

## Uma Abordagem para Detecção Automática de Sinônimos em Mapas Conceituais

Ramon A. Azeredo<sup>1</sup>, Camila Z. Aguiar<sup>1</sup>, Davidson Cury<sup>1</sup>,  
Marcos Lovati<sup>1</sup>, Wagner Perin<sup>1</sup>

<sup>1</sup>Departamento de Informática  
Universidade Federal do Espírito Santo (UFES) – Vitória, ES – Brasil

ramon.azeredo@aluno.ufes.br

{camila.zache.aguiar, dedecury, marcos.lovati}@gmail.com

wagner.perin@ufes.br

**Abstract.** *Concepts maps are used to represent knowledge of specific domain and time. Inconsistencies in concepts maps may turn hard computational interpretation. The concept repetition is one of them, sometimes occurring in the form of synonym. This work presents an approach to automatic synonyms detection applied to maps recognizing it's importance to formalism and consistency.*

**Resumo.** *Mapas conceituais são utilizados na representação do conhecimento num dado domínio e instante de tempo. Falhas e redundâncias ocorridas em sua concepção podem torná-los difíceis de serem interpretados especialmente pelo computador. Uma falha comum é a repetição de conceitos por meio de sinônimos. Esse artigo explora uma abordagem para detecção automática de sinônimos aplicada aos mapas reconhecendo sua importância para dar mais formalismo e consistência.*

### 1. Introdução

Mapas conceituais são utilizados para organizar e representar graficamente o conhecimento de um indivíduo ou um grupo de indivíduos em um dado domínio e instante de tempo. Eles são formados fundamentalmente por conceitos e relações entre conceitos contendo palavras ou frase de ligação [Novak and Cañas 2006]. De um modo geral, os mapas não possuem regras rígidas e priorizam a simplicidade e versatilidade, expandindo seu uso nas mais diversas áreas do conhecimento e com variados objetivos.

Tendo em vista o seu potencial, várias soluções computacionais têm sido desenvolvidas para ampliar suas funções e facilitar os processos que envolvem a utilização destes, tais como: mecanismo de inferência em mapas [Perin et al. 2013]; comparação de mapas [Lamas et al. 2005]; mesclagem de mapas [Vassoler et al. 2014]; geração de mapas conceituais a partir de textos [Aguiar and Cury 2016]; construção de ontologias a partir de mapas [Nicchio Pinotte et al. 2015]; e representação automática das classes de conhecimento definidas por Piaget [Rios et al. 2015].

Ao analisar tais soluções é possível perceber, em alguns casos, que o conteúdo presente no mapa é processado a fim de interpretar o conhecimento e produzirem resultados inteligentes. É fácil concluir, portanto, que mapas contendo falhas em suas

representações, produzirão respostas incorretas pelos sistemas automáticos que dependem da consistência de suas informações. Isso não é novidade, já que um estudo realizado em 2014 [Assis et al. 2014] apontou consequências negativas em sistemas computacionais quando alimentados por mapas conceituais com erros ou inconsistência, dentre eles, a ocorrência de repetição de conceitos, mesmo que na forma de sinônimos.

Este artigo propõe uma abordagem para a detecção de conceitos sinônimos no momento da construção do mapa conceitual a fim de torná-lo mais consistente e formal, tendo por objetivo investigar o problema de reconhecimento de sinônimos em mapas conceituais, analisar as principais abordagens neste contexto e, por fim, propor um modelo conceitual.

## 2. Mapas Conceituais

O mapa conceitual é uma ferramenta que procura refletir a organização conceitual de uma pessoa (ou grupo de pessoas) sobre um domínio específico [Novak and Cañas 2006]. Trata-se de uma representação gráfica bidimensional que indica relações entre um conjunto de conceitos normalmente organizados de forma hierárquica [Pérez and Vieira 2005].

Os conceitos aparecem dentro de caixas [Souza and Boruchovitch 2010] representando uma série de objetos, eventos ou situações que possuem atributos comuns [Pérez and Vieira 2005]. Já as relações são especificadas por meio de frases de ligação sobre as setas direcionadas que unem os conceitos [Souza and Boruchovitch 2010]. Quando dois ou mais conceitos são conectados por meio de frases de ligação, surge uma unidade semântica chamada proposição [Souza and Boruchovitch 2010].

O mapa conceitual é baseado na teoria da aprendizagem significativa de David Ausubel [Souza and Boruchovitch 2010]. Segundo essa teoria, uma aprendizagem é dita significativa quando uma nova informação adquire significados para o aprendiz através de uma espécie de ancoragem com conceitos já existentes em sua estrutura de conhecimento [Moreira 2010].

Neste sentido, ferramentas computacionais têm surgido para facilitar os trabalhos envolvendo uso de mapas conceituais em todas as suas etapas (construção, manutenção e avaliação). Algumas destas ferramentas de apoio vão além dos serviços básicos e utilizam técnicas de Inteligência Artificial (IA) e Processamento de Linguagem Natural (PLN) para facilitar atividades complexas, tais como: mesclagem, comparação de mapas e, ainda, recuperação de informações tendo os mapas como base de conhecimento.

Apesar de não existirem regras rígidas para a construção de mapas conceituais [Novak and Cañas 2006], inconsistências podem tornar os mapas não interpretáveis do ponto de vista computacional, dificultando a extração de informações relevantes. De acordo com [Assis et al. 2014] as inconsistências mais comuns cometidas pelos estudantes em seus mapas conceituais são: ausência de frases de ligação, formação de frases ao invés de proposições, dois conceitos em uma caixa, repetição de conceitos, falta de simplicidade e clareza proposicional, ausência de verbos nas frases de ligação e concordância verbal.

Ainda segundo os autores [Assis et al. 2014], a existência de um agente inteligente que acompanhe o processo de construção de mapas conceituais pode minimizar a

ocorrência dessas inconsistências. Portanto, entendemos que é importante apoiar o processo de construção de mapas conceituais a fim de alcançar um maior formalismo por meio de uma representação mais consistente do conhecimento presente no mapa e, conseqüentemente, tornar sua interpretação mais coerente do ponto de vista computacional, a fim de produzir resultados mais consistentes nos sistemas inteligentes que trabalham com sua interpretação e extração de informações.

### 3. O Problema de Detectar Sinônimos

Segundo [Yıldız et al. 2014], duas ou mais palavras são sinônimas quando transmitem significados idênticos ou similares. No contexto dos mapas conceituais, cada par de conceitos unidos por uma frase de ligação forma uma “unidade de conhecimento” e estabelece uma relação semântica entre os conceitos [Perin and Cury 2016]. Dessa forma, um conceito sinônimo pode ser qualquer outro conceito que preserve essa relação semântica.

Na literatura, podemos identificar duas abordagens distintas direcionadas a tratar o problema de detecção de sinônimos: baseada em dicionários e em estatísticas de coocorrência de palavras.

#### 3.1. Dicionário

A forma mais tradicional de identificar se duas palavras possuem o mesmo significado é realizar uma consulta a um dicionário de sinônimos. Na computação, isso é feito utilizando bancos de dados léxicos, que são construídos manualmente. Conseqüentemente isso agrega alto nível de qualidade, mas envolve um custo elevado de esforço humano. Além disso, esse custo é repetido para cada novo idioma que se deseja construir um banco de dados léxico.

A maior limitação dessa abordagem é em relação a termos técnicos e científicos. Segundo [Turney 2001], em uma coleção de trabalhos científicos, apenas 70% das palavras-chave estão no *WordNet* (banco de dados léxico da língua inglesa).

#### 3.2. Estatística de Coocorrência de Palavras

Outra forma de reconhecer se duas palavras são sinônimas é analisar quais palavras as acompanham, isto é, com quais palavras elas coocorrem.

A forma mais simples e direta de calcular uma coocorrência  $n$  é contar o número de vezes que uma palavra de contexto  $c$  aparece a uma certa distância da palavra alvo  $w$  [Bullinaria and Levy 2007]. Esse cálculo pode ser normalizado dividindo essa coocorrência por sua negação  $\neg n$ , ou seja, o número de vezes que a palavra de contexto  $c$  não aparece a uma certa distância da palavra alvo  $w$ . O que resulta na probabilidade condicional  $P$  de  $c$  em relação a  $w$  ocorrer, como pode ser visto na Equação 1.

$$\frac{n(c, w)}{\neg n(c, w)} = P(c|w) = \frac{P(c \cap w)}{P(w)} \quad (1)$$

Quanto mais próximo de 1 for o resultado desse cálculo, mais forte é a ligação semântica entre as palavras  $c$  e  $w$ . De forma análoga, quanto mais próximo de 0 for o resultado, mais independentes são as palavras analisadas. Torna-se importante destacar que

um resultado próximo de 1 não significa necessariamente que duas palavras são sinônimas [Bullinaria and Levy 2007].

Na literatura, abordagens estatísticas têm sido avaliadas, principalmente, em: *dataset* de questões do TOEFL (*Test of English as a Foreign Language*) constituído de 80 palavras contendo, para cada uma dessas palavras, quatro possíveis candidatos a sinônimo; e *dataset* ESL (*English as a Second Language*) constituído de 50 questões. Por fim, a abordagem deve identificar qual dos candidatos é o sinônimo correto para a palavra.

Entretanto abordagens estatísticas podem sofrer com a esparsidade dos dados, isto é, o *corpus* utilizado pode ter palavras que aparecem raramente ou até mesmo que não aparecem. Assim, quanto maior o *corpus* utilizado, mais confiáveis serão as estatísticas para termos com baixa frequência [Bullinaria and Levy 2007].

Em razão disso, alguns algoritmos passaram a utilizar a *Web* como *corpus* para calcular as medidas de coocorrência, uma vez que a *Web* acompanha mais rapidamente às mudanças do idioma, tal como o surgimento de novas palavras, dialetos e significados. A análise é feita por meio de sites de busca como o Google, Yahoo e Bing que fornecem uma interface simples do conteúdo disponível na *Web* [Bollegala et al. 2007]. Assim, a quantidade de páginas encontradas para a busca “W1 AND W2” pode ser considerada uma medida global da coocorrência das palavras W1 e W2.

#### 4. Trabalhos Relacionados

A seguir explanamos abordagens encontradas na literatura direcionadas para a detecção de sinônimos, não limitado ao contexto de mapas conceituais.

O sistema SynFinder tem como objetivo encontrar sinônimos de uma palavra baseado no domínio informado [Lombardi and Marani 2015]. Trata-se de uma solução híbrida que usa: (i) a *WordNet* para recuperar o conjunto de sinônimos referentes a um termo em diferentes contextos e (ii) os resultados de pesquisas na *Web* para filtrar os sinônimos de acordo com o domínio especificado. A performance do sistema foi em média 0.94 para a medida *precision* e o tempo médio de execução foi menos de um segundo. Mas o sistema não é testado no *dataset* TOEFL ou ESL.

Em [Bullinaria and Levy 2007] os autores apresentam o método PPMIC (*Positive Pointwise Mutual Information with Cosine*) que utiliza a medida de similaridade PMI (*Pointwise Mutual Information*) e o *corpus* BNC (*British National Corpus*) para extrair as informações das palavras analisadas. Para calcular a similaridade de duas palavras W1 e W2, o método calcula o valor de coocorrência entre a palavra W1 e as palavras que aparecem com ela (contexto). Os valores gerados são utilizados como coordenadas de um vetor. O mesmo processo é feito para a palavra W2. Por fim, é medido a distância entre esses dois vetores calculando o ângulo do cosseno para detectar se as palavras são sinônimas. Esse método foi aplicado ao *dataset* TOEFL e obteve 85% de acertos.

Em [Islam and Inkpen 2006] os autores apresentam o método SOC-PMI (*Second Order Co-occurrence PMI*) que usa o PMI para detectar se duas palavras são sinônimas com base nas informações extraídas do *corpus* BNC. Primeiro, é realizado um pré-processamento do *corpus* eliminando *stop words* (palavras consideradas irrelevantes) e aplicando lematização (reduzir uma palavra flexionada a sua parte essencial). Depois disso, cada termo é inserido numa lista com sua respectiva frequência no *corpus*. Para

calcular a similaridade de duas palavras W1 e W2, o algoritmo faz duas listas selecionando as palavras que coocorrem com W1 e W2 e calcula um valor de PMI para cada palavra. Em sequência, as listas são ordenadas de forma decrescente de acordo com o valor do PMI. As palavras que estão no topo e que aparecem em ambas as listas têm o seu valor de PMI somados para gerar um valor único que represente a similaridade semântica entre as palavras W1 e W2. Esse método foi aplicado ao *dataset* TOEFL e obteve 76.25% de acerto e ao *dataset* ESL com 68% de acerto.

Em [Turney 2001] o autor apresenta o algoritmo PMI-IR para identificar sinônimos e assume que palavras similares tendem a ocorrer próximas umas das outras. Nesse caso a distância máxima é de dez palavras. A abordagem usa o PMI para calcular a coocorrência de palavras utilizando o número de páginas retornadas no site de busca. O PMI-IR obteve 73.75% e 74% de acerto no *dataset* TOEFL e ESL, respectivamente.

Em [Higgins 2004] o autor apresenta o algoritmo LC-IR (*Local Context - Information Retrieval*) que é inspirado no PMI-IR e usa o número de páginas retornadas por um site de busca para identificar sinônimos. O LC-IR assume que sinônimos tendem a ocorrer em formas gramaticalmente similares ao invés de proximidade (como o PMI-IR). O LC-IR obteve 81.3% e 78% de acerto no *dataset* do TOEFL e do ESL, respectivamente.

#### 4.1. Análise

Para fins de análise e avaliação, foram implementados alguns algoritmos para detectar sinônimos em questões dos *datasets* TOEFL e ESL utilizando as duas abordagens: dicionário e estatísticas de coocorrência.

Para a abordagem de Dicionário, implementamos um algoritmo usando o *WordNet*. Este é um dos mais utilizados banco de dados léxico da língua inglesa, contendo aproximadamente 117.000 conjuntos de sinônimos (*synsets*) [Miller 1995]. Apesar da grande quantidade de *synsets*, este algoritmo acertou apenas 6 (7,5%) e 8 (16%) das questões do TOEFL e ESL, respectivamente. Entretanto a execução do algoritmo foi rápida, demorando, em média, 2,8 (TOEFL) e 2,7 (ESL) segundos.

O algoritmo PPMIC [Bullinaria and Levy 2007], que calcula as estatísticas de coocorrência utilizando o *corpus* BNC, acertou 59 (73,75%) e 22 (44%) das questões do TOEFL e ESL. O tempo médio de execução do algoritmo foi de 1,8 (TOEFL) e 1,6 (ESL) segundos. Cabe destacar que é preciso indexar o *corpus* BNC antes de executar o algoritmo. Antes da indexação, foi feito um pré-processamento do *corpus* eliminando *stop words*.

Por último, foi implementado o algoritmo LC-IR [Higgins 2004], que calcula as estatísticas de coocorrência utilizando a *Web* como *corpus*. Esse algoritmo acertou 52 (65%) e 20 (40%) das questões do TOEFL e ESL. Entretanto o tempo gasto para resolver cada questão é em média de 9,6 segundos. O algoritmo utilizou o site de busca Bing<sup>1</sup>. Em [Higgins 2004] foi utilizado o site de busca AltaVista que não está mais disponível. Isso provavelmente explica a diferença entre o resultado encontrado em [Higgins 2004] e no algoritmo implementado nesse artigo.

Portanto, segundo os experimentos apresentados, podemos concluir que o algoritmo baseado em Dicionário obteve o menor número de acertos, dado sua limitação

---

<sup>1</sup>www.bing.com.br

de palavras. O algoritmo LC-IR que trabalha fazendo consultas em um site de busca, alcançou um número razoável de acertos, entretanto, o tempo de execução foi extremamente demorado devido as requisições que são feitas para o site de busca. Por fim, o algoritmo PPMIC [Bullinaria and Levy 2007] obteve o maior número de acertos e menor tempo de execução. Entretanto mesmo trabalhando com um *corpus* de aproximadamente 100 milhões de palavras, algumas palavras não foram encontradas como por exemplo “*bipartisanly*” e “*unequaled*”. A Tabela 1 apresenta a síntese dos resultados obtidos pelos três algoritmos.

**Tabela 1. Resultados dos algoritmos aplicados aos *dataset* TOEFL e ESL**

Algoritmo	TOEFL	ESL	Tempo Médio de Execução
Baseado em Dicionário	6 (7,5%)	8 (16%)	2,8s e 2,7s
PPMIC	59 (73,75%)	22 (44%)	1,8s e 1,6s
LC-IR	52 (65%)	20 (40%)	12 min e 8 min

## 5. Uma Abordagem Aplicada aos Mapas Conceituais

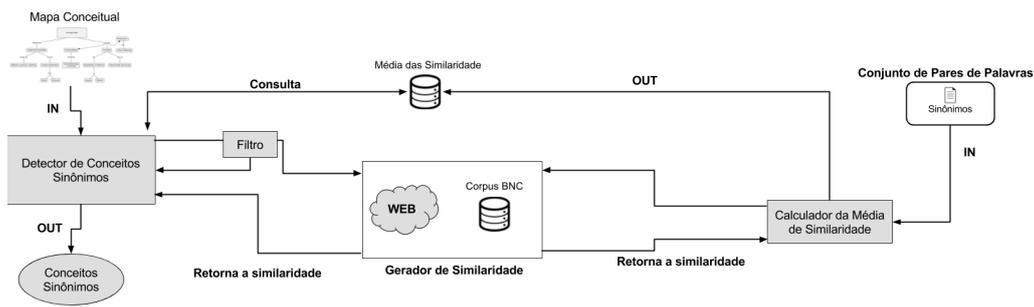
Em [Assis et al. 2014] os autores apresentam a proposta conceitual de um agente inteligente chamado VMap que atua de forma proativa sobre a estrutura e conteúdo do mapa conceitual, sugerindo aos autores adaptações e melhorias a fim de tornar os mapas mais formais e interessantes do ponto de vista do processamento computacional. Dessa forma, ao invés de trabalhar o mapa depois de construído para que seja consumido por ferramentas de IA, utiliza-se a IA para apoiar o autor no processo de construção do mapa, seguindo algumas diretrizes [Cañas and Carvalho 2005], mas sem limitar a expressividade característica dos mapas [Assis et al. 2014].

Uma das funcionalidades do VMap é que seja verificado se um novo conceito inserido no mapa já está presente, mesmo que na forma de um conceito sinônimo. Ao contrário dos *datasets* do TOEFL e ESL que possuem quatro opções e espera-se que apenas duas palavras tenham uma forte relação semântica, em um mapa conceitual todos os conceitos possivelmente vão ter uma forte relação semântica, fazendo com que a detecção automática de sinônimos seja uma tarefa ainda mais difícil.

Analisando as abordagens encontradas na literatura, as características do VMap e as particularidades dos mapas conceituais, propomos uma abordagem para detecção de conceitos sinônimos em um mapa conceitual seguindo duas principais tarefas: (i) cálculo do valor médio da similaridade de um conjunto de pares de sinônimos; e (ii) detector de conceitos sinônimos. A Figura 1 apresenta uma síntese do processo proposto para detecção de sinônimos em mapas conceituais explanadas a seguir.

Na primeira tarefa, “calcular valor médio de similaridade”, dado um conjunto S de pares de palavras sinônimas, para cada par é calculada a similaridade para se obter uma média. A similaridade é calculada por meio de um *corpus* como o BNC, por exemplo. Caso a palavra não esteja presente, o cálculo é feito utilizando a *Web* como *corpus*. Espera-se que ao calcular a similaridade de novas palavras, que ocorre na segunda tarefa, o valor seja maior ou igual a essa média caso as palavras sejam sinônimas.

Na segunda tarefa, a cada novo conceito inserido no mapa será verificado se esse conceito já não está presente no mapa na forma de um conceito sinônimo. Essa segunda



**Figura 1. Modelo Conceitual da Proposta**

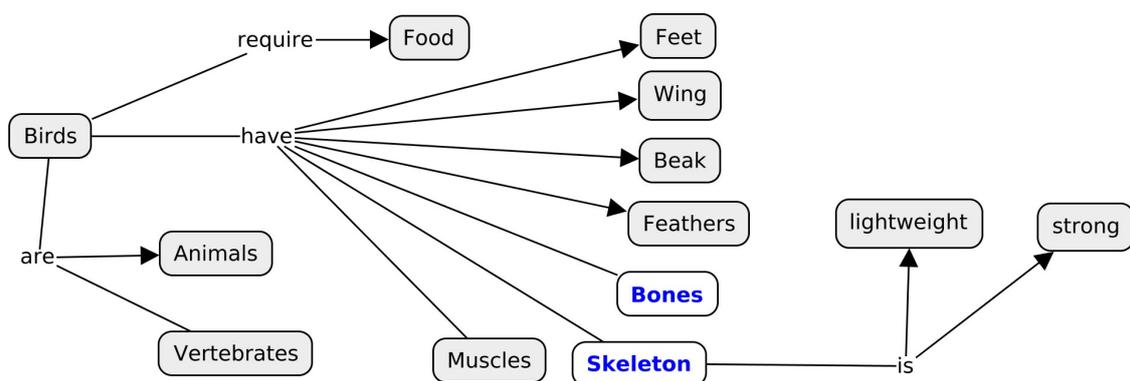
tarefa é subdividida em dois passos: (a) o par de conceitos passa por um filtro; e, se passar pelo filtro, (b) é calculado a similaridade entre as palavras.

O filtro é utilizado para identificar as relações semânticas de hierarquia (hiperonímia e hiponímia) e inclusão (holonímia e meronímia). Dado que um mapa conceitual possui como característica ter uma estrutura onde os conceitos mais gerais ficam posicionados no alto dos mapas e os mais específicos na parte inferior [Assis et al. 2014], essas relações (hierarquia e inclusão) são encontradas entre os conceitos e indicam que não se tratam de conceitos sinônimos. A identificação dessas relações é feita por meio de consultas ao *WordNet*.

Se o par de conceitos passar pelo filtro, então é calculado a similaridade entre os conceitos. O valor gerado é comparado com a média de similaridade calculada na primeira tarefa (i). Se esse valor for maior ou igual a média, será informado ao autor que o novo conceito inserido é um possível sinônimo de um conceito que já está no mapa.

### 5.1. Análise e Resultados

Para fins de análise e avaliação, foi feito um mapa conceitual, em inglês, que procura responder a questão focal “o que são pássaros?”. O mapa é apresentado na Figura 2.



**Figura 2. Mapa conceitual sobre pássaros**

O mapa gerado possui conceitos formados por apenas uma palavra, frases de ligação curtas e proposições autônomas. Na Tabela 2, são apresentadas as proposições do mapa.

Primeiro precisamos calcular o valor médio de um conjunto de palavras

**Tabela 2. Proposições do mapa sobre pássaros**

Proposições		
Conceito	Frase de Ligação	Conceito
Birds	Are	Animals
Birds	Are	Vertebrates
Birds	Require	Food
Birds	Have	Feathers
Birds	Have	Skeleton
Birds	Have	Muscles
Birds	Have	Bones
Birds	Have	Feet
Birds	Have	Beak
Birds	Have	Wing
Skeleton	Is	Strong
Skeleton	Is	Lightweight

sinônimas. Para formar esse conjunto foram utilizados 127 pares de palavras sinônimas extraídos do *dataset* TOEFL e ESL. Resultando em uma média de 0.57.

Na segunda parte do algoritmo, os conceitos do mapa conceitual são colocados em pares e formam um conjunto de 78 pares. Desses 78, apenas um par de conceitos são possíveis sinônimos: “*skeleton*” e “*bones*” (conceitos em azul no mapa). Depois esses pares passam pelo filtro para eliminar os casos de relação semântica de hierarquia e inclusão. Por exemplo, o conceito “*animals*” está numa relação de hiperonímia com “*birds*”, ou seja, parte-se de um conceito mais genérico para um mais específico. O filtro é responsável por reconhecer esse tipo de relação entre dois conceitos e indicar que não se trata de um caso de sinonímia.

Ao calcular a similaridade dos pares de conceitos e verificar se são maiores que a média, o algoritmo mostrou como resultado que 4 pares são possíveis sinônimos, como é possível ver na Tabela 3.

**Tabela 3. Resultado do algoritmo aplicado a um mapa conceitual**

Conceito	Conceito	Similaridade
Feather	Beak	0.74934429
Skeleton	Bones	0.59735242
Bones	Beak	0.60154832
Foot	Beak	0.61510560

O algoritmo trouxe 75% de falsos positivos. Essa baixa precisão mostra como é difícil reconhecer sinônimos em mapas conceituais. O par de conceitos que teve a maior similaridade, por exemplo, “*Feather*” e “*Beak*” não são sinônimos e o filtro, que utiliza o *WordNet*, não conseguiu identificar o tipo de relação entre essas duas palavras.

## 6. Conclusão e Trabalhos Futuros

Este artigo apresentou um estudo sobre técnicas computacionais para a detecção de sinônimos do qual resultou em um modelo conceitual aplicado aos mapas conceituais.

A maioria das técnicas computacionais para a detecção de sinônimos são testadas em *datasets* de questões de múltipla escolha. Nesse cenário a utilização de estatísticas de coocorrência consegue identificar a maioria dos sinônimos, já que se espera que apenas duas palavras tenham uma forte relação semântica.

No contexto dos mapas conceituais, a utilização apenas de estatísticas de coocorrência entre os conceitos presentes no mapa não é suficiente para identificar quais são sinônimos. O mapa representa o conhecimento de um dado domínio e, por isso, os conceitos vão ter uma forte relação semântica e isso torna o reconhecimento da sinonímia ainda mais difícil. Assim, a utilização de estatísticas de coocorrência precisam estar acompanhadas de outras técnicas para reconhecer sinônimos e, também, para eliminar os casos onde os conceitos possuem outras relações semânticas.

O algoritmo implementado com base no modelo conceitual proposto mostrou que é possível identificar conceitos sinônimos durante a construção de um mapa conceitual, mas apresentou o valor 0.25 para a medida *precision*.

Em trabalhos futuros, espera-se avaliar o algoritmo proposto de detecção de sinônimos em mapas conceituais na língua portuguesa, verificar se os mapas conceituais gerados se tornaram mais consistentes com a utilização do serviço e investigar os outros serviços que compõem o VMap.

## Referências

- Aguiar, C. Z. and Cury, D. (2016). A categorization of technological approaches to concept maps construction. In *2016 XI Latin American Conference on Learning Objects and Technology (LACLO)*, pages 1–9.
- Assis, D. V. d., Perin, W. d. A., Cury, D., and Vassoler, G. A. (2014). VMAP: Caracterização de uma Abordagem Para Verificação Sintática e Semântica de Mapas Conceituais. *Proceedings of the Sixth Int. Conference on Concept Mapping*, pages 514–521.
- Bollegala, D., Matsuo, Y., and Ishizuka, M. (2007). Measuring semantic similarity between words using web search engines. In *WWW '07: Proceedings of the 16th international conference on World Wide Web*, pages 757–766, New York, NY, USA. ACM.
- Bullinaria, J. A. and Levy, J. P. (2007). Extracting semantic representations from word co-occurrence statistics: A computational study. *Behavior Research Methods*, 39(3):510–526.
- Cañas, A. and Carvalho, M. (2005). Mapas conceituais e ia: Uma união improvável? *Brazilian Journal of Computers in Education*, 13(1).
- Higgins, D. (2004). Which statistics reflect semantics? rethinking synonymy and word similarity.

- Islam, A. and Inkpen, D. (2006). Second order co-occurrence pmi for determining the semantic similarity of words. In *Proceedings of the International Conference on Language Resources and Evaluation (LREC 2006)*, pages 1033–1038.
- Lamas, F., Boeres, C., Cury, D., and Menezes, C. S. d. (2005). Comparando mapas conceituais utilizando correspondência de grafos. *XVI Simpósio Brasileiro de Informática na Educação*, pages 24–27.
- Lombardi, M. and Marani, A. (2015). Synfinder: A system for domain-based detection of synonyms using wordnet and the web of data. pages 15–28.
- Miller, G. A. (1995). Wordnet: A lexical database for english. *Commun. ACM*, 38(11):39–41.
- Moreira, M. A. (2010). O que é afinal aprendizagem significativa? Technical report, Instituto de Física - UFRGS.
- Nicchio Pinotte, G., Cury, D., and Zouaq, A. (2015). Ontomap: De mapas conceituais a ontologias owl. 1:172–182.
- Novak, J. D. and Cañas, A. J. (2006). The theory underlying concept maps and how to construct them. Technical report, Institute for Human and Machine Cognition.
- Pérez, C. C. C. and Vieira, R. (2005). Mapas conceituais: geração e avaliação. *XXV Congresso da Sociedade Brasileira de Computação*, (4).
- Perin, W. d. A. and Cury, D. (2016). Uma Plataforma de Serviços para Mapas Conceituais. *XXVII Simpósio Brasileiro de Informática na Educação*, pages 230–239.
- Perin, W. d. A., Junior, R. G. M., Cury, D., and Menezes, C. S. d. (2013). Construindo mapas conceituais utilizando a abordagem imap. *Computer on The Beach*, pages 208–218.
- Rios, P. T. G., Cury, D., and Dutra, I. M. (2015). Automatizando uma argumentação construtivista por meio dos mapas conceituais. 1:157–162.
- Souza, N. A. d. and Boruchovitch, E. (2010). Mapas conceituais: estratégia de ensino/aprendizagem e ferramenta avaliativa. *Educação em Revista*, 26:195 – 217.
- Turney, P. D. (2001). Mining the web for synonyms: Pmi-ir versus lsa on toefl. In *ECML*, pages 491–502.
- Vassoler, G. A., Perin, W. d. A., and Cury, D. (2014). Mergemaps – a computational tool for merging of concept maps. *Proceedings of the Sixth Int. Conference on Concept Mapping*, pages 717–720.
- Yıldız, T., Yıldırım, S., and Diri, B. (2014). An integrated approach to automatic synonym detection in turkish corpus. pages 116–127.