

## Abordagem linguística para identificação da dimensão afetiva expressa em textos de Ambientes Virtuais de Aprendizagem – um Léxico da Emoção

Sandro J. Rigo<sup>1</sup>, Isa M. Alves<sup>2</sup>, Oliver Gazzola<sup>1</sup>, Francini Belau<sup>2</sup>, Jorge L. V. Barbosa<sup>1</sup>, Cristiano Costa<sup>1</sup>

<sup>1</sup>Programa de PósGraduação em Computação Aplicada – PIPCA - UNISINOS  
São Leopoldo – RS – Brazil

<sup>2</sup>Programa de PósGraduação em Linguística Aplicada – PPGLA - UNISINOS  
São Leopoldo – RS – Brazil

{rigo, ialves, jbarbosa, cac}@unisinis.br, {oliver.gazzola, francini.belau}@gmail.com

**Resumo.** *A mediação digital na Educação a Distância dificulta ao professor a percepção do estado de ânimo de alunos. Entretanto, o gênero textual é amplamente estimulado e utilizado neste contexto, fato que tem desencadeado esforços para sua utilização com objetivo de identificar sentimentos expressos pelos alunos. Este trabalho relata uma abordagem utilizando teorias da Semântica Lexical Computacional para representação linguístico-computacional do léxico da emoção. São descritos e analisados estudos realizados na análise de corpus, descrição de unidades lexicais e sua utilização em identificação de sentimentos como apoio à atividade docente. A abordagem adotada possibilita a obtenção de ganhos com relação a outras abordagens similares.*

**Abstract.** *Digital mediation in distance education hinders the teacher's perception of the state of mind of students. However, the text is widely encouraged and used in this context, a fact that has triggered efforts to use it in order to identify feelings expressed by students. This paper reports an approach using theories of Computational Lexical semantics for computational linguistic representation of the lexicon of emotion. Are described and analyzed studies in corpus analysis, description of lexical units and their use in identifying feelings as support for teaching activities. The approach adopted allows gains with respect to other similar approaches.*

### 1. Introdução

A dinâmica descrita pela mediação digital utilizada na Educação a Distância implica na possibilidade de situações que podem se configurar como contextos de isolamento ou de pouca integração. Porém, a utilização dos subsídios gerados neste processo de interação a partir de Ambientes Virtuais de Ensino e Aprendizagem (AVEA) permite a configuração de ferramentas e implementação de recursos capazes de promover relações de maior proximidade. São exemplos destas situações o uso das produções textuais e os dados de interações usados na geração de informações descritoras do perfil dos alunos,

no diagnóstico de suas dificuldades, no apoio aos procedimentos de avaliação e na identificação de aspectos de sua subjetividade no processo de ensino-aprendizagem.

De particular interesse para este trabalho, a identificação do estado emocional de alunos em AVEAs vem sendo estudada como uma possibilidade de ampliação da efetividade e qualidade das interações entre professores e alunos (Bercht 2001; Jaques & Vicari 2005; Desmarais & Baker 2012). De forma muito categórica, estudos sobre a relação entre emoção e memória indicam uma relação positiva entre o estado de ânimo de alunos e atividades cognitivas como memorização e aprendizagem (Le Doux 2001; Pergher *et al.* 2006).

Dado o seu grande e recente desenvolvimento e expansão, sendo que de modo muito particular no contexto da Educação a Distância (EAD), as ferramentas de análise de textos e interações gerados pelos alunos podem contribuir significativamente na identificação de manifestações portadoras de sentimentos, de opiniões ou de relatos de fatos significativos para os alunos e seu processo de aprendizado. A análise de sentimentos com base na expressão textual dos alunos possui iniciativas desenvolvidas neste sentido, que se apresentam com diferentes finalidades e compõem um cenário bastante diversificado. Em alguns casos tratam-se de iniciativas para identificação de estados emocionais dos alunos (Longhi, 2009, Azevedo 2011, Edécio 2011), sendo utilizadas abordagens com base em termos específicos. Em outros casos, trata-se da utilização de ferramentas estatísticas ou de aprendizagem de máquina para identificação de elementos textuais importantes e apoio à produção textual (Macedo 2009; Klemman 2009). Estas iniciativas apresentam resultados interessantes e motivadores, que permitem considerar esta linha de atuação como positiva e com possibilidade de promover a geração de recursos para apoio efetivo aos professores.

Para ampliar o escopo de tratamento dos dados textuais em AVEAs, este trabalho descreve resultados obtidos a partir de uma abordagem de representação linguístico-computacional dedicada à análise de emoções registradas nas diversas ferramentas de interação destes ambientes. Este trabalho é motivado na identificação de melhorias possíveis no tratamento textual considerado a partir dos trabalhos estudados. As abordagens baseadas no tratamento textual sem apoio de elementos linguísticos apresentam erros na identificação de situações envolvendo figuras de linguagem, sinônimos, estilos de escrita diferentes dos previstos e modelados. Uma destas melhorias obtidas no uso de elementos linguísticos está relacionada com a identificação mais precisa dos relatos textuais em linguagem natural.

Este trabalho resulta de uma pesquisa interdisciplinar envolvendo a Semântica Lexical Computacional e recursos de Processamento de Linguagem Natural. No que tange a tarefas de cunho linguístico, tomam-se os princípios da Linguística Cognitiva como base teórica para a descrição do *corpus*. Desse modo, compõem esta descrição informações do tipo relações léxico-conceituais e frames semânticos. Essa organização do léxico integrando camadas com descrição de relações de sentidos e *frames* torna a representação compatível com as propostas das *wordnets*<sup>1</sup> e *framenets*<sup>2</sup>,

---

<sup>1</sup> <http://wordnet.princeton.edu/>

<sup>2</sup> <https://framenet.icsi.berkeley.edu/fndrupal/>

respectivamente, o que facilita futuros estudos multidomínios e multilíngues. Do ponto de vista computacional, os desafios centrais foram o desenho e a validação de uma ferramenta compatível com AVEAs que sirva como um apoio para a identificação e classificação de declarações subjetivas feitas por alunos da EAD.

Os resultados obtidos consistem na realização de estudos linguísticos, tendo como foco tarefas de análise e descrição do corpus coletado, que consiste em mensagens textuais em disciplinas na modalidade de EAD. Com base nestes estudos linguísticos foram definidos elementos de um léxico da emoção. Por sua vez, este conjunto de recursos fornece os subsídios para a ferramenta computacional desenvolvida, o que permitiu a implementação e realização de testes de validação. Destaca-se como principal contribuição deste trabalho a construção de um léxico da emoção e de um sistema computacional que implementa a análise de mensagens textuais a partir de um ponto de vista inovador, incluindo fundamentalmente os recursos linguísticos como determinantes para a geração dos resultados finais de identificação de aspectos afetivos contidos nestas mensagens.

## **2. Aspectos da Emoção e Aprendizagem**

Autores como Picard e Healey (1997), Le Doux (2001) e Pergher et al. (2006), apresentam reflexões que nos são úteis para compreender a importância de mapear emoções e suas relações com aprendizagem. Estas abordagens permitem a atuação no sentido do acompanhamento dos interlocutores em ambientes virtuais (sejam eles consumidores, leitores ou alunos), para ampliar as informações de como eles são, o que eles querem e o que os sensibiliza a refletir, agir e discutir. Logo, surgem elementos essenciais para a reflexão sobre quem são os alunos que chegam ao ensino universitário a distância (ou que podem chegar) e o que eles buscam. Há também aí elementos que auxiliam na compreensão das razões para que certas experiências nos AVEAs não sensibilizem o aluno, não o motivem ou, até mesmo, não facilitem a compreensão dos conceitos trabalhados por determinada atividade acadêmica.

Não há como negar que é essencial que o professor identifique como seus alunos estão se sentindo com os estímulos que têm recebido no AVEA e que busque elementos para compreender esse sentimento não somente a partir de conhecimentos específicos trabalhados pela atividade acadêmica, mas também a partir do conhecimento sobre o perfil desse aluno. Vale destacar que se entende aqui perfil do aluno como algo que antecede o conhecimento sobre o que tradicionalmente se denomina a realidade do aluno, trata-se de uma informação pessoal, mas que tem reflexos culturais e que, portanto, é de certa forma coletiva. Esse tipo de reflexão é aqui empreendido no contexto da EAD, situação em que o contato do professor com o aluno ocorre quase que exclusivamente através de mediação digital, ou seja, a partir do que é registrado no AVEA, em grande parte através de textos.

É nesse cenário que a EAD existe hoje e é nessa cultura que vivem nossos alunos, tenham eles mais ou menos desenvoltura com as ferramentas de interação disponíveis na web. Desse modo, mapear sentimentos torna-se tarefa essencial para a (re)elaboração das atividades do professor. Discutir estratégias de ensino-aprendizagem não é o foco neste trabalho, entretanto, contribuir para a qualificação desse processo está

entre as motivações para a proposição da ferramenta e para a definição do objeto semântico em análise, o léxico da emoção.

### **3. Abordagens relacionadas para Análise de Sentimentos em Textos**

Segundo Pang e Lee, (PANG e LEE, 2008) a utilização da Web com base nas opiniões expressas por pessoas é um fato que vem crescendo de forma bastante evidente. Um dos contextos motivadores para tal está ligado à análise automatizada da reputação de empresas visando o entendimento da aceitação desta e de seus produtos entre seus consumidores (LIU, 2012). No contexto da educação observa-se um processo similar, em parte possibilitado pela ampla adoção de AVEAs, em especial na EAD. Os alunos são estimulados, durante os processos de ensino-aprendizagem, ao uso de diversos mecanismos de interação que permitem a geração de textos. Além das situações de cunho objetivo, onde se observam mensagens textuais relacionadas às respostas de perguntas ou indicação de resultados obtidos, são observadas situações nas quais textos descrevendo subjetividades são gerados. Estes podem corresponder a uma diversidade de situações, variando entre comentários dos alunos sobre seu desempenho, observações pessoais sobre comentários de colegas e professores, indicações sobre o grau de satisfação e aceitação com os processos de ensino, ou a descrição de dificuldades encontradas pelos alunos ao participar deste tipo de modalidade de ensino.

Dado o grande volume de dados neste formato, a realização deste tipo de análise sobre aspectos de cunho subjetivo, quando de forma individualizada e manual, tende a se tornar inviável e custosa. Ao mesmo tempo, o tratamento destes aspectos é de fundamental importância para que se obtenham avanços na utilização do material gerado em interações e expressões dos alunos ou usuários de AVEAs (LONGUI et al, 2009). As técnicas utilizadas para tal objetivo são ainda alvo de pesquisa e experimentação, sendo conhecidas com diversas nomenclaturas, tais como Mineração de Opinião (*Opinion Mining*), Análise de Sentimento (*Sentiment Analysis*), Análise da Subjetividade (*Subjectivity Analysis*) ou Análise de Julgamento (*Appraisal Extraction*) (PANG e LEE, 2008).

Estas técnicas visam, resumidamente, implementar processos que permitem a extração de frases que denotem conteúdo opinativo para que seja possível categorizá-las e, de forma geral, associar a esta categorização atributos de polaridade. Deste modo, são possíveis identificações de textos contendo opiniões positivas ou negativas a respeito de determinados assuntos ou objetos. Além desta, outras abordagens buscam descrever de forma mais específica o tipo de sentimento expresso no texto, sendo que em determinados contextos esta tarefa possui uma complexidade considerável, tal como em textos onde diferentes opiniões são expostas e comparadas, ou então em determinados estilos de redação.

As abordagens conhecidas para tratamento deste problema utilizam etapas bem identificadas, nas quais são realizadas inicialmente operações de identificação de componentes ou sentenças no texto que possuam expressões de subjetividade. Estes resultados são então tratados com etapas de extração de características e classificação, sendo que se observa, em diferentes trabalhos, a experimentação de diversas técnicas aplicadas nestas etapas, de acordo com a finalidade e abordagem em questão. Nas etapas de identificação de expressões e extração de características são empregados diferentes

conjuntos de informações de apoio, tais como léxicos específicos contendo termos e polaridade, conjuntos de expressões de interesse ou a descrição de classes gramaticais dos termos tratados. Em geral estas etapas são tratadas com o apoio de técnicas de pré-processamento que possibilitam o tratamento de *stop-words* e *stemmer*, entre outras características (PANG et al, 2002). Já a etapa de classificação pode ser observada com uma grande variação de técnicas oriundas da área de Inteligência artificial (CHEN, 2011; TANG, 2009).

Algumas abordagens adotam o uso de léxicos descrevendo a polaridade de termos encontrados em textos (LIU, 2012). Apesar dessa abordagem permitir a obtenção de resultados adequados em algumas organizações de texto, em muitas situações observa-se uma interpretação inadequada da informação relacionada com o termo e sua polaridade, sendo que estas situações são associadas com a falta de tratamento linguístico mais amplo.

É consensual em Processamento de Linguagem Natural (PLN) a importância da existência de bases léxicas como a proposta nesse artigo para o processamento automático da língua. Uma das dificuldades para a prática desta abordagem está no custo da construção desse recurso linguístico e na identificação das melhores estratégias teórico-metodológicas para sua construção. No Brasil, em especial, a mineração de sentimentos é tarefa altamente desafiadora tendo em vista a escassez de recursos linguísticos para o português, evidenciando a importância de pesquisa na área e a relevância deste trabalho para a comunidade científica.

#### **4. Metodologia empregada**

A proposta aqui apresentada tem como objetivo facilitar a tarefa de identificação de opiniões, sentimentos e emoções expressos pelos alunos nos diferentes espaços disponíveis em um AVEA. Inicialmente realizou-se a construção de uma base de dados lexicais, contendo informações sintático-semânticas de itens lexicais indicadores de sentimentos e opiniões. Para tal optou-se pela adoção da Semântica Lexical Cognitiva devido ao fato de ser esta uma teoria suficientemente abrangente e com poder descritivo adequado à representação da semântica do léxico da emoção de maneira útil aos fins computacionais. A partir desta opção, a descrição do léxico da emoção permite a representação linguística, pois caracteriza um fenômeno que se manifesta na superfície da língua; possibilita a representação linguístico-computacional, pois permite uma descrição formalizável das informações linguísticas; e, ainda, permite a representação linguístico-cognitiva, pois tem o status de isomorfismo funcional, entendendo-se que a realidade psicológica de uma representação não é imediata, mas pode ser representativa do que temos na mente (Dominiek e Rice 1995).

Os textos analisados para a construção deste léxico foram obtidos a partir da consulta a dados reais gerados por alunos de cursos na modalidade de EAD em uma ferramenta de avaliação do curso e do processo de aprendizagem do aluno. Para a análise dos textos as seguintes atividades são realizadas: (i) compilação de corpus com textos extraídos do AVEA Moodle; (ii) pré-processamento dos textos, envolvendo a conversão para linguagem compatível com as ferramentas de gerenciamento de corpora; a extração de substantivos, adjetivos e verbos a partir do parser PALAVRAS (Bick, 2000); o uso de concordanceador para a manipulação do corpus, especificamente, o

Word Smith Tools; (iv) descrição das unidades lexicais que comporão a base de dados lexicais.

Esse léxico consiste no componente fundamental da ferramenta computacional dedicada a mapear, analisar automaticamente e disponibilizar ao professor um relatório das emoções e sentimentos expressos por seus alunos. O presente trabalho tomou como orientação a exploração dos recursos das bases lexicais construídas, porém de forma a incluir no tratamento computacional aspectos mais abrangentes, tais como algumas informações contextuais, além das informações de polaridade utilizadas em grande parte dos trabalhos estudados. Este aspecto está sendo explorado com o objetivo de verificar a possibilidade de diferenciação, com a utilização de regras sintáticas em conjunto com as informações de polaridade. Acredita-se que esta composição possibilita a melhoria de alguns resultados, evitando situações específicas, quando os termos de polaridade estão associados com componentes modificadores. Para tal encaminhamento ser validado em uma abordagem exploratória, foram implementados os componentes descritos a seguir.

Em primeiro lugar, foram delimitados componentes para possibilitar a comparação de resultados entre abordagens baseadas em informações de polaridade descritas em um léxico e a abordagem aqui proposta, utilizando componentes adicionais tais como regras sintáticas no processo de análise dos documentos. Para tal foi escolhido o SentiLex-PT3, que consiste de um léxico de sentimentos para o português. Ele é composto por lemas adjetivais e formas flexionadas. No momento da consulta e utilização, esta base continha 6321 lemas adjetivais e 25406 formas flexionadas. Cada entrada deste léxico apresenta informações sobre polaridade e alvo deste sentimento, além do método de identificação de polaridade utilizado. Esta última informação é importante, pois parte deste léxico é anotada manualmente, porém parte possui seus atributos gerados automaticamente (SILVA, 2010).

Além deste componente, foram selecionados textos originados em um AVEA (MOODLE) que estão relacionados com expressão de sentimentos de alunos em relação ao seu desempenho em disciplinas em EAD. Estes textos, para utilização de regras sintáticas, são tratados pelo parser PALAVRAS (Bick, 2000) que possibilita a obtenção de documentos com as informações sintáticas dos textos analisados. Estes conjuntos são finalmente tratados com um léxico desenvolvido dentro do contexto deste trabalho, a partir do qual são descritos elementos tais como adjetivos com potencial importante para a identificação de sentimentos, associados por sua vez com regras que possibilitam uma identificação mais precisa dos contextos de sua ocorrência.

Este conjunto de recursos foi organizado a partir de uma arquitetura geral para que pudessem ser realizados os experimentos iniciais de validação da abordagem. As regras sintáticas consideradas são geradas pela etapa de análise linguística e análise linguística-computacional, sendo este um trabalho ainda em andamento. As regras foram identificadas para que alguns componentes textuais, tais como os adjetivos, por exemplo, fossem tratados adequadamente nas etapas de identificação e classificação. Os exemplos que concretizam estas duas regras podem ser vistos nas frases abaixo, na Tabela 1. Nestas situações, pode ser destacado o potencial de associação entre o adjetivo e o substantivo das frases. Em algumas situações em que o contexto das frases

---

<sup>3</sup> [http://dmir.inesc-id.pt/reaction/SentiLex-PT\\_01](http://dmir.inesc-id.pt/reaction/SentiLex-PT_01)

apresenta-se com maior complexidade é possível configurar regras com os componentes adequados para esta análise.

Tabela 1 – Exemplos de regras sintáticas e frases associadas

(DET) + N + VCOP + ADJ	A disciplina é excelente.
(DET) + N + VCOP + ADV + ADJ	O material foi muito bom.
(DET) + N + ADV NEGAÇÃO + VCOP + (ADV) + ADJ/NOME	O material não é claro.
DET + N + QUE + VCOP + (ADV) + ADJ	O professor, que foi muito bacana durante o semestre, recebeu uma homenagem.

Por sua vez, a utilização das regras se baseia na análise sintática realizada pelo parser PALAVRAS, que permite que sejam identificados corretamente os componentes de cada oração. Como o resultado da operação deste parser é disponibilizado em um arquivo no formato XML, torna-se viável o seu processamento automático, como base para a implementação de algoritmos de localização e aplicação de regras sintáticas. A aplicação desenvolvida para testes desta abordagem foi desenvolvida em linguagem PHP, utilizando o framework MVC (*Model-View-Controller*) de alto desempenho designado Yii (Yes, It Is!)<sup>4</sup> sobre o servidor Apache, utilizando banco de dados MySQL. O funcionamento do sistema desenvolvido está intimamente ligado ao tratamento das regras cadastradas, de modo a identificar a informação correta, sobre os textos analisados, de acordo com as informações linguísticas cadastradas no ambiente. Para tal foi definido no ambiente uma interface de cadastramento de regras que possibilita a indicação dos elementos que as compõe e também a indicação de atributos para maior flexibilidade, no caso de indicação de itens opcionais e de itens que identificam inversão de polaridade do termo.

O corpus final contém as informações a seguir: (a) 376 trechos classificados conforme os temas: autoavaliação, avaliação das ferramentas, avaliação do Moodle, avaliação do chat, avaliação do professor, avaliação da disciplina, expectativas (do aluno), avaliação da modalidade de ensino e avaliação do grupo (colegas, participantes da comunidade); (b) identificação de 219 adjetivos; após a classificação como qualificadores ou classificadores, 122 foram considerados na pesquisa e classificados em positivos, negativos e objetivos e 97 foram excluídos da análise.

## 5. Estudo de caso, resultados e avaliação

A implementação do sistema descrito e a construção do léxico da emoção proporcionou o ambiente para a realização de testes exploratórios para apoiar a validade da utilização de regras com base linguística como elementos de apoio na análise de sentimentos em textos. A seguir estão descritos alguns dos resultados obtidos, com base na experimentação de análise de documentos coletados no AVEA MOODLE, descrevendo a avaliação de alunos de cursos de graduação na modalidade de EAD, para disciplinas cursadas. Foram identificadas e comparadas especificamente as polaridades e sentimentos obtidas com as duas abordagens: (a) com uso de regras sintáticas e do léxico da emoção; (b) com a análise baseada na polaridade descrita pelo léxico Sentilex-

<sup>4</sup> <http://www.yiiframework.com/>

PT. Deste modo, alguns resultados são descritos a seguir, na figura 1. Eles descrevem inicialmente uma situação quanto à proporcionalidade dos resultados.

	Polaridade Adjetivos		Regras	
	Ocorrências	%	Ocorrências	%
<b>Positivo</b>	246	42,34	92	74,19
<b>Negativo</b>	58	9,98	22	17,74
<b>Neutro</b>	142	24,44	8	6,45
<b>Não classificado</b>	135	23,23	2	1,61
<b>Total</b>	581	100	124	100

Figura 1 - Comparação entre abordagens utilizadas

Observando os dados acima, na figura 1, observa-se que a análise por regras tem um índice maior de carga opinativa polarizada, ou seja, a porcentagem de ocorrências positivas e negativas é de 91,93% contra 52,32% na análise por polaridade de adjetivos (com uso do SentiLex-PT). Este resultado pode ser considerado como um indicativo que um conjunto de regras capazes de detectar assinaturas de qualificadores em uma sentença são mais efetivas que a simples e pura análise dos adjetivos.

Um outro cenário de teste foi desenhado para visualizar melhor as diferenças de resultados entre as análises por polaridade e por regras. Foram verificadas apenas as sentenças detectadas por regras (primeira linha na Figura 1) e os resultados foram comparados com os obtidos apenas com a polaridade dos adjetivos para estas mesmas sentenças detectadas nas regras (segunda linha na Figura 1). Como observado, houve apenas a diferença de um resultado positivo a menos, e conseqüentemente, um negativo a mais, na análise por regras. Para possibilitar uma visualização mais detalhada da diferença entre as duas análises, foi necessária a verificação de todos os casos por um especialista, para determinar quantas análises divergiram, conforme demonstrado a seguir na análise detalhada de um exemplo de sentença que diverge no cálculo de polaridade ao utilizar-se regras. A análise proporcional dos resultados mostrou que houve divergência no cálculo de polaridades de pelo menos uma sentença do corpus escolhido. Observando detalhadamente cada sentença analisada, constatou-se que 43 sentenças tiveram cálculos divergentes, 2 sentenças consideradas positivas na análise por adjetivos foram classificadas como negativas na análise por regras, enquanto que uma sentença considerada negativa na análise por adjetivos foi considerada positiva na análise por regras. Como pode ser observado na figura 2 abaixo, a análise por regras gerou o resultado considerado correto, considerando o teor dos comentários expressado nas sentenças. Observa-se que a coluna “Sentilex” identifica a polaridade negativa (-1) associada com o adjetivo “difícil”, enquanto que a regra “R-2” descreve a polaridade corretamente, como positiva.

Id	Word	Lemma	Pos	Sentilex	Match	Polarity
304_1	vi	ver	v-fin			
304_2	que	que	conj-s			
304_3	não	não	adv		R-2 - Advérbio (não) [O] [PI] + Verbo + Advérbio + Adjetivo	1
304_4	era	ser	v-fin		R-2 - Advérbio (não) [O] [PI] + Verbo + Advérbio + Adjetivo	1
304_5	tão	tão	adv		R-2 - Advérbio (não) [O] [PI] + Verbo + Advérbio + Adjetivo	1
304_6	difícil	difícil	adj	-1	R-2 - Advérbio (não) [O] [PI] + Verbo + Advérbio + Adjetivo	1
304_7	de	de	prp			
304_8	acompanhar	acompanhar	v-inf			
304_9	a	o	art			
304_10	distância	dist=ncia	n			



Figura 2 - Detalhamento da análise de regras e léxico

Deste modo, foi possível verificar que a utilização das classificações baseadas em regras, de certo modo, restringe o número de ocorrências encontradas para o universo de sentenças que são identificadas com o padrão cadastrado, mas também aumenta a confiabilidade dos resultados obtidos e permitem um cálculo de polaridade mais coerente em relação à técnicas de verificação baseadas apenas nos adjetivos.

## 6. Conclusões

O principal problema abordado neste trabalho é, do ponto de vista linguístico e linguístico-computacional, a construção de uma base de dados lexicais do domínio da emoção construída a partir de itens lexicais extraídos de textos de alunos postados nos diferentes recursos do ambiente Moodle. As discussões teóricas empreendidas para a construção da base léxica contribuirão para o avanço dos estudos sobre a semântica das emoções e, além disso, servirão como referência para a organização do léxico da emoção – no contexto da EAD - em português. Esta é uma área carente de recursos desenvolvidos para o português, sendo que esta consiste em uma contribuição destacada para este trabalho.

De forma teórica, alguns problemas de pesquisa foram enfatizados neste trabalho. Sob o viés linguístico e linguístico-computacional, a descrição da semântica das emoções com vistas à construção de uma base de dados lexicais contendo informações sintático-semânticas do léxico da emoção. Sob o viés linguístico-computacional e computacional, o estudo sobre técnicas de mineração de sentimento e estratégias para a implementação de uma ferramenta com esse fim. Por fim, destaca-se o estudo sobre a relevância da identificação e tratamento de aspectos afetivos nas mensagens textuais em ambientes voltados ao contexto da EAD.

A ferramenta desenvolvida apresentou resultados adequados, conforme apresentados na seção anterior, tendo sido concebida para oferecer ao professor informações com diferentes graus de interferência automática, com a realização automática do reconhecimento de expressões de emoção registradas no ambiente e a disponibilização dessa informação ao professor através de uma interface específica a partir da qual o professor possa acessar o léxico enriquecido com informações semânticas e o contexto original da fala do aluno.

Como trabalhos futuros, são identificadas melhorias na ferramenta, para que seja possível o reconhecimento automático da existência de relatos de satisfação, a classificação dessa opinião de acordo com a descrição semântica codificada no léxico que compõe o sistema e a disponibilização dessas informações ao professor, acompanhadas de estatísticas sobre a frequência de cada tipo de relato, além da disponibilização de acesso rápido ao contexto original da fala do aluno.

## 7. Referências

- AZEVEDO, B. F. T. (2011) MINERAFÓRUM: Um recurso de apoio para análise qualitativa em fóruns de discussão. Programa de Pós-Graduação em Informática na Educação. UFRGS. Tese de Doutorado.
- Bercht, M.. (2001) “Em Direção a Agentes Pedagógicos com Dimensões Afetivas”. Instituto de Informática. UFRGS. Tese de Doutorado.

- Bick, E. 2000. *The Parsing System Palavras: Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework*, Aarhus University Press.
- CHEN, L. S.; LIU, C. H.; CHIU, H. J. A neural network based approach for sentiment classification in the blogosphere. *Journal of Informetrics*, [S.1], v.5, n.2, p.313-322, 2011.
- Desmarais, M.C., Baker, R.S.J.d. (2012) A Review of Recent Advances in Learner and Skill Modeling in Intelligent Learning Environments. *User Modeling and User-Adapted Interaction*, 22 (1-2), 9-38.
- Dominiek, S.Rice, S.1995 Network analyses of prepositional meaning: mirroring whose mind – the linguist’s or the language user’s? *Cognitive Linguistics* 6(1):89–130
- Edécio, E. F.; Bercht, M.; Reategui, E.. *Deteção e Tratamento do Estado Afetivo Frustração do Aluno na Disciplina de Algoritmos*. Anais do XXII SBIE - XVII WIE. 2011.
- Jaques, P. A., & Vicari, R. (2005) “Considering Student's Emotions in Computer Mediated Learning Environments”. In: Z. Ma (Ed). *Web-based Intelligent e-Learning Systems: Technologies and Applications* (pp.122-138). Hershey, PA: Information Science Publishing.
- KLEMMANN, M.; LORENZATTI, A.; REATEGUI, E. O Emprego da Ferramenta de Mineração de Textos SOBEK como Apoio à Produção Textual. In: *Simpósio Brasileiro de Informática na Educação, 2009, Florianópolis*. Anais do SBIE, 2009.
- LIU, B. *Sentiment Analysis and Opinion Mining*. Morgan and Claypool Publishers, 2012. 165p.
- Le Doux, J. (2001). *o cérebro emocional: os misteriosos alicerces da via emocional*. Rio de Janeiro: Objetiva.
- LONGHI, M. T.; BEHAR, P. A. e BERCHT, M.. O Desafio de Reconhecer a Dimensão Afetiva em Ambientes Virtuais de Aprendizagem. In: *Anais do SBIE2008 - XIX Simpósio Brasileiro de Informática na Educação, Fortaleza, CE*. Novembro, 2008.
- Macedo, A. L.; Reategui, E.; Lorenzatti, A.; Behar, P. Using Text-Mining to Support the Evaluation of Texts Produced Collaboratively. *Education and Technology for a Better World, IFIP Advances in Information and Communication Technology, Volume 302*. ISBN 978-3-642-03114-4. Springer Berlin Heidelberg, 2009, p. 368.
- Oliveira Jr. R.; et al. Uma Ferramenta de Monitoramento Automático de Mensagens de Fóruns em Ambientes Virtuais de Aprendizagem. *Anais do XXII SBIE - XVII WIE*. 2011.
- PANG, B. et al. Thumbs up? Sentiment Classification using Machine Learning Techniques. In: *Conference on empirical methods in natural language processing*. 2002. *Proceedings...[S.l.:s.n.]*, 2002, p.78-86.
- Picard, R. W. Healey J.,(1997), "Digital Processing of Affective Signals,"*Proceedings of ICASSP*, May 12-15, 1998, Seattle, WA
- PERGHER, G. K. et al. Memória, humor e emoção. *Revista de Psiquiatria do Rio Grande do Sul, Porto Alegre*, v. 28, n. 1, p. 61.68, jan./abr. 2006.
- SILVA, M. *Automatic Expansion of a Social Judgment Lexicon for Sentiment Analysis*. University of Lisbon. [S.l.]. 2010.
- TANG, H. TAN, S.; CHENG, X. A survey on sentiment detection of reviews. *Expert Systems Applications*. NY, USA, p.10760-10773, September, 2009.