

UNIVERSIDADE DO VALE DO RIO DOS SINOS

Programa Interdisciplinar de Pós-Graduação em
Computação Aplicada

Paulo Roberto Pasqualotti

**Reconhecimento de expressões de emoções na
interação mediada por computador**



Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

UNIVERSIDADE DO VALE DO RIO DOS SINOS
CIÊNCIAS EXATAS E TECNOLÓGICAS
PROGRAMA INTERDISCIPLINAR DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO
APLICADA – PIPCA

Paulo Roberto Pasqualotti

Reconhecimento de expressões de emoções na interação mediada por computador

Dissertação de Mestrado submetida à avaliação como
requisito parcial para obtenção do grau de Mestre em
Computação Aplicada

Orientadora:

Prof^ª. Dr^ª. Renata Vieira

São Leopoldo

2008

Dados Internacionais de Catalogação na Publicação

P284r Pasqualotti, Paulo Roberto

Reconhecimento de expressões de emoções na
interação mediada por computador / Paulo Roberto
Pasqualotti. – São Leopoldo, 2008.

94 p.

Dissertação (Mestrado) – Universidade do Vale
do Rio dos Sinos. Programa Interdisciplinar de Pós-
graduação em Computação Aplicada – PIPCA, São
Leopoldo, 2008.

Orientadora: Prof^a Dr^a. Renata Vieira

1. Inteligência artificial. 2. Computação –
Linguagem natural. I. Vieira, Renata. II. Título.

CDU 004.8:800.8

Bibliotecário responsável:
Silvane Ruas Manhago
CRB 10/1783

Dedico este trabalho a todas as pessoas que, de alguma forma, contribuíram para a minha caminhada e que sempre estiveram presentes com palavras, com energia positiva, em sintonia, no pensamento e oração.

Agradecimentos

À minha orientadora, Prof^ª. Dr^ª. Renata Vieira, pela dedicação, comprometimento, confiança e pelo exemplo como profissional, demonstrando, sempre, conhecimento e atitude de educadora.

À Prof^ª. Dr^ª. Patrícia Augustin Jaques, pelo apoio, incentivo e orientações sobre os assuntos envolvendo minha dissertação, assim como sua prestatividade e comprometimento com a qualidade do trabalho.

À minha esposa Fátima, pelos incontáveis momentos de apoio e incentivo para que eu conseguisse realizar meus sonhos, estando ao meu lado em cada momento do meu mestrado.

Ao Pablo e Pietro, meus filhos amados, pelo fato de existirem, pela paciência e por entenderem o porquê das minhas intermináveis horas de estudo ao invés do futebol.

Aos meus pais, Deoclécio e Jandira, pois esse foi mais um passo na minha vida e só foi possível porque estiveram presentes e foram os responsáveis pelos meus primeiros e principais passos.

Aos meus irmãos, sobrinhos, cunhadas e todos aqueles que, mesmo na distância, torceram e rezaram pelos meus estudos.

Às tradutoras Caroline e Lílian, pela demonstração de responsabilidade e comprometimento com o trabalho de tradução e validação da base de dados.

Aos meus colegas do laboratório de Engenharia da Linguagem, em especial os colegas Luiz Carlos e Patrícia, pela amizade e pela ajuda, iniciada no nivelamento e que se estendeu por durante esses dois anos de muito estudo.

À direção e colegas do Colégio Mãe de Deus, em especial à Irmã Elenar, que sempre compreendeu a importância do mestrado na minha caminhada como educador e me incentivou a seguir em frente.

Aos colegas da Caixa Federal da Gerência de Filial do FGTS, em especial à Gerente de Filial Angela Pianca Scangarelli, pelo apoio. Aos colegas da Gerência das Ações Judiciais,

especialmente ao Gerente de Serviços Jose Carlos Friedrich, pelo incentivo e compreensão. A todos os demais colegas pelo carinho, apoio e incentivo e, também, aqueles que prontamente responderam ao formulário da pesquisa realizada neste trabalho.

Aos colegas, professores e alunos do curso de Licenciatura em Computação, Ciência da Computação e Sistemas de Informação da Feevale, pelo incentivo e demonstração de carinho frente a esse desafio.

Aos demais colegas do Mestrado, pelo companheirismo, troca de idéias e ajuda nos momentos de encontros nos laboratórios e em sala de aula.

Ao corpo docente do PIPCA pela disposição; aos administradores de rede, em especial ao Fábio, pela prestatividade. À secretária Sandra, pelas informações e pelo atendimento dedicado e atencioso nos momentos de dúvidas.

À CAPES, pelo auxílio financeiro durante o transcorrer do Mestrado, custeando e oportunizando-me os estudos.

À força divina que, certamente, esteve presente nos momentos de angústia, pressão, cansaço e intermináveis madrugadas e finais de semana. Pela sua luz, que sempre ajudava a resolver problemas, decifrar relatórios e encontrar as melhores palavras.

E àqueles que, mesmo no pensamento, estiveram comigo e me ajudaram a acreditar que era possível.

Resumo

No Processamento da Linguagem Natural (PLN) os avanços recentes na análise automática de texto conduziram ao surgimento de uma nova área que é o tratamento e reconhecimento dos aspectos subjetivos no texto, considerando fatores como opiniões, sentimentos, emoções e afetividade.

Entre as aplicações dessa área podemos citar: medir o nível de satisfação de clientes, reconhecer a diferença entre emissão de opiniões e descrição de fatos, reconhecer perfil empreendedor e posturas interpessoais, identificar possibilidades de estados depressivos em pessoas, entre outras possibilidades, envolvendo a subjetividade.

A pesquisa na área requer o desenvolvimento de métodos, recursos e técnicas, que, integrados, possibilitarão a sistemas computacionais serem capazes de manipular significado afetivo no discurso. Porém, por tratar-se de estudos recentes, os recursos ainda são bastante escassos para a língua portuguesa.

Dessa forma, essa dissertação propõe o desenvolvimento de um léxico afetivo para a língua portuguesa, denominado de WordNet Affect BR. Essa base foi construída como um recurso para aplicações que requerem um vocabulário de palavras de emoções, como um movimento inicial ao estudo do reconhecimento da subjetividade em discurso da língua portuguesa. Além do desenvolvimento da base, este trabalho apresenta uma ferramenta de chat, denominada de “Emoticon”, desenvolvida com o objetivo de utilizar a base lexical WordNet Affect BR e validar seu conteúdo como um recurso lingüístico adequado ao reconhecimento das palavras de emoções.

Abstract

In the area of Natural Language Processing (NLP) the recent advances in text automatic analysis have led to the emergence of a new area which is the management and recognition subjective aspects in texts, considering factors such as opinion, sentiments, emotions and affection.

Among the applications of this area we enumerate: to determine the client satisfaction level, to recognize the difference between opinion emission and description of facts, to recognize entrepreneur profile and interpersonal posture, to identify the possibility of depressive states in people, among other possibilities related to subjectivity.

The research in this area requires the development of methods, resources and techniques, that, integrated, can make computational systems able to manage the affective meaning in discourse. However, since these studies are recent, the resources are scarce for the portuguese language.

In this way, this dissertation proposes the development of an affective lexicon for the portuguese language, named WordNet Affect BR. This database was developed as a resource for applications that demand a vocabulary of emotion, as a initial step to the study of the recognition of subjectivity in portuguese language discourse. Besides the development of the database, this work presented a chat tool, named “Emoticon” which was developed in order to validate the content of the WordNet Affect BR database as a linguistic resource proper for the recognition of words related to emotions.

Lista de figuras

Figura 2.1.2.1: Modelo OCC	19
Figura 2.2.3.1: Tela on-line da WordNet com o resultado da palavra "happy"	25
Figura 2.2.5.1: Estrutura da base Affect	27
Figura 3.1.1: Modelo ER do relacionamento "synset_tem_gloss" da WordNet	34
Figura 3.1.2: Modelo ER dos relacionamentos entre os synsets da WordNet	35
Figura 3.1.3: Interface de consulta para a emoção "happy"	36
Figura 3.1.4: Interface do resultado da consulta da palavra "happy"	37
Figura 3.2.1: Estrutura do trabalho de criação da base WordNet Affect BR	38
Figura 3.3.1: Interface de tradução das palavras para a emoção "distress"	41
Figura 3.3.2: Interface de tradução da palavra e do glossário	42
Figura 3.3.3: Modelo ER das relações entre a WordNet Affect BR e a WordNet	42
Figura 3.4.1: Resultado quanto à concordância das fases de tradução e validação	44
Figura 4.1: Formulário de validação dos grupos de palavras e imagens	51
Figura 4.1.1: Gráfico da distribuição dos participantes por sexo	52
Figura 4.1.2: Gráfico da distribuição dos participantes por faixa etária	52
Figura 4.1.3: Gráfico com percentual das respostas dos usuários	53
Figura 5.1: Modelo da ferramenta	59
Figura 5.1.1: Estrutura da tabela "chatscript"	61
Figura 5.1.2: Tela de login da ferramenta de chat	62
Figura 5.1.3: Interface da tela da ferramenta de chat	63
Figura 5.2.1: Modelagem da relação entre as tabelas de emoções e de imagens	66
Figura 5.2.2: Imagens relacionadas à emoção raiva	67
Figura 5.4.1: Simulador para reconhecer e representar imagens de emoção	69
Figura 6.2.1: Tela do diálogo da turma de graduação	73

Lista de tabelas

Tabela 2.2.2.1: A estrutura da WordNet com relação às classes gramaticais	23
Tabela 2.2.5.1: Classificação com palavras baseada em Ortony e colegas	28
Tabela 2.2.6.1: Lista de "a-labels" com os respectivos estados afetivos e exemplos	29
Tabela 2.2.6.2: Novos synsets obtidos a partir das relações da WordNet	29
Tabela 2.2.6.3: Estrutura da WordNet Affect e exemplo de registros da base	30
Tabela 3.2.1: Emoções do modelo OCC com a tradução para a língua portuguesa	39
Tabela 3.2.2: Estrutura proposta para os níveis da base WordNet Affect BR	39
Tabela 3.3.1: Distribuição das palavras para a tradução	43
Tabela 3.4.1: Resultado quantitativo da fase de validação da tradução das palavras	45
Tabela 3.4.2: Exemplos de resultados quanto ao consenso para validar a tradução	45
Tabela 3.5.1: Tradução de "relief", com número do sentido e glossário	46
Tabela 3.5.2: Multi-palavras da WordNet Affect BR	48
Tabela 4.1: "Emoticons" representando uma emoção	49
Tabela 4.1.1: Faixa etária dos participantes	52
Tabela 4.1.2: Resultados quanto às respostas do formulário	53
Tabela 4.1.3: Grupo 10 com maior percentual de concordância	54
Tabela 4.1.4: Grupo 3 com menor percentual de concordância	55
Tabela 4.1.5: Registros dos usuários quanto às imagens de emoção	57
Tabela 5.1.1: Campos da tabela "chatscript"	61
Tabela 5.2.1: Etapas de tratamento das mensagens do diálogo	64
Tabela 5.2.2: Exemplo da emoção "amor" e possibilidades para estender seu uso	66
Tabela 6.2.1: Resultados da avaliação da ferramenta de chat	73
Tabela 6.3.1: Resultado do processamento da valência de um log do chat(1)	74
Tabela 6.3.2: Resultado do processamento da valência de um texto aleatório(2)	75

Lista de abreviaturas

AR	Affective Reasoner
CAT	Categoria Afetiva
CSS	Cascading Style Sheets
EMO	Emoção
ER	Entidade-Relacionamento
HTML	Hiper-Text Markup Language
IA	Inteligência Artificial
ID	Identificação Do Usuário
IHC	Interação Homem-Computador
ITC-IRST	Instituto Trentino de Cultura
LIBRAS	Língua Brasileira de Sinais
LOG	Relatório em arquivo ou impresso
OCC	Modelo definido por Ortony, Collins E Clore
PHP	Hypertext Preprocessor: Linguagem de Programação da Web
PLN	Processamento da Linguagem Natural
POS	Part Of Speech
SQL	Structured Query Language
TCC	The cognitive and Communication Technologies
TICS	Tecnologias da Informação e da Comunicação
WEB	World Wide Web ou Www: Internet Gráfica

Sumário

1. INTRODUÇÃO	12
1.1. MOTIVAÇÃO	12
1.2. OBJETIVOS	14
1.2.1. <i>Objetivo Geral</i>	14
1.2.2. <i>Objetivos específicos</i>	14
1.3. ORGANIZAÇÃO DO TEXTO.....	15
2. FUNDAMENTOS TEÓRICOS.....	16
2.1. ESTADOS AFETIVOS E EMOÇÕES	16
2.1.1. <i>Computação Afetiva</i>	18
2.1.2. <i>Modelo cognitivo de emoções OCC</i>	19
2.2. BASE LEXICAL AFETIVA.....	22
2.2.1. <i>WordNet</i>	22
2.2.2. <i>Synsets</i>	23
2.2.3. <i>Relações da WordNet</i>	23
2.2.4. <i>WordNet Domains</i>	25
2.2.5. <i>Base Affect</i>	26
2.2.6. <i>WordNet Affect</i>	28
2.3. TRABALHOS RELACIONADOS	30
3. CRIAÇÃO DA BASE LEXICAL DE EMOÇÕES - WORDNET AFFECT BR.....	34
3.1. INFORMAÇÕES DAS BASES WORDNET E WORDNET AFFECT.....	34
3.2. BASE LEXICAL DE EMOÇÕES NA LÍNGUA PORTUGUESA	37
3.3. TRADUÇÃO DAS PALAVRAS DA BASE	40
3.4. VALIDAÇÃO DA TRADUÇÃO DAS PALAVRAS.....	44
3.5. COMPOSIÇÃO FINAL DA BASE WORDNET AFFECT BR	46
4. IMAGENS PARA A REPRESENTAÇÃO DE EMOÇÕES	49
4.1. FORMULÁRIO DE EXPRESSÃO VISUAL DE EMOÇÕES.....	51
5. FERRAMENTA	59
5.1. A FERRAMENTA DE CHAT “EMOTICON”	60
5.2. MÓDULO DE RECONHECIMENTO DE PALAVRAS DE EMOÇÕES	64
5.3. MÓDULO DE REPRESENTAÇÃO VISUAL DAS PALAVRAS DE EMOÇÕES	67
5.4. MÓDULO DE ADMINISTRAÇÃO DA FERRAMENTA.....	67
6. EXPERIMENTOS.....	70
6.1. PERFIL DOS USUÁRIOS.....	71
6.2. RESULTADOS E CONSIDERAÇÕES	71
6.3. VALÊNCIA.....	73
7. CONSIDERAÇÕES FINAIS.....	76
7.1. TRABALHOS FUTUROS.....	77
REFERÊNCIAS	79
ANEXO A - ARQUIVOS DA WORDNET	83
ANEXO B - RELAÇÕES ENTRE SYNSETS DA WORDNET.	87
ANEXO C – A BASE WORDNET AFFECT BR.....	88

ANEXO D – GRUPOS DE PALAVRAS E IMAGENS DE EMOÇÃO	90
ANEXO E – AUTORIZAÇÃO PARA USO DA WORDNET AFFECT	92

1. Introdução

1.1. Motivação

Diversas áreas do conhecimento dedicam-se a estudar e propor técnicas e métodos de enriquecimento do ato comunicativo com o objetivo de ultrapassar dificuldades e necessidades dos interlocutores. Entre as dificuldades, podemos citar diferenças idiomáticas, necessidades especiais como surdez, baixa ou nenhuma visão, entre outras, que podem ser contempladas por mecanismos técnicos ou recursos computacionais que ajudem na solução de tais dificuldades. Na área da computação, da eletrônica e das tecnologias da informação e comunicação (TICs), muitos trabalhos têm sido apresentados com a intenção de mostrar soluções para as dificuldades apontadas acima, como por exemplo, leitores de tela para pessoas com dificuldades de visão, equipamentos e próteses que respondem a impulsos cerebrais para pessoas com dificuldades motoras, sistemas de tradução automática, sintetizadores de voz, entre outras. Dessa forma, a computação tem propiciado o desenvolvimento de recursos que permitem a dispositivos tecnológicos atender às finalidades adequadas ao processo comunicativo.

Outras questões envolvendo a comunicação entre sujeitos são as subjetivas que estão envolvidas no ato de enviar e receber mensagens. Na comunicação, perceber o significado e a intenção do autor da mensagem é importante, pois é da percepção e da interpretação que seguirão as ações como “feedback” ao que foi entendido, bem como a sequência da troca de mensagens será apoiada nas intenções percebidas.

Por esse motivo, pesquisas recentes na área da análise de linguagem natural têm por objetivo classificar de forma automática os aspectos subjetivos do discurso (textos ou diálogos), tais como opiniões e afetividade. Exemplos de tais trabalhos são (Gregory et al., 2006; Genereux and Evans, 2006; Liu et al., 2003; Ma et al., 2005; Prendinger and Ishizuka, 2001), entre outros. Estas pesquisas unem as áreas da Computação Afetiva e do Processamento da Linguagem Natural (PLN), as quais têm desenvolvido métodos, técnicas e recursos que, integrados, tornam sistemas computacionais capazes de manipular significado afetivo e de sentimentos no discurso.

Picard (1997) define a computação afetiva como “a computação que se relaciona, surge de, ou deliberadamente influencia as emoções”. Pesquisas recentes na área da computação afetiva têm utilizado métodos, técnicas e abordagens para dotar os sistemas computacionais, programas e dispositivos de hardware, com capacidade de detectar e externalizar condições e situações que envolvam sentimentos, emoções e afetividade. O PLN apresenta uma característica multidisciplinar, integrando lingüística e computação para apresentar soluções computacionais em diversas áreas de aplicação, entre elas a de reconhecimento e representação de informações contidas em textos.

Léxicos computacionais são recursos essenciais para o desenvolvimento de trabalhos de PLN. Da integração do PLN e da Computação Afetiva surgem os léxicos especializados em termos afetivos que dão suporte ao desenvolvimento de trabalhos nessas áreas. Não existem ainda recursos desse tipo para a língua portuguesa. Este trabalho visa superar esta lacuna e desenvolver a primeira base lexical de palavras de emoção para a língua portuguesa, utilizando-a em uma aplicação voltada para a comunicação entre usuários.

Neste trabalho foi desenvolvida uma ferramenta de chat, como uma plataforma para os testes da base lexical, que faz uso de recursos lingüísticos com a capacidade de reconhecer as palavras de emoção utilizados no diálogo. O desenvolvimento do trabalho explorou técnicas propostas pela computação afetiva, que é de representar visualmente outras formas de comunicação, em específico a forma textual, com o objetivo de enriquecer a comunicação (Picard, 1997). A ferramenta permite que as palavras reconhecidas nas mensagens dos usuários sejam representadas como imagens na interface do chat.

O chat é um ambiente para discussão em grupo na web, utilizado, principalmente, para conversas e discussões por grupos de pessoas dispersas geograficamente. Utiliza uma forma de comunicação designada de síncrona, ou seja, todos os participantes devem estar conectados no mesmo instante. Sendo assim, essa ferramenta mostra-se uma escolha interessante para o registro do diálogo, bem como da análise das expressões que ocorrem na conversação.

Uma base lexical de palavras de emoção para a língua portuguesa foi construída a partir das estruturas e das relações presentes nas bases lexicais WordNet (Miller, 1995; Fellbaum, 1998) e WordNet Affect (Valitutti et al, 2004), apresentadas na Seção 2.2. O desenvolvimento dessa base apresentou-se como um desafio ao criar e disponibilizar um recurso ainda não disponível para o português.

1.2. Objetivos

Os objetivos do presente trabalho estão descritos abaixo, como Objetivo Geral e Objetivos Específicos:

1.2.1. Objetivo Geral

Este trabalho tem como objetivo geral a construção de uma base lexical de emoções adequada para o reconhecimento de expressões em mensagens trocadas entre dois ou mais sujeitos, utilizando mecanismos computacionais de comunicação. Nesse sentido, é proposta uma ferramenta integrada a um sistema de chat, baseada em um recurso lexical relacionado às palavras de emoção para a língua portuguesa denominada de WordNet Affect BR. A ferramenta deve reconhecer e representar imagens associadas às palavras, apresentando-as na interface do usuário.

1.2.2. Objetivos específicos

Os seguintes objetivos específicos estão presentes neste trabalho:

- a) Realizar o estudo da fundamentação teórica sobre o enfoque psicológico das emoções, sentimentos e estados afetivos, assim como os estudos da área da Computação Afetiva, baseada nos trabalhos de Picard (1997) e sua relação com a área do Processamento da Linguagem Natural (PLN);
- b) Construir a WordNet Affect BR, uma base lexical com palavras do domínio das emoções a partir das estruturas da WordNet e WordNet Affect. Essa base será criada tomando como referência o modelo cognitivo psicológico de emoções OCC;
- c) Propor uma técnica de reconhecimento de palavras de emoções nas mensagens em discursos e sua representação visual por meio de imagens que transmitam o significado dessas palavras;
- d) Implementar um protótipo da ferramenta, conforme definição do objetivo geral do presente trabalho;
- e) Aplicar a ferramenta com grupos de usuários, avaliando e validando a ferramenta e a representação de palavras de emoções com imagens.

1.3. Organização do Texto

Esta proposta de dissertação organiza-se da seguinte forma: no capítulo 2 é realizada uma revisão bibliográfica apresentando a fundamentação do trabalho com definições sobre emoções, sentimentos e estados afetivos, além dos conceitos da computação afetiva relacionados ao trabalho e o modelo cognitivo psicológico de emoções OCC. Nesse capítulo também é realizado um estudo introdutório às bases lexicais afetivas.

No capítulo 3 são detalhadas as fases que envolveram a criação da base WordNet Affect BR, assim como a metodologia para estender a base quanto à quantidade de palavras de emoções. Também são apresentados, nesse capítulo, os estudos realizados com a WordNet e a WordNet Affect, além de apresentar e detalhar a modelagem que permitiu desenvolver a WordNet Affect BR com os recursos de consulta e de tradução das palavras de emoções para a língua portuguesa.

No capítulo 4 é discutido e apresentado o uso de imagens na representação de significados de emoção e sua aplicabilidade nesse trabalho.

No capítulo 5 é descrita a ferramenta desenvolvida como plataforma de testes e validação da base WordNet Affect BR. São apresentadas as funcionalidades do chat “Emoticon”, bem como os módulos de reconhecimento das palavras de emoções e a representação visual. Também é apresentado o módulo de administração da ferramenta, sua aplicação e suas funcionalidades.

No capítulo 6 são descritos os experimentos com o chat, a análise dos resultados do uso da ferramenta, assim como são analisados os resultados de uma pesquisa realizada com usuários quanto à adequação do agrupamento de palavras de emoções e sua relação com as imagens que as representam. Nesse capítulo também são descritos alguns experimentos com um recurso da ferramenta em processar textos e apresentar resultados relativos ao cálculo da valência quanto à emoção presente nesses textos.

E, finalmente, no capítulo 7 são apresentadas as considerações finais desse trabalho com as possibilidades de trabalhos futuros em aplicações envolvendo a base WordNet Affect BR.

2. Fundamentos teóricos

2.1. Estados afetivos e emoções

Com o objetivo de esclarecer sobre o papel das emoções no presente trabalho, primeiramente são apresentadas algumas noções sobre emoções e estados afetivos. Em seguida é apresentada a classificação dos estados afetivos, baseados no trabalho de Scherer (2000).

Emoções, humor, sentimentos e outros estados afetivos desempenham um importante papel no comportamento, e, portanto, na forma como as pessoas convivem e se relacionam. Conforme Picard (1997, pág. 47), “emoções são importantes para a inteligência humana, na tomada de decisões racionais, na interação social, percepção, memória, aprendizagem e criatividade”. Para Damásio (1996), as emoções são indispensáveis para a vida racional das pessoas, pois é o comportamento emocional e a forma de cada indivíduo reagir às situações de forma singular que diferenciam um indivíduo do outro. Dessa forma, as emoções ajudam a avaliar situações como sendo boas ou más, seguras ou perigosas e, assim, atuam diretamente na tomada de decisão.

Embora o termo emoção seja popularmente utilizado para muitos fenômenos de ordem afetiva, esses fenômenos devem ser denominados de forma mais genérica pelo termo “estado afetivo” (Jaques e Vicari, 2005). Um estado afetivo pode ser visto como um termo mais abrangente, o qual inclui outros estados além das emoções, como por exemplo, o humor (Scherer, 2000; Fridja, 1994, apud Jaques e Vicari, 2005). Scherer (2000) apresenta uma classificação de estados afetivos em emoções, humor, atitudes interpessoais, atitudes e traços de personalidade, relacionados abaixo:

Emoções: é um episódio relativamente breve de resposta sincronizada para a maioria ou todos os seres vivos na avaliação de um evento interno ou externo como sendo de maior significância. Alguns exemplos de emoção são: raiva, tristeza, alegria, vergonha, orgulho, entre outros. Ortony e colegas propõem uma definição similar, porém mais precisa para emoções (Ortony et al., 1988). De acordo com os autores, emoções são reações com valência (positiva ou negativa) a eventos, agentes ou objetos, cuja natureza particular é determinada pela maneira que a situação disparadora é construída. De acordo com essa definição para

Jaques e Vicari (2005), surpresa não é uma emoção, pois não possui uma valência. Da mesma forma, a emoção é considerada um estado mental intencional, porque é direcionada a um objeto intencional. Por exemplo, estar com raiva de alguém, mas, ao mesmo tempo, admirar outra pessoa (Fridja, 1994 apud Jaques e Vicari, 2005).

Humor: o humor é um estado afetivo difuso que consiste em mudanças no sentimento subjetivo, possui baixa intensidade, mas longa duração e não tem causa aparente. Alguns exemplos de humor são: irritado, deprimido, chateado e nervoso.

Posturas interpessoais: são as atitudes interpessoais ocorridas durante as relações com outras pessoas em determinada interação, definindo a forma como as trocas interpessoais ocorrem. Ou seja, é a maneira como nos posicionamos afetivamente em relação a alguém com quem estamos interagindo. Aparentar frieza, distanciamento e indiferença são exemplos de posturas interpessoais.

Atitudes: são posturas suportadas por crenças de cunho afetivo, ou seja, preferências e pré-disposição nas relações com outras pessoas ou objetos. Gostar, amar, desejar, valorizar ou odiar são exemplos de atitudes.

Traços de personalidade: são tendências de comportamento e disposição de personalidade estáveis, típicos de uma pessoa. São mais estáveis no tempo que as posturas interpessoais. São exemplos de traços de personalidade: confiável, nervoso, ansioso, relaxado, lento e hostil.

Em uma abordagem cognitivista da emoção, os estudos focam principalmente o processo de “appraisal”. De acordo com Scherer (2000), a idéia central é a de que as emoções são disparadas e diferenciadas em função da interpretação ou avaliação subjetiva de um acontecimento, situação ou objeto. O processo de “appraisal” busca explicar as diferentes respostas emocionais de indivíduos para o mesmo acontecimento. Por exemplo, um jogo de futebol pode ser avaliado de formas distintas pelo torcedor do time vencedor e por quem torce para o time que perde a partida. Dessa forma, essa avaliação pode disparar emoções diversas, tais como: alegria, frustração, raiva, temor por algo que possa acontecer (seu time não ganhar o campeonato, por exemplo). Portanto, diz-se que o “appraisal” é um processo cognitivo, uma avaliação de situações que pode disparar emoções.

2.1.1. Computação Afetiva

O uso cada vez mais intenso do computador nas formas de comunicação e interação entre as pessoas tem levado a estudos no ramo da Inteligência Artificial (IA) para permitir a máquinas e sistemas computacionais reconhecer e expressar emoções. O campo da IA que pesquisa emoções em computadores é chamado de Computação Afetiva (“Affective Computing” em inglês) e está dividido em dois principais campos (Jaques e Vicari, 2005): (1) Emoções na Interação Homem-Computador e (2) na Síntese de Emoções.

O primeiro estuda mecanismos para reconhecer e expressar emoções por máquinas na interação homem-computador (IHC). O segundo ramo investiga a simulação de emoções por máquinas (síntese de emoções) a fim de descobrir mais sobre as emoções humanas e também construir robôs que pareçam mais reais, com características mais próximas da realidade. O presente trabalho situa-se no primeiro ramo, mais especificamente no reconhecimento de expressões de emoções no discurso.

De acordo com Picard (1997), devido ao seu poder de processamento, computadores possuem maior habilidade para processar padrões, embora as pessoas possuam recursos superiores para interpretar o significado em padrões. Por exemplo, o reconhecimento de expressões ocorre pela observação e processamento dos padrões de expressões e comportamentos emocionais.

Para sistemas computacionais interagirem efetivamente com o usuário, eles devem ser capazes de reconhecer emoções do usuário e responder de forma apropriada. Atualmente, existem diversas formas de reconhecimento de emoções do usuário por meio de padrões, podendo ser citadas como as principais o reconhecimento de expressões faciais, comportamentos observáveis, entonação vocal e sinais fisiológicos (Picard, 1997). Muitos trabalhos foram desenvolvidos com o propósito de reconhecer e expressar afetividade (Liu et al, 2003), (Ma et al, 2005), (Wang et al., 2005), entre outros.

Outros trabalhos apontam para a análise automática de estados afetivos a partir de documentos escritos em linguagem natural, do inglês “automatic sentiment analysis”, baseados em categorias de emoção, tais como “positivo” ou “negativo” (Gregory et al, 2006). Esses textos podem estar em grandes volumes de documentos, um Corpus, ou serem partes de diálogos analisados em tempo real, ou seja, durante a conversação, o que requer também um sistema de comunicação efetivo.

2.1.2. Modelo cognitivo de emoções OCC

O modelo OCC, definido por Ortony e colegas (Ortony et al, 1988) é um modelo psicológico cognitivo de emoções que apresenta como surgem as emoções a partir da descrição dos processos cognitivos elicitados.

Esse modelo, assim chamado pela combinação das letras iniciais de seus autores, Ortony, Clore e Colins, foi organizado em uma estrutura com 22 tipos de emoções, listadas abaixo e mantidos na língua inglesa: “happy for”, “resentment”, “gloating”, “pity”, “joy”, “distress”, “pride”, “shame”, “admiration”, “reproach”, “love”, “hate”, “hope”, “fear”, “satisfaction”, “fears-confirmed”, “relief”, “disappointment”, “gratification”, “remorse”, “gratitude” e “anger”. O modelo OCC é apresentado na Figura 2.1.2.1.

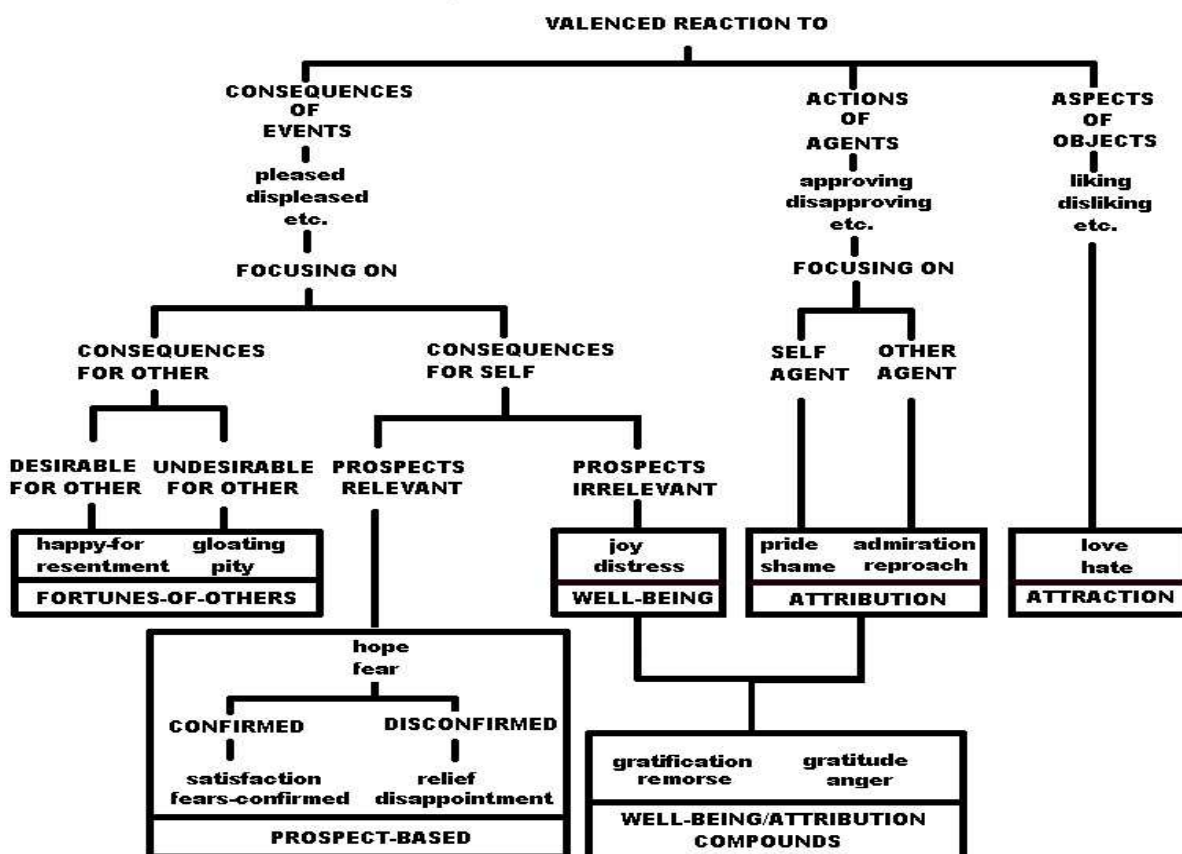


Figura 2.1.2.1: Modelo OCC

Baseado na teoria cognitiva das emoções, o modelo OCC considera que as emoções surgem ao serem disparadas por meio de uma avaliação cognitiva, denominada de “appraisal”, que um indivíduo faz baseado nos três aspectos do mundo: eventos, agentes e

objetos. Eventos são as formas, as ocorrências pelas quais as pessoas percebem aquilo que acontece ao seu redor e quanto à sua percepção de mundo diante das conseqüências geradas. Agentes podem ser pessoas, animais ou, em alguns casos, outros objetos¹, ou abstrações como instituições. E por fim, os objetos, que são percebidos pelas pessoas como objetos presentes no mundo, exercendo uma atração, gerando um sentimento positivo ou negativo.

De acordo com a estrutura cognitiva da “appraisal”, há três componentes interlaçados que fundamentam as percepções de bom ou ruim: objetivos, padrões e atitudes. A variável associada com reações aos eventos é a desejabilidade, ou seja, se eles promovem ou impedem alguém de atingir seus objetivos. As ações de um agente são avaliadas de acordo com a sua obediência a padrões e normas. E, por último, objetos são avaliados como atraentes de acordo com a compatibilidade de seus atributos quanto à preferência de alguém (Jaques e Vicari, 2005).

Resumidamente, de acordo com o modelo OCC, as emoções podem surgir da: (1) avaliação das conseqüências de eventos de acordo com a sua desejabilidade em relação aos objetivos de uma pessoa, (2) avaliação (aprovação ou desaprovação) das ações de uma pessoa de acordo com padrões morais de um indivíduo, e (3) o apelo atrativo dos aspectos de objetos.

As emoções alegria, tristeza, medo, esperança, satisfação, alívio, frustração e medo-confirmado surgem pela avaliação da desejabilidade das conseqüências de um evento para si próprio.

A **alegria (felicidade)** pode surgir devido à ocorrência de algum evento desejado. Por outro lado, a **tristeza** surgirá pela ocorrência de um fato ou evento não desejado. Além disso, quando o evento é esperado, ou seja, há uma expectativa por parte do indivíduo que o evento aconteça ou não, as emoções satisfação, frustração, esperança, medo, medo-confirmado e alívio podem surgir.

O **medo** e a **esperança** podem surgir quando não houve ainda a confirmação de que o evento aconteceu, ou seja, quando o indivíduo ainda está experimentando expectativa. O medo ocorre quando as conseqüências do evento são indesejáveis de acordo com os objetivos da pessoa, e esperança, caso contrário. Quando esse evento esperado acontece, as emoções

¹ Um objeto pode ser um agente se gerar uma emoção de desaprovação e um objeto se causar desagrado. Por exemplo, um mau funcionamento de um automóvel pode causar desagrado em seu proprietário, ou desaprovação pela falha inesperada.

satisfação ou medo-confirmado podem surgir. A **satisfação** acontece quando um evento esperado que já ocorreu é desejável, já o **medo-confirmado** quando o evento é indesejável. Quando o evento esperado acabou por não se realizar, pode haver frustração ou alívio. A **frustração** acontece quando há a confirmação de que o evento esperado e desejável não aconteceu. O **alívio** acontece para a mesma situação, mas sendo o evento não desejável.

Todas as emoções explicadas até agora ocorrem pela avaliação da desejabilidade do evento para si próprio. Mas, os eventos podem, também, ser avaliados como indesejáveis ou desejáveis para outras pessoas e, nesse caso, o indivíduo pode sentir as emoções feliz-por-outro, ressentimento, pena ou prazer-pelo-mal-do-outro.

Quando o evento é desejável para uma outra pessoa, um indivíduo pode sentir **alegria pela outra pessoa** ou **ressentimento**. Quando o evento é indesejável para o outro, ele pode experimentar **prazer pelo mal alheio** ou **pena**.

Quando o objeto de avaliação é a ação de uma pessoa, pode ocorrer orgulho, vergonha, admiração ou reprovação.

Orgulho e **vergonha** acontecem quando o indivíduo aprova (ou desaprova) a sua própria ação. Quando é aprovado (ou desaprovado) a ação de outra pessoa, **admiração** ou **reprovação** podem ser experimentados.

Existe ainda uma outra categoria de emoções que surgem quando uma pessoa avalia as ações de outra ou de si próprio em relação à interferência na realização de seus objetivos. Dessa maneira, essas emoções resultam em focar simultaneamente a ação de um agente e o evento resultante e suas conseqüências.

Quando o indivíduo avalia as suas próprias ações, ele pode sentir **gratificação** quando ele aprova a sua ação e ela tem conseqüências positivas para si, e **remorso** quando ele desaprova a sua ação e esta tem conseqüências negativas. Quando ele avalia as ações de outra pessoa, ele pode sentir **gratidão** por essa pessoa quando ele aprovar as ações dessa outra pessoa e, também, quando o evento resultante dessa ação tiver conseqüências positivas para si. Do mesmo modo, a **raiva** surgirá quando ele desaprovar a ação de outro e essa ainda tiver conseqüências negativas para si.

Por último, um indivíduo pode gostar ou não dos aspectos de um objeto. Quando o objeto exerce uma atração positiva ele sente **afeição** ao objeto, caso contrário, sente **aversão**.

2.2. Base Lexical afetiva

Conforme Specia e Nunes (2004), os léxicos computacionais são recursos lexicais criados, geralmente, de forma manual, especificamente para o tratamento computacional. São também chamadas de bases de dados lexicais (Lexical Databases). De acordo com Jurafsky (2000), um léxico pode ser entendido como uma estrutura altamente sistemática que define o significado das palavras e como elas podem ser usadas. Cada termo de um léxico é identificado como um lexema, que consiste de uma forma ortográfica e fonológica com uma forma de representação de significado.

A WordNet é uma base de dados composta de entradas, ou palavras, que são os lexemas com uma estrutura hierárquica baseada nos synsets e nas relações lexicais e semânticas, conforme apresentado na Seção 2.2.2. Da mesma forma, a WordNet Affect, descrita na seção 2.2.6, é uma base cuja organização está relacionada com a estrutura dos synsets da WordNet, acrescida de informações dos estados afetivos. Dessa forma a WordNet Affect mostrou-se adequada para atender às necessidades desse trabalho.

Por sua vez, diversos autores, Liu e colegas (2003), Ma e colegas (2005) entre outros, apresentaram o uso de um léxico em trabalhos de reconhecimento de emoções baseadas em texto. Ortony e colegas (1987) definem a importância de um “léxico afetivo” afirmando que o léxico afetivo permite utilizar não somente palavras que se referem diretamente a emoções, mas a muitas outras palavras, que mesmo não se referindo a emoções, implicam em diversas outras formas delas.

2.2.1. WordNet

A WordNet é uma base de dados disponível de forma eletrônica, considerada como um importante recurso da área da lingüística. Ela surgiu em 1984, como um experimento lingüístico com o objetivo de ser uma base lexical para a língua inglesa, desenvolvida na Universidade de Princeton² (Miller, 1995; Fellbaum, 1998). Atualmente conta com mais de 200.000 palavras, conforme Tabela 2.2.2.1, entre todos os conjuntos de categorias lexicais: substantivos, verbos, adjetivos e advérbios.

² Atualmente a WordNet está na sua versão 2.1 liberada em março de 2005 para o sistema operacional Windows e na versão 3.0, liberada em dezembro de 2006 para os sistemas operacionais Unix, Linux e Solaris.

A WordNet é uma organização das palavras pelo seu significado e forma, por isso é considerada uma “matriz lexical” bi-dimensional quanto às suas relações. Uma dimensão, das relações lexicais, ocorre entre as palavras, e uma outra, denominada de semântica-conceitual, ocorre entre os conjuntos de sinônimos, denominados de synsets (Valitutti et al, 2004).

A semântica lexical relacional³, presente na WordNet, assume que o significado de um item lexical deriva dos significados dos demais itens, de acordo com certas possibilidades de relação ao longo da árvore semântica (Fernandes, 2005).

A estrutura hierárquica da WordNet permite sua utilização por diversas aplicações que exploram as relações lexicais e semânticas entre os synsets. As classes gramaticais que identificam as palavras na WordNet (substantivos, verbos, adjetivos e advérbios) são agrupadas nos conjuntos de sinônimos, cada um expressando um conceito distinto e que representam, por sua vez, um conceito ou categoria léxica diferente (Botelho, 2002).

2.2.2. Synsets

Um synset é um conjunto de sinônimos que compõem a estrutura da base da WordNet e são descritos por uma “glossa”, ou glossário, que é uma descrição semelhante à encontrada nos dicionários. Um synset pode apontar para outros synsets por meio de relações de hierarquia.

A Tabela 2.2.2.1 relaciona as categorias lexicais que integram a base da WordNet. Uma palavra pode aparecer em mais de um synset e em mais de uma classe gramatical.

Tabela 2.2.2.1: A estrutura da WordNet com relação às classes gramaticais

Classe gramatical	Palavras	Synsets
Substantivos	117.097	81.426
Adjetivos	11.488	13.650
Verbos	22.141	18.877
Advérbios	4.601	3.644

2.2.3. Relações da WordNet

Relações lexicais

Na WordNet as relações lexicais ocorrem entre as palavras e podem ser classificadas em:

³ Ao estudo lingüístico das palavras, suas relações e significados, dá-se o nome de Semântica Lexical, de acordo com Jurafsky (2000).

- a) Sinonímia: é a principal relação da WordNet. Ocorre entre palavras diferentes que possuem o mesmo significado. A sinonímia, geralmente, atende ao princípio da substituição, onde duas palavras são sinônimas se elas podem ser substituídas com sucesso uma pela outra (Jurafsky, 2000). As palavras que possuem uma relação lexical de sinonímia estão no mesmo synset.
- b) Antonímia: é o contrário da sinonímia, onde o significado de uma palavra é o oposto de outro. Essa relação ocorre, principalmente, entre a categoria dos adjetivos.

Relações semânticas

Na WordNet as relações semânticas, também denominadas de semântico-conceituais, ocorrem entre os grupos de sinônimos, os synsets, e podem ser classificadas em:

- a) Hipernímia: é uma relação semântica que ocorre entre substantivos e verbos, conectando um conceito mais abrangente e um mais específico. Por exemplo, o synset “Mamífero” é uma superclasse do synset “Cachorro”.
- b) Hiponímia: é uma relação semântica que ocorre entre substantivos e verbos, conectando uma classe subordinada e sua superordenada. Por exemplo Figueira é uma sub-classe de árvore.
- c) Meronímia: refere-se à relação entre substantivos, da parte para o todo. Por exemplo, árvore é um membro de floresta. Pode-se dizer então que árvore é um merônimo com relação à floresta. Da mesma forma, cozinha é “parte do” apartamento, portanto cozinha é merônimo com relação a apartamento.
- d) Holonímia: refere-se à relação entre substantivos, do todo para a parte. Por exemplo, hidrogênio é substância de água. Pode-se dizer, então, que água é holônimo com relação a hidrogênio. Da mesma forma, disciplina é “parte da” grade curricular, portanto grade curricular é holônimo com relação à disciplina.
- e) Causalidade: refere-se à relação de causalidade entre os verbos, onde o primeiro elemento é o causativo e o segundo resultativo. Por exemplo, mostrar-ver e dar-receber.
- f) Troponímia: refere-se à relação entre os verbos, onde o primeiro implica no segundo. Por exemplo, caminhar e mover.

A relação de “pertencer a” aponta para o substantivo com o qual o adjetivo possui uma relação. Os advérbios, por sua vez, geralmente apontam para o adjetivo do qual é derivado. A lista completa das relações entre os synsets está relacionada no Anexo B.

A Figura 2.2.3.1 corresponde à tela de consulta da WordNet on-line com o resultado para a palavra “happy”.

<p>Adjective</p> <ul style="list-style-type: none"> •S: (adj) happy (enjoying or showing or marked by joy or pleasure) "a happy smile"; •S: (adj) felicitous, happy (marked by good fortune) "a felicitous life"; "a happy outcome" <ul style="list-style-type: none"> ▪similar to <ul style="list-style-type: none"> •S: (adj) fortunate "other, less fortunate, children died"; "a fortunate choice" ▪derivationally related form ▪antonym •S: (adj) glad, happy (eagerly disposed to act or to be of service) "glad to help" •S: (adj) happy, well-chosen (well expressed and to the point) "a happy turn of phrase";
--

Figura 2.2.3.1: Tela on-line da WordNet com o resultado da palavra “happy”.

Na figura é possível visualizar a estrutura relacional que define a organização dessa base lexical. A palavra “happy” possui, no exemplo, quatro sentidos na WordNet. A relação lexical pode ser percebida entre as palavras “felicitous” e “happy”, no mesmo synset. E a relação semântica de “similar to” entre esse synset com o da palavra “fortunate”.

2.2.4. WordNet Domains

Desenvolvida pelos pesquisadores do grupo TCC⁴, a “WordNet Domains” pode ser descrita como uma base lexical multi-lingual, criada para uso da WordNet com as palavras nas línguas inglesa e italiana. Essa base é um recurso para fins de pesquisa e está disponível no site do grupo na internet⁵.

Uma hierarquia de “labels” ou etiquetas foi organizada com, inicialmente, 250 etiquetas de domínios. A hierarquia adotada identifica níveis por especificidade, onde, por exemplo, o domínio denominado de “religião”, pode ter subníveis como: *teologia, mitologia, ocultismo*, etc.

⁴ A divisão de pesquisa TCC (“The cognitive and Communication Technologies”) integra o ITC-Irst, um grupo de pesquisa europeu das áreas do Processamento da Linguagem Natural e da Computação. Disponível em <http://tcc.itc.it/>

⁵ A WordNet Domains está disponível em <http://wndomains.itc.it>

Um domínio pode incluir synsets que originalmente possuem diferentes categorias sintáticas. Por exemplo, o domínio *medicina* agrupa os synsets *doutor*, *hospital* e *operar*, de diferentes categorias lexicais (os dois primeiros são substantivos; o último, um verbo).

Um domínio pode incluir synsets de diferentes hierarquias da WordNet. Por exemplo, o domínio *esporte* contém sentidos para *atleta* que deriva de *estilo de vida*; *equipamentos de jogos* deriva de *objetos físicos*; *esporte* deriva de *ações de praticar* e *campo de jogo* deriva de *localização*.

Uma aplicabilidade para as etiquetas de domínio é reduzir a ambigüidade, podendo ser utilizadas em algoritmos de desambiguação no sentido das palavras, reduzindo o problema de polissemia⁶. Por exemplo, a palavra *banco*, do inglês “bank”, na WordNet apresenta dez sentidos para a classe dos substantivos e oito para a classe dos verbos. Isso significa que tal palavra é polissêmica. Uma etiqueta de domínio soluciona o problema, pois classificaria essa palavra no domínio *economia* se a mesma fosse uma referência à entidade ou instituição financeira, ou no domínio *geografia* se fosse uma referência a um banco de área presente na maré baixa, ou ainda no domínio *arquitetura*, se fosse relacionado ao objeto que se dispõe ao redor de uma mesa, utilizado para as pessoas sentarem.

A WordNet Affect é parte integrante da base WordNet Domains, pois ela contém os synsets anotados em uma hierarquia adicional chamada de “domínio afetivo”.

2.2.5. Base Affect

A base Affect é um recurso lingüístico criado manualmente e contém os termos que se referem aos estados afetivos. Sua parte principal é composta das classes gramaticais da WordNet: substantivos, verbos, adjetivos e advérbios. Cada termo possui um “frame”, com “slots”, contendo informações lexicais e afetivas, conforme mostrado na Figura 2.2.5.1.

⁶ Um problema de polissemia ocorre pela propriedade que uma mesma palavra tem de apresentar vários significados. A polissemia permite jogos de palavras e trocadilhos, o âmbito da ambigüidade e da conotação, pois muitas palavras possuem múltiplos significados (Jurafsky, 2000).

```

[name]: anger
[ita]: <rif src=c> rabbia, collera </rif>
<rif src=wn sense=1> ira, collera, arrabbiatura, rabbia</rif>
<rif src=wn sense=2> collera, ira, bile, furia, rabbia</rif>
[def]: <rif src=wn sense=1> (Psychology) a strong emotion; a feeling
that is oriented toward some real or supposed grievance </rif>
<rif src=wn sense=2> (Physiology) the state of being angry</rif>
[synonyms]: <rif src=wn sense=1>choler, ire</rif>
<rif src=wn sense=2>angriness</rif>
<rif src=mw> fury, indignation, ire, mad, rage, wrath</rif>
[antonyms]: <rif src=mw>forbearance</rif>
[pos]: n
[posr]: <v>anger</v> <a>angry</a> <r>angrily</r>
[fundamental]: <rif src=d>a</rif>
[elliot]: anger
[valence]: -
[arousal]: 2
[ortony]: emotion
[notes]:

```

Figura 2.2.5.1: Estrutura da base Affect

As informações lexicais incluem correlações entre as línguas inglesa e italiana (slot “ita”); a informação da classe gramatical a qual a palavra pertence (slot “pos”); relações de sinonímia (slot “synonyms”) e antonímia (slot “antonyms”); a definição (slot “def”), semelhante à “glossa” da WordNet; e o slot “posr” que relaciona os termos que possuem diferentes classes gramaticais, mas que apontam para a mesma categoria psicológica. No exemplo o adjetivo “cheerful” está semanticamente ligado ao substantivo “anger”, ao verbo “anger” e ao advérbio “angrily”.

As informações afetivas referem-se às teorias de emoção baseadas no conceito da avaliação cognitiva, teorias das emoções básicas e teorias dimensionais. As informações afetivas estão relacionadas aos slots “Elliot”, “valence”, “arousal” e “ortony”.

O conteúdo do slot “Elliot” identifica uma das 24 categorias de emoção identificadas por Elliot (1992). Nos slots “valence” e “arousal” estão as informações dimensionais que representam, respectivamente, valores afetivos positivo ou negativo e o nível de excitação da emoção. O slot “ortony” é baseado no trabalho de Ortony e colegas (Ortony et al., 1987), que classificam os termos em: termos emocionais, termos afetivos não-emocionais e termos de estados mentais não-afetivos. Além disso, outros termos podem ser identificados como: traços de personalidade, de comportamento, atitudes mentais, estados físicos, estados psicológicos e sentimentos, conforme relacionados na Tabela 2.2.5.1.

Parte da informação da base Affect foi coletada com o uso de dicionários e documentos científicos que tratam sobre a psicologia das emoções. Outras informações foram inseridas baseadas de forma intuitiva e arbitrária pelos autores, com revisões e validação por psicólogos e lexicógrafos.

Tabela 2.2.5.1: Classificação com palavras baseada em Ortony e colegas.

Termos	Palavras
Emoção	Raiva
Estados Cognitivos	Dúvida
Personalidade	Competitivo
Comportamento	Choro
Atitude mental	Ceticismo
Sentimento	Prazer

2.2.6. WordNet Affect

A WordNet Affect é um recurso lexical, desenvolvida pelo mesmo grupo da “WordNet Domains” (TCC), que contém um conjunto synsets, gerados a partir da WordNet, que representam conceitos afetivos.

A base WordNet foi disponibilizada para esse trabalho de pesquisa pelo grupo de criador, o ITC-Irst, conforme mail de solicitação e a resposta autorizando o uso, constante no Anexo E.

Para o desenvolvimento dessa base, os trabalhos foram divididas pelo grupo em dois estágios distintos (Valitutti et al, 2004):

O primeiro consistiu da projeção dos termos da base Affect para os respectivos synsets na base WordNet Affect. A informação afetiva da base Affect foi acrescentada aos synsets correspondentes da WordNet Affect, com uma informação afetiva identificada pela marca “a-label”, conforme Tabela 2.2.6.1. A informação afetiva projetada nos “a-labels” é o conteúdo do “slot” “ortony” da base Affect, que conforme descrito na Seção 2.2.5, foi utilizado para diferenciar entre os diferentes tipos de estados afetivos definidos por Ortony e colegas (Ortony et al., 1987).

Tabela 2.2.6.1: Lista de “a-labels” com os respectivos estados afetivos e exemplos.

a-label	Estado afetivo	Exemplos
emo	emotion	substantivo “anger”, verbo “fear”
moo	mood	substantivo “animosity”, adjetivo “amiable”
tra	trait	substantivo “aggressiveness”, adjetivo “competitive”
cog	cognitive state	substantivo “confusion”, adjetivo “dazed”
phy	physical state	substantivo “illness”, adjetivo “all_in”
eds	edonic signal	substantivo “hurt”, substantivo “suffering”
sit	emotion- situation eliciting	substantivo “awkwardness”, adjetivo “out_of_danger”
res	emotional response	substantivo “cold_sweat”, verbo “tremble”
beh	behaviour	substantivo “offense”, adjetivo “inhibited”
att	attitude	substantivo “intolerance”, noun “defensive”
sen	sensation	substantivo “coldness”, verbo “feel”

Numa segunda etapa foram realizadas as correlações com os synsets da WordNet, onde relações semânticas foram utilizadas para alimentar a WordNet Affect. A cada relação semântica verificada na WordNet (antônimo, similaridade, “pertencer a”, atributos, “ver-também” e “derivado-de”), foi verificado se o significado afetivo é preservado. As relações de hiperônimo, implicação, causas e grupos-de-verbos, onde o significado afetivo é parcialmente preservado, não foram consideradas para compor a base afetiva proposta.

A Tabela 2.2.6.2 apresenta o número de novos synsets afetivos obtidos a partir das relações da WordNet, num acréscimo de 1.433 synsets à base WordNet Affect.

Tabela 2.2.6.2: Novos synsets obtidos a partir das relações da WordNet

	Substantivos	Adjetivos	Verbos	Advérbios
Similar to	0	668	0	0
Antonym	64	106	23	6
Pertains-to (direct)	0	2	0	0
Pertains-to	16	0	0	0
Derived-from	0	0	12	12
Derived-from	0	308	0	0
Also-see	0	148	0	0
Attribute	38	0	0	0
Is-value-of	0	30	0	0

A Tabela 2.2.6.3, abaixo, apresenta alguns registros exemplos da base WordNet Affect. A coluna “CAT” é a etiqueta afetiva (“a-label”), que contém a informação oriunda do “slot” “ortony” da base Affect.

Tabela 2.2.6.3: Estrutura da WordNet Affect e exemplo de registros da base

POS	ID do Synset	Origem	CAT
N	#10972097	core	emo
N	#03848510	core	beh, att, tra
V	#00548199	antonym	emo, cog
A	#02475653	pertains-to	eds,emo
A	#00117308	derived-from	emo
A	#02218350	also-see	tra,att
A	#00196128	attribute	emo
N	#03793754	is-value-of	beh,att,tra

Os valores para os campos da WordNet Affect, descritos na Tabela 2.2.6.3 são:

- a) POS (Part Of Speech), que identifica a classe gramatical do synset entre adjetivo, advérbio, substantivo ou verbo;
- b) ID do SYNSET, que é o código de localização do synset na base;
- c) ORIGEM, que identifica se o synset foi obtido diretamente da base Affect (“core”) ou pelas relações da WordNet;
- d) CAT, que é a “etiqueta afetiva” do synset, descritos na Tabela 2.2.6.1.

2.3. Trabalhos relacionados

Abaixo são listados alguns trabalhos estudados que estão relacionados às pesquisas desenvolvidas neste trabalho, no que se refere à identificação de termos e expressões com sentido afetivo e naqueles que utilizam bases lexicais como recurso para o reconhecimento dessas expressões.

Wang e colegas (2005) desenvolveram um chat que utiliza uma técnica de animação de texto associado à informação afetiva para expressar o estado emocional do usuário. O sistema obtém o seu estado afetivo por meio de sensores fisiológicos acoplados no corpo do

usuário. O foco da abordagem é a detecção, em tempo real, de emoções de forma não-verbal de um dos usuários por meio de sensores fisiológicos e, após estimar seu estado afetivo, apresenta para outro usuário na forma de textos animados, ou seja, ocorre uma mudança automática da aparência do texto do segundo usuário, tais como: tamanho, cor, velocidade de apresentação na tela, posição, etc. Os autores basearam-se na estimação dos estados afetivos por meio de uma estrutura bi-dimensional de valores, com um eixo representando a intensidade de excitação afetiva medida (“arousal”) e outro indicando se a emoção é positiva ou negativa (valência). De acordo com os estudos e experimentos, os autores afirmam que o uso de sensores fisiológicos no trabalho facilitou a identificação do nível de excitação (“arousal”), porém o eixo “valência” somente foi obtido de forma satisfatória pela especificação manual.

Ma e colegas (2005) apresentam experimentos com um sistema de comunicação síncrona entre dois ou mais usuários, no formato de um chat baseado num módulo que avalia o conteúdo afetivo a partir da análise textual das mensagens trocadas durante o diálogo. Avatares são associados a cada um dos participantes do diálogo e apresentam, de forma visual, as emoções percebidas, por meio de gestos ou por mensagens faladas através de sintetizadores de voz. As mensagens são processadas com o objetivo de reconhecer seu conteúdo: Para isso os autores utilizaram algumas técnicas em nível de palavras e sentenças eliminando aquelas consideradas sem valor emocional: questões, “stop-words”, sentenças sem o uso do pronome da primeira pessoa, entre outras. Para criar a base lexical afetiva na qual as palavras do diálogo eram avaliadas quanto à afetividade, os autores iniciaram a partir das seis emoções básicas, segundo Ekman (1993): felicidade, tristeza, raiva, medo, surpresa e aversão. Utilizando a base WordNet Affect foi possível estender a base de palavras de emoções a partir das relações de sinonímia dos synsets.

Prendinger e Ishizuka (2001) propõem um framework baseado na comunicação afetiva entre pessoas e sistemas computacionais. O principal objetivo é capacitar agentes a reconhecerem afetividade na comunicação e a expressarem emoções por meio de um comportamento afetivo adequado. Durante a conversação, os agentes participam e tentam influenciar os estados mentais (crenças, objetivos, etc.) e estados afetivos (por exemplo, emoções) de outros agentes. No presente trabalho foram utilizados o modelo das emoções básicas de Ekman (1993) para expressar emoções, e o modelo cognitivo de emoções OCC (Ortony et al 1988) para capacitar os agentes a raciocinar sobre as emoções percebidas do usuário ou de outro agente.

Também motivados pela melhoria da comunicação, Liu e colegas (2003) consideram que uma questão a ser trabalhada é o reconhecimento do estado afetivo do usuário, pois o sucesso na interação social está intrinsecamente ligado ao sucesso na comunicação afetiva. Os autores utilizaram uma base de conhecimento do senso comum com fatos do dia-a-dia, denominada de “Open Mind”, para reconhecer emoções, em nível de sentença, nas situações comuns do cotidiano das pessoas. As emoções são classificadas de acordo com as categorias das emoções básicas propostas por Ekman (1993), utilizando um modelo que usa a estrutura de seis emoções básicas para dar um valor afetivo às mensagens trocadas no diálogo.

Elliot (1992) desenvolveu um framework denominado de “Affective Reasoner”, ou AR, baseado no modelo OCC, para projetar agentes capazes de responder emocionalmente. As respostas e interfaces de comunicação entre os agentes do AR e o mundo real podem ocorrer por meio de expressões faciais ou por respostas em diálogos, o que determina e define as formas de respostas que os agentes utilizarão para comunicar-se e externalizar sua emoção ao responder ao usuário. A capacidade dos agentes em representar e responder emocionalmente ocorre pela modelagem de um conjunto de frames “appraisal” que representam os objetivos individuais do agente em relação aos eventos, os princípios em relação às ações intencionais do agente, as preferências relacionadas aos objetos, assim como um conjunto de 440 possibilidades de expressar as emoções. Para a comunicação com usuários, o agente AR utiliza vários canais multimídia, entre eles 70 expressões faciais modeladas em tempo real, fala sintetizada que permite construir sentenças dinamicamente. Além disso, o agente AR tem um recurso que permite o reconhecimento de fala (Jaques e Vicari, 2005).

Gregory e colegas (Gregory et al., 2006) apresentam considerações sobre a existência de inúmeros trabalhos na literatura voltados para a medição de emoções baseadas em texto, tais como técnicas de análise automática do nível de satisfação de clientes frente a determinado produto, ou para identificar atitudes de empregados baseados em respostas dadas em formulários, questionários, entre outros. Com o objetivo de apresentar uma forma de como a medição pode ser apresentada de forma inteligível e adequada aos usuários, são discutidas abordagens de visualização de emoções e conteúdo afetivo de documentos. Por fim é apresentada uma ferramenta visual que apresenta as informações geradas a partir da análise dos documentos.

Outros trabalhos também têm sido desenvolvidos voltados para a análise de emoções em textos por meio da classificação automática. Genereux e Evans (2006) apresentam um trabalho de classificação de weblogs, conforme seu conteúdo afetivo, em uma abordagem baseada em um modelo bi-dimensional, posicionando as emoções reconhecidas no espaço formado pelos eixos “positivo ou negativo” e “ativo ou passivo”.

De uma forma geral, todos esses trabalhos apontam para a necessidade de capacitar sistemas computacionais a reconhecer e expressar emoções.

3. Criação da base lexical de emoções - WordNet Affect BR

Um dos objetivos desse trabalho é disponibilizar uma base lexical com palavras de emoções para a língua portuguesa denominada de WordNet Affect BR, baseada na estrutura da WordNet e da WordNet Affect. Essa base foi criada tomando como referência o modelo psicológico cognitivo de emoções OCC, apresentado na Seção 2.1.2., detalhado na Figura 3.2.1 quanto às fases de criação e definição da base WordNet Affect BR

3.1. Informações das bases WordNet e WordNet Affect

O desenvolvimento da base lexical de emoção para a língua portuguesa iniciou a partir da estrutura hierárquica dos synsets da WordNet na sua versão 1.6. Essa versão foi utilizada para manter a compatibilidade com a WordNet Affect. Seu conteúdo quanto aos “a-labels” afetivos, aos quais estão vinculados cada synset, apontam para os códigos da versão 1.6.

A consulta à WordNet está disponível “on-line” na internet, porém para o desenvolvimento das atividades planejadas para a pesquisa esse formato apresentou-se inadequado, pois era necessário vincular os synsets e as relações lexicais e semânticas, presentes na consulta, às informações afetivas dos “a-labels” da WordNet Affect.

A estrutura da base de dados da WordNet adaptada para este trabalho é apresentada nas figuras abaixo.

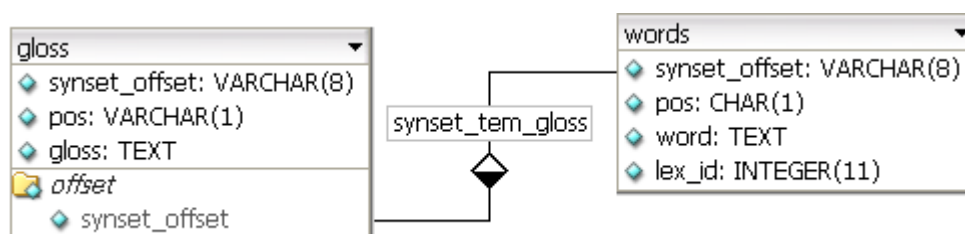


Figura 3.1.1: Modelo ER do relacionamento “synset_tem_gloss” da WordNet.

Na Figura 3.1.1 o relacionamento “synset_tem_gloss” vincula cada synset à sua respectiva definição, que na WordNet é denominada de “gloss”. O glossário é uma breve descrição, com exemplos, que permite ao usuário conhecer a definição do synset.

O conteúdo dos glossários permite identificar, compreender e diferenciar os diversos sentidos que uma mesma palavra possui na WordNet.

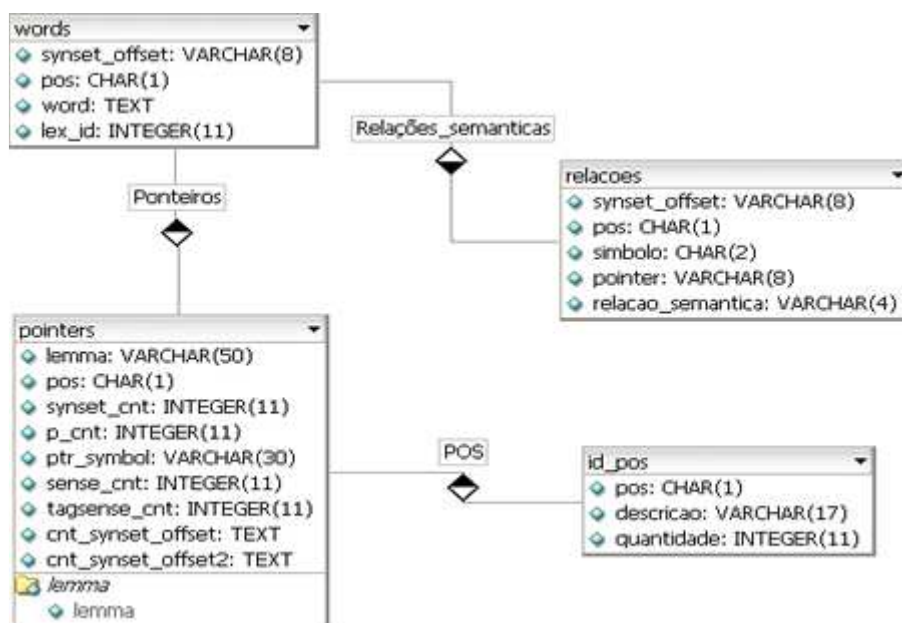


Figura 3.1.2: Modelo ER dos relacionamentos entre os synsets da WordNet.

Na WordNet as relações de sinonímia ocorrem **dentro** do synset, entre as palavras, ou seja, palavras de um mesmo synset são sinônimas entre si.

Por outro lado, as relações semânticas ocorrem **entre** os synsets. Conforme detalhado na Figura 3.1.2, o relacionamento “ponteiros” permite vincular todos os sentidos de uma palavra aos synsets para os quais cada um desses synsets possui uma relação semântica, conforme pode ser observado e melhor detalhado na Figura 3.1.3.

Da mesma forma, o relacionamento “relações_semântica” vincula cada synset às relações semânticas com os demais synsets. E, por fim, o relacionamento “POS” permite relacionar os synsets à classe gramatical respectiva: substantivo, adjetivo, verbo ou advérbio.

Outras tabelas e relacionamentos foram criados para viabilizar o processamento e uso das informações adequadamente, contemplando e mantendo a estrutura hierárquica presente na WordNet.

Inicialmente, os arquivos foram tratados de modo a refletir os resultados obtidos “on-line”, ou seja, a partir da consulta de uma palavra, o sistema deveria apresentar toda a estrutura hierárquica entre os synsets com as respectivas relações lexicais e semânticas.

Um módulo de consulta foi desenvolvido e está disponível para acesso on-line⁷ permitindo navegar pelas informações da estrutura hierárquica da WordNet, enriquecido com a informação da WordNet Affect.

As Figuras 3.1.3. e 3.1.4. apresentam as interfaces com o resultado da consulta para a emoção “happy”.

Figura 3.1.3: Interface de consulta para a emoção “happy”

A interface de consulta oferece as opções para o usuário definir qual o formato e as informações que serão apresentadas na tela, conforme Figura 3.1.3:

- a) código do synset: é um número de identificação que cada synset possui na WordNet. O formato é de oito dígitos, utilizado como chave para os diversos relacionamentos entre as tabelas;
- b) categoria de POS: é a informação quanto à classe gramatical do synset: substantivo, adjetivo, verbo ou advérbio;
- c) número do sentido na base: é a identificação numérica do sentido do synset;
- d) glossário: se marcado, mostrará ao lado de cada synset a definição do synset, com exemplos;
- e) label afetivo: é a informação do “a-label” afetivo da WordNet Affect, detalhado na Tabela 2.2.6.1 sobre a WordNet Affect;

⁷ A consulta pode ser feita on-line no endereço <http://www.paulo.pasqualotti.nom.br/basewordnet/consulta.php>

f) código synset pointer: é o código do synset que possui alguma relação com os synsets da palavra consultada.

A Figura 3.1.4 demonstra o resultado da consulta para “happy”, onde as opções selecionadas pelo usuário, detalhado na Figura 3.1.3, são apresentadas na tela, destacando-se que, no exemplo “happy”, possui seis sentidos na WordNet, tendo quatro deles uma relação com a WordNet Affect.

<input checked="" type="checkbox"/>	01099314 (Adjetivo) happy
	Base afetiva: Emoção
<input checked="" type="checkbox"/>	01308490 (Adjetivo) happy, pleased
	Base afetiva: Emoção
<input checked="" type="checkbox"/>	00998186 (Adjetivo) felicitous, happy
	Base afetiva: Emoção
<input type="checkbox"/>	00556743 content(p), happy
	Base afetiva: Emoção Glossário: satisfied; enjoying well-being and contentment; "felt content with her lot"; "quite happy to let things go on as they are"
	Relações entre os synsets
	00556147 Similar a(a) (emo) <u>contented</u> (a) (emo) <u>content(p)</u>
<input checked="" type="checkbox"/>	00863471 (Adjetivo) euphoric, happy
<input checked="" type="checkbox"/>	00948198 (Adjetivo) happy, well-chosen

Figura 3.1.4: Interface do resultado da consulta da palavra “happy”

As relações semânticas entre os synsets e a definição do glossário podem ser observadas na Figura 3.1.4. A informação do “a-label” afetivo da WordNet Affect pode ser observado na linha “Base afetiva: emoção”. No exemplo, dois synsets de “happy” não possuem essa relação, ou seja, esses dois sentidos não são considerados como um estado afetivo na WordNet Affect.

3.2. Base lexical de emoções na língua portuguesa

Para a criação da base lexical de emoções para a língua portuguesa, denominada de WordNet Affect BR, foram definidas algumas ações que permitiram, a partir das emoções presentes no modelo OCC, expandir para uma quantidade de palavras suficiente para torná-la aplicável como um léxico.

Dessa forma, o modelo OCC de emoções surge no presente trabalho como uma importante referência que permitiu, a partir das palavras presentes na sua estrutura, definir o escopo das palavras da base lexical a ser criada. Para tanto, o modelo OCC norteou todo o processo das etapas apresentadas na Figura 3.2.1., desde a definição das palavras iniciais da estrutura, assim como a sua extensão por meio do processo de tradução, considerando-se o

significado e escopo das palavras de emoções, de acordo com a teoria cognitiva presente no modelo OCC, conforme detalhado nas seções seguintes.

- a) as emoções do modelo OCC⁸ foram analisadas e uma relação entre substantivos e adjetivos foi criada, gerando uma lista com 42 palavras, listadas na Tabela 3.2.1;
- b) para cada uma dessas palavras foram analisadas suas relações com as bases WordNet e WordNet Affect. Para isso, foi utilizada a interface de consulta detalhada na Seção 3.1;
- c) a partir da lista de palavras resultante, foram realizadas as atividades de tradução e validação para o português, dando origem à base WordNet Affect BR.

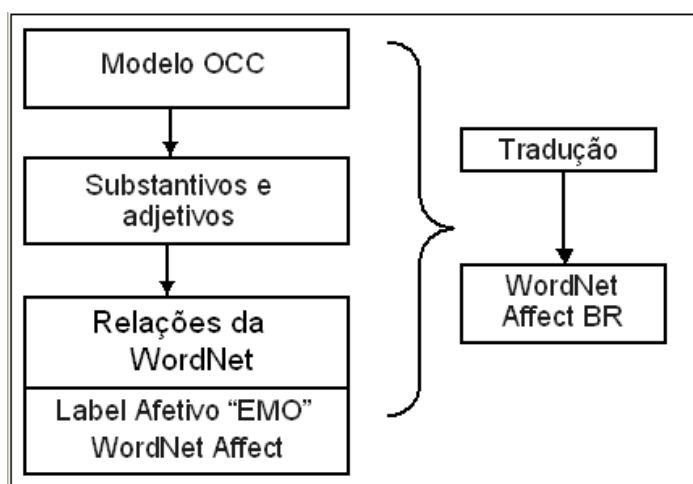


Figura 3.2.1: Estrutura do trabalho de criação da base WordNet Affect BR.

A Tabela 3.2.2. apresenta o escopo da base lexical de emoções, com sua estrutura dividida em níveis. No primeiro nível estão as emoções do modelo OCC. No segundo, encontram-se os sinônimos das palavras de primeiro nível, ou seja, palavras do mesmo synset com o registro correspondente na base WordNet Affect. E no terceiro estão todas as palavras que possuem relações semânticas com o synset do primeiro e segundo nível: de similaridade e “ver também...” para adjetivos e hipernímia e hiponímia para substantivos.

⁸ O modelo OCC apresenta 22 emoções. Para este trabalho foram consideradas 21, pois as emoções “fear” e “fears confirmed” foram agrupadas em uma única emoção definida como “fear”.

Tabela 3.2.1: Emoções do modelo OCC com a tradução para a língua portuguesa.

Modelo OCC das emoções	
Original inglês Substantivos/Adjetivos	Tradução Substantivos/Adjetivos
Happiness-for /Happy-for	Feliz por/felicidade por
Resentment/resentful	Ressentimento/ressentido
Gloating	Feliz pelo mal do outro/ regozijado
Pity/pityful	Pena/penalizado
Joy/joyful	Alegria/alegre
Distress/distressed	Tristeza/triste
Pride/proud	Orgulho/orgulhoso
Shame/shameful	Vergonha/envergonhado
Admiration/admired	Admiração/admirado
Reproach/reproached	Reprovação/reprovado
Love/loving/loved	Afeição/amável/amado
Hate/hateful	Ódio/ odiado
Hope/hopeful	Esperança/esperançoso
Fear/fearful	Medo/medroso/apavorado
Satisfaction/satisfied	Satisfação/satisfeito
Relief/relieved	Alívio/aliviado
Disappointment/disappointed	Decepção/desapontado/decepcionado
Gratification/gratificated	Gratificação/gratificado
Remorse/remorseful	Remorso/arrependido
Gratitude/grateful	Gratidão/grato
Anger/angry	Raiva/irado

Tabela 3.2.2: Estrutura proposta para os níveis da base WordNet Affect BR.

Nível	Origem
1	Palavra de emoções do modelo OCC (adjetivos e substantivos)
2	Palavras do mesmo synset do primeiro nível
3	Palavra com relações de “similar to”, “ver também” para os adjetivos e hipernímia e hiponímia para os substantivos.

Exemplificando, a base contém as seguintes informações, geradas a partir da palavra “happy”, estendida nos três níveis, de acordo com os critérios estabelecidos na Tabela 3.2.2:

- a) Primeiro nível: a palavra “happy”;
- b) Segundo nível: a palavra “happy” está em seis synsets da WordNet, sendo que desses, quatro estão marcados com a “etiqueta afetiva” (“a-label”) na WordNet Affect. No segundo nível da base estarão as palavras “pleased”, “felicitous” e “content”.

c) Terceiro nível: as palavras identificadas através das relações de “similar to”, “ver também”, hipernímia e hiponímia para os quatro synsets da WordNet, a partir da palavra “happy”, são: “blessed”, “blissful”, “bright”, “elysian”, “paradisiacal”, “paradisiac”, “paradisiacal”, “paradisaic”, “paradisal”, “paradisial”, “golden”, “halcyon”, “prosperous”, “joyful”, “joyous”, “laughing”, “riant”, “glad”, “fortunate” e “contented”.

Como resultado dessa fase de extensão da lista de palavras de emoção, foi obtida uma nova lista totalizando 403 palavras, que foram traduzidas para a língua portuguesa conforme detalhado nas seções seguintes.

3.3. Tradução das palavras da base

Uma tradução para a língua portuguesa foi realizada por duas alunas selecionadas a partir de entrevistas, ambas formadas no curso de Letras da Universidade do Vale dos Sinos, com experiências em trabalhos de tradução e vivência em países da língua inglesa, com boa fluência e conhecimento dessa língua.

Definido quem executaria a tradução, foram realizados encontros de nivelamento para o entendimento da proposta de pesquisa e a metodologia a ser seguida, bem como para o planejamento das ações atendendo às premissas definidas para essa fase.

A orientação que norteou todo o processo foi de que a tradução não deveria contemplar somente o sentido isolado das palavras, mas sim considerar o seu significado presente nas definições da WordNet e na estrutura dos synsets. Por isso, além das palavras, foram traduzidos os glossários de cada synset.

Da mesma forma, o modelo OCC foi apresentado e explicado às tradutoras, para que houvesse o entendimento sobre a base teórica do modelo e também para que o mesmo fosse utilizado como referência para a tradução das palavras e dos respectivos glossários.

Para facilitar o processo de tradução foi construída uma interface disponibilizada na internet, conforme a Figura 3.3.1, permitindo que todo o trabalho fosse realizado a distância. Questões de segurança envolvendo o acesso à informação foram contempladas com a criação de mecanismos de controles com a identificação do usuário e senha, cujo objetivo principal foi o de preservar e manter a integridade e a segurança dos dados.

Quanto à apresentação em tela, a interface contempla a forma como a base de palavras foi criada, dividida por níveis, apresentando no primeiro nível os synsets com as emoções do modelo OCC; no segundo nível, as palavras do mesmo synset, ou seja, os sinônimos, e no terceiro nível, as palavras daqueles synsets que possuem as relações de hipernímia, hiponímia para os substantivos e de similaridade e da relação “ver também...” para os adjetivos.

Foi disponibilizada a leitura do glossário, em tela, permitindo aos tradutores conhecerem as definições e significados dos synsets. Na Figura 3.3.1 é apresentado um exemplo da interface de tradução para a emoção “distress”.

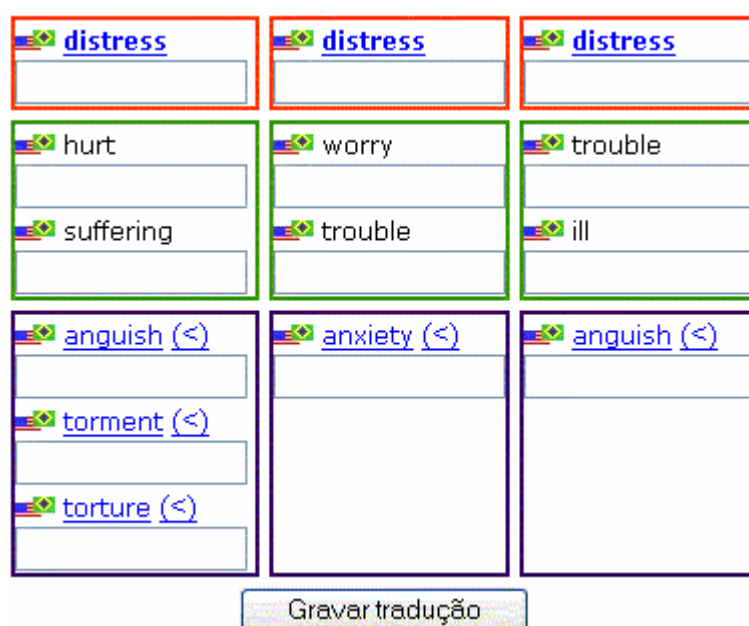


Figura 3.3.1: Interface de tradução das palavras para a emoção “distress”

Na WordNet essa palavra apresenta cinco sentidos, porém somente três têm significado de emoção e fazem parte da base WordNet Affect. Dessa forma, os três synsets são apresentados, por coluna, com as demais palavras que derivaram das relações, por níveis conforme detalhado na Tabela 3.2.2.

Ao clicar na imagem das bandeiras, presentes ao lado de cada palavra, o sistema abrirá uma janela, conforme Figura 3.3.2., permitindo a tradução da palavra e do glossário daquele synset.

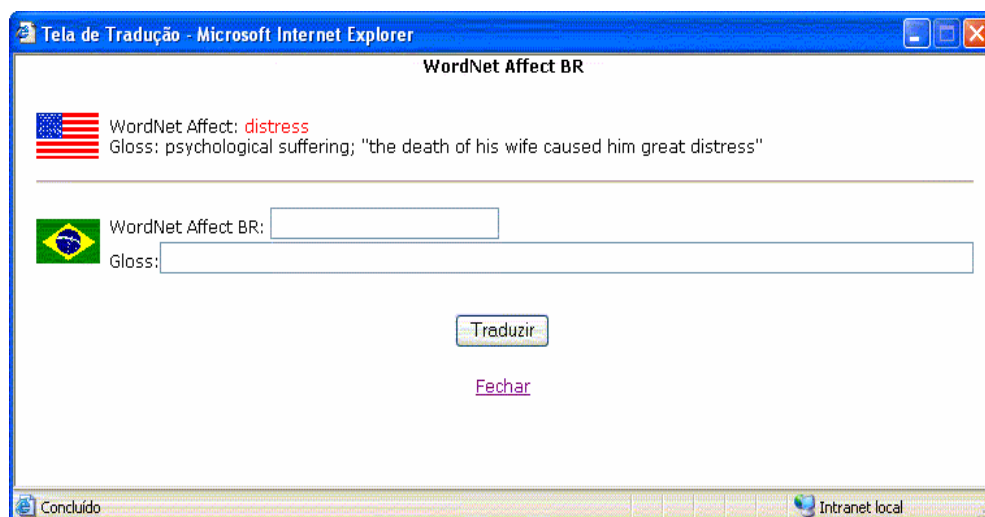


Figura 3.3.2: Interface de tradução da palavra e do glossário.

A compreensão e entendimento pelas tradutoras do significado presente em cada synset foram fundamentais para a atividade de tradução, pois na WordNet uma palavra pode fazer parte de mais de um synset, com significados diferentes.

A definição obtida pela tradução do glossário auxiliou a tradução da palavra, pois conhecer a sua descrição foi importante para identificar a melhor tradução.

Por meio do relacionamento entre as bases, apresentado na Figura 3.3.3., é possível relacionar a WordNet Affect BR com a WordNet. Essa relação permitirá, futuramente, utilizar a base WordNet Affect BR em trabalhos “multi-lingue” inglês-português.

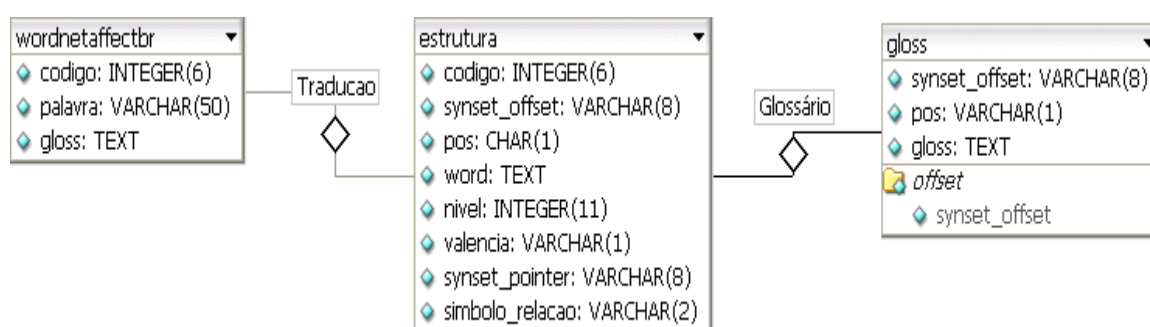


Figura 3.3.3: Modelo ER das relações entre a WordNet Affect BR e a WordNet.

O relacionamento “traducao” entre as tabelas “wordnetaffectedbr” e “estrutura” é responsável por integrar e manter a vinculação das informações da base WordNet Affect BR, com a WordNet, permitindo, por exemplo, relacionar cada palavra a um dos sentidos presente na WordNet.

Isso é possível por meio do campo “synset_offset” da tabela “estrutura” cujo conteúdo vincula cada registro à base da WordNet. Da mesma forma, o conteúdo do campo “código” relaciona-a à tabela “wordnetaffectedbr”.

A tabela “gloss” possui a definição para cada synset. A sua relação com as outras tabelas permitiu a tradução dessas definições conforme o significado e sentido dos synsets.

As atividades de tradução e validação foram realizadas pelas tradutoras diretamente nas bases de dados da proposta. Para isso foram criadas interfaces adequadas para esse trabalho. A metodologia definida para esse trabalho, prevendo as fases de tradução e validação, exigiu a criação de controles adicionais, que identificassem e relacionassem cada palavra à respectiva tradutora. Dessa forma foi possível atender à premissa de que toda a palavra traduzida por uma obrigatoriamente deveria ser validada pela outra e vice-versa, conforme detalhado na Seção 3.4.

Para a atividade de tradução e validação, as 403 palavras em inglês foram divididas em duas outras listas, conforme Tabela 3.3.1.

Tabela 3.3.1: Distribuição das palavras para a tradução

LISTA 1				
SUBSTANTIVOS				
palavra	nível 1	nível 2	nível 3	total
admiration	3	4	3	10
anger	2	4	1	7
disappointment	1	0	0	1
distress	3	6	4	13
fear	2	4	13	19
gloating	1	2	0	3
gratification	2	0	0	2
gratitude	1	0	0	1
happiness	2	1	9	12
hate	1	1	8	10
hope	4	1	3	8
joy	2	4	8	14
ADJETIVOS				
palavra	nível 1	nível 2	nível 3	total
admiring	1	0	0	1
angry	1	0	26	27
ashamed	1	0	8	9
content	2	1	5	8
disappointed	1	5	0	6
distressed	2	6	1	9
fearful	4	15	7	26
glad	2	1	0	3
grateful	2	1	2	5
gratified	1	0	1	2
TOTAL	41	56	99	196

LISTA 2				
SUBSTANTIVOS				
palavra	nível 1	nível 2	nível 3	total
love	4	1	9	14
pity	3	4	4	11
pride	4	2	0	6
relief	7	13	2	22
remorse	1	2	10	13
reproach	1	0	0	1
resentment	1	4	9	14
satisfaction	4	3	7	14
shame	3	2	6	11
ADJETIVOS				
palavra	nível 1	nível 2	nível 3	total
happy	4	2	3	9
hateful	0	0	0	0
hopeful	1	0	4	5
joyful	3	4	2	9
loving	1	0	38	39
pitiful	2	13	1	16
proud	2	3	3	8
relieved	1	2	1	4
remorseful	1	3	2	6
reproachful	0	0	0	0
resentful	1	0	4	5
TOTAL	44	58	105	207

3.4. Validação da tradução das palavras

Após finalizada a tradução, uma outra interface foi utilizada para a fase de validação, mantendo-se a premissa usada na fase de tradução, ou seja, considerar a descrição e o significado da palavra de acordo com a definição do seu glossário.

As partir dos resultados das fases de tradução e validação, foi gerado um processamento cruzando as duas bases que contém as palavras traduzidas em cada fase. Essas bases foram listadas e avaliadas quanto à diferença de conteúdo, conforme Figura 3.4.1. abaixo.










Código	Word	Tradução	Validação	Resultado
1	<u>admiration</u>	<u>afeição</u>	admiração	
2	<u>admiration</u>	<u>estranheza</u>	admiração	
3	<u>admiration</u>	<u>admiração</u>	admiração	
4	<u>anger</u>	<u>raiva</u>	raiva	
5	<u>anger</u>	<u>raiva</u>	raiva	
6	<u>disappointment</u>	<u>decepção</u>	decepção	
7	<u>distress</u>	<u>sofrimento</u>	sofrimento	
8	<u>distress</u>	<u>adversidade</u>	perigo	
9	<u>distress</u>	<u>aflição</u>	aflição	

Figura 3.4.1: Resultado quanto à concordância das fases de tradução e validação

Com o resultado da validação da tradução apresentando discordância de 42,86%, conforme Tabela 3.4.1, foi necessário acrescentar mais uma fase no processo de tradução, que foi a análise e discussão, pelas tradutoras, das palavras que apresentaram diferenças no seu resultado.

Essa etapa permitiu a correção e a escolha da palavra mais adequada para a tradução, a partir do consenso entre as suas tradutoras.

Tabela 3.4.1: Resultado quantitativo da fase de validação da tradução das palavras

Resultado	Quantidade	Percentual
Concordância	232	57,14%
Discordância	171	42,43%
Total	406	

Outro objetivo dessa atividade de análise e validação da tradução foi de verificar os motivos pelo qual as palavras traduzidas apresentaram diferenças, possibilitando estender a base quando fosse identificada sinonímia entre as palavras. Em algumas situações, o consenso foi por uma das palavras, descartando a outra.

Por exemplo, a palavra “hope” possui, considerando adjetivos e substantivos, cinco sentidos na WordNet, tendo quatro deles significado de emoção, ou seja, que constam da base WordNet Affect, como pode ser observado na Tabela 3.4.2. Cada um dos quatro sentidos possui uma definição, pelo glossário, que orientou todo processo de tradução e validação, permitindo, também, analisar e eliminar discordâncias de entendimento para cada um desses sentidos.

Nos dois primeiros sentidos houve discordância com a opção pela palavra que, no consenso das tradutoras, é a mais adequada de acordo com seu significado. Para os dois outros sentidos houve concordância e ambas realizaram a mesma tradução.

Tabela 3.4.2: Exemplos de resultados quanto ao consenso para validar a tradução

Palavra	Glossário	Tradução 1	Tradução 2	Consenso
“hope”	Um exemplo específico sobre sentir-se esperançoso	expectativa	esperança	expectativa
“hope”	O sentimento universal de que algum desejo será realizado	confiança	esperança	confiança
“hope”	Razões para sentir-se esperançoso sobre o futuro	esperança	esperança	esperança
“hope”	Alguém ou algo em quem/que as expectativas estão centradas	esperança	esperança	esperança
“distress”	Um forte sentimento de ansiedade	agonia	aflição	agonia e aflição

Em outras situações a mesma palavra gerou duas traduções distintas, porém ambas as palavras foram aceitas como válidas, originando sinônimos e palavras adicionais na base. Esse fato aumentou a base, de 403 palavras iniciais, traduzidas e validadas, para 457 palavras

na composição final da tradução. Por exemplo, a emoção “distress”, cujo glossário em um dos seus sentidos define-a como sendo “um forte sentimento de ansiedade” foi traduzida para agonia e aflição, sendo considerada, em ambas as situações, válidas, sendo inseridas como sinônimos, conforme Tabela 3.4.2.

Com as fases de tradução e validação das bases de emoções finalizadas e as palavras avaliadas e validadas, fez-se necessário analisar individualmente cada ocorrência da base resultante, adequando a lista com a exclusão de palavras duplicadas, conforme detalhado na Seção 3.5.

3.5. Composição final da base WordNet Affect BR

As fases de tradução e validação das palavras de emoção para a língua portuguesa, apresentaram algumas ocorrências de duplicidades e de divergências quanto à sua aplicação, cujos detalhes são relatados a seguir:

- a) Como exemplo de duplicidades, a emoção “relief” possui sete sentidos na WordNet. Porém, em quatro deles a tradução foi para a mesma palavra, alívio. Todos os glossários indicam uma sensação de sentir-se aliviado pela mudança de um estado desconfortável para um confortável, como pode ser observado na Tabela 3.5.1. Para a composição final da base de emoções, as palavras duplicadas foram inseridas somente uma vez.

Tabela 3.5.1: Tradução de “relief”, com número do sentido e glossário.

Nº	Palavra	Glossário
1	assistência humanitária	ajuda aos velhos ou indigentes ou deficientes físicos;
2	alívio	o ato de reduzir algo desagradável como a dor;
3	alívio	sensação que ocorre quando algo incômodo é removido ou reduzido;
4	alívio	a condição de estar confortável ou aliviado, depois de ter seu sofrimento aliviado
5	auxílio	assistência em tempo de dificuldade;
6	descanso	uma pausa para relaxamento;
7	alívio	uma mudança para melhor

- b) Quanto às ocorrências de aplicabilidade na língua portuguesa, algumas situações exigiram uma maior discussão e estudo para que a tradução mantivesse o escopo dado pelo modelo OCC. Todas as palavras e expressões originadas nessas ocorrências foram mantidas na base de emoções.

Por exemplo, a emoção “gloating”, que, de acordo com o modelo OCC, indica satisfação por um evento indesejável para outra pessoa, não possui uma tradução adequada para uma palavra da língua portuguesa. A tradução resultou na expressão “prazer pelo sofrimento alheio” e “alegria pelo sofrimento alheio”, mantendo, dessa forma, o sentido, mas com uma diferença na sua estrutura, pois a palavra em inglês é traduzida na língua portuguesa para uma expressão mais complexa, no caso, uma frase.

Da mesma forma, a emoção “satisfaction” possui o synset “gloat” com uma relação de hiponímia, com o significado de “satisfação maldosa”, porém essa satisfação é por um fato indesejado a outra pessoa, representando a emoção de sentir satisfação, que no vocabulário da língua portuguesa não encontra uma correspondência, pois as emoções de sentir prazer e alegria estão relacionadas a eventos desejáveis e positivos, para si próprio ou para outra pessoa.

Também um dos sentidos da emoção “pity” possui como sinônimo a palavra “pathos” com a definição de “sentimento de simpatia com tristeza pelos infortúnios dos outros”, que foi traduzido para *piedade*.

As palavras traduzidas para contente e feliz, para adjetivos, e contentamento e felicidade, para substantivos, apresentam significados que, na língua portuguesa, são utilizados no mesmo contexto, portanto muitas vezes são empregadas para representar a mesma emoção.

No caso de “happiness” um dos synsets, traduzido para a palavra alegria, possui uma relação de hipernímia que está melhor relacionado a contente (integração, contentamento, satisfação e regozijo). Já o synset traduzido para “felicidade” possui as relações de hipernímia que estão relacionados a bem-aventurança e beatitude.

Esse mesmo synset possui a relação de similaridade com a emoção “tenderhearted”, que na tradução significa “coração mole”, representando uma metáfora, que também não será tratada pela ferramenta no reconhecimento de palavras de emoção.

“Loving” possui similaridade à “fatherly”, significando “amor fraterno”, isto é, um sentimento fraternal com relação a outra pessoa.

A emoção “remorse” possui uma relação de sinonímia com o synset “self-reproach”, traduzido para auto-reprovação, tendo o mesmo uma relação de hipernímia com “guilty

conscience”, ou seja, “consciência pesada” que é uma combinação de duas palavras significando uma emoção própria a partir de um evento ocorrido.

O synset “angry” possui uma relação de similaridade com “mad”, que traduzido significa “louco de raiva”, não significando um estado de loucura, mas sim, um emprego de uma metáfora que significa sentir raiva. Assim como o synset “the_green-eyed_monster” é traduzido para “olho grande”, relacionado ao emprego de uma metáfora que remete à emoção de sentir inveja despertado por ciúmes.

A Tabela 3.5.2 mostra a listagem das multi-palavras e as palavras originais, obtidas a partir da tradução, onde podem ser percebidas as ocorrências de metáforas e anigramas.

Tabela 3.5.2: Multi-palavras da WordNet Affect BR.

Multi-palavras da WordNet Affect BR	
"gloating"	alegria pelo mau alheio
"relief"	assistência humanitária
"physical_attraction"	atração física
"guilty_conscience"	consciência pesada
"tenderhearted"	coração mole
"pitiful"	de dar pena
"sexual_desire"	desejo sexual
"mad"	louco de raiva
"in a bad way"	muito mal
"anticipant"	na expectativa
"the_green-eyed_monster"	olho grande
"proud of"	orgulhoso de
"gloating"	prazer no sofrimento alheio
"guilt_feelings"	sentimento de culpa
"uxorious"	submisso à esposa
"smoldering"	vermelho de raiva

Dessa forma, a base original da WordNet Affect BR, que, na fase final da tradução apresentou 457 palavras, após as fases de retirada das duplicidades e análise dos sentidos de emoção, ficou com um total de 289 palavras, listadas no Anexo C desse trabalho.

4. Imagens para a representação de emoções



Diversos trabalhos na literatura apresentam exemplos onde o reconhecimento e a representação por meio de alguma forma visual permitiram qualificar e enriquecer o processo comunicativo. Por exemplo, Ma et al (2005) com representação de imagens e avatares, Wang et al (2005) com textos animados, Elliot (1992) com a representação de respostas por meio de expressões faciais, entre outros, utilizam a representação visual para transmitir significados que são reconhecidos por quem recebe as mensagens.

O “emoticon” é uma forma usada, basicamente, na comunicação baseada na internet. Seu uso popularizou e tornou-se uma espécie de padrão na troca de mensagens, permitindo representar e transmitir o estado emocional e afetivo das pessoas de uma forma diferente da textual.

A denominação “emoticon” deriva da contração do inglês “emotion” e “icon”, ou seja, uma emoção representada por uma imagem, um ícone. Diversas são as descrições e significado, para emoticon. Conforme a Wikipedia (2007), um emoticon é considerado uma forma de comunicação paralingüística, também conhecido por “smiley”, cuja tradução significa sorriso.

Os “emoticons” geralmente aparecem como uma seqüência de caracteres, tais como: :), ou ^-^ e :-); ou também como uma imagem. Ambas as formas traduzem e transmitem o estado psicológico, emotivo, de quem os emprega, por meio de ícones ilustrativos de uma expressão facial, conforme exemplos apresentados na Tabela 4.1.

Tabela 4.1: “Emoticons” representando uma emoção

Imagem	Significado e representação	Seqüência de caracteres
	um sorriso, um estado de alegria, ...	:-) ou :)
	tristeza, estar chorando, ...	:-(ou :(

A partir do pressuposto que imagens podem representar o significado contido em palavras de emoções, a ferramenta de chat desenvolvida nesse trabalho reconhece essas palavras contidas em mensagens trocadas por usuários e apresenta imagens relacionadas.

Para viabilizar a representação na ferramenta, foi criada uma relação entre a base de palavras e imagens. Para a criação da base de imagens as palavras da base WordNet Affect BR foram agrupadas por afinidade, inicialmente contemplando a metodologia que permitiu estender a base a partir das emoções do modelo OCC, apresentada na Seção 3.1.1.

Após, os sinônimos e as demais palavras presentes nas relações de hipernímia, hiponímia, similaridade e “ver também” foram inseridas manualmente no mesmo agrupamento. Além disso, foram agrupadas as palavras que podem ser representadas pela mesma imagem.

Algumas emoções do modelo OCC, mesmo que surgidas a partir de situações diferentes, na representação de seu significado, mostraram-se adequadas para serem agrupadas e vinculadas à mesma imagem. Por exemplo, alegre, feliz e contente, podem estar vinculadas a uma imagem que represente a sensação de bem estar, de alegria e vinculadas a um sorriso.

As palavras da base de emoções foram analisadas e agrupadas em 15 diferentes grupos, relacionadas a uma ou mais imagens. A partir do agrupamento foram identificadas as imagens que, ao transmitir seu significado, mantinham a representação da emoção do grupo de palavras, à qual está vinculada⁹.

Com o objetivo de validar o agrupamento e o uso de imagens para representar as palavras de emoções presentes no diálogo, foi disponibilizado um formulário na internet para que um grupo de pessoas analisasse e desse o parecer quanto à concordância, ou não, do agrupamento das palavras e as respectivas imagens. O formulário e os grupos de palavras com as respectivas imagens estão no Anexo D desse trabalho.

No formulário é possível ao usuário responder se concorda, concorda parcialmente ou se discorda das imagens para o respectivo grupo de palavras. Da mesma forma, no formulário o usuário pode justificar sua resposta, bem como sugerir a retirada ou a inclusão de palavras no agrupamento. Também foi incluído no formulário os campos para que possa se identificar com nome, sexo e idade.

A Figura 4.1. apresenta um recorte do formulário. A listagem completa dos grupos de palavras e as imagens relacionadas estão no Anexo D.

⁹ As imagens foram obtidas na internet do site www.postsmile.com na versão de download “free”.


<p>ADJETIVOS: afortunado, alegre, beatífico, elevado, exultante, feliz, jubiloso, venturoso</p> <p>SUBSTANTIVOS: alegria, alegria pelo mau alheio, beatitude, bem-aventurança, deleite, elação, elevação, entusiasmo, esplendor, excitação, exultação, felicidade, jovialidade, júbilo, prazer, prazer no sofrimento alheio</p>		<p>RESPOSTA</p> <p><input checked="" type="radio"/> CONCORDO</p> <p><input type="radio"/> CONCORDO PARCIALMENTE</p> <p><input type="radio"/> DISCORDO</p>
<p>Caso você concorde parcialmente, ou discorde, das imagens desse grupo, digite no espaço abaixo os motivos.</p> <p>500</p> <div style="border: 1px solid #ccc; height: 50px; width: 100%;"></div>	<p>Caso você considere que alguma das palavras não faça parte do grupo, liste-as no espaço abaixo.</p> <p>500</p> <div style="border: 1px solid #ccc; height: 50px; width: 100%;"></div>	<p>Caso alguma palavra não conste desse grupo de emoções, mas que você acredita que possa ser incluída, sugira no espaço abaixo. 500</p> <div style="border: 1px solid #ccc; height: 50px; width: 100%;"></div>

Figura 4.1: Formulário de validação dos grupos de palavras e imagens.

No formulário, as palavras à esquerda, conforme apresentado na figura acima, estão divididas entre adjetivos e substantivos. Em alguns casos duas ou mais imagens são utilizadas para representar o grupo de palavras.

À direita são apresentadas as opções de escolha, conforme concordância do usuário quanto à relação entre as palavras e a respectiva imagem.

Na parte inferior, para cada agrupamento, o usuário pode escrever os motivos pelos quais ele concorda parcialmente ou discorda, se for o caso, ou também de indicar palavras que ele considere que não devam fazer parte do grupo e também sugerir palavras que não constem daquelas apresentadas na tela, mas que no seu entendimento possam fazer parte do agrupamento.

4.1. Formulário de expressão visual de emoções

O formulário foi respondido por 55 usuários, dos quais 32 do sexo feminino e 23 do masculino, conforme gráfico da Figura 4.1.1.

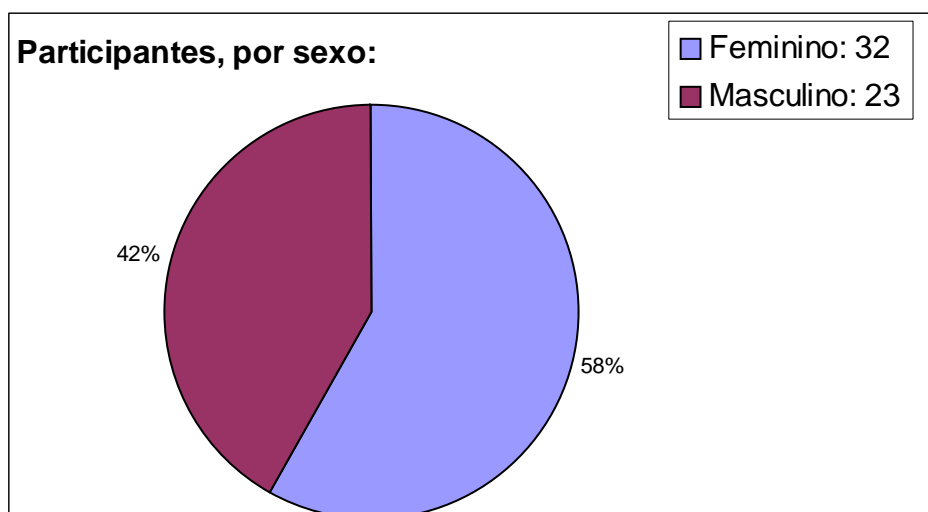


Figura 4.1.1: Gráfico da distribuição dos participantes por sexo

Quanto à faixa etária, a média de idade ficou em 33,7 anos, sendo o participante mais novo com 17 anos e o mais velho com 64. Essa distribuição, por idade, está representada no gráfico da Figura 4.1.2. e na Tabela 4.1.1.

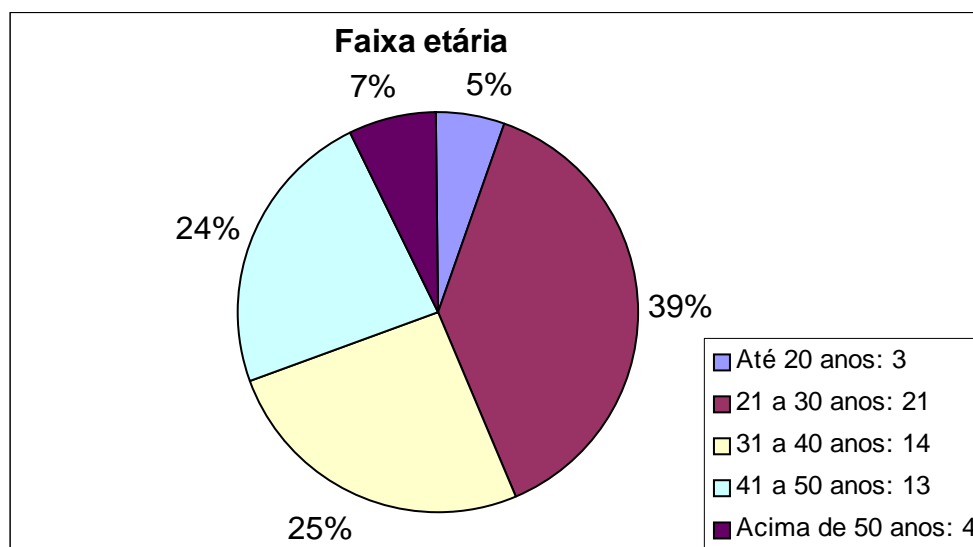


Figura 4.1.2: Gráfico da distribuição dos participantes por faixa etária.

Tabela 4.1.1: Faixa etária dos participantes

Faixa etária	Quantidade
Até 20 anos	3
21 a 30 anos	21
31 a 40 anos	14
41 a 50 anos	13
Acima de 50 anos:	4
Total	55

Quanto à metodologia da pesquisa, para cada grupo de palavras o usuário podia escolher se concordava, concordava parcialmente ou discordava do agrupamento proposto.

No total, as palavras formavam 15 grupos, considerando que 55 participantes responderam, foram, então, 825 respostas possíveis. Dessas, 603 respostas, 73% do total, concordam com a relação entre imagens e os grupos de palavras. Outras 179 respostas, com o índice de 22%, concordaram parcialmente e 43 respostas, ou 5% do total, discordaram.

A Tabela 4.1.2. apresenta as quantidades, por grupo de palavras, de cada uma das opções de resposta: concorda, concorda parcialmente e discorda. A Figura 4.1.3, por sua vez, apresenta o gráfico com os percentuais totais para cada opção de resposta.

Tabela 4.1.2: Resultados quanto às respostas do formulário

Grupos	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Total
Concorda	31	36	26	32	33	48	41	44	46	51	39	42	38	47	49	603
Parcial	24	17	21	15	17	4	11	8	9	0	13	13	17	4	6	179
Discorda	0	2	8	8	5	3	3	3	0	4	3	0	0	4	0	43

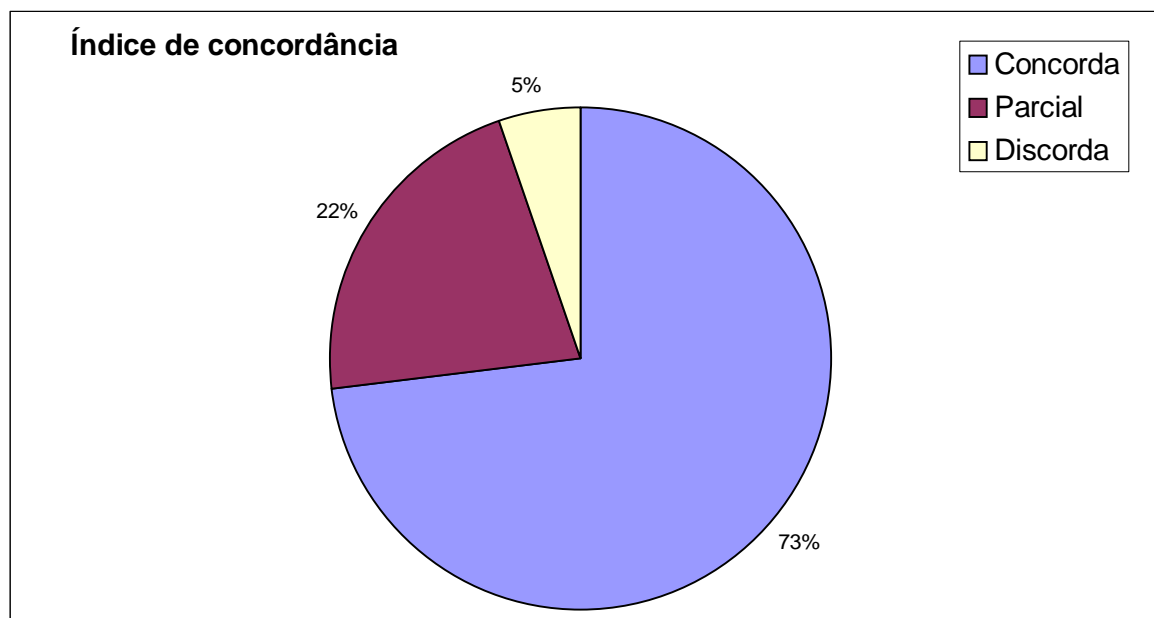



Figura 4.1.3: Gráfico com percentual das respostas dos usuários

O maior percentual de concordância foi no grupo 10, com 93% das respostas, ou seja, 51 usuários concordam com o agrupamento e a relação com a imagem, apresentados na

Tabela 4.1.3. Para esse grupo não houve nenhuma resposta com concordância parcial, apresentando 4 com discordâncias, com índice de 7%.

Tabela 4.1.3: Grupo 10 com maior percentual de concordância


Grupo 10	
ADJETIVOS: arrependido, contrito, penitente, pesaroso	
SUBSTANTIVOS: arrependimento, auto-reprovação, compunção, consciência pesada, culpa, penitência, remorso, sentimento de culpa.	

Onde houve discordância, as respostas dos usuários referem-se à representação de emoção que a imagem exprime, não à retirada ou inclusão de palavras, pois nenhum usuário que discordou sugeriu inserir ou retirar palavras desse grupo.

Quanto aos motivos da discordância, um dos usuários não registrou nenhum motivo ou alteração nas palavras, apenas discordou do agrupamento. As respostas dos demais usuários que discordaram foram: “A imagem me parece não concordar com alguma coisa, somente isso.”, “Acho que a figura exprime sentimento de bah sou muito burro, só faço coisa errada...”, “A imagem não me passa estes sentimentos”.

O grupo 3 apresentou o menor percentual de concordância, com 47% das respostas, ou seja, 26 que concordaram com o agrupamento e a relação com a imagem, apresentados na Tabela 4.1.4. Para esse grupo 26 respostas concordaram parcialmente, ou seja, 38% das respostas. E 8 delas foram de discordância, com um percentual de 15%.

Tabela 4.1.4: Grupo 3 com menor percentual de concordância

Grupo 3	
ADJETIVOS: ativo, galante, impressionante, majestoso, nobre, orgulhoso, orgulhoso de...	
SUBSTANTIVOS: orgulho, soberba	

Para esse grupo as opiniões mostraram-se divididas entre a alteração do agrupamento de palavras com a representação das mesmas pela imagem.

Dos usuários, 11 responderam que a imagem remete a outro significado diferente do sugerido. Por exemplo: “... a imagem retrata algum tipo de atenção. Quer ser visto de alguma forma.”, “A imagem está muito simpática para representar orgulho”, “A imagem deveria ser mais pomposa.”, “Uma imagem que se refira a uma superioridade alegre seja mais adequada.”, “Não me parece uma imagem que representa arrogância ou presunção.”, “Não concordo que a imagem representa soberba. Para mim, trata-se apenas de comemoração por uma vitória.”, “Nobre, galante, majestoso não combinam muito com a imagem apresentada.”, “As mãos para cima balançando não possuem “nobreza” e “galanteza” alguma.”.

Quanto às palavras, de todas as exclusões sugeridas, a emoção *soberba* foi a que teve o maior número de sugestões nesse sentido. Foram 17 sugestões para excluí-la do agrupamento. Em seguida, as palavras *majestoso* e *orgulho*, com 7 e 6 sugestões de exclusão, respectivamente.

Quanto à sugestão para inserir palavras no agrupamento, destaca-se a relação que os usuários fizeram da imagem com a sensação de sucesso, de vitória por algum feito realizado. A palavra *vitioso* foi citada 3 vezes juntamente com a palavra *conquista*. Da mesma forma, foi sugerido incluir palavras cuja imagem repassa uma representação significativa, como é o caso para *entusiasmo*, *motivação* e *vibrante*. E por fim, a mesma imagem também pode ser relacionada às emoções *alegria*, *felicidade*, *contentamento*, *entusiasmo* e *maravilhado*.

O formulário foi utilizado com professores e estudantes de graduação de uma universidade gaúcha e um grupo de empregados de uma empresa. O fato dos participantes

conviverem próximos entre si e, conseqüentemente, utilizarem expressões regionais no seu linguajar, refletiu nas sugestões de palavras não contempladas nos agrupamentos apresentados.

Por exemplo, para o grupo das emoções *feliz* e *alegre*, foi sugerida a palavra *faceiro*. Para o grupo 2, das emoções *contente* e *animado*, foi sugerido a inclusão das expressões “*legal*” e “*tri-legal*” e a expressão “*bah*” foi sugerida para representar a concordância com alguma situação, porém não são expressões que descrevem emoções, mas de alguma forma expressam um estado afetivo. O estudo particular das interjeições é interessante no contexto deste trabalho. Também apareceram sugestões de inclusão de gírias, como por exemplo: “*numa boa*”, “*na paz*”.

As metáforas também apareceram, indicando o uso dessa estrutura da linguagem na comunicação. No grupo da emoção *amor*, foram sugeridas as expressões “*louco de amor*”, “*enlouquecido pelo amor*”, “*flechado pelo cupido*” e “*babando de amor*”.

Nos registros e comentários quanto às mensagens visuais que as imagens remetem, muitos usuários apresentaram uma análise mais elaborada do que somente relacionar a imagem com o grupo de palavras. Pode-se constatar que uma imagem possui um significado forte quanto à mensagem que ela carrega consigo. No caso daquelas selecionadas para esse trabalho, o fato de serem animadas permitiu a percepção das sutilezas do movimento, o que colaborou nas constatações que os usuários registraram.

A Tabela 4.1.5 apresenta alguns exemplos de respostas dos usuários em situações onde não concordaram com a relação entre imagem e o agrupamento das palavras, cuja relação está no Anexo D.

Tabela 4.1.5: Registros dos usuários quanto às imagens de emoção

Situação	Registros dos usuários
Grupo 5: ao inserir a palavra <i>concordante</i> .	“É pelo movimento afirmativo que a cabeça realiza”
Grupo 4: ao justificar a discordância com a imagem.	“A mãozinha abanando me parece estar relacionada à despedida de algo.”
Grupo 4: ao analisar o movimento da imagem.	“O rosto triste acenando dá a impressão de estar comovido pela situação triste de uma outra pessoa.”
Grupo 4: ao justificar a discordância com a imagem.	“Aparentemente o emoticon quer dizer um adeus com tristeza.”
Grupo 4: ao discordar do movimento da imagem e sua relação com palavras do grupo.	“O gesto da mão que o boneco faz significa chamar alguém para mim. Penso que o polegar deveria estar apenas apontado para baixo.”
Grupo 5: ao discordar da relação entre imagem e grupo de palavras.	“Não percebo na imagem o sentimento expressado pelas palavras relacionadas.”
Grupo 14: ao analisar o movimento de uma imagem.	“A imagem remete à reverência, sendo que o personagem reverencia a alguém. Sendo assim, o adjetivo ideal, na minha opinião, seria ADMIRADOR.”
Grupo 11: ao justificar não perceber relação de uma palavra com a imagem.	“A figura passa a impressão de estar ansioso, com expectativa, esperando por algum contato ou resposta. Não vejo nenhuma confiança no emoticon.”

Depois de analisados, quantitativamente, os dados e as respostas dos participantes, conclui-se que o percentual de concordância alcançado, conforme consta na Tabela 4.1.2. e demonstrado no gráfico da Figura 4.1.3 foi satisfatório para fins de realização do experimento.

As sugestões de alterações e adequações nas imagens e na base de palavras de emoções apontam para uma otimização e melhoria que serão realizadas em um trabalho futuro.

5. Ferramenta

Como aplicação da base afetiva, foi desenvolvida uma ferramenta de chat que reconhece palavras de emoção presentes no diálogo e representa-as na tela por meio de imagens.

Para a representação visual da palavra de emoção reconhecida no diálogo, o algoritmo utiliza os dados e relacionamentos entre as tabelas das bases de palavras, denominada de “wordnet affect br” e de imagens, denominada de “emoticon”, retornando para o módulo de interface com o usuário a identificação da imagem que deve ser apresentada na tela. O modelo da ferramenta está detalhado na Figura 5.1.

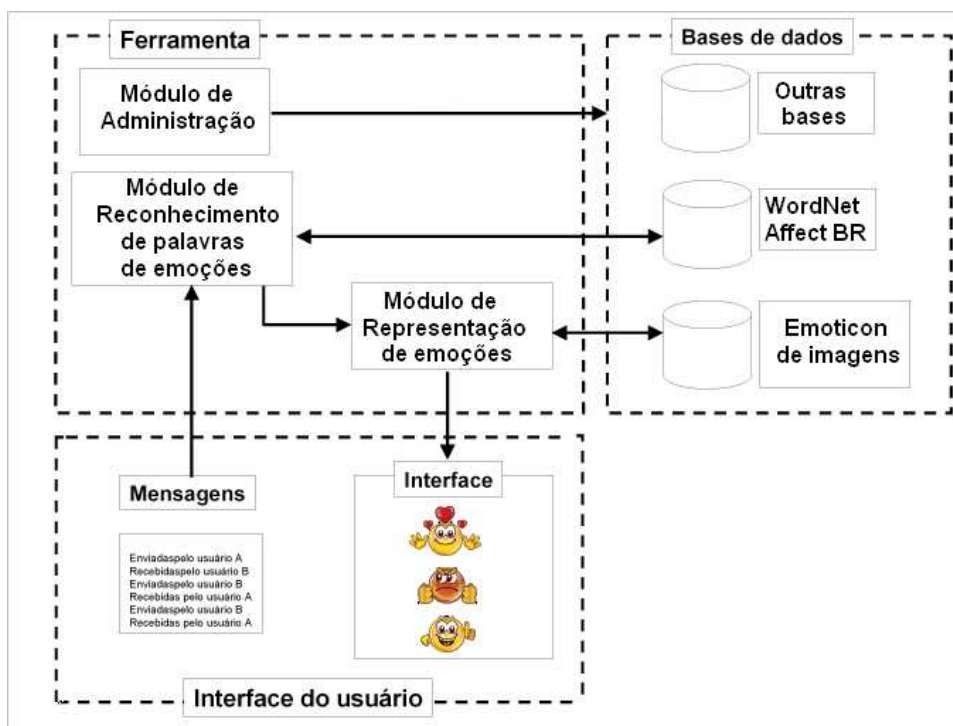


Figura 5.1: Modelo da ferramenta

Na Figura 5.1, o modelo é apresentado com um destaque para o fluxo da mensagem e das ações dos módulos de reconhecimento e representação das imagens.

A mensagem enviada pelos usuários é tratada pelo módulo de reconhecimento, permitindo que se possam identificar as ocorrências das palavras de emoção na mensagem. Elas são enviadas pelos usuários e tratadas como uma expressão completa, ou seja, um

conjunto de palavras separadas por espaços em branco, podendo conter sinais de pontuação e apresentar palavras escritas em maiúsculas ou minúsculas.

Por exemplo, a expressão “Olá! FIQUEI FELIZ em encontrar você hoje!”, para fins de análise, é tratada da mesma forma que a expressão “olá, fiquei feliz em encontrar você hoje”.

Todas as alterações e adequações nas mensagens ocorrem somente no formato. As mensagens são apresentadas aos usuários salientando as palavras de emoções em negrito e a respectiva imagem, como pode ser observado na Figura 5.1.3 que apresenta uma tela da ferramenta de chat durante um diálogo.

5.1. A ferramenta de chat “emoticon”

Durante o desenvolvimento do protótipo da ferramenta de chat, foram realizados inúmeros testes para identificar a melhor lógica para o algoritmo, tornando-o eficiente relacionado ao tempo de resposta, à facilidade no armazenamento e rapidez na recuperação e quanto à apresentação das informações na tela do usuário.

Alguns critérios mínimos de usabilidade foram atendidos para minimizar dificuldades no uso da ferramenta de chat, reduzindo, dessa forma eventuais problemas que pudessem alterar os resultados dos experimentos. Dessa forma a ferramenta apresenta-se com uma interface intuitiva, de fácil aprendizado e uso. Além disso, alguns requisitos de acessibilidade (Dias, 2003) também foram atendidos, tais como: caixas de digitação com tamanho adequado, fontes de texto com tamanho e cores adequados à leitura em tela e uma diagramação da página que atendessem a questões de ergonomia quanto à apresentação das informações e sua disposição na tela. Procurou-se, para tanto, dar às formas de interação presente na ferramenta, a capacidade de serem utilizadas pela maioria das pessoas usuários da internet.

Quanto ao armazenamento, a ferramenta de chat utiliza uma única tabela de dados, denominada de “chatscript”, com uma estrutura funcional que permite agilidade tanto na gravação quanto na recuperação dos registros do diálogo, conforme pode ser observado na Figura 5.1.1.

chatscript	
pk_Id	INTEGER(10)
theText	VARCHAR(100)
theNick	VARCHAR(20)
timestamp	INTEGER(15)
text_emocao	TEXT

Figura 5.1.1: Estrutura da tabela “chatscript”

Os campos da tabela são descritos na Tabela 5.1.1. e foram definidos para armazenar os dados durante o diálogo e recuperar informações posteriormente, para análise quantitativa e qualitativa.

Tabela 5.1.1: Campos da tabela “chatscript”

Nome do campo	descrição
pk_Id	Número sequencial identificando cada registro do diálogo, incrementado automaticamente conforme as mensagens são enviadas pelo usuário
theText	mensagens originalmente digitadas pelo usuário
theNick	identificação do usuário informada na tela de “login”
timestamp	horário no qual a mensagem é enviada. * No “login” armazena o horário de ingresso no diálogo
text_emocao	mensagem formatada após ter sido tratada pelos módulos de reconhecimento e representação das palavras de emoção

Quanto à programação, a ferramenta foi desenvolvida com rotinas de código aberto, cujo funcionamento é baseado em um sistema de chat¹⁰, com identificação do usuário, envio de mensagens, leitura em tela de todas as mensagens do diálogo, atualização em tempo real da lista de usuários participantes do diálogo e a identificação do usuário que enviou a mensagem.

As rotinas específicas de tratamento das informações requeridas pela ferramenta e pelo modelo de reconhecimento e representação das palavras de emoção foram acrescentadas.

Toda programação e desenvolvimento utilizaram tecnologia WEB, disponibilizando, aos usuários, acesso à ferramenta de chat em um endereço internet, de livre acesso sem exigência de senha. Quanto à tecnologia, foram utilizadas a linguagem de programação de

¹⁰ O sistema de chat Microcyb está disponível em <http://www.microcyb.com> sob licença GNU.

scripts PHP, versão 5.1.6 . O servidor WEB é o Apache, versão 5, com a base de dados em MySQL, versão 5.0.25^a.

Nas páginas construídas a partir da linguagem de marcação de hiper-texto, HTML, foram incorporadas folhas de estilos, em CSS com funções em Javascript na validação de campos de formulário e eventos da interação entre a ferramenta e o usuário.

A ferramenta de chat, inicialmente, exige a digitação da identificação do usuário. A Figura 5.1.2 apresenta a tela de “login”¹¹.

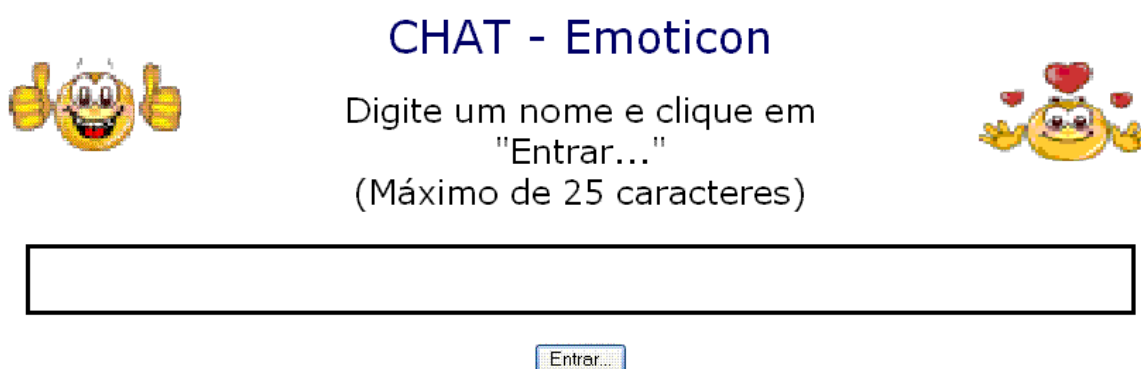


Figura 5.1.2: Tela de login da ferramenta de chat.

Depois da identificação, o sistema registra o usuário na base de dados, com a data e hora de ingresso no diálogo. Após, a interface do chat é apresentada ao usuário, permitindo que o mesmo interaja com os demais usuários presentes no diálogo, envie e receba mensagens, apresentando-as na tela.

As mensagens enviadas no diálogo são recebidas e enviadas pela interface do usuário para o módulo de reconhecimento de palavras de emoção, que, por sua vez, processa a sentença e avaliar cada palavra quanto à sua existência na base lexical de emoções. Caso localize, o módulo de reconhecimento envia a informação para o módulo de representação de emoções, para que o mesmo mostre na tela a respectiva palavra, em destaque, com a imagem relacionada à palavra de emoção.

A Figura 5.1.3 apresenta a interface completa do chat, com todas as partes que a compõem: interface de mensagens, lista de participantes e o campo de envio de mensagens.

¹¹ O chat está disponível na internet no site <http://www.paulo.pasqualotti.nom.br/emoticon>

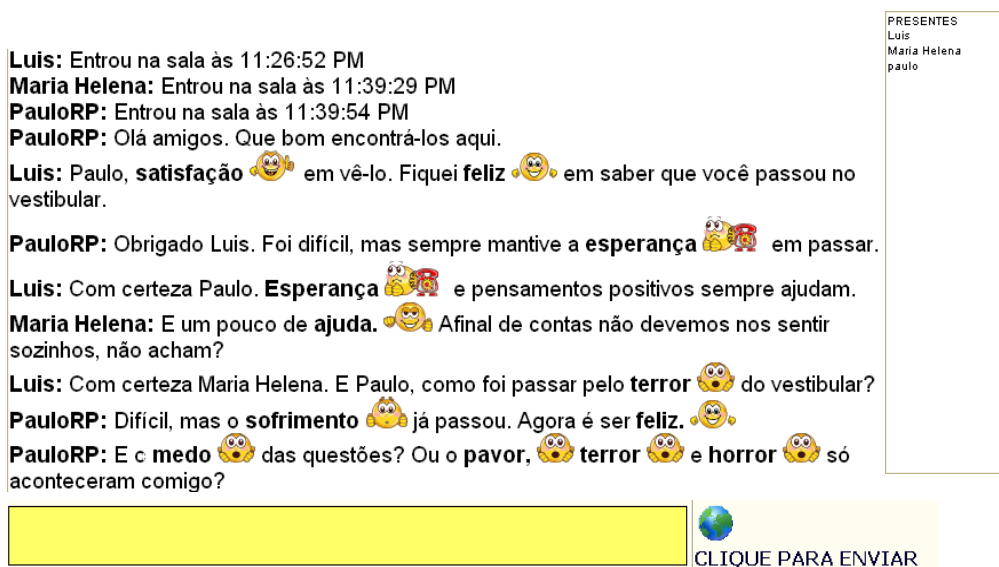


Figura 5.1.3: Interface da tela da ferramenta de chat

O espaço para as mensagens, conforme Figura 5.1.3 é onde as mensagens são apresentadas. As palavras em negrito são aquelas que foram tratadas pela ferramenta e identificadas pelo módulo de reconhecimento de palavras de emoção. As imagens relacionadas ao grupo de palavras de emoção, conforme detalhado no Capítulo 4, são apresentadas ao lado das palavras pelo módulo de representação de emoções.

A ferramenta permite que o usuário visualize todo o conteúdo do diálogo por meio da rolagem da tela. À direita da interface do chat são listadas as identificações dos usuários. A identificação é o nome informado pelo usuário na tela de login.

Na parte inferior da tela está o espaço de interação do usuário, onde o mesmo digita a mensagem e clica no botão com a figura do globo para enviá-la.

Normalmente, programas de chat eliminam qualquer informação sobre o diálogo após o último usuário sair do bate-papo. Porém, para fins de pesquisa, a ferramenta desenvolvida mantém a informação do usuário, bem como as mensagens.

Como complemento às ações de gerenciamento das bases de dados, foi criada uma interface de administração das bases, onde é possível excluir as mensagens e registros dos usuários, bem como salvar os registros em arquivos separadamente para trabalhos posteriores, conforme detalhado na Seção 5.4.

5.2. Módulo de reconhecimento de palavras de emoções

O módulo de reconhecimento das palavras de emoções processa as sentenças em três etapas: tratamento da expressão, análise das palavras e verificação na base WordNet Affect BR.

As mensagens são recebidas e processadas de acordo com as regras de extração das informações definidas e programadas no algoritmo, conforme etapas descritas na Tabela 5.2.1.

Para a ferramenta desenvolvida nesse trabalho, as fases nas quais as sentenças são processadas e tratadas ocorrem, respectivamente, nos níveis: da sentença, das palavras, dos caracteres especiais, dos caracteres de pontuação e dos espaços em branco. As sentenças interrogativas e negativas não são tratadas de forma diferenciada daquelas frases afirmativas, não sendo identificados os sinais de pontuação para interrogação “?” nem palavras de negação.

Tabela 5.2.1: Etapas de tratamento das mensagens do diálogo.

Etapas
Separação por palavra
Processamento individual
Grava formato original
Formata para minúscula
Retira pontuação: ponto final, interrogação, vírgula, exclamação, ponto-e-vírgula, dois pontos, apóstrofo e parênteses

A etapa de *separação por palavra* e o *processamento individual* têm como objetivo principal viabilizar o reconhecimento individual da palavra, ou seja, palavras compostas ou de expressões mais complexas não são tratadas na ferramenta atual.

A etapa de *grava formato original* armazena a mensagem originalmente enviada pelo usuário. Isso permite, por exemplo, recuperar essa informação e todo o diálogo, processando-o posteriormente como um “log”. Essa etapa, apesar de não gerar processamento para o módulo de reconhecimento, foi inserida nessa fase do algoritmo, pois sua ação permite manter a integridade e a estrutura dos dados, sendo, portanto, necessária sua execução na fase inicial do módulo de reconhecimento.

A etapa de *formata para minúscula* formata a mensagem para caracteres minúsculos, independente da forma que o usuário enviou. Essa etapa é importante caso a linguagem de programação seja “case sensitive”, ou seja, que faça diferenciação entre caracteres maiúsculos e minúsculos. Dessa forma as expressões “amável” e “Amável” são tratadas de forma diferente se a linguagem faz essa diferenciação. Na base WordNet Affect BR as palavras de emoção estão gravadas na sua forma minúscula.

Quanto à etapa “Retira pontuação”, ela elimina os caracteres que não representam uma palavra, nesse caso, os sinais de pontuação. Geralmente eles são digitados junto às palavras sem espaços em branco entre ambos. Por isso essa etapa é responsável pelo refinamento da expressão, restando, no final do processamento, somente as palavras que serão tratadas e analisadas pela fase seguinte.

Na Tabela 5.2.1 estão descritos os caracteres de pontuação a serem extraídos pelo algoritmo. Caso opte-se por reconhecer os caracteres de “emojicons” no diálogo, o algoritmo deve ser alterado.

Após cada palavra ser separada para o processamento, a mesma é enviada para a análise da palavra. Nessa fase, o algoritmo considera que a digitação da palavra está correta, não tratando quanto à existência de erros de digitação nem o uso indevido de termos pelos usuários.

É comum ocorrer que usuários de chat usem uma comunicação informal com a presença de mensagens concisas e abreviadas, sem acentuação devida, com abreviações de palavras, e, em alguns casos, de sentenças inteiras. Ocorre, também, a substituição de palavras por outras inexistentes na língua portuguesa, criando sua própria linguagem e forma de comunicação. Por exemplo, é possível encontrar a substituição da palavra “não” pela expressão “naum” em diálogos presentes na internet. Piltcher e colegas apresentam uma proposta de correções de palavras em chat prevendo a ocorrência dessas situações em ambientes de troca de mensagens, propondo o uso de um corretor ortográfico na comparação de termos por meio de uma função de similaridade (Piltcher et al, 2002).

Dessa forma é possível estender a análise da palavra com o uso de outros recursos como: corretores ortográficos, bases de palavras que, integradas ao algoritmo, possibilitem estender a abrangência da ferramenta e reconhecer outras palavras além daquelas apresentadas na base de emoções, etc.

O presente trabalho, por limitação de tempo, não contempla a análise de palavras, tais como, verbos, palavras derivadas, variantes em número (singular e plural) e em gênero (masculino e feminino), conforme pode ser observado na Tabela 5.2.2.

Tabela 5.2.2: Exemplo da emoção “amor” e possibilidades para estender seu uso.

Emoção	Amor
Verbo	Amarei, amarás, amavam, ...
Derivada	Amorzinho, amoreco, amorzão, ...
Número (singular e plural)	Amores: na base consta “amor”.
Gênero (masculino, feminino)	Amada, pois na base consta “amado”.

Tal tratamento é importante para estender as possibilidades de uso da ferramenta no diálogo da língua portuguesa ou para atender às demandas em reconhecer padrões no processamento de textos.

Por fim, a fase de verificação da palavra e sua ocorrência na base WordNet Affect BR utiliza o relacionamento denominado de “rel_emoticon”, entre os campos da base “wordnetaffectbr” e da base “emoticon”, conforme Figura 5.2.1., para análise e localização da correspondência entre a palavra e a respectiva imagem.

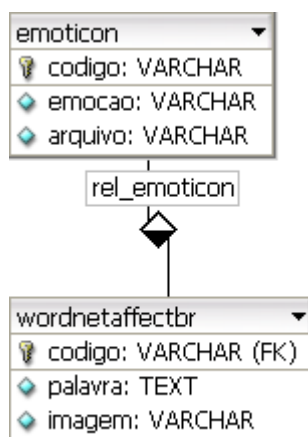


Figura 5.2.1: Modelagem da relação entre as tabelas de emoções e de imagens

Inicialmente a palavra é verificada na WordNet Affect BR. Caso seja localizada, a informação sobre o código da imagem é armazenada e processada por meio do relacionamento entre as bases. A partir do registro na base “emoticon”, é identificada a imagem que representará, visualmente, o significado da palavra na interface do usuário.

Conforme listado no Anexo D, com os grupos de palavras e respectivas imagens, algumas emoções podem ser representadas por duas ou mais imagens. Nesse caso, o algoritmo escolhe, aleatoriamente, uma das imagens vinculadas ao mesmo grupo para apresentar ao usuário. A Figura 5.2.2. mostra as imagens selecionadas para representar a emoção “raiva”. Nesse caso, se essa palavra estiver presente na mensagem, qualquer uma das cinco imagens pode ser apresentada.



Figura 5.2.2: Imagens relacionadas à emoção raiva.

Caso o resultado das fases do módulo de reconhecimento das palavras de emoções seja positivo, uma informação é repassada ao módulo de representação visual que executa os procedimentos relativos para apresentar ao usuário as imagens relacionadas.

5.3. Módulo de representação visual das palavras de emoções

O módulo de representação das palavras de emoções recebe a informação da imagem relativa à palavra da base de emoções, montando o registro de uma forma que, ao ser processado pela ferramenta de chat, as mensagens sejam apresentadas adequada e rapidamente aos usuários.

O algoritmo que monta as informações da imagem utiliza o recurso das “tags” presente na linguagem de marcação para páginas internet, o HTML. Dessa forma, as informações utilizadas pela ferramenta de chat são gravadas de duas formas distintas: na sua forma original, mantendo a mensagem do usuário conforme foi enviada e, também, no formato compatível com as tags de formatação que apontam para as respectivas imagens. Esse método agiliza o processamento das mensagens e diminui o tempo de resposta entre a leitura da base e sua apresentação na tela do usuário.

5.4. Módulo de administração da ferramenta

O módulo de administração da ferramenta foi desenvolvido com o objetivo de permitir a administração e o gerenciamento das bases de dados e as diferentes interfaces e recursos da ferramenta de chat.

As diversas funções do módulo de administração da ferramenta, desenvolvidas para esse trabalho, foram algumas das implementações que viabilizaram o desenvolvimento deste, criadas na forma de recursos e outras, como funcionalidades. São elas: manutenção das bases, geração de arquivos de “log”, consultas à base de dados, interface de tradução e validação, consulta da validação da tradução, consultas quantitativas da base de dados, inclusão de novas palavras na base WordNet Affect BR, cálculo da valência de arquivos, gerenciamento das bases de imagens de emoção e o simulador do reconhecimento e representação das imagens de emoção.

A *Manutenção das bases* permite gerenciar as tabelas e relacionamentos quanto ao conteúdo, ou seja, excluir, alterar ou inserir dados. A *geração de arquivos de “log”* é utilizada para criar arquivos com os registros relativos às mensagens do chat, gravando no formato texto para posterior processamento. Quanto à *consulta à base de dados*, é um recurso criado inicialmente para o cruzamento entre as bases WordNet e WordNet Affect para recuperar informações dos synsets e relacioná-los à informação afetiva. Os recursos *interface de tradução e validação* permitiram consultar e efetuar alterações nas bases de tradução e, também, consultar e processar os dados quanto à validação. A consulta quantitativa da base de dados é um recurso que apresenta em tela as informações quanto às quantidades dos registros das bases de dados e da ferramenta de chat. A *administração de synset* permite incluir ou excluir palavras à base da WordNet Affect BR, mantendo a integridade e relacionamento à estrutura de synsets da WordNet, atendendo às relações hierárquicas dessa base lexical. O *cálculo da valência* permite processar arquivos externos à ferramenta, como por exemplo, o “log” gravado do diálogo, gerando informações relativas às ocorrências de palavras de emoção no texto. O *gerenciamento das bases de imagens de emoção* possibilita administrar a base que relaciona as imagens com as palavras de emoção. Esse recurso é útil, pois permite alterar ou incluir novas. Por fim, o *simulador do reconhecimento e representação das imagens de emoção* é um recurso útil na validação do algoritmo que trata as expressões enviadas pelos usuários. Por meio da interface mostrada na Figura 5.4.1, o sistema permite ao usuário digitar uma mensagem, que é tratada de forma análoga ao processamento dos módulos de reconhecimento e representação existentes na ferramenta de chat. O resultado é apresentado na tela, com as palavras de emoção em negrito e as respectivas imagens ao lado.

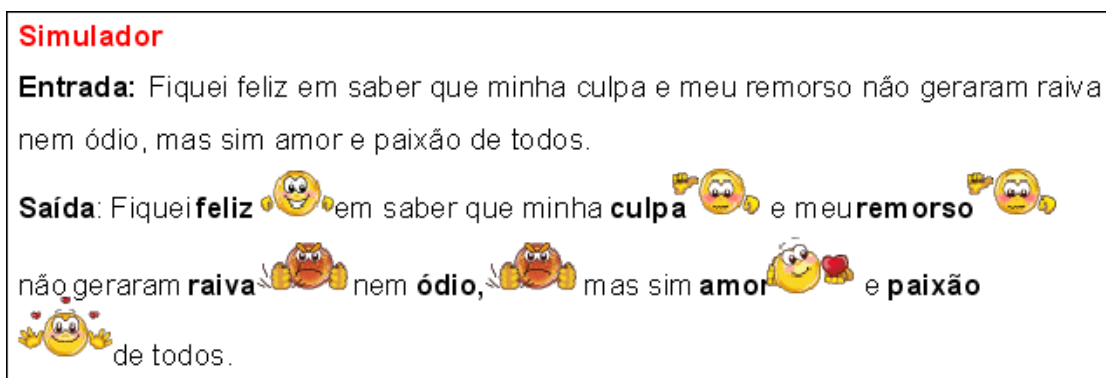


Figura 5.4.1: Simulador para reconhecer e representar imagens de emoção

A ferramenta, bem como todas as funcionalidades e recursos apresentados, foram desenvolvidos para a internet. Dessa forma, todas as facilidades implementadas são possíveis de serem executadas a distância, estendendo a possibilidade para manutenção das bases e da ferramenta.

6. Experimentos

Dentre os experimentos desse trabalho, foram realizados encontros virtuais, utilizando a ferramenta de chat, com usuários de uma turma da terceira idade e com uma turma de alunos de um curso de graduação, ambos da Feevale¹².

Esses experimentos visaram validar a base WordNet Affect BR como um recurso lexical quanto ao seu conteúdo e também para validar a ferramenta de chat como um recurso para a comunicação entre os usuários, adequado e melhorado pelas capacidades de reconhecer e representar as palavras de emoção do diálogo.

O método definido para a pesquisa foi avaliar a validade de ambos os recursos, da base e da ferramenta, com um questionário onde os usuários deveriam responder a algumas questões após terem participado dos encontros. Após, as respostas foram avaliadas, cujos resultados e considerações são apresentados na Seção 6.2.

Para o questionário foram elaboradas três questões, destacando-se que as duas primeiras visavam avaliar quanto ao uso das imagens e a última, fazer uma avaliação quanto à ferramenta de chat. As questões foram:

- a) Quanto às imagens apresentadas no chat, você consegue relacioná-las a emoções na medida que elas aparecem na tela?
- b) Você concorda que imagens possuem a capacidade de dar o mesmo significado das palavras digitadas nas mensagens?
- c) Que avaliação você faz do chat como uma ferramenta para a comunicação e integração de pessoas na internet? Gostou, não gostou ou, para você, é indiferente? Comente.

Os resultados apontaram para uma significativa concordância dos usuários com os agrupamentos de palavras e suas imagens.

¹² Centro Universitário Feevale é uma instituição de nível superior do município de Novo Hamburgo/RS. O site oficial na internet é www.feevale.br

6.1. Perfil dos usuários

A turma da terceira idade, do projeto de extensão “Inclusão Digital para a Terceira Idade” da Feevale, é composta por 18 alunos, sendo que desses, 82% são do sexo feminino; 41% dos alunos estão na faixa etária de 50 e 60 anos e 37% de 60 a 70 anos. Quanto à escolaridade, 42% dos participantes possuem ensino superior (Pasqualotti et al, 2007).

Esse grupo participa há três semestres do projeto de inclusão digital. Com relação ao conhecimento de Informática, 61% dos participantes não possuíam conhecimento da tecnologia e 39% possuíam somente um conhecimento básico.

Atualmente, esse grupo usa o computador normalmente, com aulas semanais, todos eles já tendo participado dos módulos básicos, intermediários e avançados de informática do projeto. Mesmo com alguma experiência, esse grupo apresentou algumas dificuldades iniciais relacionadas ao uso da ferramenta e à leitura na tela das informações apresentadas. Também apresentou um problema relativo à habilidade no uso do teclado e do mouse.

Foi definido que essa turma teria um contato com a ferramenta de chat para que pudessem conhecer e habituar-se à leitura e escrita das mensagens, bem como compreender o funcionamento e as regras do chat. Após uma fase de adaptação e conhecimento, foram realizados dois encontros por meio da ferramenta.

A turma de graduação, composta de 15 alunos, além da professora, é de um curso da área da Computação. Os alunos têm contato diário com computador e usam outras aplicações de comunicação, principalmente programas de mensagens instantâneas. Com essa turma não foram identificadas dificuldades em usar o chat.

6.2. Resultados e considerações

Abaixo são destacadas algumas considerações dos usuários quanto aos recursos de representação de imagens e sua relação com as palavras de emoção, avaliados nas questões “a” e “b” do questionário. As respostas fizeram uma referência positiva indicando que as imagens escolhidas estavam adequadas para representarem as palavras que foram digitadas. Essa constatação vai ao encontro dos resultados apresentados na Seção 4.1 na pesquisa realizada.

Em algumas respostas percebe-se que o uso de imagens prendeu a atenção e facilitou o entendimento do conteúdo das mensagens quanto ao seu significado. Foi destacada também a curiosidade que despertava nos usuários em descobrir quais palavras resultariam em imagens na tela e também quais seriam essas imagens. Esse foi um fator que prendeu a atenção dos usuários. Curiosamente esse relato apareceu nas avaliações das duas turmas.

O relato a seguir é de um aluno da turma de graduação e representa a capacidade que as imagens têm em transmitir significado: *“Ao olhar a imagem já podia perceber, mesmo sem ler a palavra em negrito, o que foi escrito, pelo menos o que a imagem queria transmitir. Numa das frases um colega escreveu sobre a preocupação com uma prova. Ele escreveu a palavra nervoso e a palavra apavorado, e mais uma que acho que foi que tinha batido um pavor um dia antes. As figurinhas na tela ficaram bem legais, pois como eu sabia qual o assunto que nós estávamos falando, eu já imaginei como ele estava se sentindo. Foi muito legal mesmo.”*.

Quanto à ferramenta de chat, baseado nas respostas dos usuários à questão “c”, podem ser destacadas e feitas as seguintes considerações:

A interface foi considerada simples e limpa. A atenção do usuário ficou dirigida à digitação e envio das mensagens, dispensando-o de preocupar-se em clicar em outras opções na tela. Também foi salientada a acessibilidade da interface, principalmente quanto ao tamanho das fontes e à distribuição das informações na tela.

Outro detalhe salientado pelos usuários como um dificultador foi a ocorrência de tempos de resposta muito altos. Isso ocorreu em alguns momentos dos experimentos, independente da quantidade de participantes e de mensagens enviadas. Esse fato, conforme se constatou, ocorreu na performance da rede sob diversas circunstâncias, inclusive com outras aplicações, descartando-se que esses problemas tenham sido causados pela ferramenta ou devido ao acesso à base de dados. Para minimizar o problema, foram feitos alguns testes com o tempo que a ferramenta utiliza para atualizar as mensagens na tela. Assim, a cada intervalo desse tempo, que pode ser reconfigurado, a ferramenta de chat efetuava nova leitura na base de dados, atualizando as mensagens e a lista de participantes na tela.

Quanto aos resultados dessa questão, se o usuário gostou, não gostou ou ficou indiferente à ferramenta de chat, a maioria das respostas foi positiva, conforme Tabela 6.2.1,

indicando que a mesma está adequada para ser utilizada em encontros virtuais para comunicação entre os usuários como uma ferramenta de chat.

Tabela 6.2.1: Resultados da avaliação da ferramenta de chat

	Terceira	Graduação	Total
Gostou	15	16	31
Não gostou	0	0	0
Indiferente	3	0	3

A Figura 6.2.1 apresenta uma cópia da tela de um dos encontros com a turma de graduação, com alguns recortes do diálogo.

João: Entrou na sala às 07:05:48 PM

Luis Antônio: Entrou na sala às 07:05:48 PM

Elen: Entrou na sala às 07:05:48 PM

Prof. Paulo: he he espero que estejam felizes.

João: E aí professor Paulo. Estou **feliz** 😊 em revê-lo.

Prof. Paulo: João, você tá bem? continua aquela pessoa **amável** 😊 e **feliz**? 😊

Luis Antônio: quanta **alegria**... 😊

Elen: Que lindo!

Elen: Que **agradável** 😊

João: Sempre muito **amável**. 😊 Mas quando lembro de nossas aulas fico **triste** 😞 de saudade a saudade.

Elen: Adorei os emoticons

João: E o Sr? Como vai? Não tem **vergonha** 😊 de estar brincando de chat com esses amáveis pupilos.

Luis Antônio: Paulo.. tinha q ter uma lista dos emoticons comportados... que **triste**... 😞

Elen: Tchau. Fiquei tão **triste** 😞 por ter que parar.

Figura 6.2.1: Tela do diálogo da turma de graduação

6.3. Valência

De acordo com Ortony e colegas (Ortony et al., 1988), emoções são reações com valência (positiva ou negativa) baseadas na situação que as dispara, a partir da avaliação cognitiva e psicológica que o sujeito faz diante das consequências dos eventos, das ações de agentes e da atratividade exercida pelos objetos.

A ferramenta desenvolvida no presente trabalho reconhece as palavras de emoção utilizadas pelos usuários no diálogo, representando-as por meio de imagens na interface das mensagens do chat. Um dos objetivos do trabalho é trabalhar a palavra no momento em que

ela ocorre no diálogo, porém sem determinar ou inferir sobre os estados emotivos dos sujeitos, pois tal trabalho enseja uma análise detalhada de todo o contexto no qual o diálogo ocorre, além de estudos mais aprofundados, principalmente no campo das emoções.

Utilizando medidas quantitativas quanto à ocorrência das palavras de emoção em um texto pode-se indicar a sua carga da valência, considerando-se o eixo dimensional positivo e negativo. Esses resultados são possíveis por meio do processamento de arquivos contendo as mensagens trocadas entre os usuários do chat, estendendo as possibilidades para outros tipos de texto.

Trabalhos na literatura estudada apresentam o uso das teorias dimensionais para a medição e determinação de valores de emoção baseadas em textos. Gregory e colegas (Gregory et al., 2006) e Genereux e Evans (2006) direcionaram seus esforços para o uso de técnicas para análise e identificação da carga de valência de textos e formulários para identificar a carga afetiva do seu conteúdo.

Da mesma forma, Wang e colegas (2005) utilizaram, em seus experimentos, uma estrutura bi-dimensional de valores, sendo uma delas para indicar a valência das emoções percebidas por meio de sensores fisiológicos.

Das possibilidades para o uso da base lexical de emoções, o algoritmo para identificação da valência em arquivos texto foi acrescentado à ferramenta, conforme apresentado na Seção 5.4. O algoritmo faz a leitura e a análise das palavras de um texto quanto à valência do conteúdo, apresentando dados quantitativos e estatísticos, conforme demonstrado abaixo nas Tabelas 6.3.1. 6.3.2., que apresenta os resultados do processamento de um “log” do chat em um dos experimentos desse trabalho(1) e do processamento de um texto, com uma linguagem poética, escolhido aleatoriamente para esses testes(2)

Tabela 6.3.1: resultado do processamento da valência de um log do chat(1)

TOTALIZADORES
Total de palavras lidas: 1.020
Total de palavras de emoções localizadas: 195
Valência das palavras de emoções localizadas
Positiva (+) = 147 (75,38 %)
Negativa (-) = 48 (24,62 %)

Tabela 6.3.2: resultado do processamento da valência de um texto aleatório(2)

TOTALIZADORES
Total de palavras lidas: 5640
Total de palavras de emoções localizadas: 250
Valência das palavras de emoções localizadas
Positiva (+) = 99 (39.60%)
Negativa (-) = 151 (60.40%)

O cálculo da valência é apresentado nessa seção como um recurso para uso e validação da base lexical de emoções, mas o tratamento dos resultados deverá estar relacionado aos objetivos da aplicação desse recurso em implementações específicas, não sendo objeto do presente trabalho detalhar o formato e a adequação dos resultados apresentados nas tabelas acima.

A modelagem dos dados, conforme apresentado na Seção 3.3., torna possível a recuperação da informação da valência das palavras de emoções, permitindo que o processamento possa ser feito a partir das bases WordNet Affect BR, da língua portuguesa, ou da base com as palavras de emoções na língua inglesa, conforme apresentado e detalhado no modelo ER da Figura 3.3.3.

7. Considerações finais

Dotar sistemas computacionais com a capacidade de perceber e representar significados de emoções é fundamental para a comunicação e o relacionamento entre as pessoas, pois é no entendimento dos significados, muitas vezes subjetivos do diálogo, que a interação se desenvolve. Portanto, ao pensar em sistemas computacionais para a comunicação on-line, cujo papel seja o da mediação do diálogo, deve-se pensar e projetá-los com a capacidade de darem o devido tratamento à presença de emoções nas reações e ações dos usuários.

Para isso é necessário que haja o entendimento de como os sujeitos se comportam diante de determinada situação e quais as formas e intensidades nas quais as emoções se apresentam. Dessa forma será possível desenvolver interfaces que representem emoções, bem como criar sistemas computacionais que reconheçam esses estados emocionais, semelhante à forma como ocorre com as pessoas durante uma interação. Dessa forma a base lexical desenvolvida neste trabalho pode ser utilizada em trabalhos que têm o objetivo de reconhecer palavras de emoções em recursos baseados em texto, assim como representá-las por meio de imagens.

A base WordNet Affect BR poderá ser usada no desenvolvimento de outras aplicações que façam uso de um léxico afetivo, seja quanto à sua ocorrência, como realizado no presente trabalho, ou em trabalhos que busquem inferir quanto ao estado emocional das pessoas.

A WordNet Affect BR foi desenvolvida a partir de uma teoria que apresenta as emoções, bem como justifica as formas de surgimento de cada uma delas pela capacidade cognitiva das pessoas em interpretar significados diante da sua interpretação de mundo. Da mesma forma, a metodologia utilizada para estender a abrangência da base de palavras utiliza a WordNet e a WordNet Affect, recursos lexicais amplamente validados e utilizados pelas áreas da Computação Afetiva e o PLN.

Essa ferramenta de chat utiliza a base lexical de emoções agregando recursos adicionais àqueles usualmente encontrados em programas desse tipo. Além de permitir aos usuários enviar e receber mensagens, a ferramenta incorpora capacidades de reconhecimento e representações de palavras de emoção numa forma visual, com imagens que transmitem e representam os significados das palavras.

A funcionalidade da ferramenta em utilizar imagens para representar palavras de emoções agrega à capacidade de comunicação outras formas pelo qual as pessoas percebem e reconhecem os significados presentes no diálogo.

7.1. Trabalhos futuros

O trabalho desenvolvido, em especial, os experimentos realizados apontaram para outras possibilidades tanto para as funcionalidades da ferramenta de chat quanto para o aumento da abrangência da base WordNet Affect BR, permitindo atender a outras aplicações.

Quanto à base lexical, podem ser destacados como trabalho futuro:

- a) inclusão dos verbos afetivos, considerando os diversos tempos e formas verbais da língua portuguesa;
- b) tratamento de outras formas gramaticais nas quais as palavras são formadas, como: palavras derivadas, de gênero e número, palavras compostas, anigramas, entre outras;
- c) uso de expressões regionais, gírias e metáforas, palavras típicas de uso localizado, tais com a expressão “faceiro” que define uma pessoa que está alegre ou feliz;
- d) adequar a base para armazenar expressões na forma de “emoticons”, conforme detalhado no Capítulo 4;
- e) estudar outras expressões que possam indicar emoções como satisfação ou insatisfação, por exemplo: interjeições “bah”, “tri” e “legal!”;
- f) adequar uso de imagens para representar significado de emoções. Estudar e implementar novas possibilidades, baseados nas sugestões dos usuários que responderam ao formulário de avaliação;
- g) incluir na base lexical outros estados afetivos, baseados nas categorias elicitadas pelos autores estudados (Ortony e colegas, 1987) e presentes na WordNet Affect, tais como: traços de personalidade, humor, comportamento, entre outros.

Quanto às possibilidades para outras implementações, seja adequando a ferramenta de chat, ou criando outras aplicações computacionais, podem ser destacadas:

- a) atender o uso da Língua Brasileira de Sinais, as LIBRAS, para representar ao usuário as palavras de emoção com animações nessa língua;
- b) utilizar em aplicações que trabalhem com emoções específicas para aquela demanda. Por exemplo, analisar nível de satisfação do cliente onde as emoções contente, feliz, raiva, alegre, satisfeito, possam ser tratadas adequadamente para os propósitos da aplicação.
- c) implementar a avaliação da ferramenta pelos usuários com a metodologia “likert scale”, onde o usuário possa avaliar adequadamente a ferramenta considerando as respectivas escalas de valor para “gostar”, conforme apresentado na Seção 6, alínea “c”;
- d) implementar o cálculo da valência para aplicações que incorporam a análise de textos como entrada do processamento.

O trabalho apresentado foi o primeiro passo para o desenvolvimento de trabalhos relacionados à subjetividade e afetividade no discurso da Língua Portuguesa. A base lexical WordNet Affect BR é um recurso que pode ser expandido na medida que novas implementações forem surgindo. Essa base poderá servir para outros propósitos de estudos e pesquisas envolvendo o uso de um léxico de emoções, estando disponível para propósitos dessa natureza.

Referências

- (Baségio, 2006). Baségio, Túlio Lima. Uma abordagem Semi-Automática para identificação de Estruturas Ontológicas a partir de Textos da Língua Portuguesa do Brasil. Dissertação (Mestrado em Computação) – Faculdade de Informática – PUC-RS, 2006.
- (Botelho et al., 2002) Botelho, F. C., Torres, K. S., Braga, J. L. Reengenharia e Visualização de Conceitos no WordNet. Revista Eletrônica de Iniciação Científica - REIC, REIC, v. II, p. 01-16, 2002.
- (Damásio, 1996). Damásio, Antônio R. O Erro de Descartes: Emoção, Razão e o Cérebro Humano, São Paulo: Companhia das Letras, 1996.
- (Dias, 2003). Dias, Claudia. Usabilidade na WEB: criando portais mais acessíveis. Rio de Janeiro, Editora Alta Books, 2003.
- (Elliot, 1992) Elliot, C. D. The Affective Reasoner: a process model of emotions in a multi-agent system. Ph.D. thesis, Northwestern University, Evanston, Illinois, 1992.
- (Ekman, 1993) Ekman, P.: Facial Expression and Emotion. American Psychologist, 1993. 48, 384-392
- (Fellbaum, 1998) Fellbaum, C. WordNet: an electronic lexical database. MIT Press. 1998
- (Fernandes, 2005) Fernandes, Rodrigo Panchiniak. A integração dos papéis qualia para redes semânticas. Dissertação (Mestrado em Linguística) – Centro de comunicação e expressão – Universidade Federal de Santa Catarina, 2005.
- (Fridja, 1994) Fridja, N. Varieties of affect: emotions and episodes, moods, and sentiments. In: EKMAN, P.; DAVIDSON, R. J. (Ed.). The Nature of Emotion. New York: Oxford University Press, 1994. p. 59-67.
- (Genereux and Evans, 2006) Genereux, Michel; Evans, Roger. Towards a validated model for affective classification of texts Proceedings of the Workshop on Sentiment and Subjectivity in Text, pages 55–62, Sydney, July 2006. c 2006 Association for Computational Linguistics.

- (Gregory et al., 2006) Gregory, Michelle L.;Chinchor, Nancy;Whitney, Paul; Carter, Richard; Hetzler,Elizabeth; Turner, Alan. User-directed Sentiment Analysis: Visualizing the Affective Content of Documents. Proceedings of the Workshop on Sentiment and Subjectivity in Text, pages 23–30, Sydney, July 2006. c2006 Association for Computational Linguistics
- (Jaques e Vicari, 2005) Jaques, Patrícia Augustin. Vicari, Rosa Maria. Estado da arte em ambientes inteligentes de aprendizagem que consideram a afetividde do aluno. Artigo publicado na revista Informática na Educação: Teoria & Prática, vol.8, número 1, 2005. Disponível em <http://www.pgie.ufrgs.br/revista>)
- (Jurafsky e Martin, 2000) Jurafsky, D.; Martin, J. Speech and Language Processing: an Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition, Upper Saddle. 2000. River, New Jersey, Prentice Hall. (pp. 1-18)
- (Liu et al., 2003) Liu. Hugo, Lieberman Henry, Selker Ted. A model of textual affect sensing using real-world knowledge, Proceedings of the 8th international conference on Intelligent user interfaces, January 12-15, 2003, Miami, Florida, USA
- (LREC, 2004) LREC 2004. International Conference on Language Resources and Evaluation Disponível em <http://www.lrec-conf.org/lrec2004/index.php>
- (Ma et al., 2005). Ma, Chunling. Prendinger, Helmut. Ishizuka, Mitsuru. Emotion Estimation and Reasoning Based on Affective Textual Interaction, in Affective Computing and Intelligent Interaction (First Int'l Conf. ACII 2005) (J. Tao, T. Tan and R. W. Picard (Eds.)), Springer LNCS 3784, Beijing, China, pp.622-628 (2005.10)
- (Ma et al., 2005). Ma, Chunling. Prendinger, Helmut. Ishizuka, Mitsuru. A Chat System Based on Emotion Estimation from Text and Embodied Conversational Messengers, in Entertainment Computing -- ICEC 2005 (Proc. IFIP 4th Int'l Conf. on Entertainment Computing-2005, Kobe-Sanda, Japan), Springer LNCS 3711, pp.535-538 (2005.9)
- (Miller, 1995) Miller, A.G. WordNet: A Lexical Database for English, Communications of the ACM, Vol. 38, No.11, November 1995, pp. 39 - 41.
- (Ortony et al., 1987) Ortony, A.; Clore, G.L.; Foss, M.A. The Referential Structure of the Affective Lexicon. Universiry of Illinois at Urbana-Champaign (1987)

- (Ortony et al., 1988) Ortony, A.; Clore, G.L.; Collins, A. "The Cognitive Structure of Emotions", Cambridge University Press. 1988
- (Pasqualotti, et al, 2006). Pasqualotti, P.R; Pérez, C.C.C; Bez, R.; Klein, C. Inclusão Digital para Terceira Idade: oportunidades, possibilidades e propostas inovadoras. CIIE-Congresso ibero-americano de inclusão. Argentina, 2007
- (Picard, 1997) Picard, Rosling. Affective computing. Cambridge, Massachusetts Institute of Technology. The MIT Press, 1997.
- (Piltcher et al, 2002) Pitcher, Gustavo. Borges, Thyago. Loh, Stanley. Lichtnow, Daniel. Simões, Gabriel. Correção de palavras em chats: Avaliação de bases para dicionários de referência. XXV Congresso da Sociedade Brasileira de Computação. Unisinos. São Leopoldo-RS, 2002.
- (Prendinger and Ishizuka, 2001) Prendinger, Helmut; Ishizuka, Mitsuru. Simulating Affective Communication with Animated Agents Proc. Eighth IFIP TC.13 Conference on Human-Computer Interaction (INTERACT 2001), pp.182-189, Tokyo, Japan (2001.7).
- (Prendinger and Ishizuka, 2001) Helmut Prendinger and Mitsuru Ishizuka: Affective and Social Behavior in Animated Agents (Invited Paper), CD-Rom Proc. 9th IFSA World Congress and 20th NAFIPS Int'l Conf., Session on Kansei Systems, pp.2918-2923, Vancouver, Canada (2001.7).
- (Scherer, 2000). Scherer, K.R. Psychological models of emotion. In Joan C. Borod (Ed.), The Neuropsychology of Emotion (pp. 137–162). New York: Oxford University Press.
- (Specia and Nunes, 2004). Specia, Lucia. Nunes, Maria das Graças Volpe. Desambiguação Lexical Automática de Sentido: Um Panorama. Série de Relatórios do Núcleo Interinstitucional de Linguística Computacional. NILC - ICMC-USP, Caixa Postal 668, 13560-970 São Carlos, SP, Brasil, 2004
- (Valitutti et al., 2004) Valitutti, Alessandro, Strapparava, Carlo e, Sttock, Oliveira. Developing Affective Lexical Resources. PsychNology Journal, 2004 Volume 2, Number 1, 61 – 83

- (Valitutti et al., 2004) Valitutti, Alessandro, Strapparava, Carlo. WordNet-Affect: an affective extension of WordNet. In Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004), Lisbon, May 2004, pp. 1083-1086.
- (Wang et al., 2005) Hua Wang, Helmut Prendinger, Mitsuru Ishizuka, and Takeo Igarashi: Affective Communication in Online Chat Using Physiological Sensors and Animated Text, Trans. Human Interface Society, Vol.7, No.1, pp.39-45. 2005
- (Wikipédia, 2007). WEB. Disponível em <http://pt.wikipedia.org/wiki/Emoticon>
Acessado em 03 de outubro de 2007.
- (WordNet, 2007) WordNet: a lexical database for the English language. Cognitive Science Laboratory. Princeton University. 221 Nassau St. Princeton, NJ 08542
Disponível em <http://WordNet.princeton.edu/>

ANEXO A - arquivos da WordNet

A partir dos arquivos de dados da WordNet, foram extraídos os registros para criar as demais bases lexicais da base de dados da ferramenta. Os arquivos abaixo descritos deram origem às tabelas do sistema, de forma a contemplar as necessidades de organização da estrutura de acordo com os objetivos, ou seja, cada tabela criada contém elementos necessários para localizar e identificar os termos afetivos, bem com as demais relações entre eles, contemplados na WordNet.

Arquivo DAT

O arquivo Dat contém informações correlacionando os códigos dos synsets aos termos definidos e especificados para a base da WordNet.

Cada synset possui um ou mais termos lexicalmente categorizados pelo sentido do termo. O registro inclui ainda informações sobre a categoria gramatical (POS) e as relações entre o synsets e os demais synsets da WordNet.

Exemplo de um registro do arquivo DAT para o código de synset 01308490 (happy)

01308490 00 s 02 happy 0 pleased 0 001 & 01308059 a 0000 | experiencing pleasure or joy; “happy you are here”; “pleased with the good news”

Formato do registro DAT

synset_offset lex_filenum ss_type w_cnt word lex_id [word lex_id...] p_cnt [ptr...] [frames...] | gloss

Onde:

- a) synset_offset: 8 dígitos representando o código do synset na base;
- b) lex_filenum: dois dígitos que identificam o arquivo léxico que contém o synset;
- c) ss_type: um caracter indicando o código de POS (“v” para verbo; “a” para adjetivo; “s” para adjetivo satélite; “n” para substantivo e “r” para advérbio) do synset;

- d) w_cnt: dois dígitos indicando a quantidade de palavras (sentidos) do synset;
- e) word: é cada sentido do synset. A quantidade de termos está definida em w_cnt;
- f) lex_id: identifica cada sentido a um arquivo léxico. O padrão é zero;
- g) p_cnt: quantidade de ponteiros aponta para esse synset. Ponteiro pode ser entendido como um synset que possui alguma relação lexical com o synset, onde essa relação é representada por um símbolo (pointer symbol);
- h) pointer symbol (no exemplo, o símbolo “&” representa uma sinonímia entre o synset e aquela apontada pelo id). Cada pointer possui um “pointer symbol” correspondente;
- i) ptr: é o id do synset que possui uma relação com o synset relacionado;
- j) pos: um caracter indicando o código de POS (“v” para verbo; “a” para adjetivo; “s” para adjetivo satélite; “n” para substantivo e “r” para advérbio) do ponteiro;
- k) source/target: quatro dígitos que identificam a relação entre os dois synsets (o ponteiro e aquele apontado), de acordo com o símbolo dessa relação;
- l) gloss: um texto contendo a glossa, ou seja, a descrição, semelhante ao dicionário, do termo representado por aquele synset, de acordo com as palavras (Word) que ele representa.

A relação entre o synsets e os ponteiros com os quais existe uma relação forma um dos recursos mais importantes da WordNet, que é capacidade da base em apontar os relacionamentos entre os synsets, portanto, permite uma expansão da base para uma quantidade e abrangência além da sua capacidade, limitada pela estrutura física, mas aumentada pela estrutura das relações.

A tabela A.1 apresenta as quantidades físicas de cada categoria (POS) e as relações com outros synsets, definidas pela relação lexical encontrada no arquivo DAT.

Tabela A.1: relação entre quantidade de synsets e relações entre os ponteiros

Tabela de relações entre os synsets da base WordNet		
POS	Quantidade lida	Relações (ponteiros)
Adjetivos	17.915	33.455
Advérbios	3.575	3.861
Substantivos	66.025	175.091
Verbos	12.127	25.748
Total	99.642	238.155

Fonte: WordNet 1.6 Arquivo DAT

O total mostra que dos 99.642 synsets (sentidos) encontrados na base, existem possibilidades de 238.155 reações lexicais, seja de sinonímia, antonímia, meronímia, holonímia ou qualquer outra definida e contemplada na estrutura da WordNet.

Arquivo IDX

Outro arquivo utilizado na criação das bases de dados foi o arquivo IDX. Esse arquivo é um índice listado em ordem alfabética com todas as palavras encontradas na WordNet, relacionando os sentidos da palavra na base.

Exemplo de um registro do arquivo IDX:

happy a 6 4 ! & ^ = 6 4 01099314 01308490 00998186 00556743 00863471 00948198

Formato do registro IDX

lemma pos synset_cnt p_cnt [ptr_symbol...] sense_cnt tagsense_cnt synset_offset
[synset_offset...], onde:

- lemma: é a palavra/expressão da base WordNet;
- pos: um caracter indicando o código de POS (“v” para verbo; “a” para adjetivo; “s” para adjetivo satélite; “n” para substantivo e “r” para advérbio) do ponteiro;
- synset_cnt: número de synsets que a palavra possui. Este é o número de sentidos da palavra na WordNet, de acordo com o constante no arquivo DAT.

- d) `p_cnt`: número de ponteiros que a palavra possui;
- e) `ptr_symbol`: seqüência de símbolos que identificam as relações dos synsets;
- f) `sense_cnt`: repete o número de sentidos do synset;
- g) `tagsense_cnt`: mantido por compatibilidade com outras versões da WordNet;
- h) `synset_offset`: Seqüências de synsets onde cada 8 algarismos identificam o código synset da palavra na WordNet.

ANEXO B - relações entre synsets da WordNet.

Tabela B.1. Principais relações da WordNet

	Relações
Substantivos	antonímia, hipernímia, hiponímia, holonímia, meronímia e atributo de...
Verbos	antonímia, hipernímia, hiponímia, implicações, causa, ver também... e grupo de verbo
Adjetivos	antonímia, similar a, pertencer a ...(relação com substantivo), atributo, ver também...
Advérbios	antonímia e derivado de adjetivo





ANEXO C – a base WordNet Affect BR








A base lexical de emoções para a língua portuguesa, desenvolvida no presente trabalho, é denominada de WordNet Affect BR e possui na sua versão atual 289 palavras.

abalado	caloroso	execração	nobre
abatido	carinhoso	expectante	obsessão
abominação	cativado	expectativa	ódio
abrandamento	coitado	expição	ofendido
acanhado	colérico	exultação	olho grande
acanhamento	comiseração	exultante	orgulho
acrimonioso	compaixão	felicidade	orgulhoso
admiração	compensação	feliz	orgulhoso de
admirado	complacente	fraco	otimista
admirável	compunção	frustrado	paixão
adorável	concupiscência	fúria	pânico
adversidade	confiança	furioso	paternal
afeição	conflito	galante	patético
afeiçoado	consciência pesada	gana	patos
afetuoso	consternação	gosto	pavor
aflição	constrangido	gratidão	pavoroso
aflito	constrangimento	gratificação	pena
afortunado	contentamento	gratificante	penitência
agonia	contente	grato	penitente
agoniado	contrariado	horrendo	perturbado
agonizante	contrito	horrível	pesaroso
agradável	coração mole	horror	piedade
agradecido	covarde	horroroso	pobre
ajuda	covardemente	hostilidade	prazer
alarmante	covardia	humilhação	prazer no sofrimento alheio
alegre	cuidadoso	humilhado	preocupação
alegria pelo mau alheio	culpa	idolátrico	preocupado
alegria	culpado	ignóbil	presunção
aliviado	de dar pena	ignomínia	presunçoso
alívio	decepção	impressionante	problema
altivo	decepcionado	incitado	radiante
amado	dedicado	indenização	raiva
amargo	deleite	indignado	raivoso
amargura	delicado	infâmia	rancor
amável	deplorável	infeliz	rancoroso

amor	derrotado	inferior	realização
amoroso	desafortunado	inflamado	receoso
angústia	desapontado	inimizado	regozijo
angustiado	descanso	inquieto	remorso
angustiante	desconsolado	insegurança	repugnância
animado	desejo	inseguro	ressentido
ansiedade	desejo sexual	integração	ressentimento
antecipativo	desembaraço	inveja	reverência
apaixonado	desgraça	ira	revoltado
apavorado	desgraçado	irado	romântico
apavorante	desonra	irascível	satisfação
aprazível	desonrado	irritação	satisfeito
apreciação	desprezível	irritadiço	sentimento de culpa
apreciativo	devotadíssimo	irritado	soberba
apreço	devotado	irritável	socorro
ardor	dívida	jovialidade	sofrimento
arrependido	elação	júbilo	solidariedade
arrependimento	elevação	jubiloso	sossego
assistência humanitária	elevado	lamentável	suavizado
assombro	enamorado	lastimoso	submisso a esposa
assustador	encabulado	louco	temeroso
atenção	encantado	louco de raiva	temível
atenuado	enfezado	mágoa	temor
atração	enfurecido	majestoso	terno
atração física	enlouquecido	mal	terrível
atraído	entusiasmo	malevolência	terror
auto-complacência	envergonhado	malignidade	tormento
autodepreciação	espanto	medo	tortura
auto-reprovação	esperança	medonho	triste
auto-satisfação	esperançoso	medroso	tristeza
auxílio	esplendor	miserável	vaidoso
aversão	estima	misericórdia	veneração
azarado	estranheza	moderação	venturoso
beatífico	estressado	muito mal	vergonha
beatitude	exaltado	na expectativa	vermelho de raiva
bem-aventurança	exasperado	namorico	zangado
brabo	excitação	namoro	zeloso
bravo			

Anexo D – grupos de palavras e imagens de emoção

Grupo de palavras	Imagem
<p>Grupo 1:</p> <p>ADJETIVOS: afortunado, alegre, beatífico, elevado, exultante, feliz, jubiloso, venturoso</p> <p>SUBSTANTIVOS: alegria, alegria pelo mau alheio, beatitude, bem-aventurança, deleite, elação, elevação, entusiasmo, esplendor, excitação, exultação, felicidade, jovialidade, júbilo, prazer, prazer no sofrimento alheio</p>	
<p>Grupo 2:</p> <p>ADJETIVOS: animado, complacente, contentamento, contente, presunçoso, radiante, satisfeito</p> <p>SUBSTANTIVOS: auto-complacência, auto-satisfação, compensação, expiação, gratificação, indenização, integração, presunção, realização, regozijo, satisfação</p>	
<p>Grupo 3:</p> <p>ADJETIVOS: altivo, galante, impressionante, majestoso, nobre, orgulhoso, orgulhoso de...</p> <p>SUBSTANTIVOS: orgulho, soberba</p>	
<p>Grupo 4:</p> <p>ADJETIVOS: azarado, coitado, de dar pena, deplorável, desafortunado, desconsolado, desgraçado, desprezível, infeliz, lamentável, lastimoso, miserável, patético, pobre</p> <p>SUBSTANTIVOS: comiserção, compaixão, misericórdia, patos, pena, piedade, solidariedade</p>	
<p>Grupo 5:</p> <p>ADJETIVOS: abrandamento, aliviado, atenuado, suavizado</p> <p>SUBSTANTIVOS: ajuda, alívio, assistência humanitária, auxílio, descanso, desembaraço, moderação, socorro, sossego</p>	
<p>Grupo 6:</p> <p>ADJETIVOS: decepcionado, derrotado, desapontado, frustrado</p> <p>SUBSTANTIVOS: decepção</p>	
<p>Grupo 7:</p> <p>ADJETIVOS: abalado, abatido, aflito, agoniado, agonizante, angustiado, angustiante, contrariado, estressado, inquieto, perturbado, preocupado, triste</p> <p>SUBSTANTIVOS: adversidade, aflição, agonia, angústia, conflito, mágoa, problema, sofrimento, tormento, tortura, tristeza</p> <p>* A tradução gerou ainda dois advérbios para esse grupo: mal e muito mal.</p>	
<p>Grupo 8:</p> <p>ADJETIVOS: acanhado, constrangido, culpado, desonrado, encabulado, envergonhado, humilhado</p> <p>SUBSTANTIVOS: acanhamento, autodepreciação, constrangimento, desgraça, desonra, humilhação, ignomínia, infâmia, vergonha</p>	

<p>Grupo 9:</p> <p>ADJETIVOS: acrimonioso, amargo, brabo, bravo, colérico, enfezado, enfurecido, enlouquecido, exaltado, exasperado, furioso, ignóbil, incitado, indignado, inflamado, irado, irascível, irritadiço, irritado, irritável, louco, louco de raiva, ofendido, raivoso, rancoroso, ressentido, revoltado, vermelho de raiva, zangado</p> <p>SUBSTANTIVOS: abominação, amargura, aversão, dívida, execração, fúria, gana, hostilidade, inimizade, inveja, ira, irritação, malevolência, malignidade, ódio, olho grande, raiva, rancor, repugnância, ressentimento</p>	
<p>Grupo 10:</p> <p>ADJETIVOS: arrependido, contrito, penitente, pesaroso</p> <p>SUBSTANTIVOS: arrependimento, auto-reprovação, compunção, consciência pesada, culpa, penitência, remorso, sentimento de culpa</p>	
<p>Grupo 11:</p> <p>ADJETIVOS: antecipativo, esperançoso, expectante, otimista</p> <p>SUBSTANTIVOS: ansiedade, confiança, esperança, expectativa, na expectativa</p>	
<p>Grupo 12:</p> <p>ADJETIVOS: adorável, afeiçoado, afetuoso, amado, amável, amoroso, apaixonado, atraído, caloroso, carinhoso, cativado, cuidadoso, dedicado, delicado, devotadíssimo, devotado, enamorado, encantado, idolátrico, paternal, romântico, submisso a esposa, terno, zeloso</p> <p>SUBSTANTIVOS: amor, ardor, atração, atração física, concupiscência, coração mole, desejo, desejo sexual, namorico, namoro, obsessão, paixão</p>	
<p>Grupo 13:</p> <p>ADJETIVOS: alarmante, apavorado, apavorante, assustador, covarde, covardemente, fraco, horrendo, horrível, horroroso, inferior, inseguro, medonho, medroso, pavoroso, receoso, temeroso, temível, terrível, vaidoso</p> <p>SUBSTANTIVOS: assombro, atenção, consternação, covardia, espanto, estranheza, horror, insegurança, medo, pânico, pavor, preocupação, temor, terror</p>	
<p>Grupo 14:</p> <p>ADJETIVOS: admirado, admirável</p> <p>SUBSTANTIVOS: admiração, afeição, apreciação, apreço, estima, gosto, reverência, veneração</p>	
<p>Grupo 15:</p> <p>ADJETIVOS: agradável, agradecido, aprazível, apreciativo, gratificante, grato</p> <p>SUBSTANTIVOS: gratidão</p>	

Anexo E – autorização para uso da WordNet Affect

Abaixo o mail do grupo de pesquisa do Itc-Irst autorizando o uso da base WordNet Affect e WordNet Domain, recebido em 07/11/2006, às 15h10min, via mail:

Dear Paulo,

we have received your request for a WordNet-Domains licence and we thank you for your interest.

Please find attached WordNet-Domains 3.1.

The distribution of WordNet Domains includes WordNet Affect, an additional hierarchy of “affective domain labels”, with which the synsets representing affective concepts are further annotated.

The README file provides a short description of all single files.

In particular, you will find three data files:

- “wn-domains-2.0-20050210” contains the mapping between WordNet 1.6 synsets and their corresponding domains (based on the new Domain Hierarchy, see below). Each line of the file reports:

- (i) synset offset - part of speech
- (ii) the list of domains associated to the synset.

Example: 00081298-n baseball play

- “wn-domains-3.1beta-20050907” [NEW with regard to WN-Domains-2.0] contains a beta version of the mapping between Princeton WordNet 2.0 synsets and their corresponding domains (based on the new Domain Hierarchy, see below). The format is as above.

- “wn-affect-labels-1.0.txt” contains the mapping between some WordNet 1.6 synsets and their corresponding affective domains.

Each line of the file reports:

- (i) part of speech # synset offset
- (ii) label that indicates if the synset belongs to the manually annotated core or if it was automatically annotated applying the specified wordnet relation
- (iii) one or more affective labels, separated by “|”.

Examples:

n#05611157|core|emo|moo

a#00479718|similar-to|emo

Since version 2.0, WordNet Domains has adopted a revised Domain Hierarchy. The updated version is described in the following paper (included in the distribution along with other documents comparing the old and new version of the hierarchy):

Luisa Bentivogli, Pamela Forner, Bernardo Magnini and Emanuele Pianta, "Revising WordNet Domains Hierarchy: Semantics, Coverage, and Balancing", Proceedings of COLING 2004 Workshop on "Multilingual Linguistic Resources", Geneva, Switzerland, August 28, 2004, pp. 101-108.

Please contact us for any problem or question.

Regards,

Manuela Speranza

>-----

>PERSONAL DATA

>-----

>

>First Name: Paulo

>Last Name: Pasqualotti

>Affiliation: University Center Feevale and University Vale dos Sinos

>Title in the Institution / Company: Researcher and teacher

>Advisor:

>Usage of WordNet Domains : Research and use of PLN for getting lexical information in texts and identification of the person's affective states, with aged peoples in special case.

>Country: Brasil

>State:

>City: NOVO HAMBURGO

>Zip: 93352-000

>Phone: 055 (51) 3586-8800

>Fax:

>E-mail: ppasqualotti@feevale.br

>Personal homepage:

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)