

Metodologia de Apoio ao Processo de Ensino- Aprendizagem baseada em Extração de Informações Abertas e Teoria dos Grafos

Jonatas Vasconcelos, Vlândia Pinheiro

Programa de Pós-Graduação em Informática Aplicada
Universidade de Fortaleza
Av. Washington Soares, 1321, Fortaleza, Ceará, Brasil

jntvsc@gmail.com, vladiacelia@unifor.br

Resumo. *Este trabalho descreve uma metodologia automática para avaliação do aprendizado em domínios do conhecimento não específicos, utilizando um técnicas de Processamento de Linguagem Natural (PLN) e da Teoria dos Grafos. Objetiva através de Mapas Conceituais gerados a partir dos textos livres, em português, escritos por alunos e professores, avaliar o nível de assimilação e similaridade entre os mapas gerados. O principal diferencial deste trabalho é a utilização de métricas para avaliação automática dos alunos, consistindo em ferramenta de auxílio do professor. No experimento realizado concluiu-se que é possível, e com grau de acurácia da ordem de 75%, extrair triplas conceituais de textos descritivos de alunos.*

Abstract. *This paper describes an automatic methodology for assessment of learning in non-specific knowledge domains, using a Natural Language Processing (PLN) and Graph Theory techniques. Objective through Concept Maps generated from the free texts written in Portuguese, written by students and teachers, to evaluate the level of assimilation and similarity between the generated maps. The main differential of this work is the use of metrics for automatic assessment of students, consisting of a teacher aid tool. In the experiment it was concluded that it is possible, and with accuracy of 75%, to extract conceptual triples from descriptive texts of students.*

1. Introdução

Avaliar o aprendizado sempre foi um desafio na área da docência. Novak (2010) propôs uma ferramenta gráfica denominada Mapa Conceitual (MC) capaz de exibir cadeias de conceitos relacionados, formando um mapa dos conhecimentos adquiridos sobre uma determinada temática. Essa ferramenta fundamenta-se na Teoria da Aprendizagem Significativa de Ausubel (1968), segundo a qual, um indivíduo aprende quando consegue fazer relações com conteúdos aprendidos anteriormente. Segundo Novak, o docente deve criar experiências educacionais com foco na facilitação da aprendizagem significativa dos discentes.

A criação de proposições nos Mapas Conceituais precisa ser inteligível, ou seja, os conceitos e a relação semântica entre eles devem ser explicitamente expressos e significativos para o leitor. Redes Conceituais (RC) são estruturas gráficas formadas por

conceitos interligados entre si na forma de uma rede com o objetivo de representar uma estrutura de conhecimento. Entretanto, diferem de forma sutil dos Mapas Conceituais (MC), cujas estruturas gráficas expressam proposições, isto é, conceitos são ligados por relação verbal que agregam um significado para o conjunto: conceito inicial - relação verbal - conceito final. Na representação gráfica, os nós representam os conceitos, e as arestas rotuladas com frases curtas representam as relações verbais (Moreira, 1998).

Diversos trabalhos propõem metodologias, ambientes e ferramentas com uso de Mapas Conceituais como apoio ao processo de ensino-aprendizagem (BITTENCOURT et al., 2011; CARABETTA, 2013; LIMA, 2017; VIEIRA, 2018), sem a preocupação de automatizar as etapas do método. Há exemplo de trabalhos que procuram melhorar a coerência dos conceitos e relações presentes nos Mapas Conceituais (RIBEIRO et al., 2011), outros trabalhos focam a construção automática ou semiautomática dos Mapas Conceituais (KOWATA et al., 2011, 2010, 2009), e outros propõem utilizar mecanismos de inteligência artificial para favorecer a interação com mapas conceituais (HANDRICK, 2005; PERIN et al., 2012).

Neste trabalho, propomos uma metodologia para auxiliar o processo de ensino-aprendizagem que se baseia na construção automática de Mapas Conceituais, a partir de textos livres escritos por alunos e professores, na Língua Portuguesa do Brasil. A metodologia utiliza ferramentas de Extração Aberta de Informações (*Open Information Extraction* – Open IE) de textos e métricas da Teoria dos Grafos, as quais automatizam e otimizam a avaliação do processo ensino-aprendizagem. Ferramentas de Open IE não necessitam da definição *a priori* de quais conceitos e relações semânticas devem ser extraídas. Existem ferramentas de extração de informação aberta para a língua portuguesa - ExtroutMAP [Aguiar et al., 2017(b)] e DependentIE [Oliveira et al. (2017)], com nível de precisão e cobertura da ordem de 60%. Nos experimentos realizados neste trabalho, a ferramenta que apresentou melhor desempenho foi a DependentIE. De outro lado, propomos métricas da área de Teoria dos Grafos como suporte à avaliação e à comparação entre MCs. Grafos são estruturas gráficas representadas por pontos e linhas que formam pares ordenados de conjuntos no plano, se consideramos o conjunto $C = (V, A)$, os elementos $v \in V$ são chamados vértices ou nós de C , e os elementos $a \in A$ são chamados arestas ou links (Borba, 2013).

As questões de pesquisa que nortearam as investigações deste trabalho foram: (QP1) é possível e com que grau de acurácia extrair relações semânticas entre conceitos a partir de textos dissertativos de alunos? (QP2) As métricas da Teoria dos Grafos capturam a evolução do conhecimento do aluno? Foram realizados experimentos visando responder às questões de pesquisa acima e os resultados indicaram que a metodologia, aqui proposta, é viável e promissora.

2. Trabalhos Relacionados

O custo da tarefa de elaboração e análise dos Mapas Conceituais é alto, especialmente em novos domínios de conhecimento, exigindo experiência e esforço cognitivo do elaborador. Textos complexos e extensos também irão exigir do leitor um maior esforço para entendimento. Alguns autores pesquisaram as vantagens do uso de MC como ferramenta de apoio no processo de ensino-aprendizagem, bem como a geração automática dos mesmos. Pérez & Vieira (2005) apresentam uma proposta de sistema para extração de informações de textos livres em Língua Portuguesa do Brasil. Nesta proposta, para a extração dos argumentos das proposições foi utilizado o Palavras Xtractor, o qual extrai as palavras em sua forma básica, sem flexões de gênero, número ou grau. Lee & Segev (2012) propõem também um método automático para a construção

de mapas de domínio específico, aplicando técnicas de mineração de dados a um corpus para extrair palavras chaves relevantes com base na métrica TF/IDF.

Com uma abordagem diferente, Wang et al. (2016) focaram sua pesquisa na representação gráfica do conhecimento, a partir de livros didáticos e não nas respostas de alunos ou professores. O sistema proposto utilizava os índices dos livros e a hierarquia dos títulos e subtítulos para extração dos conceitos fundamentais, sendo o conteúdo enriquecido com informações da Wikipédia. Aguiar et al. (2017(b)) propõem automatizar o processo de construção de MC utilizando técnicas de mineração de dados e sumarização de textos. Eles usaram a ferramenta OpenIE do ExtroutMap para extração de informações abertas, a partir de estruturas linguísticas.

O trabalho de Koury (2018) utiliza o software *Wordlist* para listar e obter a frequência das palavras, e a ferramenta Hamlet II para gerar a rede conceitual e as relações diretas entre os conceitos (BRIER, 2011). As palavras chaves de textos eram classificadas manualmente através dos textos dos alunos e do professor, e, posteriormente, identificada de forma manual as semelhanças entre os MCs gerados de textos dos alunos, professor, e material didático. Utilizou uma métrica que indica o percentual de acertos dos conceitos-chaves para avaliar a assimilação dos alunos sobre o material didático e os conceitos do professor.

2.1. Sistemas de Extração de Informação Aberta (OpenIE)

A extração de informações abertas (Open IE) visa extrair, de textos em linguagem natural, as relações e seus argumentos, sem a necessidade de se determinar anteriormente quais informações devem ser extraídas. Em geral, os sistemas de OpenIE utilizam conhecimento linguístico, como etiquetas morfossintáticas e relações de dependências sintáticas, para identificar proposições sujeito-predicado-objeto e extrair triplas no formato (argumento1, relação, argumento2) (Aguiar, 2017a). A extração aberta de informação (Open IE) mostrou-se útil em várias aplicações de PLN (Processamento de Linguagem Natural), tais como Perguntas e Resposta (Fader et al., 2014), Extração de relações (Soderland et al., 2010) e Recuperação de informações (Etzioni, 2011).

Um dos desafios para os sistemas Open IE é que relações relevantes são frequentemente distribuídas entre as cláusulas ou apresentadas de uma forma não canônica. Por exemplo, no texto “Nascido em Honolulu, no Havaí, Obama é um cidadão dos EUA.”, o desafio está em não extrair triplas falsas como (Honolulu; nascer em; Havaí).

Oliveira et al. (2017) apresentam o sistema de OpenIE, denominado DependentIE, o qual utiliza uma abordagem de pesquisa em profundidade (DFS) para identificar os argumentos das frases. A técnica toma como base um Analisador de Dependência (DP) específico para o Português do Brasil. O estado da arte dos sistemas de Open IE apresenta precisão em torno de 57,34% [Oliveira et al. (2017); Santos, (2015); Xavier et al. (2015)] para língua portuguesa, o que indicia que estes sistemas ainda têm muito a avançar. No entanto, neste trabalho, prevemos o uso de sistemas de Open IE na metodologia proposta, dada a vantagem de realizar extrações automáticas, mas avaliamos o impacto das imprecisões e incompletude de tais sistemas, buscando responder à QP1.

3. Metodologia de Apoio ao Ensino-Aprendizagem

A metodologia proposta consiste em etapas para extração, geração e avaliação de Mapas Conceituais, obtidos a partir de textos livres na Língua Portuguesa do Brasil, escritos por alunos e professor. A geração é realizada por sistemas de OpenIE e a comparação é

realizada por meio de métricas da área de Teoria dos Grafos, sempre visando apoiar o processo ensino-aprendizagem. A Figura 1 ilustra o fluxo do macroprocesso.

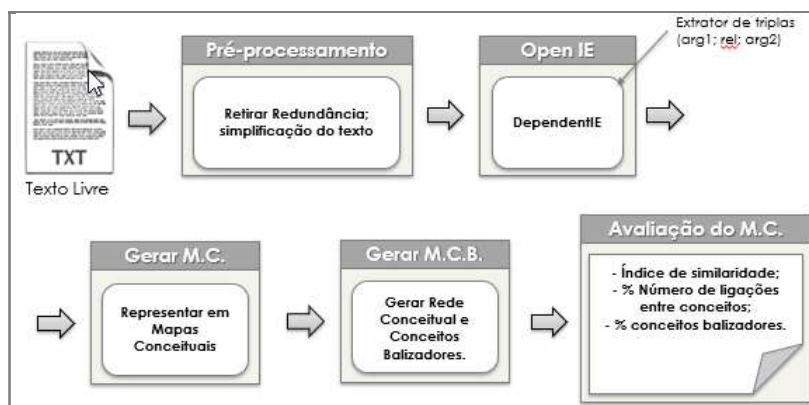


Figura 1 – Visão geral da metodologia proposta.

Em linhas gerais, um texto livre é previamente processado, de forma manual, para tratar redundância e simplificar as orações. A etapa de pré-processamento não é obrigatória, entretanto, nos experimentos realizados se mostrou necessária, devido à deficiência dos sistemas Open IE tradicionais em lidar com textos complexos e estruturas de orações coordenadas, subordinadas e apositivas. Por exemplo, considerando o primeiro parágrafo do texto livre escrito pelo Aluno 03 “*Pragmática: é o entendimento da sentença como todo, a comunicação que foi transmitida.*”, o sistema DependentIE extraiu erroneamente a tripla (arg1: Pragmático é | rel: foi transmitida | arg2: o entendimento, da sentença como todo). Após a etapa de pré-processamento, o texto foi transformado para “*Pragmática é o entendimento da sentença como todo, a comunicação que foi transmitida*”, e a nova tripla extraída (arg1: Pragmática | rel: é o entendimento | arg2: da sentença como todo.) é mais coerente.

Na etapa seguinte, a extração de triplas relacionais no formato (arg1, rel, arg2) é realizado por um sistema de OpenIE, a partir do texto livre.

Em seguida, as triplas de cada texto são usadas para geração automática do correspondente MC, cuja representação gráfica é usada como recurso avaliativo. Cada texto livre de cada aluno i corresponderá a um MC_i . Na sequência, é gerado um Mapa Conceitual Balizador (MCB), a partir do texto do professor, o qual servirá de baseline (referência) para avaliar a evolução do aprendizado do aluno. A partir do MCB serão extraídos os conceitos indutores do assunto objeto de estudo. Tais conceitos e seus sinônimos são os conceitos fundamentais que o professor pretende ensinar e compõem o lastro de todo o processo ensino-aprendizagem (Koury, 2018). A Figura 2 apresenta trechos de um texto original escrito por um professor, a lista dos conceitos indutores extraídos do texto original (argumentos das triplas), e a lista de conceitos sinônimos, gerada por busca em dicionários.

A Figura 3 apresenta um exemplo de MC_i (aluno i) e MCB (professor), extraído dos textos “*A Pragmática é o entendimento da sentença como todo, a comunicação que foi transmitida.*” e “*A análise pragmática, considera elementos contextuais para definir o significado do texto. A pragmática concentra na análise do falante, no tempo da fala, nas intenções/crenças, também em conhecimento de senso comum.*”.

Texto Original: “A análise léxica ou morfológica consiste em identificar, no texto de entrada, as palavras ou expressões e os morfemas das palavras. Morfemas são partes da palavra, por exemplo, raiz, prefixo, sufixo, etc. Geralmente, são usados tokenizadores para realizar a análise léxica e identificar os tokens do texto de entrada.....Por fim, a análise pragmática considera elementos contextuais para definir o significado do texto. Aqui o foco está na análise do falante, do tempo da fala, das intenções e crenças, e também em conhecimento de senso comum...”

Lista dos conceitos indutores:

- análise léxica, morfológica, palavra, morfema, tokenizador, token, análise pragmática, elementos contextuais, análise do falante, tempo da fala, intenções e crenças, conhecimento de senso comum.

Lista dos sinônimos:

- léxica = morfológica = morfologia; intenção = vontade; morfema = parte da palavra.

Figura 2 – Parâmetros usados na avaliação. Fonte: Autor.

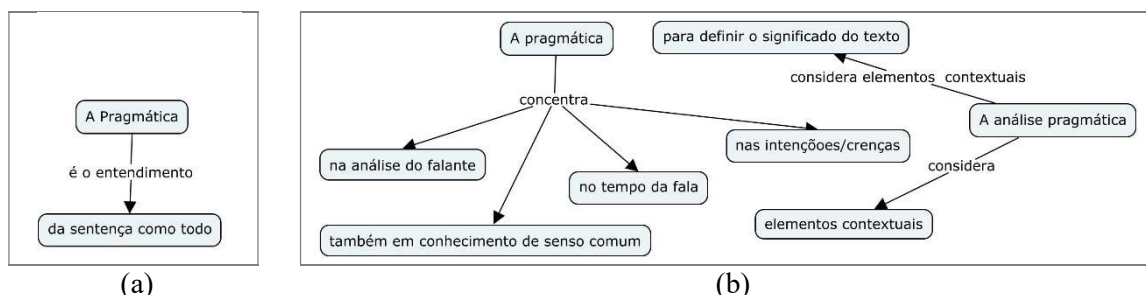


Figura 3 – Mapas Conceituais do Aluno (a) e do Professor (b). Fonte: Autor.

Na etapa final do processo, serão utilizadas métricas da Teoria dos Grafos para mensurar a similaridade entre os diferentes MC_i e o MCB. As métricas estão definidas nas subseções seguintes.

3.1. Métrica 01 – Assertividade de Conceitos (AC)

O **percentual de assertividade dos conceitos** pode ser usado para avaliar o grau de assimilação dos conceitos do tema estudado (Koury, 2018). A fórmula (1) é usada para calcular a assertividade dos conceitos presentes no MC_i (Mapa Conceitual do Aluno i) em relação ao MCB (Mapa Conceitual do professor).

$$(1) AC = \frac{|NCA|}{|NCI|} * 100$$

Seja o MC_1 do aluno 1 que contém 7 conceitos comuns ao MCB, que, possui ao todo, 17 conceitos indutores. Neste exemplo, o valor da assertividade do MC_1 é $AC = (8 / 17) * 100 = 47,05\%$. Ou seja, o aluno citou 47% dos conceitos que o professor propôs no ensino do tema.

3.2. Métrica 02 – Assertividade das Relações (AR)

O **percentual de assertividade das relações** captura quais relacionamentos entre os conceitos assimilados pelo aluno foram apreendidos pelo mesmo (Koury, 2018). A fórmula (2) é usada para calcular a assertividade dos relacionamentos presentes no MC_i (Mapa Conceitual do Aluno i) em relação ao MCB (Mapa Conceitual do professor).

$$(2) AR = \frac{|NRA|}{|NRCI|} * 100$$

Seja o MC_1 do aluno 01 que contém 7 relacionamentos comuns ao MCB, que, possui ao todo, 30 relacionamentos coerentes extraídos entre os conceitos indutores. Calculando a métrica 02 deste exemplo tem-se que a $AR = (7 / 30) * 100 = 23,33\%$.

Dessa forma, podemos considerar que o aluno entendeu em torno de 23% dos relacionamentos entre os conceitos que o professor propôs no ensino do tema.

3.3. Métrica 03 – Similaridade entre Grafos

Uma forma de aferir a similaridade entre grafos é compará-los por suas identidades através de isomorfismo¹. Para que possamos comparar os grafos com essa estratégia é necessário que os vértices (conceitos) e as arestas (relacionamentos entre conceitos) do grafo do aluno sejam idênticos aos do professor, ou seja, que o grafo do aluno esteja contido no grafo do professor. Seja **A** o grafo correspondente ao MC_i de um Aluno *i*, e **P** o grafo correspondente ao MCB do Professor. A similaridade entre dois grafos é calculada a partir do tamanho do Máximo Subgrafo Comum (MaSC) e Mínimo Supergrafo Comum (MiSC).

3.3.1. Máximo Subgrafo Comum (MaSC)

Supondo a existência de um subgrafo **A'** (grafo contido em **A** e **P** simultaneamente) comum aos grafos **A** e **P**, então **A'** é isomorfo por subgrafo a **A** e **P**. O subgrafo **A'** é considerado o **Máximo Subgrafo Comum** caso não exista outro grafo que seja isomorfo por subgrafo à **A** e **P** com um número maior de vértices.

A fórmula (3) é usada para calcular a similaridade entre os grafos **A** e **P**.

$$(3) \textit{Similaridade} (A, P) = \frac{|MaSC(A,P)|}{\max(|A| |P|)}$$

Sendo, **|A|** o número de relações entre os conceitos dos alunos e **|P|** o número de relações entre conceitos do professor. O MC₁ do aluno 1 contém 7 relações comuns ($|MaSC(A, P)|$) ao MCB, o qual possui 36 relações totais ($\max(|A| |P|)$). Calculando no exemplo dado, tem-se que $\textit{Similaridade} (A, P) = (7 / 36) = 0,19$ ou 19%. A equação (3) aplica a definição MaSC e retorna um valor de similaridade entre [0,1].

3.3.2. Mínimo Supergrafo Comum (MiSC)

O Mínimo Supergrafo Comum possui lógica contrária ao MaSC. Dessa forma o MiSC é um grafo de dois grafos, **A** e **P**, com menor número possível de vértices (Bunke & Shearer, 1998). A similaridade também pode ser calculada utilizando a distância entre os grafos e o MiSC (fórmulas 4 e 5).

$$(4) \textit{Distância} (A, P) = |MiSC(A, P)| - |MaSC(A, P)|$$

$$(5) \textit{Similaridade} (A, P) = 1 - \left(\frac{\textit{Distância}(A,P)}{|MiSC(A,P)|} \right)$$

Sendo **|A|** o número de relações entre os conceitos dos alunos e **|P|** o número de relações entre conceitos do professor. O MC₁ do aluno 1 contém no mínimo 1 vértices comuns ($|MiSC(A, P)|$) ao MCB, utilizando o ($|MaSC(A, P)|$) calculado 0,19 teríamos: $\textit{Distância} (A, P) = 1 - 0,19 = 0,81$, logo a $\textit{Similaridade} (A, P)$ seria: $1 - (0,81 / 1) = 0,19$ ou 19%. Enquanto o MaSC pode ser considerado um tipo de operador de interseção em grafos, o MiSC pode ser interpretado como união de grafos.

¹ Dois grafos **A** e **P** são isomorfos se podem ser representados pelo mesmo diagrama, isto é, se existe uma correspondência perfeita entre os vértices e arestas dos dois grafos (Neuhaus & Bunke, 2007).

4. Avaliação Experimental

A metodologia proposta foi aplicada em uma turma da disciplina de Introdução ao PLN, do curso de Pós-graduação em Informática Aplicada, composta de 9 alunos. Os alunos eram homens e mulheres entre 25 a 40 anos, todos cursando pós-graduação (em nível de mestrado e doutorado). No início da disciplina, o professor aplicou uma avaliação diagnóstica que continha a questão “Explique com suas palavras os objetivos de cada um dos níveis de processamento linguístico, identificando a interação entre eles.”. Como material de apoio, foi apresentada aos alunos uma figura da arquitetura genérica de um sistema de interpretação em linguagem natural. Todos os alunos e o professor responderam à questão de forma discursiva.

A metodologia foi aplicada em todos os textos usando as seguintes ferramentas: *DependentIE* – sistema de *OpenIE* que extraiu as triplas no formato (arg1, rel, arg2); *CmapTools* – ferramenta para geração dos MCs a partir das triplas (*CmapTools* 2019); Rotinas desenvolvidas para cálculo das métricas propostas (seções 3.1 a 3.3).

Para verificar a importância da etapa de pré-processamento, proposta nesse trabalho, avaliamos a precisão na extração das relações verbais (triplas arg1-rel-arg2) pelo sistema de *OpenIE* (*DependentIE*), a partir dos textos originais e dos textos com ajustes (após etapa de pré-processamento). Para os textos sem ajustes, o sistema extraiu 80 triplas, sendo 40 corretas, atingindo 50% de precisão. Para os textos com ajustes, a precisão subiu para 74,44%, pois o sistema conseguiu extrair 67 triplas corretas das 90 extraídas no total.

Os nove MCi foram gerados usando a ferramenta *CmapTools*, a partir das triplas extraídas pelo *DependentIE*. Em seguida as métricas de avaliação M01, M02 e M03 foram calculadas. A Tabela 2 apresenta o resultado da aplicação das métricas para cada MCi dos alunos. Importante verificar a melhoria nos índices de assertividade e similaridade, quando consideramos os textos originais ajustados (fase de pré-processamento). O Aluno 7 foi o que apresentou melhores índices e foi também o que obteve a nota máxima atribuída pelo professor na avaliação diagnóstica realizada.

Tabela 2: Métricas por aluno dos textos com e sem ajustes.

Alunos	Cálculo p/ texto s/ ajuste (%)			Média s/ ajuste	Cálculo p/ texto c/ ajuste (%)			Média c/ ajuste
	M01 ²	M02 ³	M03 ⁴	Méd.	M01	M02	M03	Méd.
Aluno 01	24,47	13,33	11,11	16,30	47,06	23,33	19,44	29,94
Aluno 02	27,53	10,00	8,33	15,29	52,94	20,00	16,67	29,87
Aluno 03	33,65	16,67	13,89	21,40	64,71	33,33	19,44	39,16
Aluno 04	18,35	10,00	8,33	12,23	35,29	20,00	13,89	23,06
Aluno 05	27,53	13,33	11,11	17,32	52,94	30,00	22,22	35,05
Aluno 06	33,65	10,00	8,33	17,33	64,71	20,00	16,67	33,79
Aluno 07	36,71	20,00	16,67	24,46	70,59	40,00	22,22	44,27
Aluno 08	33,65	16,67	13,89	21,40	64,71	33,33	19,44	39,16
Aluno 09	24,47	6,67	5,56	12,23	47,06	16,67	13,89	25,87

No final da disciplina, a mesma questão foi respondida pelos alunos e a metodologia, aqui proposta, foi novamente aplicada nos textos dos alunos. A Tabela 3 apresenta o resultado das métricas calculadas para os novos MCi de cada aluno. Mais

² Métrica 01, Assertividade de Conceitos (AC).

³ Métrica 02, Assertividade das Relações (AR).

⁴ Métrica 03, Similaridade entre Grafos.

uma vez, ocorre uma melhoria nos índices de assertividade e similaridade quando consideramos os textos originais ajustados (fase de pré-processamento). O Aluno 07 foi, novamente, o que apresentou melhores índices. Considerando a questão proposta pelo professor, observa-se que todos os alunos caíram de rendimento ao final da disciplina, segundo a metodologia de avaliação. Este fato também foi observado nas notas finais atribuídas pelo professor aos alunos.

Tabela 3: Métricas por aluno dos textos com e sem ajustes, no final da disciplina.

Alunos	Cálculo p/ texto s/ ajuste (%)			Média s/ ajuste	Cálculo p/ texto c/ ajuste (%)			Média c/ ajuste
	M01	M02	M03	Méd.	M01	M02	M03	Méd.
Aluno 01	41,18	6,67	5,56	17,80	35,29	16,67	13,89	21,95
Aluno 02	41,18	10,00	8,33	19,84	41,18	10,00	8,33	19,84
Aluno 03	23,53	6,67	5,56	11,92	23,53	10,00	8,33	13,95
Aluno 04	29,41	3,33	2,78	11,84	35,29	6,67	5,56	15,84
Aluno 05	29,41	10,00	8,33	15,91	35,29	10,00	8,33	17,87
Aluno 06	29,41	6,67	8,33	14,80	29,41	10,00	8,33	15,91
Aluno 07	35,29	16,67	13,89	21,95	47,06	13,33	11,11	23,83
Aluno 08	29,41	10,00	8,33	15,91	35,29	13,33	11,11	19,91
Aluno 09	41,18	10,00	8,33	19,84	41,18	10,00	8,33	19,84

4.1. Discussão dos resultados

As questões de pesquisas que nortearam este trabalho foram: QP1 “É possível e com que grau de acurácia, extrair relações semânticas entre conceitos a partir de textos descritivos de alunos?”; e QP2 “As métricas da Teoria dos Grafos capturam a evolução do conhecimento do aluno?”.

Como relação a QP1, concluímos que é possível a extração das relações de forma automatizada utilizando uma ferramenta OpenIE, contudo, com grau de acurácia baixo se considerarmos os textos originais, sem ajustes. A acurácia melhora quando são feitos ajustes nos textos no sentido de formalização com a gramática e simplificação das sentenças. Os sistemas de Open IE apresentam dificuldades ao lidar com textos complexos e estruturas de orações coordenadas, subordinadas e apositivas.

Analisando a QP2, concluímos que as métricas propostas consistem em indicadores do nível de assimilação e evolução do conhecimento do aluno. No experimento realizado, houve involução na assimilação de conceitos balizadores e seus relacionamentos, pelos alunos da disciplina. Ao apresentarmos este resultado ao professor, o mesmo analisou que o foco da disciplina foi mais prático do que teórico, o que ocasionou a perda do processo de ensino-aprendizagem. As métricas também capturam o padrão de resultado de alguns alunos em relação aos demais, como é o caso do Aluno 7 que permaneceu com a maior média.

5. Conclusão

Neste artigo, propõe-se uma metodologia automática para avaliação do processo de ensino-aprendizagem, baseada em técnicas de Processamento de Linguagem Natural e Teoria dos Grafos. Basicamente, Mapas Conceituais são gerados a partir dos textos livres, em português, de alunos e professores, e são avaliados o nível de assimilação e similaridade entre os mapas gerados. O principal diferencial da metodologia, aqui proposta, é a utilização de métricas para avaliação automática dos alunos, consistindo em poderosa ferramenta de auxílio do professor, guiando decisões e melhorias na estratégia de ensino-aprendizagem. No experimento realizado concluiu-se que é

possível, e com grau de acurácia da ordem de 75%, extrair triplas conceituais de textos descritivos de alunos, e que as métricas de avaliação capturam a evolução de conhecimento dos alunos sobre os conceitos balizadores da área em estudo

Como trabalhos futuros, tem-se a aplicação da metodologia em disciplinas diversas e uma análise mais profunda dos casos de erros, visando melhorar o processo de extração dos conceitos e relações. Outro desenvolvimento desta pesquisa será a combinação de MCs, visando estabelecer a conexão com conceitos já sedimentados pelos alunos.

Referências

- AGUIAR, C. Zaché. Concept Maps Mining for Text Summarization. Dissertação (Dissertação em Informática – Universidade Federal do Espírito Santo, Vitória-ES, p. 96-101, 2017(a).
- AGUIAR, C. Z., CURY, D. & ZOUAQ, A. Mineração de Mapas Conceituais para Sumarização de Textos. In: Anais dos Workshops do Congresso Brasileiro de Informática na Educação (CBIE). p. 57. 2017(b).
- AUSUBEL, D. P. Educational Psychology: A Cognitive View. New York: Holt, Rinehart and Winston. 1968.
- BITTENCOURT, G.K.G.D., SCHAURICH D., MARINI M., CROSSETTI M.G.O. Aplicação de mapa conceitual para identificação de diagnósticos de enfermagem. Rev. Bras. Enferm. 2011; 64(5):963-7.
- BRIER, A.; HOPP, B. Quantitative Analysis of Textual Data with HAMLET II 3.0 for Windows: A Multidimensional Scaling Approach to Quantitative Textual Analysis. 2011.
- BUNKE, H. & SHEARER, K. A graph distance metric based on the maximal common subgraph. Pattern Recognition Letters, 19(3–4):255 – 259. 1998.
- CARABETTA, V. A Utilização de Mapas Conceituais como recurso didático para a construção e inter-relação de conceitos. Revista Brasileira de Educação Médica. v. 37, n. 3, p.441–447, 2013.
- CmapTools: knowledge modeling kit. Produced by Institute for Human and Machine Cognition. 2019. Version 6.03.01. Disponível em: <http://cmap.ihmc.us>. Acessado em: 13 de maio de 2019.
- ETZIONI, Oren. Search needs a shake-up. Nature, 476(7358):25–26. 2011.
- FADER, Anthony; ZETTLEMOYER, Luke, and OREN, Etzioni. Open question answering over curated and extracted knowledge bases. In KDD. 2014.
- HANDRICK, F. ROSAI: Uma Proposta de Representação do Modelo ROSA em Linguagem Lógica. Rio de Janeiro, RJ: IME - Instituto Militar de Engenharia, Dissertação de Mestrado, 2005.
- KOURY, Alessandra Franchi. O uso de redes e mapas conceituais para comparar o conhecimento estabelecido, o processo de ensino e o aprendizado dos aspectos macroscópicos e submicroscópicos do equilíbrio químico. Dissertação (Dissertação em Ensino de Ciências – Universidade de São Paulo, São Paulo, p. 34, 2018.
- KOWATA, J. H.; CURY, D.; BOERES, M. C. S. Caracterização das Abordagens para Construção (Semi) Automática de Mapas Conceituais. Simpósio Brasileiro de Informática na Educação. Florianópolis, SC - Brasil: Anais do XX SBIE. 2009.
- KOWATA, J. H.; CURY, D.; BOERES, M. C. S. Uma abordagem computacional para construção de Mapas Conceituais. XXI Simpósio Brasileiro de Informática na Educação. João Pessoa, PB: Anais do XXI SBIE. 2010.

- KOWATA, J. H.; CURY, D.; BOERES, M. C. S. Construindo Mapas Conceituais a partir de Textos: Uma abordagem computacional aplicada à Língua Portuguesa do Brasil. Simpósio Brasileiro de Informática na Educação. Aracaju, Sergipe - Brasil: Anais do XXI SBIE. 2011.
- LEE, J. H., & SEGEV, A. Knowledge Maps For E-Learning. *Computers & Education*, 59(2), p. 353-364. 2012.
- LIMA, J. A.; SAMPAIO, C. G.; BARROSO, M. C. S. B.; VASCONCELOS, A. K. P. e SARAIVA, F. A. S. Avaliação da aprendizagem em química com uso de mapas conceituais. *Revista Thema*, v. 14, n. 2, p. 37-44, 2017.
- MOREIRA, M. A. Mapas Conceituais e Aprendizagem Significativa. *Cadernos do Aplicação*, v.11, n. 2, p. 143-156, 1998.
- NOVAK, J. D. Learning, Creating, and Using Knowledge. Concept Maps as Facilitative Tools in Schools and Corporations. Routledge, 2010.
- OLIVEIRA, Leandro Souza; GLAUBER, Rafael; CLARO, Daniela Barreiro. DependenceIE: An Open Information Extraction system on Portuguese by a Dependence Analysis. In: Anais do XIV Encontro Nacional de Inteligência Artificial e Computacional – ENIAC. 2017.
- PÉREZ, C. C. C., & VIEIRA, R. Mapas Conceituais: Geração e Avaliação. In Anais do III Workshop em Tecnologia da Informação e da Linguagem Humana (TIL). p. 2158-2167. 2005.
- PERIN, W. D. A.; CURY, D.; MENEZES, C. S. D. Construindo Mapas Conceituais Utilizando a Abordagem iMap. XVII Versão Congresso Internacional de Informática Educativa. Santiago - Chile: Anais do TISE. 2012.
- PERIN, W. d. A. iMap - Um Mecanismo de Inferência para Mapas Conceituais. Dissertação (Mestrado) — Departamento de Informática, Universidade Federal do Espírito Santo, 2014.
- RIBEIRO, E. F. et al. Um estudo sobre o incremento da Coesão e Coerência (Expressividade) em Mapas Conceituais. Simpósio Brasileiro de Informática na Educação. Juiz de Fora, MG - Brasil: Anais do XVI SBIE. 2011. p. 233-242.
- SANTOS, V. and PINHEIRO, V. Report é um sistema de extração de informações aberta para língua portuguesa. In Proceedings of the x Brazilian Symposium in Information and Human Language Technology (STIL), Natal, RN, Brazil. SBC. 2015.
- SODERLAND, Stephen; ROOF, Brendan; QIN, Bo; XU, Shi; MAUSAM; and ETZIONI, Oren. Adapting open information extraction to domain-specific relations. *AI Magazine*, 31(3):93–102. 2010.
- VIEIRA, M. Luis; CASTAMAN, A. Sara; INOCENTE, Luciane. Mapa Conceitual: Instrumento de Avaliação Educacional no Curso de Agronomia. *rev. Unoesc & Ciência ACHS*. v. 9 n. 1. p. 57-64. 2018.
- WANG, S., ORORBIA, A., WU, Z., WILLIAMS, K., LIANG, C., PURSEL, B., & GILES, C. L. Using Prerequisites to Extract Concept Maps From textbooks. In Proceedings of the 25th International Conference on Information and knowledge Management. p. 317-326. ACM. 2016.
- XAVIER, Clarissa Castellã; LIMA, Vera Lúcia Strube; SOUZA, Marlo. Open Information Extraction Based on Lexical Semantics. *Journal of the Brazilian Computer Society* 2015, 21:4 doi:10.1186/s13173-015-0023-2. 2015.