

Uma Abordagem Computacional para Construção de Mapas Conceituais a partir de Textos em Língua Portuguesa do Brasil

Juliana Kowata¹, Davidson Cury¹, M. Claudia Boeres¹

¹ Depto de Informática – Universidade Federal do Espírito Santo (UFES)
Av. Fernando Ferrari, 514 – 29075-910 – Vitória – ES – Brasil

{juliana.kowata,dedecury}@gmail.com, boeres@inf.ufes.br

***Resumo.** Este artigo descreve uma abordagem computacional para a construção de mapas conceituais a partir de documentos em Língua Portuguesa do Brasil. Investigamos o problema da ressignificação de conteúdos textuais por meio do uso de mapas conceituais. Propomos uma solução amparada na aplicação de recursos computacionais por considerar que a independência do auxílio humano no processo de construção de mapas conceituais contribui para a superação das dificuldades de sua construção a partir do “zero”.*

***Abstract.** This article describes a computational approach to build concept maps from documents expressed in Portuguese of Brazil. We investigate the usage of concept maps as a way of giving new meanings to textual contents. We propose a solution supported by computational resources, because we believe that the independence of human aid in the construction process of concept maps helps to overcome the difficulties of build them from scratch.*

1 Introdução

A integração de diferentes tecnologias no currículo escolar fomenta o surgimento de maneiras diversas de ensinar e aprender, seja pela introdução de novas tecnologias, seja por meio da ressignificação das chamadas tecnologias convencionais. Compete ao professor decidir quais veículos e linguagens privilegiar para levar os estudantes à melhor compreensão dos conteúdos [Silva, 2005].

Nesta pesquisa, consideramos que os mapas conceituais são ferramentas que possibilitam a ressignificação de conteúdos, apoiando, desta forma, práticas educacionais voltadas para a aprendizagem significativa. Por isso, ao longo deste artigo, investigamos o problema da ressignificação de textos por meio do uso de mapas conceituais, buscando, em última instância, a construção de conteúdos dinâmicos a partir de conteúdos estáticos.

Propomos uma solução para construção de mapas conceituais amparada na aplicação de recursos computacionais por considerar que a independência do auxílio humano no processo de construção de mapas conceituais contribui para a superação das dificuldades de construção a partir do “zero” [Chang et al. 2008], na redução de tempo e esforços despendidos na aquisição de conhecimento [Lee et al. 2009] [Tseng et al. 2007], sobretudo em domínios extensos [Valerio e Leake, 2006] e dependentes de especialistas [Chang et al., 2008].

A pesquisa propõe uma abordagem computacional para a construção de mapas conceituais a partir de textos. Propusemo-nos a tratar textos em português do Brasil, em conformidade com as recomendações da UNESCO que orientam para a criação e

aplicação de recursos de aprendizagem em língua materna [UNESCO, 2009]. Assim, nossa pesquisa teve como escopo a **definição de uma abordagem computacional para a construção de mapas conceituais a partir de textos em Língua Portuguesa do Brasil**.

Formalmente, o problema resume-se na necessidade de tomar um documento d , com conteúdo expresso em Língua Portuguesa (Brasil), e representá-lo por meio de um mapa conceitual mc . O conteúdo do documento d é constituído por um conjunto de sentenças $s_1...s_n$, ou seja, $d = \{s_1, s_2, \dots, s_n\}$. Para cada sentença s_i definida por meio de linguagem natural, em que i designa o índice da sentença no documento d , é necessário identificar proposições $p_1...p_n$ a serem extraídas, com o intuito de possibilitar a construção de um mapa conceitual mc . Uma proposição p_i é definida por um conjunto de três elementos ordenados, $c1_i$, r_i e $c2_i$, nos quais $c1_i$ e $c2_i$ são conceitos e r_i uma relação entre os conceitos.

A transformação de uma sentença s_i em triplas proposicionais no formato $\langle \text{conceito} - \text{relação} - \text{conceito} \rangle$ requer a identificação prévia dos elementos candidatos a conceitos e relações. Sob o ponto de vista computacional, enfrentamos a questão da descoberta de elementos candidatos a conceitos e relações de mapas conceituais a partir de uma sentença s_i , como um problema de reconhecimento de padrões linguísticos e de entidades. A construção de proposições a partir de conceitos e relações é abordada como um problema de otimização que pode ser tratado por meio da seleção de um arranjo adequado entre elementos com o objetivo de formar uma tripla.

Este trabalho está estruturado nas seguintes seções: a Seção 2 exibe uma visão geral da nossa abordagem para o problema. A Seção 3 mostra o conjunto de atividades relativas à construção de mapas conceituais a partir de textos. A Seção 4 descreve o estudo de caso da abordagem e, finalmente, na Seção 5 apresentamos nossas considerações finais.

2 Visão Geral da Abordagem

A abordagem que propomos busca resguardar a fidelidade do mapa conceitual construído ao conteúdo do texto original, além de assegurar a compreensibilidade da nova representação. Dado os objetivos expostos e em face às dificuldades intrínsecas da manipulação da linguagem natural, além das particularidades da Língua Portuguesa, a adoção de métodos e técnicas que restrinjam a perda semântica no mapeamento texto-mapa conceitual, sem comprometer a essência do conteúdo e sem limitar a leitura humana, é crucial para o êxito desta abordagem.

Dentre os achados resultantes do levantamento que realizamos na literatura [Kowata, Cury e Boeres, 2009], encontramos que: 10 das 15 abordagens, 66,67%, usam fontes de dados não estruturadas; destas, a metade delas constroem mapas conceituais ditos completos por meio de métodos híbridos e métodos linguísticos. Levando-se em consideração os resultados apresentados por pesquisas similares, desenvolvemos uma predileção natural pelos métodos linguísticos e híbridos para tratar do problema.

Com vistas à auxiliar a resolução do problema, lançamos mão da analogia [Nonaka e Takeuchi, 1997] entre o documento e o brinquedo de encaixe, no formato do

famoso sistema LEGO¹. No brinquedo de encaixe, peças podem ser arranjadas de diversas maneiras, produzindo diferentes módulos. Analogamente, no documento, as sentenças são módulos constituídos por peças, no caso, por peças linguísticas, tais como: palavras, números e sinais de pontuação. Estas peças linguísticas são organizadas observando-se as regras da sintática, semântica e pragmática. Tanto no conteúdo de um documento como em módulos construídos a partir das peças de encaixe, pode-se observar a existência de unidades atômicas de construção. O rearranjo das peças do brinquedo permite novas construções, assim como a recombinação das peças linguísticas admite outras representações, sendo a representação proposicional a peremptória para a construção de mapas conceituais.

A partir da definição da analogia, como exposta anteriormente, propomos uma abordagem em que o conteúdo de um documento deve ser fragmentado até que suas peças, de natureza linguística, tornem-se disponíveis. Posteriormente, as peças linguísticas são reconstruídas na forma de conceitos e relações em mapas conceituais. A expressão síntese da abordagem consiste na dicotomia “desconstruir-reconstruir”. Assim, projetamos um modelo de processo composto por um conjunto de atividades que se concatenam por meio de suas entradas e saídas, formando um túnel, de sentido único, linear, como ilustrado na Figura 1.

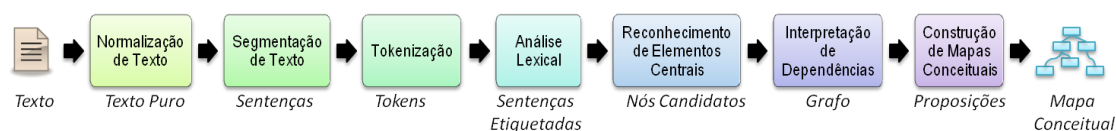


Figura 1. Modelo de Processo para Construção de Mapas Conceituais a partir de Textos

Na próxima seção, cada uma das atividades será explicada em mais detalhes.

3 Atividades

São sete as atividades do processo para a construção de mapas conceituais a partir de textos: (i) *Normalização de Texto*; (ii) *Segmentação de Texto*; (iii) *Tokenização*; (iv) *Análise Lexical*; (v) *Reconhecimento de Elementos Centrais*; (vi) *Interpretação de Dependências*; e (vii) *Construção de Mapas Conceituais*. As atividades são orquestradas sequencialmente de forma que a cada uma produza um único artefato que será tomado como insumo para a atividade seguinte, desde o documento de origem até a construção do mapa conceitual.

3.1 Normalização de Texto

Esta atividade é responsável por eliminar os marcadores de formatação existentes no conteúdo de um documento, tais como *tags*, estilos de fontes, metadados. A eliminação destas marcas permite a manipulação de arquivos em diferentes formatos de texto. A existência de inúmeros formatos de documentos aliada à velocidade de surgimento de novas extensões e de atualizações de versão que não guardam compatibilidade com as versões anteriores são fatores complicadores para a atividade que depende,

¹ LEGO é a designação comercial de uma família de brinquedos, muito popular, composta por peças que se encaixam (www.lego.com).

fundamentalmente, da definição de filtros apropriados para tratamento das diferentes codificações.

3.2 Segmentação de Texto

A manipulação de textos geralmente pressupõe a habilidade de dividi-lo em sentenças individuais [Bird, Klein e Loper, 2009]. A segunda atividade, denominada Segmentação de Texto, é justamente o mecanismo que provê a segmentação do conteúdo de texto puro em sentenças. Cada sentença limita um conjunto semântico mínimo para definição de uma proposição. A segmentação de textos é realizada por meio da identificação de caracteres finalizadores de sentenças, principalmente, dos sinais de pontuação. A dificuldade desta atividade está na correta associação dos sinais de pontuação ao fim de sentença, uma vez que tais sinais também podem demarcar abreviações de nomes, separação de dígitos em datas, horas, números de telefones e números ordinais [Cimiano, 2006].

3.3 Tokenização

A terceira atividade é a atividade de identificação de *tokens*, ou seja, de elementos de cada oração, inclusos neste conjunto: palavras, números e sinais de pontuação. Essa etapa utiliza-se, basicamente, de sinais gráficos como espaços e algoritmos para o reconhecimento de entidades limítrofes de um *token*. Segundo Cimiano (2006), o uso dos espaços em branco nem sempre definem de forma adequada as fronteiras das palavras, sobretudo de palavras compostas e de nomes próprios. Por isso, é recomendada a utilização conjunta *tokenizador* juntamente com técnicas de reconhecimento de nomes de entidades.

3.4 Análise Lexical

A atividade “*Análise Lexical*” consiste em determinar, para cada *token* a etiqueta morfológica correspondente. A etiqueta morfológica, também conhecida como classe de palavras identifica um: substantivo, verbo, pronome, preposição, advérbio, conjunção, artigo etc.

A saída desta atividade é um arquivo contendo *tokens* anotados por suas respectivas etiquetas morfológicas. A dificuldade da atividade está na desambiguação semântica do *token*, pois é bem comum que uma mesma palavra seja classificada em diferentes classes gramaticais, de acordo com o contexto de seu uso.

3.5 Reconhecimento de Elementos Centrais

A quinta atividade, “*Reconhecimento de Elementos Centrais*”, tem por objetivo produzir um conjunto de elementos que serão candidatos a conceitos e relações de mapas conceituais. Para isso, recebe como entrada, o conjunto de *tokens* etiquetados morfológicamente e realiza um processamento utilizando técnicas de *chunking*.

O *chunking* – também conhecido como uma análise sintática parcial ou superficial – aplica técnicas de processamento superficial, tipicamente caracterizados por autômatos finitos e expressões regulares, para manter o agrupamento de palavras, formando constituintes com significados, tipicamente com um núcleo que é modificado por outras palavras da unidade [Cimiano, 2006]. Em português, as principais sequências de palavras ordenam-se em torno de um nome (Sintagma Nominal), verbo (Sintagma

Verbal) ou são precedidas por preposições (Sintagma Preposicional). Esses arranjos são denominados genericamente de “*chunks* linguísticos”.

Novak e Gowin (1984) afirmam que, na construção de mapas conceituais, é importante que palavras adequadas sejam escolhidas para rotular conceitos e relações. Considerando que um documento é formado por sentenças, que por sua vez são listas de palavras dispostas segundo regras linguísticas, configura-se um grande desafio a identificação de um único subconjunto de palavras para nomear de forma adequada os elementos centrais dos mapas conceituais. De acordo com as orientações de construção de um mapa conceitual, podemos inferir que Sintagmas Nominais (NP), Sintagmas Verbais (VP) e Sintagmas Prepositivos (PP) são candidatos primários ao mapeamento dos elementos centrais de mapas conceituais. O Sintagma Nominal (NP) é o principal candidato ao conceito, o Sintagma Verbal (SV), por sua vez, é um aspirante a se tornar uma relação e o Sintagma Preposicional (PP) pode ser tanto um conceito já associados com a relação, tomando como essência o Sintagma Nominal precedido de preposição, ou uma relação, se considerarmos verbos ou locuções verbais precedidos de preposição.

Os *chunks* identificados são, portanto, mapeados para nós candidatos a conceitos e relações. O mapeamento de um *chunk* para um conceito ou uma relação dependerá, primordialmente, da possibilidade de estabelecer conexões entre estes elementos na forma de proposições.

3.6 Interpretação de Dependências

Tomando cada um dos candidatos a elementos de mapas conceituais, o interpretador de dependências transforma o candidato a elemento de mapa conceitual em um nó, uma aresta de um grafo, ou um elemento híbrido composto simultaneamente de nó e aresta.

Na atividade, verifica-se o tipo do *chunk*: aqueles com núcleos verbais ou prepositivos são mapeados em arestas e os *chunks* com núcleos nominais são mapeados em nós. À medida que os elementos são transformados em nós ou arestas, o interpretador busca a posição mais adequada no grafo para subsumir o novo elemento. A subsunção ocorre por meio da identificação da proximidade dos nós, unindo aqueles que possuem maior afinidade, de acordo com regras de aproximação pré-definidas.

3.7 Construção de Mapas Conceituais

A última atividade do processo é aquela que consiste na definição de proposições, baseadas em um grafo de entrada, denominada de “Construção de Mapas Conceituais”. Para isso, caminha-se no grafo, de forma a encontrar proposições de mapas conceituais. A cada caminho que parte de um conceito, passando por uma relação que leva a outro conceito, é criada uma proposição. Ao fim da atividade, um conjunto de proposições é formado, determinando a representação do conteúdo original por meio de conexões em um mapa conceitual.

4 Estudo de Caso

Para a validação do processo proposto para a construção de mapas conceituais a partir de textos, a pesquisa realizou um estudo de caso baseado em duas etapas: (i) implementação de um protótipo computacional e (ii) condução de testes por meio do protótipo.

4.1 Protótipo Computacional:

O protótipo computacional desenvolvido foi denominado de *Text2Cmap*. A versão α do *Text2Cmap* é composta por cinco módulos distintos que realizam as atividades do modelo de processo proposto, conforme ilustra a Figura 2.

O protótipo foi codificado e compilado em *Python 2.6.5*, [Python, 2010], usando o módulo de processamento de linguagem natural chamado de *Natural Language Processing Toolkit (NLTK)* [Bird, Klein e Loper, 2009].

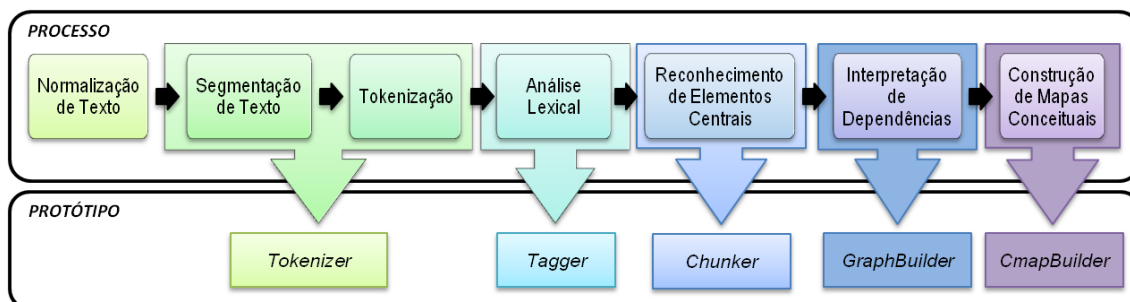


Figura 1. Módulos do Protótipo e Atividades do Processo

O módulo *Tokenizer* realiza as atividades de leitura do documento, segmentação do texto em sentenças e a tokenização, todas por meio de recursos do módulo *NLTK*.

O módulo *Tagger* associa as etiquetas morfológicas para os *tokens* por meio do *tagger* open-source *Freeling* [Atserias, 2006] associado a um arquivo de configurações para a Língua Portuguesa definido por Garcia e Gamallo (2010).

O módulo *Chunker* utiliza técnicas de *chunking* a partir um conjunto de expressões regulares constituídos por meio de etiquetas morfológicas para identificar candidatos a elementos de mapas conceituais, tais como: Sintagmas Nominais (SN), Verbais (SV), Sintagmas Preposicionados (SP), entre outros.

O módulo *GraphBuilder* toma cada um dos candidatos e os transforma em nós de grafos. Por meio de um algoritmo de busca que implementa a heurística *Best-First*, o módulo constrói as arestas entre os dois nós identificando aqueles que se “aproximam” sintaticamente, produzindo um grafo. A proximidade sintática é calculada por meio de pesos binários arbitrados que determinam a possibilidade de ligação de dois nós por meio de uma aresta. A geração de arquivos de grafos compatíveis com *Graphviz* é feita por meio do módulo *GvGen* [GvGen, 2007] e para visualização dos mesmos, usamos formato “*dot*” do *Graphviz* [Graphviz, 2010].

O módulo *CmapBuilder* recebe o grafo cujos nós são candidatos a conceitos e as arestas, candidatas a relações, e realiza uma busca em profundidade. No caminhar pelo grafo, sempre que houver a possibilidade de traçar um caminho de um nó a outro, extrai-se uma proposição. A visualização do conjunto de proposições extraídas é feita por meio do *CmapTools* [Novak e Cañas, 2006].

4.2 Exemplo de Processamento

Nesta Seção, apresentamos algumas saídas decorrentes da execução do *Text2Cmap*. O processamento de um documento qualquer denominado de “[*documento*].txt” leva à produção de um conjunto de saídas, resultados intermediários do processo, enumeradas na Tabela 2.

Tabela 2. Módulos e Saídas Esperadas

Módulo	Saídas
Tokenizer	[documento]_token.txt
Tagger	[documento]_tag.tmp
	[documento]_tag.txt
Chunker	[documento]_chunk.txt
GraphBuilder	[documento]_dot.txt
	[documento].gif
CmapBuilder	[documento]_cmap.txt
	[documento].cmap

É por meio destas saídas que o processo subjacente ao protótipo permite o acompanhamento e a rastreabilidade das ações realizadas. Na Figura 2 estão ilustrados trechos e algumas das saídas produzidas do processo.

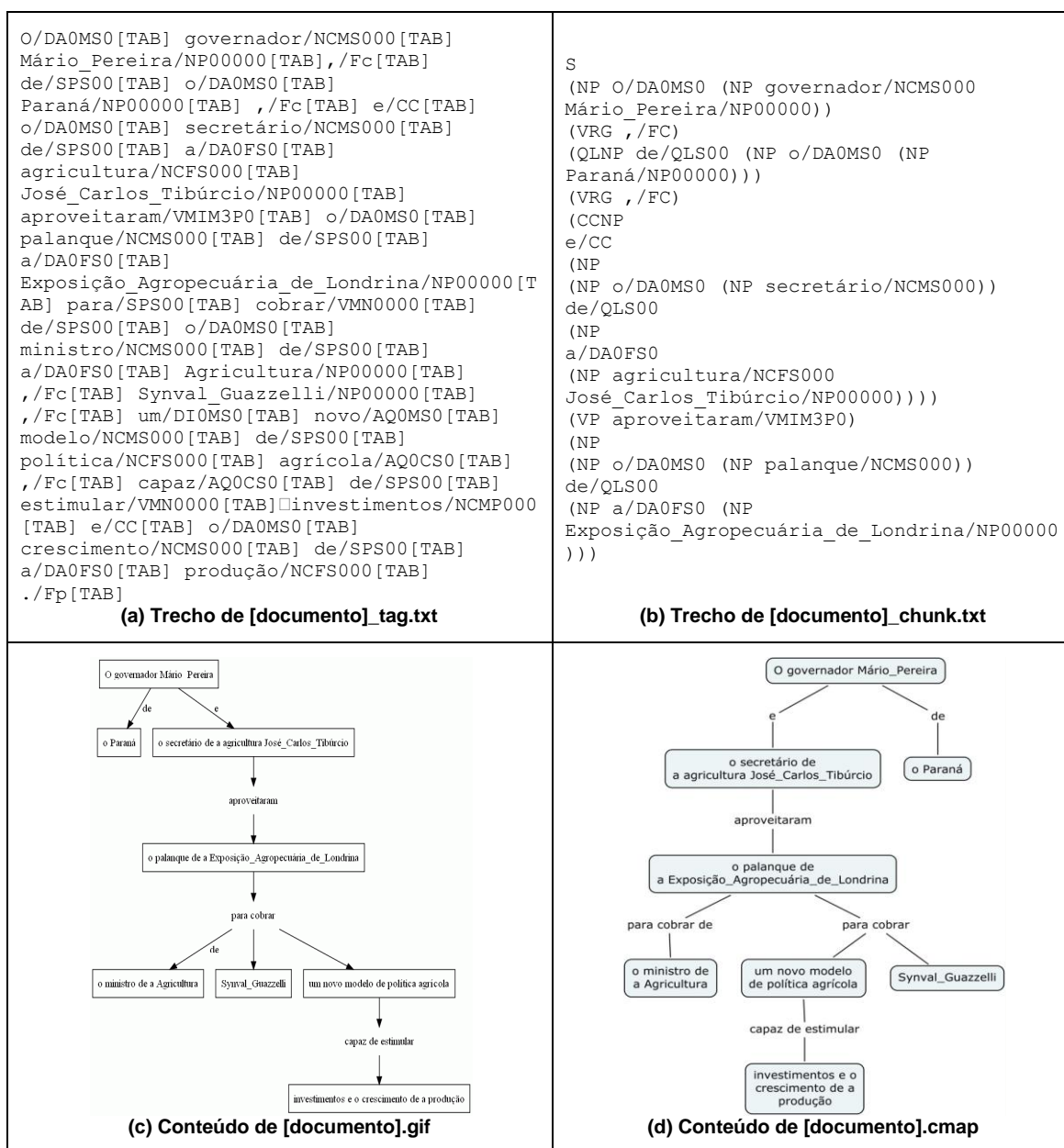


Figura 2. Exemplos de saídas produzidas no processo

4.3 Avaliação com Corpus Linguístico

O Mac-Morpho é um corpus fechado e anotado, em português do Brasil, desenvolvido no contexto do projeto Lácio-Web [Aluisio et al., 2004]. Mac-Morpho contém aproximadamente 1,2 milhões de palavras que foram etiquetadas pelo *parser* Palavras e, posteriormente, cada etiqueta foi mapeada para o elemento correspondente no conjunto de etiquetas do Projeto Lácio-Web. A anotação morfossintática foi validada manualmente. O corpus é composto por 109 arquivos em texto puro, que por sua vez traduzem-se em 51.397 sentenças. Do corpus Mac-Morpho, utilizamos uma amostra, denominada de A1, que contém aproximadamente 2% do volume total dos textos, selecionados de forma aleatória.

4.4 Resultados do Experimento

Considerando exclusivamente o resultado quantitativo do processamento da amostra A1, é possível definir quatro grupos, distintos, conforme ilustrado na Tabela 3:

- i. G1: Documentos que levaram à formação de mapas com aumento da quantidade de *tokens*;
- ii. G2: Documentos que levaram à formação de mapas conceituais **com redução parcial** da quantidade de *tokens*;
- iii. G3: Documentos que levaram à formação de mapas conceituais **sem redução** da quantidade de *tokens*; e
- iv. G4: Documentos que não levaram à formação de mapas conceituais, ocasionado pela **redução total** do quantitativo de *tokens*.

Tabela 3. Formação de clusters de acordo com a quantidade de *tokens* do documento original e do mapa conceitual construído

Grupos		Qtde	Tokens Válidos Média	Chunks							
				Média							
G1	Formação de mapas conceituais COM AUMENTO de quantidade de <i>tokens</i>	173	24,32	Nominais	NP	CCNP	PRNP	CSNP	QLNP	SPNP	Média
									2,73	0,05	0,05
				Verbais	VP	SPVP	CCVP	PRVP	CSV	-	Média
					1,38	0,19	0,06	0,25	0,01		1,89
				Outros	RP	VRG	PNT	DPT	PTV	NA	Média
					0,07	1,82	1,00	0,04	0,00	0,72	3,64
G2	Formação de mapas conceituais COM REDUÇÃO de quantidade de <i>tokens</i>	442	22,51	Nominais	NP	CCNP	PRNP	CSNP	QLNP	SPNP	Média
					2,55	0,09	0,03	0,05	0,24	1,38	4,34
				Verbais	VP	SPVP	CCVP	PRVP	CSV	-	Média
					1,47	0,23	0,08	0,20	0,28		2,27
				Outros	RP	VRG	PNT	DPT	PTV	NA	Média
					0,14	1,45	1,01	0,03	0,00	1,61	4,24
G3	Formação de mapas conceituais SEM ALTERAÇÃO de quantidade de <i>tokens</i>	262	16,29	Nominais	NP	CCNP	PRNP	CSNP	QLNP	SPNP	Média
					2,15	0,05	0,01	0,00	0,11	1,19	3,51
				Verbais	VP	SPVP	CCVP	PRVP	CSV	-	Média
					1,20	0,18	0,08	0,22	0,01		1,69
				Outros	RP	VRG	PNT	DPT	PTV	NA	Média
					0,04	0,75	0,99	0,01	0,00	0,22	2,00
G4	Não houve formação de mapas conceituais	123	9,33	Nominais	NP	CCNP	PRNP	CSNP	QLNP	SPNP	Média
					0,85	0,04	0,02	0,01	0,06	0,46	1,44
				Verbais	VP	SPVP	CCVP	PRVP	CSV	-	Média
					0,94	0,08	0,02	0,05	0,15		1,24
				Outros	RP	VRG	PNT	DPT	PTV	NA	Média
					0,14	0,31	1,30	0,06	0,00	0,43	2,24

Observou-se que documentos com poucos *chunks* verbais e nominais dificultam a criação de proposições, como vistos no grupo G4, que apresenta uma baixa densidade desses elementos e perda total de *tokens* no processo. Neste grupo, com média de *tokens*

reduzidos em comparação aos demais grupos, os *chunks* nominais ocorrem em média, 1,44 para 1,24 verbais. Como a relação mínima para a formação de proposições é de dois conceitos para uma relação, é requerido ao menos dois *chunks* nominais para um verbal.

Nos grupos G1, G2 e G3 houve a formação de mapas conceituais. Notadamente, o grupo G1 foi o que mais produziu proposições em relação aos outros dois: 3,89 comparados aos 2,88 do grupo G2 e aos 2,40 do grupo G3. Coincidentemente, o grupo G1 é o que apresenta a melhor relação entre *chunks* nominais e verbais: 5,01 para 1,88.

5 Considerações Finais

Nesta pesquisa, propusemos uma abordagem computacional para a construção de mapas conceituais a partir de textos em português do Brasil combinada à implementação de um protótipo. O estudo de caso com uma amostra de 1.000 documentos resultou na identificação de quatro grupos distintos de resultados do processamento, utilizando o critério da perda de *tokens* durante o processo: G1 (com aumento de *tokens*), G2 (com redução parcial de *tokens*), G3 (sem redução de *tokens*) e G4 (com redução total de *tokens*).

Observamos que o processo não apresentou perdas semânticas em 40,56% da amostra. Em 9,33% houve perda total do conteúdo durante o processamento. Entre as causas que originaram a perda de elementos no processo estão: falhas do *tagger*, parcela da amostra inadequada para processamento e os erros de *chunker*.

Trabalhos futuros devem considerar a definição de mecanismos mais robustos no tratamento de anáforas, número, gênero, sentenças interrogativas e imperativas, além de refinar heurísticas para definição dos grafos e identificação das proposições. Outro aspecto igualmente relevante é a seleção de textos adequados para a extração de proposições.

6 Referências

- Aluisio, S., Pinheiro, G. M., Manfrim, A. M. P., Oliveira, L. H. M. de., Genoves Jr., L. C., Tagnin, S. E. O. (2004). The Lácio-Web: Corpora and Tools to advance Brazilian Portuguese Language Investigations and Computational Linguistic Tools. In: Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004).
- Atserias, J., Casas, B., Comelles, E., González, M., Padró, L., Padró, M. (2006). FreeLing 1.3: Syntactic and semantic services in an open-source NLP library. In: Proceedings of the fifth international conference on Language Resources and Evaluation (LREC 2006), ELRA. Genoa, Italy. Disponível em <http://www.lsi.upc.edu/~nlp/freeling>.
- Bird, S., Klein, E., Loper, E. (2009). Natural Language Processing with Python – Analyzing Text with the Natural Language Toolkit. CA: O'Reilly Media.
- Chang, T.-H., Tam, H.-P., Lee, C.-H., Sung, Y.-T. (2008). Automatic Concept Map Constructing using top-specific training corpus. In: Proceedings of the Asia-Pacific Educational Research Association Board Meeting (APERA'2008). Singapore.

- Cimiano, P. (2006). *Ontology Learning and Population from Text: Algorithms, Evaluation and Applications*. Springer Science.
- Garcia, M., Gamallo, P. (2010). Using Morphosyntactic Post-processing to Improve POS-tagging Accuracy. In: *Proceedings of the International Conference on Computational Processing of the Portuguese Language (PROPOR'2010)*. Porto Alegre, RS.
- Graphviz 2.26.3 (2010). Graphviz Project Website. Acesso em: 22-03-2010, disponível em http://www.graphviz.org/Download_windows.php.
- GvGen 0.9. (2007). GvGen Project Website. Acesso em: 2010-03-22, disponível em <http://software.inl.fr/trac/wiki/GvGen>.
- Kowata, J. H., Cury, D., Boeres, M. C. S. (2009). Caracterização das Abordagens para Construção de Mapas Conceituais. Paper presented at the XX Brazilian Symposium on Computer in Education (SBIE 2009).
- Lee, C.-H., Lee, G.-G., Leu, Y. (2009). Application of automatically constructed concept map of learning to conceptual diagnosis of e-learning. In: *Expert Systems with Applications*, 36 (2), 1675-1684.
- Mercado, L. P. L. (2002). Formação docente e novas tecnologias. In: *Novas tecnologias na educação: reflexões sobre a prática*. Luís Paulo Leopoldo Mercado (Org.). Maceió: EDUFAL, 210 p.
- Nonaka, I., Takeuchi, H. (1997). *Criação de Conhecimento na Empresa*, Elsevier Editora.
- Novak, J. D.; Gowin, D. B. (1984). *Learning how to learn*. Cambridge University Press, 1984.
- Novak, J. D., Cañas, A. J. (2006). The Origins of the Concept Mapping Tool and the Continuing Evolution of the Tool. *Information Visualization Journal*, 5 (3), 175-184.
- Python 2.6.5. (2010). Python Programming Language Official Website. Acesso em: 2010-03-22, disponível em <http://www.python.org/download/releases/2.6.5/>.
- Silva, E. T.. (2005). Revalorização do livro diante das novas mídias. Veículos e linguagens do mundo contemporâneo: a educação do leitor para as encruzilhadas da mídia. In: *Integração das Tecnologias na Educação. Salto para o Futuro*. Almeida, Maria Elizabeth Bianconcini de., Moran, José Manuel. (Org). 204 p. Disponível em <http://portal.mec.gov.br/seed/arquivos/pdf/1sf.pdf>
- Tseng, S.-S., Sue, P.-C., Su, J.-M., Weng, J.-F., Tsai, W.-N. (2007). A new approach for constructing the concept map. In: *Computers & Education*, 49 (3), 691-707.
- UNESCO, Organização das Nações Unidas para Educação, Ciência e Cultura. (2009). Padrões de Competência em TIC para professores. Marco Político. Disponível em <http://unesdoc.unesco.org/images/0015/001562/156210por.pdf>.
- Valerio, A., Leake, D. (2006). Jump-Starting Concept Map Construction with Knowledge Extracted from Documents. In: A. J. Cañas, J. D. Novak, F. M. González (Ed.), *Proceedings Second International Conference on Concept Mapping (CMC'06)*, 1, pp. 296-303. San José, Costa Rica.