M}$  é dado por:

$$\llbracket \phi \rrbracket_{\mathfrak{M}} = \{ \langle w, p, \theta \rangle \mid \mathfrak{M}, \langle w, p, \theta \rangle \Vdash \phi \}. \quad (34)$$

O conjunto interpretação (Definição 11) é o conjunto de mundos possíveis, precisificações, teorias de definição de predicado em que a sentença  $\phi$  é considerada como verdade. Por exemplo, para a sentença  $\phi =$  “a maçã está no cesto”, o conjunto interpretação de  $\phi$  pode ser definido por  $\llbracket \phi \rrbracket_{\mathfrak{M}} = \{ \langle w_1, p_1, \theta_1 \rangle, \langle w_2, p_2, \theta_1 \rangle, \langle w_2, p_3, \theta_1 \rangle \}$  onde  $w_1$  e  $w_2$  são os mundos possíveis que representam, respectivamente, as situações de inclusão total e inclusão parcial,  $p_1, p_2$  e  $p_3$  são as precisificações exemplificadas acima e  $\theta_1$  é a TDP definida para a preposição “em” (predicado EM), dada por:  $\theta_1 : \text{EM}(tr, lm) \Leftrightarrow [\text{Container}(lm) \wedge \neg \text{Liquido}(tr)]$  onde  $tr$  é “maçã” e  $lm$  é “cesto”.

A função de interpretação para essa linguagem é tal que para cada tripla  $\langle$ esquema imagético, característica semântica, eixo espacial $\rangle$  corresponde aos mundos possíveis, e cada instanciação das relações REQ corresponde a uma precisificação. Em termos dessa semântica, cada fórmula proposicional é interpretada como um conjunto de triplas  $\langle$ mundos possíveis, precisificações, teorias de definição de predicado $\rangle$  em que a proposição é considerada como verdade. Concernente ao conjunto interpretação, Bennett (2011) diz:

A fim de levar em conta a compreensão de uma informação proposicional pelos agentes, precisamos relacionar o estado cognitivo dos agentes com a nossa semântica formal das proposições, que dá o significado de uma proposição em termos de um conjunto de interpretações. Dois aspectos do estado cognitivo são claramente relevantes:

o que o agente acredita sobre o estado do mundo e o que o agente considera um uso aceitável da terminologia (especialmente os predicados vagos). Se o agente considera uma proposição como verdadeira, dependerá de ambos os aspectos.

De acordo com a função de interpretação  $\llbracket \phi \rrbracket_{\mathfrak{M}}^{w,p,\theta}$  especificada acima, a atitude do agente para o significados dos termos é modelada em relação às TDP e as escolhas das precisificações que o agente considera ser aceitável. A atitude de um agente é uma estrutura denominada *standpoint* (ou estrutura de crença).

**Definição 12.** Uma *estrutura de crença* caracteriza a gama de mundos possíveis e interpretações que são consideradas plausíveis ou aceitáveis. Formalmente, uma estrutura de crença é uma tupla<sup>1</sup>

$$\langle B, A, \Psi \rangle \quad (35)$$

onde:

- 1)  $B \subseteq W$  é o conjunto de mundos possíveis que são compatíveis com o significado de uma expressão espacial;
- 2)  $A \subseteq P$  é o conjunto de precisificações que são admissíveis para uma expressão espacial;
- 3)  $\Psi \subseteq \Theta$  é o conjunto de TDP que caracteriza todas as possíveis definições de predicados que são considerados como aceitáveis.

Para um modelo  $\mathfrak{M}$ , é definida a condição que uma sentença  $\phi$  possui em relação a uma crença  $\langle B, A, \Psi \rangle$ . Formalmente

$$\mathfrak{M}, \langle B, A, \Psi \rangle \Vdash \phi \text{ sse } (B \times A \times \Psi) \subseteq \llbracket \phi \rrbracket_{\mathfrak{M}}. \quad (36)$$

Bennett (2011) diz que  $\phi$  é válida para uma estrutura de crença (*standpoint*) se é verdade em todos os mundos no conjunto de crença para todas as precisificações admissíveis e todas as Teoria de Definição de Predicado aceitáveis. Em outras palavras, o agente considera que, para qualquer interpretação razoável de predicado ambíguos e escolhas de limiares,  $\phi$  é verdade em todos os mundos possíveis consistentes com a crença do agente. Neste trabalho, a visão de Bennett (2011) sobre a crença do agente foi modificada para que ela esteja ligada com as características semânticas dos termos e as escolhas das relações de REQ.

Em suma, a ideia central desse capítulo foi estabelecer um modelo para que os significados das expressões espaciais pudessem ser caracterizadas. No geral, pode-se destacar quatro pontos principais: i) estabelecer os mundos possíveis; ii) identificar as Teoria de Definição de Predicado; iii) estabelecer as precisificações; iv) identificar as estruturas de crença. Com os mundos possíveis é possível verificar os sentidos de uma expressão espacial de forma individual, enquanto que as TDP e as precisificações levam em consideração as características e

<sup>1</sup>A Definição 12 e as fórmulas 35 e 36 foram retiradas de (BENNETT, 2011).

localização do trajetor e do marco. Por fim, os *standpoints* estão ligados a uma sentença específica, isto é, quais as interpretações possíveis para uma determinada sentença em termos de mundos possíveis, TDP e precisificações.

No próximo capítulo serão apresentadas a implementação do SSP para a preposição “in” e uma proposta de implementação para a preposição “em”.



## 6 IMPLEMENTAÇÃO DA SEMÂNTICA STANDPOINT PARA POLISSEMIA

Neste capítulo serão apresentados exemplos de implementação utilizando as preposição inglesa “*in*” e a preposição portuguesa “*em*”. Inicialmente serão estabelecidos os modelos de precisificação para cada uma delas, e em seguida haverá a aplicação desses modelos, através das estruturas de crença, em sentenças contendo tais preposições.

### 6.1 MODELO DE PRECISIFICAÇÃO ESPACIAL PARA A PREPOSIÇÃO *IN*

Seguindo a Definição 8 (Capítulo 5), o modelo de precisificação espacial  $\mathfrak{J}$  para a preposição *in* é dado por

$$\mathfrak{J} = \langle \mathcal{F}, T, R, V, \Theta, K, \xi, P \rangle \quad (37)$$

onde  $\mathcal{F} = \langle D, \mathcal{S}, \Sigma, S, X, W \rangle$  é um quadro espacial. Os conjuntos  $\mathcal{S}$  (esquemas imagéticos) e  $\Theta$  (características semânticas) foram obtidos com base nos trabalhos de Herskovits (1987) e Tyler e Evans (2003).

Herskovits (1987) define o significado ideal da preposição “*in*” como sendo a inclusão de um construto geométrico em um outro construto geométrico, podendo ser esse último unidimensional, 2D ou 3D. A seguir, serão mostrados os outros significados que Herskovits (1987) atribui à preposição “*in*”, e no final, como esses significados se relacionam com os mundos possíveis estabelecidos.

- 1) Entidade espacial em um contentor: o objeto referência tri-dimensional tem um interior em que o objeto localizado (também tri-dimensional) está total ou parcialmente contido.
  - O interior pode ser totalmente cercado pelo objeto: “the preserves in the sealed jar”;
  - Pode ser definido por uma concavidade em forma de xícara: “the milk in the glass”;
  - Pode ser dois planos, linhas ou objetos quase-cilíndricos reunidos em um ângulo: “the garbage in the gutter”;
  - Pode consistir em alguns espaços limitados por dois planos em um ângulo: “the dried flowers in the book”;
  - Um buraco de um objeto em forma de toroide: “the foot in the stirrup”;
  - O objeto referência pode ser um buraco em si mesmo: “the man in the doorway”;
  - O fechamento completo é relativamente raro, portanto algumas fronteiras do interior geralmente são imaginárias, mas bem definidas. Pode-se ter que imaginar os limites em mais de um ou dois lugares: “the bird in a cage”.

- 2) Lacuna ou objeto incorporado em um objeto físico.
  - O objeto localizado está incluído em uma região normalizada definida pelo objeto referência: “the nail in the board”;
  - Uma superfície conceitualizada como uma lâmina bem fina: “the crack in the surface”;
  - O objeto localizado é uma lacuna ou um buraco: “the hole in the wall”;
  - O objeto localizado está dissolvido ou misturado dentro do objeto referência: “sugar and milk in the coffee”.
- 3) Objeto físico no ar: “the bird in the air”.
- 4) Objeto físico no contorno de outro, ou de um grupo de objetos: “the bird in the tree”;
- 5) Entidade espacial localizada numa parte de um espaço ou ambiente: “There is a chair in the middle of the room”.
- 6) O objeto localizado pode ser uma parte física do objeto referência: “a page in a book”.
- 7) Pessoa em uma roupa: “a man in a red hat”;

Através dos casos de uso elencados acima, definiu-se os mundos possíveis para a preposição “in”. Foram detectados dois esquemas imagéticos principais: o de contenção, como na frase “the milk in the glass” e o de suporte, como na frase “the nails in the board”. As características semânticas de “in” foram divididas em quatro: i) inclusão, como na frase “the preserves in the sealed jar”; ii) cercamento, como em “the foot in the stirrup”; iii) encapsulamento, como em “the bird in the cage”; iv) aderência, como por exemplo “the nails in the board”. Os eixos espaciais foram o de conteúdo (dentro/fora) e verticalidade (cima/baixo). Ademais, tendo em mãos definições de Herskovits (1987), Tyler e Evans (2003) e Zelinsky-Wibbelt (1993) para a preposição “in”, foi feita uma busca no corpus COCA para a obtenção de frases que contivessem a preposição “in” empregada em sentido espacial. Após a análise dessas frases, estabeleceu-se o conjunto de características semânticas para os termos (trajetor e marco) que estão ligados pela preposição, dados pelo conjunto  $S$ . A definição de cada um dos termos foi obtida na *WordNet*. As Teoria de Definição de Predicado foram definidas através das características semânticas do trajetor e do marco, e são uma contribuição desta tese. Nas TDP, a preposição “in” é representada pelo predicado  $IN(x,y)$  e definida através de funções booleanas que representam cada uma das características semânticas definidas no conjunto  $S$ . Por fim, as precisificações representam a localização do trajetor e do marco, por meio de fórmulas de Raciocínio Espacial Qualitativo. Portanto, o modelo  $\tilde{\mathcal{T}}$  é definido como:

- 1)  $D$  são os termos representando os objetos do cenário, definidos de acordo com a sentença;
- 2)  $\mathcal{S} = \{\text{Contenção, Suporte}\}$  é o conjunto de esquemas imagéticos;
- 3)  $\Sigma = \{\text{Inclusão, Cercamento, Encapsulamento, Aderência}\}$  é o conjunto de características semânticas da preposição *in*;

- 4)  $S = \{$
- Container: qualquer objeto que pode ser usado para conter coisas;
  - Vessel: um objeto usado como um contentor, geralmente para líquidos;
  - Receptacle: um contentor que é usado para colocar ou guardar coisas dentro;
  - Liquid: substância que é líquida em temperatura ambiente;
  - Beverage: qualquer líquido que pode ser ingerido;
  - Clamp: dispositivo que mantém as coisas firmemente juntas;
  - Stirrup: suporte consistindo de argolas para cavaleiros colocarem os pés;
  - Enclosure: uma estrutura consistindo de uma área que pode ser encapsulada para algum propósito;
  - Extremity: uma parte externa do corpo humano, como pés ou mãos;
  - Pocket: um bolso pequeno dentro de uma vestimenta para guardar ou carregar artigos;
  - Bag: um contentor flexível com uma única abertura;
  - Pouch: um contentor pequeno ou médio para guardar ou carregar coisas;
  - Closet: armário usado para armazenamento;
  - Cabinet: um móvel que lembra um armário com portas, prateleiras e gavetas;
  - Car: veículo automotor com quatro rodas;
  - Tube: uma forma cilíndrica oca;
  - Trench: uma vala cavada como fortificação tendo um parapeito da terra escavada;
  - Ditch: um pequeno caminho natural para a água;
  - Gutter: um canal ao longo do beiral ou no telhado; recolhe e escoia a água da chuva;
  - Seat: peça de mobília designada para assento;
  - Stool: um assento simples sem braços ou encosto;
  - Water: um ambiente aquático;
  - Fastener: dispositivo que fixa ou anexa um objeto a outro;
  - Pan: utensílio de cozinha composta por uma vasilha geralmente de metal;
  - Dish: uma peça de louça normalmente usado como um recipiente para manter ou servir comida;
  - Tray: um receptáculo aberto para exibir ou servir comida (bandeja);
  - Oven: utensílio de cozinha para assar ou cozinhar;
  - Fridge: um refrigerador em que o refrigerante é bombeado por um motor;
  - Home Appliance: um utensílio que executa um trabalho em uma casa;
  - Support: qualquer dispositivo que suporta o peso de outro objeto;
  - Wall: uma divisória arquitetônica com altura e comprimento maior que sua espessura; usado para dividir ou delimitar uma área ou para suportar outra estrutura;
  - Equipment: um instrumento necessário para a realização de um serviço;

- **Compartment**: um espaço interior em que uma área é subdividida;
  - **Luggage**: itens utilizados para transportar pertences em viagem;
  - **Bin**: contentor, geralmente com uma tampa.
- } é o conjunto de características semânticas dos termos da sentença;
- 5)  $X = \{\text{Dentro, Fora, Em cima}\}$  é o conjunto de eixos espaciais;
- 6)  $W = \{$
- $w_1 = (\text{Contenção, Inclusão, Dentro})$ : o objeto está contido em um meio contentor.
  - $w_2 = (\text{Contenção, Inclusão, Fora})$ : o objeto está fora ou parcialmente contido em um meio contentor.
  - $w_3 = (\text{Contenção, Cercamento, Dentro})$ : o objeto está cercado por um contentor.
  - $w_4 = (\text{Contenção, Encapsulamento, Dentro})$ : um objeto está envolvido por um outro objeto.
  - $w_5 = (\text{Suporte, Aderência, Em cima})$ : um objeto está sendo suportado por outro objeto.
- } é o conjunto de mundos possíveis;
- 7)  $T = \{\text{DC, P, PP, EQ, O, PO, DR, EC, TPP, NTPP, DENTRO, P-DENTRO, FORA, TOP-DENTRO, GEO-DENTRO, ESP-DENTRO, NTC, CIMA}\}$  é o conjunto de predicados REQ;
- 8)  $R = \{\text{IN}(tr, lm)\}$  é o conjunto de predicados que representam os usos da preposição *in*;
- 9)  $V = \{tr, si, lm\}$  é o conjunto de símbolos variáveis que representam os papéis que os termos da sentença pode exercer;
- 10)  $\Theta$  é o conjunto de TDP, dado por:
- $\theta_1 : \text{IN}(tr, lm) \Leftrightarrow [\text{Container}(lm) \vee \text{Vessel}(lm) \vee \text{Receptacle}(lm)] \wedge [\neg \text{Liquid}(tr) \wedge \neg \text{Beverage}(tr)]$ : o marco é um contentor e o trajetor é um objeto sólido.
  - $\theta_2 : \text{IN}(tr, lm) \Leftrightarrow [\text{Container}(lm) \vee \text{Vessel}(lm) \vee \text{Receptacle}(lm)] \wedge [\text{Liquid}(tr) \vee \text{Beverage}(tr)]$ : o marco é um contentor e o trajetor é um líquido.
  - $\theta_3 : \text{IN}(tr, lm) \Leftrightarrow [\text{Wall}(lm)]$ : o marco é uma parede.
  - $\theta_4 : \text{IN}(tr, lm) \Leftrightarrow [\text{Stirrup}(lm) \vee \text{Clamp}(lm)]$ : o marco possui a forma de um anel, tal como um estribo, pinça ou alicate.
  - $\theta_5 : \text{IN}(tr, lm) \Leftrightarrow [\text{Enclosure}(lm)]$ : o marco é uma estrutura que consiste em uma área que foi fechada para algum propósito.
  - $\theta_6 : \text{IN}(tr, lm) \Leftrightarrow [\text{Extremity}(lm)]$ : o marco é uma extremidade do corpo humano.
  - $\theta_7 : \text{IN}(tr, lm) \Leftrightarrow [\text{Pocket}(lm) \vee \text{Bag} \vee \text{Pouch}(lm)]$ : o marco é um contentor flexível com uma única abertura.

- $\theta_8 : \text{IN}(tr, lm) \Leftrightarrow [\text{Water}(lm)]$ : o marco é um meio líquido.
  - $\theta_9 : \text{IN}(tr, lm) \Leftrightarrow [\text{Fastener}(tr)] \wedge [\neg \text{Container}(lm) \wedge \neg \text{Vessel}(lm) \wedge \neg \text{Receptacle}(lm)]$ : o trajetor é um elemento fixador e o marco não é um contâiner.
  - $\theta_{10} : \text{IN}(tr, lm) \Leftrightarrow [\text{Support}(lm)]$ : o marco é um suporte tal como um gancho de parede.
  - $\theta_{11} : \text{IN}(tr, lm) \Leftrightarrow [\text{Cabinet}(lm) \vee \text{Closet}(lm)]$ : o marco é um compartimento de armazenagem.
  - $\theta_{12} : \text{IN}(tr, lm) \Leftrightarrow [\text{Car}(lm)]$ : o marco é um veículo.
  - $\theta_{13} : \text{IN}(tr, lm) \Leftrightarrow [\text{Tube}(lm) \vee \text{Pipe}(tm)]$ : o marco é um tubo ou um cano.
  - $\theta_{14} : \text{IN}(tr, lm) \Leftrightarrow [\text{Seat}(lm) \wedge \neg \text{Stool}(lm)]$ : o marco é um assento, exceto quando for um banquinho.
  - $\theta_{15} : \text{IN}(tr, lm) \Leftrightarrow [\text{Trench}(lm) \vee \text{Ditch}(lm) \vee \text{Gutter}(lm)]$ : o marco é uma passagem para líquido ou uma longa e estreita escavação no solo.
  - $\theta_{16} : \text{IN}(tr, lm) \Leftrightarrow [\text{Pan}(lm) \vee \text{Dish}(lm) \vee \text{Tray}(lm)]$ : o marco é um recipiente aberto onde normalmente os alimentos são preparados ou servidos.
  - $\theta_{17} : \text{IN}(tr, lm) \Leftrightarrow [\text{Oven}(lm) \vee \text{Fridge}(lm) \vee \text{HomeAppliance}(lm)]$ : o marco é um aparelho eletrodoméstico.
  - $\theta_{18} : \text{IN}(tr, lm) \Leftrightarrow [\text{Equipment}(lm)]$ : o marco é um máquina ou um equipamento mecânico.
  - $\theta_{19} : \text{IN}(tr, lm) \Leftrightarrow [\text{Compartment}(lm)]$ : o marco é um compartimento dentro de uma área fechada.
  - $\theta_{20} : \text{IN}(tr, lm) \Leftrightarrow [\text{Bin}(lm)]$ : o marco é um contentor geralmente tampado.
- 11)  $K(d)$  é o conjunto de todas as características semânticas de um termo  $d \in D$ ;
- 12)  $\xi$  é o mapeamento dos termos de uma sentença aos papéis que eles exercem na frase;
- 13)  $P$  é o conjunto de precisificações, dado por:
- $p_1 \equiv_{def} \text{GEO} - \text{DENTRO}(x, y)$
  - $p_2 \equiv_{def} \text{P} - \text{DENTRO}(x, y)$
  - $p_3 \equiv_{def} \text{CONT} - \text{DENTRO}(x, y)$
  - $p_4 \equiv_{def} \text{FORA}(x, y) \wedge \exists z [[\text{GEO} - \text{DENTRO}(z, y) \vee \text{P} - \text{DENTRO}(z, y)] \wedge \text{EC}(x, z)]$
  - $p_5 \equiv_{def} \text{PO}(x, y)$
  - $p_6 \equiv_{def} \text{TPP}(x, y) \vee \text{NTPP}(x, y)$
  - $p_7 \equiv_{def} \neg \text{DENTRO}(x, y) \wedge \text{CIMA}(x, y) \wedge \text{EC}(x, y)$
  - $p_8 \equiv_{def} \text{TOP} - \text{DENTRO}(x, y)$

No modelo de precisificação espacial para a preposição “in” mostrado acima foram estabelecidos cinco mundos possíveis, que representam os casos de uso da preposição. Os mundos possíveis mostram o caráter polissêmico da preposição. Por exemplo, as frases “the bowl in the

microwave” e “the window in the wall” representam, respectivamente, os mundos possíveis  $w_1$  e  $w_4$ . No primeiro caso, a preposição “in” é usada no sentido de contenção, isto é, o “bowl” está dentro do “microwave”, enquanto que na segunda sentença a preposição é utilizada no sentido de envolvimento, ou seja, não pode-se dizer que “window” está dentro de “wall”, mas que “window” é uma abertura em “wall”.

Outra contribuição desta tese foi estabelecer um segundo nível de polissemia. Considere os seguintes exemplos:

- 1) the water in the bottle;
- 2) the pears in the bowl.

Em ambos os casos, o mundo possível evocado é o de contenção ( $w_1$ ). Contudo, no primeiro exemplo o trajetor (“water”) está totalmente contido no marco. Já no segundo exemplo, existe a possibilidade do trajetor “pear” estar fisicamente fora do “bowl”, mas em contato com outras peras que estão dentro do “bowl”. Assim, foram estabelecidas as TDP, baseadas nas características semânticas do trajetor e do marco. Portanto, cada mundo possível tem um conjunto de TDP associados a ele. Por fim, cada TDP possui um conjunto de precisificações. No exemplo de “the pears in the bowl”, há três possíveis localizações de “pear” em relação ao “bowl”: totalmente dentro do “bowl”, parcialmente dentro do “bowl” ou fora do “bowl” mas em contato com outras peras dentro do “bowl”. Assim, ao estabelecer os mundos possíveis, as TDP e as precisificações, pode-se entender os diferentes sentidos de uma preposição, que é a principal contribuição desta teste. Na próxima seção, será apresentado o algoritmo para a implementação de um sistema computacional e os resultados obtidos para o modelo de precisificação espacial da preposição “in”.

### 6.1.1 Algoritmo para o SSP

O algoritmo apresentado nesta seção tem por objetivo implementar o modelo de precisificação espacial dada uma determinada preposição. O algoritmo tem como entrada uma frase contendo uma expressão espacial e a saída são os *standpoints*, isto é, o conjunto de triplas  $\langle$ mundos possíveis, TDP, precisificações $\rangle$  compatíveis com a sentença. O algoritmo trabalha com uma tabela de Teoria de Definição de Predicado para a preposição, e cada TDP possui um conjunto de mundos possíveis e precisificações associados. Tal tabela é previamente determinada e inserida manualmente de acordo com o modelo de precisificação espacial.

Em seguida, o conjunto de características semânticas do trajetor e do marco são determinados e verificados na tabela de TDP quais delas são compatíveis com as características semânticas. Esses conjuntos são obtidos através de uma função de similaridade, que será explicada posteriormente. Por fim, uma vez que as TDPs compatíveis com a sentença são determinadas, o conjunto de mundos possíveis e precisificações compatíveis com a sentença são estabelecidos. Como cada TDP tem uma lista de precisificações e outra de mundos possíveis relacionados,

simplesmente são verificados nestas listas quais mundos possíveis e precisificações são compatíveis com uma TDP em particular.

As entradas para o trajetor, marco e indicador espacial são feitas manualmente, pois não foi automatizado o processo de extração desses elementos de uma frase. Os conjuntos de TDP, mundos possíveis e precisificações são pré-determinados de acordo com o modelo de precisificação espacial.

Para a implementação do Algoritmo 1 foi utilizada a *WordNet*, que está presente na biblioteca NLTK para Python. De acordo com a *WordNet*, um *synset* ou conjunto de sinônimos é definido como um conjunto de um ou mais sinônimos que são intercambiáveis em um contexto sem a mudança do valor verdade da proposição em que eles estão envolvidos.

Os conjuntos  $K_1$  e  $K_2$  são determinados por uma função de similaridade entre os *synsets* dos termos e os *synsets* pré-determinados pelo modelo de precisificação espacial. Essa função é dada por

`synset1.path.similarity(synset2)`: Retorna um valor que indica como dois sentidos de palavras são semelhantes, com base no caminho mais curto que conecta os sentidos em uma taxonomia (hiperônimos/hipônimos). A pontuação está no intervalo de 0 a 1. Por padrão, há um nó raiz falso adicionado aos verbos, portanto, nos casos em que anteriormente não foi possível encontrar um caminho - e Nenhum é retornado - ele ainda deve retornar um valor. O comportamento antigo pode ser alcançado definindo `simular` que a raiz seja falsa. Uma pontuação 1 representa identidade.<sup>1</sup>

Foram escolhidos valores acima de 0.33, que significa que um termo  $S$  está, no máximo, a dois níveis na árvore de hiperônimos (um hiperônimo é um vocábulo de sentido mais genérico em relação a outro. Por exemplo, assento é hiperônimo de cadeira, de poltrona, etc.). Os termos “container” e “basket” possuem um grau de similaridade 0.5, pois o *synset* de “container” está diretamente acima do *synset* de “basket”. A função de similaridade entre “vase” e “vessel” retorna 0.33 pois está a dois níveis na árvore.

O conjunto  $S$  é definido pelas características semânticas pré-estabelecidas para o trajetor e o marco. Como visto anteriormente, essas características semânticas dependem do indicador espacial. Nas linhas 9 a 17, para cada característica semântica é testada a similaridade entre o trajetor e o marco. O símbolo  $\odot$  nas linhas 25 e 37 é utilizado para verificar quais as TDP são verdade para uma determinada característica semântica. Assim, para cada característica semântica do trajetor e do marco são estabelecidas quais são as TDP compatíveis, isto é, as TDP que retornam o valor verdade. Na linha 44 é feita uma chamada para o Algoritmo 2, que verifica as possíveis interpretações para o conjunto de TDP  $\Psi$  e o indicador espacial  $R$ .

As linhas 5 e 6 do Algoritmo 2 fornecem os conjuntos de mundos possíveis e precisificações, respectivamente. Esses conjuntos são especificados manualmente, baseado no modelo de precisificação espacial (5). O símbolo  $\odot$  nas linhas 10 e 11 é usado para verificar os mundos possíveis e as precisificações associadas com a TDP. Para tanto, cada TDP tem um conjunto de mundos possíveis e precisificações correlacionadas e são informadas previamente. A saída

<sup>1</sup>Retirado de <http://www.nltk.org/howto/wordnet.html>.

deste algoritmo são os *standpoints*, isto é, o conjunto de triplas ⟨mundos possíveis, Teoria de Definição de Predicado (TDP), precisificações⟩ compatíveis com uma sentença  $\phi$ , ou uso incorreto se não há nenhuma TDP compatível com a sentença. Há duas interpretações para o caso onde  $\Psi$  é *NULL*:

- 1) O uso do esquema, como em “The bottle is in the lid”, e
- 2) Uso não previsto na Teoria de Definição de Predicado (TDP), como na frase “The knife is in the stone”.

### 6.1.2 Testes

O algoritmo foi testado com 71 frases preposicionais (FP), e os resultados foram divididos em 20 Teorias de Definição de Predicados, definidas no modelo de precisificação espacial  $\mathfrak{J}$ , visto na seção anterior. A Tabela 6 mostra esses resultados e foi obtida dos testes com cada uma das sentenças e mostra, em adição às TDP obtidas, os mundos possíveis e as precisificações associadas com cada TDP. É importante observar que metade das TDP estão correlacionadas com o mundo possível  $w_1$ , que significa a inclusão de um objeto em um meio contentor. Assim, o mundo possível  $w_1$  refere-se ao significado ideal de Herskovits (1987), que diz que cada preposição espacial possui um significado principal e um conjunto de significados derivados. Outro fato relevante é que 41 sentenças testadas estão também correlacionadas com o mundo possível  $w_1$ . Assim, pode-se associar  $w_1$  com o significado ideal da preposição “in”. Tyler e Evans (2003) propõem em seu modelo de polissemia baseado em princípios que cada preposição possui um significado principal, que os autores chamam de proto-cena, e uma rede semântica de outros significados. Seguindo o conceito de Tyler e Evans (2003), pode-se dizer que  $w_1$  refere-se a cena de “in”. Contudo, neste trabalho estendemos o conceito de proto-cena para representar não somente um significado, mas todas as possíveis interpretações para uma preposição em uma sentença, e será abordado com mais detalhes na Seção 6.1.3.

A Tabela 6 foi obtida testando cada uma das 71 FP no sistema implementado a partir dos Algoritmos 1 e 2. Antes, porém, foram definidos os mundos possíveis e as precisificações associados com cada uma das TDP. Essas associações foram feitas com base nos trabalhos de Herskovits (1987) e Tyler e Evans (2003) e são uma das contribuições desta tese.

Das FP analisadas, sete apresentaram uma ambiguidade na Teoria de Definição de Predicado. Isso ocorre pois uma TDP pode ter mais de um significado para uma palavra, como por exemplo a palavra “stirrup”, que pode ser caracterizada como um objeto em formato de anel ou um suporte. Na sentença “the jar in the bin”, a palavra “bin” foi classificada como um contentor simples (TDP  $\theta_1$ ) ou um contentor com uma tampa (TDP  $\theta_{20}$ ). Nove outras não foram classificadas em nenhuma TDP, pois não houve características semânticas compatíveis com os trajetores e marcos destas frases. A seguir, as frases que não foram classificadas:

- 1) “Sword in the stone”
- 2) “Wires in the cable”

Tabela 6 – Resultados por TDP

TDP	Quantidade	Mundos possíveis	Precisificações
$\theta_1$	15	$w_1, w_2$	$p_1, p_2, p_4$
$\theta_2$	3	$w_1$	$p_3$
$\theta_3$	4	$w_4$	$p_6$
$\theta_4$	1	$w_4$	$p_5, p_6$
$\theta_5$	1	$w_4$	$p_8$
$\theta_6$	2	$w_3$	$p_5$
$\theta_7$	1	$w_1$	$p_1$
$\theta_8$	4	$w_3$	$p_5, p_6$
$\theta_9$	4	$w_4$	$p_5$
$\theta_{10}$	3	$w_4$	$p_5$
$\theta_{11}$	4	$w_1$	$p_1$
$\theta_{12}$	3	$w_1$	$p_1$
$\theta_{13}$	1	$w_1$	$p_3$
$\theta_{14}$	1	$w_5$	$p_7$
$\theta_{15}$	2	$w_2$	$p_2$
$\theta_{16}$	1	$w_1, w_2$	$p_1, p_2$
$\theta_{17}$	10	$w_1$	$p_1$
$\theta_{18}$	1	$w_3$	$p_5$
$\theta_{19}$	2	$w_3$	$p_1$
$\theta_{20}$	3	$w_1$	$p_3$
Sem classificação	9	-	-

- 3) “Jewelry in the center of the table”
- 4) “Key in the cartouche”
- 5) “Blade in table saw”
- 6) “Jar in the mattress”
- 7) “Pans in the sink”
- 8) “Bottle in the towel”
- 9) “Dried flowers in the book”

Por fim, três sentenças testadas mostraram uso incomum da preposição “in”:

- 1) “Hat in the head”
- 2) “Head in the hat”
- 3) “Bottle in the cap”

Nos exemplos acima, o esquema imagético invocado não está correspondendo com o uso comum da preposição. Nos dois primeiros casos, a preposição mais indicada seria a “on”, enquanto que no terceiro exemplo, não é comum “bottle” estar em “cap”, pois “cap” não exerce um suporte em “bottle”.

Foi utilizado um *gold standard* para analisar os resultados obtidos. Esse padrão foi baseado em dicionários de uso, e também em trabalhos sobre semânticas de preposições, tais como Herskovits (1987), Tyler e Evans (2003) e Wunderlich (1985). Ademais, não foi possível

a comparação com outros sistemas pois não foi encontrado outro software na literatura para esse propósito. Assim, 12.67% das sentenças não foram classificadas. Porém, o algoritmo teve um uso positivo de 87.33% de sentenças corretamente classificadas. Para verificar a correção do SSP, foram criadas representações visuais para cada interpretação provida pelo algoritmo SSP. Essas representações visuais são figuras que representam esquematicamente o uso da preposição na sentença. Na próxima seção será mostrado como essas figuras são obtidas.

### 6.1.3 Representação visual

Baseado na teoria de coremas de Brunet (1986), que estabelece esquemas para representações de direção, o trabalho realizado por Ligozat, Nowak e Schmitt (2007) utiliza os coremas para representar de forma visual uma narrativa e também no conceito de proto-cenas de Evans (2006). Será apresentada nesta seção uma proposta de modificação destes conceitos para representar as interpretações de uma sentença através de ilustrações.

Originado no trabalho de Brunet (1986), coremas são componentes icônicos que podem ser utilizados para representar configurações espaciais específicas, bem como processos prototípicos em topografia. Também é importante pontuar que na teoria de Tyler e Evans (2003), o sentido central para uma preposição é diretamente fundamentado em um tipo específico de cena espacial. Essa cena espacial, que relaciona um trajetor e um marco em uma configuração espaço-geométrica é chamada de proto-cena. Ademais, o modelo de polissemia baseado em princípios diz que cada preposição da língua inglesa é representada por uma cena espacial primária ou proto-cena (TYLER; EVANS, 2003).

No escopo desta tese, o termo “cena” é usado para representar um arranjo entre trajetor e marco. A implementação das proto-cenas foi feita através da biblioteca OpenCV. Um conjunto de formas geométricas foi pré-definido para cada elemento do conjunto  $S$  e, para uma determinada sentença, o programa gera as formas do trajetor e do marco e suas localizações são determinadas pelas precisificações. A Figura 12 ilustra a proto-cena para a sentença  $\phi =$  “the pear in the bowl”. Neste caso, o conjunto de características semânticas de “bowl” é dado por  $K(bowl) = \{\text{Container}\}$  e o conjunto interpretação para a sentença é dado por  $[[\phi]]_S = \{\langle w_1, p_1, \theta_1 \rangle, \langle w_2, p_2, \theta_1 \rangle, \langle w_2, p_4, \theta_1 \rangle\}$ . Esse conjunto interpretação foi obtido através da aplicação do Algoritmo 1 na sentença. É importante observar que o conjunto de características semânticas do termo “bowl” possui apenas um elemento, logo o marco terá o mesmo formato nas três figuras. A posição do trajetor varia de acordo com a precisificação: a Figura 12(a) corresponde à interpretação  $\langle w_1, p_1, \theta_1 \rangle$  onde o trajetor está inteiramente contido no marco. Na interpretação  $\langle w_2, p_2, \theta_1 \rangle$ , o trajetor está parcialmente contido no marco (Figura 12(b)) enquanto que na interpretação  $\langle w_2, p_4, \theta_1 \rangle$ , representada pela Figura 12(c), o trajetor está fora do marco mas em contato com algum outro objeto dentro do marco.

A importância das cenas para este trabalho é que elas permitem visualizar esquematicamente cada interpretação de uma sentença. As demais proto-cenas para a preposição “in” estão

inseridas no Apêndice A.4.1. Na próxima seção será apresentado o modelo de precisificação espacial para a preposição portuguesa “em”.

## 6.2 MODELO DE PRECISIFICAÇÃO PARA A PREPOSIÇÃO “EM”

Nesta seção é apresentado o modelo de precisificação para a preposição “em” bem como a proposta de implementação, uma vez que não há uma interface implementada para a *WordNet* em língua portuguesa. Como o foco do trabalho não é desenvolver tal ontologia, o modelo de precisificação para a preposição “em” será apenas teórico. Vale ressaltar que o formalismo proposto nesse trabalho serve para expressões em mais de um idioma. Optou-se pela preposição “em” pois além de ser polissêmica, apresenta sentidos parecidos com as preposições inglesas “in”, “on” e “at”. A seguir, tem-se o modelo  $\mathfrak{E}$  para a preposição “em”.

$$\mathfrak{E} = \langle \mathcal{F}, T, R, V, \Theta, K, \xi, P \rangle \quad (38)$$

onde  $\mathcal{F} = \langle D, \mathcal{S}, \Sigma, S, X, W \rangle$  é quadro espacial. Portanto, o modelo  $\mathfrak{E}$  é definido como:

- 1)  $D$  são os termos representando os objetos do cenário, definidos de acordo com a sentença;
- 2)  $\mathcal{S} = \{\text{Contenção, Suporte, Verticalidade}\}$ ;
- 3)  $\Sigma = \{\text{Inclusão, Cercamento, Encapsulamento, Aderência, Sobreposição}\}$ ;
- 4)  $S = \{\text{contentor, Recipiente, Líquido, Pinça, Invólucro, CorpoHumano, Bolsa, Armário, Tubo, Vala, Assento, Fixador, Bandeja, Eletrodoméstico, Parede, Veículo, Adesivo, Plataforma}\}$  é o conjunto de características semânticas dos termos da sentença;
- 5)  $X = \{\text{Dentro, Fora, Em cima, Embaixo}\}$  é o conjunto de eixos espaciais;
- 6)  $W = \{$ 
  - $w_1 = (\text{Contenção, Inclusão, Dentro})$ : o objeto está contido em um meio contentor.
  - $w_2 = (\text{Contenção, Inclusão, Fora})$ : o objeto está fora ou parcialmente contido em um meio contentor.
  - $w_3 = (\text{Contenção, Cercamento, Dentro})$ : o objeto está cercado por um contentor.
  - $w_4 = (\text{Contenção, Encapsulamento, Dentro})$ : o objeto está envolvido por um outro objeto.
  - $w_5 = (\text{Suporte, Aderência, Em cima})$ : o objeto está aderido a um outro objeto.
  - $w_6 = (\text{Verticalidade, Sobreposição, Em cima})$ : o objeto está localizado sobre um outro objeto.

- $w_7 = (\text{Verticalidade, Sobreposição, Embaixo})$ : o objeto está localizado sob um outro objeto.
- } é o conjunto de mundos possíveis;
- 7)  $T = \{\text{DC, P, PP, EQ, O, PO, DR, EC, TPP, NTPP, DENTRO, P-DENTRO, FORA, TOP-DENTRO, GEO-DENTRO, ESP-DENTRO, NTC, CIMA, BAIXO}\}$  é o conjunto de predicados REQ;
- 8)  $R = \{\text{EM}\}$  é o conjunto de predicados que representam os usos da preposição “em”;
- 9)  $V = \{tr, si, lm\}$  é o conjunto de símbolos variáveis que representam os papéis que os termos da sentença pode exercer;
- 10)  $\Theta$  é o conjunto de TDP, dado por:
- $\theta_1 : \text{EM}(tr, lm) \Leftrightarrow [\text{contentor}(lm) \vee \text{Recipiente}(lm)] \wedge [\neg \text{Líquido}(tr)]$ : o marco é um recipiente e o trajetor é um objeto sólido.
  - $\theta_2 : \text{EM}(tr, lm) \Leftrightarrow [\text{contentor}(lm) \vee \text{Recipiente}(lm)] \wedge [\text{Líquido}(tr)]$  o marco é um recipiente e o trajetor é um meio líquido.
  - $\theta_3 : \text{EM}(tr, lm) \Leftrightarrow [\text{Líquido}(lm)]$ : o marco é um meio líquido.
  - $\theta_4 : \text{EM}(tr, lm) \Leftrightarrow [\text{Vala}(lm)]$ : o marco não possui os lados delimitados, como uma vala.
  - $\theta_5 : \text{EM}(tr, lm) \Leftrightarrow [\text{Pinça}(lm)]$ : o marco possui o formato de uma pinça, tal como um alicate ou um estribo.
  - $\theta_6 : \text{EM}(tr, lm) \Leftrightarrow [\text{Invólucro}(lm)]$ : o marco é um invólucro.
  - $\theta_7 : \text{EM}(tr, lm) \Leftrightarrow [\text{CorpoHumano}(lm)]$ : o marco é uma parte do corpo humano
  - $\theta_8 : \text{EM}(tr, lm) \Leftrightarrow [\text{Bolsa}(lm)]$ : o marco é um contâiner flexível usualmente com um lado aberto.
  - $\theta_9 : \text{EM}(tr, lm) \Leftrightarrow [\text{Armário}(lm)]$ : o marco é um espaço utilizado para armazenamento.
  - $\theta_{10} : \text{EM}(tr, lm) \Leftrightarrow [\text{Tubo}(lm)]$ : o marco possui o formato de um tubo.
  - $\theta_{11} : \text{EM}(tr, lm) \Leftrightarrow [\text{Assento}(lm)]$ : o marco é um assento.
  - $\theta_{12} : \text{EM}(tr, lm) \Leftrightarrow [\text{Fixador}(tr)]$ : o trajetor é um elemento fixador, como um parafuso ou prego.
  - $\theta_{13} : \text{EM}(tr, lm) \Leftrightarrow [\text{Bandeja}(lm)]$ : o marco tem o formato de uma bandeja ou prato.
  - $\theta_{14} : \text{EM}(tr, lm) \Leftrightarrow [\text{Eletrodoméstico}(lm)]$ : o marco é um eletrodoméstico, tal como uma geladeira ou fogão.
  - $\theta_{15} : \text{EM}(tr, lm) \Leftrightarrow [\text{Parede}(lm)]$ : o marco é uma parede ou objeto na posição vertical.
  - $\theta_{16} : \text{EM}(tr, lm) \Leftrightarrow [\text{Veículo}(lm)]$ : o marco é um veículo.
  - $\theta_{17} : \text{EM}(tr, lm) \Leftrightarrow [\text{Adesivo}(tr)]$ : o trajetor é um objeto aderente, como um adesivo.

- $\theta_{18} : EM(tr,lm) \Leftrightarrow [Plataforma(lm)]$ : o marco é uma plataforma horizontal.
- } é o conjunto de TDP;
- 11)  $K(d)$  é o conjunto de todas as características semânticas de um termo  $d \in D$ ;
  - 12)  $\xi$  é o mapeamento dos termos de uma sentença aos papéis que eles exercem na frase;
  - 13)  $P$  é o conjunto das precisificações, dado por:
    - $p_1 \equiv_{def} GEO - DENTRO(x,y)$
    - $p_2 \equiv_{def} TOP - DENTRO(x,y)$
    - $p_3 \equiv_{def} CONT - DENTRO(x,y)$
    - $p_4 \equiv_{def} P - DENTRO(x,y)$
    - $p_5 \equiv_{def} FORA(x,y) \wedge \exists z[[GEO - DENTRO(z,y) \vee P - DENTRO(z,y)] \wedge EC(z,x)]$
    - $p_6 \equiv_{def} PO(x,y)$
    - $p_7 \equiv_{def} TPP(x,y) \vee NTPP(x,y)$
    - $p_8 \equiv_{def} \neg DENTRO(x,y) \wedge CIMA(x,y) \wedge EC(x,y)$
    - $p_9 \equiv_{def} \neg DENTRO(x,y) \wedge BAIXO(x,y) \wedge EC(x,y)$

### 6.2.1 Sentenças com a preposição “em”

- 1)  $\phi_1 = \text{“a vasilha na geladeira”}$ 
  - $\xi(tr) = \text{vasilha}$
  - $\xi(lm) = \text{geladeira}$
  - $\xi(si) = \text{em}$
  - $K(tr) = \{\text{contentor}\}$
  - $K(lm) = \{\text{Eletrodoméstico}\}$
  - Mundo possível:  $B = \{w_1, w_6\}$
  - Precisão:  $A = \{p_1, p_8\}$
  - TDP:  $\Psi = \{\theta_{14}\}$
  - Conjunto interpretação:  $[[\phi_1]]_{\epsilon} = \{\langle w_1, p_1, \theta_{14} \rangle, \langle w_6, p_8, \theta_{14} \rangle\}$

Nesse exemplo há duas interpretações possíveis: a primeira, de acordo com a precisificação  $p_1$ , o trajetor “vasilha” está dentro do marco “geladeira”; na segunda interpretação, a precisificação  $p_8$  mostra que o trajetor está localizado em cima do marco.

- 2)  $\phi_2 = \text{“a etiqueta na geladeira”}$ 
  - $\xi(tr) = \text{etiqueta}$
  - $\xi(lm) = \text{geladeira}$
  - $\xi(si) = \text{em}$
  - $K(tr) = \{\text{Adesivo}\}$

- $K(lm) = \{\text{Eletrodoméstico}\}$
- Mundo possível:  $B = \{w_1, w_5, w_6\}$
- Precisão:  $A = \{p_1, p_7, p_8\}$
- TDP:  $\Psi = \{\theta_{14}, \theta_{17}\}$
- Conjunto interpretação:  $[[\phi_2]]_{\mathcal{E}} = \{\langle w_1, p_1, \theta_{14} \rangle, \langle w_6, p_8, \theta_{14} \rangle, \langle w_5, p_7, \theta_{17} \rangle\}$

Em relação à sentença  $\phi_1$ , a sentença  $\phi_2$  apresenta uma terceira interpretação: o trajetor “etiqueta” pode estar dentro, em cima ou aderido externamente ao marco “geladeira”.

3)  $\phi_3 = \text{“as flores no vaso”}$

- $\xi(tr) = \text{flores}$
- $\xi(lm) = \text{vaso}$
- $\xi(si) = \text{em}$
- $K(tr) = \{\}$
- $K(lm) = \{\text{Contâiner}\}$
- Mundo possível:  $B = \{w_1, w_2\}$
- Precisão:  $A = \{p_1, p_4\}$
- TDP:  $\Psi = \{\theta_1\}$
- Conjunto interpretação:  $[[\phi_3]]_{\mathcal{E}} = \{\langle w_1, p_1, \theta_1 \rangle, \langle w_2, p_2, \theta_1 \rangle\}$

Na sentença  $\phi_3$  tem-se duas interpretações possíveis: na primeira as flores estão totalmente incluídas dentro do vaso, enquanto que na segunda as flores podem estar parcialmente incluídas no vaso.

4)  $\phi_4 = \text{“a água na garrafa”}$

- $\xi(tr) = \text{água}$
- $\xi(lm) = \text{garrafa}$
- $\xi(si) = \text{em}$
- $K(tr) = \{\text{Líquido}\}$
- $K(lm) = \{\text{Contâiner}\}$
- Mundo possível:  $B = \{w_1\}$
- Precisão:  $A = \{p_3\}$
- TDP:  $\Psi = \{\theta_2\}$
- Conjunto interpretação:  $[[\phi_4]]_{\mathcal{E}} = \{\langle w_1, p_3, \theta_2 \rangle\}$

Na sentença  $\phi_4$  só há uma interpretação possível: o trajetor é um líquido e o marco um recipiente. Assim, o trajetor está preenchendo o espaço dentro do marco.

5)  $\phi_5 = \text{“o vaso na mesa”}$

- $\xi(tr) = \text{vaso}$
- $\xi(lm) = \text{mesa}$
- $\xi(si) = \text{em}$
- $K(tr) = \{\text{Contâiner}\}$
- $K(lm) = \{\text{Plataforma}\}$

- Mundo possível:  $B = \{w_6\}$
- Precisão:  $A = \{p_8\}$
- TDP:  $\Psi = \{\theta_{18}\}$
- Conjunto interpretação:  $\llbracket \phi_5 \rrbracket_{\mathcal{E}} = \{\langle w_6, p_8, \theta_{18} \rangle\}$

Na sentença  $\phi_5$  o trajector “vaso” está localizado em cima de uma plataforma (mesa), sendo a única interpretação.

Esses exemplos mostram o carácter altamente polissêmico da preposição “em”. Como o objetivo foi estabelecer somente o modelo de precisificação espacial para a preposição “em”, não foi implementado um sistema computacional para o teste automático da sentença, sendo que os exemplos mostrados nesta seção foram obtidos manualmente. Ademais, os mundos possíveis, as TDP e as precisificações foram obtidas baseados em dicionários e trabalhos sobre as preposições em português, tais como Pontes (1992), Oliveira (2009) e Garcia (2018).

### 6.3 CONSIDERAÇÕES FINAIS DO CAPÍTULO

Neste capítulo foram apresentados dois modelos de precisificações para as preposições “in”, na língua inglesa, e “em” em português. Para tanto, foram estabelecidos os mundos possíveis, ou seja, as situações onde as preposições podem ser empregadas, destacando assim o carácter polissêmico das preposições. No caso da preposição “em” foram elencados sete mundos possíveis em três esquemas imagéticos diferentes. Isso mostra o alto grau de polissemia da palavra “em”. Em seguida, baseado em dicionários de uso, gramáticas e ontologias, foram definidos os conjuntos de características semânticas dos termos que acompanham a preposição e também as precisificações, que são as possíveis localizações espaciais do trajector em relação ao marco. Além disso, foi mostrada a implementação do algoritmo SSP para a preposição “in” baseado no modelo de precisificação espacial. Os resultados obtidos foram satisfatórios, pois 87% das sentenças foram classificadas de forma correta.

Figura 9 – Algoritmo 1 - Teorias de Definição de Predicados

```

1 início
2   Entrada:  $\phi$ 
3   Saída: Interpretação
4    $tr \leftarrow$  trajetor de  $\phi$ 
5    $lm \leftarrow$  marco de  $\phi$ 
6    $R \leftarrow$  indicador espacial de  $\phi$ 
7    $\Theta \leftarrow$  TDPs de  $R$ 
8    $S \leftarrow$  características semânticas de  $R$ 
9    $n \leftarrow |S|$ 
10   $i \leftarrow 1$ 
11  enquanto  $i \leq n$  faça
12    se  $S_i.path.similarity(tr) \geq 0.33$  então
13       $K_1(tr) \leftarrow S_i$ 
14    fim
15    se  $S_i.path.similarity(lm) \geq 0.33$  então
16       $K_2(lm) \leftarrow S_i$ 
17    fim
18     $i \leftarrow i + 1$ 
19  fim
20   $i \leftarrow 1$ 
21   $t \leftarrow |\Theta|$ 
22   $m \leftarrow |K_1|$ 
23  enquanto  $i \leq m$  faça
24     $x \leftarrow 1$ 
25    enquanto  $x \leq t$  faça
26      se  $K_1(tr)_i \odot \theta_x = v$  então
27         $\Psi \leftarrow \theta_x$ 
28      fim
29       $x \leftarrow x + 1$ 
30    fim
31     $i \leftarrow i + 1$ 
32  fim
33   $j \leftarrow 1$ 
34   $m \leftarrow |K_2|$ 
35  enquanto  $j \leq m$  faça
36     $y \leftarrow 1$ 
37    enquanto  $y \leq t$  faça
38      se  $K_2(lm)_j \odot \theta_y = v$  então
39         $\Psi \leftarrow \theta_y$ 
40      fim
41       $y \leftarrow y + 1$ 
42    fim
43     $j \leftarrow j + 1$ 
44  fim
45  return Interpretação ( $\Psi, R$ )
46 fim

```

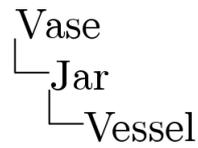


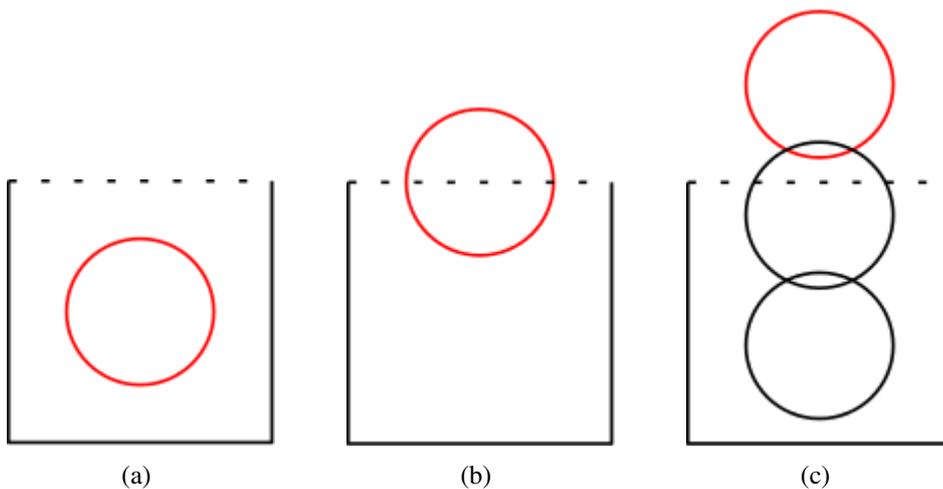
Figura 10 – Árvore de hiperônimos para “vase”

Figura 11 – Algoritmo 2 - Semântica Standpoints para Polissemia

```

1 Entrada: Interpretação  $\Psi, R$ 
2 Saída: Standpoint
3 se  $\Psi = NULL$  então
4   return Uso Incorreto
5 senão
6    $W \leftarrow$  mundos possíveis de  $R$ 
7    $P \leftarrow$  precisificações de  $R$ 
8    $n \leftarrow |\Psi|$ 
9    $z \leftarrow 1$ 
10  enquanto  $z \leq n$  faça
11     $B \leftarrow \Psi_z \odot W$ 
12     $A \leftarrow \Psi_z \odot P$ 
13     $z \leftarrow z + 1$ 
14  fim
15  return  $\{\langle \Psi, B, A \rangle\}$ 
16 fim
17 fim
  
```

Figura 12 – Exemplo de proto-cenas



(d) Fonte: Autor



## 7 DISCUSSÃO E CONSIDERAÇÕES FINAIS

Nesta tese foi apresentado um formalismo lógico denominado Semântica Standpoint para Polissemia (SSP) para o tratamento da polissemia em expressões espaciais. Inicialmente foi modificado o trabalho de Bennett (2011), originalmente utilizado para o tratamento da vagueza, para trabalhar com polissemia. O formalismo de Bennett (2011) é uma semântica de Supervaloração, onde as precisificações são atribuições de valores à limiares vagos e os predicados vagos são definidos por Teorias de Definição de Predicado (TDP), que são fórmulas de uma linguagem de primeira ordem envolvendo funções de medição e limiares. No SSP, a principal contribuição foi modificar a precisificação para que pudesse determinar a localização do trajetor em relação ao marco. Em relação às TDP, estabeleceu-se uma linguagem de primeira ordem baseada nas características semânticas do trajetor e do marco. Essas características semânticas foram definidas com base em dicionários e em trabalhos relacionados com semânticas de preposições espaciais. Na implementação do sistema computacional para o SSP, utilizou-se a *WordNet* para extrair tais características semânticas. Uma vez definidas as características semânticas, passou-se à fase de elaboração das TDP. Uma Teoria de Definição de Predicado dá o significado para um predicado da linguagem em relação às características semânticas do trajetor e marco. Os predicados da linguagem são os indicadores espaciais (os termos que relacionam trajetor e marco, normalmente representados por preposições). Assim, cada indicador espacial tem um conjunto de TDP.

Optou-se por utilizar a abordagem de semântica de supervaloração pois há alguns trabalhos que tratam a vagueza com supervaloração. Como este trabalho assume que a polissemia pode ser tratada como um fenômeno análogo à vagueza, acreditou-se que seria viável a modificação do formalismo de Bennett (2011) para a polissemia. Ademais, a vantagem dessa abordagem é que ela é tolerante à elaboração, isto é, para cada novo conhecimento basta acrescentar uma nova regra. Outra vantagem para o uso da Supervaloração é que o sistema computacional implementado a partir dela pode ser feito por um sistema baseado em regras, sem a necessidade de treinamento e base de testes. Há alguns trabalhos na literatura, como o de Soto, Olivas e Prieto (2008), que utiliza lógica *fuzzy* para recuperação de informação em sinonímia e polissemia. Nesse trabalho, os autores estabeleceram índices de sinonímia e polissemia através da quantidade de significados de um determinado termo. Ademais, são citados alguns trabalhos que tratam a questão da polissemia em preposições, tais como o de Garcia (2018), que propôs uma rede, denominada PrepNet, para os significados de preposições em Português do Brasil, o de Tyler e Evans (2003) que apresenta a semântica das preposições em Inglês e o de Herskovits (1987), que traz um estudo cognitivo sobre as preposições em Inglês. A principal desvantagem do SSP é que, para cada termo é necessário extrair as características semânticas e fazer a comparação com todas as regras determinadas pelas TDP, ocasionando uma certa lentidão, da ordem de segundos, no sistema. Assim quanto mais regras, mais demorada será a resposta dada. Por fim, não foi encontrado na literatura um outro sistema que trate computacionalmente

a polissemia de preposições espaciais e, sendo assim, não houve critérios de comparação com outros sistemas.

Para os indicadores espaciais, definiu-se um conjunto de mundos possíveis, que são as possibilidades de uso de um indicador espacial. Os mundos possíveis são uma combinação de características semânticas, esquemas imagéticos e eixos espaciais. Esse método de classificação dos indicadores espaciais foi baseado no trabalho de Ilari et al. (2008) e utilizado por Garcia (2018). Assim, os mundos possíveis definem os diversos sentidos (polissemia) que um indicador espacial pode ter. Os esquemas imagéticos definem a categorização cognitiva do indicador espacial, isto é, qual o papel cognitivo que ele exerce. Por exemplo, temos os esquemas de Contenção, Suporte, Proximidade, entre outros. Já as características semânticas são os papéis secundários exercidos pelo indicador espacial. Por exemplo, para o esquema de imagem de Contenção, o indicador espacial pode representar a característica de conteúdo ou de cercamento. Por fim, os eixos espaciais determinam a organização espacial expressada pelo indicador espacial. Por exemplo, no esquema de contenção, os eixos espaciais podem representar a situação em que o trajetor está dentro ou fora do marco.

Como o intuito do SSP é, em um momento futuro, ser implementado em um agente robótico inteligente, escolheu-se determinar os mundos possíveis para um indicador espacial utilizando, principalmente, os esquemas imagéticos. Os esquemas imagéticos são as representações do processo cognitivo que estabelecem de padrões de entendimento e raciocínio humano. Assim, para estabelecer um sistema inteligente que seja capaz de interagir com humanos é necessário que ele consiga simular seus padrões de comportamento (KURZWEIL, 2012). Por isso optou-se por categorizar os indicadores espaciais de acordo com os esquemas de imagem que eles representam.

Além das TDP e dos mundos possíveis, o conjunto interpretação de uma sentença necessita também das precisificações. Neste trabalho, utilizou-se um formalismo de Raciocínio Espacial Qualitativo (REQ) denominado Cálculo de Conexão de Regiões (CCR) para se determinar as precisificações. Assim, um indicador espacial possui um conjunto de arranjos espaciais, determinados pelo CCR, para indicar a localização no espaço do trajetor em relação ao marco. De posse desses três conceitos, pode-se então estabelecer o conjunto interpretação de uma sentença, que é o conjunto de triplas <mundos possíveis, teorias de definição de predicados, precisificações> que são admissíveis para uma sentença.

*Standpoints* são os conjuntos de interpretações possíveis para uma sentença são os principais componentes do SSP. Com ele pode-se verificar com precisão os sentidos de um indicador espacial dentro de uma frase. Assim, um agente inteligente pode, precisamente, compreender o sentido de uma frase através das interpretações plausíveis dadas pelo *standpoint*. Uma das vantagens do *standpoint* é que ele especifica quais as interpretações são possíveis, bastando que o agente escolha qual a melhor interpretação possível para uma determinada cena. Assim, decidiu-se representar de forma visual as interpretações, dando origem às proto-cenas.

Com base nos conjuntos de interpretações as proto-cenas foram definidas, que são as representações visuais de cada interpretação possível de um termo polissêmico. Em uma proto-cena, o formato do trajetor e do marco são determinados pelas TDP, enquanto a disposição no espaço serve para se determinar suas localizações. Optou-se por fazer essa representação visual para avaliar o grau de acerto do SSP em relação ao resultado esperado. Neste ponto vale ressaltar que as proto-cenas foram estabelecidas por um padrão *gold standard* de teste, pois não foi encontrado na literatura algum outro trabalho que faça tal classificação. O padrão *gold standard* foi definido com base em trabalhos sobre preposições, tais como o de Tyler e Evans (2003) e Herskovits (1987), que apresentam os casos de uso e as redes semânticas das preposições em Inglês. Como contribuição do presente trabalho, essas informações foram compiladas e também foi criado um padrão de testes para a preposição “in”. Dado que os trabalhos de Tyler e Evans (2003) e Herskovits (1987) são referências na área de linguística, pode-se basear neles para estabelecer o padrão de testes. Assim, para definir as saídas esperadas para cada frase, estabeleceu-se um conjunto de regras fundamentadas nesse padrão.

Munidos desses conceitos, passou-se à fase de implementação do SSP. Para tanto, foi escrito um script na linguagem Python e biblioteca NLTK para extrair o grau de similaridade de um termo em relação às características semânticas pré-determinadas nas TDP. A função de similaridade faz uma busca na árvore de hiperônimos de um termo  $x$  dado um termo  $y$ , retornando um valor entre 0 e 1. Quanto mais próximo de 1, maior é a similaridade e menor a distância de  $y$  na árvore de hiperônimos de  $x$ . Por exemplo, a função de similaridade para “Container” e “Basket” retorna o valor 0.5, mostrando que “Container” está posicionado no primeiro nível da árvore de “Basket”. Foi estabelecido um limiar de 0.33 para dizer que um termo é similar a um outro, ou seja, se eles pertencem a uma mesma classe de objetos. Esse limiar foi adotado pois fixou-se a busca na árvore de hiperônimos até o segundo nível. Esse corte foi feito com o intuito de eliminar discrepâncias, como por exemplo classificar o objeto “Basket” dentro da mesma classe do objeto “Pants”. Assim, após testes com um grupo de palavras, fixou-se o limiar 0.33. Assim, um limiar muito baixo classifica objetos conceitualmente muito diferentes dentro de uma mesma classe, ao passo que um limiar muito próximo a 1 necessitaria de muitas classes de objetos. Por exemplo, com um limiar acima 0.5 seria necessário criar uma classe para o objeto “Jar” e uma outra para o objeto “Basket”, porém pode-se agrupar ambos objetos dentro da classe de objetos “Container”.

Para testar o SSP, foram extraídas 71 frases preposicionais do corpus COCA contendo a preposição “in” fazendo a ligação entre os objetos trajetor e marco (Capítulo 6). O SSP classificou corretamente 87% dessas frases enquanto que para os outros 13% não houve similaridade com nenhuma característica semântica com as TDP definidas. Dois exemplos são citados para frases que não foram classificadas:

- 1) Knife in the stone;
- 2) Bottle in the mattress.

Isso ocorreu pois em ambas as frases os marcos (“stone” e “mattress”) não foram classificadas em nenhuma das 36 classes de palavras definidas no conjunto  $S$  (e utilizadas para a criação das TDP). Optou-se por não criar mais classes de palavras e tratar os 13% como exceções.

Por fim, o SSP foi capaz de prever 3 usos incomuns para a preposição “in”, entre eles a frase “the bottle in the lid”. Neste exemplo, devido às particularidades do trajecto e do marco, não é comum dizer que a Garrafa está na Tampa, pois a Tampa não oferece nenhuma relação de suporte ou de contenção sobre a garrafa. Aqui, retomando o conceito de esquema imagético, dir-se-ia que não há um esquema que associa Garrafa na Tampa, mas o inverso é totalmente plausível, pois a Garrafa exerce um esquema de suporte em relação à Tampa.

As 71 frases foram extraídas do corpus através de uma busca pela preposição “in” com substantivos colocados à uma distância de palavras à esquerda e à direita da preposição. Em seguida foram selecionadas aquelas em que os substantivos se referiam a algum objeto. Assim, foi delimitado somente frases contendo substantivos que fazem referência a objetos. Essas 71 frases foram utilizadas somente como prova do conceito, e em trabalhos futuros o formalismo proposto será testado em alguma base de testes de tamanho apropriado. Ademais, devido o SSP ser tolerante à elaboração, surgindo um novo uso, basta acrescentar uma nova regra nos moldes de uma TDP.

Além de implementar o SSP para a preposição “in”, este trabalho também traz como contribuição um modelo de precisificação espacial para a preposição “em”. Por se tratar de uma preposição com alto grau de polissemia, optou-se por estabelecer esse modelo. O SSP para a preposição “em” não foi implementado em um sistema computacional pois ainda não há uma ferramenta para calcular a função de similaridade entre *synsets* na língua portuguesa.

### Artigos publicados durante o doutorado

Durante o doutorado foram publicados quatro artigos sobre o tema polissemia em preposições espaciais. A seguir, é dado um breve resumo sobre esses artigos.

- 1) “Pinning down polysemy: A formalisation for a Brazilian Portuguese preposition”<sup>1</sup>: neste artigo foi apresentada uma proposta para o tratamento da polissemia da preposição “em” baseada em uma formalização dos significados de “em” através de Raciocínio Espacial Qualitativo. Foram definidas 14 situações onde a preposição “em” poderia ser utilizada e foi feito um comparativo entre as respostas obtidas pelo formalismo e as respostas obtidas por pessoas que responderam a um questionário sobre os usos de “em” nessas 14 situações.
- 2) “A formalism for the treatment of polysemy of vertical prepositions in Brazilian Portuguese”<sup>2</sup>: neste artigo foi apresentado um estudo para as preposições verti-

<sup>1</sup>Disponível em: <https://www.sciencedirect.com/science/article/pii/S1389041716300766>

<sup>2</sup>Disponível em: [http://qrg.northwestern.edu/qr2017/papers/QR2017\\_paper\\_3.pdf](http://qrg.northwestern.edu/qr2017/papers/QR2017_paper_3.pdf)

cais (sobre e sob) na língua portuguesa. Foram estabelecidos 22 significados para essas preposições, através de consultas à gramáticas e demais trabalhos sobre preposições, além de uma análise de 384 sentenças obtidas em corpus.

- 3) “Formalism for Treatment of the Ambiguity in Front/Back Axis Expressions”<sup>3</sup>: neste artigo foi apresentado um modelo para o tratamento da ambiguidade em expressões espaciais projetivas, tais como “em cima”, “embaixo”, “à direita de”, “à esquerda de”, “na frente”, “atrás”, entre outras. Para tanto, foi estudada as formas cognitivas de como essas expressões são empregadas. Por exemplo, questões como perspectivas déctica, intrínseca e extrínseca são levadas em consideração para o emprego de tais preposições. Por fim, foi proposto um modelo baseado em Cálculo de Direções Cardinais e uma heurística para poder definir o uso das preposições projetivas.
- 4) “Standpoints Semantics for Polysemy in Spatial Prepositions”<sup>4</sup>: esse artigo mostra uma versão resumida da formalização e implementação do SSP abordado nesta tese.

## 7.1 CONSIDERAÇÕES FINAIS

Foi apresentado neste trabalho um formalismo denominado Semântica Standpoint para Polissemia para o tratamento da polissemia em expressões espaciais, baseado em semântica de supervaloração. Classificou-se os usos polissêmicos de tais expressões em uma tripla de elementos ⟨características semânticas, esquemas imagéticos, eixos espaciais⟩ denominada mundos possíveis e estabeleceu-se o conceito de Teoria de Definição de Predicado, que é a parte principal do formalismo. Com as TDP é possível determinar a aplicação do indicador espacial tendo por base as características semânticas do trajetor e do marco. Juntamente com as precisificações, que no SSP são fórmulas de um formalismo de REQ, foi mostrado o conceito de *standpoint*, que dá as interpretações possíveis para uma determinada sentença. Por fim, para cada interpretação possível de uma sentença foi dada uma representação visual denominada proto-cena.

A principal contribuição desta tese foi modificar um formalismo para o tratamento da vagueza, baseado em semântica de supervaloração, para tratar a polissemia em expressões espaciais. Outra contribuição foi o conceito de proto-cena, que é a representação visual das interpretações possíveis para uma sentença. Além disso, foi implementado um sistema computacional para o SSP para tratar a polissemia da preposição “in”, sendo assim uma contribuição prática deste trabalho.

Como trabalhos futuros, pretende-se estabelecer modelos de precisificação espacial para outras preposições, tanto em inglês quanto em português, e também implementar uma versão do SSP para as preposições em Português (no Apêndice 7.1 há um estudo sobre as preposições

<sup>3</sup>Disponível em: <https://aaai.org/ocs/index.php/FSS/FSS17/paper/download/15957/15287>

<sup>4</sup>Em submissão

espaciais em Português, que servirá de base para a implementação do SSP para essas preposições). Para tanto, será necessário construir uma ontologia para se extrair as características semânticas dos termos em Português. Um outro trabalho a ser feito é implementar uma rede neural para se obter as interpretações possíveis para uma sentença espacial, onde as entradas seriam os índices de similaridade com uma gama de categorias de objetos pré-estabelecidas. Isso será feito também através da função de similaridade entre os *synsets* e saída seria as possíveis localizações do trajetor em relação ao marco. Outro trabalho a ser desenvolvido é estabelecer um modelo de geração de imagens 3D a partir de descrições espaciais textuais, tendo por base trabalhos como o de Chang et al. (2015), Coyne e Sproat (2001), Johansson et al. (2005), Krishnaswamy e Pustejovsky (2016) e Pustejovsky e Krishnaswamy (2016).

## **APÊNDICE A – EXPRESSÕES ESPACIAIS**

Neste apêndice serão apresentadas as preposições espaciais na língua portuguesa. O uso dessas preposições está diretamente relacionado aos conceitos de esquemas imagéticos abordados no Capítulo 4. Por exemplo, o esquema imagético de Contenção está associado aos usos das expressões espaciais “em”, “dentro de” ou “incluso”, mas não está relacionado com as preposições “sobre” ou “sob”. Por isso é necessário entender como os esquemas imagéticos e as expressões espaciais se relacionam.

Entretanto, nesse apêndice serão considerados apenas os esquemas imagéticos estáticos, ou seja, aqueles que não apresentam a ideia de movimento. Seguindo a classificação de esquemas imagéticos de Clausner e Croft (1999), os seguintes esquemas serão abordados nesse apêndice:

- 1) Contenção;
- 2) Verticalidade;
- 3) Contato;
- 4) Frente-Trás;
- 5) Esquerda-Direita.

## A.1 CONTENÇÃO

A ideia de Contenção está relacionada à noção topológica de fechamento de marco. Essa noção também possui relação com outros conceitos relacionados com a estrutura dos contentores, tais como limite e complemento, os quais permitem perceber as áreas interna e externa do marco e, conseqüentemente, é possível determinar a localização do trajetor em relação ao marco. As definições mostradas nesta seção são baseadas nos seguintes trabalhos: Neves (2011), Castilho (2014), Cuyckens (1993), Hawkins (1993), Herskovits (1987), Hottenroth (1993), Tyler e Evans (2003), Vandeloise (1991), Zelinsky-Wibbelt (1993), Coventry e Garrod (2004) e Pontes (1992).

### A.1.1 Inclusões Total e Parcial

A diferença entre inclusões total e parcial depende de itens lexicais e do conhecimento prévio do mundo. Para um melhor entendimento, veja os exemplos a seguir<sup>1</sup>:

**Exemplo 8.** “(...) depois de a água da jarra ter alcançado a temperatura ambiente, a água **na** garrafa térmica ainda estará gelada.”

**Exemplo 9.** “Entre as maiores variações da semana estão a sardinha **em** lata.”

Nos Exemplos 8 e 9 é possível extrair a noção de inclusão total pois envolvem trajetores (água, sardinha) em marcos contentores, ou seja, objetos tridimensionais fechados (garrafa térmica, lata) (OLIVEIRA, 2009).

<sup>1</sup>Todos os exemplos foram retirados do Corpus do Português (DAVIES; FERREIRA, 2016)

Já nos Exemplos 10 e 11, os marcos são objetos tridimensionais abertos (vasilha, pia), e os trajetores (leite, pratos) estão contidos nos marcos. Portanto, a relação de inclusão total pode ser observada tanto em contentores fechados como em abertos.

**Exemplo 10.** “(...) e os outros 25 quando coltar da estação depois de botar o leite **na** vasilha.”

**Exemplo 11.** “(...) e só aí vê pontas de cigarro no cinzeiro, pratos sujos **na** pia (...)”

**Exemplo 12.** “pelo menos dez homens chegaram ao local **na** caçamba do caminhão.”

No Exemplo 12, uma das interpretações possíveis é que os homens (trajetor) estão em pé na caçamba (marco). Com isso, há partes do trajetor que estão localizadas fora do marco, constituindo uma relação de inclusão parcial entre homens e caçamba. De acordo com Oliveira (2009), a distinção entre inclusão parcial e total é informada explícita ou implicitamente pelo contexto, o que inclui propriedade espaciais e inferências não codificadas sobre a situação. Isto pode ser visto no Exemplo 13, onde o tamanho do trajetor (“Senhora da Conceição”) determina se há inclusão parcial ou total. Caso seja pequeno, estará totalmente contido no marco (“sacola de lona”). Entretanto, é possível que a imagem da santa seja maior que a sacola, ficando assim uma parte fora ou, em outras palavras, parcialmente dentro da sacola (inclusão parcial).

**Exemplo 13.** “Era sem dúvida Pedro Sacristão, que vinha buscar a sua Virgem, a Senhora da Conceição. Delfino enfiou a imagem **na** sacola de lona.”

### A.1.2 Inclusão em um meio

Uma outra abordagem para a relação de inclusão é dada quando o marco inclui totalmente o trajetor e possui a característica de um meio e não de um objeto contentor. No Exemplo 14, apesar de haver uma relação de inclusão, onde o trajetor (“cálcio”) está contido em um meio (“sangue”), não é possível substituir diretamente “em” por “dentro de” sem prejuízo de significado.

**Exemplo 14.** “(...) diminuindo a concentração de cálcio **no** sangue.

### A.1.3 Inclusão trajetor vazio

Neste tipo de inclusão, o trajetor é uma entidade concreta negativa, isto é, uma ausência de matéria envolvida pelo marco. No Exemplo 15, o trajetor (“furo”) está incluído no marco (“pneu”), porém não é “palpável”, tornando-se assim um trajetor vazio. Nesse tipo de inclusão a preposição “em” não possui sinônimo.

**Exemplo 15.** “Moreno já havia passado reto numa curva no início da prova por um furo **no** pneu traseiro esquerdo.”

#### A.1.4 Outros tipos de inclusão

- 1) Inclusão em marco não contentor: no Exemplo 16, o trajetor (“hotel”) está incluído em um marco (“montanha”) que não é configurado como um contentor.

**Exemplo 16.** “Hotel construído **no meio de** uma montanha surpreende visitantes no Chile.”

- 2) Marco com lados abertos: nesse tipo de inclusão o marco não possui seus lados delimitados, como tubos ou valas (Exemplo 17).

**Exemplo 17.** (...) para recolher os fósseis expostos **numa** vala na base duma colina baixa.”

- 3) Inclusão sem controle: nesse caso o marco tem um papel de cobertura e não exerce um controle sobre o trajetor.

**Exemplo 18.** “O queijo estava **dentro da** queijeira.”

- 4) Inclusão em um limite: o marco funciona como um delimitador para o trajetor. No Exemplo 19, o trajetor (“indivíduos”) está delimitado pelo marco (“fronteiras”).

**Exemplo 19.** “**Entre** as fronteiras da cidade convivem indivíduos de diferentes camadas sociais.”

- 5) Trajetor como objeto perfurante: neste caso o trajetor perfura ou atravessa o marco (Exemplo 20).

**Exemplo 20.** “Ele quer espetar um prego **na** parede para pendurar um quadro”.

- 6) Marco com limites vagos: nesse tipo de inclusão o trajetor está caracterizado como incluído no marco mesmo estando fora dos limites geométricos. No Exemplo 11, pode ser que haja algum prato da pilha que esteja fora dos limites marco (“pia”) mas está caracterizado como dentro da pia.

- 7) Marco como anel: a borda do marco pode ser reduzida à noção de um anel, estando o trajetor incluído nessa borda.

**Exemplo 21.** “O inseto se debatia **na** pinça.”

- 8) Marco quantitativo: quando a dimensão do marco não é percebida com clareza ou não é necessária para compreender um cenário. No Exemplo 22, o observador dificilmente contextualizará o marco (“estrada”) através da sua topologia.

**Exemplo 22.** “Criança perdida **na** estrada é resgatada por polícia.”

- 9) Marco como superfície: nesse tipo de inclusão, o marco possui os lados levemente levantados, comportando-se como uma superfície ou um contentor. No Exemplo 23 o marco (“prato”) é caracterizado como uma superfície.

**Exemplo 23.** “Estevão pôs a comida **sobre** o prato.”

A Tabela 7 mostra a relação entre os tipos de inclusão e os termos usados para descrevê-los.

Tabela 7 – Tipos de contenção e expressões relacionadas.

Tipo de contenção	Expressões
Inclusão total	em, dentro de
Inclusão parcial	em, dentro de
Inclusão em um meio	em
Inclusão trajetor vazio	em
Inclusão em marco não contentor	em, dentro de, em meio de, em meio a, sob
Marco com lados abertos	dentro de, em
Inclusão sem controle	em, dentro de, embaixo de
Inclusão em um limite	dentro de, entre, cercado por
Trajedor como objeto perfurante	em
Marco com limites vagos	em, dentro de
Marco como anel	em
Marco quantitativo	em, dentro de, em meio de
Marco como superfície	em, dentro de, sobre
Exclusão	fora de, em torno de

Fonte: Autor

## A.2 VERTICALIDADE

Baseado nos trabalhos de Neves (2011), Castilho (2014), Herskovits (1987), Brenda (2014), Tyler e Evans (2001) e Abella e Kender (1993) foram encontradas sete relações para o esquema de verticalidade, que são listadas a seguir.

- 1) Marco como suporte: o trajetor é suportado pelo marco se o peso do trajetor pressiona ou empurra o objeto marco. Há dois casos a serem considerados: o primeiro quando o trajetor está diretamente em contato com o marco. No Exemplo 24, o trajetor (“pote”) está localizado numa posição superior e em contato diretamente com o marco (“mesa”). O segundo caso ocorre quando há um objeto localizado entre o trajetor e o marco (suporte indireto). No Exemplo 25, o trajetor (“flores”) podem estar dentro de um vaso e, portanto, não estão em contato direto com o marco (“mesa”).

**Exemplo 24.** “Colocou o pote **sobre** a mesa.”

**Exemplo 25.** “As flores foram colocadas **em cima da** mesa.”

- 2) Marco como superfície: em geral, o marco é uma superfície externa, horizontal e voltada para cima. No Exemplo 26, o marco (“piso”) não é um objeto em si, mas uma linha contínua.

**Exemplo 26.** “(...) o sofá, posto **sobre** o piso, estava danificado.”

- 3) Trajetor acima do marco: no Exemplo 27, o trajetor (“pássaros”) está acima do marco (“cabeças”) mas não há uma relação de suporte entre eles.

**Exemplo 27.** “Os pássaros voam **sobre** nossas cabeças.”

- 4) Trajetor como suporte: no Exemplo 28 o marco (“caixinha”) pressiona o trajetor (“livro”) estabelecendo uma relação de suporte.

**Exemplo 28.** “O livro de matemática estava **embaixo da** caixinha.”

- 5) Trajetor abaixo do marco: quando o trajetor está numa posição inferior ao marco e não há uma relação de suporte entre trajetor e marco. No Exemplo 29 o trajetor (“gato”) está posicionado abaixo do marco (“cadeira”), mas o marco não exerce a relação de suporte sobre o trajetor. Um outro caso ocorre quando o trajetor está posicionado embaixo do marco, havendo uma relação de suporte entre o marco e o trajetor.

**Exemplo 29.** “O gato escondeu-se **sob** a cadeira.”

A Tabela 8 apresenta as relações de verticalidade e as expressões utilizadas para descrevê-las.

Tabela 8 – Relações de verticalidade.

Relações	Expressões
Marco como suporte (direto)	em, sobre, em cima de
Marco como suporte (índireto)	em, sobre, em cima de
Marco como superfície	em, sobre
Trajedor acima do marco	acima de, sobre
Trajedor como suporte	sob, embaixo de, debaixo de
Trajedor abaixo do marco	sob, embaixo de, debaixo de
Trajedor debaixo do marco com suporte	sob, embaixo de, debaixo de, em

Fonte: Autor.

### A.3 CONTATO

Nesta seção será mostrado as relações de contato que não implicam em contenção ou verticalidade. Nessas relações, os marcos são entidades planas (ou conceitualizadas como planas) que oferecem algum tipo de suporte aos trajetores. A Tabela 9 apresenta as relações de contato e suas respectivas expressões espaciais.

- 1) Aderência: segundo Levinson e Meira (2003), a relação de aderência é uma variante mais específica do esquema de contato, onde o trajetor está situado na frente do marco, mas implicando em suporte entre marco e trajetor. No Exemplo 30, o trajetor (“armas”) está localizado à frente do marco (“parede”), e este exerce o papel de suporte sobre o trajetor.

**Exemplo 30.** “Meu pai tinha várias armas **na** parede.”

- 2) Cobertura: de acordo com Bennett e Cialone (2014), na relação de cobertura o marco é totalmente ou parcialmente encoberto pelo trajetor. No Exemplo 31, o trajetor (“paletó”) cobre parcialmente o marco (“camisa”), deixando uma parte deste visível. Já no Exemplo 32, o trajetor (“lençol”) cobre totalmente o marco (“colchão”).

**Exemplo 31.** “ O paletó **sobre** a camisa.”

**Exemplo 32.** “O lençol colocado **sobre** o colchão”.

Tabela 9 – Relações de contato.

Relações de contato	Expressões
Aderência	em, sobre, junto a
Cobertura total	sobre, sob, por cima, por dentro, por baixo
Cobertura parcial	sobre, sob, por cima, por dentro, por baixo

Fonte: Autor.

#### A.4 ESQUEMA DIREITA-ESQUERDA E FRENTE-TRÁS

Nesta seção serão apresentadas as expressões espaciais para os esquemas imagéticos de lateralidade (esquerda-direita) e frontalidade (frente-atrás). Essas expressões têm diferente formas de interpretação: por exemplo, quando se diz que “um objeto *X* está à esquerda de um outro objeto *Y*”, o ouvinte pode interpretar essa sentença de três formas:

- 1) à minha esquerda;
- 2) à sua esquerda (referindo-se ao falante);
- 3) o lado esquerdo do objeto *Y*.

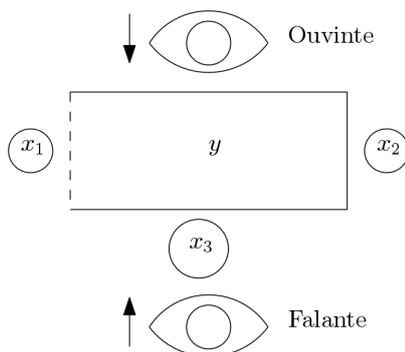
A Figura 13 ilustra o exemplo acima. Assim, uma expressão espacial do eixo lateral ou frontal (expressão projetiva) pode ser afetada por um fenômeno linguístico denominado ambiguidade pragmática. Isso ocorre pois as relações de orientação são derivadas de modelos mentais e quadros de referência que são utilizados para determinar o lado frontal de um objeto.

##### A.4.1 Quadros de referência

De acordo com Retz-Schmidt (1988), um quadro de referência para orientação determina o lado frontal de um objeto. Há três tipos básicos de quadros de referência: intrínseco, extrínseco e deítico.

- 1) Intrínseco: na perspectiva intrínseca, as regiões em torno do objeto marco (frente, atrás, direita e esquerda) são determinadas somente pelo tipo ou propriedades do

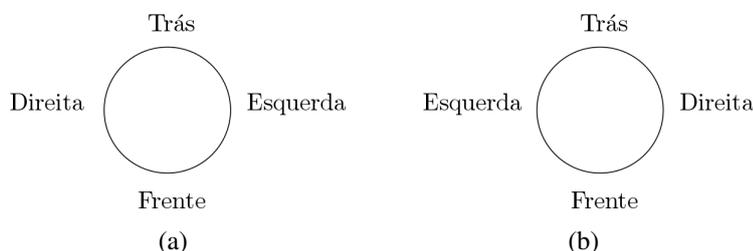
Figura 13 – Possíveis interpretações para “à esquerda”.



Fonte: Autor.

marco. Retz-Schmidt (1988) sugere critérios típicos para identificação da frente intrínseca: o lado situado na direção do movimento, o lado contendo um aparato perceptual ou o lado caracteristicamente orientado pelo observador. Uma vez que a frente intrínseca do marco é determinada, os outros lados podem ser atribuídos de duas formas: o marco pode ser visto do lado interno (Figura 14(a)) ou do lado externo (Figura 14(b)).

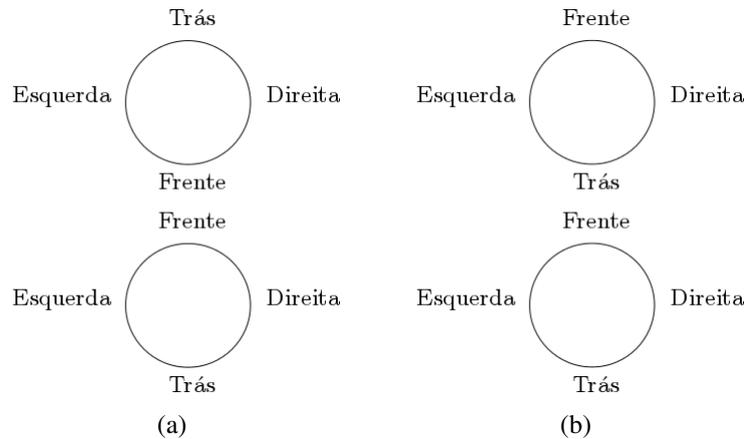
Figura 14 – Atribuições de frente intrínseca.



Adaptado de Retz-Schmidt (1988) e Herskovits (1987).

- 2) Extrínseco: quando um objeto não possui uma frente intrínseca ele recebe uma frente provisória. Vandeloise (1991) diz que um objeto não intrinsecamente orientado pode adquirir uma frente provisória através de outros objetos na vizinhança. Wunderlich (1985) argumenta que a frente provisória depende da acessibilidade: a frente de um objeto extrínseco é o lado que está acessível mais rapidamente. Teixeira (2001) cita um modelo mental de atribuição de frente denominado espelhamento. Nesse modelo, um objeto adquire uma frente provisória através do processo de espelhamento (Figura 15(a)). Ademais, baseado em experimentos cognitivos, Teixeira (2001) diz que o falante tende a identificar o lado esquerdo como o lado de trás e o lado direito como frente. Isso ocorre devido à organização discursiva (da esquerda para a direita).
- 3) Défítico: quando a orientação é dada pelo ponto de vista do observador. Retz-Schmidt (1988) diz que na perspectiva défística o quadro de referência pode ser

Figura 15 – Atribuições de frente extrínseca.

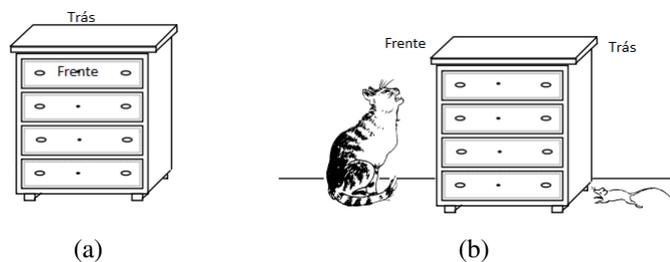


Adaptado de Retz-Schmidt (1988).

orientado de duas maneiras: a primeira por espelhamento e na segunda o marco pode ser visto pelo ponto de vista do falante, onde o lado de trás está mais próximo e a frente é o lado mais distante do marco. Essa forma de atribuição é conhecida como princípio *Tandem* (Figura 15(b)) (RETZ-SCHMIDT, 1988; WUNDERLICH, 1985). A atribuição de esquerda e direita e esquerda é a mesma em ambas as formas.

No modelo mental de visibilidade, a presença ou ausência de um traço de visibilidade/acessibilidade pode alterar a percepção de frente/trás e também pode cancelar orientação intrínseca de um objeto (TEIXEIRA, 2001). Na Figura 16(a), a cômoda possui uma frente intrínseca, mas na Figura 16(b), sob o ponto de vista do gato, pode-se dizer que “o rato está atrás da cômoda”.

Figura 16 – Alteração na percepção de frente intrínseca.



Fonte: Teixeira (2001).

No demais, Retz-Schmidt (1988) lista alguns papéis relacionados com a ambiguidade pragmática entre quadros de referência:

- 1) Entre usos intrínseco e deístico, onde o ponto de vista é omitido ou é intrínseco;
- 2) Entre usos deísticos (falante ou ouvinte como ponto de vista);
- 3) Entre usos intrínsecos (diferentes critérios na determinação de frente intrínseca);

- 4) Entre usos intrínsecos e extrínsecos (um objeto com frente intrínseca pode adquirir frente provisória);
- 5) Entre usos extrínsecos (diferentes orientações extrínsecas).

A Tabela 10 mostra a as expressões projetivas.

Tabela 10 – Expressões projetivas.

Lado	Expressão
Frente	diante de, em frente de, em frente a, na frente de, à frente de, defronte a
Trás	atrás de
Esquerda	à esquerda de
Direita	à direita de
Esquerda ou Direita	ao lado de, do lado de, no lado (direito/esquerdo) de

Fonte: Autor.

**APÊNDICE B – PROTO-CENAS PARA A PREPOSIÇÃO “IN”**

Neste apêndice serão listadas as imagens (proto-cenas) para a preposição “in”. (o trajetor é representado pelas figuras em vermelho, e marco pelas figuras em preto)<sup>1</sup>.

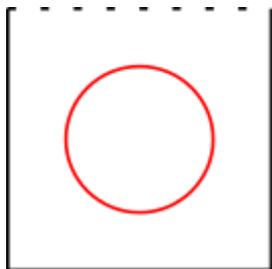


Figura 17 – Inclusão total

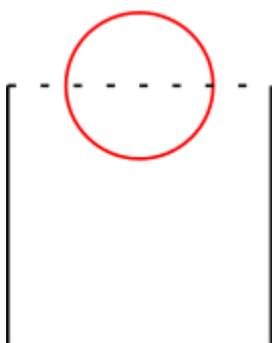


Figura 18 – Inclusão parcial

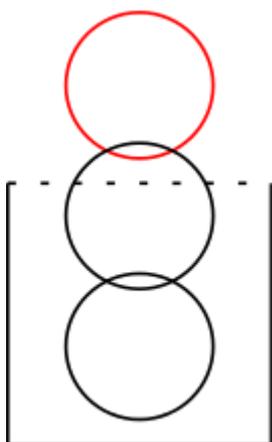


Figura 19 – Inclusão por tolerância

---

<sup>1</sup>Todas as figuras deste apêndice foram desenvolvidas pelo autor

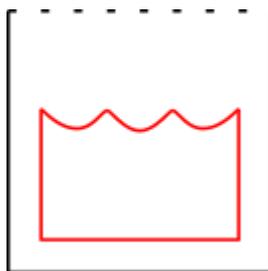


Figura 20 – Inclusão de um líquido

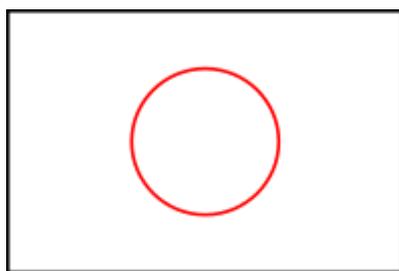


Figura 21 – Trajetor sobre uma parede

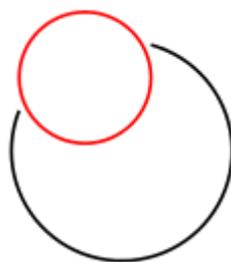


Figura 22 – Marco em forma de anel

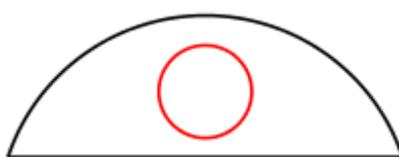


Figura 23 – Encapsulamento

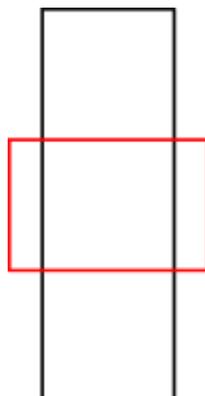


Figura 24 – Marco como extremidade do corpo-humano

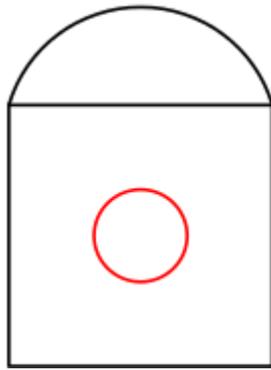


Figura 25 – Inclusão total em uma bolsa

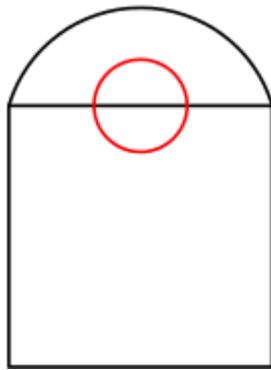


Figura 26 – Inclusão parcial em uma bolsa

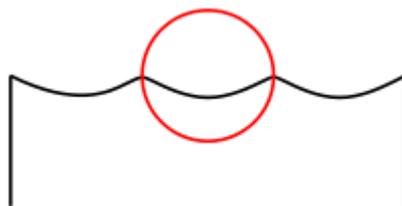


Figura 27 – Inclusão em um meio líquido

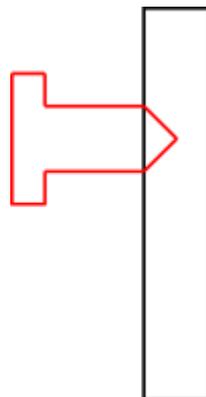


Figura 28 – Trajetor é um objeto fixador

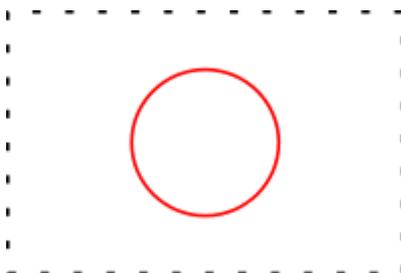


Figura 29 – Marco como suporte

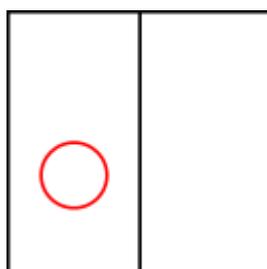


Figura 30 – Inclusão em um compartimento de armazenagem

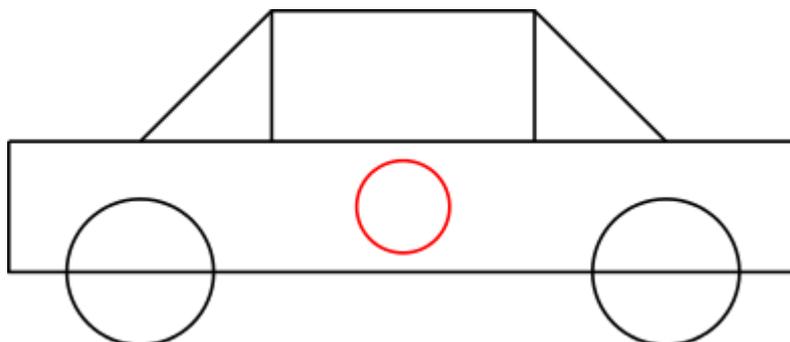


Figura 31 – Inclusão em um veículo



Figura 32 – Inclusão em um tubo

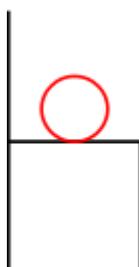


Figura 33 – Marco como assento

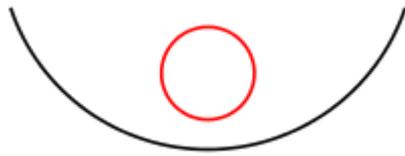


Figura 34 – Inclusão de uma vala



Figura 35 – Inclusão em um contentor aberto

## REFERÊNCIAS

- ABELLA, A.; KENDER, J. Qualitatively describing objects using spatial prepositions. In: IEEE. QUALITATIVE Vision, 1993., Proceedings of IEEE Workshop on. [S.l.: s.n.], 1993. p. 33–38.
- ARAÚJO, P. **Aspectos semântico-cognitivos de usos espaciais das preposições para e em na fala de comunidades quilombolas**. 2008. Tese (Doutorado) – Universidade de São Paulo.
- BENNETT, B. Application of supervaluation semantics to vaguely defined spatial concepts. In: SPRINGER. INTERNATIONAL Conference on Spatial Information Theory. [S.l.: s.n.], 2001. p. 108–123.
- \_\_\_\_\_. Standpoint Semantics: a framework for formalising the variable meaning of vague terms. **Understanding Vagueness-Logical, Philosophical and Linguistic Perspectives**, 2011.
- \_\_\_\_\_. What is a forest? On the vagueness of certain geographic concepts. **Topoi**, Springer, v. 20, n. 2, p. 189–201, 2001.
- BENNETT, B.; CIALONE, C. Corpus Guided Sense Cluster Analysis: a methodology for ontology development (with examples from the spatial domain). In: FOIS. [S.l.: s.n.], 2014. p. 213–226.
- BOWERMAN, M. Learning how to structure space for language: A crosslinguistic perspective. **Language and space**, MIT Press, Cambridge, p. 385–436, 1996.
- BRENDA, M. **The Cognitive Perspective on the Polysemy of the English Spatial Preposition Over**. [S.l.]: Cambridge Scholars Publishing, 2014.
- BRUNET, R. La carte-modèle et les chorèmes. **Mappemonde**, v. 86, n. 4, p. 2–6, 1986.
- CARLSON, L. Object use and object location: The effect of function on spatial relations. **Cognitive interfaces: Constraints on linking cognitive information**, Oxford: Oxford University Press, p. 94–115, 2000.
- CARNAP, R. **The logical structure of the world**. [S.l.]: Univ of California Press, 1967.
- CASTILHO, A. **Nova gramática do português brasileiro**. [S.l.]: Contexto, 2014.
- CHANG, A. et al. Text to 3d scene generation with rich lexical grounding. **arXiv preprint arXiv:1505.06289**, 2015.

CHARNIAK, E. Passing markers: A theory of contextual influence in language comprehension. **Cognitive science**, Wiley Online Library, v. 7, n. 3, p. 171–190, 1983.

CLAUSNER, T.; CROFT, W. Domains and image schemas. **Cognitive linguistics**, Mouton De Gruyter, v. 10, p. 1–32, 1999.

COBREROS, P. Supervaluationism and logical consequence: A third way. **Studia Logica**, Springer, v. 90, n. 3, p. 291–312, 2008.

COHN, A.; RANDELL, D.; CUI, Z. Taxonomies of logically defined qualitative spatial relations. **International journal of human-computer studies**, Elsevier, v. 43, n. 5-6, p. 831–846, 1995.

COHN, A. et al. Qualitative spatial representation and reasoning with the region connection calculus. **GeoInformatica**, Springer, v. 1, n. 3, p. 275–316, 1997.

COVENTRY, K. R.; GARROD, S. **Saying, seeing and acting: The psychological semantics of spatial prepositions**. [S.l.]: Psychology Press, 2004.

COYNE, B.; SPROAT, R. WordsEye: an automatic text-to-scene conversion system. In: ACM. PROCEEDINGS of the 28th annual conference on Computer graphics and interactive techniques. [S.l.: s.n.], 2001. p. 487–496.

CROFT, W.; CRUSE, A. **Cognitive linguistics**. [S.l.]: Cambridge University Press, 2004.

CUYCKENS, H. The Dutch spatial preposition—in “: a cognitive semantic analysis. **The semantics of prepositions: from mental processing to natural language processing**, v. 3, p. 27–71, 1993.

DAHL, V. Natural language processing and logic programming. **The Journal of Logic Programming**, v. 19-20, p. 681–714, 1994. Special Issue: Ten Years of Logic Programming. Disponível em: <<http://www.sciencedirect.com/science/article/pii/0743106694900361>>.

DAVIES, M.; FERREIRA, M. **O corpus do Português: 1 bilhão de palavras**. 2016. Disponível em: <[www.corpusdoportugues.org](http://www.corpusdoportugues.org)>.

DESOLNEUX, A.; MOISAN, L.; MOREL, J. Computational gestalts and perception thresholds. **Journal of Physiology-Paris**, Elsevier, v. 97, n. 2-3, p. 311–324, 2003.

DODGE, E.; LAKOFF, G. Image schemas: From linguistic analysis to neural grounding. **From perception to meaning: Image schemas in cognitive linguistics**, Mouton de Gruyter, p. 57–91, 2005.

DREYFUS, H. What computers can't do, 1972.

EVANS, V. **Cognitive linguistics**. [S.l.]: Edinburgh University Press, 2006.

FASS, D. Collative semantics. In: ASSOCIATION FOR COMPUTATIONAL LINGUISTICS. PROCEEDINGS of the 11th conference on Computational linguistics. [S.l.: s.n.], 1986. p. 341–343.

FINE, K. Vagueness, truth and logic. **Synthese**, Springer, v. 30, n. 3-4, p. 265–300, 1975.

FRANK, A. Qualitative spatial reasoning: Cardinal directions as an example. **International Journal of Geographical Information Science**, Taylor & Francis, v. 10, n. 3, p. 269–290, 1996.

FROMKIN, V.; RODMAN, R.; HYAMS, N. **An introduction to language**. [S.l.]: Cengage Learning, 2018.

GARCIA, D. **PrepNet.Br: uma proposta de representação semântica para as preposições do português**. 2018. Tese (Doutorado) – UFSCar, São Carlos.

HAMPE, B.; GRADY, J. **From perception to meaning: Image schemas in cognitive linguistics**. [S.l.]: Walter de Gruyter, 2005. v. 29.

HAWKINS, B. On universality and variability in the semantics of spatial adpositions. **The semantics of prepositions**, Walter de Gruyter, v. 3, p. 327, 1993.

HAYES, P. Naive physics: ontology for liquids (working paper 35). **Institute of Semantic and Cognitive Studies, Geneva**, 1978.

HERSKOVITS, A. **Language and spatial cognition**. [S.l.]: Cambridge University Press, 1987.

\_\_\_\_\_. Language, spatial cognition, and vision. In: SPATIAL and Temporal Reasoning. [S.l.]: Springer, 1997. p. 155–202.

HIRST, G. **Semantic interpretation and the resolution of ambiguity**. [S.l.]: Cambridge University Press, 1992.

HOTTENROTH, P. Prepositions and object concepts: A contribution to cognitive semantics. **The semantics of prepositions**, p. 179–220, 1993.

ILARI, R. et al. As preposições. In: GRAMÁTICA do português culto falado do Brasil, Vol. II: Classes de palavras e processos em construção. Campinas: Editora Unicamp, 2008. p. 623–808.

JACKENDOFF, R. **Semantics and cognition**. [S.l.]: MIT press, 1983. v. 8.

- JOHANSSON, R. et al. Automatic text-to-scene conversion in the traffic accident domain. In: IJCAI. [S.l.: s.n.], 2005. v. 5, p. 1073–1078.
- JOHNSON, M. **The body in the mind: The bodily basis of reason and imagination**. Chicago: University of Chicago Press, 1987.
- KANG, Y. **Cognitive linguistics approach to semantics of spatial relations in Korean**. 2012. Tese (Doutorado) – Georgetown University.
- KEEFE, R. Vagueness: supervaluationism. **Philosophy Compass**, Wiley Online Library, v. 3, n. 2, p. 315–324, 2008.
- KNAŚ, I. Polish equivalents of spatial at. In: ASSOCIATION FOR COMPUTATIONAL LINGUISTICS. PROCEEDINGS of the Third ACL-SIGSEM Workshop on Prepositions. [S.l.: s.n.], 2006. p. 9–16.
- KRISHNASWAMY, N.; PUSTEJOVSKY, J. Multimodal semantic simulations of linguistically underspecified motion events. In: SPATIAL Cognition X. [S.l.]: Springer, 2016. p. 177–197.
- KURZWEIL, Ray. **How to create a mind**. [S.l.]: Penguin.[RM], 2012.
- LAKOFF, G. **Women, fire, and dangerous things: What categories reveal about the mind**. Chicago: Chicago University Press, 1989.
- LAKOFF, G.; JOHNSON, M. **Metaphors we live by**. [S.l.]: University of Chicago press, 2008.
- \_\_\_\_\_. **Philosophy in the Flesh**. New York: Basic Books, 1999. v. 4.
- LANDAU, B.; JACKENDOFF, R. Whence and whither in spatial language and spatial cognition? **Behavioral and brain sciences**, Cambridge University Press, v. 16, n. 2, p. 255–265, 1993.
- LANGACKER, R. **Foundations of cognitive grammar: Theoretical prerequisites**. [S.l.]: Stanford university press, 1987. v. 1.
- LANGACKER, R.; MUNRO, P. Passives and their meaning. **Language**, JSTOR, p. 789–830, 1975.
- LEVINSON, S.; MEIRA, S. “Natural Concepts” in the Spatial Topological Domain-Adpositional Meanings in Crosslinguistic Perspective: An Exercise in Semantic Typology. **Language**, Linguistic Society of America, v. 79, n. 3, p. 485–516, 2003.
- LIAMKINA, O. Semantic structure of the German spatial particle über. **Journal of Germanic Linguistics**, Cambridge University Press, v. 19, n. 2, p. 115–160, 2007.

- LIGOZAT, G. **Qualitative Spatial and Temporal Reasoning**. [S.l.]: Wiley, 2013. (ISTE).
- LIGOZAT, G.; NOWAK, J.; SCHMITT, D. From language to pictorial representations. In: LANGUAGE and Technology Conference (L&TC'07). [S.l.: s.n.], 2007.
- LIPKA, L. **An outline of English lexicology: lexical structure, word semantics, and word-formation**. [S.l.]: Walter de Gruyter, 1990. v. 3.
- MAHPEYKAR, N.s; TYLER, A. The semantics of Farsi be: applying the principled polysemy model. In: SPRINGER. INTERNATIONAL Conference on Spatial Information Theory. [S.l.: s.n.], 2011. p. 413–433.
- MARQUES, M. **Iniciação à semântica**. 7. ed. Rio de Janeiro: Zahar, 2009.
- MEHLBERG, H. The reach of science. Yniversity of Toronto Press, 1958.
- MORATZ, Reinhard; RAGNI, Marco. Qualitative spatial reasoning about relative point position. **Journal of Visual Languages & Computing**, Elsevier, v. 19, n. 1, p. 75–98, 2008.
- MORAVCSIK, J. How do words get their meanings? **The journal of philosophy**, v. 78, n. 1, p. 5–24, 1981.
- NELSON, G. The Structure of Appearance. **Reidel Pub. Comp., Dordrecht and Boston**, 1951.
- NERLICH, B. Polysemy: past and present. **Polysemy: Flexible patterns of meaning in mind and language**, Walter de Gruyter, v. 142, p. 51–76, 2003.
- NEVES, M. **Gramática de usos do português**. 2. ed. São Paulo: Editora Unesp, 2011.
- NIDA, E. **A Componential Analysis of Meaning: An Introduction to Semantic Structures**. [S.l.]: Walter de Gruyter, 1979. v. 57.
- O'GRADY, W.; DOBROVOLSKY, M.; KATAMBA, F. **Contemporary linguistics: An introduction**. 6. ed. Boston: St. Martin's, 2010.
- OLIVEIRA, A. **Relações semântico-cognitivas no uso da preposição em no português do Brasil**. 2009. Tese (Doutorado) – UFMG, Belo Horizonte.
- PONTES, E. **Espaço e tempo na língua portuguesa**. Campinas: Pontes, 1992.
- PUSTEJOVSKY, J.; KRISHNASWAMY, N. VoxML: A visualization modeling language. **arXiv preprint arXiv:1610.01508**, 2016.
- RANDELL, D.; CUI, Z.; COHN, A. A spatial logic based on regions and connection. **KR**, v. 92, p. 165–176, 1992.

RETZ-SCHMIDT, G. Various views on spatial prepositions. **AI magazine**, v. 9, n. 2, p. 95, 1988.

RICE, S. Polysemy and lexical representation: The case of three English prepositions. In: PROCEEDINGS of the fourteenth annual conference of the cognitive science society. [S.l.: s.n.], 1992. v. 8994.

RIEGER, C.; SMALL, S. Word expert parsing. In: MORGAN KAUFMANN PUBLISHERS INC. PROCEEDINGS of the 6th international joint conference on Artificial intelligence-Volume 2. [S.l.: s.n.], 1979. p. 723–728.

ROSCH, E. Human categorization. **Studies in cross-cultural psychology**, v. 1, p. 1–49, 1977.

SANTOS, P.; BENNETT, B.; SAKELLARIOU, G. Supervaluation semantics for an inland water feature ontology. In: INTERNATIONAL Joint Conference on Artificial Intelligence. San Francisco: Morgan Kaufmann, 2005. v. 19, p. 564–569.

SCHLICK, M. Allgemeine erkenntnislehre. In: ALLGEMEINE Erkenntnislehre. [S.l.]: Springer, 2009. p. 121–822.

SEARLE, J. Literal meaning. **Erkenntnis**, Springer, v. 13, n. 1, p. 207–224, 1978.

SHAKOVA, D.; TYLER, A. 11 Taking the Principled Polysemy Model of spatial particles beyond English: the case of Russian za. **Equinox Publishing**, p. 267–292, 2010.

SKIADOPOULOS, S.; KOUBARAKIS, M. Composing cardinal direction relations. **Artificial Intelligence**, Elsevier, v. 152, n. 2, p. 143–171, 2004.

SOTO, A.; OLIVAS, J. A.; PRIETO, M. Fuzzy approach of synonymy and polysemy for information retrieval. In: GRANULAR Computing: At the Junction of Rough Sets and Fuzzy Sets. [S.l.]: Springer, 2008. p. 179–198.

TALMY, L. Figure and ground in complex sentences. In: ANNUAL Meeting of the Berkeley Linguistics Society. [S.l.: s.n.], 1975. v. 1, p. 419–430.

\_\_\_\_\_. How language structures space. In: SPATIAL orientation. [S.l.]: Springer, 1983. p. 225–282.

\_\_\_\_\_. **Toward a cognitive semantics**. [S.l.]: MIT Press, 2000. v. 2.

TARSKI, A. The semantic conception of truth: and the foundations of semantics. **Philosophy and phenomenological research**, JSTOR, v. 4, n. 3, p. 341–376, 1944.

TCHIZMAROVA, I. A Cognitive Linguistic Analysis of the Bulgarian Verbal Prefix pre-, through, over'. **Glossos**, v. 7, p. 1–49, 2006.

TEIXEIRA, J. **A Verbalização do Espaço: modelos mentais de frente/trás**. [S.l.]: Universidade do Minho. Centro de Estudos Humanísticos (CEHUM), 2001.

TRAUGOTT, E. On the rise of epistemic meanings in English: An example of subjectification in semantic change. **Language**, JSTOR, p. 31–55, 1989.

TYLER, A.; EVANS, V. Reconsidering prepositional polysemy networks: The case of over. **Language**, Linguistic Society of America, v. 77, n. 4, p. 724–765, 2001.

\_\_\_\_\_. **The semantics of English prepositions: Spatial scenes, embodied meaning, and cognition**. [S.l.]: Cambridge University Press, 2003.

VAN DER GUCHT, F.; WILLEMS, K.; DE CUYPERE, L. The iconicity of embodied meaning. Polysemy of spatial prepositions in the cognitive framework. **Language Sciences**, Elsevier, v. 29, n. 6, p. 733–754, 2007.

VAN FRAASSEN, B. Singular terms, truth-value gaps, and free logic. **The journal of Philosophy**, JSTOR, v. 63, n. 17, p. 481–495, 1966.

VANDELOISE, C. Methodology and analyses of the preposition in. **Cognitive Linguistics**, Walter de Gruyter, v. 5, n. 2, p. 157–184, 1994.

\_\_\_\_\_. **Spatial prepositions: A case study from French**. [S.l.]: University of Chicago Press, 1991.

WILKS, Y. Senses and texts. **Computers and the Humanities**, Springer, v. 31, n. 2, p. 77–90, 1997.

WINOGRAD, T. What does it mean to understand language? **Cognitive science**, Wiley Online Library, v. 4, n. 3, p. 209–241, 1980.

WINOGRAD, T.; FLORES, F. **Understanding computers and cognition: A new foundation for design**. [S.l.]: Intellect Books, 1986.

WITTGENSTEIN, L. **Philosophical investigations**. [S.l.]: John Wiley & Sons, 1953.

WUNDERLICH, D. Raum, Zeit und das Lexikon. In: **SPRACHE und Raum**. [S.l.]: Springer, 1985. p. 66–89.

ZELINSKY-WIBBELT, C. Interpreting and translating prepositions: A cognitively based formalization. **The Semantics of Prepositions**, Berlin: Walter de Gruyter & Co, p. 351–390, 1993.