

# A relação de meronímia no domínio jurídico: um estudo visando sua inserção em uma ontologia jurídica

Thaís D. Minghelli

Rua João Neves da Fontoura, 716/203. Centro. São Leopoldo/RS - Brasil.

minghelli.adv@terra.com.br

**Resumo.** *O presente artigo aborda a relação de meronímia (RM) ou parte-todo para fins de representação do domínio jurídico, mais especificamente do Direito Processual Penal (DPP). Tendo em vista a grande demanda por sistemas cada vez melhores no que tange à recuperação da informação, o grande número de sites jurídicos, bem como a relevância social em aclarar a linguagem jurídica para o leigo, ontologias têm se revelado úteis neste sentido. Por isso, objetiva-se aqui refletir de que forma os conceitos do DPP podem ser explicitados sob a luz da relação parte-todo, visando inserir os resultados da pesquisa em uma ontologia.*

## 1. Introdução

Este artigo apresenta o andamento da pesquisa de mestrado, a qual aborda a RM no domínio jurídico, mais especificamente, no DPP para fins de uma futura inclusão em uma ontologia jurídica linguística. A escolha desta relação semântica para a representação do DPP deve-se ao fato deste domínio ser constituído por um conhecimento segmentado em etapas, em que a RM se mostra recorrente, como também de que referida relação em uma ontologia linguística possibilita a associação pela máquina, por exemplo, das partes/fases do DPP a um todo/atividade tal como um jurista é capaz de fazer.

Neste sentido, a máquina é capaz de identificar que as ações, correspondentes a unidades e não a tipos, frisa-se, <depoimento do ofendido>, <inquirição de testemunhas da acusação e da defesa>, <depoimento dos peritos>, <acareação dos envolvidos> etc., *compõem* um todo, uma atividade, qual seja: a <audiência de instrução e julgamento>. Ontologias permitem a recuperação da informação com maior eficiência, bem como a unificação de ordenamentos jurídicos e a desburocratização, visto que elas ensejam o mapeamento de conceitos, o raciocínio automático e o processamento da linguagem natural (PLN). Na União Européia ontologias jurídicas vêm sendo construídas, visando a comparação e unificação da legislação, como também a recuperação da informação. A *Core Legal Ontology* (CLO), a LRI-Core e LOIS são exemplos de iniciativas neste sentido.

Tendo em vista referidos movimentos e a demanda cada vez maior de aperfeiçoar sistemas de busca em *sites* jurídicos, objetiva-se construir uma ontologia jurídica linguística. Para tanto, apresentam-se os seguintes objetivos: (i) apresentar os *corpora*; (ii) discorrer sobre considerações semânticas acerca da RM, trazendo exemplos jurídicos; (iii) definir e distinguir ontologias, léxicos computacionais e ontologias linguísticas e (iv) mencionar a metodologia e a análise dos tipos de merônimos mais pertinentes para o DPP, objetivando-se verificar em quais dos *corpora*

podem ser encontrados exemplos mais produtivos por meio da utilização de marcadores lingüísticos (ML) no uso da ferramenta *corpógrafo*.

## 2. Noções prévias

### 2.1. Os corpora

Os *corpora* correspondem: (i) Ao código de Processo Penal (CPP), ou seja, à norma jurídica aplicada na existência de um processo criminal, em que estão registradas todas as etapas que instruem o desenrolar de uma ação penal, bem como os possíveis participantes do contexto penal e processual penal, as instituições e os documentos. O CPP é composto por artigos 811 artigos, organizados em títulos e capítulos, podendo conter incisos, alíneas e letras. (ii) A dez acórdãos do Tribunal de Justiça do Rio Grande do Sul, coletados por meio de pesquisa orgânica sob as palavras-chaves: <homicídio qualificado apelação>.

### 2.2. Relação de meronímia ou parte-todo: considerações semânticas

O léxico pode ser estudado sob diferentes perspectivas. Uma delas é a análise por meio de suas relações de sentido com outras palavras, estruturando e regularizando o léxico de uma língua. Dentre os diferentes tipos de relações lexicais, há as ordenadoras do léxico em superordenados/subordinados, em parte-todo e de forma associativa (Saeed, 1997, p.63). Aqui, serão abordadas as relações parte-todo, as quais a priori são noções ontológicas pertencentes à disciplina de Mereologia.

Tendo em vista a interface Linguística do tema, traz-se autores desta linha como Cruse (2000), o qual entende que a meronímia corresponde a uma relação ou conexão entre duas entidades de uma mesma natureza ontológica, aludindo a ideia de inclusão entre dois elementos mutuamente implicados como em X está implicado no sentido de Y, podendo estar dispostos hierarquicamente. Saeed (1997, p. 70) refere que meronímia é um termo usado para descrever relações parte-todo entre dois termos lexicais, em que as partes ou merônimos são unidades de *holônimos*. Conhecida como uma *relação partitiva* é representada por ML, tais como: *parte de, contém, tem, possui*, sendo mais encontrada entre os nominais.

Estudando a relação parte-todo entre pares de unidades lexicais, percebem-se características peculiares a esta relação semântica. Nota-se que algumas partes podem ser mais essenciais diante do todo, tal como em *peça acusatória e processo*. Outras, apesar de comuns, não são obrigatórias, sendo tidas como facultativas como em *rol de testemunhas*, já que uma ação penal pode ser oferecida ainda que sem a menção de testemunhas do fato criminoso. Da mesma forma, a RM pode ser transitiva ou intransitiva, isto é, quando não há o mesmo tipo de RM em todas as partes de uma hierarquia, configurando uma assimetria. É transitiva quando há uma correspondência de *acarretamento* entre as categorias de uma hierarquia. No plano jurídico, exemplifica-se a transitividade: *A votação no Tribunal do Júri (evento jurídico) contém leitura dos quesitos. A leitura dos quesitos contém uma ordem. A votação tem uma ordem*. Lyons (1977) verifica que nem sempre a RM é transitiva, sendo a intransitividade gerada pelas variedades de relações parte-todo existentes e presentes em um mesmo silogismo meronímico.

Caracterizando a RM, faz-se a revisão teórica com base em Cruse (2000, p. 154-5). A primeira característica é a **necessidade** de algumas partes ao *holônimo* para sua

adequada formação, configurando uma relação canônica. Em oposição à *necessidade*, há a **opcionalidade** da parte diante do todo (Cruse, 1986, p. 162-3), quando a parte é facultativa na relação. O segundo traço da RM é a **integralidade** da parte ao todo. Detecção possível de ser feita pela descrição de como a parte é presa ao todo, sendo que quanto mais difícil esta descrição for, mais integrado parte-todo estarão. No domínio jurídico ilustra-se com o objeto físico *processo*, (*autos* como é chamado tecnicamente) e seus volumes. A **motivação** é outro traço que pode ser identificado na RM, implicando na *função* da parte e permitindo que o todo funcione conforme o fim proposto e inclusive que a parte seja identificada mais facilmente. A **congruência** ou harmonia entre parte e todo pode originar dois fenômenos: a **supermeronímia** e a **semimeronímia**. O primeiro corresponde às partes que são aplicadas a mais de um *holônimo*, como a <apresentação de documentos> é parte da <resposta do réu>, como também é parte da <peça acusatória>. O segundo trata da sobreposição da parte no todo.

No que tange aos tipos de merônimos, Winston et. al. (1987) trazem seis categorias. (i) **Componente-objeto-integral**: relação típica parte-todo, em que o todo inclui a parte, sendo constituído por vários componentes discretos, com limites e funções definidas em relação ao todo. (ii) **Membro-coleção**: parte como um objeto singular e membro de um todo que denota um conjunto. (iii) **Material-objeto**: parte como elemento constitutivo de um objeto concreto, ou seja, à conexão entre a matéria de que é feita e uma entidade concreta. (iv) **Porção-massa**: parte indefinida diante do todo, o qual configura uma entidade mais concreta. (v) **Lugar-área**: parte é o lugar localizado no todo ou área. (vi) **Ação-atividade**: parte como fase de uma atividade/todo, descrevendo diferentes subatividades que compõem uma maior estruturada e cronologicamente organizada. Antecipa-se que dentre aludidas categorias, as mais produtivas para a representação do DPP são a (i), (ii), (v) e (vi), reestudadas quando da análise deste artigo.

### 2.3. Ontologias, léxicos computacionais e ontologias lingüísticas

O termo ontologia tem sua origem na Filosofia, como o estudo das categorias que compõem o mundo, visando uma classificação universal. Sob o viés tecnológico, as ontologias se diferem da concepção filosófica, apesar de se relacionarem com ela, de certa maneira, já que ambas trabalham com representações de *mundos*. Ontologias ressurgem no contexto da web semântica, criada por Berners Lee, tendo em vista que o conhecimento organizado e relacionado por meio de relações semânticas facilita a recuperação da informação. Ontologias permitem a indexação de textos com maior precisão, permitindo melhores resultados quando da busca virtual. Elas têm o objetivo prático de oferecer estruturas de conhecimento para os sistemas computacionais, possibilitando a resolução de problemas de conhecimento de mundo relacionados ao PLN e o raciocínio lógico automatizado. Como produto tecnológico não buscam a verdade, mas o que existe em um domínio, almejando o melhor funcionamento de um determinado sistema computacional.

Atualmente, ontologias, como estruturas para organização do conhecimento, têm sido estudadas pela Ciência da informação, Linguística Computacional, Inteligência Artificial (IA) com ênfase em aplicações voltadas para o PLN e Web Semântica, por exemplo, visando estruturar e descrever conceitos. Considerando este fato, percebe-se uma imprecisão terminológica e a aproximação de ontologias, léxicos e

ontologias linguísticas. Hirst (2004, apud Prévot, 2010) afirma que *ontologias e recursos lexicais são aparentemente similares o suficiente para que sejam usados de forma intercambiável e combinada*, apesar de possuírem aspectos que os distingam pontualmente. Por tal razão é pertinente apontar em que medida se aproximam e se afastam.

Entende-se que ontologias devem observar a lógica formal. Para a (IA), uma ontologia é uma especificação formal e explícita de uma conceitualização, sendo que o existente é passível de representação (Gruber, 1993). O autor vincula *conceitualização* à especificação, descrição e representação dos termos em uma linguagem de programação, *formal*. *Compartilhado* relaciona ao conhecimento consensualmente compreendido, não restrito a um indivíduo. Para sistemas de informação, uma ontologia é um artefato de engenharia, constituído por um vocabulário específico usado para descrever certa realidade, organizado hierarquicamente e inter-relacionados por relações de suposição. Tais relações descrevem o significado pretendido, considerando-se a finalidade para a qual a ontologia é construída. A hierarquia básica de uma ontologia corresponde a uma taxonomia, relações *é\_um* e outras como: de meronímia, *parte\_de*, e associativas.

Outra forma de estruturação do conhecimento distinta da ontológica, mas que em vários aspectos se aproxima desta, podendo, por vezes, serem intercambiáveis ou ocorrerem concomitantemente são léxicos computacionais. Aplicações computacionais são um recurso tecnológico em prol do PLN, cujas primeiras aplicações foram para fins de tradução automática, sendo hoje usadas para a recuperação de informação. Atualmente, léxicos estão se tornando recursos robustos com muitas informações linguísticas. WordNets equivalem a importantes bases de dados para o PLN. Por estas razões, a fronteira entre léxicos e ontologias torna-se, cada vez mais tênue. Uma grande diferença entre léxicos e ontologias é que os primeiros descrevem o significado de expressões de linguagem natural; enquanto os segundos descrevem as entidades de um domínio e as relações entre as mesmas. Grosso modo, o que se percebe é um processo de *ontologização* dos léxicos, dando margem a mais um termo afim: ontologias linguísticas. Estas se situariam entre as ontologias formais e os léxicos robustos. Elas emergem da semântica e do léxico, representam um determinado conhecimento por meio de relações semânticas, mas sem o compromisso de se valer da lógica formal. Prévot et. al. (2010) explica que a conceitualização é baseada em critérios linguísticos, mais precisamente em informações encontradas em recursos lexicais.

Neste sentido, ontologias linguísticas caracterizam-se por armazenar apenas conceitos lexicalizados, conceitos expressos por uma ou mais palavras. Sob este viés, uma ontologia linguística corresponde aos sentidos de uma dada língua, aos conceitos compartilhados por uma comunidade lingüística (DI FELIPPO, 2008). O objetivo da pesquisa de mestrado é justamente este: representar o DPP por meio da RM, visando inseri-lo em uma ontologia linguística jurídica, a qual estaria comprometida com a representação do léxico e não com uma linguagem formal.

### **3. Metodologia**

Tendo em vista os *corpora* selecionados (cf.2.2) e as considerações acerca da RM (cf.2.3) pressupõe-se que: os tipos de merônimos (i), (ii), (v) e (vi) de Winston et.al (1987) são pertinentes para a representação do DPP. Assim, objetiva-se, verificar em

quais dos *corpora* podem ser encontrados exemplos mais produtivos por meio da utilização de ML como: tem, composto, constituído, fase, peça, contém, bem como frases preposicionais com preposições como "de" ("do", "da", por exemplo) e "em" ("no", "na"), fazendo uso de concordanciador. Vale esclarecer que neste artigo trouxeram-se apenas os ML que ensejaram resultados, considerando ainda a concisão do artigo. Estabelece-se, assim, a metodologia para a realização da análise: (i) pautado nos pressupostos estudados, testaram-se os ML que representam relação de parte-todo; (ii) utilizando o *corpógrafo*, testaram-se os ML nos *corpora*; (iii) analisaram-se qual *corpus* trouxe exemplos que melhor representassem o domínio do DPP e (iv) apresentaram-se, neste artigo, somente os pares de parte-todo que trouxeram resultados de meronímia para inclusão em uma ontologia jurídica lingüística.

#### 4. Análise dos dados

Testaram-se os ML acima mencionados; todavia poucos se mostraram produtivos em ambos os *corpora*. Em decorrência do limite de páginas, mencionaremos somente um exemplo de cada retorno pertinente para a representação da RM no DPP, procurando um enquadramento nas tipologias de Winston et. al. (1987) retro-referidas:

**ML composto:** recuperou 01 exemplo apenas no *corpus* do CPP e nenhum no conjunto de 10 acórdãos. Ex. O Tribunal do Júri é **composto** por 1 (um) juiz togado, seu presidente e por 25 (vinte e cinco) jurados que serão sorteados dentre os alistados, 7 (sete) dos quais constituirão o Conselho de Sentença em cada sessão de julgamento.

**ML fase:** recuperou 11 concordâncias do CPP e 3 dos acórdãos, sendo que todos expressam o mesmo tipo de RM. Os retornos obtidos ilustram a categoria ação-atividade, pois aludem às subatividades de um evento maior. Ex. O exame poderá ser ordenado na do inquérito, mediante representação da autoridade policial ao juiz competente.

**ML conterà:** recuperou 11 concordâncias do CPP e 0 dos acórdãos, sendo todas as 11 tipologias de componente-objeto-integral, porquanto elementos constitutivos de um todo em sua forma mais típica. Ex. A representação conterà todas as informações que possam servir à apuração do fato e da autoria.

#### 5. Considerações finais

Tendo em vista o limite de páginas imposto, resumem-se as considerações finais neste parágrafo. O presente artigo revisou a literatura quanto à RM, situou ontologias, léxicos computacionais e ontologias lingüísticas. Na análise, comparou dois *corpora* do domínio especializado, almejando testar se a busca automática por merônimos seria produtiva. A análise mostrou que: (i) provavelmente pela especialidade do domínio jurídico, a busca automática não restou produtiva, apesar de ter apresentado alguns resultados, conduzindo à conclusão de que a busca por merônimos deverá ser manual no caso dos *corpora* jurídicos; (ii) o *corpus* do CPP trouxe mais resultados; (iii) poucos ML trouxeram resultados precisos, a maioria deles esbarrou na polissemia dos termos jurídicos, como o sentido de *peça* como documento jurídico escrito, *parte* como autor ou réu, *tem* ou *possui* no sentido de posse e *constituído* como os poderes outorgados ao defensor para representar em juízo e (iv) a categoria ação-atividade de Winston et. al. (1987) foi o tipo de meronímia mais recorrente no DPP.

## Referências Bibliográficas

- BERTOLDI, A. A semântica dos adjetivos: como e por que incluí-la em uma ontologia de domínio jurídico. Dissertação. UNISINOS, São Leopoldo, 2007.
- BREUKER J, and WiINKEL, R. *Use and reuse of legal ontologies in knowledge engineering and information management*. ICAIL03 Wks on Legal Ontologies and Web-based Information Management, Edinburgh, <http://lri.jur.uva.nl/~winkels/legontICAIL2003.html>, 2003. (LRI-Core)
- CRUSE, D. A. *Lexical Semantics*. Cambridge: Cambridge University Press, 1986.
- CRUSE, D. A. *Meaning in language: an introduction to semantics and pragmatics*. Oxford: Oxford University Press, 2000.
- DI FELIPPO, Ariani. (2008). *Ontologias lingüísticas aplicadas ao processamento automático das línguas naturais: o caso das redes wordnets*. In: Magalhães, J. S.; Travaglia, L. C. (Orgs). *Múltiplas perspectivas em Linguística*. Uberlândia: Edufu, 2008. ISBN 978-85-7078-200-7.
- GANGEMI, Aldo. et al. *Some ontological tools to support legal regulatory compliance, with a case study*. Laboratory for Applied Ontology, ISTC-CNR, Rome Italy\*\* Università per Stranieri, Perugia, Italy \*\*\* ITTIG-CNR, Florence, Italy. s/a.
- GANGEMI, A.; SAGRI, M. T. e TISCORNIA, D. *A constructive framework for legal ontologies*. In: BENJAMINS, V. R. et al. (Eds.). *Law and the Semantic Web: Legal ontologies, methodologies, legal information retrieval, and applications*, LNAI (3369). Berlin/Heidelberg: Springer-Verlag, 2005, p.97-124.
- GRISHMAN, R. CALZOLARI, N. "Lexicons". In COLE, R. (Ed.). *Survey of the state of the art in human language technology*. Studies in Natural Language Processing. York: Cambridge University Press. 1997. p 368-370.
- GRUBER, Thomas R. *A Translation Approach to Portable Ontology Specifications*. In: *Knowledge Acquisition*, 5(2):199-220, 1993.
- MARRAFA, Palmira. *WordNet do Português: uma de dados de conhecimento linguístico*. Instituto Camões. 2001.
- MILLER, George A. et al. *Introduction to WordNet: An On-line Lexical Database*. 1993.
- PRÉVOT, L. et. al. *Ontology and the lexicon: a multidisciplinary perspective*. In: C. Huang. et al. *ontology and the lexicon: a natural language processing perspective*. New York: Cambridge University Press, 2010, p.3-24.
- RIEMER, Nick. *Introducing semantics*. Cambridge University Press. 2010.
- SAEED, John I. *Semantics*. Blackwell Publishers Inc. 1997.
- Sites:** Corpógrafo: <http://www.linguateca.pt>
- CLO: [www.ittig.cnr.it](http://www.ittig.cnr.it)
- LOIS: [www.loisproject.org](http://www.loisproject.org)
- WordNet: Disponível em: <http://wordnet.princeton.edu/perl/webwn?s=word-you-want>. Acessado em: 26/10/2010.
- EuroWordNet: Disponível em: <http://www.illc.uva.nl/EuroWordNet/>. Acessado em 27/10/2010.
- SOWA, J. F. (2006). *Building, sharing and merging ontologies*. Disponível: <http://www.jfsowa.com/ontology/ontoshar.htm>. acesso em 29/10/2010.