

# Inferência em Redes Semânticas: Investigação de Métodos de Inferência de Conhecimento de Tradução em ConceptNets Paralelas

Laís Augusta da Silva Meuchi<sup>1</sup>, Helena de Medeiros Caseli<sup>1</sup>  
Junia Coutinho Anacleto<sup>1</sup>

<sup>1</sup>Departamento de Computação, Universidade Federal de São Carlos (UFSCar)  
Rod. Washington Luís, Km 235, CP 676, CEP 13565-905 São Carlos-SP

{laismeuchi@comp,helenacaseli@dc,junia@dc}.ufscar.br

**Abstract.** *This paper describes a research proposal to extract translation knowledge from two ConceptNets derived from bases of common sense in different languages. Common sense can be defined as the knowledge shared by a group of people in a given time, space and culture. Given a ConceptNet in Portuguese and another ConceptNet in English, our goal is to infer knowledge that could be used to fix/filter automatically generated translations. To do so, in this paper we present the initial ideas, a background method and an example of inference in the ConceptNets under study.*

**Resumo.** *Esse artigo descreve uma proposta de inferência de conhecimento útil para a tradução automática a partir de duas redes de conceitos (ConceptNets) obtidas a partir de bases de senso comum em idiomas distintos. O senso comum pode ser definido como o conhecimento compartilhado por um determinado grupo de pessoas em um dado tempo, espaço e cultura. A partir de uma rede de conceitos em português e de outra rede de conceitos em inglês, pretende-se inferir conhecimento que poderá ser aplicado para ajustar/filtrar traduções geradas automaticamente. Para tanto, neste artigo são apresentadas as ideias iniciais, um método que poderá servir de base e um exemplo de aplicação da inferência nas redes em questão.*

## 1. Introdução

A Tradução Automática (TA) é uma subárea da Inteligência artificial que pode ser entendida como a tradução de uma língua natural (fonte) para outra língua natural (alvo) por meio de programas de computador. Um grande desafio da TA é produzir traduções que preservem o sentido da versão original.

Este trabalho sugere o uso de senso comum para ajustar/filtrar as traduções utilizando o contexto no qual as palavras se encontram para auxiliar a escolha da melhor tradução. O senso comum pode ser definido como o conhecimento compartilhado por um determinado grupo de pessoas em um dado tempo, espaço e cultura.

Para embasar essa proposta, na seção 2 são apresentados dois métodos de inferência em redes semânticas, propostos na literatura e que poderão ser considerados nesse trabalho. A seção 3 apresenta a estrutura e o funcionamento das duas redes de conceitos (ConceptNets) a serem utilizadas nessa pesquisa, bem como os objetivos da mesma. Por fim, a seção 4 apresenta as considerações finais e os próximos passos previstos para esse trabalho.

## 2. Levantamento Bibliográfico

Poucos trabalhos podem ser encontrados na literatura a respeito da inferência de conhecimento a partir de redes semânticas. Em [Shapiro 1978], são apresentados 2 métodos de inferência em redes semânticas: o método baseado em caminho (*path-based*) e o método baseado em nó (*node-based*).

O método baseado em nó parte do pressuposto de existência de instâncias de padrões de nós na rede semântica. Um exemplo de inferência baseado em nó encontrado em [Cravo e Martins 1989] e apresentado na Figura 1 explica como funciona esse método. Se existir, na rede semântica, as conexões e respectivas relações como na Figura 1-a, por meio de regras de inferência é possível inferir o resultado exposto na Figura 1-b (que *Rover* é membro da classe *mammals*). Essa inferência segue o princípio de que se um elemento é membro de uma classe que está contida em uma superclasse, então esse elemento também é um membro dessa superclasse.

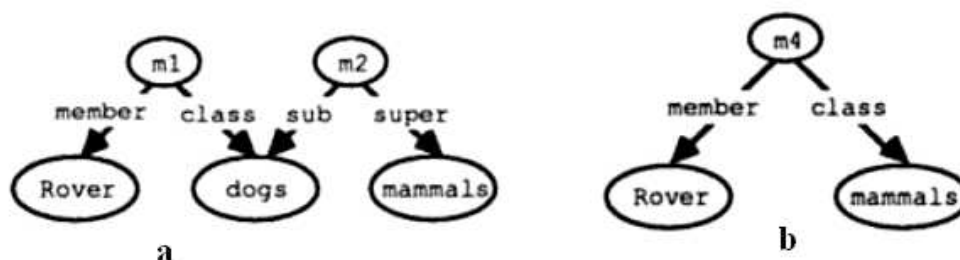


Figura 1. *Rover* é membro da classe *dogs* que, por sua vez, é uma subclasse da superclasse *mammals* [Cravo e Martins 1989]

Já o método baseado em caminho assegura que se existe um arco rotulado  $R$  entre dois nós  $a$  e  $b$ , pode-se assegurar que existe uma relação entre o nó  $a$  e o nó  $b$  (indicada por  $aRb$ ) mesmo que esta relação não esteja representada na rede, ou seja, podendo ser obtida por inferência através de uma outra informação presente na rede. Usando o exemplo anterior, podemos chegar ao mesmo resultado por meio desse método analisando-se a rede da seguinte maneira: quando encontramos uma sequência de arcos do tipo member-/class/sub-/super que liga um nó  $a$  a um nó  $b$  podemos inferir a existência de uma outra sequência de arcos do tipo member-/class entre esses nós podendo, assim, chegar ao nó  $m4$  (Figura 1-b).

Segundo [Shapiro 1978], o método baseado em caminho é mais eficiente, porém é restrito a relacionamentos binários, enquanto o método baseado em nó é mais geral. Uma junção desses dois métodos para que se aproveite as vantagens de ambos também é apresentada em [Shapiro 1978] e poderá ser considerada neste trabalho.

## 3. Inferência em Redes de Conceitos (ConceptNets)

A estrutura da ConceptNet é baseada em: (a) nós que contém os conceitos da base de senso comum e (b) arcos que conectam dois nós representando as relações entre eles. Existem vinte relações mapeadas pela ConceptNet, as quais foram definidas baseadas nos estudos sobre a teoria de Minsky de como o pensamento humano funciona [Minsky 1986]. Por exemplo, algumas relações encontradas nessas ConceptNets são: IsA, UsedFor, CapableOf, PartOf e DefinedAs.

Uma ilustração da estrutura das ConceptNets brasileira [Anacleto et al. 2008] e norte-americana [Singh 2002], com fragmentos semelhantes nas duas redes, é dada pelas Figuras 2 e 3, respectivamente.<sup>1</sup> Como pode-se observar nas Figuras 2 e 3, existe um relacionamento bem visível entre as duas redes. Pode-se achar muitos casos em que as relações entre nós de expressões em inglês possuem um correspondente em português, como no exemplo dado em que *book* está ligada com *learn* pela relação *UsedFor*, e *livro* também está ligada com *aprender* pela relação *UsedFor*.

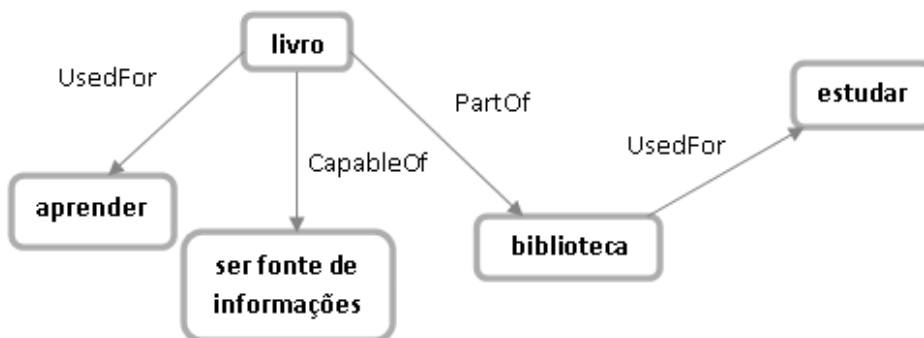


Figura 2. Representação gráfica da ConceptNet brasileira

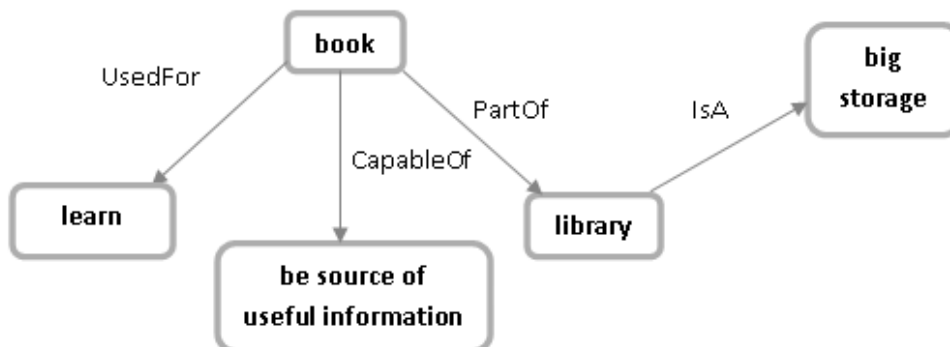


Figura 3. Representação gráfica da ConceptNet americana

A partir do conteúdo exposto, esse trabalho visa estudar métodos de inferência em redes semânticas, mais especificamente, nas ConceptNets brasileira e americana. Tais redes poderão sofrer um alinhamento prévio, ou seja, a definição das correspondências (ou links) entre os nós das ConceptNets em português e em inglês com base em diversas informações como, por exemplo: as correspondências entre as palavras que representam os conceitos em cada nó (obtidas de um dicionário bilíngue) ou mesmo a estrutura hierárquica da árvore.

As ligações entre os nós de uma mesma ConceptNet e entre as ConceptNets serão usadas para permitir a inferência de informações por meio de técnicas como as citadas na

<sup>1</sup>Outras informações sobre os projetos de senso comum brasileiro e americano podem ser obtidas, respectivamente, em: <http://www.sensocomum.ufscar.br> e <http://commons.media.mit.edu/en/>.

seção 2. Espera-se, com isso, que tais informações inferidas possam ser úteis para a tarefa de TA.

#### **4. Considerações Finais e Trabalhos Futuros**

Nesse artigo apresentou-se a proposta de inferência automática de informação útil para a TA a partir de conhecimento de senso comum representado por meio de ConceptNets. Não se tem conhecimento, até o momento, de pesquisas que empreguem conhecimento de senso comum no auxílio à TA da maneira sugerida nesse trabalho.

Os próximos passos dessa pesquisa se concentram na implementação de algoritmos para a inferência de “conhecimento de tradução” a partir das ConceptNets brasileira e norte-americana (por exemplo aqueles apresentados na seção 2). O conteúdo inferido será avaliado no auxílio da tarefa de TA propriamente dita.

#### **Referências**

- J. C. Anacleto, A. F. P. Carvalho, E. N. Pereira, A. M. Ferreira e A. J. F. Carlos (2008). Machines with good sense: How can computers become capable of sensible reasoning? Em *IFIP AI*, páginas 195–204.
- M. R. Cravo e J. P. Martins (1989). Path-based inference in snebr. Em *EPIA 89: Proceedings of the 4th Portuguese Conference on Artificial Intelligence*, páginas 97–106, London, UK. Springer-Verlag.
- M. Minsky (1986). *The Society of Mind*. Simon and Schuster, New York.
- S. C. Shapiro (1978). Path-based and node-based inference in semantic networks. Em *Proceedings of the 1978 workshop on Theoretical issues in natural language processing*, páginas 219–225, Morristown, NJ, USA. ACL.
- P. Singh (2002). The OpenMind Commonsense project. KurzweilAI.net. Disponível em: <<http://web.media.mit.edu/push/OMCSProject.pdf>>.