

Mineração de Mapas Conceituais para Sumarização de Textos

Camila Z. Aguiar¹, Davidson Cury¹, Amal Zouaq²

¹Departamento de Informática – Universidade Federal do Espírito Santo (UFES)
Vitória – ES – Brasil

² School of Electrical Engineering and Computer Science – University of Ottawa
Ontario, Canada

{camila.zacche.aguiar, dedecury, amalzouaq}@gmail.com

Abstract. *This article presents the development of a master's research directed to automatically construct concept maps for the summarization of text in Portuguese and English languages. To this purpose, a public online tool has been developed and, consequently, a public library for extracting information on concept maps. The tool was applied in a real controlled scenario for conceptual and technological analysis and validation of the proposal. The results have shown the importance of the research as support for the knowledge construction.*

Resumo. *Este artigo apresenta o desenvolvimento de uma pesquisa de mestrado que objetiva a construção automática de mapas conceituais para a sumarização de textos em idiomas português e inglês. Para este fim, uma ferramenta pública on-line foi desenvolvida e, conseqüentemente, uma biblioteca para extração de informações em mapas conceituais. A ferramenta foi aplicada em um cenário real controlado para análise e validação da proposta. Os resultados obtidos foram satisfatórios e apresentaram evidências da importância da pesquisa como apoio à construção do conhecimento.*

1. Introdução

Mapas conceituais são ferramentas gráficas para a representação e construção do conhecimento [Novak & Cañas 2010]. Conceitos e relações formam a base para o aprendizado e, portanto, os mapas conceituais têm sido amplamente utilizados em diferentes situações e propósitos na educação, sendo uma delas a representação do texto escrito. Mesmo um gramático e complexo texto pode ser representado por um mapa conceitual contendo apenas conceitos e relações que representem o que foi expresso no texto de uma forma mais complexa.

No entanto, a construção manual de um mapa conceitual exige bastante tempo e esforço na identificação e estruturação do conhecimento, especialmente quando o mapa não deve representar os conceitos existentes na estrutura cognitiva do autor. Em vez disso, o mapa deve representar os conceitos expressos em um texto, geralmente de um outro autor. Assim, várias abordagens tecnológicas têm sido propostas para facilitar esse processo de construção.

Este artigo apresenta uma nova abordagem para a construção automática de mapas conceituais como sumarização de textos acadêmicos em idiomas português e inglês. A sumarização produz um mapa conceitual como uma representação resumida do texto,

mantendo suas diversas e mais importantes características. A abordagem compreende técnicas de processamento de linguagem natural e recuperação de informação aplicadas a partir de uma ferramenta pública on-line e uma biblioteca para extração de informações em mapas conceituais desenvolvidos como artefatos da pesquisa. A ferramenta será integrada a uma plataforma baseada em serviços, em desenvolvimento.

2. Caracterização do Problema e Motivação

A sociedade da informação está constantemente acessando informações de forma muito rápida e ampla, e novas informações são produzidas, refletidas, publicadas ou compartilhadas quase que instantaneamente. Embora nos permita aprofundar nessa vasta rede de informações, ela também produz uma sobrecarga cognitiva reconhecidamente prejudicial aos nossos processos perceptivos e cognitivos [Toffler 1970].

A partir de uma análise quantitativa sobre todas as coleções de algumas bases de dados científicas (Springer, IEEE Xplore e ACM), verificamos que o número de artigos publicados nos últimos dezesseis anos é superior ao do último século. Assim, um estudante é confrontado constantemente com grande quantidade de informação, em fluxo incontrolável, para se manter informado sobre um assunto em particular. Esse fato traz diferentes desafios para o processo de aprendizagem do estudante, dos quais destacamos: (i) O estudante deve selecionar os documentos relevantes para um determinado assunto dentre todos aqueles disponíveis. Naturalmente, ele precisa compreender as informações contidas nesses documentos para depois selecioná-los ou não; (ii) Os documentos são compostos por grande quantidade de informação, geralmente escrita de forma complexa e, na maioria dos casos, em idioma diferente do idioma nativo do estudante, o que interfere na sua capacidade de ler e entender; (iii) Depois de selecionar o documento relevante, o estudante também deve investir um grande esforço cognitivo para identificar e entender as informações descobertas.

A pesquisa da literatura sugere que a representação gráfica pode reduzir os problemas de sobrecarga de informação e desorientação de aprendizagem para estudantes [Chen et al. 2008]. Assim, mapas conceituais, enquanto uma representação gráfica, podem ser utilizados como uma representação mais significativa da informação. Portanto, consideramos a hipótese de que a sumarização de um texto representado por um mapa conceitual pode atribuir características importantes para assimilar o conhecimento do texto, bem como diminuir a sua complexidade e minimizar tempo e esforço cognitivo.

Mapas conceituais utilizados como ferramenta para a representação gráfica de textos proporciona uma maneira visual e holística de representação do conhecimento. Usando mapas conceituais, o estudante poderia conhecer os principais conceitos do assunto antes de se aprofundar no texto. Isso favoreceria a assimilação de novos conhecimentos, especialmente em textos cujo idioma não é o idioma nativo do estudante. Consequentemente, olhando para a representação gráfica, o estudante poderia reduzir o tempo necessário para analisar a relevância do documento para o assunto. Assim, um mapa conceitual que represente efetivamente a sumarização de um documento, permite aos usuários obter uma certa compreensão independente do documento [Karannagoda et al. 2013].

3. Método de Pesquisa

O processo investigativo aplicado durante esta pesquisa é extenso e, portanto, será sintetizado nas principais contribuições para a área e para a educação.

A pesquisa se iniciou por meio de uma **questão de investigação** definida como: “Abordagens tecnológicas podem automaticamente construir mapas conceituais sumarizados a partir de textos?”. Desta questão, definimos um **processo de mineração de mapas conceituais** cobrindo quatro eixos de interesse: Descrição da Fonte de Dados: define qual o tipo de fonte e técnica utilizadas para processamento; Definição de Domínio: identifica o domínio da fonte de dados; Extração de Elementos: considerado o núcleo do processo, faz uso dos passos anteriores para extrair conceitos e relações; e Visualização do Mapa: especifica a representação gráfica das proposições.

Devido à necessidade de melhor identificar e analisar as funcionalidades e características inerentes a questão de investigação, realizamos um **estudo detalhado** sobre abordagens tecnológicas para construção automática de mapas conceituais publicados entre 1994 e 2016 nas bases de dados IEEE Xplore, ACM e Elsevier Science Direct. A partir deste estudo, elaboramos uma **categorização** definida em duas perspectivas, Fonte de Dados e Representação Gráfica, contendo 14 categorias. O estudo selecionou 30 artigos relevantes, que foram aplicados à categorização proposta para identificar os principais aspectos e limitações de cada abordagem. Destes, apenas cinco foram identificados como **trabalhos relacionados**.

A partir das informações coletadas, elaboramos um **modelo conceitual**, do qual resultou o desenvolvimento de uma **arquitetura** orientada a serviços para construção automática de mapas conceituais a partir de textos. Tal arquitetura foi implementada como a ferramenta on-line CMBuilder¹, elaborada a partir da biblioteca ExtroutMap² que foi produzida como um artefato da pesquisa e disponibilizada publicamente para utilização, extensão e requisição de serviços.

Finalmente, para analisar o contexto e os artefatos produzidos, realizamos uma série de **experimentos e provas de conceito**: (i) Para verificar as dificuldades e limitações na construção manual de mapas a partir de textos, realizamos um experimento com especialistas aplicando a construção manual e analisando as dificuldades e qualidade dos mapas construídos; (ii) Para validar a exatidão da biblioteca ExtroutMap-OpenIE na extração de proposições, realizamos uma análise objetiva sobre o dataset English Wikipedia³ comparando o resultado com outras duas bibliotecas, OLLIE⁴ and ClausIE¹, cujo resultado alcançou precision de 0.65; (iii) Para validar a qualidade do mapa conceitual construído pela ferramenta realizamos uma análise subjetiva comparando o mapa construído pelo CMBuilder com o mapa construído manualmente por especialistas, a ferramenta alcançou 0.75/0.45 para precision/recall dos conceitos e 0.57/0.23 para precision/recall das relações em idioma inglês e alcançou 0.68/0.38 para precision/recall dos conceitos e 0.41/0.19 para precision/recall das relações em idioma português; e (iv) Para verificar se o mapa conceitual sumarizado pelo CMBuilder tem influência para a compreensão do assunto abordado em um texto, realizamos um experimento com

¹ <http://cmpaas.inf.ufes.br/cmbuilder>

² <http://extroutmap.lied.inf.ufes.br>

³ www.mpi-inf.mpg.de/departments/d5/software/clausie

⁴ <https://knowitall.github.io/ollie>

estudantes de pós-graduação utilizando mapas conceituais sumarizados como fonte de dados para um questionário. Os estudantes atingiram 60% de acertos para mapas extraídos de pequenos textos com questões de múltipla escolha e 77% de acertos para mapas extraídos de textos extensos com questões discursivas.

4. Trabalhos Relacionados

A partir de uma revisão da literatura sobre abordagens tecnológicas para construção automática de mapas conceituais entre os anos de 1994 e 2016, nas bases IEEE Xplore, ACM e Elsevier Science Direct (detalhada em [Aguiar et al. 2016]), aplicamos a categorização proposta para obter uma análise objetiva sobre essas abordagens e selecionar os trabalhos relacionados, apresentados a seguir.

A abordagem [Wang et al. 2008] gera mapas conceituais a partir de resumos em inglês. Utiliza análise morfológica e sintática, identificando os elementos com base na estrutura das frases e regras sintáticas. Aplica a normalização para corrigir erros ortográficos, depende da detecção de sinônimos e da resolução anáfora. Usa análise estatística para verificar a relevância das proposições, fazendo uso da interação com o usuário para definir proposições incertas.

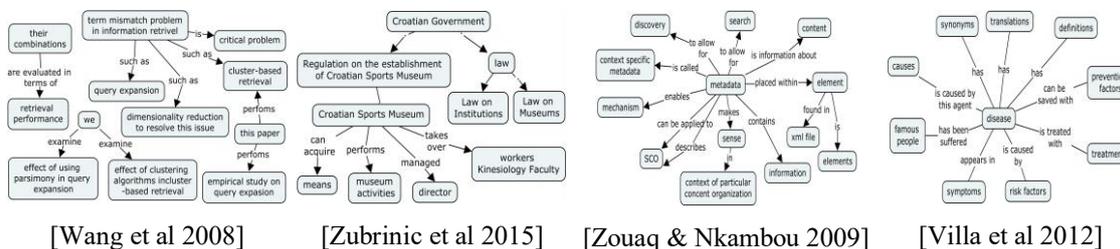
A abordagem [Zubrinić et al. 2015] gera mapas a partir de documentos legais em língua croata como um resumo do texto. Cria mapas hierárquicos de uma área específica usando o thesaurus do domínio. A partir de um corpus de domínio, os documentos são pré-processados e os metadados são mapeados. Usa técnicas linguísticas para lematização, reconhecimento de entidades, resolução de co-referência, análise léxica e sintática. Os conceitos são identificados a partir dos metadados e da frequência dos termos no texto. As proposições são extraídas a partir do padrão sujeito-predicado-objeto que contenha os conceitos identificados e relações estabelecidas entre os conceitos em um thesaurus. Uma estrutura de árvore formada por 25-30 conceitos das proposições é construída hierarquicamente atribuindo o título do texto como nó raiz.

A abordagem [Zouaq & Nkambou 2009] gera mapas conceituais de textos em inglês como etapa intermediária para gerar uma ontologia. Para isso, utiliza técnicas linguísticas de segmentação, normalização, análise estatística e sintática. Aplica o aprendizado de máquina para identificar palavras-chave e cria um mapa conceito semântico de frases contendo essas palavras-chave. As triplas são extraídas a partir das regras sintáticas e dependências gramaticais entre as palavras na frase. Os padrões léxico-semântico interpretam essa estrutura para extrair conceitos e relações. Finalmente, realiza análise estatística para definir a relevância de conceitos e relações.

A abordagem [Villa et al. 2012] gera mapas conceituais de texto clínico em língua inglesa. Esta abordagem usa conceitos e uma ontologia para obter ricas informações sobre o domínio. O sistema pré-processa um conjunto de termos médicos compilados em uma lista e busca por termos do domínio no texto. O usuário escolhe um conceito e consultas são realizadas na base de conhecimento para recuperar informações sobre o conceito.

Olhando para os mapas gerados pelas abordagens (Figura 1), podemos observar: mapa fragmentado em porções [Wang et al. 2008]; rótulo de conceitos longo [Wang et al. 2008] e formado por pronome [Wang et al. 2008]; rótulo de relação ausente [Zubrinić et al. 2015] ou formado por preposição [Wang et al. 2008]; utiliza outras fontes de dados além do texto como ontologia [Zouaq & Nkambou 2009], base de conhecimento [Villa et al. 2012] e thesaurus [Zubrinić et al. 2015]; mapa criado a partir de um conjunto de

documentos [Zouaq & Nkambou 2009], ou pequeno texto contendo algumas sentenças [Villa et al. 2012]; e mapa que representa apenas um domínio específico [Villa et al. 2012; Zubrinić et al. 2015].



[Wang et al 2008]

[Zubrinić et al 2015]

[Zouaq & Nkambou 2009]

[Villa et al 2012]

Figura 1. Mapa conceitual construído pelos trabalhos relacionados

De acordo com a análise realizada nas abordagens, podemos identificar os seguintes desafios: (i) definir rótulos pequenos e significativos; (ii) identificar conceitos relevantes relacionados a um determinado domínio; (iii) estabelecer links entre conceitos que não são evidenciados diretamente no texto; e (iv) identificar o domínio de um documento. Ressaltamos que nenhuma das abordagens apresentaram a sumarização de um texto inteiro, limitando a representação a partir de um fragmento. Ademais, nenhuma delas está direcionada à construção de mapas a partir de textos em português e pouquíssimas não estão limitadas a um domínio específico.

4. Descrição da Solução

A seguir apresentamos uma visão geral do modelo conceitual elaborado para atingir o objetivo proposto por esta pesquisa. O modelo foi desenvolvido sobre um ambiente web que permite interações do usuário [Aguiar et al. 2015]. A Figura 2 apresenta uma visão geral do modelo e seus componentes.

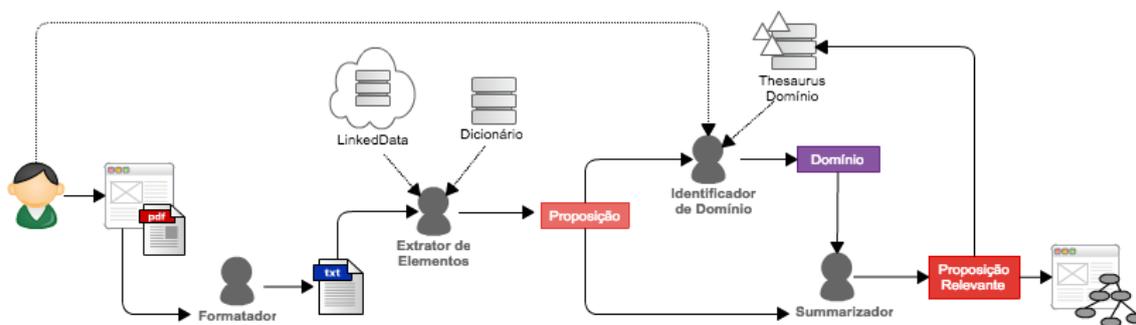


Figure 2. Visão geral do modelo conceitual

Uma síntese do processo mostrado na figura pode ser descrita da seguinte forma: o usuário acessa uma aplicação web e carrega uma fonte de dados em formato pdf. Em seguida, o **Servidor Formatador** transforma esse pdf em um texto não formatado. A partir desse texto, o **Servidor Extrator de Elementos** extrai um conjunto de proposições usando um dicionário léxico e um *Linked Data* para sua rotulação. Em seguida, as proposições são usadas pelo **Servidor Identificador de Domínio** para encontrar uma referência de domínio no Thesaurus de Domínio. As proposições e, opcionalmente, os conceitos do domínio de referência são usados pelo **Servidor Sumarizador** para gerar um mapa conceitual contendo proposições relevantes. Tais proposições são devolvidas a aplicação web onde o mapa conceitual é apresentado. As seções subsequentes abordam mais detalhes de cada etapa.

4.1. Servidor Formatador

O servidor é responsável por transformar o arquivo pdf de entrada em um texto não formatado. Para isso, são usadas atividades para (i) extrair texto a partir do pdf; (ii) eliminar marcadores de rótulos, referências, *tags* e estilo da fonte; e (iii) remover caracteres especiais.

4.2. Servidor Extrator de Elementos

O servidor aplica a biblioteca OpenIE do ExtroutMap que implementa um modelo para extração de informações abertas a partir de estruturas linguísticas. O processo inicia com a etapa de normalização da fonte de dados removendo frases não-proposicionais e resolvendo anáforas. Em seguida, o texto é tokenizado e recebe análise morfológica. Por fim, o texto é segmentado em sentenças e aplicada análise sintática, construindo uma árvore parse.

A etapa de *Identificação de Estrutura Independente* aplica uma nova segmentação para a árvore parse de cada sentença a fim de identificar estruturas independentes completas contendo uma estrutura menos complexa. Definimos por estrutura independente completa, aquela formada por sintagmas completos seguindo o padrão (1) e (2). Os sintagmas completos são: (i) sintagma NP, contém um núcleo NN ou os seus derivados, (ii) sintagma VP, contém um núcleo VB, ou derivado, e um sintagma completo NP, e (iii) sintagma PP, contém um núcleo IN, ou derivado, e um sintagma completo NP. Estruturas intermediárias, sintagmas incompletos e *tags* existentes entre os sintagmas completos, são ignorados.

$$S < ((NP < (NN+)) \$ (VP < (VB+ \$ (NP < (NN+)))) \quad (1)$$

$$S < ((NP < (NN+)) \$ (PP < (IN \$ (NP < (NN+)))) \quad (2)$$

Na etapa de *Ajuste de Estrutura* são aplicados alguns ajustes sobre as estruturas independentes a fim de tornar mais compreensível e clara as futuras proposições. Para isso, (i) regras morfológicas são aplicadas para identificar o núcleo dos sintagmas, adotamos tokens {JJ, NN} para nomes e {VB, IN, RB} para relações; (ii) todos os *tokens* que pertencem ao núcleo NP são lematizados; (iii) cada sintagma preposicional é transformado em um sintagma verbal por meio de um mapeamento; e (iv) relação de especialização é identificada por meio de nomes compostos e estrutura gramatical.

A etapa de *Extração de Triplos* objetiva extrair proposições que representem o fato expresso na estrutura independente completa. Para isso localizamos o primeiro sintagma verbal da estrutura e extraímos: sujeito, o sintagma nominal localizado antes do VP; objeto, o sintagma nominal localizado dentro do VP; e predicado, os *tokens* localizados entre o sujeito e objeto. A partir do sujeito, predicado e objeto é formada a proposição conceito-relação-conceito.

4.3. Servidor Identificador de Domínio

O servidor faz uso de técnicas de aprendizado supervisionado de clusterização e classificação para identificar o domínio do texto a partir da lista de proposições extraídas na etapa anterior. Um tesouro de domínio é construído automaticamente nesta etapa, armazenando as proposições do texto em grupos de domínio de forma incremental à medida que novos textos do mesmo domínio são processados para criar mapas.

4.4. Servidor Sumarizador

O servidor aplica uma sumarização sobre as proposições extraídas a fim de selecionar apenas as mais relevantes ao domínio do texto.

A etapa de Ranking é responsável por atribuir um peso para os conceitos segundo algum parâmetro. Para isso, representamos a lista de conceitos na forma de um grafo, considerando que cada vértice possui um *score hub*, número de conexões de saída; e *score authority*, número de conexões de entrada. Assim, o peso W de cada conceito k é computado pela fórmula (3), cujo peso $W(k)$ máximo é igual a 1. A fórmula associa o *score hub* $H(k)$ e *authority* $A(k)$ com a frequência dos conceitos no texto $TF_d(k)$ e no domínio $TFIDF_{\Omega}(k)$. Os melhores parâmetros de ajuste do modelo HARD [Leake et al. 2004] foram atribuídos para $\rho=2.235$ e $\sigma=1.764$. Os parâmetros $\beta=0.1$, $\alpha=0.2$ e $\gamma=0.7$ foram adotados em fase de experimento.

$$W(k) = [\beta.TFIDF_{\Omega}(k)] + [\alpha.TF_d(k)] + [\gamma.(\rho.A(k) + \sigma.H(k))] \quad (3)$$

A etapa de Sumarização é responsável por identificar as proposições relevantes diante do conjunto de triplas extraídas. Para isso aplicamos o conceito de quartis à topologia do grafo a fim de classificar os vértices, cujo peso de cada um é atribuído de acordo com a etapa de Ranking. Cada vértice é classificado como *heavy*, caso esteja localizado no terceiro quartil; *interjacent*, caso esteja localizado no caminho entre dois vértices *heavy*; *adjacent*, caso o peso do vértice seja superior ou igual ao menor peso dos vértices *interjacent*; e *light*, caso não se enquadre em nenhuma das classificações anteriores.

5. Prova de Conceito

No decorrer dessa pesquisa foram implementadas todas as funcionalidades, exceto o servidor formatador, do modelo conceitual proposto na Seção 4 em uma ferramenta publica on-line CMBUILDER. Diante de uma série de experimentos e provas de conceitos realizados, sintetizados na Seção 3, abordamos neste artigo a prova de conceito que objetiva verificar se os mapas construídos automaticamente pela ferramenta influenciam a construção do conhecimento sobre o assunto endereçado no texto. O experimento foi conduzido com doze estudantes de pós-graduação do curso de Ciência da Computação da Universidade Federal do Espírito Santo. Uma vez que o experimento tratou de assuntos conhecidos, os estudantes foram instruídos a não tirar proveito de seus conhecimentos prévios, mas a manter apenas a informação apresentada no experimento.

A coleta de dados foi realizada por meio de um questionário preparado com os seguintes recursos: **Texto I** contendo 123 palavras e **Texto II** contendo 302 palavras, apresentados na Figura 3. Esses textos foram submetidos ao CMBUILDER que os sumarizou automaticamente no **Mapa I** e no **Mapa II**, apresentados na Figura 4.

Texto I: “Biodiversity is the sum of all species on the planet. Some of these species contain important substances that treat several diseases. The most relevant thing about biodiversity is that the rich North needs biodiversity and the poor South has biodiversity. One of the ways to promote a sustainable development is to pay the poor nations to save the forests that they still have. Resources can be extracted but not exhausted. Thus, the environment can be preserved. The Earth belongs to all mankind. Everybody needs to help in the protection of the planet. And there is much to do. We have to fight pollution in all its forms to avoid acid rain, the greenhouse effect, and the death of species, rivers, lakes and seas.” [Adaptado de Newsweek June 1992]

Texto II: “Concept maps are graphical tools for organizing and representing knowledge. Concept maps include concepts, usually enclosed in circles or boxes of some type, and relationships between concepts indicated by a connecting line linking two concepts. Words on the line, referred to as linking words or linking phrases, specify the relationship between the two concepts. We define concept as a perceived regularity in events or objects, or records of events or objects, designated by a label. The label for most concepts is a word, although sometimes we use symbols such as + or %, and sometimes more than

Figura 5. Score alcançado pelo grupo A (texto) e B (mapa) na Etapa 1

A Figura 6 mostra os resultados alcançados pelos estudantes nos Grupos A e B durante a Etapa 2, questão discursiva. Usamos a seguinte distribuição de conteúdo para avaliar as respostas discursivas: O que é um mapa conceitual? (20%); Do que é composto? (20%); De onde veio a ideia principal? (20%); Quem criou? (20%); acrescido dos seguintes conceitos relevantes: mapa conceitual, ferramenta, representação, conceito, relação, Novak, Ausubel, conhecimento, psicologia, pesquisa e criança (20%).

Conforme mostrado pelo gráfico, o Grupo A, usando o mapa, apresentou resultados superiores (77%) do que o Grupo B, usando texto (54%). Este resultado é justificado pelo mapa conceitual sumarizado apresentar informações suficientes e objetivas para a interpretação do assunto abordado. A pontuação inferior do Grupo B vem do fato de que os estudantes não conseguiram extrair todas as informações relevantes do texto, pois o texto é gramaticalmente complexo e um pouco extensivo.

**Figura 6. Score alcançado pelo grupo A (mapa) e B (texto) na Etapa 2**

Observando os mapas construídos pela ferramenta e as questões respondidas pelos estudantes, podemos concluir que a qualidade do mapa conceitual construído automaticamente pelo CMBUILDER possibilitou aos estudantes responderem adequadamente, em sua maioria, as perguntas do questionário, o que fundamenta a solução proposta na Seção 4. Os estudantes atingiram 60% de acertos para mapas extraídos de pequenos textos com questões de múltipla escolha e 77% de acertos para mapas extraídos de textos extensos com questões discursivas. Além disso, verificamos uma grande diferença nas respostas elaboradas pelos estudantes dos dois grupos. O Grupo A, usando o mapa, elaborou respostas mais objetivas e completas, enquanto o Grupo B, usando texto, elaborou respostas mais dispersas.

Embora a pesquisa seja limitada, evidencia que o uso de um mapa conceitual sumarizado a partir do texto pode contribuir para a construção do conhecimento, porém enfatizamos que o seu sucesso depende da qualidade do mapa conceitual utilizado. Além disso, podemos considerar que um texto, especialmente extenso, apresenta grandes dificuldades ao leitor para identificar informações relevantes e consequentemente assimilá-las, caso não observado nos mapas conceituais.

6. Conclusões

O desenvolvimento de abordagens tecnológicas para construção automática de mapas conceituais a partir de textos tem mostrado resultados razoavelmente promissores. Para contribuir com os esforços em superar este desafio, apresentamos neste artigo o trabalho desenvolvido em uma pesquisa de mestrado direcionada a construir uma nova abordagem para construção automática de mapas baseada em técnicas linguísticas que processa textos em idioma português e inglês. Com o experimento descrito na Seção 5 concluímos que o uso de mapas conceituais sumarizados automaticamente a partir de texto pode ser um recurso de apoio à construção do conhecimento. Podemos observar grande diferença

na informação construída pelos estudantes que utilizaram mapas daqueles que utilizaram textos, o que também oferece fortes indícios da sua contribuição. Ressaltamos que os mapas construídos automaticamente pela ferramenta conseguiram representar, em sua maioria, a informação relevante do assunto abordado no texto e, conseqüentemente, validar a proposta e o objetivo da dissertação. Trabalhos futuros estão direcionados à melhoria da ferramenta no que diz respeito aos requisitos de usabilidade e a expansão da biblioteca em recursos de extração de informações em mapas conceituais.

Referências

- Aguiar, C. Z., Cury, D., & Zouaq. (2015). A. Automatic Construction of Concept Maps from Texts. In Proceedings of the 7th Concept Mapping Conference.
- Aguiar, Z. C., & Cury, D. (2016). A categorization of technological approaches to concept maps construction. In Learning Objects and Technology (LACLO). IEEE.
- Chen, N. S., Wei, C. W., & Chen, H. J. (2008). Mining e-Learning domain concept map from academic articles. *Computers & Education*, 50(3), 1009-1021.
- Karannagoda, E. L., Herath, H. M. T. C., Fernando, K. N. J., Karunarathne, M. W. I. D. (2013). Document analysis based automatic concept map generation for enterprises. In *Advances in ICT for Emerging Regions (ICTer)*. IEEE.
- Leake, D., Maguitman, A., & Reichherzer, T. (2004). Understanding knowledge models: Modeling assessment of concept importance in concept maps. In Proceedings of the 26th CSS.
- Novak, J. D., & Cañas, A. J. (2010). A teoria subjacente aos mapas conceituais e como elaborá-los e usá-los. *Práxis Educativa*, 5(1), 9-29.
- Toffler, A. (1970). *Future shock*. Amereon Ltd., New York.
- Villa, M., Aparicio, F., Maña, M. J., & de Buenaga, M. (2012). A learning support tool with clinical cases based on concept maps and medical entity recognition. In Proceedings ACM International conference on Intelligent User Interfaces.
- Wang, W. M., Cheung, C. F., Lee, W. B., & Kwok, S. K. (2008). Mining knowledge from natural language texts using fuzzy associated concept mapping. *Information Processing & Management*, 44(5), 1707-1719.
- Zouaq, A., & Nkambou, R. (2009). Evaluating the generation of domain ontologies in the knowledge puzzle project. *IEEE Transactions on Knowledge and Data Engineering*.
- Žubrinić, K., Obradović, I., & Sjekavica, T. (2015). Implementation of method for generating concept map from unstructured text in the Croatian language. In *Software, Telecommunications and Computer Networks (SoftCOM)*. IEEE.