

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

**ESCOP:
Estratégia de Suporte
à Solução Cooperativa de Problemas**

por

ANTÔNIO RODRIGO DELEPIANE DE VIT

Dissertação submetida à avaliação como requisito parcial para a obtenção do
grau de Mestre em Ciência da Computação



Dra. Liane Margarida Rockenbach Tarouco
Orientadora

UFRGS
Instituto de Informática
Biblioteca

Porto Alegre, junho de 2000.

CIP - CATALOGAÇÃO DA PUBLICAÇÃO

VIT, Antônio Rodrigo Delepiane de

ESCOPE: Estratégia de Suporte à Solução Cooperativa de Problemas / por Antônio Rodrigo Delepiane de Vit. – Porto Alegre : PPGC da UFRGS, 2000.

116f. : il.

Dissertação (mestrado) – Universidade Federal do Rio Grande do Sul. Programa de Pós-graduação em Computação, Porto Alegre, BR – RS, 2000. Orientadora: Tarouco, Liane Margarida Rockenbach.

1. CSCW. 2. PBL. 3. Trabalho Cooperativo. 4. Aplicações Internet. I. Tarouco, Liane Margarida Rockenbach. III. Título.

Aplicação da Computação - SBO
Informática Educacional
Ensino a distância
Trabalho Cooperativo
ENP 1 05.01.01.00

INSTITUTO DE INFORMÁTICA BIBLIOTECA	
N.º CHAMAD.	N.º REG.
341.681.32(043) V335e	39357
ORIGEM:	DATA:
<i>D</i>	05.01.01. R\$20,00
FUNDO:	CLASS.
II	II

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitora: Profa. Wrana Panizzi

Pró-Reitor de Pós-Graduação: Prof. Franz Rainer Semmelmann

Diretor do Instituto de Informática: Prof. Philippe Olivier Alexandre Navaux

Coordenadora do PPGC: Profa. Carla Maria Dal Sasso Freitas

Bibliotecária-Chefe do Instituto de Informática: Beatriz Regina Bastos Haro

Agradecimentos

A Deus, por Sua presença constante e protetora.

Aos meus pais, Renato e Alizene, minha querida "nona" Irma, minha estimada tia Nilsa e a meus irmãos Patrícia e Estevão, que são as pessoas que sempre me acompanharam, mesmo que a distância, dando-me todo o apoio e incentivo.

À Sandra, minha namorada, por ter estado ao meu lado, pelos incentivos, apoio, carinho e dedicação.

À professora Liane Margarida Rockenbach Tarouco, que além de me orientar durante todo o desenvolvimento deste trabalho, foi a pessoa que me possibilitou a primeira oportunidade de trabalhar com pesquisa.

Ao estimado amigo João Carlos Cavalcanti da Silveira, pela disposição de sempre me ajudar e orientar, pela amizade e pela presença nos "momentos complicados" que sempre surgiram.

Aos amigos que fiz e/ou que se fizeram presentes durante o mestrado e que, de alguma forma, tornaram este período da minha vida bem menos conturbado através de sua presença alegre, "tragos" e momentos memoráveis. São pessoas com quem espero não perder mais o contato (*in alphabetical order*): Adriana Pereira, Alessandro Copetti, Alexandre Dalepiane, Cléo Junior, Fauzi Shubeita, Fábio Schneider, Luciano Ferreira, Maurício Fiorese e Régis Guitel.

Aos professores Rosa Maria Viccari e Simão Sirineu Tocani, do PPGC da UFRGS, pela ajuda prestada na realização deste trabalho.

À UNICRUZ - Universidade de Cruz Alta, pelo auxílio financeiro concedido através da bolsa de estudos e aos colegas Evandro Kruehl, Gilnei Luis de Moura, José Carlos Nunes Medina e Marco Aurélio Spohn pela amizade e ajuda técnica dispensadas.

Ao colega do PPGC, Leandro Krug Wives, pelo desprendimento em ajudar-me sempre que necessário e pela cooperação quando da utilização prática de seus trabalhos; e ao colega Juliano Malacarn que, além da ajuda técnica, sempre mostrou-se disponível e prestativo quando solicitado.

Aos funcionários do PPGC, pela competência, presteza e boa vontade que sempre demonstraram.

Sumário

Lista de Abreviaturas	6
Lista de Figuras.....	8
Lista de Tabelas	10
Resumo	11
Abstract	12
1 Introdução.....	13
1.1 MOTIVAÇÃO.....	14
1.2 O TRABALHO DESENVOLVIDO	14
2 Suporte ao trabalho cooperativo	16
2.1 CARACTERÍSTICAS DO TRABALHO COOPERATIVO INTERMEDIADO POR COMPUTADOR	16
2.2 FERRAMENTAS DE SUPORTE AO TRABALHO COOPERATIVO.....	19
2.2.1 <i>Telnet</i>	20
2.2.2 <i>FTP</i>	20
2.2.3 <i>E-Mail</i>	21
2.2.4 <i>Newsgroup</i>	22
2.2.5 <i>IRC</i>	22
2.2.6 <i>ICQ</i>	22
2.2.7 <i>CU-SeeMe</i>	23
2.2.8 <i>WWW</i>	23
2.2.9 <i>BBS</i>	24
2.3 ANÁLISE COMPARATIVA DAS FERRAMENTAS EXISTENTES.....	24
3 Suporte à Cooperação em EAD.....	26
3.1 A EDUCAÇÃO A DISTÂNCIA VIA INTERNET.....	29
3.2 MOTIVAÇÕES PARA O TRABALHO COOPERATIVO EM EAD	32
3.3 CARACTERÍSTICAS DESEJÁVEIS NAS FERRAMENTAS DE SUPORTE À COOPERAÇÃO	32
3.4 SISTEMAS DE APOIO À COLABORAÇÃO E SEU USO EM PBL	33
3.5 FERRAMENTAS PARA SOLUÇÃO COOPERATIVA DE PROBLEMAS.....	36
3.5.1 A Ferramenta <i>WebSaber</i>	37
3.5.2 A Ferramenta <i>TANGO Interactive</i>	40
3.5.3 A Ferramenta <i>Interloq</i>	45
3.5.4 A Ferramenta <i>SAACI</i>	51
3.6 ANÁLISE COMPARATIVA DOS SISTEMAS INSPECIONADOS.....	53
4 Proposta de um Sistema para Apoio ao Trabalho Cooperativo em EAD.....	56
4.1 TRABALHOS PRÉVIOS CONSIDERADOS.....	58
4.1.1 Língua Natural: Classificação de Orações	59
4.1.2 Um protótipo da ferramenta " <i>Issue Analyser</i> ": projeto, implementação e validação.....	66
4.1.3 Um estudo sobre Agrupamento de Documentos Textuais em Processamento de Informações não Estruturadas Usando Técnicas de " <i>Clustering</i> "	75
4.2 DESCRIÇÃO DO MODELO DO SISTEMA ESCOP	87
5 Descrição da Implementação	89

5.1 VISÃO GERAL DO SISTEMA ESCOP	89
5.2 AMBIENTE DE IMPLEMENTAÇÃO	90
5.3 DESCRIÇÃO FUNCIONAL DOS MÓDULOS	92
5.4 DESCRIÇÃO DA INTERFACE	95
5.5 CONSIDERAÇÕES GERAIS SOBRE A IMPLEMENTAÇÃO	104
6 Conclusões	106
6.1 CONTRIBUIÇÕES DESTE TRABALHO	109
6.2 SUGESTÕES E TRABALHOS FUTUROS	109
Bibliografia	112

Lista de Abreviaturas

3D	<i>3 Dimension</i>
ADS	<i>Árvore de Derivação Sintática</i>
BBS	<i>Bulletin Board System</i>
C/C++	<i>C Programming Language / C++ Programming Language</i>
CCITT	<i>Comité Consultatif International Téléphonique et Télégraphique</i>
CGI	<i>Common Gateway Interface</i>
CMC	<i>Comunicação Mediada por Computador</i>
CS	<i>Coeficiente de Similaridade</i>
CSCW	<i>Computer Supported Cooperative Work</i>
DFD	<i>Diagrama de Fluxo de Dados</i>
DRS	<i>Discourse Representation Structures</i>
DRT	<i>Discourse Representation Theory</i>
EAD	<i>Educação a Distância</i>
FTP	<i>File Transfer Protocol</i>
GDSS	<i>Group Decision Support Systems</i>
GUI	<i>Graphical User Interface</i>
HTML	<i>HyperText Markup Language</i>
HTTP	<i>HyperText Transfer Protocol</i>
IP	<i>Internet Protocol</i>
IRC	<i>Internet Relay Chat</i>
ISP	<i>Internet Service Provider</i>
IT	<i>Information Tecnology</i>
JDK	<i>Java Development Kit</i>
JVM	<i>Java Virtual Machine</i>
LAN	<i>Local Area Network</i>
PBL	<i>Problem Based Learning</i>
PC	<i>Personal Computer</i>
PCS	<i>Percentual do Coeficiente de Similaridade</i>
RFC	<i>Request For Comment</i>
RPC	<i>Remote Procedure Call</i>
SADGV2	<i>Sistema de Apoio à Decisão para Grupos versão 2</i>
SDL	<i>Specification and Description Language</i>
SGBD	<i>Sistema Gerenciador de Banco de Dados</i>

SRI	Sistema de Recuperação de Informações
TCP/IP	<i>Transmission Control Protocol/Internet Protocol</i>
TPF	Total de Palavras da Frase
URL	<i>Uniform Resource Locator</i>
VRML	<i>Virtual Reality Modeling Language</i>
WWW	<i>World Wide Web</i>

Lista de Figuras

FIGURA 3.1 - A tela principal de conferências.....	47
FIGURA 3.2 - A tela de criação de novas conferências	48
FIGURA 3.3 - A tela de conferências.....	48
FIGURA 3.4 - A tela de abertura de conferências: <i>Conference Organizer</i>	49
FIGURA 3.5 - A tela de abertura de conferências: <i>Associate</i>	50
FIGURA 3.6 - A tela de abertura de conferências: <i>WWW</i>	50
FIGURA 4.1 - Exemplo de DRS	60
FIGURA 4.2 - ADS da Figura 4.1	61
FIGURA 4.3 - Aplicação da primeira regra sobre a figura 4.2.....	62
FIGURA 4.4 - Aplicação da segunda regra sobre a figura 4.3	62
FIGURA 4.5 - Aplicação da terceira regra sobre a figura 4.4	62
FIGURA 4.6 - Árvore obtida com a redução de "a lição"	63
FIGURA 4.7 - Redução final da Árvore da figura 4.6.....	63
FIGURA 4.8 - Diagrama de Funcionamento do Sistema	64
FIGURA 4.9 - Os Módulos de Comunicação e <i>Interface</i> do SADGV2.....	70
FIGURA 4.10 - Os Submódulos de Comunicação do SADGV2	71
FIGURA 4.11 - Módulos Principais do SADGV2	71
FIGURA 4.12 - O primeiro passo do algoritmo de agrupamento.....	72
FIGURA 4.13 - O segundo passo do algoritmo de agrupamento	73
FIGURA 4.14 - O terceiro e o quarto passos do algoritmo de agrupamento.....	73
FIGURA 4.15 - O quinto passo do algoritmo de agrupamento	74
FIGURA 4.16 - Resultado final do algoritmo de agrupamento	74
FIGURA 4.17 - Objetivo do agrupamento de informações textuais	77
FIGURA 4.18 - Etapas do processo de agrupamento	79
FIGURA 4.19 - A <i>interface</i> inicial da ferramenta <i>Eureka</i>	83
FIGURA 4.20 - A tela de seleção de documentos para agrupamento	84
FIGURA 4.21 - A tela de seleção de <i>Stopwords</i>	84
FIGURA 4.22 - A tela de seleção e definição de novas classes de <i>Stopwords</i>	85
FIGURA 4.23 - A tela do processo de Progresso da Identificação de grupos.....	85
FIGURA 4.24 - A tela de grupos identificados	85
FIGURA 4.25 - A tela de visualização de grupos	86
FIGURA 4.26 - O Modelo Proposto.....	88
FIGURA 5.1 - O Sistema ESCOP - <i>Environment</i>	90
FIGURA 5.2 - O módulo Filtro I.....	92
FIGURA 5.3 - O módulo Agrupador.....	92
FIGURA 5.4 - O módulo de Frequência de Idéias	93
FIGURA 5.5 - O módulo Indexador	93
FIGURA 5.6 - O módulo Votador	94
FIGURA 5.7 - O módulo Ordenador	94
FIGURA 5.8 - O módulo Filtro II.....	95
FIGURA 5.9 - O Sistema ESCOP - Tela Principal	96
FIGURA 5.10 - A chamada de "Informações" e de "Executar Módulo"	96
FIGURA 5.11 - O módulo "Filtro I"	97
FIGURA 5.12 - O módulo "Agrupador"	98
FIGURA 5.13 - O módulo "Frequência de Idéias"	100
FIGURA 5.14 - O módulo "Votador"	101
FIGURA 5.15 - O módulo "Ordenador"	103

FIGURA 5.16 - O módulo "Filtro II" 103

Lista de Tabelas

TABELA 2.1 - As duas dimensões primárias das tecnologias de <i>groupware</i>	18
TABELA 3.1 - Os meios de transmissão de imagem segundo a taxa de apresentação	28
TABELA 3.2 - Os meios de avaliação e seu peso sobre a nota final do aluno.....	31
TABELA 3.3 - Análise comparativa dos sistemas inspecionados.....	54
TABELA 4.1 - Comparação entre as técnicas <i>Brainstorming</i> , Grupo Nominal e <i>DELPHI</i>	68
TABELA 4.2 - Sistemas de Apoio à Decisão em Grupo em pequenos/grandes grupos.....	69

Resumo

Na atualidade, a incessante busca pela melhor forma de acesso a informações mistura-se com as dificuldades inerentes à falta de tempo, comum em nosso cotidiano. Tal realidade cria dificuldades para que comunidades isoladas, geográfica ou economicamente, possam dispor das atualidades tecnológicas.

Surgindo como um meio de reduzir tais contrastes, a Internet apresenta-se como uma das maneiras de integrar a Educação a Distância e o trabalho cooperativo de forma ágil e sem demandar excessos de recursos, além de prover grandes facilidades para os mais variados tipos de atividades em grupo.

Visando tirar proveito destes recursos, busca-se com este estudo integrar paradigmas de cooperação para o trabalho e o ensino com esta rede. Procura-se por estratégias de suporte à solução cooperativa de problemas, através do estudo inicial de conceitos básicos para o entendimento de termos e definições metodológicas de trabalho cooperativo à distância.

A seguir, procura-se conciliar o uso destes paradigmas com a Educação a Distância e com Sistemas de Informação. Avalia-se também ferramentas existentes de *groupware*, com o objetivo de obter diferenças e comparações que possam justificar o desenvolvimento de um modelo híbrido, capaz de ser utilizado em um processo decisório.

Como principal contribuição, este trabalho inclui o projeto e a implementação de um sistema capaz de automatizar a sumarização e a votação de idéias com foco ou temas comuns, através de uma *interface WWW (World Wide Web)* de forma a facilitar a tomada de decisões em grupo.

Palavras-chave: CSCW, PBL, Trabalho Cooperativo, Aplicações para Internet.

TITLE: "ESCOP: Strategy of Support to the Cooperative Solution of Problems"

Abstract

Presently a continuous rush towards the best form to access information coexists with difficulties due to the lack of time so common nowadays. Isolated, both geographically or economically, communities are handicapped in this race.

Stepping up as a form of reducing these contrasts, Internet presents itself as an agile and non excessive resources demanding way of integration between Distance Learning and cooperative work. Besides that, Internet provides great facilities to the most diverse kinds of group activities.

Seeking to take profit from these resources it is aimed with this study to integrate the paradigms of cooperation for work and education with the net. It is looked for strategies that can support the cooperative solving of problems, through the initial study of basic concepts for the understanding of terms and methodological definitions of cooperative work at distance.

Next, it is tried to reconcile the use of these paradigms with Distance Learning and with Information Systems. The existing groupware tools are also evaluated with the objective of revealing differences and comparisons that can justify the development of a hybrid model, capable of being used in a decision-taking process.

As main contribution, this work includes the project and the development of a system capable to automate the summarization and the voting of ideas with common focus or themes, through an WWW (*World Wide Web*) interface in order to facilitate the taking of decisions in a group.

Key-words: CSCW, PBL, Cooperative Work, Applications for Internet.

1 Introdução

Percebe-se que o trabalho em grupo detém valiosas atenções relativas ao desenvolvimento de aplicações complexas, que visem facilitar o trabalho cooperativo suportado por computador. Tais aplicações revelam-se fundamentais também para o entendimento e a aplicação de metodologias relativas ao ensino descentralizado.

As empresas, buscam nas estratégias de colaboração, um subsídio para o aumento da qualidade das informações que disponibilizam. É através destas ferramentas que torna-se possível a criação de bases de dados que permitem a observação de detalhes importantes à competitividade de uma organização. Também, procura-se por uma redução nos custos associados à produção e ao gerenciamento, visto que, reuniões virtuais, podem significar uma economia substancial de tempo e recursos financeiros.

Já a Educação a Distância apresenta-se como uma inovação educativa, apesar de não ser uma prática de uso recente. Tem sua origem na educação por correspondência, onde, até hoje, a maioria dos cursos à distância utilizam em larga escala os materiais impressos e a distribuição pelos correios, chegando também aos meios de comunicação de massa, como a televisão e o rádio. A utilização da Internet procura acrescentar mais uma possibilidade, considerando-se esta mídia já popularizada e acessível, para a democratização do ensino.

Com este trabalho busca-se avaliar estudos e possíveis ferramentas, que permitam realizar automaticamente o agrupamento de idéias com foco ou temas comuns para a organização de tópicos discutidos no decorrer de um processo decisório em grupo, bem como algumas metodologias de apoio ao processo de tomada de decisão em grupo, situações estas comuns na Educação a Distância e no trabalho cooperativo.

Tais características, procuradas nestes sistemas, os tornam flexíveis para a coordenação de tarefas conjuntas, principalmente aquelas responsáveis pela integração de módulos, com *interfaces* claras e intuitivas, que possibilitem chegar a aplicações de utilização genérica, como as utilizações necessárias a Educação a Distância, cujas razões de utilização são vistas em Tarouco [TAR 98a]:

- Falta de tempo para assistir às aulas presenciais;
- Distância: dificuldade de deslocamento;
- Finanças: custo;
- Oportunidade de fazer cursos não oferecidos no local ou período de tempo apropriados;
- Possibilidade de entrar em contato com outros estudantes, de diferentes classes sociais, culturais, econômicas e experimentais.

Como conseqüência, os participantes ganham não só conhecimento, mas também novas habilidades sociais, incluindo a habilidade de comunicar-se e colaborar com colegas amplamente dispersos, a quem eles podem nunca ter visto.

A ênfase, ao relacionamento dos temas propostos com Internet, deriva-se do fato de que considera-se esta uma das formas mais democráticas de alcançar a disseminação de informações, através de uma mídia que certamente alcança tal objetivo.

1.1 Motivação

Tomando-se por referências iniciais os estudos de Souza [SOU 94] e percebendo-se o espaço existente para a continuidade deste trabalho, desenvolve-se um novo modelo, cujo objetivo principal é o projeto e a implementação de um sistema capaz de automatizar a votação e a sumarização de idéias com foco ou temas comuns.

São analisadas e avaliadas, para fins de utilização, ferramentas de gerência de bases de dados relacionais e também de tratamento de linguagem natural, além de servidores de serviços HTTP (*HyperText Transfer Protocol*) e FTP (*File Transfer Protocol*). Testes práticos ou consultas de referência com as ferramentas WebSaber, TANGO *Interactive*, Interloq, e SAACI também fazem parte do escopo desta proposta.

Implementa-se tal proposta através da construção de uma *interface* WWW para o usuário final, um módulo de ligação entre esta e um sistema voltado ao trabalho de tratamento de informações de um mediador das discussões grupais, para ambiente Windows, finalizando-se com a utilização do chamado módulo de agrupamento, que é o responsável pela manipulação das idéias geradas pelo grupo.

Com relação ao sistema implementado por Souza [SOU 94], propõe-se a sua continuidade aliada às seguintes inovações:

- Ofertar flexibilidade para a utilização com outras ferramentas, como *chat* e *email*;
- Exigir pouco tempo de instrução para uso;
- Utilizar a técnica DELPHI para o agrupamento de idéias;
- Implementar um gerador de relatórios para as reuniões, com informações gerais e parciais, segundo critérios definidos pelo coordenador de uma sessão.

1.2 O Trabalho Desenvolvido

O foco principal, deste estudo, é a definição de um modelo de sistema que suporte a tomada de decisões em grupo na Internet, aliado à implementação de um protótipo deste modelo.

Para isto, utiliza-se de um módulo agrupador de documentos textuais, procurando-se ainda agregar em uma proposta híbrida o potencial de múltiplas ferramentas síncronas e assíncronas de apoio ao trabalho cooperativo.

Visando-se uma abordagem funcional e gradativa dos temas e conceitos apresentados, estruturou-se esta dissertação da forma que se segue.

O capítulo 2 dedica-se a conceitos introdutórios do trabalho cooperativo intermediado por computador, seguido da descrição de funcionamento de nove ferramentas e serviços que suportam o trabalho cooperativo de forma síncrona ou assíncrona. A seguir, analisa-se comparativamente estas propostas, relacionando-se e explicitando-se suas características de sincronismo.

O capítulo 3 descreve o suporte à cooperação em EAD (Educação a Distância), com foco na utilização da Internet. Seguidos das motivações que levaram a estudar-se tal proposta, são comentadas as características desejáveis nas ferramentas de suporte à cooperação. Um tópico que considera-se muito importante para o embasamento da proposta apresentada é a caracterização do método de aprendizagem PBL (*Problem*

Based Learning), que encontra-se exemplificado com um caso prático de uma disciplina do PPGC da UFRGS, ministrada pela orientadora deste estudo. Apresenta-se, ainda neste capítulo, a descrição de quatro ferramentas para a solução cooperativa de problemas - WebSaber, TANGO *Interactive*, Interloq e SAACI - finalizando-se com a análise comparativa destas propostas, também com o objetivo de justificar certos pontos levantados no capítulo anterior, relativos à integração de ferramentas e serviços isolados.

O capítulo 4 introduz a proposta de um sistema para apoio ao trabalho cooperativo em EAD. São apresentadas as conclusões referentes a três trabalhos prévios considerados - Língua Natural: Classificação de Orações, trabalho de diplomação de Arthur V. Magalhães [MAG 94]; Um protótipo da ferramenta "*Issue Analyser*": projeto, implementação e validação, dissertação de mestrado de Dávison W. de Souza [SOU 94]; e Um Estudo sobre Agrupamento de Documentos Textuais em Processamento de Informações, dissertação de mestrado de Leandro K. Wives [WIV 99] - que servem de base aos levantamentos iniciais para a proposição de um novo modelo de tratamento de informações para trabalhos cooperativos em grupo. A partir destas propostas, busca-se integrar, ponderar a respeito, utilizar ou descartar certas funcionalidades destes estudos, visando fornecer justificativas teóricas e práticas para a descrição de um novo modelo, que é mostrado ao final deste capítulo.

O capítulo 5 descreve a implementação proposta, detalhando seu ambiente de implementação, a descrição funcional de seus módulos e *interface*, culminando com as considerações gerais acerca da implementação.

O capítulo 6, finalmente, apresenta os resultados e conclusões obtidos no decorrer deste estudo, além de perspectivas de trabalhos futuros.

2 Suporte ao trabalho cooperativo

Para que se possa usufruir dos benefícios oferecidos pelos métodos de suporte ao trabalho cooperativo, enfocando-se aqueles disponibilizados pela Internet, deve-se considerar o entendimento básico de alguns de seus conceitos fundamentais.

Segundo constatações de Ribeiro [RIB 98], têm-se que atualmente cresce a pressão de utilização da Web como plataforma padrão para o desenvolvimento e utilização de ferramentas para o suporte ao trabalho cooperativo. Em seu trabalho, sugerem-se os seguintes pressupostos:

- Presença de um grupo de usuários inicial bastante numeroso;
- Integração com as informações existentes disponíveis na Web;
- Uso de padrões consolidados ("de fato") abertos para todos usuários;
- Novos padrões são incorporados de forma aberta e simples. Por exemplo, o próprio HTML (*HyperText Markup Language*) foi criado com uma filosofia aberta e o mais simples possível;
- Os *softwares* para a Internet são, em grande parte, de domínio público e extensíveis a qualquer plataforma;
- O ambiente é, em geral, independente da estrutura organizacional, ultrapassando tais barreiras.

Inicia-se o aprofundamento do estudo relacionado ao tema pelo entendimento conceitual de CSCW (*Computer Supported Cooperative Work*, ou Trabalho Cooperativo Suportado/Intermediado por Computador) e *groupware*, seguindo-se de alguns dos principais serviços disponíveis na Internet para apoio ao *groupware* e a EAD, aliados a maneira como seus conceitos inerentes são projetados com vistas a oferecer *interfaces* que possibilitem o manuseio e a troca de informações dispostas em diferentes localidades geográficas. Desta forma, é possível dispor tais dados a partir de várias plataformas de *hardware* e *software*, sem que isto afete a integridade e a disponibilidade dos mesmos.

2.1 Características do trabalho cooperativo intermediado por computador

O trabalho cooperativo apresenta-se como uma forma de solucionar problemas, através da divisão destes em partes que possam ser resolvidas por grupos, cujo objetivo é diminuir o tempo e os custos envolvidos através deste processo colaborativo.

A utilização de ferramentas computacionais por estas equipes de trabalho é fundamental. O computador fomenta uma forma prática e indispensável de aumentar o desempenho dos indivíduos envolvidos nestas tarefas.

Para obter-se um bom aproveitamento do conceito da divisão de problemas e da utilização desta ferramenta, supõe-se, como necessária, uma forma de coordenar o desenvolvimento de tais atividades. É desta forma que busca-se apoio na proposta CSCW.

Pode-se definir CSCW como um conjunto de conceitos sobre o suporte a múltiplos indivíduos, trabalhando juntos, com o auxílio de sistemas computacionais. O

significado das palavras individuais no termo deve ser especialmente destacado. Não é um conceito facilmente definido, em parte devido ao fato de que os limites do campo de atuação são difíceis de circunscrever e de uma definição principal não existir - visto que esta definição pode chegar a um campo que cobre qualquer coisa que suporta a inserção de computadores para atividades nas quais mais de uma pessoa está envolvida [BAN 92].

Pode-se, ainda, optar pelas definições seguidas pela linha que aponta CSCW como os sistemas computacionais que dão suporte para possibilitar pessoas a interagir cooperativamente. O termo ainda pode ser utilizado para designar todas as pesquisas na área de trabalho cooperativo com suporte por computador, enquanto que o termo *groupware* pode ser utilizado para designar, em termos de *hardware* e *software*, os programas que oferecem apoio aos sistemas de teleconferências, sistemas de suporte à decisão, correio eletrônico, editores de texto colaborativo etc. [DIE 96].

— *Groupware* é normalmente considerado sinônimo de CSCW. Isto ocorre porque os temas relacionados a CSCW também pertencem ao domínio de *groupware*. Cita-se CSCW como a área que estuda as atividades colaborativas coordenadas e assistidas por computador, como a comunicação e resolução de problemas feitas por um grupo de indivíduos, e *groupware* como o *software* multi-usuário que apoia CSCW [RIB 98].

As várias formas existentes de trabalho cooperativo, segundo Bannon [BAN 92], e as distinções entre tais termos, como trabalho cooperativo, trabalho colaborativo e grupo de trabalho, não estão bem estabelecidas na comunidade CSCW. Porém, um grupo é definido como um conjunto relativamente fechado e fixo de pessoas compartilhando o mesmo objetivo e engajadas em uma direta e incessante comunicação. Os processos cooperativos de tomada de decisão em um grupo são muito diferentes dos que envolvem a interação de múltiplos objetivos de escopo e natureza variados, assim como de diferentes heurísticas, estruturas conceituais etc.

Para fins de simplificação, pode-se trabalhar com a idéia de que CSCW é o estudo de como as pessoas trabalham juntas, utilizando tecnologia computacional. Tais aplicações incluem *email*, sistemas de notificação, videoconferência, sistemas de *chat*, jogos interativos e aplicações compartilhadas em tempo real, como escrita compartilhada.

Considera-se importante buscar uma definição para *groupware*, o qual está inerentemente associado a CSCW. Basicamente, trata-se de uma tecnologia implementada para facilitar o trabalho de grupos. Esta pode ser utilizada para comunicar, cooperar, coordenar, resolver problemas, competir ou negociar. O termo é também usado para referir uma classe específica de tecnologias confiadas a modernas redes de computadores, como *email*, *newsgroups*, *videophones* ou *chat*.

As tecnologias de *groupware* são tipicamente divididas em duas dimensões primárias: tempo e espaço.

A primeira, considera se os usuários estão trabalhando juntos ao mesmo tempo (*groupware* em tempo real ou síncrono) ou em tempos diferentes (*groupware* assíncrono). A segunda, considera se os usuários estão trabalhando juntos no mesmo local ("face-a-face") ou em lugares diferentes (distantes). Estas considerações podem ser observadas na tabela 2.1.

Destacam-se aqui as vantagens significativas do *groupware* sobre sistemas monousuários. Estas são algumas das mais comuns:

descritos acima. Pode-se considerar as ferramentas de correio eletrônico e bancos de dados distribuídos como ferramentas cooperativas.

Banco de dados e servidores de redes podem ser confundidos com *groupware*, por permitem o compartilhamento de recursos, porém não o são. O papel destes sistemas é fornecer suporte aos *groupwares*, para que estes sejam construídos, provendo as funcionalidades necessárias. Os sistemas de bancos de dados, por exemplo, não consideram as necessidades de comunicação existentes em um grupo.

É preciso também estabelecer a diferença entre sistemas multi-usuário e os voltados para o trabalho em grupo. Os primeiros já existem há bastante tempo, como por exemplo, os sistemas de gerenciamento de bancos de dados, mas possibilitam somente a interação indireta, que se dá através da visualização de objetos comuns. Não estão presentes nesses sistemas a comunicação direta entre os usuários ou o monitoramento de ações dos mesmos, características importantes no trabalho cooperativo.

Além dos termos *groupware* e CSCW, outras siglas e expressões freqüentemente são associadas a esta área, mas todas tentam expressar como os computadores e a tecnologia de redes podem facilitar a comunicação entre os membros de um grupo. Algumas são mais restritivas, na medida em que definem sistemas de suporte a uma atividade específica. Por exemplo, a sigla GDSS (*Group Decision Support Systems*) representa os sistemas que apoiam a atividade de tomada de decisão. Já os Sistemas de Suporte a Reuniões são mais abrangentes, incluindo desde salas eletrônicas até os próprios sistemas de tomada de decisão.

A sigla CMC, que significa Comunicação Mediada por Computador, é ainda mais abrangente, englobando o universo computador/comunicação. Enquanto a área CSCW preocupa-se especificamente com a aplicação do computador à comunicação dentro do contexto do trabalho, CMC significa somar computador à comunicação em qualquer contexto.

Aliado a esta proposta, estuda-se a seguir as mais triviais ferramentas de suporte ao trabalho cooperativo, vistas segundo suas utilizações mais fundamentais.

2.2 Ferramentas de Suporte ao Trabalho Cooperativo

Dentre os principais serviços e ferramentas disponibilizados na Internet, têm-se alguns que surgiram junto com a rede. Através de décadas de aperfeiçoamento, suas *interfaces* foram aprimoradas, mas seu funcionamento básico continua sendo o mesmo. No texto seguinte, vê-se detalhadamente, segundo os conceitos disponibilizados no *web site* [WEB 98a], tais características.

Alguns dos serviços citados encontram-se em fase de substituição por outros mais atualizados, mas são igualmente abordados, como forma de se buscar um entendimento histórico/temporal destes.

Salienta-se a importância de tais ferramentas, como forma de fomento ao apoio à colaboração. Uma equipe que deseje desfrutar destes benefícios deve dispor de um ou mais grupos de tais serviços. Seus usos são os mais variados, abordando utilizações que vão de um simples *chat* entre um grupo de amigos ao final de um dia de trabalho, até uma importante reunião virtual entre diretores de uma empresa. Os benefícios de tais métodos também aplicam-se a um professor, que necessite estender seus métodos de ensino a alunos geograficamente distantes.

2.2.1 Telnet

É um programa de emulação de terminal para redes TCP/IP (*Transmission Control Protocol/Internet Protocol*) como a Internet. Ao ser executado em seu computador, o programa *Telnet* faz uma conexão com um servidor através da rede. Por meio desta ligação, os comandos digitados localmente serão executados como se o fossem feitos na máquina remota, ou seja, o console do servidor. Fica-se desta forma, apto a assumir o controle de um servidor e manter comunicação com outros servidores da rede.

Para iniciar uma sessão *Telnet*, deve-se obter permissão de acesso a um servidor, através do uso de um *username* e *password* válidos - em um processo conhecido como *logon/login* - onde, após a identificação ser validada, o usuário pode realizar suas tarefas ou, em caso contrário, este não terá seu acesso permitido ao computador desejado.

Este serviço é uma das maneiras mais comuns para controlar remotamente vários servidores, como por exemplo, *web servers*.

2.2.2 FTP

O FTP é um dos serviços mais clássicos disponíveis para a transferência de arquivos pelas redes de computadores. É definido pela RFC (*Request For Comment*) 959, publicada em 1985 e provê as facilidades necessárias para transferência "para" e "de" sistemas remotos de computadores.

Normalmente, um usuário que precise transferir um arquivo necessita certos privilégios para obter o acesso desejado a sistemas de arquivos remotos. A facilidade, conhecida como *anonymous* FTP, trabalha via um tipo especial de conta de usuário para visitantes públicos, implementada no sistema remoto.

As conexões estabelecidas utilizam o protocolo TCP.

Geralmente os servidores FTP utilizam a porta 21 para a espera de requisições de conexão. A escolha do número de portas para as conexões de dados depende de comandos emitidos pela conexão de controle. Convencionalmente, o cliente envia uma mensagem de controle a qual indica o número de porta na qual o cliente está preparado para aceitar um pedido de conexão de dados de entrada.

O uso de conexões separadas, para controle e dados, oferece as vantagens de que duas conexões podem selecionar diferentes qualidades de serviço, como por exemplo o *delay* mínimo para a conexão de controle e o máximo *throughput* para a conexão de dados, o que evita problemas de provisão de escape e transparência para comandos embutidos com o fluxo de dados.

Quando uma transferência está sendo estabelecida, esta é sempre iniciada pelo cliente, porém ou o cliente ou o servidor podem escolher aquele que irá enviar os dados. Assim como o usuário pode transferir os arquivos desejados, o mecanismo de transferência de dados é também usado para a transferência de listas de diretórios do servidor para o cliente.

Para fins de colaboração, um servidor FTP pode servir como repositório centralizado de informações para um grupo de trabalho ou de estudos. Através do uso de *download*, é possível obter-se dados deste servidor na forma de arquivos, ou, ao

invés disso, através de um *upload* é possível gravar informações nesta máquina, de forma que, sempre que necessário, estas estejam disponíveis.

Como será visto ao final do capítulo, é possível implementar esta proposta para uma aula virtual.

2.2.3 EMail

Abrevia a descrição de *Electronic Mail*, que é a transmissão de mensagens através de redes de comunicação. As mensagens podem ser notas inseridas pelo teclado ou arquivos eletrônicos armazenados em disco. Muitos *mainframes*, minicomputadores e redes de computadores possuem um sistema de *email*. Alguns sistemas de correio eletrônico são confinados a um único sistema de computador ou rede, mas outros possuem *gateways* para outros sistemas de computadores, habilitando os usuários a enviar correio eletrônico para qualquer lugar do mundo. Companhias que são completamente computadorizadas fazem uso intensivo de *email*, porque é rápido, flexível e confiável.

A maioria dos sistemas de *email* incluem um editor de textos rudimentar, para a composição de mensagens, mas muitos permitem que se edite as mensagens utilizando qualquer tipo de editor. Pode-se enviar uma mensagem para um repositório pela especificação do endereço deste. Também é possível enviar a mesma mensagem para vários usuários ao mesmo tempo. Isto é chamado *broadcasting*.

As mensagens enviadas são armazenadas em caixas postais eletrônicas até que o destinatário vá buscá-las. Para obter acesso a um correio qualquer, deve-se checar a caixa postal eletrônica periodicamente, embora muitos sistemas emitam um alerta quando chega uma nova correspondência. Esta, após ser lida, pode ser armazenada, respondida ou apagada. As cópias de memorandos podem ser impressas em papel, se considerados importantes.

Todos os serviços *on-line* e de ISP (*Internet Service Provider*) oferecem *email*, e a maioria também suporta *gateways* que permitem a troca de correspondência com usuários de outros sistemas. Usualmente, em poucos segundos ou minutos esta chega a seu destino. Esta é uma maneira particularmente efetiva de comunicação com um grupo, devido à possibilidade de enviar uma mensagem ou documento através de *broadcast* para todos do grupo, imediatamente.

Embora diferentes sistemas de *email* utilizem diferentes formatos, existem alguns padrões emergindo, que tornam possível para usuários de todos os sistemas trocar mensagens. O CCITT (*Comité Consultatif International Téléphonique et Télégraphique*) desenvolveu o padrão X.400, o qual tenta prover uma maneira universal de endereçamento de mensagens. O padrão de endereçamento de fato utilizado é o usado pelo sistema Internet, porque quase todos os sistemas de *email* tem um *gateway* Internet.

Nos anos recentes, o uso de *email* têm crescido. De acordo com algumas estimativas, existem 25 milhões de usuários enviando 15 bilhões de mensagens por ano. Tais estimativas possuem um crescimento impressionante, com tais valores sendo constantemente incrementados [Web 98a].

Uma das aplicações importantes relacionadas à utilização do *email*, são as listas de discussão (*mailing list*), identificadas por um único nome, como mail-

list@sandybay.com. Quando uma mensagem de *email* é enviada para a lista, esta é automaticamente enviada para todos os endereços desta lista.

Muitos clientes de *email* suportam listas, as quais permitem o envio de mensagens para grupos definidos pelo usuário. Em adição, existem servidores de listas que as administram de forma centralizada para grupos de usuários.

Este recurso permite a grupos que possuam relações através de assuntos de interesse a possibilidade de partilhar informações que somente são interessantes a esta comunidade. Um exemplo disto, são as listas de discussões de assuntos técnicos ou aquelas disponibilizadas em uma escola para um grupo específico de alunos ou para uma disciplina em particular. Para fazer parte destes tipos de lista é necessário enviar um pedido de inscrição e, em determinados casos, a algum processo de seleção.

2.2.4 Newsgroup

Possui o mesmo sentido de um fórum, um grupo de discussões *on-line*. Na Internet, existem milhares de *newsgroups* cobrindo cada interesse concebível. Para receber e enviar mensagens para um *newsgroup*, é necessário um *news reader*, que é um programa executado no computador cliente, sendo conectado a um servidor de *news* na Internet.

Possui utilidade parecida com a das listas de discussão, sendo que a diferença entre ambos está no tipo de aplicativo utilizado.

2.2.5 IRC

Significa *Internet Relay Chat*, um sistema de conversa desenvolvido por Jarkko Oikarinen, na Finlândia, nos anos 80. Tornou-se muito popular com um maior número de pessoas tendo acesso à Internet, devido ao fato de que habilita pessoas conectadas, a partir de qualquer lugar, a participar em discussões de forma instantânea. Além disso, não possui limite quanto a um número máximo de conexões.

Para participar de uma discussão em um IRC, é necessário um cliente IRC e acesso à Internet. Um cliente IRC é um programa que funciona localmente e envia e recebe mensagens "para" e "de" um servidor IRC. O servidor IRC, por sua vez, é responsável por enviar todas as mensagens para todos os participantes de uma discussão. Podem haver muitas discussões ao mesmo tempo, sendo que cada uma delas transita por um único canal.

É uma das ferramentas de colaboração mais simples e acessível, visto que, na maior parte dos casos, a troca de informações baseia-se em pequenos textos enviados pelas partes que estão conectadas a um canal. Sua utilização atende a necessidades bastante variadas, como um departamento de uma empresa, uma aula virtual e até mesmo como meio de descontração, pois existem canais de *chat* dedicados a assuntos de vão de carros e motos até a namoros virtuais.

2.2.6 ICQ

É um programa de mensagens instantâneas fácil de utilizar desenvolvido pela companhia Mirabilis LTD. Deve ser pronunciado em letras separadas, de forma que o som fica parecido com "*I-Seek-You*", sendo similar a outros programas, como alguns da

empresa America OnLine. É utilizado como uma ferramenta de conferência por indivíduos em rede para conversar, enviar correios eletrônicos, executar transferência de arquivos, participar de jogos de computador, entre outros.

Uma vez instalado no computador cliente, pode-se criar listas de amigos, familiares, associados em negócios etc., os quais também devem possuir o ICQ em seus computadores. É através destas listas que as pessoas são achadas e notificadas quando entram ou saem da rede. Também é permitido o envio de mensagens, conversas em tempo real etc.

O ICQ pode ser considerado como sendo uma forma avançada de IRC, visto que, disponibiliza mais opções que o segundo. Tais incrementos tendem a oferecer melhor suporte à colaboração através da Internet.

2.2.7 CU-SeeMe

Corresponde a um programa para videoconferência, que utiliza a Internet para transmitir sinais de áudio e vídeo. Uma versão gratuita deste *software* é disponibilizada pela Cornell University e uma versão comercial é vendida pela empresa WhitePine, Inc. O produto pode ser utilizado para videoconferência, ponto-a-ponto ou multiponto. Exige placa de som, auto-falantes e câmera de vídeo.

Esta ferramenta agrega as funções *chat*, video e audioconferência. Tais recursos podem ser utilizados conjuntamente ou em separado. A qualidade da linha de comunicação entre as partes conectadas é que irá definir o tipo de serviço a ser utilizado.

A instalação do produto é bastante fácil e, geralmente, para fins de videoconferência, é utilizado um refletor, que é um servidor responsável por retransmitir as informações enviadas pelos clientes. Em uma ligação ponto-a-ponto não é necessário a existência deste, bastando para isso, que a segunda parte envolvida possua um endereço IP válido.

Várias universidades e empresas utilizam este produto, visto sua flexibilidade e praticidade de utilização. Existem, atualmente, indicativos de que o processo de Educação a Distância, com o auxílio da videoconferência, está seguindo um caminho irreversível e, se já não o é, brevemente será uma tendência. Exemplos disto são projetos existentes nos países da América do Norte e Europa que, entre outras, visam aumentar significativamente as velocidades das conexões de rede a fim de dar suporte ao tráfego de dados com características de tempo real, provenientes de videoconferências com motivos educacionais. Percebe-se cada vez mais que se está utilizando da Internet para auxiliar e aprimorar o ensino.

2.2.8 WWW

É uma definição ampla, que pode ser resumida como sendo um sistema de servidores Internet, que suportam especialmente documentos formatados. Estes documentos são formatados em uma linguagem chamada HTML, que suporta referências para outros documentos, que podem ser arquivos gráficos, de áudio e vídeo. Isto permite a navegação de um documento a outro através dos chamados *links* selecionáveis. Não são todos os servidores Internet que fazem parte da *World Wide Web*.

As aplicações chamadas *web browsers* facilitam o acesso através da rede WWW. Dois dos mais populares navegadores são o *Netscape Navigator* e o *Microsoft's Internet Explorer* [Web 98a].

Geralmente, as aplicações para *groupware* e EAD utilizam a WWW como meio para a troca de informações, visto esta oferecer baixos custos e grande variedade de meios multimídia.

2.2.9 BBS

Significa *Bulletin Board System*, onde um computador é conectado a linhas telefônicas e a um *modem*, atuando como servidor para intermediar comunicação por mensagens ou transmissão de arquivos entre os computadores de seus usuários. Atualmente este tipo de aplicação é utilizado em *links* de baixa velocidade, por exigir poucos recursos para seu funcionamento.

Devido a esta última característica, pode-se considerar a utilização de uma BBS como uma opção de baixo custo para a implementação de uma lista de discussões e repositório de informações assíncronas.

2.3 Análise Comparativa das Ferramentas Existentes

Percebe-se, neste capítulo, pelas descrições das ferramentas e conceitos apresentados, as diversas variedades de opções para o suporte ao trabalho cooperativo intermediado por computador.

Analisando-se as ferramentas referenciadas anteriormente, segundo suas características de sincronismo, pode-se obter a seguinte classificação:

- Síncronas: *Telnet*, *FTP*, *IRC*, *ICQ*, *CU-SeeMe*
- Assíncronas: *EMail*, *Newsgroup*, *BBS*

Considera-se a WWW como um meio que mescla estas duas características, pela sua grande variedade de utilizações, como já mencionado.

Desta forma, pode-se utilizar tais ferramentas de tal modo que, ao final, obtenha-se, de maneira simplificada, estratégias de suporte à solução cooperativa de problemas para o apoio às atividades de trabalho cooperativo mais triviais, como discussões *on-line* (utilizando *IRC* ou *ICQ*), videoconferência (utilizando o *CU-SeeMe*) e discussões através de listas (*EMail*, *Newsgroup* e *BBS*). Ritzel [RIT 98] mostra em seu trabalho que tais tecnologias de sistemas multimídia apoiam o ensino a distância, possibilitando a elaboração de cursos à distância com melhores resultados.

Como forma de exemplificar tal utilização, supõem-se, para aplicação em EAD, de uma sala de aula virtual, onde o aluno irá receber, através do *CU-SeeMe*, áudio e vídeo do professor, que desta forma, apresenta o conteúdo da aula.

Esta estratégia somente é disponibilizada ao professor, de forma a baratear o processo e também porque a comunicação bidirecional de áudio e vídeo determina um tráfego muito intenso de comunicações na rede, o que tende a crescer muito de acordo com o número de alunos presentes em uma sessão.

Ao surgir uma dúvida a respeito de determinado tópico, o aluno usa a ferramenta de *chat* para apresentá-la, de forma que todos os participantes desta aula possam visualizá-la.

Ao perceber uma questão em seu terminal, o professor pode escolher a mídia para respondê-la: ou pelo próprio canal de *chat* ou então pelo *CU-SeeMe*, onde todos os presentes acompanham tal explicação.

Como forma de documentar e disponibilizar o material utilizado para as aulas, o professor pode utilizar-se de servidores WWW e FTP. Desta forma, ao perder uma aula *on-line*, basta ao educando acessar tal material para inteirar-se de suas atividades.

O servidor WWW pode armazenar as sessões de vídeo com aulas ou debates gravados, bem como o arquivo com as mensagens enviadas via *chat*, para acesso posterior, da mesma maneira que também podem ser armazenadas as mensagens de uma lista.

O servidor FTP pode ser usado pelos alunos para enviar e/ou receber arquivos com suas tarefas.

Para o estudo extra-classe das tarefas dadas aos alunos, pode ser disponibilizada uma lista de discussões assíncrona, sempre supervisionada pelo professor responsável, que tem o papel de mediador das discussões. Assim, têm-se uma boa integração entre os participantes, já que estes tem condições de trabalhar em grupo sem a necessidade de um deslocamento físico para tais atividades, o que também resolve o problema de compatibilidade de horário.

Estas ferramentas podem ser combinadas em sistemas de apoio ao trabalho de grupo com fins educacionais, resultando em sistemas integrados, tal como será descrito no capítulo seguinte.

Esta característica de integração é obtida quanto tais propostas agregam todas as características vistas até aqui, adicionadas a outros conceitos, como funcionalidades para a organização de informações distribuídas e o acesso por diferentes tipos usuários para diferentes tipos de aplicações, por exemplo, o que forma inicialmente a base teórica para a justificativa deste estudo.

3 Suporte à Cooperação em EAD

Conforme observou-se no capítulo anterior, a utilização da Internet como ferramenta de apoio ao ensino pode ser considerada fundamental, considerando-se as vantagens oferecidas por este meio, proporcionadas pela ampla gama de serviços que podem ser utilizados.

Ao construir-se a ponte que une os conceitos relativos a CSCW e PBL, - conceito que será discutido ao longo deste capítulo -, percebe-se que a educação passa a contar com um variado leque de vantagens oferecidas por tais possibilidades.

Os métodos tradicionais de ensino, caracterizados pelo professor presente em uma sala de aula repleta de alunos, passa a ter uma conotação figurativa com tais ferramentas disponíveis.

Mesmo sem considerar-se necessária uma comparação entre os métodos de ensino tradicionais e os de EAD, percebe-se, pela conclusão do capítulo dois, que o embasamento técnico, já apresentado, permite a aplicação destes conceitos em sala de aula.

O foco principal deste capítulo é fazer o enlace entre o EAD e as técnicas vistas até aqui, atendo-se principalmente às vantagens disponibilizados pela Internet, pois considera-se que tal proposição pode ser embasada pelas considerações de Tarouco [TAR 98b], que diz ser a EAD caracterizada:

- Pela separação do professor e aluno no espaço e/ou tempo;
- Pelo controle do aprendizado realizado mais intensamente pelo aluno do que pelo instrutor distante;
- Pela comunicação entre alunos e professores mediada por documentos impressos ou alguma forma de tecnologia.

A primeira forma de EAD foram os cursos, por correspondência, na Europa. Este meio de ensino a distância foi muito utilizado até a metade deste século, quando o rádio e televisão instrucional tornaram-se populares.

No entanto, com o surgimento de tecnologias interativas sofisticadas, os educadores passaram a utilizar ferramentas como: *email*, BBS's, Internet, audioconferência baseada em telefone e videoconferências com 1 ou 2 caminhos de vídeo e 2 caminhos de áudio. Uma ferramenta da Internet que tem sido muito utilizada é o WWW, o qual possibilita a elaboração de cursos à distância com avançados recursos de multimídia.

Com o desenvolvimento destas tecnologias interativas que possibilitam contato em tempo real entre locais espalhados geograficamente, começam a surgir as chamadas salas de aula virtuais. Dentre as principais características destas novas salas de aula, pode-se citar a possibilidade de contato com um maior número de colegas, com os quais os estudantes podem colaborar numa quantidade superior ao que podem encontrar em sua própria região. Além de permitir o acesso a um quadro bastante extenso de professores e mentores, numa dimensão impossível para uma única instituição educacional local.

Em poucos anos, computadores e telecomunicações de alto desempenho serão utilizados como material didático. Do mesmo modo, comunidades virtuais e ambientes artificiais compartilhados farão parte da rotina do "dia-a-dia", como o telefone, televisão, rádio e jornais são hoje. Por esta razão, as experiências de aprendizagem à distância serão tidas como vitais para todos os estudantes, ainda quando o mesmo conteúdo puder ser ensinado "face-a-face". Por esta razão, todo o ato de ensinar terá alguns atributos da EAD [CRU 98].

A EAD é um recurso de incalculável importância como modo apropriado para atender a grandes contingentes de alunos, de forma mais efetiva que outras modalidades, e sem riscos de reduzir a qualidade dos serviços oferecidos em decorrência da ampliação da clientela atendida.

Souza [SOU 97] mostra que a LDB, através da lei número 9394/96, dedica um de seus artigos ao ensino a distância, o que, segundo ele, mostra-a bastante conectada com a modernidade educacional. Anterior a esta lei, Luckesi [LUC 89] aponta que o ensino a distância é uma opção de proceder a ação educativa, através da instrução.

A escolha da modalidade da EAD, como meio de dotar as instituições educacionais de condições para atender às novas demandas por ensino e treinamento ágil, célere e qualitativamente superior, tem por base a compreensão de que, a partir dos anos 60, a EAD começou a distinguir-se como uma modalidade não-convencional de educação, capaz de atender, com grande perspectiva de eficiência, eficácia e qualidade, aos anseios de universalização do ensino, e, também, como meio apropriado à permanente atualização dos conhecimentos gerados, de forma cada mais intensa, pela ciência e cultura humana [NUN 98].

Cattani [CAT 98] cita que, atualmente, várias universidades brasileiras têm enveredado por esta forma de ensino, aplicando-o inclusive em cursos de pós-graduação.

As estratégias consideradas por Tarouco [TAR 98a] como eficientes na Educação a Distância são:

- O desenvolvimento de métodos apropriados de *feedback* e reforços;
- A adaptação aos diferentes estilos de aprendizado dos alunos;
- O uso de estudo de casos e exemplos (*Problem Based Learning*);
- O complemento dos cursos com informações impressas.

Sob estes aspectos vê-se, em Cruz [CRU 98], que pode ser fundamental a qualquer instituição interessada em começar a produzir EAD, o esforço de criação de um modelo que sirva de base e dê coerência para todas as ações direcionadas neste sentido. Entretanto, para se chegar ao produto educativo é preciso que as etapas de planejamento, produção e aplicação da Educação a Distância sejam detalhadamente trabalhadas pela equipe executiva do projeto de forma a atender da melhor maneira possível as necessidades do público alvo. Isso quer dizer que o desenvolvimento de um produto (para Educação a Distância) deve ser adequado ao atendimento das necessidades de um usuário (cliente e usuário final). Dessa maneira, a EAD pode ser considerada um sistema sob o qual interagem os elementos necessários para que se desenvolva integralmente o processo de ensino/aprendizagem.

Acrescendo-se à definição de um modelo os modos de aprendizado, Tarouco [TAR 98b] observa que outra variável importante na eficácia do aprendizado é a preferência do aluno por um modo particular de aprendizado:

- Cooperativo;
- Competitivo;
- Individualizado.

Muitos projetos de EAD incorporam aprendizado cooperativo, projetos colaborativos e interatividade entre grupos de alunos e entre locais.

Para a implementação de tais paradigmas, segundo Tarouco [TAR 98b], são utilizadas as seguintes tecnologias:

- **Audioconferência:** estabelece uma conexão por voz apenas entre duas ou mais pessoas usando telefone convencional ou fonia na Internet. É necessário usar uma ponte (*audio bridge*) para a comunicação entre muitos pontos;
- **Conferências Audiográficas:** usa-se linhas telefônicas para transmitir áudio e gráficos (desenhos e diagramas). Pode usar fax ou sistemas de quadro eletrônico;
- **Computer Conference:** comunicação de grupo usando o computador para compartilhar arquivos. Vale-se de comunicação bidirecional;
- **Teleconferência Comutada:** utiliza telefones para comunicação entre pessoas em vários locais;
- **Quadro Eletrônico (*Electronic Blackboard*):** vem a ser um quadro branco que transmite o que é escrito ou desenhado sobre ele.
- **Full Motion Video:** transmissão de imagens de vídeo animadas de modo a tornar possível visualizar uma ação completa, segundo uma taxa de *frames* por segundo apropriada;
- **Fully Interactive Audio/Video:** vídeo interativo bidirecional com voz e imagem.

TABELA 3.1 - Os meios de transmissão de imagem segundo a taxa de apresentação

Meio de Apresentação	Frames por segundo
Cinema	24
TV	30
Videoconferência	15

Observa-se que a taxa de *frames* por segundo, apresentada na tabela 3.1, pode variar conforme o protocolo utilizado ou de acordo com diferentes valores de configuração do *software* utilizado.

O estudo prossegue através de um breve roteiro para a implementação e validação destas propostas, através da utilização da Internet para EAD.

3.1 A Educação a Distância via Internet

Como forma de integração de todas as tecnologias vistas até este ponto, têm-se a utilização da Internet que visa promover uma maior interação aluno-professor e aluno-aluno, como um espaço de troca e produção coletiva de conhecimento e informação, fora dos horários de aula, pela videoconferência. Essa interação acontece através de *sites* WWW, disponível para os alunos e demais envolvidos no curso. Nos *sites* estão as ferramentas necessárias para o aluno se comunicar com seus professores ou colegas, comentar as aulas, discutir temas relacionados às disciplinas em andamento, enviar sua produção ao professor e acessar ementas de disciplinas, bibliografias de referência, artigos e outras informações importantes para um bom desempenho no curso [CRU 98].

Tais *sites* devem ser mantidos e atualizados, já que neles estão disponíveis todas as informações sobre o curso, sendo que, normalmente, as informações gerais são de acesso público e as informações específicas de cada disciplina podem ser acessadas somente através de senhas individuais pelos alunos regularmente matriculados. Estas informações são disponibilizadas com antecedência, permitindo aos alunos uma preparação para as aulas.

Ao mesmo tempo, através do uso de ferramentas específicas que são disponibilizadas aos alunos, estes podem discutir em tempo real, utilizando um tipo de sala de reuniões, assuntos relacionados às disciplinas em andamento; inteirar-se das últimas novidades de cada disciplina, como mudanças de datas, horários etc., através de um serviço de notícias (*news*); obter os textos, transparências e artigos das aulas seguintes (ou anteriores), disponíveis em uma biblioteca virtual. Além destes serviços, também pode-se utilizar outros, como uma sala de produção, que pode ser vista como um "escaninho virtual", onde o aluno deposita seus trabalhos (exercícios, artigos, roteiros de seminários e, eventualmente, provas) para a avaliação posterior do professor. Esta sala pode ser vista como um serviço com dupla função, uma vez que o aluno a utiliza para disponibilizar materiais para o professor, e este a utiliza para verificar a produção de seus alunos fora do horário de aula via videoconferência, permitindo que se avalie o nível de responsabilidade e envolvimento, assim como o próprio nível de aprendizagem dos alunos em determinados assuntos. Outras ferramentas que podem ser disponibilizadas são bancos de casos, agendas, salas de discussão e um *mailbox*, as quais oferecem informações detalhadas sobre a programação das disciplinas, espaço para fóruns de discussão entre os alunos e professores e acesso a um serviço de *email*.

Leite [LEI 98] ilustra o seguinte princípio:

Quando se fala em EAD não mais se pode limitar seu escopo ao uso do material impresso ou da televisão. Os sistemas de EAD comportam e até solicitam a utilização de mais de uma tecnologia de maneira integrada.

Considera-se que hoje a tecnologia permite que se tome contato com a realidade indiretamente. A relação do educando com a realidade não se limita mais à sua experiência pessoal e ao que a escola e a família lhe proporcionam, administrando a informação e os modelos de interpretação da realidade. As fontes de informação estão muito mais diversificadas e a escola tem o dever de estimular novas formas de experimentação e criação dos educandos; para que essa função seja cumprida, os professores devem estar capacitados para tal, principalmente quando esse ensino for feito a distância via rede de computadores, porque suas características são diferentes daquelas apresentadas pelo ensino presencial.

As tecnologias da comunicação já permitem que profissionais se atualizem mediante cursos de EAD via rede de computadores, recebendo materiais escritos e audiovisuais pelo WWW. Deve-se lembrar que, como visto anteriormente neste trabalho, o desenvolvimento tecnológico já possibilita inclusive a utilização de videoconferências na rede.

A rede de computadores possui atributos que, segundo Leite [LEI 98], a caracterizam como um meio distinto de ensino-aprendizagem. São eles: provê acesso de maneira econômica e as informações são apresentadas em formatos variados e não encontrados em nenhuma outra combinação de meios; a maior parte do conteúdo da rede em geral não está disponível em nenhum outro formato, a não ser no original dos autores; a rede permite que o trabalho do professor e dos alunos possa ser compartilhado com o mundo, de maneira diferente da que o aluno pode encontrar no ambiente tradicional de ensino; alunos abordam a rede com vontade, motivação, respeito e receio, sabendo que é uma tecnologia de ponta, utilizada por profissionais atualizados e adultos de sucesso.

Dentre as mudanças utilizadas pela informatização via rede, identifica-se a necessidade de manejo de múltiplas fontes de referência, mediante intervenção ativa do usuário, que tenderá a aplicá-las de modo cada vez mais autônomo. E, certamente, esse tipo de construção de conhecimento, não linear e não seqüencial, possibilitados pelos sistemas de hipertexto e hiperímídia, irão requerer dos atuais professores/instrutores novas aprendizagens, principalmente no que diz respeito ao planejamento, desenvolvimento e avaliação de programas de EAD que utilizem este meio.

Leite [LEI 98] sugere que é importante ressaltar algumas considerações, relacionadas à construção de novos espaços, para a aprendizagem via EAD:

- Um curso de EAD através da rede deve ser planejado, desenvolvido e avaliado por um grupo interdisciplinar. Devido a complexidade do próprio processo educativo, aliada à complexidade do domínio atualizado das informações e dos mecanismos de interação com a rede, dificilmente um único profissional desenvolverá um trabalho de EAD de qualidade, se trabalhar isoladamente;
- Para que um curso via rede seja desenvolvido é fundamental que seja feito previamente um plano instrucional detalhado do curso;
- Os professores, pessoal administrativo e de apoio envolvidos em um curso via rede precisam desejar aprender uma maneira totalmente nova de comunicar a mensagem e de garantir que a aprendizagem aconteça;
- O professor ou equipe de professores responsáveis pelo desenvolvimento de um curso via rede devem ter experiência de sala de aula e terem dado o curso presencialmente;
- O professor ou professores que desenvolveram um curso via rede devem ser responsáveis pelo seu oferecimento, pelo menos na primeira vez que o curso for oferecido via rede;
- Os alunos que se inscreverem em cursos via rede devem ter experiência prévia de navegação na Internet, ou o curso deve incluir uma unidade introdutória de modo a familiarizar o aluno com esta tecnologia;
- A tecnologia e o pessoal técnico de apoio devem estar disponíveis para que um curso via rede possa ser oferecido;

- A seleção das novas tecnologias a serem utilizadas em programas de capacitação devem ser orientadas pelo conhecimento da estratégia de ensino a ser adotada, do nível educativo do programa a ser desenvolvido, da proposta de formação e reciclagem dos professores e das estratégias de acompanhamento e avaliação do programa;
- Em cursos via Internet, sugere-se que sejam feitos exercícios ou testes curtos semanais para que os alunos se mantenham atualizados em relação ao curso;
- Sugere que a nota final de cursos via rede seja resultante de diversas atividades de avaliação realizadas durante o curso, conforme mostra a tabela 3.2;
- Os sistemas administrativos precisam estar estruturados para este tipo de curso para que ele tenha sucesso;
- Devido aos custos elevados deste tipo de curso, é indicado que sejam feitos convênios entre instituições de capacitação de professores públicas e privadas e empresas;
- As instituições devem desenvolver projetos e programas cooperativos de EAD, devendo se comunicar nos níveis local, regional, nacional e internacional;
- Qualquer que seja o curso via rede, ele só terá chance de sucesso se tiver apoio da administração da instituição.

TABELA 3.2 - Os meios de avaliação e seu peso sobre a nota final do aluno

Tipo de Avaliação	% sobre a Nota Final
Exame Final	35
Trabalhos realizados durante o curso	35
Testes e contribuições nas aulas	30

Como forma de sintetizar suas idéias, Leite [LEI 98] finaliza focalizando a atenção no professor, onde aquele que se propuser a ensinar em sistemas de EAD deve refletir sobre alguns aspectos fundamentais, que são:

- Contexto de ensino: que é alterado devido a separação física entre os participantes do processo e mediatizado pelo uso da tecnologia; o ambiente de aprendizagem assume nova configuração. O professor, para atuar efetivamente, precisa reconhecer essa mudança no ambiente e sua influência no contexto. Mais especificamente, o professor precisa trabalhar com as potencialidades do meio e adaptá-lo às limitações impostas a sua abordagem instrucional;
- Alunos: em programas de EAD eles vivenciam a aprendizagem de maneira diferente do ensino presencial, portanto, têm uma perspectiva diferente daqueles que não estão separados do *locus* de instrução. O professor precisa estar atento e sensível aos obstáculos psicológicos, sociais e técnicos a serem enfrentados pelo aluno de cursos via EAD;
- Métodos: profissionais de EAD deverem ser cuidadosos em simplesmente não reapplicarem métodos tradicionais de ensino presencial, pois precisam reconhecer que eles não podem ser simplesmente utilizados em situações de EAD. Há

necessidade de serem exploradas estratégias alternativas de ensino, contextualizadas no ambiente de EAD. Os métodos de ensino de EAD devem, em geral, buscar reduzir a distância interpessoal, promover a interação, aumentar o *feedback* e garantir a aprendizagem e a transferência da mensagem.

As instituições de ensino que optarem pela EAD e pela manutenção da sua credibilidade e respeito usando tecnologia inovadora para chegar aos alunos em lugares distantes e atendendo às suas necessidades, além de observar os aspectos acima destacados, não podem se intimidar pelos obstáculos apresentados por esta modalidade de ensino.

Desta forma, pode-se sintetizar este item destacando-se as vantagens anteriormente mostradas de EAD sobre os métodos tradicionais de ensino, aliados ao desenvolvimento de tecnologias interativas que permitem o contato em tempo real entre locais e pessoas espalhados geograficamente.

Finalmente deve-se perceber que tais condições permitem uma maior integração do educando com seu meio, haja visto o aumento da informação disponível, o que vem a modificar a percepção e a interpretação dos fatos e da própria realidade.

3.2 Motivações para o Trabalho Cooperativo em EAD

O trabalho desenvolvido por Otsuka [OTS 98], abordou um estudo sobre necessidades de ferramentas para apoio à cooperação. Utiliza como termo sinônimo de Trabalho Cooperativo a sigla CMC, que, como visto no capítulo 2, significa Comunicação Mediada por Computador. Fundindo as duas propostas, considera-se que CMC engloba o universo computador/comunicação, enquanto que a área de Trabalho Cooperativo Suportado por Computador preocupa-se especificamente com a aplicação do computador à comunicação dentro do contexto do trabalho. Considera-se ainda que CMC significa somar computador à comunicação em qualquer contexto.

Seguindo as relações propostas por Otsuka [OTS 98], observa-se que a CMC possibilita uma comunicação rápida, intensa e eficiente, com a introdução de um grande número de novos recursos, o que provê um maior enriquecimento nas comunicações. Destaca-se destas idéias as principais lições obtidas durante os experimentos e pesquisas sobre o uso de CMC no ensino:

- A CMC pode expandir os meios tradicionais de ensino e aprendizado, tanto face a face quanto à distância, além de prover novos modelos e oportunidades para aprendizagem;
- Os educadores podem projetar novos modelos de aprendizagem e também projetar ambientes de rede que suportem uma aprendizagem efetiva;
- O uso educacional da CMC oferece uma estrutura conceitual e tecnológica na aprendizagem para a vida toda (*lifelong learning*).

3.3 Características Desejáveis nas Ferramentas de Suporte à Cooperação

A EAD que faz uso da CMC como meio de comunicação apresenta novas características que distinguem-se de outras formas tradicionais, como apostilas, rádio e

televisão, por exemplo. Desta forma, pode-se apresentar como características desejáveis nas ferramentas de suporte à cooperação o seguinte:

- O incremento da interatividade entre aluno-aluno e/ou aluno-professor, já que, nos meios tradicionais esta interação pode ser muito pequena ou ineficiente. Considera-se que, em uma sala de aula com um número indeterminado de alunos, as interações entre aluno e professor são mínimas, haja visto que aproximadamente 10% do total de alunos dominam as interações e os demais permanecem relativamente passivos;
- A eliminação da barreira da timidez, aliada a maior flexibilidade oferecida aos alunos que trabalham ou possuem outras responsabilidades;
- Participação distribuída, com a diminuição de grupos dominantes e incremento do número de alunos participantes;
- Participação assíncrona, o que além da independência geográfica, propicia ao aluno um tempo maior para formular suas idéias e conseqüentemente realizar uma contribuição mais rica, através de ferramentas como *email*, *news* e *WWW*;
- Aumento da qualidade dos trabalhos produzidos, principalmente porque estes serão vistos por várias pessoas;
- Rapidez e abrangência, porque um documento disponibilizado em um servidor *WWW* pode ser instantaneamente visualizado em qualquer parte do mundo. Uma mensagem de correio eletrônico, em condições normais de funcionamento da rede, pode chegar a qualquer parte do mundo em alguns segundos. Além disso, serviços como *talk*, *chat* e áudio/videoconferência permitem a comunicação em tempo real;
- Variedade de ferramentas disponíveis, que vão desde as ferramentas em modo texto assíncronas, como é o caso do correio eletrônico, até as ferramentas multimídia síncronas, como as ferramentas de videoconferência, garantindo uma maior flexibilidade na escolha da forma mais rápida de comunicação a cada caso particular de ensino;
- Eliminação dos atrasos de disponibilização do material instrucional, já que, nos meios tradicionais a distribuição de material impresso é realizada pelo sistema postal e exige a impressão do número de cópias necessárias, seguida da postagem de cada uma delas, implicando em muitos gastos, esforços e na perda de um tempo considerável para estas tarefas. Através das ferramentas de CMC, é possível a disponibilização instantânea do material instrucional, eliminando os atrasos de publicação e distribuição.

A estas características desejáveis somam-se as necessidades de infra-estrutura e conhecimento tecnológico adequados a tais atividades. Desta forma, deve ser seguida a aplicação de uma metodologia de aprendizagem que pode ser utilizada para implementar estas propostas.

3.4 Sistemas de Apoio à Colaboração e seu uso em PBL

Nos estudos elaborados por VIT [VIT 99a], observa-se que várias são as metodologias possíveis para buscar formas de treinamento e cooperação entre indivíduos.

Busca-se embasamento para a aplicação de propostas de treinamento e ensino com uma metodologia que propicie a descentralização da solução de problemas, principalmente quando não é possível dispor de um instrutor constantemente presente. Ao mesmo tempo, procura-se acrescentar as vantagens percebidas na aplicação da solução cooperativa de problemas.

Ao contrário da instrução tradicional, a qual é freqüentemente conduzida na forma de leitura, a forma de ensinar em PBL normalmente ocorre dentro de pequenos grupos de discussão formados por estudantes orientados por um instrutor [LEW 98].

Devido à reduzida quantidade de informações diretas disponibilizadas em PBL, os estudantes assumem grande responsabilidade por seus próprios aprendizados. O papel do instrutor torna-se o de perito no assunto, guia de recursos e consultor das tarefas do grupo. Este arranjo promove processamento de informações em grupo em lugar de somente o compartilhamento de resultados. O papel do instrutor é o de encorajar a participação do estudante, prover a informação apropriada para manter os estudantes em seu rumo, evitar repetir os erros cometidos anteriormente e assumir o papel de estudante da mesma categoria.

A habilidade em resolver problemas é maior do que somente acumular conhecimentos e regras; é o desenvolvimento de estratégias cognitivas flexíveis que ajudem a analisar o inesperado em situações mal estruturadas para produzir soluções significativas. Embora muitos dos assuntos complexos atuais estejam dentro do universo de conhecimentos dos estudantes, a habilidade necessária para resolver estes problemas está freqüentemente fora da instrução. A resolução típica de problemas ensinada em escolas freqüentemente tende a ser de situações específicas com parâmetros de problemas bem definidos com resultados predeterminados com uma resposta correta. Nestas situações, são freqüentes a existência dos procedimentos requeridos para resolver o problema, que é o foco da instrução. Infelizmente, os estudantes orientados por este método não estão adequadamente preparados para encontrarem problemas em que necessitem transferir seu conhecimento para novos domínios, uma habilidade requerida para a vivência em sociedade.

Dentre as vantagens destacadas pelo método PBL, a mais visível é a ênfase dada ao significado, não aos fatos. Considera-se que a maioria dos estudantes retém e utiliza pouco daquilo que é memorizado em sala de aula. E este é o método comumente utilizado nos programas tradicionais. O PBL tenta quebrar este ciclo através do engajamento dos estudantes na estruturação de soluções para problemas reais, que sejam relevantes e contextualizados. Substituindo-se leituras por grupos de discussão e pesquisa colaborativa, os estudantes tornam-se ativamente engajados na aprendizagem.

Como os participantes procuram soluções para seus problemas de sala de aula, estes tendem a assumir o incremento de responsabilidade para o seu aprendizado. Estes estudantes utilizam recursos selecionados por si mesmos, como jornais, buscas *on-line* e outros recursos disponíveis em bibliotecas, sendo os livros textos mais freqüentemente acessados pelos primeiros, em relação ao estudantes considerados tradicionais. Eles também mudam sua visão a respeito dos instrutores, que passam de uma origem de respostas aos testes, para um possível recurso que auxilie na resolução de problemas relevantes. Juntos, estes processos e habilidades de aprendizagem ajudam os estudantes a tornarem-se mais competentes na busca de informações que os estudantes tradicionais.

Poucas pessoas na sociedade trabalham isoladamente. Devido à interação social ser um importante aspecto do trabalho, PBL incorpora times colaborativos na solução de problemas relevantes. Esta metodologia promove a integração de estudantes e

equipes de trabalho, aumentando assim as habilidades interpessoais dos estudantes, através da dinâmica de grupo, da avaliação de estratégias e das formas de apresentar e defender idéias.

As mudanças necessárias à aplicação de PBL, todavia, não são consideradas de fácil aceitação. Conforme Lewis [LEW 98], as pessoas vêem as mudanças negativamente, como um trabalho adicional e não como um benefício adicional. Estas modificações são consideradas como uma mudança de experiência de vida e requerem um período de adaptação.

Esta transição não é considerada difícil somente para os instrutores, já que é também vista como uma grande mudança para os estudantes. O PBL requer mais tempo disponível pelos alunos e espera que estes sejam responsáveis e independentes para seus aprendizados. Considera-se que o sucesso deste paradigma irá depender de comunicação efetiva e orientação.

Encontraram-se apontamentos para alguns problemas relativos à aplicação desta técnica:

- A realização acadêmica de estudantes envolvidos em PBL;
- A quantidade de tempo requerida para implementação;
- A mudança de papel dos estudantes no processo;
- A mudança de papel dos professores no processo;
- Problemas relativos ao que foi efetivamente aprendido pelo estudante;
- Tamanho do programa *versus* o aprendizado do estudante.

Também devem ser avaliados os custos envolvidos em tal proposta, visto que estes são considerados mais altos que os envolvidos nos métodos tradicionais, principalmente porque envolvem várias pequenas salas de estudo equipadas para o ensino, bem como número adequado de cópias de recursos de biblioteca disponíveis.

Para enfatizar o elo de ligação entre a proposta do método PBL e os sistemas de apoio à colaboração, descritos no capítulo 2 e novamente discutidos neste capítulo, utiliza-se do exemplo ilustrado por VIT [VIT 99a] como forma de exemplificar tais proposições, utilizando-se da metodologia aplicada na disciplina de Gerência de Redes do Programa de Pós-Graduação em Computação da Universidade Federal do Rio Grande do Sul, pela professora titular da disciplina, Liane Margarida Rockenbach Tarouco. As referências podem ser encontradas em www.penta.ufrgs.br.

Observou-se que as aulas foram apresentadas através de formulários HTML, divididos em partes iniciadas pela aula número 1 seguidas em seqüência até a aula número 15. O material disponível para estas aulas encontrava-se exclusivamente nas páginas referenciadas pela professora, numa aplicação prática dos conceitos apresentados anteriormente no item 3.1 deste estudo.

A utilização da Internet deu-se desde os conteúdos programáticos da disciplina, até as bases de referências utilizadas, as quais sempre referenciavam-se através de *hiperlinks*.

A aplicação da metodologia PBL deu-se na solução dos problemas propostos em aula. No início de cada uma destas, apresentava-se aos alunos as tarefas que deveriam ser resolvidas. Inicialmente, estes dispunham de um tempo para leitura, onde eram apresentados aos conceitos que seriam mais tarde utilizados.

Logo após este período inicial, os alunos passavam à solução de problemas práticos, como por exemplo, a análise das *interfaces* de um *router*. Para solucionar tais tarefas, os estudantes dispunham da ajuda do orientador presente em sala de aula e, principalmente, do material ao qual haviam tido contato anteriormente.

Ao final da solução do problema, os alunos encarregavam-se de enviar uma página HTML para um endereço específico, onde um repositório destes era atualizado a cada semana com as novas tarefas executadas.

Muitas destas tarefas eram divididas em grupos, geralmente pequenos, tal como é sugerido pela proposta PBL. Para aliar os conceitos de CSCW, vê-se a distribuição geográfica dos alunos, que muitas vezes encontravam-se em locais diferentes daqueles onde eram dadas as aulas.

Discussões e avisos referentes aos temas de aula eram participados através de uma lista de discussões, administrada pela professora titular da disciplina.

Os alunos tinham sua participação registrada na atividade através de um formulário que armazenava seus nomes e os locais de onde estavam participando da aula.

Através de murais virtuais, os estudantes podiam saber quem estava presente à aula e também buscar referências para seus trabalhos.

Ao final da disciplina, efetuou-se uma avaliação acerca da metodologia utilizada e do rendimento percebido pelo aluno. Tal questionário foi aplicado pela professora titular da disciplina.

Como conclusão desta proposta prática, observou-se que a aplicação dos conceitos percorridos ao longo do capítulo encaixaram-se de forma a concretizar os objetivos a que se propuseram. Aos estudantes, foram aumentadas as exigências relativas ao conteúdo programático, haja vista a elevada carga de tarefas extra-classe. Como pontos positivos, aponta-se o trabalho em grupo, onde notou-se uma maior integração entre os colegas da disciplina, acrescido do apoio necessário às tarefas realizadas em grupo. Considera-se, com base nestes dados aqui apresentados, possível a utilização desta mesma metodologia em aulas e trabalhos cooperativos à distância.

Trata-se a seguir da referência feita no capítulo 2, a respeito da abordagem das ferramentas de suporte ao trabalho cooperativo, na forma de sistemas complexos e de alto nível de utilização, ou seja, quando integram as características aqui discutidas, adicionadas a outros conceitos, como funcionalidades para a organização de informações distribuídas e o acesso por diferentes tipos usuários para diferentes tipos de aplicações.

3.5 Ferramentas para Solução Cooperativa de Problemas

Através dos estudos de VIT [VIT 99a], percebe-se que, dentre os vários serviços disponíveis via Internet para a solução cooperativa de problemas percorridos no capítulo 2, poucos apresentam-se na forma de soluções integradas. Na maioria das vezes, o próprio usuário precisa descobrir meios de integrar tais serviços.

Para atender as necessidades específicas de EAD, existem várias soluções disponíveis, como será visto neste item. Destas, algumas apresentam características que

somente servem aos propósitos de ensino, enquanto que outras propõem soluções para problemas mais genéricos.

Analisadas as ferramentas, notou-se que, como características comuns, todas elas disponibilizam acesso a listas de discussões, salas de *chat* e algum tipo de serviço de *email*. A característica que mostrou-se mais destacada foi a utilização da Internet, como mídia de interação e troca de informações.

3.5.1 A Ferramenta WebSaber

Bucando-se ferramentas que validem os paradigmas apresentados, considerou-se conveniente apresentar o ambiente WebSaber que, devido as suas características baseadas em resolução de problemas que englobem o enfoque visto anteriormente pela proposta PBL, apresenta-se como convenientemente prático para a utilização em EAD [VIT 99a].

WebSaber [WEB 98b] é um ambiente construído sobre a Internet, voltado para a resolução cooperativa de problemas. O ambiente está organizado segundo um modelo de hipertexto e é apoiado em um editor cooperativo, um bloco de notas e em algumas ferramentas de comunicação e cooperação da Internet.

O sistema direciona-se, preferencial mas não exclusivamente, para alunos do 2º. segmento do Ensino Básico e do Ensino Fundamental, coordenados por um tutor. Dadas algumas dificuldades relacionadas ao trabalho cooperativo, ou seja, à confiança mútua, à reciprocidade, à interação, à responsabilidade individual e ao compartilhamento e à socialização das informações entre alunos e tutores, a necessidade de segurança em ambientes distribuídos e a busca de eficiência na execução das atividades, o número ideal de participantes envolvidos na resolução de problemas gira em torno de 10/15 alunos.

O ambiente sustenta-se, do ponto de vista pedagógico, nos cenários educacionais inovadores, mediados por computadores: aprendizagem baseada em problemas (PBL), aprender fazendo, aprender explorando e navegando para encontrar respostas.

Considera-se importante destacar que o *software* WebSaber diferencia-se de outros ambientes distribuídos para aprendizagem cooperativa através das seguintes características [WEB 98b]:

- Trazer para o *browser* ferramentas de comunicação e cooperação disponíveis na Internet;
- Utilizar ferramentas síncronas. Além de *software* para *chat*, o WebSaber põe à disposição de seus usuários um editor cooperativo;
- Oferecer ao tutor suporte para tutoria;
- Propor e apoiar uma forma abrangente de trabalho no ambiente. A forma de trabalho proposta para solução de problemas em WebSaber é uma variante dos passos utilizados no método científico clássico. O ambiente propõe que o problema a ser resolvido seja analisado, que algumas hipóteses sejam levantadas, ainda que de forma não explícita, que se procure chegar à solução, seguindo as etapas de coleta exaustiva de informações, exclusão de informações não contingentes ao problema, pré-seleção de idéias potencialmente relevantes, depuração destas idéias, seleção de idéias promissoras, teste destas idéias através

das etapas de planejamento e execução da solução. A fase de planejamento é crucial, pois nesta fase, a solução é formalizada em metas materializáveis, vinculadas a uma dimensão temporal, aos participantes envolvidos e ao compromisso efetivo dos participantes com a execução da solução encontrada.

3.5.1.1 Requisitos funcionais do ambiente

O sistema, em sua 1ª. versão, visa fornecer um ambiente seguro e amigável para trabalho educacional cooperativo, em *web browser*, com as seguintes funcionalidades:

- Fornecer *homepage* do ambiente;
- Fornecer ambiente de trabalho para o tutor, através do Modo Tutor;
- Prover *links*, a partir da *homepage*, para problemas disponíveis, para ajuda, para o Modo Tutor, para o ambiente, bem como para textos selecionados;
- Fornecer locais para a solução cooperativa de problemas: *Hall*, *SittingRoom* e *WorkRoom*;
- Fornecer um mural com problemas existentes;
- Fornecer formulário para a entrada de novos problemas;
- Dispor de cronograma, onde o tutor inclui, exclui e altera atividades e datas;
- Prover espaço virtual onde o tutor informa ao grupo datas de *chat* e alterações no cronograma;
- Fornecer formulário para cadastro de tutores e estudantes;
- Gerar automaticamente *homepage* dos participantes, a partir dos dados do cadastro;
- Tornar disponível as *homepages* geradas;
- Fornecer um local de trabalho (*WorkRoom*) onde são apresentadas as etapas e sub-etapas a serem seguidas para a solução de problema com as ferramentas de apoio correspondentes;
- Por a disposição dos participantes um bloco de notas individual, onde são anotadas e posteriormente recuperadas as anotações;
- Fornecer um editor cooperativo, para que os participantes construam e compartilhem passo a passo a solução do problema;
- Dar acesso a serviços de comunicação da Internet. Na 1ª. versão do ambiente, os serviços postos disponíveis são lista de discussão e *chat*.

3.5.1.2 Modelo de dados

No ambiente, um processo é definido como um conjunto de etapas e sub-etapas. Os processos em WebSaber são encadeados, ou seja, um processo só é iniciado quando as etapas e sub-etapas do processo anterior são concluídos. Ocorrem três processos no ambiente:

- Processo 1: processo de socialização, composto de etapas informais de conversas;
- Processo 2: processo de pré-solução do problema;
- Processo 3: processo de solução do problema propriamente dito.

3.5.1.3 Modelo de cooperação

O modelo de cooperação envolve a definição do suporte à interação e à coordenação. A interação em WebSaber tem somente a dimensão temporal. As interações temporais podem ser síncronas ou assíncronas. As interações síncronas requerem *hardware* robusto e eficiente controle de concorrência e de segurança de dados. Em WebSaber, são usadas interações síncronas e assíncronas.

Para encontros, planejamento, argumentação e execução da maior parte das tarefas, WebSaber oferece suporte para interações assíncronas. Para geração de idéias, co-execução de tarefas chave e encontros iniciais, o ambiente prevê o uso de interações síncronas. A coordenação dos trabalhos no WebSaber é sempre realizada pelo tutor.

3.5.1.4 Implementação

A característica marcante do WebSaber é sua implementação baseada em reutilização de ferramentas disponíveis na Internet e em soluções encontradas pela equipe de desenvolvimento do Projeto AulaNet, em andamento no Departamento de Informática da PUC-Rio (DI/PUC-Rio). O ambiente WebSaber, em *web browser*:

- Permite o acesso, através de *script CGI (Common Gateway Interface)*, ao editor cooperativo *NetMeeting* e ao *software* de *chat Conferencing Room*;
- Reutiliza diversas soluções desenvolvidas e encontradas pela equipe Aulanet para acesso via *web browser*: lista de discussão e gerador de *homepage* dos participantes;
- Reutiliza as rotinas de programação desenvolvidas pela equipe AulaNet para implementar a agenda de *chat*;
- Põe a disposição de cada participante um bloco de notas individual, implementado como um formulário HTML e acessado por senha individual.

O ambiente é composto de:

- *Homepage*: uma página de abertura, com *links* para ajuda e os locais de Websaber: *Hall*, *SittingRoom* e *WorkRoom* e para textos selecionados sobre o ambiente;
- *Hall*: página com formulário HTML apresentando a relação de problemas disponíveis;
- *SittingRoom*: página HTML com *links* para:
 - cadastro de participantes, gerando as *homepages* dos participantes;
 - bloco de notas (formulário HTML);
 - agenda para sessões de *chat* com apoio de *Conferencing Room*;

- lista de discussões;
- cronograma de atividades.
- *WorkRoom*: página HTML com *links* para:
 - *NetMeeting* (editor cooperativo);
 - sessões de *chat* com apoio de *Conferencing Room*;
 - lista de discussões;
 - bloco de notas: formulário HTML acessado por senha individual.
- Ajuda: páginas HTML com dupla função: explicar aos usuários sobre a utilização do ambiente e fornecer sugestões de atividades educacionais ao tutor.

Um protótipo operacional do ambiente WebSaber foi implementado, incorporando todas as suas funcionalidades, e está disponível em <http://beatles.les.inf.puc-rio.br/websaber/ambiente/indice.htm>.

3.5.2 A Ferramenta TANGO *Interactive*

Buscando-se uma estratégia para o suporte à solução cooperativa de problemas, tema que norteia este trabalho, procurou-se por uma ferramenta que pudesse agregar os conceitos já ilustrados anteriormente neste trabalho e que fossem conjugados a uma *interface* de fácil entendimento e utilização, com exigências simples de plataforma e recursos computacionais [VIT 99a].

Considerou-se conveniente apresentar o ambiente TANGO *Interactive* que, devido as suas características baseadas na independência de plataforma de *hardware* e *software*, mostra-se convenientemente prático para a utilização em EAD e CSCW.

Seguindo-se as medidas indicadas em suporte *on-line* [TAN 98] e também aquelas conseguidas através de avaliações práticas, pôde-se chegar aos resultados que serão a seguir relacionados.

3.5.2.1 Avaliação

O sistema TANGO *Interactive* é uma ferramenta de colaboração, para a Web, que estende as capacidades dos *web browsers*, permitindo uma completa interatividade, rica em recursos multimídia e que também oferece um ambiente colaborativo multi-usuário.

É também um *framework* para a construção de sistemas completamente colaborativos já que possibilita a construção de sistemas colaboradores de complexidade arbitrária utilizando este *framework*.

O sistema é escrito em Java. Seus módulos são usualmente implementados como *applets* que interagem umas com as outras e podem controlar o comportamento umas das outras. A interação das *applets* em TANGO vai muito mais além da comunicação trivial entre algumas *applets* em uma página HTML. *Applets* TANGO podem vir de espaços de nomes completamente diferentes. Não há exigência de que diferentes instâncias da mesma *applet* venham do mesmo servidor HTTP.

As *applets* podem ser carregadas quando necessárias e descarregadas a qualquer momento, assegurando que o sistema seja leve e ágil. Segundo suporte técnico obtido em [TAN 98], TANGO *Interactive* é o primeiro e, aparentemente, o único sistema que implementa esta arquitetura muito flexível e poderosa., trabalhando de uma forma muito simples. Inicia-se o sistema a partir de um *web browser*, o qual, depois disso, pode continuar sendo utilizado "para" e "com" outras aplicações. O sistema faz a conexão a um servidor TANGO da escolha do usuário. Uma vez no sistema, pode-se usar uma das mais de três dúzias de aplicações colaboradoras para as seguintes tarefas:

- Participar de trabalhos em projetos que integrem equipes de colegas;
- Assistir aulas em universidades virtuais;
- Criar e utilizar uma sala de conversas pública ou privada;
- Criar uma videoconferência;
- Assistir a um filme;
- Surfar na Internet;
- Jogar um *game*, como xadrez.

Pode-se fazer tudo isso ao mesmo tempo, em qualquer combinação, em tantas salas de conversa quanto se queira. Estas são as demonstrações de flexibilidade e escalabilidade que conferem a esta ferramenta características únicas. Pode ser esta uma de suas maiores vantagens, portanto.

Seguindo-se a estes fatos, as descrições obtidas de [TAN 98], consideram que Java *applets* são leves para fins de consumo de recursos computacionais, mas leva-se tempo para carregá-las, além de poucos desenvolvedores estarem aptos a escrevê-las. Sendo assim, Javascript é uma alternativa para implementação de aplicações clientes, o que viabiliza ao sistema TANGO poder tomar uma página HTML e automaticamente torná-la uma entidade compartilhada.

Além disso, aplicações Web escritas em C ou C++, podem ser utilizadas como parte de um sistema colaborativo pessoal. O sistema TANGO *Interactive* simplifica esta tarefa, além de disponibilizar suporte para conexões entre aplicações, o que alivia o desenvolvedor de pensar em qualquer infra-estrutura de colaboração, dispensando-o para concentrar-se em visualizar a aplicação.

A flexibilidade do produto chega ao ponto de poder-se construir um mundo 3D (3 Dimension) VRML (*Virtual Reality Modeling Language*), povoá-lo com *avatars* e lhes deixar interagir, disponibilizando-se estas facilidades através de nodos JavaScript VRML.

TANGO *Interactive* pode ser considerado um avançado, poderoso e extensível colaborador Web. As aplicações disponíveis com esta versão exemplificam cada uma das tecnologias que foram mencionamos anteriormente e estão disponíveis para uso de teste, no momento da elaboração das referências deste trabalho.

3.5.2.2 Conceitos Básicos

As aplicações TANGO *Interactive* são usadas em sessões. Considera-se que existe uma sessão quando um ou mais usuários estiverem conectados ao mesmo servidor, comunicando-se um com o outro e utilizando aplicações TANGO. As

aplicações em uma sessão podem se comunicar, enquanto que aplicações rodando em sessões separadas não podem.

Os usuários do TANGO *Interactive* são conectados a um servidor colaborativo. O servidor é completamente invisível para os usuários e estes podem estar conectados a somente um servidor a cada vez.

Qualquer usuário pode iniciar qualquer número de sessões de aplicação de qualquer tipo. Não existem limites no número de usuários, número de sessões ou número de usuários em uma sessão. Em particular, um usuário pode iniciar ou participar em múltiplas sessões da mesma aplicação.

Pode-se iniciar sessões selecionando-se o ícone da aplicação. Participa-se de (ou deixa-se de) uma sessão existente selecionando-a da lista de sessões ativas e pressionando-se o botão "*Join session*" / "*Local close*". Convida-se alguém para uma sessão pressionando o botão "*Remote open*" e selecionando seu nome da lista *pop-up*. Para aplicações que requerem níveis de controle, o usuário concede/recebe controle selecionando uma sessão da lista e pressionando o botão "*Become master*" / "*Grant master*".

As aplicações TANGO não possuem botões "*Exit*". Fecha-se aplicações utilizando o botão "*Local close*" no TANGO *Control Application*.

O TANGO *Control Application* (CA) é a "peça central" do sistema. Constitui-se de uma *applet*, a qual controla todas as outras *applets* no sistema, iniciando-as, fechando-as e comunicando-se com estas quando necessário. Existem muitas formas de TANGO CA, algumas delas com GUIs (*Graphical User Interfaces*) completamente diferentes. Todas elas são compatíveis e completamente capazes de conversar umas com as outras. A GUI que se vê pode ser diferente em versões posteriores, mas ainda irá suportar a mesma funcionalidade básica.

3.5.2.2.1 TANGO *Interactive Control Application* (CA)

O painel principal do sistema é o "centro de controle" para o sistema colaborativo inteiro. Entre outras coisas, ele é usado para:

- Exibir nomes de aplicações colaborativas disponíveis;
- Permitir o controle de aplicações colaborativas: lançar e terminar instâncias de aplicações em máquinas locais e/ou remotas, unir e partir de sessões de aplicações existentes;
- Exibir o estado do sistema: usuários do sistema e listas de sessões ativas;
- Controlar certas características globais do sistema tais como modelos de controle de sessões ou mensagens de voz.

3.5.2.2.2 Configuração do CA

O CA possui algumas opções de configuração. Elas são controladas através da tabela "*Configurations*" na pasta correspondente. Pode-se manipular o comutador de comunicação de voz, expor o console do sistema e ajustar a sessão e o nível do modo de

controle. Esta etiqueta também exibe a tabela de configuração de aplicação. A própria tabela é configurada externamente pelo administrador do sistema.

3.5.2.2.3 TANGO *Interactive Mail*

O TANGO *Interactive* suporta seu próprio sistema de correio eletrônico assíncrono. No momento, este correio interno é disponível somente se o CA do usuário está ativo. Pode-se utilizar as facilidades do correio para enviar mensagens públicas ou privadas para outros usuários correntes. O correio é facilmente acessível pressionando-se a etiqueta *MailBox* na pasta correspondente. A entrega da correspondência em TANGO é imediata. O recipiente é notificado via o ícone de correio, bem como com mensagens de áudio. O correio do TANGO *Interactive* possuirá *interface* para sistemas externos de correio padrão no próximo lançamento oficial do sistema [TAN 98].

3.5.2.2.4 Pastas de aplicações

A pasta de Ferramentas Colaborativas exibe múltiplas pastas de aplicações. Elas são usadas pelos grupos de aplicações colaborativas TANGO *Interactive*, em categorias lógicas. Pressiona-se na etiqueta da pasta de aplicações para ver as aplicações disponíveis, pressionando-se então em um ícone de aplicação para lançar a aplicação localmente. Pastas de aplicações podem ser configuradas pelo administrador do sistema.

3.5.2.2.5 Botões de funções

Botões de funções (como "*Join session*" ou "*Become master*") são usados para lançar instâncias locais ou remotas de aplicações colaborativas e para controlar o nível de sessão das aplicações (ou seja, para decidir quais usuários controlam a aplicação a qualquer determinada instância). Nota-se que não se usa botões de função para iniciar uma sessão de aplicação. Para iniciar uma sessão, simplesmente pressiona-se o ícone da aplicação em uma das pastas de aplicações. Os botões de funções são (semi)inteligentes e não deixarão que se entre em ações obviamente insensatas (por exemplo, como excluir alguém em outra sessão).

3.5.2.2.6 Pasta de usuários

A pasta "*Users*" lista todos os usuários de um servidor TANGO *Interactive*. Qualquer um pode tornar-se um usuário iniciando uma das muitas formas do *Control Application*. O painel do usuário também provê informação sobre o estado dos usuários: escolhas correntes são disponibilizadas, informações sobre ocupado ou ausente, bem como sobre o tempo da última mudança. Um usuário pode, mas não tem que ter uma sessão de aplicação ativa. Um usuário pode participar em múltiplas sessões. Para ver a informação completa do estado de um usuário, seleciona-se e pressiona-se duplamente a entrada deste no próprio painel. Do painel *pop-up* descobre-se as sessões nas quais o usuário está participando, e se houver o desejo de comunicar-se com este, pode-se "pingar" para o usuário (basta verificar se o CA do usuário ainda está ativo), e "grita-se" com ele (para chamar sua atenção). Também pode-se convidar um usuário para unir-se a uma sessão de conversa ou a um jogo (ou qualquer outra aplicação disponível) iniciando

ambas as instâncias locais e remotas da aplicação que se gostaria de utilizar conjuntamente.

3.5.2.2.7 Sessões

A pasta "*Collaboratory Tools*" possui o painel que lista todas as sessões de aplicações ativas.

O sistema TANGO *Interactive* permite múltiplas sessões de uma aplicação (ou seja, pode-se ter múltiplas janelas de conversa, falando em cada uma delas com um grupo diferente de usuários).

Cada sessão possui um usuário mestre, listado no painel de sessão. O estado mestre pode ser dinamicamente transferido de um usuário para outro. Para algumas aplicações, o estado mestre define quem controla a aplicação. Muitas aplicações de sistemas permitem que múltiplos usuários operem concorrentemente a aplicação. Para tais aplicações, o estado mestre somente controla a habilidade de um usuário para escolher uma sessão de aplicação e/ou permitir que outros usuários unam-se a esta aplicação.

Seleciona-se uma sessão existente quando se deseja unir-se a ela, sair dela, pegar o nível (ou seja, tornar-se mestre desta sessão), ou fechar a sessão inteira. Os botões de funções destacados dizem que ações podem ser tomadas em determinado momento.

3.5.2.3 Requisitos para instalação

3.5.2.3.1 Um servidor TANGO

- Requer JDK (*Java Development Kit*) 1.1.2 ou posterior;
- É estável e escalável;
- É independente de plataforma: testado (segundo [TAN 98]) em IRIX 5/6, Solaris, Linux, Windows NT e também é executável em Windows 95 (mas isto não é recomendado);
- Disponível para *download* em <http://trurl.npac.syr.edu/tango>;
- Na versão para Windows NT, exige JVM (*Java Virtual Machine*) e também o Microsoft *NetMeeting*.

3.5.2.3.2 Um cliente TANGO

- Download:
 - Versão UNIX:
 - http://trurl.npac.syr.edu/tango/use_it/Download/download.html;
 - Versão Win32: <http://trurl.npac.syr.edu/tango>.
- Tipo de pacote:
 - Versão UNIX: *gzipped tar file*;

- Versão Win32: .exe file. (válido para Win NT/95/98).
- Utilização:
 - Versão UNIX: não exige privilégios de *root* para a instalação, cria sua própria estrutura de diretórios, após a reinicialização do *browser* o sistema já pode ser utilizado;
 - Versão Win32: não exige diretório específico para instalação, cria sua própria estrutura de diretórios, atualiza o diretório de *plug-ins* do *browser* e também o diretório de classes Java, traz junto ao pacote o *software* BuenaVista. A versão analisada exige *Netscape Navigator*.

3.5.2.4 Utilização do sistema TANGO *Interactive*

Utilizou-se a seguinte plataforma de *hardware* e *software* para testes:

- Servidor:
 - *Software*: Windows NT 4.0 e TANGO *Interactive Server*
 - *Hardware*: PC Pentium 200 MHz
- Cliente:
 - *Software*: Windows 95 e TANGO *Interactive Client 1.0*
 - *Hardware*: PC Pentium 200 MHz
- Rede:
 - LAN (*Local Area Network*) Ethernet
 - Protocolo TCP/IP

Um tutorial detalhado de utilização desta ferramenta pode ser obtido em [VIT 99a], em português e, pela Internet, em inglês, em [TAN 98].

3.5.3 A Ferramenta Interloq

De acordo com as especificações obtidas em [INT 99], caracteriza-se o *software* Interloq como um programa de computador utilizado através de uma rede de computadores, para auxiliar colaboradores a beneficiar-se de atividades de aprendizado cooperativo em salas de aula ou em EAD. O produto utiliza desenhos para criar discussões complexas que possam ser utilizadas de forma a obter-se simplicidade para a sua revisão, organização e entendimento.

É um produto desenvolvido para a satisfação das seguintes necessidades próprias de trabalhos cooperativos:

- *Brainstorm*;
- Debates;
- Tomada de decisões;
- Escritas;

- Solução de problemas;
- Pesquisa.

Através da utilização de uma *interface* amigável, a ferramenta permite que estudantes "novatos" obtenham simples discussões após poucos minutos de treinamento, enquanto que estudantes mais "experientes" podem utilizar características avançadas, como uma discussão pela Internet. Através desta possibilidade os alunos podem beneficiar-se das vantagens de *hiperlinks* entre notas e *web pages*, da procura de textos e das múltiplas opções de visão permitidas por tal paradigma.

O pacote Interloq pode ser utilizado na plataforma Windows 3.1 ou 95, em redes Novell, em computadores Macintosh e pela Internet, contanto que possua-se conexão WWW. O programa requer menos espaço em disco que um processador de textos (segundo informações de [INT 99]).

Os benefícios da utilização desta proposta, descritos em [INT 99], são os seguintes:

- Organizar uma larga faixa de atividades cooperativas;
- Salvar discussões para posterior revisão e impressão;
- Dar aos estudantes uma visão dos fatos de seus trabalhos;
- Permitir nivelamento entre os tipos de participações dos estudantes;
- Dar a qualquer participante iguais oportunidades de contribuição;
- Desenvolver a criatividade, senso de análise crítica e de escrita, além de habilidades de leitura.

3.5.3.1 Requisitos para instalação

3.5.3.1.1 Um servidor Interloq

O programa Interloq, quando utilizado sem o *Interloq Server*, permite que cinco ou dez usuários (dependendo do tipo de licença) comuniquem-se em uma LAN. Com a utilização do servidor, mais usuários podem comunicar-se simultaneamente sobre a Internet, conseguindo a combinação adequada para EAD e grandes instalações de LANs. O *software* servidor é executado em um computador servidor, utilizando a plataforma Windows ou Macintosh, recebendo e enviando notas para outros computadores que estejam utilizando o *software* cliente.

- Servidor Macintosh:
 - *Software*: System 7.x e TCP/IP (MacTCP) *networking*
 - *Hardware*: Power Mac, 68020 ou superior
- Servidor Windows:
 - *Software*: Windows (3.1/95/NT) e TCP/IP (WinSock) *networking*
 - *Hardware*: PC com processador Intel i386 ou superior

3.5.3.1.2 Um cliente Interloq

Por ser um programa baseado em *interface* gráfica, o cliente basicamente requer um teclado e um *mouse*. O *software* está disponível em versões para inglês e francês.

- Cliente Macintosh:
 - *Software*: System 7.x e TCP/IP (MacTCP) *networking*
 - *Hardware*: qualquer Macintosh
- Cliente Windows:
 - *Software*: Windows (3.1/95/NT) e TCP/IP (WinSock) *networking*
 - *Hardware*: PC com processador *Intel* i 286 ou superior

3.5.3.1.3 Utilização do sistema Interloq

Através da série de figuras mostradas a seguir, pode-se acompanhar os passos envolvidos na utilização do sistema, especificamente no que refere-se a manutenção de uma conferência. Os passos referentes a discussões, notas, árvores de discussões, visualização de conferências, *linking* dinâmico e utilização de menus, podem ser obtidos em [INT 99].

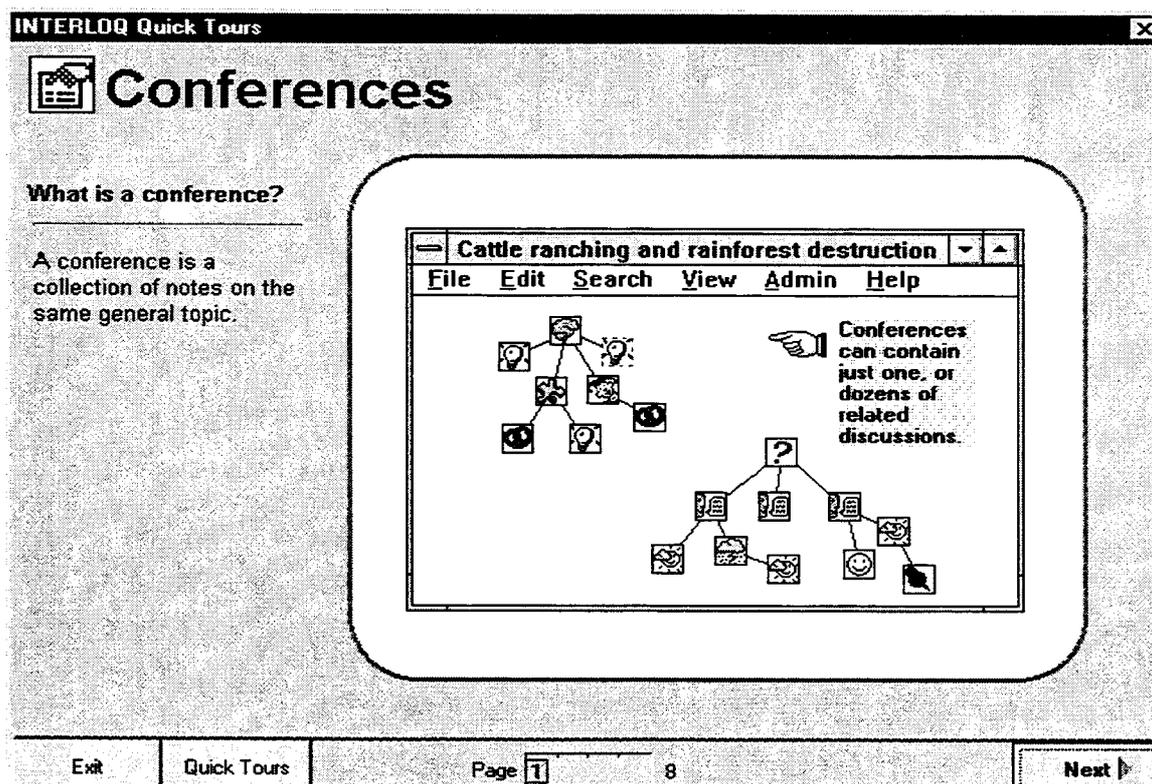


FIGURA 3.1 - A tela principal de conferências

Uma conferência pode conter somente uma ou dúzias de discussões relacionadas. Para criar-se uma nova conferência, deve-se escolher a opção *New* do

menu de arquivos, definindo-se então a categoria da conferência e o título, conforme observa-se na figura 3.2.

Na figura 3.3 observa-se que uma nova conferência é simplesmente uma janela em branco, que mostra o título da primeira na barra de títulos. Para iniciar uma discussão, utiliza-se o *mouse* para selecionar o lugar onde deseja-se que esta apareça e então seleciona-se o tipo de discussão.

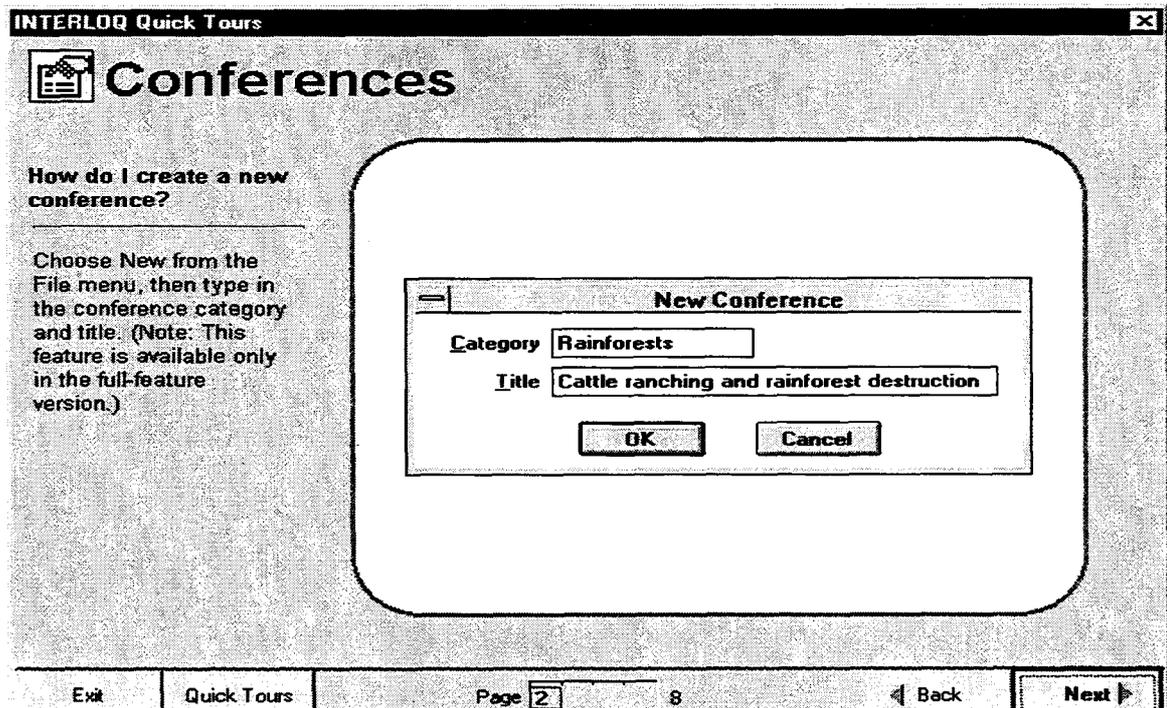


FIGURA 3.2 - A tela de criação de novas conferências

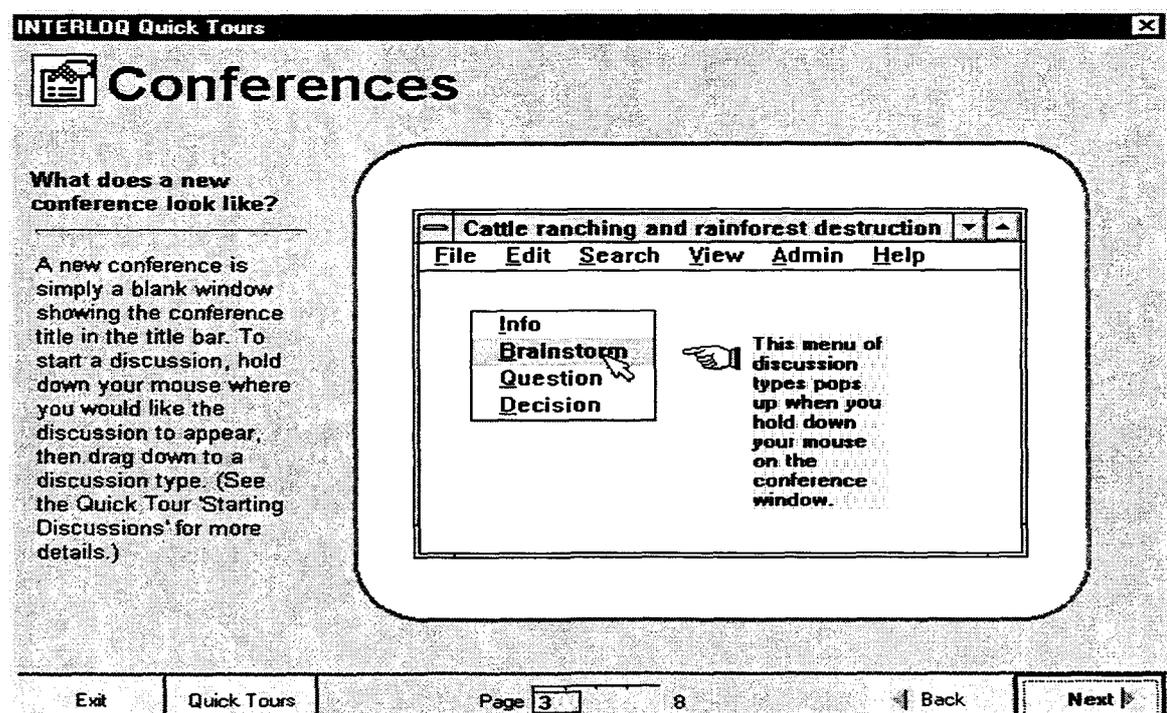


FIGURA 3.3 - A tela de conferências

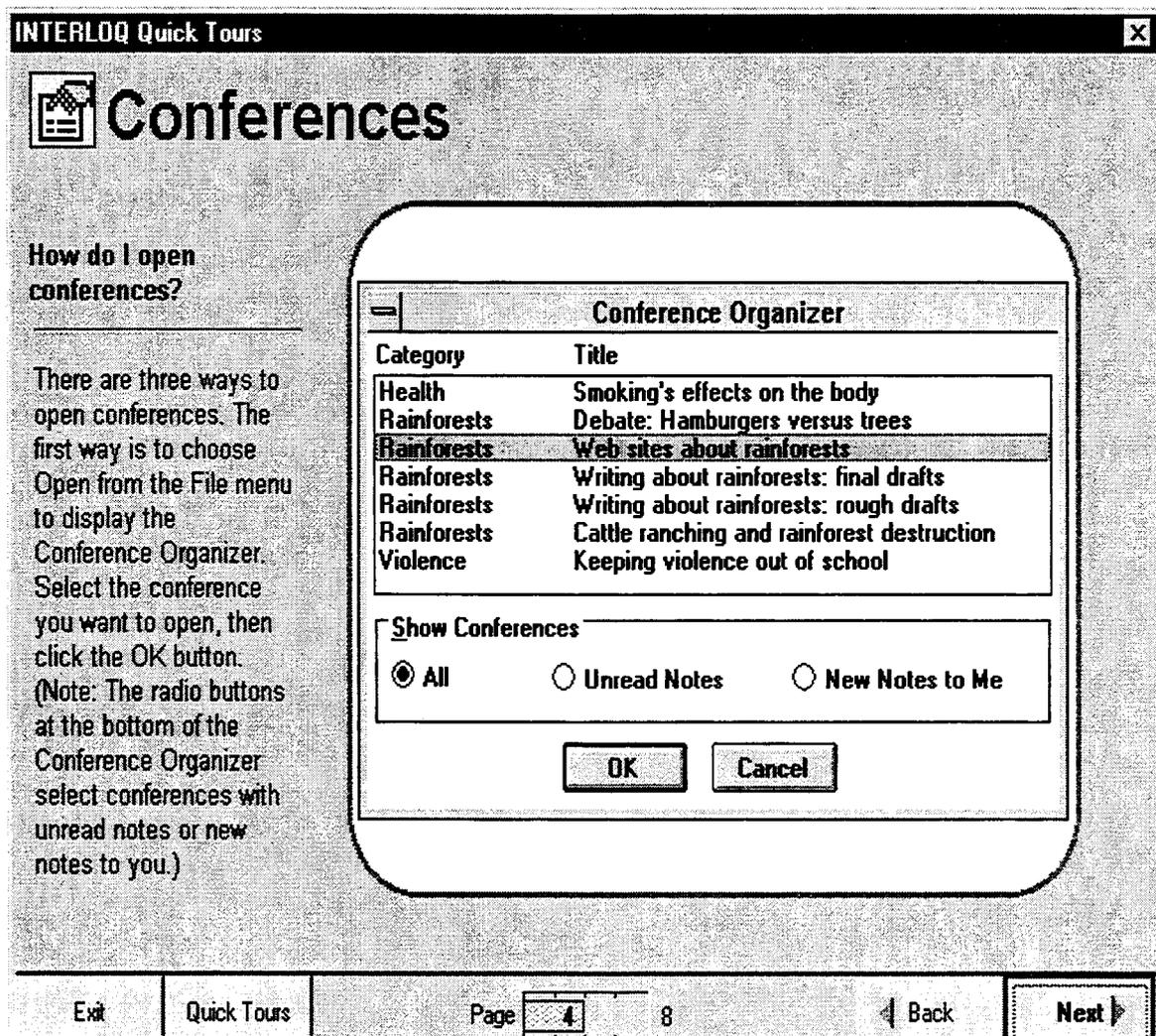


FIGURA 3.4 - A tela de abertura de conferências: *Conference Organizer*

Existem três maneiras disponíveis para abrir-se uma conferência. A primeira, demonstrada pela figura 3.4, é escolher a opção *Open* do menu de arquivos para mostrar o *Conference Organizer*. Seleciona-se a conferência desejada e pressiona-se o botão OK.

A segunda maneira, é através do *File Manager*. Escolhe-se *Associate* no menu de arquivos, digita-se "ARG" na caixa *Files with Extension*, selecionando-se em seguida o *path* para o programa "INTERLOQ.EXE". Pode-se, então, abrir arquivos de conferências, os quais possuem a extensão ".ARG", através de uma dupla pressão no *mouse*, conforme observa-se na figura 3.5.

A figura 3.6 ilustra a terceira possibilidade para manipular-se uma conferência, através da utilização de *links* para conferências na WWW. Primeiramente, deve-se configurar o *software* Interloq como um "*Netscape Helper Application*".

As conferências a que o usuário liga-se via WWW aparecerão no *Conference Organizer*, e é através deste que pode-se acessar tais conferências, sem a necessidade de retornar para outra página.

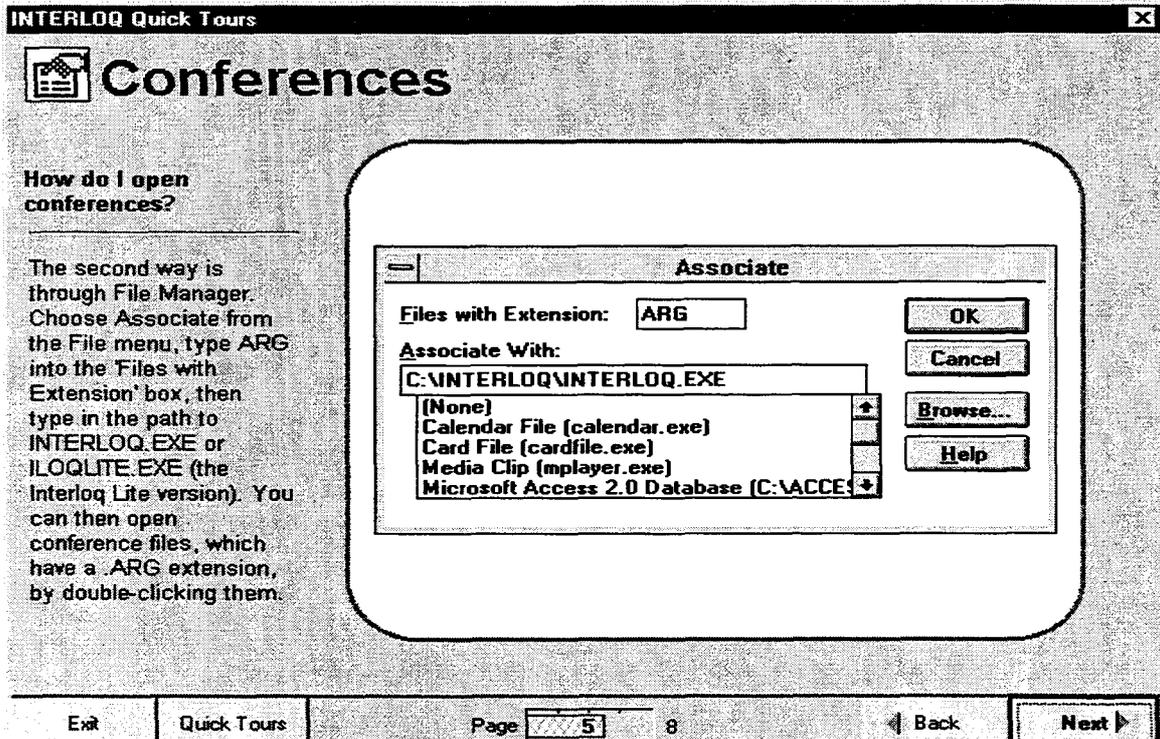


FIGURA 3.5 - A tela de abertura de conferências: Associate

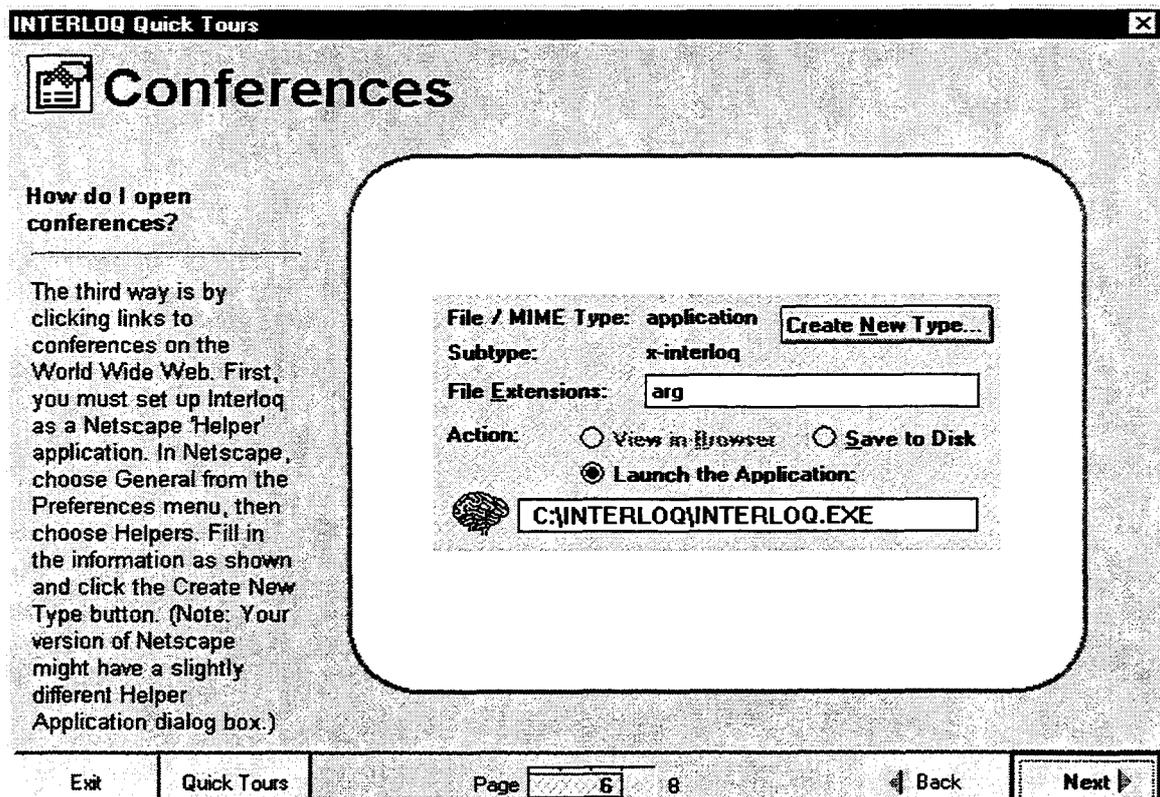


FIGURA 3.6 - A tela de abertura de conferências: WWW

Maiores informações a respeito desta ferramenta, podem ser obtidas, em inglês, em [INT 99].

3.5.4 A Ferramenta SAACI

Baseando-se nos trabalhos de Otsuka [OTS 98], observa-se que a principal contribuição do modelo proposto é integrar, em um único sistema, as principais funcionalidades julgadas necessárias para o apoio à aprendizagem colaborativa, através de redes de computadores.

3.5.4.1 Comunicação

A comunicação entre os participantes é organizada em sessões, fóruns de discussões e conversas informais.

As sessões permitem a realização de dois tipos de comunicação: as reuniões e as conferências. As sessões abrangem os encontros virtuais mais formais, marcados com antecedência, e que possuem objetivos específicos. As reuniões são encontros virtuais do grupo, que ocorrem durante todo o processo de colaboração realizados para o tratamento de questões relacionadas ao projeto de colaboração do grupo, tais como o planejamento das atividades, debate de temas relacionados ao trabalho desenvolvido e avaliação do trabalho realizado. As reuniões virtuais podem ser realizadas com o auxílio de ferramentas de *chat*, videoconferência ou listas de discussões. Observa-se que tais facilidades são explanadas e relacionadas neste trabalho, no capítulo 2.

O fórum de discussões é um espaço de discussões aberto ao grupo durante todo o processo de colaboração. As participações não são marcadas e tampouco possuem objetivos específicos, como nas sessões. Neste espaço, as contribuições ocorrem de forma espontânea, à medida que os participantes percebem a necessidade de tirar dúvidas, dar contribuições ou trocar idéias sobre assuntos relacionados ao trabalho desenvolvido pelo grupo. O fórum de discussões pode ser implementado através de listas de discussões, *newsgroup* ou ferramentas de conferência eletrônica textual.

As conversas informais ocorrem naturalmente entre os participantes do grupo, e possibilitam uma maior integração entre estes. Os assuntos tratados nestas conversas não estão necessariamente relacionados ao trabalho desenvolvido pelo grupo e podem ser comparados aos "bate-papos de corredor", que ocorrem nas interações presenciais. As conversas informais podem ser realizadas através de *email* ou através de ferramentas de *chat* ou videoconferência.

3.5.4.2 Tomada de decisão

O suporte a tomada de decisões é considerado necessário durante as reuniões, onde são realizadas as discussões de vários temas, sendo que os participantes muitas vezes tem de tomar decisões, escolhendo uma proposta que satisfaça a maioria.

Para facilitar a tomada de decisões, o modelo proposto por Otsuka [OTS 98] prevê a criação de uma ferramenta de votação, que possibilita o cadastro de propostas a serem votadas, a votação destas, a contagem dos votos e a divulgação dos resultados.

3.5.4.3 Coordenação

A coordenação das atividades do grupo é dividida em duas etapas: o planejamento e o acompanhamento das atividades planejadas. Na fase de planejamento o grupo necessita discutir sobre a definição de metas, prazos e a divisão e distribuição das tarefas. Esta fase é executada através de sessões de reuniões do grupo, destinadas a este fim. Quando o grupo chega a um acordo sobre a divisão e distribuição de tarefas, todas as definições sobre as tarefas (data, hora inicial e final e a descrição da atividade) de cada participante devem ser registradas na agenda do grupo.

3.5.4.4 Percepção

No suporte à percepção a proposta preocupou-se em prover formas de percepção que possibilitassem, tanto a detecção da presença de cada participante e suas últimas ações, como também a percepção de ações passadas do grupo.

Desta forma, todas as ações importantes de cada participante são registradas automaticamente pelo sistema e podem ser consultadas, por todos, a partir de duas opções: ações atuais e relatório de ações passadas.

3.5.4.5 Compartilhamento

Para possibilitar a formação, do que Otsuka [OTS 98] chama de uma "memória organizacional do grupo", o SAACI possui uma base de dados onde são armazenadas todas as informações necessárias sobre o grupo. A cada atividade do grupo procura-se registrar os principais resultados, o que é de grande importância para a orientação do desenvolvimento das atividades futuras. Dessa forma, são registradas as atas de sessões, dúvidas de conferências passadas, documentos e anotações do grupo, os *bookmarks* registrados durante a navegação de participantes do grupo pela Web, além das ações e agenda do grupo.

3.5.4.6 Construção de conhecimentos

O processo de construção de conhecimentos dá-se principalmente durante a comunicação entre os participantes do grupo e durante a "navegação" do aluno pelas páginas da Web.

Observa-se que, durante o processo de comunicação implementado nesta proposta, os alunos comportam-se de acordo com os conceitos discutidos no item 3.4 deste capítulo.

3.5.4.7 Representação dos conhecimentos

A representação dos conhecimentos é suportada através da autoria e anotação colaborativa de hiperdocumentos. Cada participante fica responsável pela autoria de um documento ou partes de um documento, que após concluído deve ser lido e analisado

pelos demais participantes, os quais também podem representar seus conhecimentos através da anotação de sugestões e observações ao documento.

3.5.4.8 Avaliação colaborativa

Nos programas de aprendizagem colaborativa é bastante comum que cada participante tenha de avaliar a participação dos demais e a qualidade do produto resultante do trabalho destes. A avaliação de todo o processo de colaboração pode ser realizado através da constante observação das ações dos participantes do grupo e da realização de reuniões, com a finalidade de discutir esta questão. É possível também a avaliação dos documentos criados pelos participantes do grupo, que representam o produto do trabalho destes. A avaliação dos documentos é realizada através de anotações contendo sugestões e críticas a estes.

3.5.4.9 Considerações sobre a implementação

A implementação do protótipo SAACI foi realizada utilizando a linguagem HTML, *scripts* CGI, a linguagem Javascript e o banco de dados MiniSQL 2.0. Todas estas referências podem ser encontradas nos textos de Otsuka [OTS 98], que considera que, através da implementação do protótipo, foi possível a confirmação da viabilidade do uso do WWW como base para a implementação de um sistema de apoio à aprendizagem colaborativa. Dentre as principais vantagens observadas pode-se citar:

- O usuário não precisa ter uma cópia do SAACI instalado em sua máquina, basta ter um cliente WWW (*browser*). O conteúdo da aplicação torna-se disponível ao usuário através do cliente WWW, que, durante a execução, realiza as requisições da URL (*Uniform Resource Locator*) de cada "página" que compõe a aplicação, ao servidor WWW. Desta forma é fácil a distribuição e atualização do sistema;
- Devido ao vasto alcance da Internet e independência de plataforma do serviço WWW, o sistema pode ser acessado por qualquer pessoa conectada à Internet, independentemente da plataforma de *hardware* (PC (*Personal Computer*), Macintosh ou Sun etc) ou sistema operacional usado (Windows, Unix, OS/2 etc);
- A comunicação entre os clientes WWW e o servidor WWW é realizada através do protocolo HTTP, o que torna os detalhes de comunicação transparentes ao desenvolvedor do sistema;
- A realização de chamadas às aplicações externas ao sistema pode ser realizada facilmente através da configuração destas como *helper applications* do cliente WWW.

3.6 Análise Comparativa dos Sistemas Inspecionados

Baseando-se no que foi descrito nos itens 3.2 e 3.3 deste trabalho, buscou-se, nos sistemas apresentados no item 3.5, o referencial prático que pudesse embasar o que

foi teoricamente explanado sobre utilização de CMC e seu uso em EAD, como suporte a PBL.

Para esta análise, conta-se com referências práticas observadas através da avaliação de produtos [TAN 98, INT 99] e referências teóricas, obtidas através de material de suporte [OTS 98, WEB 98b]. Considera-se importante ressaltar que os trabalhos de Otsuka [OTS 98] referenciam [WEB 98b], porém, de forma mais sucinta que a apresentada no item 3.5.1 deste trabalho.

A tabela 3.3 apresenta uma comparação entre os sistemas analisados, em relação ao suporte destes às atividades descritas por Otsuka [OTS 98] e sumarizadas a seguir:

TABELA 3.3 - Análise comparativa dos sistemas inspecionados

	WebSaber	TANGO <i>Interactive</i>	Interloq	SAACI
Comunicação síncrona e assíncrona	•	•	•	•
Suporte à tomada de decisão	○	○	•	•
Suporte à coordenação do grupo	•	○	•	•
Suporte à percepção do grupo	○	•	•	•
Suporte ao compartilhamento de informações	•	•	•	•
Suporte à construção do conhecimento	•	•	•	•
Suporte à representação de conhecimentos	•	•	•	•
Suporte à avaliação colaborativa	•	•	•	•
<p>• suporta completamente</p> <p>○ suporta parcialmente</p>				

- Comunicação síncrona e assíncrona: a primeira é utilizada em casos que necessitem interação em tempo real e a segunda, quando esta situação não é crucial;
- Suporte à tomada de decisão: considera que a ferramenta permita a resolução de conflitos, através da negociação de propostas entre os participantes para a tomada de uma decisão em conjunto;
- Suporte à coordenação do grupo: envolve o planejamento das atividades, a distribuição de tarefas e o acompanhamento da execução destas;
- Suporte à percepção do grupo: cada participante deve ter percepção das ações dos demais para, em um contexto individual, contribuir para uma maior sinergia do grupo;
- Suporte ao compartilhamento de informações: permite o acesso de todas as informações a todos os participantes;
- Suporte à construção do conhecimento: o sistema provê meios para possibilitar e estimular a participação ativa dos envolvidos através da pesquisa, da troca de idéias, da argumentação e da reflexão;

- Suporte à representação de conhecimentos: o sistema provê meios para que, após a construção colaborativa de conhecimentos o grupo materialize e compartilhe o que produziu, através da autoria de documentos;
- Suporte à avaliação colaborativa: os participantes podem avaliar seu próprio trabalho e também o trabalho de seus parceiros.

Analisando-se a tabela 3.3, percebe-se que os sistemas avaliados atendem aos requisitos considerados, por Otsuka [OTS 98], como essenciais para a aprendizagem colaborativa através de redes de computadores. Desta forma, a utilização de qualquer um destes produtos garantirá, em maior ou menor grau, a satisfação destes requisitos, com exceção de que alguns pontos propostos pelo sistema TANGO *Interactive*, como total suporte ao desenvolvimento de módulos de aplicações em Java, não mostraram-se tão flexíveis na prática. Em contato com membros da equipe de desenvolvimento do sistema [POD 98], não obteve-se o resultado esperado quanto às facilidades para uma parceria na produção de novos objetos. Talvez tal resultado seja devido ao fato de que a ferramenta possa vir a tornar-se disponível em versão comercial em pouco tempo.

O ponto que considera-se mais importante é a possibilidade de que todos os produtos podem beneficiar-se dos recursos disponibilizados pela Internet, seja para interação entre os usuários, seja para a utilização das várias mídias disponibilizadas por este meio.

Percebe-se, desta forma, que o trabalho em grupo detém valiosas atenções relativas ao desenvolvimento de aplicações complexas que visem facilitar o trabalho cooperativo suportado por computador. Tais aplicações revelam-se fundamentais, tanto para o entendimento e a aplicação de metodologias relativas ao ensino descentralizado, quanto para quaisquer outros tipos de atividades, como gerência e coordenação de trabalhos em equipe.

Um dos aspectos mais interessantes verificados nesta análise é que, durante um processo cooperativo suportado por computador, são gerados e manipulados grandes volumes de informações, muitas das quais devem passar por um processo de "filtragem" para evitar um acúmulo de dados pouco importantes, em detrimento de informações valiosas.

Neste contexto, busca-se, na continuidade deste trabalho, avaliar estudos e possíveis ferramentas que permitam realizar automaticamente o agrupamento de idéias, com foco ou temas comuns para a organização de tópicos discutidos no decorrer de um processo decisório em grupo, situação esta comum na Educação a Distância e no trabalho cooperativo.

4 Proposta de um Sistema para Apoio ao Trabalho Cooperativo em EAD

Para introduzir os conceitos que serão discorridos neste capítulo, buscou-se embasamento teórico nos trabalhos de Barua [BAR 99], cujo artigo, a partir deste ponto comentado, mostra que as aplicações para Internet, como WWW criam a possibilidade de desenvolvimento de plataformas colaborativas globais para suportar interações entre profissionais e acadêmicos em várias disciplinas. Considera-se que a utilização de *web browsers* revolucionou a forma de utilização da Internet e faz-se necessária a definição teórica, para o desenvolvimento de aplicações colaborativas para esta mídia.

A ausência de padrões proprietários na Internet torna-a um sistema aberto, onde usuários de qualquer lugar do mundo podem comunicar-se e interagir um com o outro, sem a necessidade de possuir aplicações específicas ou proprietárias. Em virtude da habilidade de ligação de recursos de informação de qualquer posição geográfica global, a Web é considerada como um passo revolucionário na direção da disseminação de informações e interações ao redor do mundo.

Apesar deste extraordinário potencial, Barua [BAR 99] prevê um sério problema de excesso de informações (Web "*infoglut*"), em um futuro próximo, quando da ausência de uma fundamentação conceitual compartilhada para a organização das informações na Web. Além disso, para a WWW tornar-se um fórum eletrônico para interação produtiva entre pesquisadores e profissionais de várias disciplinas, considera-se a necessidade de existir capacidades interativas integradas com repositórios de informação distribuídos sobre a Internet.

De acordo com esta visão, um sistema colaborativo (denominado *Collaboratory*) provê uma plataforma eletrônica aberta para grupos ou indivíduos com interesses comuns para a eficiente troca, disseminação e criação de assuntos, idéias e conhecimentos. Ao contrário de uma aplicação ou sistema proprietário, que provê somente algumas escolhas customizadas para os usuários através de *interfaces* limitadas, um sistema aberto como a Internet provê uma grande faixa de escolhas, as quais requerem uma cuidadosa consideração por parte dos desenvolvedores.

Na análise, do texto de Barua [BAR 99], o primeiro tópico que considera-se essencial para a fundamentação deste trabalho é o que discorre sobre o gerenciamento de interação, pois define-se que os sistemas colaborativos devem prover mecanismos para gerenciar interações de forma ordenada. Para fins de exemplificação, ilustra-se que, se um artigo é posto em discussão e se os usuários fazem comentários e sugestões, suas introduções serão menos significativas na ausência de uma maneira de organizar tais interações. Entretanto, se os comentários puderem ser organizados de acordo com suas posições em relação ao tema central e a metodologia de trabalho, um novo usuário irá achar relativamente fácil saber o que os outros pensam a respeito de tal artigo. Da mesma forma, o próprio objeto da discussão deve ser escrito de uma forma estruturada, de forma que os autores deixem claras as suas posições referentes ao assunto e prevejam um suporte analítico ou empírico para isto. Sem estes mecanismos para a gerência das discussões e argumentações, pode ser impossível obter-se os melhores resultados que um sistema colaborativo pode oferecer.

É também importante a capacidade presente no sistema colaborativo de ligar assuntos e idéias relativas à discussão. Conforme cresce o número de discussões, esta

característica permitirá que o usuário possa focar uma discussão para examinar conteúdos e pontos de vista que tenham sido expressos em determinado tópico. Isto determina que um ambiente, composto por uma base de dados, seja invocado e possa unir as discussões relacionadas, além de dinamicamente reorganiza-las como *links* baseados em certos critérios semânticos. Nota-se que estas ligações envolvem a criação automática de *links* entre tópicos ou assuntos de discussão, baseados em sistemas de regras específicas. Isto é diferente de usuários de sistemas colaborativos fazendo combinações de *links* dentro de seus documentos ou mensagens.

O segundo tópico, que considera-se essencial para este trabalho, é o que exemplifica os benefícios de agregar várias características colaborativas em um mesmo sistema.

A possibilidade de ligar dinamicamente uma fonte de informações para dar suporte a um assunto ou a um modelo proposto é a característica chave de uma interação culta e construtiva. Ilustra-se, para fins de exemplo que, em uma discussão a respeito do valor empresarial de IT (*Information Technology* - Tecnologia de Informação), alguém que toma a posição referente a medida de problemas que conduzem ao paradoxo de produtividade em IT poderia criar um *link* para um ou mais artigos relevantes a seus comentários. Enquanto a Web permite *links* de hipertextos, por si só não possibilita uma plataforma para comunicação síncrona e assíncrona. Os benefícios de se possuir a plataforma e as capacidades de criação de *links* juntas em uma só ferramenta é maior que a soma de valores derivadas de se possuir dois sistemas, um com capacidade de criar *links* e outro com capacidade de gerenciamento de fóruns.

O segundo exemplo ilustrativo fala de sistemas colaborativos que suportam multimídia, mas não possuem uma característica interativa, como por exemplo, capacidades de comunicação síncrona ou assíncrona. Tais sistemas são considerados menos úteis ao usuário do que um que suporte ambas as características. Considera-se uma página Web tradicional, a qual possui capacidades multimídia mas não suporta interações com *newsgroups*. De outro lado, aos *newsgroups* faltam as capacidades multimídia de aplicações baseadas na Web. Os benefícios de se possuir ambas as características são bem maiores do ponto de vista do usuário.

Como forma de sumarizar suas teorias, Barua [BAR 99] cita que, para entender o custo envolvido na complementaridade, é teoricamente possível possuir todas ou a principais características utilizadas na Internet de maneira isolada. Como exemplo, cita que *sites* Web possuem um grande volume de informações relevantes sobre virtualmente qualquer assunto. A Web também suporta documentos multimídia. As BBS's permitem interações assíncronas em uma base global, assim como os IRC's provêem possibilidades de comunicação síncrona. Porém, da perspectiva de um usuário, torna-se tedioso o uso de uma aplicação para localizar um documento e de outra para postá-lo. Em outras palavras, o custo de tempo e esforço serão bem maiores para usar aplicações separadas para cada uma destas interações. Por outro lado, possuindo-se todas estas capacidades em um único sistema pode reduzir significativamente o tempo e o custo gastos pelo usuário no sistema colaborativo. Para especificação, toma-se o caso de capacidades de organização e busca de informações. A posse destas duas características juntas reduzirá bem mais o custo para o usuário do que se este utilizasse aplicações isoladas.

Desta forma, fundamentando-se nas propostas de Barua [BAR 99] e buscando desenvolver adiante as idéias de Souza [SOU 94], objetiva-se utilizar um sistema que, a partir de uma aula ou reunião virtual, possibilite realizar automaticamente o

agrupamento de idéias com foco ou temas comuns, através da utilização de técnicas específicas, como por exemplo DELPHI, técnica que é baseada em questionários – envio pelo coordenador e resposta dos participantes [BUI 87, COR 91].

Como objeto específico de estudo neste capítulo, têm-se o projeto de um sistema capaz de automatizar a sumarização de idéias com foco ou temas comuns, através da construção de uma *interface* WWW para o usuário final, seguida de um módulo de tratamento de *log* de reuniões e finalizando-se com a manipulação das informações geradas pelo chamado módulo de agrupamento implementado por Wives [WIV 99], que é o responsável pela manipulação das idéias geradas pelo grupo.

4.1 Trabalhos Prévios Considerados

Inicialmente, considerou-se os aspectos mais importantes dos trabalhos de Souza e Wives [SOU 94, WIV 99], sendo que, muitas das características e propostas de tais obras foram aproveitadas, haja visto os muitos métodos e objetivos comuns de tais estratégias com o objeto deste estudo. Também avaliou-se o trabalho desenvolvido por Magalhães [MAG 94], para uma possível integração com o modelo proposto.

Atendo-se ao desenvolvimento de Souza [SOU 94], observa-se que neste trabalho são estudadas técnicas para organização de tópicos discutidos no decorrer de um processo decisório em grupo, bem como algumas metodologias de apoio ao processo de tomada de decisão (em grupo). A partir deste estudo, Souza [SOU 94] utiliza-se de uma técnica específica - Técnica de Grupo Nominal - para o projeto e a construção de um protótipo de uma ferramenta denominada "*Issue Analyser*", que realiza automaticamente a etapa correspondente ao agrupamento de idéias - com foco ou temas comuns - geradas num momento anterior.

Este protótipo foi implementado para uso em estações SUN e tem a função de agrupar automaticamente as idéias similares geradas na fase de "*Brainstorming*", organizando-as de acordo com a definição do grupo. Este trabalho também realiza a validação de todo o sistema - SADGV2 (Sistema de Apoio à Decisão para Grupos versão 2) - implementado.

Nos trabalhos desenvolvidos por Wives [WIV 99], considera-se que técnicas de recuperação e análise de informações, principalmente textuais, são de extrema importância em um ambiente heterogêneo como é o caso da Internet. No âmbito deste trabalho, os problemas relacionados à sobrecarga de informações, que decorre devido ao grande volume de dados a disposição de uma pessoa, são os mais importantes.

Visando minimizar estes problemas, o trabalho de Wives [WIV 99] apresenta um estudo sobre métodos de agrupamento de objetos textuais, onde os objetos são organizados automaticamente em grupos de objetos similares, facilitando sua localização, manipulação e análise.

Considerou-se também relevante o estudo de Magalhães [MAG 94] devido ao fato deste trabalhar com frases em língua natural, ou seja, seqüências de palavras digitadas por pessoas, componentes de uma reunião, que devem ser processadas para posterior utilização por outros módulos do sistema proposto.

Resumindo, no processamento da língua natural, um dos aspectos mais importantes é a representação formal da linguagem dentro de um sistema de informações. Este sistema enfrenta um grande desafio: a representação dos diversos

aspectos lingüísticos da língua natural e o tratamento de "problemas" como ambigüidade, referências a nomes, pronomes, etc. Um dos aspectos abordado nesta proposta é a classificação das orações da língua portuguesa.

Com a proposição deste novo trabalho, busca-se avaliar estudos e possíveis ferramentas que permitam a administração e a análise do trabalho cooperativo entre membros de equipes dispostos em pontos geográficos diferentes e que disponham de plataformas de *hardware* e *software* básico heterogêneas e não necessariamente compatíveis entre si.

Tais características procuradas nestes sistemas os tornam flexíveis para a coordenação de tarefas conjuntas, principalmente aquelas responsáveis pela integração de módulos, com *interfaces* claras e intuitivas, que possibilitem chegar a aplicações de utilização genérica, como as utilizações necessárias a Educação a Distância [VIT 99a] ou a Gerência de Redes de Computadores, por exemplo, aproveitando-se das características já validadas pelas propostas anteriores, para a prototipação de um novo modelo, de características híbridas.

Continua-se o estudo com uma descrição individual dos sistemas que foram avaliados e utilizados para a implementação do protótipo proposto.

4.1.1 Língua Natural: Classificação de Orações

Conforme mencionado anteriormente, os trabalhos iniciais preocuparam-se com o tratamento da língua natural, já que pretende-se manusear frases da língua portuguesa, digitadas em contexto livre, por usuários sem uma especificação rígida de dicionários.

Para tanto, optou-se por testar uma ferramenta que pudesse receber uma frase nestas condições (livre de especificações) e que devolvesse uma análise sintática desta como resultado, para posterior tratamento por outros módulos. Por atender a estas especificações, escolheu-se o sistema desenvolvido por Magalhães [MAG 94]. Este sistema utiliza a representação formal, baseada em DRT (*Discourse Representation Theory*).

Inicialmente, define-se que a representação da Língua Natural em uma linguagem formal pode ser realizada sobre dois elementos: O Discurso e o Período. Quando o tratamento se dá sobre o Discurso, a representação em uma linguagem formal pode ser usada para analisar alguns aspectos do mesmo, como por exemplo a ambigüidade. Outras aplicações deste tipo de análise podem ser a resolução de anáforas, quantificação universal e existencial, condicionais, implicações e negações. Quando se trata de Período, a análise encontra alguns problemas, pois as referências a agentes envolvidos no Período poderão não ser resolvidas com as informações disponíveis no mesmo. A representação de Língua Natural em uma linguagem formal tem como pré-requisito básico a construção de uma Gramática, que definirá a estrutura gramatical das sentenças.

Magalhães [MAG 94], em seu trabalho, não conceitua precisamente o que é DRT. O que obtém-se é que:

- Em um extremo, a DRT é caracterizada como sendo uma teoria que faz previsões definidas sobre as possibilidades das anáforas nominais e sobre as condições de verdade de certa classe de sentenças, nas quais os Sintagmas Nominais tem um papel fundamental;

- No outro extremo, a DRT é tida como uma proposta geral para os significados lingüísticos, na qual as características e requisitos das teorias baseadas em modelos e da semântica formal são combinados com um conceito procedural, de modo a proporcionar um significado lingüístico correlacionado com a forma lingüística.

Entende-se por anáfora um tipo de ambigüidade encontrada na comunicação entre agentes (indivíduos de uma mesma comunidade, Sistema de Informação ou qualquer outra entidade capaz de realizar comunicação). Tais ambigüidades devem ser resolvidas e, para isso, o elemento que insere a ambigüidade deve ser ligado ao elemento por ele referenciado.

No trabalho avaliado, a DRT é utilizada para a representação das Orações do(s) Período(s). A DRT utiliza como estruturas para a representação semântica das sentenças simples as DRSs (*Discourse Representation Structures* - Estruturas de Representação do Discurso).

4.1.1.1 Estruturas de Representação do Discurso (DRS)

DRS é a estrutura utilizada na DRT para a representação do Discurso e, nesta avaliação, para a representação das Orações. Uma DRS é formada por dois elementos:

- Um conjunto de referentes do discurso;
- Um conjunto de condições.

Os referentes representam os objetos (elementos) indicados nas orações. As condições são predicados cujos argumentos são referentes. Assim, existem objetos e condições que representam as relações entre estes objetos e os qualificam. O significado lingüístico numa DRS é determinado pelo conjunto: referentes e condições.

Em uma DRS há a possibilidade de existir como condição uma outra DRS. Estas DRSs que aparecem como condições são chamadas de sub-DRSs, e a DRS que engloba todas as outras recebe o nome de DRS-principal. Esta última representa o discurso completo que está sendo analisado.

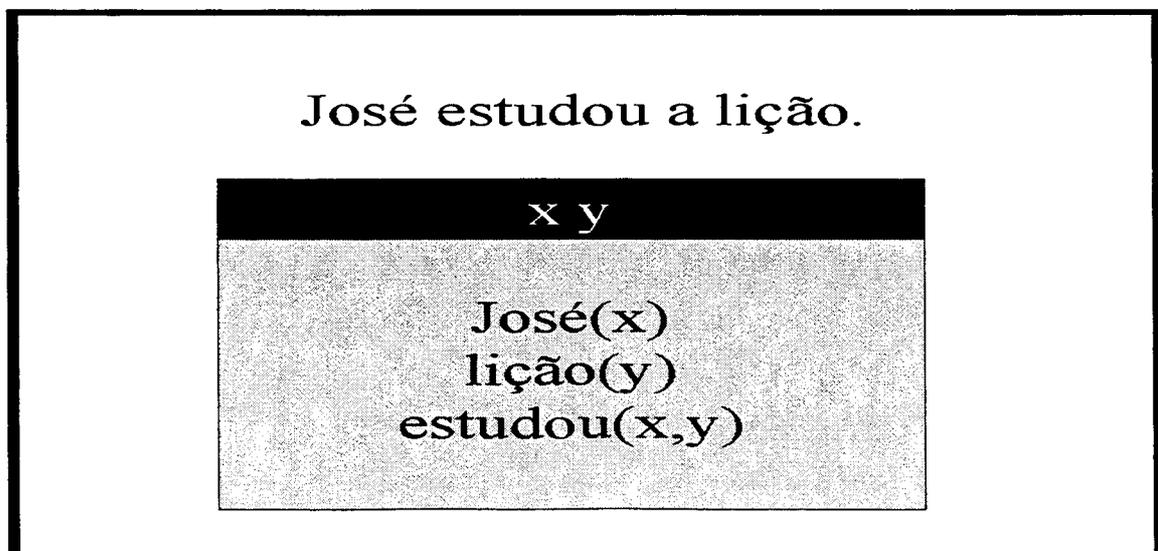


FIGURA 4.1 - Exemplo de DRS

A figura 4.1 mostra um exemplo de representação gráfica de uma DRS. Nota-se que os referentes estão representados pelas letras "x" e "y" e a condição expressa é "estudou(x,y)". "José(x)" e "lição(y)" associam, respectivamente, os referentes "x" e "y" aos referenciados "José" e a "lição". Os componentes da DRS são agrupados em uma caixa (Universo), que serve também para determinar o escopo e a acessibilidade dos referentes.

4.1.1.2 Construção de uma DRS

Antes de representar o Discurso, é necessário fazer-se uma análise sintática das orações. Esta análise pode ser feita com o uso de regras gramaticais (gramática) e, como resultado, obtém-se uma ADS (Árvore de Derivação Sintática), que representa as funções sintáticas de cada elemento (palavra) que compõe as sentenças.

A DRS é obtida a partir de reduções de sub-árvores da ADS. Esta redução é feita com base nas seguintes regras:

- Sintagmas nominais criam novos referentes e devem ser introduzidos no Universo da DRS-principal;
- A redução da sub-árvore cria uma condição formada pelo sintagma nominal, seguido do referente entre parênteses;
- O referente criado substituirá a sub-árvore na ADS.

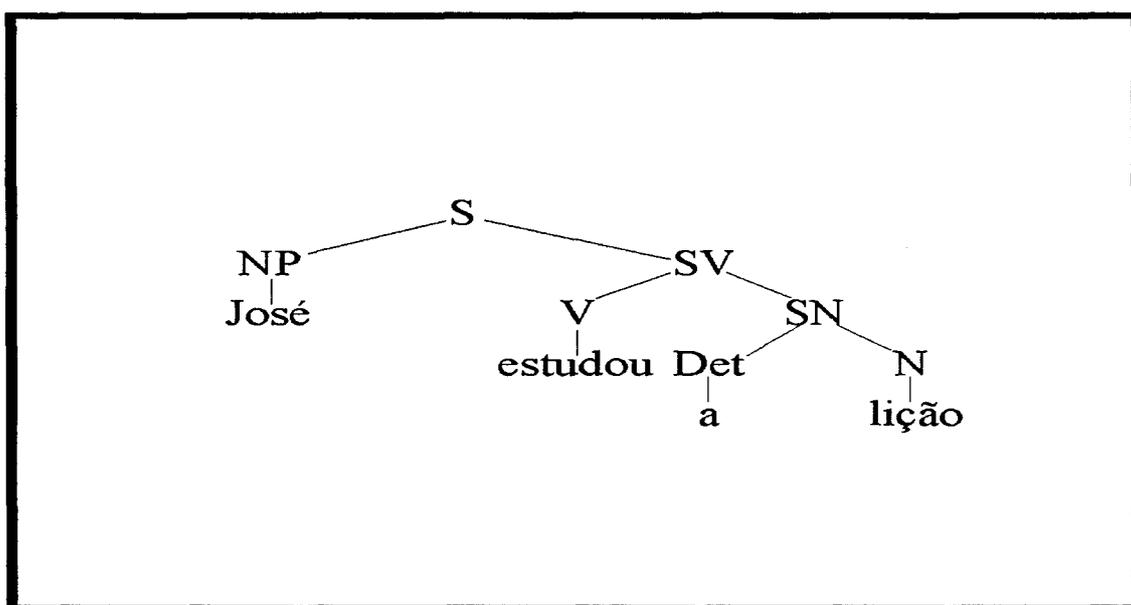


FIGURA 4.2 - ADS da Figura 4.1

A figura 4.2 mostra a Árvore de Derivação Sintática da sentença da figura 4.1. As figuras 4.3, 4.4 e 4.5 representam a aplicação das regras acima descritas para a árvore mostrada na figura 4.2 onde os elementos da primeira sub-árvore foram reduzidos, formando o conjunto de referentes e condições mostrados na figura 4.5. As sub-árvores que ainda não foram reduzidas recebem o nome de condições redutíveis.

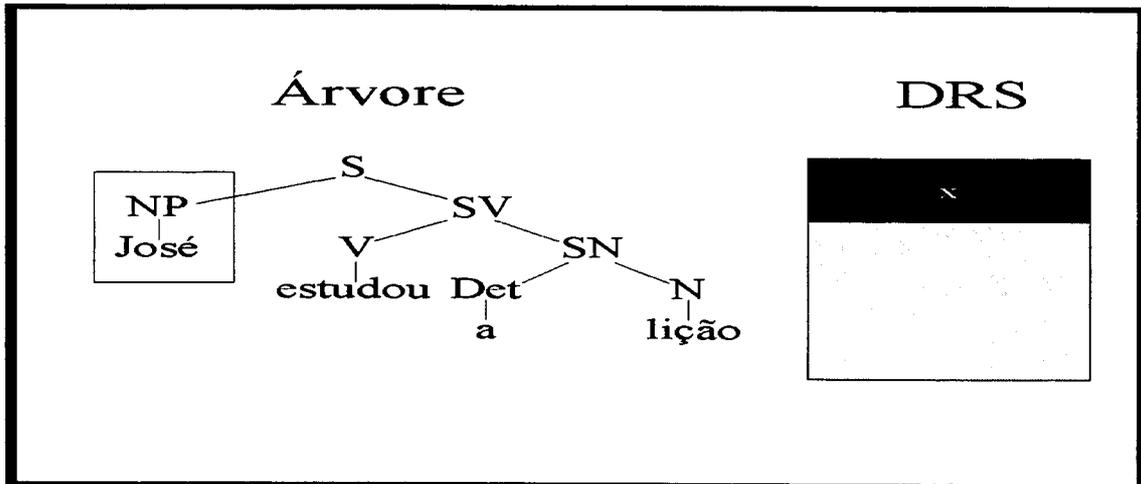


FIGURA 4.3 - Aplicação da primeira regra sobre a figura 4.2

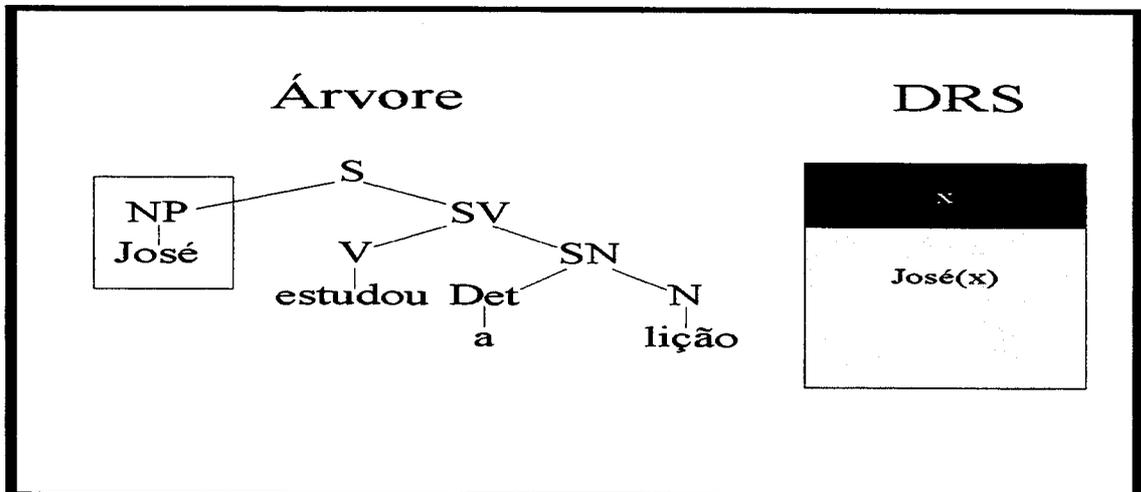


FIGURA 4.4 - Aplicação da segunda regra sobre a figura 4.3

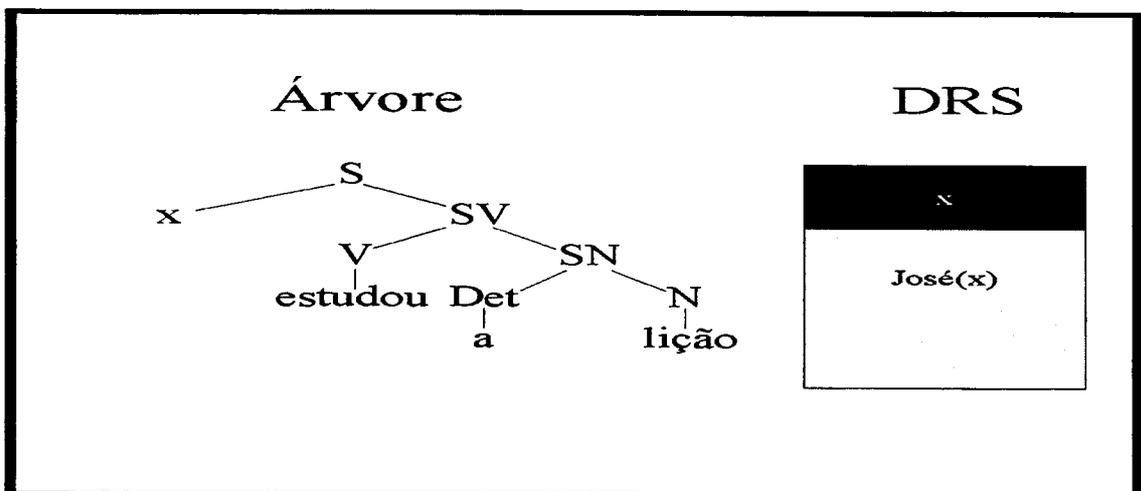


FIGURA 4.5 - Aplicação da terceira regra sobre a figura 4.4

Após a redução de "José", deverá ser feita a redução do outro sintagma nominal "a lição", fazendo-se uso dos mesmos passos. Assim, obtém-se a seguinte árvore (figura 4.6):

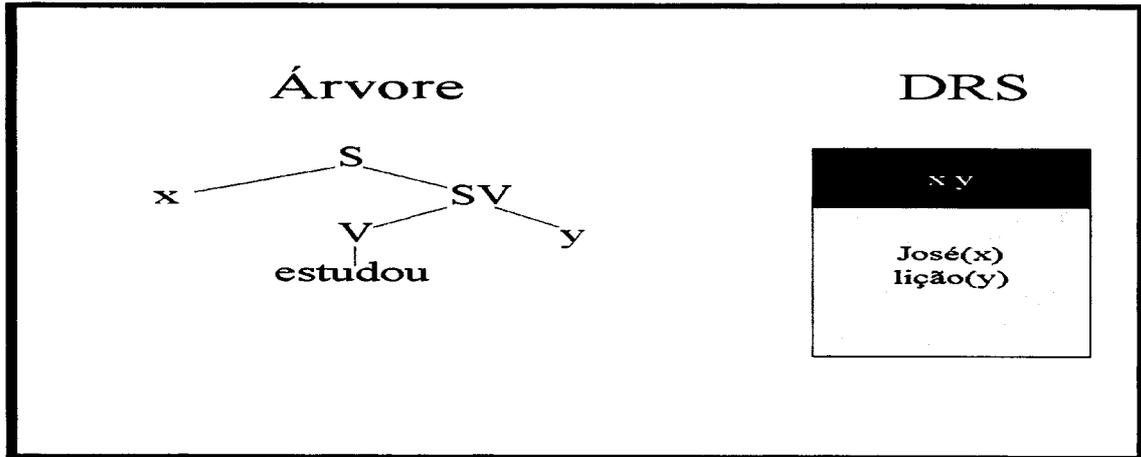


FIGURA 4.6 - Árvore obtida com a redução de "a lição"

A seguir, deve-se reduzir "x estudou y", o que gera a DRS da figura 4.7.

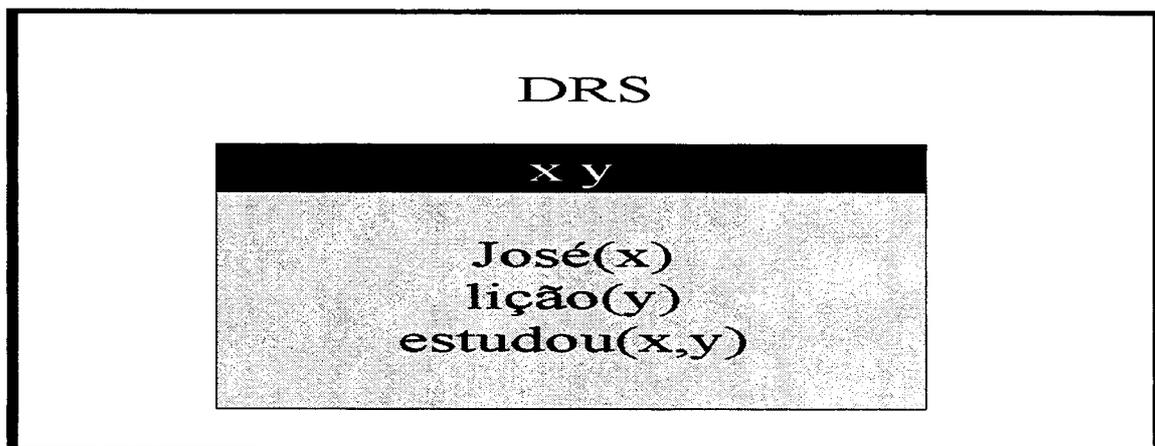


FIGURA 4.7 - Redução final da Árvore da figura 4.6

Segundo os resultados obtidos do trabalho avaliado, percebe-se que as figuras 4.1 e 4.7 são idênticas, mostrando como a DRS foi gerada a partir da Árvore de Derivação Sintática da sentença.

Conforme mencionado anteriormente, no início deste capítulo, considerou-se, *a priori*, essencial a utilização de um módulo de classificação da linguagem natural. Devido a esta consideração inicial, estudou-se detalhadamente o funcionamento deste sistema, conforme demonstra-se a seguir.

4.1.1.3 O Sistema de Classificação de Orações Implementado por Magalhães [MAG 94]

O objetivo deste sistema é a classificação de orações coordenadas sindéticas e as suas representações em DRSs simples. É representado de acordo com a figura 4.8.

Orações coordenadas sindéticas são orações que começam por conjunção coordenativa expressa. Recebem o nome das conjunções que as iniciam, havendo, portanto, cinco tipos: aditivas, adversativas, alternativas, conclusivas e explicativas [KAS 97].

Conjunções são os vocábulos gramaticais que servem para relacionar duas orações ou dois termos semelhantes da mesma oração. A conjunção coordenativa não se altera com a mudança de construção, pois liga elementos independentes [CUN 85].

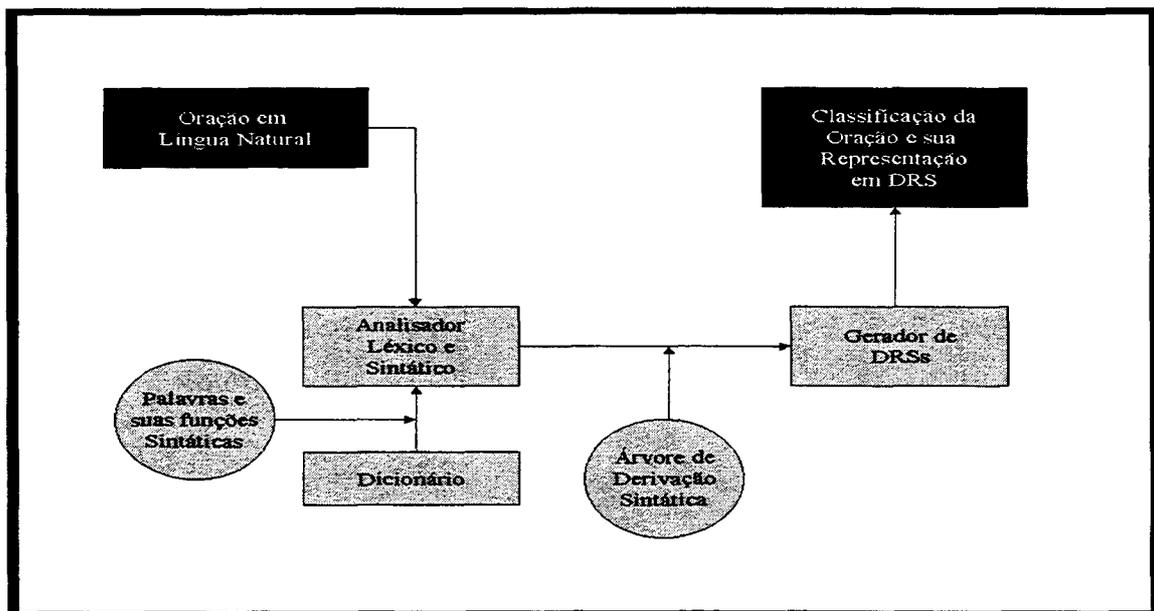


FIGURA 4.8 - Diagrama de Funcionamento do Sistema

Ilustrado pela figura 4.8, percebe-se que o sistema recebe uma oração do usuário e a passa para o Analisador Léxico e Sintático. Este faz uso das informações contidas no Dicionário para gerar a Árvore de Derivação Sintática da oração e a passa para o Gerador de DRS. Este último, por sua vez, executa a montagem da DRS com base nas regras e algoritmos anteriormente descritos e fornece como saída a classificação da oração, a DRS gerada e as ligações entre os adjetivos e os nomes a que se referem.

Devido a excelente documentação encontrada neste sistema, proposto por Magalhães [MAG 94] e pelas informações recebidas de Padilha [PAD 99], pôde-se recriar o sistema a partir de seu código-fonte e executar os programas que o constituem, em um ambiente de estações SUN, utilizando-se o compilador/interpretador C-Prolog 1.5 [PER 87, WIE 90, YOU 89]. Dentre seus módulos constituintes, destacam-se:

- **Dicionário:** módulo cujo objetivo é servir como vocabulário de palavras que podem ser inseridas no sistema. Tais palavras estão agrupadas de acordo com sua função gramatical, em cláusulas Prolog. Seu funcionamento inicial é verificar a existência da palavra consultada (pelo Analisador Léxico e Sintático) e retornar a função gramatical da palavra e informações como número, pessoa e gênero, por exemplo;
- **Analisador Léxico e Sintático:** analisa as frases de acordo com uma gramática, culminando por gerar a Árvore de Derivação Sintática da oração digitada pelo

usuário através do teclado. Somente uma sentença é processada a cada vez, palavra a palavra, sendo cada uma destas consultada no dicionário;

- Gerador de DRS: tem como entrada a ADS gerada pelo módulo anterior e seu objetivo é reduzir as sub-árvores em condições e referentes.

Na utilização prática do sistema, foram avaliados os exemplos disponíveis na referência consultada em Magalhães [MAG 94] e também foram inseridas no sistema frases que não constavam de tal documentação.

4.1.1.4 Avaliação do Sistema proposto por Magalhães [MAG 94]

O sistema comportou-se, durante sua avaliação, exatamente conforme descrito em sua documentação. Ao testar-se a frase "Katia não gosta de Sorvete.", obteve-se os seguintes resultados:

Classificação da Oração:

Período Simples

ADS gerada:

[sn(np(s,f,3,[katia])),sv(ev(s,f,3,negativa,[gosta]),sn(prep(de),sn(nome(s,m,3,[sorvete])))))]

DRS:

ref(0,1)

ref(1,2)

np(0,1,katia)

nc(1,2,sorvete)

carac(1,s,f,3)

carac(2,s,m,3)

v_drs(1,[1,2],[gosta]).

neg_universo(0,1).

Como este exemplo consta da documentação, tanto a entrada quanto a saída possuem resultados já esperados. Ao modificar-se poucas características da frase, neste caso "Katia gosta de limão.", percebe-se que o Sistema demora aproximadamente 30 minutos para retornar uma mensagem de erro, pois a palavra "limão" não consta de seu dicionário.

Desta forma, observa-se que, limitado ao tamanho de seu dicionário de palavras, o sistema somente funciona bem de acordo com esta constituição. Tal aspecto não pode ser modificado em tempo de execução, ou seja, o dicionário deste sistema "não aprende" novas palavras, o que dificulta sua utilização na prática, haja visto que pretende-se disponibilizar ao usuário uma discussão livre de vocabulários específicos, na forma de discussões via *chat*. Tais dificuldades são, portanto, resumidas ao tamanho do dicionário e à falta de capacidade de "aprendizagem" do sistema, aliados à sua implementação, que somente suporta a análise de uma única frase a cada vez, o que é outro empecilho, já que pretende-se trabalhar com centenas de frases ao mesmo tempo.

Por estes motivos, buscou-se novas alternativas a este sistema, conforme será descrito nos próximos itens.

4.1.2 Um protótipo da ferramenta "*Issue Analyser*": projeto, implementação e validação

Este sistema, desenvolvido por Souza [SOU 94], possui muitas características em comum com os objetivos previstos para o Sistema que será proposto. É um trabalho complexo e bem elaborado, mas que está limitado a características próprias de sua implementação e ambiente de utilização, como será visto neste item.

A proposta inicia-se da necessidade de criar uma nova ferramenta para o tratamento de idéias geradas no "*Brainstorming*" eletrônico de forma automatizada. Foi observado por Souza [SOU 94] que, na medida em que se possui grande quantidade de frases para agrupamento (nesta referência, chamadas idéias) mais difícil torna-se o processo de agrupar "manualmente" tais frases que possuem assuntos em comum. Além disso, o tempo consumido nesta tarefa manual é muito longo, podendo, muitas vezes, ocasionar a insatisfação e o desinteresse dos membros da reunião. Resumidamente, o objetivo deste trabalho inspecionado é o estudo/pesquisa de metodologias de apoio ao processo de tomada de decisão (em grupo).

Muitos conceitos, abordados por Souza [SOU 94], já foram ilustrados na apresentação dos trabalhos de Magalhães [MAG 94], especialmente os que tratam do processamento da língua natural. Desta forma, segue-se para a análise comparativa de técnicas de decisão em grupo, explicitadas pelo primeiro autor.

4.1.2.1 Técnicas de Decisão em Grupo: *Brainstorming*, Grupo Nominal e *DELPHI*

Entre as diversas técnicas de decisão em grupo, Souza [SOU 94] considera três destas como as de maior importância pela relevância mostrada para a compreensão de seu trabalho e que este considera utilizada em diversos trabalhos de pesquisa:

- Técnica *Brainstorming*;
- Técnica de Grupo nominal;
- Técnica *DELPHI*.

4.1.2.1.1 Técnica *Brainstorming*

É uma técnica utilizada por grupos para geração de um grande número de idéias. Tenta-se evitar julgamentos, avaliações ou críticas dessas idéias no momento em que estão sendo geradas. Seu princípio básico é o de que o grande número de idéias produzidas aumenta a probabilidade de se encontrar uma solução melhor para o problema. Seu procedimento básico é o seguinte:

- O problema é apresentado/descrito geralmente como uma questão aos participantes do grupo;
- É providenciada a privacidade dos participantes (separação física ou de outra forma);
- Os participantes elaboram e escrevem as suas idéias (em aproximadamente 10 minutos);
- Apresentação das idéias ao grupo;
- Avaliação das alternativas, com o uso de uma escala de 0 (pouca contribuição) a 4 (máxima contribuição para a solução do problema).

4.1.2.1.2 Técnica de Grupo Nominal

É uma técnica de decisão em grupo onde seus membros estão agrupados fisicamente, mas não interagem (diretamente) entre si. Foi esta a técnica utilizada por Souza [SOU 94] em seu trabalho. Os passos desta técnica são os seguintes:

- Apresentação do problema aos participantes da reunião, por exemplo, a colocação de uma questão a respeito do problema;
- Geração individual de idéias, sem diálogo ou troca de idéias entre os participantes;
- Apresentação das idéias ao grupo. Utiliza-se a técnica "*round-robin*", onde cada membro apresenta uma idéia por vez, numa manifestação seqüencial de todos os integrantes, até que cada um não tenha mais idéias a apresentar;
- Discussão e avaliação de todas as idéias geradas;
- Hierarquização de idéias. É a fase de votação destas (atribui-se individualmente o grau de importância de cada idéia);
- Obtenção do consenso do grupo, através da discussão dos resultados. É um passo que nem sempre é realizado, apesar de ser extremamente importante.

4.1.2.1.3 Técnica *DELPHI*

É uma técnica em que os participantes da reunião estão fisicamente dispersos. Todo o processamento desta técnica está baseado em questionários, que são enviados pelo coordenador e respondidos pelos participantes. A seqüência de tomada de decisões nesta técnica ocorre na seguinte ordem:

- Geração independente de idéias pela resposta ao primeiro questionário (quem responde são os membros do grupo);
- Análise e resumo do conteúdo dos questionários;
- Envio dos resumos dos questionários aos participantes, contendo a média do grupo e a posição individual do participante, seguido do envio do segundo questionário;
- Análise e agregação do conteúdo do segundo questionário com possível reformulação da posição dos participantes em função da opinião média do grupo, seguindo da convergência de idéias do grupo.

Os últimos dois passos podem ser repetidos até a obtenção de um consenso final. É uma técnica utilizada, principalmente, quando os membros do grupo não podem reunir-se em um mesmo local. Esta técnica permite estruturar a comunicação e a procura de convergência de opiniões. A tabela 4.1 apresenta a comparação entre as três técnicas estudadas por Souza [SOU 94].

TABELA 4.1 - Comparação entre as técnicas *Brainstorming*, Grupo Nominal e *DELPHI*

Parâmetros de Avaliação	<i>Brainstorming</i>	Grupo Nominal	<i>DELPHI</i>
Metodologia	Encontro Estruturado	Encontro Estruturado	Questionário Estruturado
Comportamento Normativo	Tolerância a não-conformidade	Tolerância a não-conformidade	Liberdade para discordar
Participação do membro do grupo	Igualdade entre os membros	Igualdade entre os membros	Igualdade entre os que respondem ao questionário
Qualidade das idéias geradas	Alta	Muito Alta	Alta
Grau de consenso na solução	Alto	Muito Alto	Alto
Satisfação de cada membro com a solução	Alta	Muito Alta	Alta

4.1.2.2 Eficiência de trabalho em grupos pequenos e grandes

Souza [SOU 94] revela que, com base em suas pesquisas e experiências em laboratório, existem diferenças entre o uso de um Sistema de Apoio à Decisão em Grupo em pequenos grupos (com tamanho menor ou igual a 17 participantes) e em grandes grupos (mais de 17 participantes). A tabela 4.2 apresenta um resumo destes aspectos.

4.1.2.3 O Sistema "Issue Analyser" Implementado por Souza [SOU 94]

O objetivo deste sistema é fornecer apoio ao processo de tomada de decisão em grupo, conforme menção anterior. Em documentação, explicitamente ilustrada em [SOU

94], através de descrição em representação gráfica SDL (*Specification and Description Language*), define-se que a função principal do protótipo desenvolvido, chamado "*Issue Analyser*" é o agrupamento e a organização de idéias geradas na fase "*brainstorming*", preparando-as para uma etapa posterior do processo de decisão (votação dessas idéias).

O projeto desta ferramenta foi baseado em conceitos segundo os quais um sistema desta natureza deve propiciar:

- Necessidade de pouco tempo de instrução para o uso, boa produtividade e facilidade para seu usuário;
- "*Feedback*", que significa que, após um comando, o usuário (membro do grupo) fica sabendo do efeito do mesmo;
- Uma *interface* mais natural e multifuncional;
- Uma resposta rápida aos comandos;
- Estabelecimento de um diálogo amigável, sem frustrações.

TABELA 4.2 - Sistemas de Apoio à Decisão em Grupo em pequenos/grandes grupos

	Pequenos Grupos	Grandes Grupos
Vantagens	<ul style="list-style-type: none"> • Maior eficiência; • Maior satisfação individual; • Maior oportunidade para contribuição de idéias; • Maior nível de coesão. 	<ul style="list-style-type: none"> • Maior exploração de alternativas; • Decisão de melhor qualidade; • Grande número de idéias geradas; • Liberdade para o membro expressar suas idéias (discordar).
Desvantagens	<ul style="list-style-type: none"> • Pontos de vista muito limitados; • Menor exploração de alternativas; • Maior pressão para participar consistentemente; • Decisão de menor qualidade. 	<ul style="list-style-type: none"> • Menor satisfação individual; • Inibição de alguns membros; • Maior pressão sobre o líder do grupo; • Dificuldades para alcançar o consenso.

O Sistema foi implementado através de diferentes linguagens de programação, dentre as quais, optou-se pela implementação de partes do protótipo em linguagem C e em C-Prolog. Também foi utilizado o Sistema Gerenciador de Banco de Dados INGRES.

A parte implementada em C-Prolog executa as seguintes funções:

- Realiza a análise léxica e sintática das mensagens dos participantes;
- Realiza o aprendizado de termos desconhecidos, que não estão no dicionário de palavras;
- Realiza a retirada dos termos negativos (de pouco significado, como artigos, preposições e conjunções) das mensagens;

- Realiza o agrupamento de mensagens com idéias semelhantes, com base em palavras-chave e expressões que apresentem em comum;
- Manutenção de um dicionário de sinônimos, denominado "*Thesaurus*";
- Indica os possíveis erros ortográficos contidos nas mensagens, tais como omissão, troca e/ou aumento do número de letras em uma palavra.

A parte implementada em C é responsável pelas seguintes funções:

- *Interface* gráfica com os participantes da reunião;
- Interação com o programa em C-Prolog;
- Suporte à alteração/consulta de dados de mensagens contidas na base de dados INGRES;
- Gerenciamento de janelas;
- Apresentação de resultados (agrupamento/síntese das mensagens) aos membros do grupo.

Como plataforma de *hardware*, o protótipo foi desenvolvido para uso em qualquer tipo de configuração de estações SUN. Tais estações devem estar conectadas em rede e precisam suportar serviços de RPC (*Remote Procedure Call*), além de ser necessário um meio físico para a conexão entre as estações.

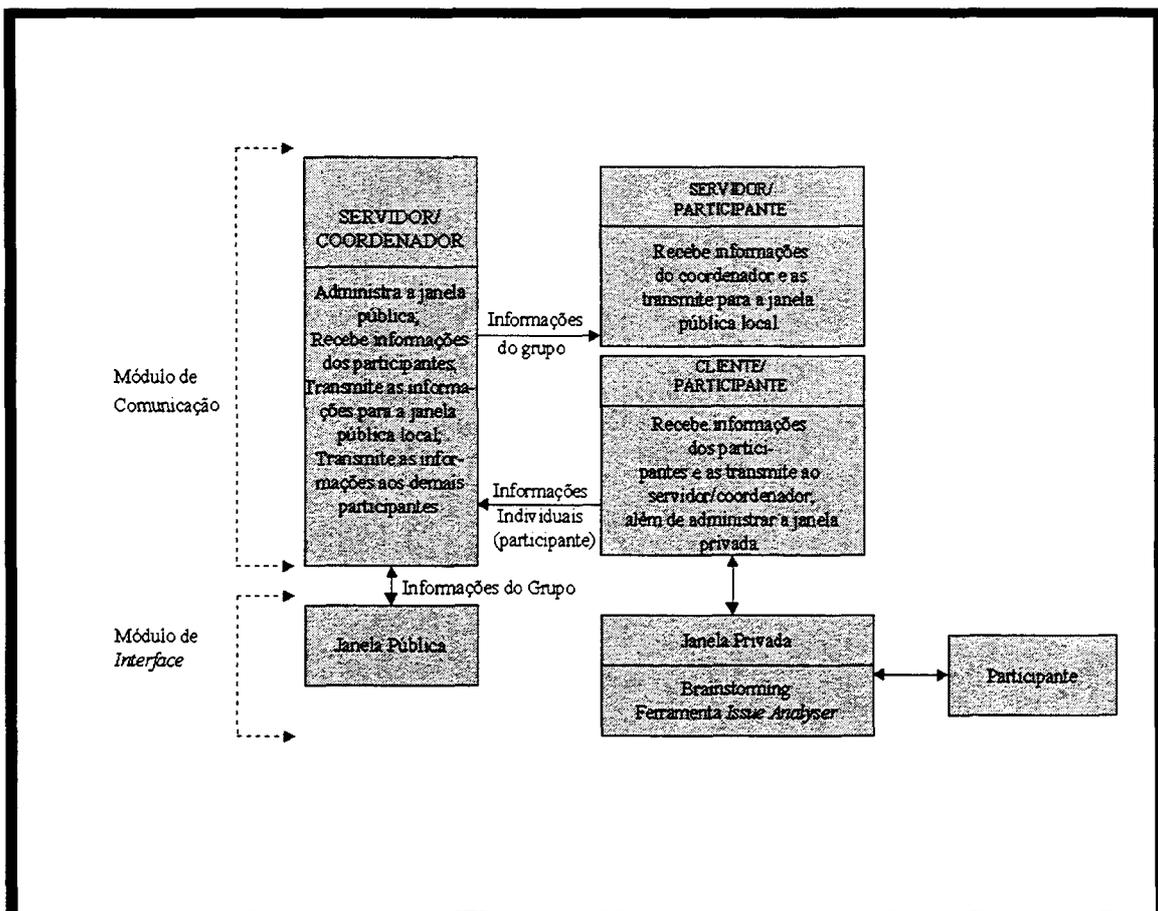


FIGURA 4.9 - Os Módulos de Comunicação e Interface do SADGV2

Conforme visualiza-se na figura 4.9, o módulo de *interface* é responsável pela troca de informações entre os participantes do encontro e o módulo de comunicação. Este último é responsável por efetuar a troca de mensagens entre os processos do protótipo, tanto locais (onde são utilizados *sockets*) quanto remotos (onde são utilizadas chamadas RPC).

A figura 4.10 ilustra os submódulos de comunicação entre servidor/coordenador, servidor/participante e cliente/participante, relacionados através dos respectivos mecanismos *sockets* ou RPC.

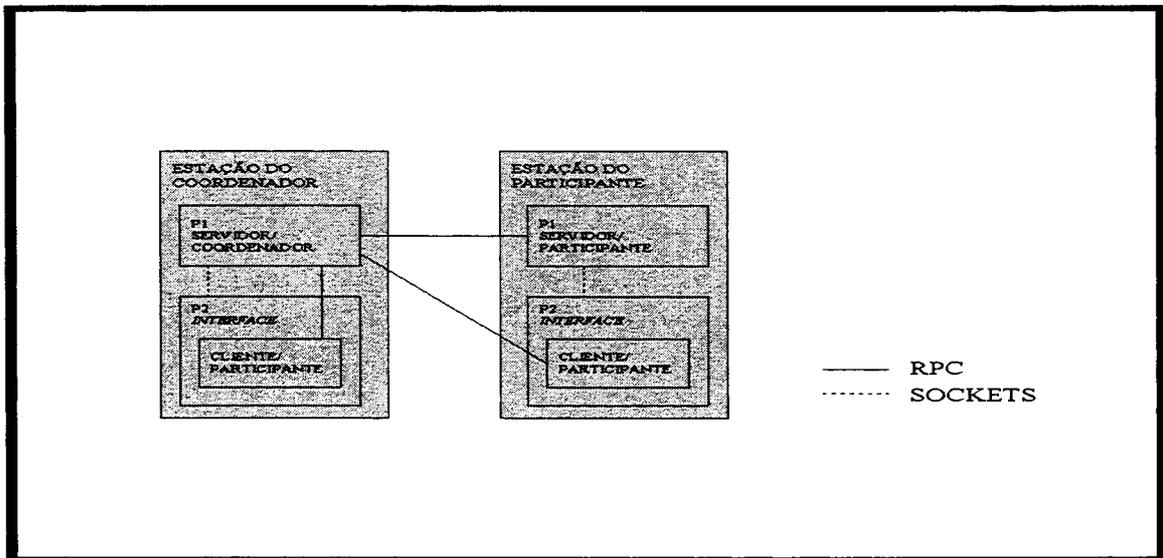


FIGURA 4.10 - Os Submódulos de Comunicação do SADGV2

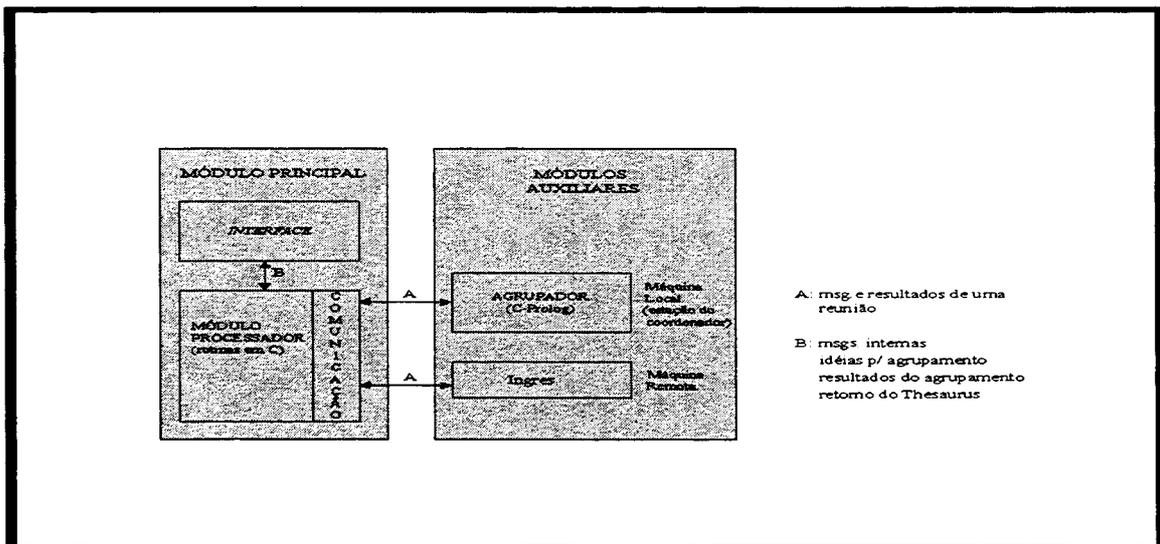


FIGURA 4.11 - Módulos Principais do SADGV2

A figura 4.11 ilustra os módulos principais da ferramenta, com destaque para o módulo agrupador, que é o responsável pelo agrupamento das idéias geradas na fase "*brainstorming*", sendo também responsável pela manutenção de um dicionário de sinônimos.

A seguir, analisa-se o algoritmo de agrupamento implementado pela ferramenta, módulo este que considera-se de extremo interesse e de maior contribuição para o Sistema que se pretende propor, devido às suas características que serão a partir deste ponto comentadas.

4.1.2.3.1 O Algoritmo de Agrupamento

O algoritmo de agrupamento de idéias similares foi implementado na linguagem C-Prolog [YOU 89, PER 87] em vista dos seguintes motivos relacionados por Souza [SOU 94]:

- Facilidades oferecidas pelo C-Prolog para o processamento de linguagem natural;
- Existência do sistema de Freitas [FRE 93], implementado em C-Prolog, que possui diversos predicados utilizáveis pelo SADGV2; por exemplo, os predicados responsáveis pela manutenção do dicionário de palavras, pela análise léxica e sintática de frases na Língua Portuguesa;
- Relativa facilidade de comunicação entre processos em C-Prolog e a linguagem C.

Os passos descritos no algoritmo de agrupamento são os seguintes:

- Passo 1: As mensagens são decompostas em frases, que por sua vez, são decompostas em palavras;



FIGURA 4.12 - O primeiro passo do algoritmo de agrupamento

- Passo 2: As palavras de função ou negativas são retiradas, obtendo-se como resultado somente as palavras de conteúdo, que possuem maior significado;
- Passo 3: Aplicação da técnica de "Stemming" [AIK 92], onde somente o radical da palavra de conteúdo permanece, sendo o sufixo retirado da palavra. O objetivo da aplicação desta técnica é o de reduzir o número de formas morfológicas em que a palavra possa ocorrer;
- Passo 4: Utilização de um dicionário de sinônimos - "Thesaurus" - para o agrupamento de idéias similares escritas com palavras diferentes (graficamente) mas de igual significado;

- Passo 5: Geração da Tabela de PCS (Percentual do Coeficiente de Similaridade), que representa o número de palavras em comum entre duas frases, sendo sua fórmula indicada por $PCS = (CS * 100) / TPF$, onde CS é o Coeficiente de Similaridade e TPF é o Total de Palavras da Frase;

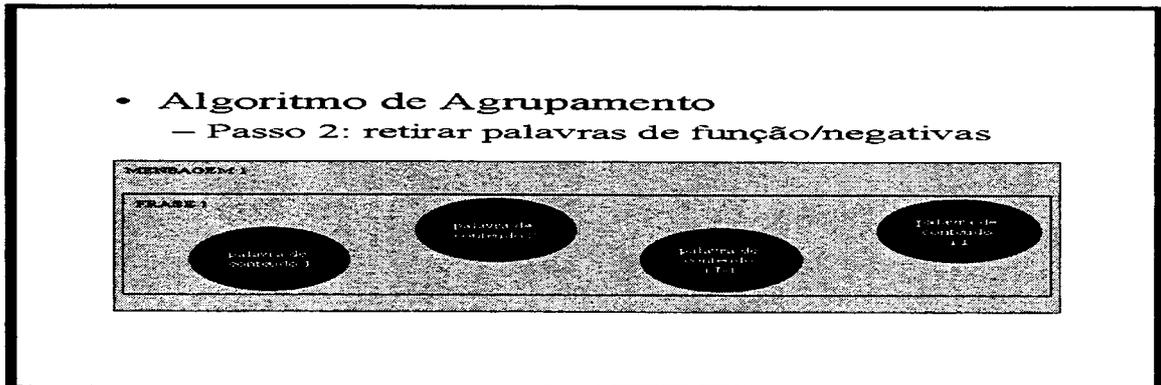


FIGURA 4.13 - O segundo passo do algoritmo de agrupamento



FIGURA 4.14 - O terceiro e o quarto passos do algoritmo de agrupamento

- Passo 6: Ignora relações onde o PCS é menor do que a percentagem mínima de palavras em comum definida pelo usuário;
- Passo 7: Gera as relações: (mensagem a, frase b) "está contida em" (mensagem c, frase d);
- Passo 8: Gera as relações: [lista de frases] "está contida em" (mensagem c, frase d);
- Passo 9: Gera o resultado final composto pelas seguintes informações: frequência da idéia, número da mensagem, número da frase, palavras de conteúdo, radicais das palavras de conteúdo e números das mensagens similares;

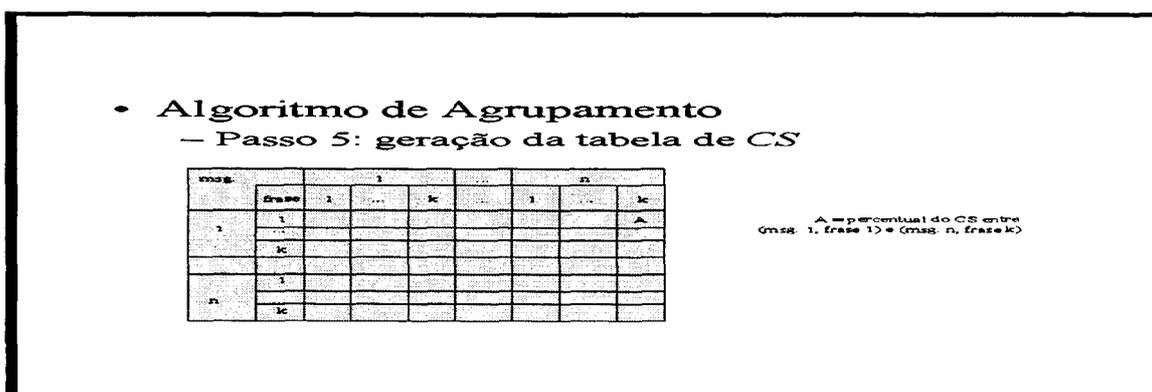


FIGURA 4.15 - O quinto passo do algoritmo de agrupamento

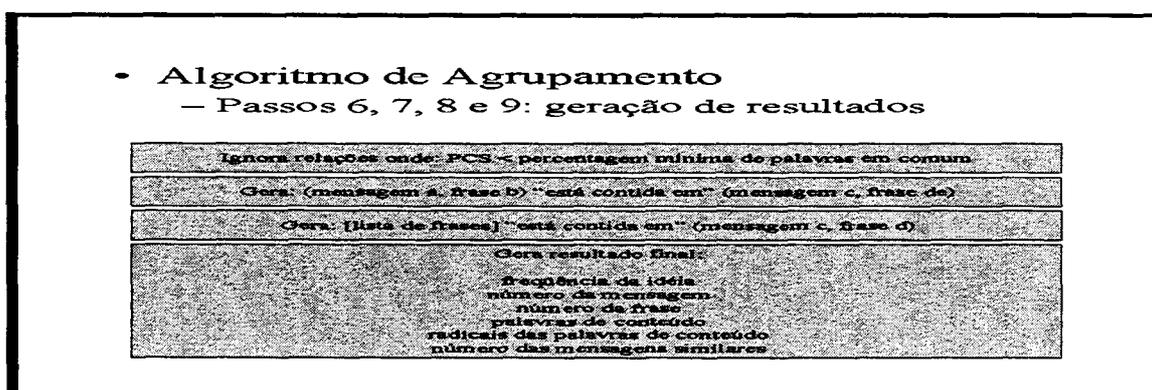


FIGURA 4.16 - Resultado final do algoritmo de agrupamento

4.1.2.4 Avaliação do Sistema SADGV2

Comparativamente ao sistema inspecionado no item 4.1.1, o sistema de Souza [SOU 94] apresenta uma boa documentação, entretanto, não foi possível reconstruir o sistema e executá-lo a partir de sua documentação, como no caso anterior. Isto porque somente foram encontrados fragmentos de código fonte do sistema (em linguagem C e em C-Prolog). Devido ao estado deste material, somente conseguiu-se executar alguns dos módulos principais, mas o funcionamento restringiu-se a telas de exibição, haja visto que faltavam muitos dos módulos que eram chamados externamente. Outro fator que impediu a execução do sistema foi que as bibliotecas de visualização gráfica do ambiente SUN e as bibliotecas de chamadas RPC não estavam disponíveis ou estavam em versões diferentes daquelas utilizadas quando dos testes do sistema SADGV2. Esta foi uma das falhas que encontrou-se no sistema: alta dependência de uma plataforma específica de *hardware* e *software*, conforme observa-se ao longo das descrições expostas no item 4.1.2..

Este foi um fato importante para a proposta de um novo sistema: não depender exclusivamente de uma plataforma de *hardware* ou *software*. Outro ponto desfavorável é que o sistema não apresenta uma *interface* WWW, seja para os usuários seja para o administrador do sistema, o que impede que este seja executado ou tenha seus dados exportados para um *Web browser* ou outro tipo de ferramenta, como um *chat*, por exemplo. Esta foi uma das características que fez com que a implementação deste sistema tenha-se tornado tão complexa (a falta de utilização e integração com *web browsers*), de acordo com o que foi apresentado por Souza [SOU 94] em seus relatos.

Justamente por não obter-se a possibilidade de executar em uma situação real o SADGV2, não pôde-se observar seu desempenho prático. Entretanto, embasado nos trabalhos de Magalhães [MAG 94], com seu módulo em C-Prolog, prevê-se que, sendo o sistema de Souza [SOU 94] desenvolvido nos mesmos moldes (acrescido da capacidade de "aprender" novas palavras), deva-se ter um desempenho razoável para pequenos grupos de frases, o que não deve acontecer quando o número destas cresce, como em uma conversação livre em um *chat*, que é o objeto de estudo da nova proposta que será apresentada.

O suporte a um SGBD (Sistema Gerenciador de Banco de Dados) via chamadas SQL foi um ponto favorável que avaliou-se em favor do sistema SADGV2, tendo inclusive imaginado-se a necessidade de um módulo deste tipo para uma nova proposta, justamente pela forma de implementação apresentada no *software* de Souza [SOU 94]. Entretanto, com o andamento dos novos trabalhos, chegou-se à conclusão de que este módulo era dispensável, como será justificado adiante.

Os estudos apresentados pelo módulo de agrupamento de frases proposto por Souza [SOU 94] e que aproveita-se dos trabalhos de Freitas [FRE 93] foram assimilados como importantes para uma nova ferramenta. Entre os tópicos que justificam esta escolha, encontram-se:

- A possibilidade de o sistema "aprender" novas palavras para o seu dicionário, que, pelos motivos acima explicitados, não foram testados na prática;
- A retirada de palavras de pouca importância das frases para melhorar a qualidade das informações agrupadas, que podem ser chamadas *StopWords*, de acordo com outros trabalhos, como o de Wives [WIV 99];
- A maneira com que foram aproveitados os trabalhos de Freitas [FRE 93], tornaram justificável a procura por um módulo que trata unicamente do agrupamento de frases. Este foi a maior contribuição de Souza [SOU 94] para a proposta de um novo modelo.

Desta forma, observa-se que o sistema de Souza [SOU 94] apresentou grandes avanços para a sua época de implementação, principalmente nos aspectos relativos ao aproveitamento de código de terceiros (neste caso, os trabalhos de Freitas [FRE 93], que vieram a conduzir as novas propostas) e na implementação de módulos de comunicação, que eram de difícil implementação quando da sua construção e tipo de constituição.

Por estes motivos, considerou-se as propostas teóricas do módulo de agrupamento deste sistema e buscou-se novas alternativas a seu modo de execução e *interface* com dados e usuário, conforme será descrito nos próximos itens.

4.1.3 Um estudo sobre Agrupamento de Documentos Textuais em Processamento de Informações não Estruturadas Usando Técnicas de "*Clustering*"

Adiantando-se aos comentários que serão feitos ao longo das descrições seguintes, considera-se que o trabalho defendido por Wives [WIV 99], foi de suma importância para a proposta que será feita por este estudo. Em verdade, utilizando-se do mesmo "modo de trabalho" proposto por Souza [SOU 94] ao beneficiar-se dos trabalhos de Freitas [FRE 93], buscou-se por uma ferramenta que fizesse toda a tarefa de agrupamento das frases dos usuários, tarefa esta que considera-se intermediária, mas de elevada importância para a implementação de uma proposta de suporte à solução

cooperativa de problemas como a que pretende-se trabalhar. Portanto, no decorrer deste item 4.1.3, discorre-se sobre os trabalhos de Wives [WIV 99] de forma detalhada e enfocando-se também as contribuições que lhe foram dadas - a criação de novas classes de *StopWords* - e as que foram obtidas do estudo do módulo de agrupamento de informações, implementado através da ferramenta *Eureka*.

No início deste capítulo 4, mostra-se que os trabalhos de Barua [BAR 99] referem-se ao termo "*Infoglut*" para dizer respeito ao excesso de informações a que se está exposto ao trabalhar-se com mídias como a Internet, por exemplo. Wives [WIV 99], inicia seu estudo deste ponto, propondo que quando uma pessoa tem facilidade de acesso a grande quantidade de informações, também deve possuir meios que indiquem quais caminhos deve tomar para obter a informação de que necessita. Em seus estudos anteriores Wives [WIV 97], mostra que os meios de acesso a SRI (Sistemas de Recuperação de Informações), quando existem, não conseguem indicar de forma eficiente o local onde podem ser encontradas as informações desejadas pelo usuário, isto porque as formas de acesso às informações disponibilizadas por estes sistemas não são as ideais.

Com isto, muitos documentos irrelevantes (de assuntos correlatos ou não ligados diretamente) são retornados. Desta forma, as pessoas (usuários) recebem uma grande quantidade de informações, geralmente desordenadas e sem muito sentido, que devem ser analisadas exaustivamente, até que alguma coisa relevante seja realmente encontrada. Esta grande quantidade de informações, além de dificultar e tornar o processo de refinamento de dados demorado, pode levar o usuário a um problema conhecido como sobrecarga de informações (*information overload*). A sobrecarga de informações ocorre quando recebe-se uma quantidade muito grande de informações, mesmo que relevantes, e não consegue-se tratá-la ou assimilá-la [CHE 96].

O objetivo primário, do trabalho de Wives [WIV 99], é estudar técnicas de agrupamento de objetos (informações) textuais, proporcionando uma forma de acesso facilitado à informação. Deste modo é possível particionar uma grande coleção de documentos, isolando aqueles pertencentes a um mesmo assunto e facilitando a identificação de documentos relevantes para o usuário. Com isso, aplicando-se técnicas adicionais, é possível identificar o assunto ou conhecimento específico de cada grupo, facilitando o processo de recuperação de informações ou descoberta de conhecimento.

Um dos assuntos, pouco explorados por Souza [SOU 94] em seus trabalhos, foi o funcionamento do algoritmo de agrupamento textual por ele utilizado. No estudo de Wives [WIV 99], entretanto, este assunto é bastante explicitado. Ao considerar-se estes conceitos essenciais como contribuição a este novo trabalho, discorre-se com detalhes sobre este tópico no próximo sub-item.

4.1.3.1 O processo e as técnicas de agrupamento de informações

A tarefa de agrupar objetos, também conhecida por *clustering* não é recente. O conceito de aglomerados (*cluster*) é tão antigo quanto as bibliotecas. Muitos anos antes da criação dos primeiros computadores, já utilizava-se este processo manualmente, pois agrupar objetos similares facilita a localização de informações. Poucos anos depois que os computadores passaram a ser utilizados por órgãos governamentais americanos, os primeiros algoritmos de agrupamento de objetos auxiliado por computador surgiram e foram implementados. Com isso, atualmente, existem muitos algoritmos de agrupamento de objetos.

No contexto de informações textuais, existem basicamente dois tipos de agrupamento: um para agrupar termos (palavras) e outro para agrupar documentos. No primeiro caso, grupos de termos similares são identificados a fim de construir dicionários de palavras (denominados *Thesaurus*, conforme visto anteriormente em [SOU 94]) que definam um mesmo assunto. Assim, estas palavras similares podem ser utilizadas em ferramentas de recuperação de informações a fim de expandir a consulta do usuário com o intuito de obter melhores resultados (documentos mais relevantes).

No escopo dos trabalhos de Wives [WIV 99] e que vêm de encontro aos objetivos de utilização para uma nova proposta, a técnica de agrupamento consiste em organizar uma série desorganizada de objetos em grupos de objetos similares. Este tipo de técnica é recomendada quando não há uma discriminação prévia de classes, sendo útil em casos onde não há a possibilidade de alocar um especialista na tarefa de separação de objetos em classes. Em muitos casos, pode não ser possível utilizar um especialista humano, por não existir na instituição alguém com conhecimento adequado sobre o domínio das informações em questão. É justamente esta a solução procurada como uma proposta mais atual e de utilização mais prática em lugar dos trabalhos de Freitas e Magalhães [FRE 93, MAG 94].

4.1.3.1.1 Agrupamento de informações textuais

O objetivo do agrupamento de informações textuais é separar uma série de documentos dispostos de forma desorganizada em um conjunto de grupos que contenham documentos de assuntos similares, conforme ilustra a figura 4.17.

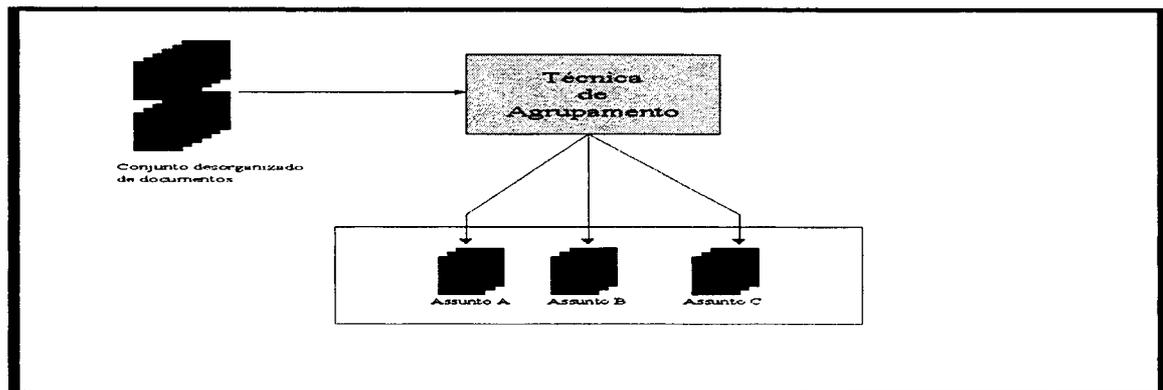


FIGURA 4.17 - Objetivo do agrupamento de informações textuais

Um dos problemas observados por Wives [WIV 99] é que a maioria das técnicas de agrupamento foram desenvolvidas para atuar sobre dados estruturados, ou seja, aqueles armazenados em SGBDs, considerados mais fáceis de serem tratados por meios computacionais e, somente há pouco tempo, principalmente com o surgimento da Internet, os chamados dados não estruturados, formados por imagens, textos e gráficos, ganharam maior atenção.

Outro ponto importante destacado é que cada algoritmo de agrupamento possui um ou mais parâmetros próprios que são difíceis de sintonizar corretamente. Logo, pelo fato destes parâmetros serem muito dependentes do algoritmo ou do domínio dos objetos utilizados, ainda não há um padrão estabelecido e torna-se muito difícil definir um

procedimento geral de agrupamento automático de informações. Este fato foi observado através de experimentos práticos com a ferramenta *Eureka*, como será discutido neste trabalho.

4.1.3.1.2 Tipos de agrupamentos textuais

O agrupamento pode ser classificado de acordo com a forma com que os grupos são constituídos e em relação à complexidade do tempo de execução do algoritmo.

Quanto à forma, existem dois tipos de agrupamento:

- Agrupamento por partição: os objetos são distribuídos em classes distintas, não havendo relação direta entre as classes. É denominado agrupamento de partição total (*flat partition*) e os documentos são separados exaustivamente e colocados em grupos totalmente diferentes;
- Partição hierárquica (*hierarchic partition*): o processo de identificação dos grupos é geralmente realimentado recursivamente, utilizando-se tanto objetos quanto grupos já identificados previamente como entrada para o processamento, constituindo-se desta forma, uma hierarquia de grupos de objetos, semelhantes a uma árvore.

Quanto à complexidade em relação ao tempo de processamento, obtém-se:

- Algoritmos de tempo constante: tem o objetivo de limitar o tempo máximo de processamento, ou ao menos descobrir o tempo gasto por cada elemento e estimar um tempo ou número máximo de comparações a ser feito, geralmente limitando o tamanho ou o número de *clusters*;
- Algoritmos de tempo linear: possuem a característica de aumentar seu tempo de processamento de acordo com o número de elementos, tempo este que, ao aumentar, cresce de modo sutil e de forma gradativa (não exponencial), já que nem todos os elementos necessitam comparações exaustivas;
- Algoritmos exponenciais de ordem quadrática: crescem exponencialmente, pois sempre que um elemento é adicionado ele deve ser comparado com todos os outros, aumentando a quantidade de comparações entre os elementos.

4.1.3.1.3 Técnicas de agrupamento

Na maioria das técnicas de agrupamento existentes são utilizados algoritmos que realizam três etapas básicas e distintas, conforme mostra a figura 4.18:

- Identificação e seleção de características: identifica palavras nos documentos e após seleciona (globalmente ou localmente) aquelas que possuem maior grau de discriminação (que caracterizam melhor o objeto). Como resultado desta etapa, são geradas listas de palavras (características) relevantes, que identificam cada documento;
- Cálculo de similaridade: identifica os graus de similaridade entre os objetos (documentos), utilizando para isto as listas de características identificadas na etapa anterior. Como resultado, obtém-se uma matriz que contém os valores de similaridade entre todos os objetos. Quanto mais características em comum

possuírem os objetos, mais similares eles são. Observa-se que esta matriz possui muitas semelhanças com a matriz obtida no trabalho de Souza [SOU 94];

- Identificação de aglomerados (*clusters*): consiste em identificar correlações entre os elementos da matriz, de acordo com as restrições impostas por cada algoritmo. Dependendo da técnica utilizada para a montagem da matriz e do algoritmo de agrupamento, todos os elementos devem ser comparados com todos, o que pode classificar o algoritmo como sendo de ordem quadrática. Ao final desta etapa, obtém-se os grupos e seus respectivos elementos.

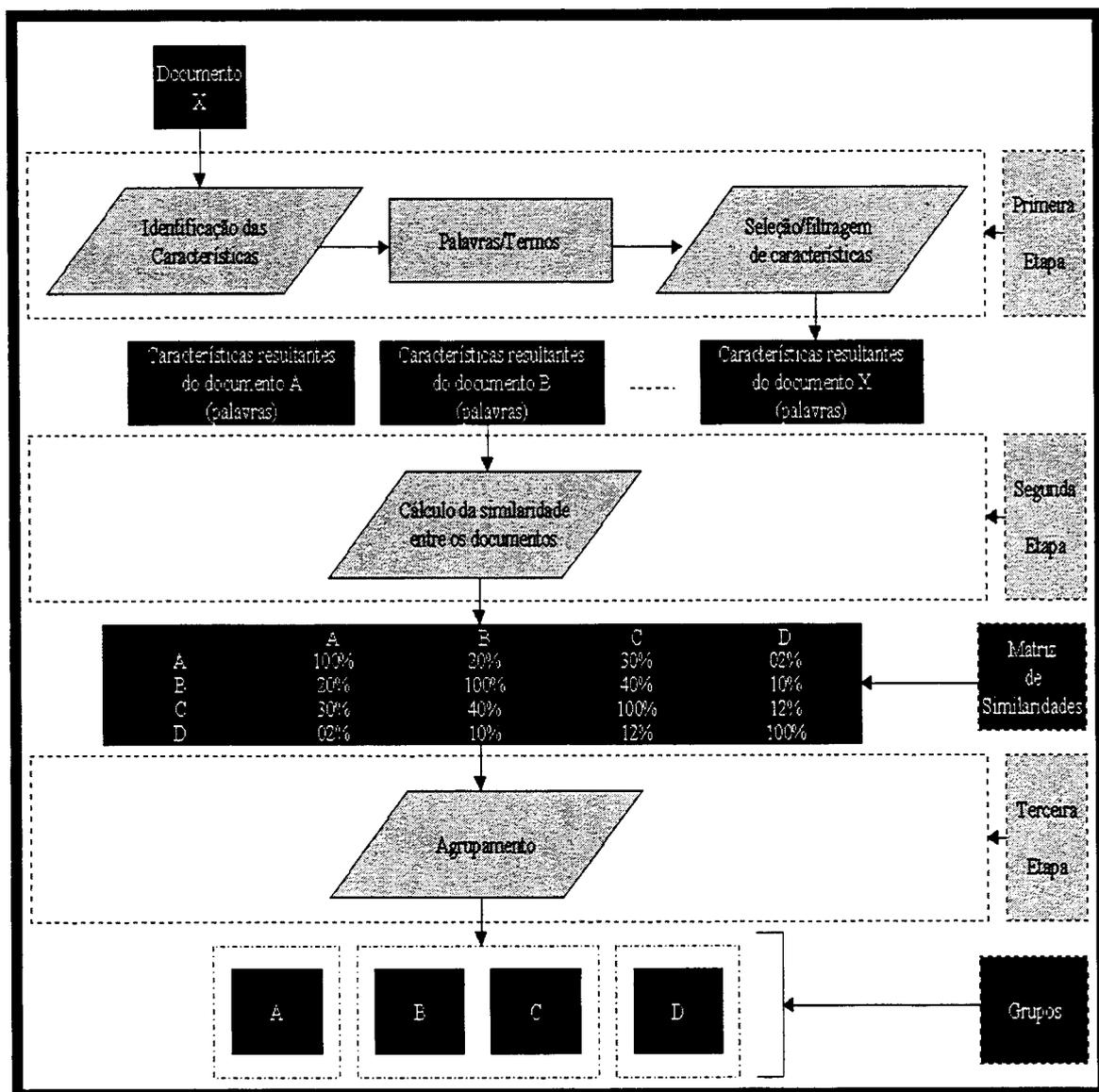


FIGURA 4.18 - Etapas do processo de agrupamento

Para detalhar as etapas mostradas na figura 4.18, Wives [WIV 99] inicia reafirmando que nas técnicas de agrupamento tradicionais, os objetos em questão já possuem atributos (características) bem definidos (em um banco de dados relacional tradicional, os campos das entidades indicam as características).

Nos textos, porém, Wives [WIV 99] considera que não é fácil identificar um atributo (campo), simplesmente porque não há algum sinal ou local predeterminado que indique os atributos. Pode ocorrer, ainda, o fato de um atributo ser identificado em um documento e nem ao menos existir em outro (os textos não são obrigados a conter as mesmas palavras). Em decorrência deste fato, as características marcantes dos objetos podem ser variadas e ambíguas (o problema do vocabulário, que foi comentado tanto por Wives [WIV 99] quanto por Souza [SOU 94]).

É necessário em primeiro lugar, estabelecer um método que identifique as características mais marcantes de cada objeto (texto). Em uma informação deste tipo, as palavras é que caracterizam-na. Portanto, nas técnicas de agrupamento de informações textuais as palavras são utilizadas como atributos. Porém nem todas as palavras podem ser utilizadas. É o caso de palavras inerentes à linguagem, que auxiliam na construção das orações e não precisam ser incluídas no processo (preposições, conjunções e artigos, por exemplo). Estas, e algumas outras palavras, geralmente possuem baixo valor de discriminação, pois aparecem em vários elementos (quase todos) e não ajudam a distinguir o conteúdo dos textos. Palavras deste tipo são chamadas de palavras negativas (*StopWords*), e são retiradas em uma etapa de pré-processamento, conhecida como remoção de palavras negativas.

Como comentário, que considera-se relevante, no sub-item 4.1.2.4 deste estudo, ao comentar-se os trabalhos de Souza [SOU 94], referencia-se o mesmo processo, ou seja, os dois trabalhos preocupam-se com esta classe específica de palavras. Em testes práticos com a ferramenta *Eurekha*, notou-se que quanto maior o conjunto de classes e de elementos nestas classes, melhor o resultado do processo de agrupamento, além de obter-se uma redução substancial do tempo de processamento do *software*.

Como contribuição ao trabalho de Wives [WIV 99], criou-se, como parte deste trabalho de pesquisa, várias classes de *StopWords*, conforme referências extraídas de Kaspary e Cunha [KAS 97, CUN 85], que lhe foram posteriormente enviadas para possível utilização em seus trabalhos futuros.

Em um documento textual, nem todas as palavras aparecem com a mesma frequência, o que torna possível selecionar aquelas palavras (atributos ou características) que são mais marcantes. Isto torna-se necessário porque a utilização de todas as palavras do documento torna o processo de agrupamento muito mais demorado. Em aplicações de agrupamento reais e práticas (que é justamente o tipo de situação em que encontra-se a nova ferramenta a ser proposta) os usuários necessitam de um tempo de resposta muito curto, sendo que o maior gargalo para as técnicas de agrupamento de informações está justamente na etapa de seleção de características, que tem como objetivo diminuir o número de características a serem processadas.

Segundo Wives [WIV 99], a técnica mais comum de identificação de atributos (palavras) marcantes é a frequência relativa, que indica o quanto determinada palavra é importante para um documento de acordo com o número de ocorrências desta palavra no documento.

O resultado das técnicas de agrupamento não é muito influenciado pela função de identificação escolhida, mas sim, pelo algoritmo de agrupamento utilizado. Além disso, fórmulas muito complexas podem ser muito demoradas quando realizam o processo de análise. Deste modo, o ganho que poderiam obter com uma análise mais acurada acaba não sendo vantajoso, pois o agrupamento torna-se muito mais demorado e a diferença nos resultados pode não ser tão significativa.

Após terem sido identificadas as características dos objetos, é possível aplicar-se uma técnica de seleção ou detecção de características importantes. Estas técnicas muitas vezes, são independentes da função de cálculo de frequência escolhida. Uma destas técnicas propõem a análise sintática dos documentos a fim de identificar semanticamente (ou morfológicamente) os atributos (palavras) mais importantes de um documento, onde seleciona-se somente os substantivos das orações. A não ser que já possua-se uma base de conhecimentos sintáticos e semânticos, aliados a uma ferramenta para tal fim, o processo pode ser muito trabalhoso.

Adaptar o vocabulário, através de uma adaptação de termos (a exemplo do que propõem Souza [SOU 94], com a utilização de técnicas de "*Stemming*"), unificando-o, facilita a identificação de documentos similares. Isto porque existem muitas palavras de grafia diferente que possuem o mesmo significado (sinônimos). Neste caso, trocar as palavras de mesmo significado por uma única facilitaria o processo de agrupamento. Do mesmo modo, é interessante eliminar as diferentes variações morfológicas de uma palavra (singular ou plural, por exemplo), utilizando somente o seu radical, porém, algumas vezes, pode tornar um texto muito abrangente, já que muitas palavras tornar-se-ão parecidas, além de que criar ferramentas para tal processamento é demasiadamente caro, já que são necessários dicionários (morfológicos e de sinônimos) específicos da linguagem utilizada nos documentos. Desta forma, segundo Wives [WIV 99], nem todos os trabalhos com este escopo utilizam-nas, pois, na maioria das vezes, assume-se a premissa de que documentos similares possuem termos comuns e, mesmo que apareçam poucas vezes, haverão termos em comum por se tratarem de textos baseados em um mesmo contexto.

A segunda etapa do processo de agrupamento consiste em analisar todos os objetos como objetivo de identificar semelhanças entre eles. O grau de semelhança entre dois objetos é dado, em geral, por uma fórmula ou função de similaridade. Esta fórmula analisa todas as características (palavras) semelhantes que os objetos possuem e retorna um valor, um grau, indicando o quanto estes objetos são similares. As principais funções, que foram utilizadas são medidas de distância e medidas difusas (*fuzzy*).

Após concluída a etapa de cálculo de similaridades (ou distância) entre os objetos, obtém-se uma matriz que indica os valores de similaridade entre os objetos. A partir desta tabela, é possível identificar os grupos de objetos, especificando algum tipo de regra de relacionamento (ou algoritmo de agrupamento). Os que apresentaram-se disponíveis na versão da ferramenta *Eurekha*, que foi disponibilizada para testes, foram os seguintes:

- *Cliques*: dentre os algoritmos apresentados, foi considerado o que oferece um melhor resultado, pois gera grupos de elementos muito coesos (se o grau de similaridade mínimo for elevado). Isto porque para cada elemento adicionado ao grupo, verifica-se seu grau de similaridade com todos os outros elementos do mesmo grupo. Caso este elemento não tenha grau de similaridade maior do que o mínimo, com um dos objetos, ele não é adicionado ao grupo. Porém, devido a esta verificação, este algoritmo torna-se mais demorado;
- *Stars*: oferece boa coesão e é muito rápido, já que não necessita realizar análises complementares como o algoritmo anterior. Apresenta dois problemas. O primeiro diz respeito ao fato dos objetos serem atribuídos ao primeiro grupo cujo grau de similaridade seja maior do que o mínimo exigido. Já que os objetos atribuídos a um grupo não são mais processados, não há garantia de que um objeto tenha sido colocado no grupo de maior similaridade. O segundo problema

diz respeito ao fato do processo não indicar todos os grupos a que o objeto poderia fazer parte;

- *Best Star*: foi desenvolvido com o intuito de solucionar o primeiro problema do algoritmo *Stars*. Neste caso, os objetos identificados como sendo mais similares a determinado grupo, mesmo que já agrupados, são reorganizados e atribuídos ao grupo de maior afinidade;
- *Full Stars*: atribui os documentos a todos os grupos cujo grau de similaridade ultrapassa o valor mínimo estabelecido. Seu funcionamento é idêntico ao do algoritmo *stars*, porém, com a diferença de que os documentos já atribuídos continuam no processo, não sendo ignorados.

4.1.3.2 O Sistema "Eureka" Implementado por Wives [WIV 99]

Todas as etapas descritas (seleção das características, cálculo de similaridade e agrupamento) foram modeladas e refinadas por Wives [WIV 99], que, com isso, conseguiu identificar alguns problemas e propor alternativas como solução.

Sua técnica de agrupamento consiste nas seguintes etapas:

- Identificação de palavras;
- Remoção de palavras negativas (*StopWords*);
- Cálculo de frequência relativa;
- Remoção de palavras com frequência inferior a um limiar estabelecido pelo usuário. Esta frequência mínima pode ser ignorada pelo usuário;
- Utilização das x palavras mais frequentes ou método de truncagem, onde x é estabelecido pelo usuário, podendo ser todas as palavras;
- Utilização de uma fórmula de similaridade *fuzzy* definida especialmente para o processo;
- Aplicação de um algoritmo de agrupamento.

Uma das mais importantes contribuições de Wives [WIV 99] e que foram de crucial importância para a proposição do novo modelo, as modificações às considerações presumidas do modelo de Souza [SOU 94] e o descarte dos trabalhos de Magalhães [MAG 94] foi que em nenhum momento sua ferramenta *Eureka* realiza alguma etapa de pré-processamento de linguagem natural. Wives [WIV 99] considera que os documentos já devam ter sido corrigidos por um corretor ortográfico, considerando também interessante que os textos tenham sofrido algum tipo de processamento de normalização de termos. Entretanto, considera que estas etapas não são obrigatórias, já que não foi realizado nenhum estudo comprovando que sua utilização torna o processo mais eficiente. Estas considerações serviram de fundamentos para a proposição do novo modelo, como será visto no item 4.2.

4.1.3.3 Avaliação do Sistema "Eurekha"

Considerando-se que a ferramenta *Eurekha* é responsável pelo agrupamento das informações manipuladas pela nova ferramenta a ser proposta, considera-se importante ater-se a detalhes do seu funcionamento, que será avaliado neste sub-item.

O *software* foi analisado em testes práticos, realizados em ambiente Windows e computadores PC. A partir de um *log* (devidamente tratado) formado pela discussão através de uma ferramenta de *chat*, as frases resultantes são agrupadas de acordo com sua similaridade para posterior manipulação.

Várias versões foram entregues pelo desenvolvedor, que sempre mostrou-se disponível para a correção dos *bugs* encontrados e possíveis modificações requeridas para a importação/exportação de resultados entre as diferentes ferramentas utilizadas.

Conforme os relatos teóricos já mencionados, o comportamento dos algoritmos mostrou-se de acordo com o item 4.1.3.1.3., sendo desta forma o algoritmo *Cliques* escolhido entre os demais por suas características já citadas e por comportar-se melhor nos testes práticos de agrupamento.

Um fato importante que deve ser mencionado é quanto a criação de novas classes de *StopWords*, que tornaram o tempo de processamento do algoritmo reduzido de 30 minutos para 20 minutos e em seguida para aproximadamente 10 minutos, em um computador equipado com um processador Pentium de 200MHz, com 64 MB de memória RAM e 4GB de disco rígido. Considera-se que os testes foram realizados com um grupo de 189 frases, considerando-se 3 fases distintas, onde, na primeira, optou-se por utilizar todas as frases do arquivo de *log* da ferramenta de *chat* (que somavam 212 frases); na segunda, todas as frases sem o nome do usuário a que pertence a frase e na terceira, frases selecionadas (189 frases), sem o nome do usuário a que pertenciam. Percebe-se que, quanto menor o número de frases e maior a seleção destas, melhores os resultados de tempo e agrupamento das frases.

Quanto a seu funcionamento inicial, a ferramenta fornece uma *interface* simples e intuitiva, conforme demonstra a figura 4.19. Nesta tela inicial estão disponíveis todas as opções passíveis de utilização.

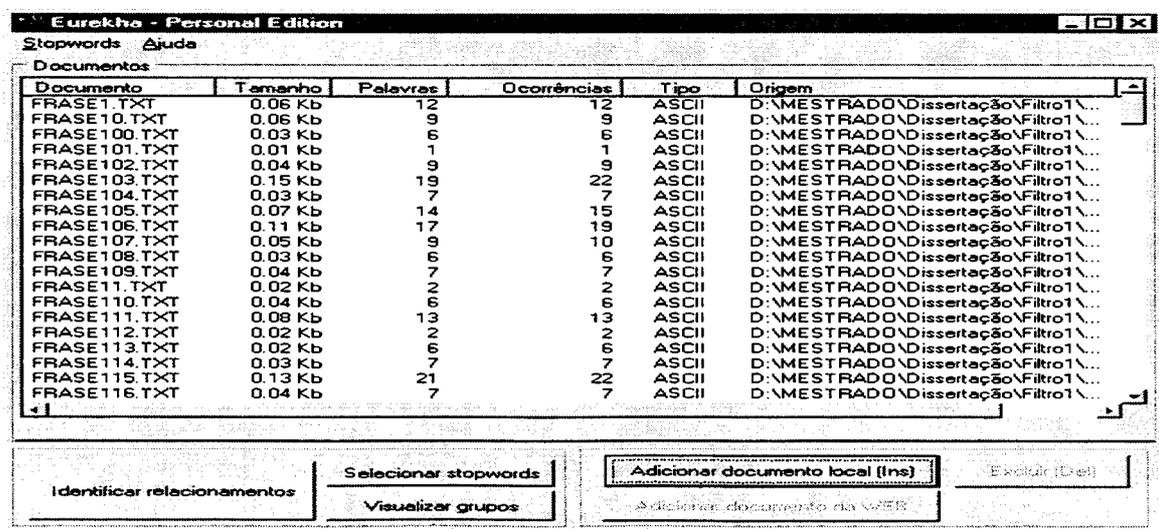


FIGURA 4.19 - A interface inicial da ferramenta *Eurekha*

O primeira opção que deve ser considerada é aquela apresentada pelo botão "Adicionar documento local (Ins)". É esta opção que permite adicionar os documentos texto que serão agrupados. A janela de seleção permite que se escolha por um diretório específico na árvore de diretórios, cujos arquivos são selecionados através do botão "Adicionar >>", e retirados da seleção através do botão "<< Remover". Observa-se que somente os arquivos que estiverem na janela da direita "Documentos Seleccionados" realmente farão parte do processo de agrupamento. Este procedimento é ilustrado pela figura 4.20.

A seguir, nesta ordem, deve-se selecionar as classes de *StopWords* (através do botão "Selecionar stopwords") que serão utilizadas para fins de agrupamento, conforme mostra a figura 4.21. Já a figura 4.22 mostra como definir e selecionar novas classes de *StopWords* através da opção "Stopwords".

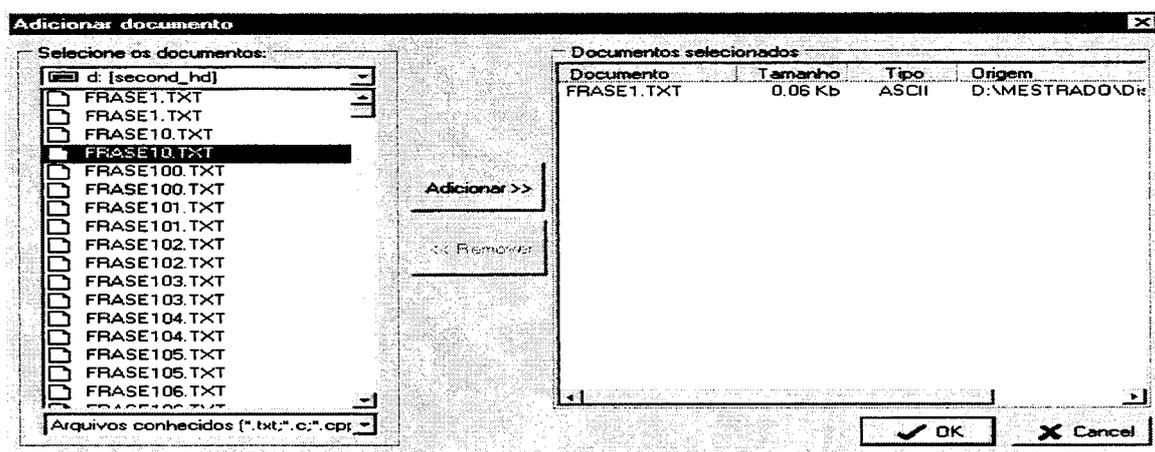


FIGURA 4.20 - A tela de seleção de documentos para agrupamento

O botão "Identificar relacionamentos" inicia o procedimento de construção da matriz de similaridades. Conforme mostra a figura 4.23, este processo somente é interrompido quando estiver terminado. É este o procedimento mais demorado da ferramenta, que, conforme dados mencionados, pode demorar entre 10 minutos e 30 minutos, de acordo com as informações de entrada. Deve-se salientar que, até esta opção, ainda não selecionou-se nenhum tipo de algoritmo de agrupamento. Este procedimento também identifica as palavras relevantes de cada frase, de acordo com o que foi comentado anteriormente. Após a identificação dos relacionamentos, é apresentada a tela da figura 4.24, com o tempo decorrido para o processo.

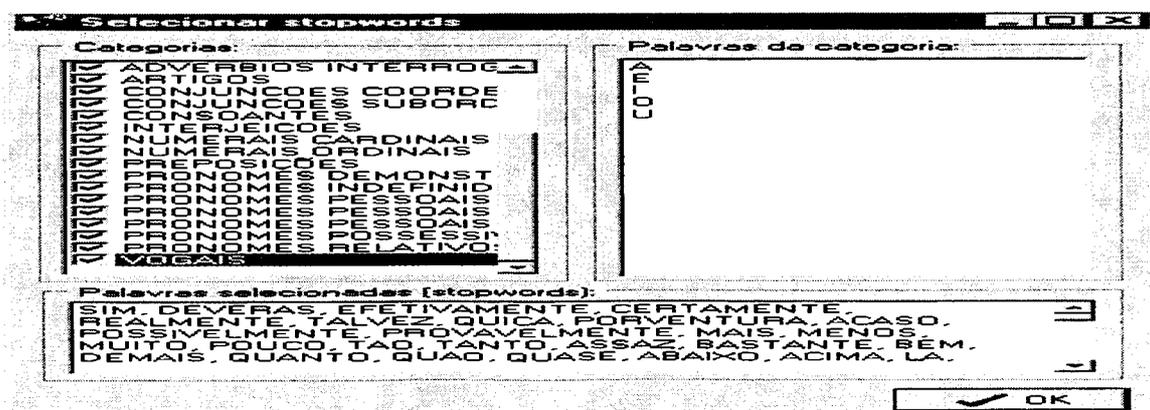


FIGURA 4.21 - A tela de seleção de *Stopwords*

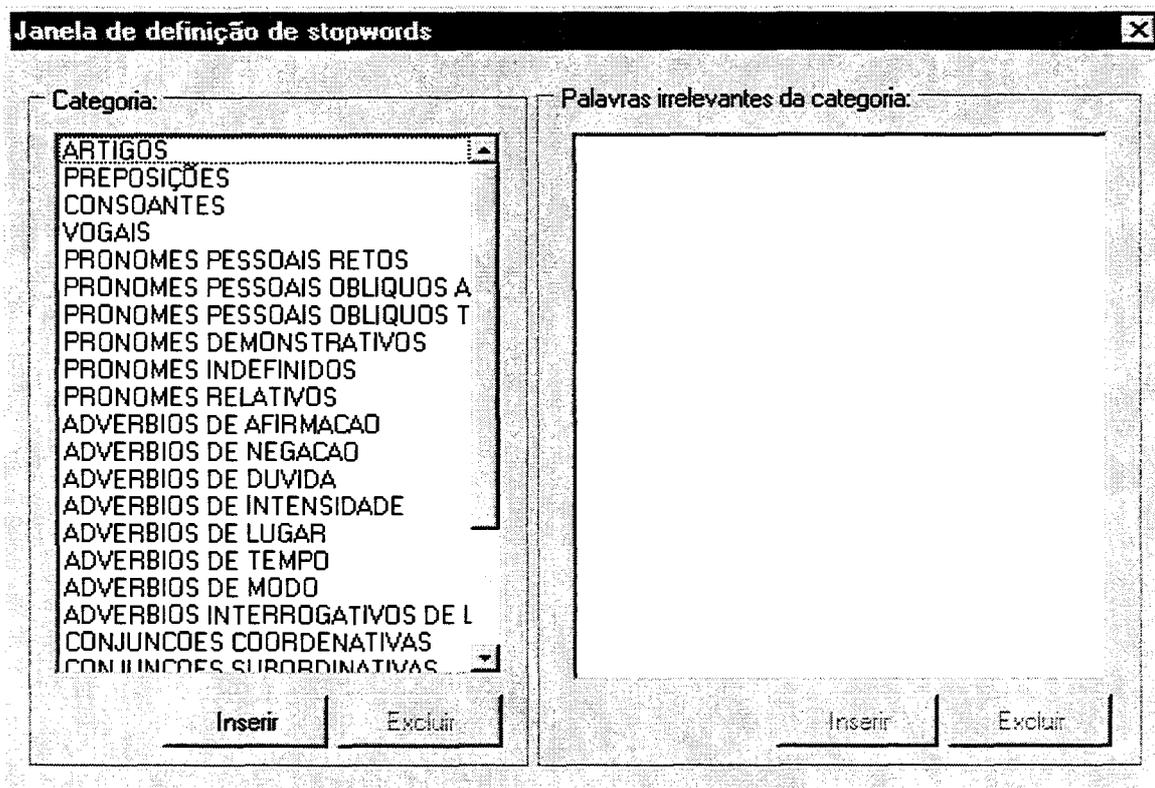


FIGURA 4.22 - A tela de seleção e definição de novas classes de *Stopwords*

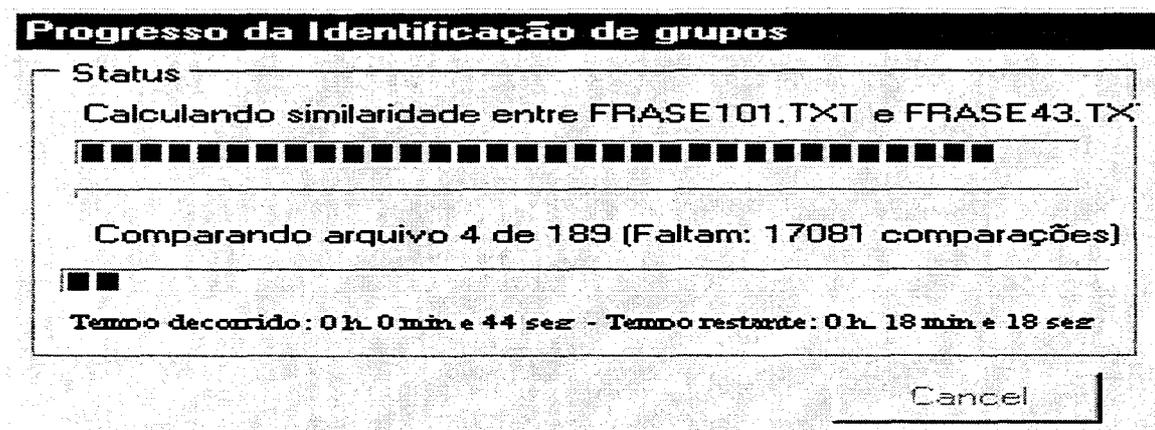


FIGURA 4.23 - A tela do processo de Progresso da Identificação de grupos

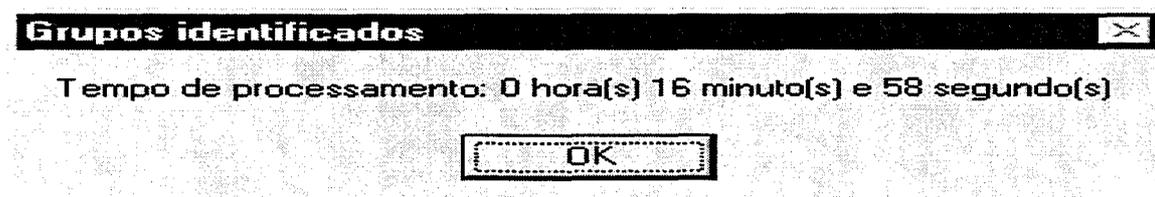


FIGURA 4.24 - A tela de grupos identificados

A seguir, seleciona-se o botão "Visualizar grupos", que é responsável pela identificação dos grupos gerados, através de um dos algoritmos mencionados ao final do item 4.1.4.1.3. Observa-se que o algoritmo "Cliques" apresentou melhor desempenho nos testes efetuados, mas dispõem-se das outras 3 opções, segundo a opção "Algoritmo".

Através da figura 4.25, observa-se na janela "Grupos" os *clusters* gerados durante o processo de agrupamento. Ao pressionar-se o sinal "+" de determinado grupo desta janela, obtém-se a constituição de tal grupo. O gráfico à direita mostra uma visualização gráfica dos *clusters* gerados.

A barra deslizante intitulada "Coeficiente de Similaridade" é a responsável pelo nível que o usuário irá requerer para a similaridade entre os documentos. Quanto mais próximo de 0%, maior será o tamanho dos grupos gerados, entretanto, menor será a similaridade entre os documentos que compõem tais grupos. Ao contrário, quanto mais próximo de 100%, menores serão os tamanhos dos *clusters* gerados e menor a quantidade de documentos destes.

De acordo com os experimentos práticos, a posição da barra de rolagem que formou grupos mais coesos ficou entre 15% e 25%. Ao selecionar-se valores fora destes limites, os grupos gerados não forneceram informações significativas. Entretanto, isto foi o que observou-se em casos específicos. O usuário deve estar familiarizado com a ferramenta para então poder escolher entre um ou outro algoritmo e também definir o grau de similaridade mais adequado para seu caso particular. O botão "Reagrupar" deve ser utilizado sempre que modificar-se o algoritmo de agrupamento ou o valor do coeficiente de similaridade.

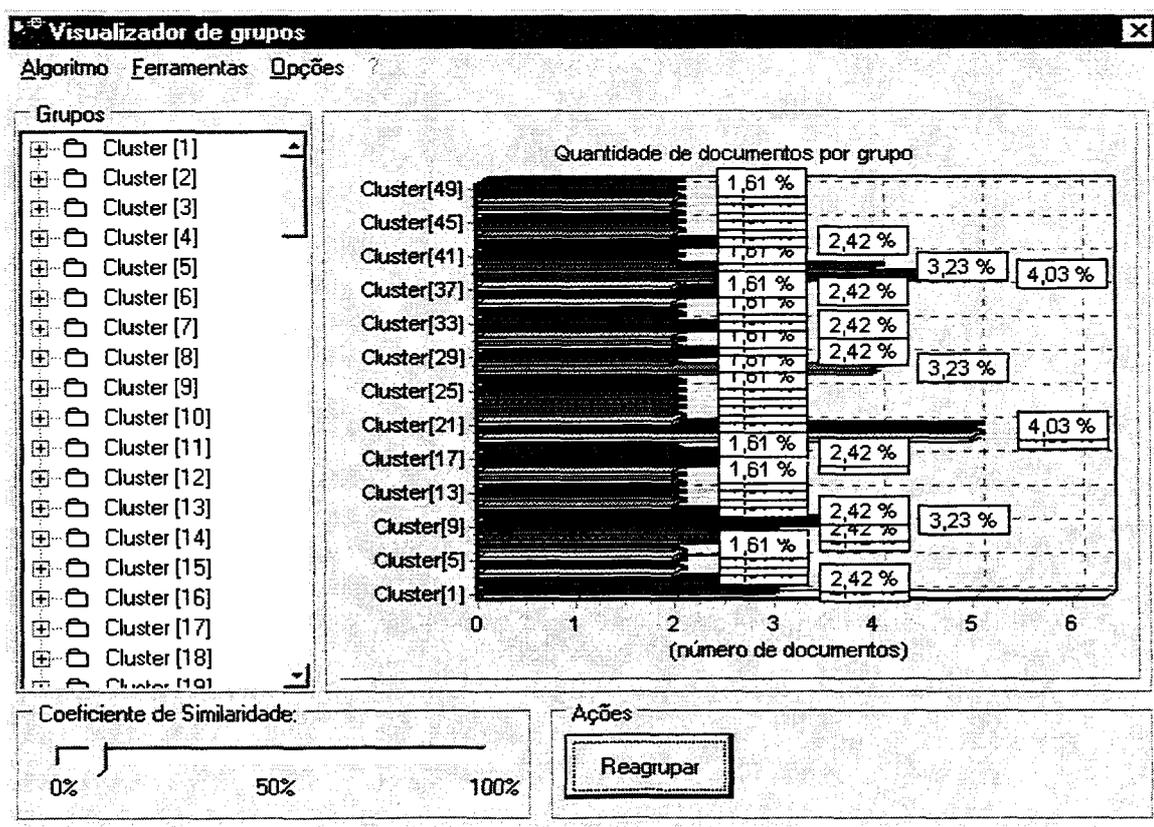


FIGURA 4.25 - A tela de visualização de grupos

O último passo a ser executado com este software é a opção "Ferramentas", onde deve-se escolher a opção "Exportar resultado do agrupamento...", que será responsável por gerar um arquivo texto chamado "*clusters.txt*" no diretório corrente da ferramenta *Eurekha*, arquivo este que será utilizado pelo protótipo proposto.

Como observa-se através deste tutorial, a intervenção do usuário para com o *software* de agrupamento é mínima e concebida da maneira mais simplificada possível, justamente para diminuir o tempo gasto em treinamento e possíveis erros de entrada de dados. O novo protótipo, que será apresentado a seguir, busca oferecer ao usuário as mesmas características de *interface* (tanto para as exigências de *hardware* como *software*), com o intuito de manter um padrão de compatibilidade de operação para o usuário final.

4.2 Descrição do Modelo do Sistema ESCOP

Pode-se interpretar e dividir o esquema proposto na Figura 4.26 - que será denominado ESCOP - segundo funções aproveitadas e funções novas. A partir do *log* gerado por ferramentas como *Palace* e *CU-SeeMe* em sessões de *chat*, o módulo denominado Filtro I separa tal documento (em formato texto puro) em arquivos de frases separadas onde, a primeira linha do arquivo de *log* denomina-se "*Frase1.txt*", a segunda "*Frase2.txt*" e assim, sucessivamente. Tal processo torna-se necessário devido às convenções de funcionamento dos módulos seguintes, onde, dentre as funções aproveitadas, destaca-se o módulo II (*Eurekha*) [VIT 99b].

Conforme observa-se, a função aproveitada, denominada módulo I (PLN), encontra-se fora de destaque em relação aos demais módulos do sistema. Considerando-se as avaliações do sub-item 4.1.1.4 e 4.1.3.2 deste estudo em relação à sua real necessidade no sistema proposto (necessidade esta que ainda não estava clara quando da proposta de VIT [VIT 99b]), tornou-se claro que, após experimentos práticos com os trabalhos de Magalhães e Wives [MAG 94, WIV 99], tal processo tornou-se dispensável, sendo então, após justificada, realizada sua retirada, como módulo constituinte, da nova proposta.

Para o módulo II, aproveitam-se, integralmente, os trabalhos desenvolvidos por Wives [WIV 99], onde, passados um conjunto de "*n*" frases, o *software* denominado *Eurekha*, realiza o agrupamento deste conjunto, de acordo com a similaridade entre as palavras que compõem cada frase. Neste trabalho, em nenhum momento realiza-se alguma etapa de pré-processamento de linguagem natural. Portanto, considera-se que os documentos já tenham sido corrigidos por um corretor ortográfico (tarefa que seria desempenhada pelo módulo I, mas, conforme observou-se, desnecessária). Além disso, é interessante que os textos tenham sofrido algum tipo de processamento de normalização de termos. Porém, estas etapas não são obrigatórias, segundo Wives [WIV 99].

A partir do módulo III, iniciam-se as chamadas funções novas. Neste ponto, a ferramenta *Eurekha* devolve como resultado de seu processamento as frases agrupadas na forma de *clusters* de frases similares, em um arquivo texto denominado "*clusters.txt*". A partir destes resultados, para cada *cluster* deve-se identificar a frase que contém a maior quantidade de palavras importantes, quando então, será determinada a "frase consolidadora", que sumariza as demais frases do *cluster*, que será então denominada Idéia. O símbolo "*" no modelo, impresso logo após a palavra Idéias, conforme a figura 4.26 é assim justificado.

O módulo IV é o responsável pela referência de uma frase a quem a formulou, considerando-se este módulo importante para fins de organização local por parte do coordenador das reuniões.

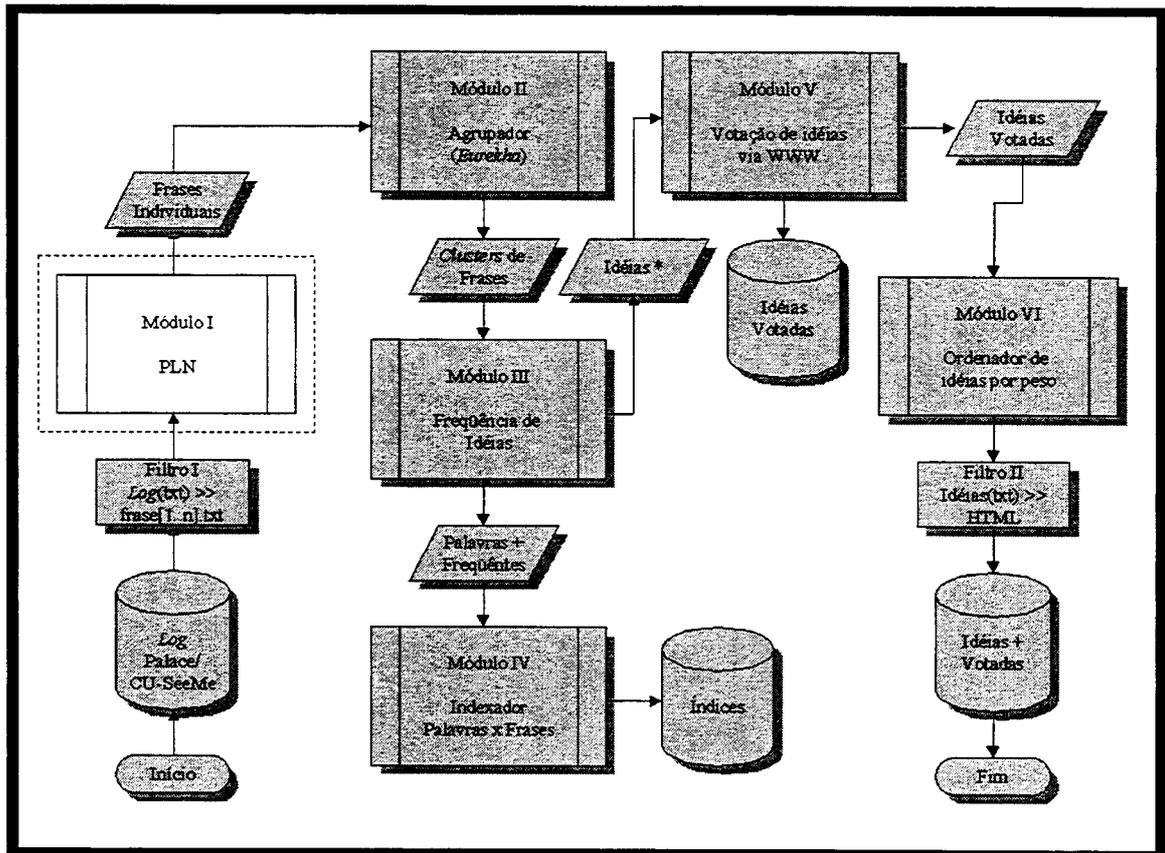


FIGURA 4.26 - O Modelo Proposto

O módulo V responde pela votação das idéias geradas nos módulos anteriores pelos participantes da reunião, através de um programa CGI ou compatível que é responsável pela contabilidade dos votos dos usuários. A partir deste ponto, o módulo VI, em conjunto com o Filtro II, tornam-se responsáveis pela edição de uma página HTML com os resultados desta fase da reunião virtual, vindos do módulo anterior.

Como caracterização do sistema, buscou-se obter uma ferramenta que venha a ser utilizada na prática de reuniões à distância, com seu escopo abrangendo áreas de EAD e trabalhos em grupo como Gerência de Redes de Computadores. Também desejava-se acrescentar um aumento de desempenho de processamento em relação às propostas anteriores, como por exemplo, a de Souza [SOU 94].

5 Descrição da Implementação

A partir do modelo proposto no capítulo anterior foi possível a implementação do Sistema ESCOP (Estratégia de suporte à Solução Cooperativa de Problemas), que é um protótipo que provê as principais funcionalidades do modelo proposto. A implementação do ESCOP possibilitou a verificação da viabilidade da utilização do serviço WWW como base para a interação de um ambiente CSCW, através da utilização de recursos de programação para Web e da integração de ferramentas já existentes na Internet e também para o agrupamento de informações textuais. Através do ESCOP foi possível também a realização da experimentação do modelo proposto.

Este capítulo apresenta os principais resultados da implementação do protótipo. Na seção 5.1 é apresentada uma visão geral do Sistema, através da explicação dos seus principais módulos. A seção 5.2 apresenta algumas considerações sobre o ambiente de implementação utilizado no desenvolvimento do protótipo. Na seção 5.3 é apresentada uma análise específica dos módulos componentes da ferramenta. Na seção 5.4, é explorada sua *Interface* com o usuário e, na seção 5.5 as considerações gerais a respeito da implementação.

5.1 Visão Geral do Sistema ESCOP

O ESCOP está dividido em sete módulos principais: Filtros I e II, Agrupador, Frequência de Idéias, Indexador, Votador e Ordenador conforme a representação da figura 5.1:

- Filtro I: recebe um arquivo de *log* de uma ferramenta qualquer de *chat* que apresente este recurso. Transforma cada linha deste arquivo em um novo arquivo texto, denominado "*frase[i].txt*", em um diretório especificado pelo coordenador da reunião (vide figura 5.2);
- Agrupador: recebe os arquivos gerados pelo Filtro I e agrupa-os de acordo com sua similaridade, através de parâmetros definidos pelo coordenador e, ao final, gera um arquivo para exportação de dados denominado "*clusters.txt*" (vide figura 5.3);
- Frequência de Idéias: recebe um arquivo de *clusters* do módulo agrupador e, através de um algoritmo específico, determina a "frase consolidadora ou sumarizadora" de cada *cluster*, que vem a ser a chamada "idéia" daquele grupo de frases, que pode ser determinada pela frase que contém a maior quantidade de palavras importantes daquele *cluster* (vide figura 5.4);
- Indexador: recebe os resultados do módulo de Frequência de Idéias e gera um arquivo que contém índices referentes aos dados do proprietário de uma determinada frase. É utilizado somente quando da manipulação de *logs* de *chats* e desativado quando da utilização de outros documentos (vide figura 5.5);
- Votador: é o responsável por gerar a página HTML com todas as idéias geradas e submetê-las ao usuário, que irá atribuir um peso ou voto àquela que considerar mais importante no escopo da discussão que segue-se e retornar este voto para o Sistema. Este é o módulo responsável pela interação entre o coordenador e o usuário através da Web (vide figura 5.6);

- Ordenador: recebe um arquivo contendo os votos dos usuários participantes da discussão e ordena-os de acordo com os pesos a estes atribuídos. Desta forma são escolhidos os assuntos para a geração de um consenso entre os usuários (vide figura 5.7);

- Filtro II: recebe um arquivo contendo os resultados do módulo Ordenador, ficando responsável por divulgar os resultados gerais da eleição aos usuários através de uma página HTML (vide figura 5.8).

A descrição geral do sistema ESCOP, em SDL [GRA 99, SAR 87], é a seguinte.

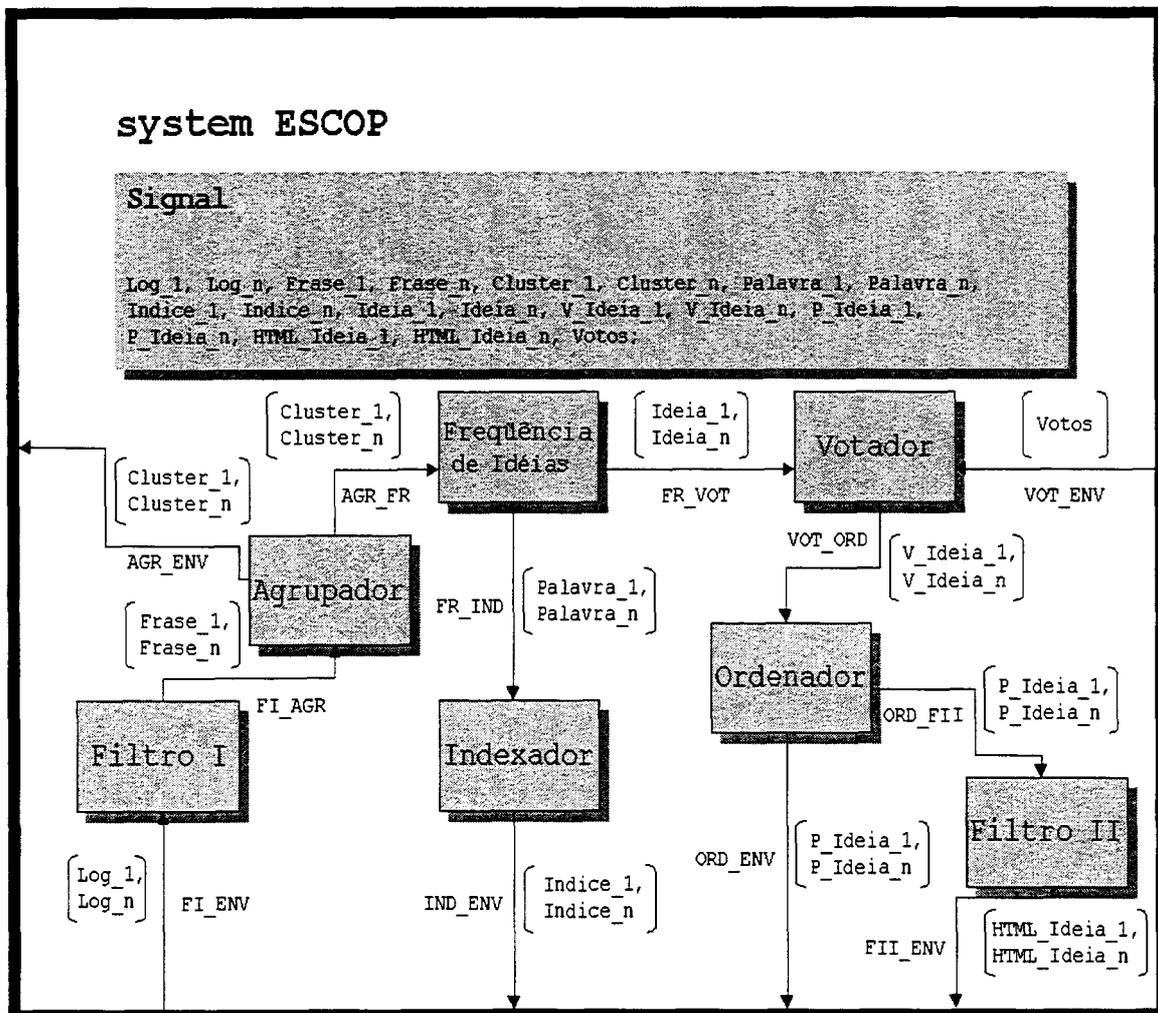


FIGURA 5.1 - O Sistema ESCOP - *Environment*

5.2 Ambiente de Implementação

Para a implementação do sistema ESCOP, utilizou-se de várias linguagens e estratégias de programação de computadores, além de programas servidores de *software* para serviços específicos de HTTP e FTP:

- Linguagem C++ [WEI 93, PER 94] e *scripts* CGI [WEI 97], para a implementação de processos auxiliares na votação de idéias através de páginas

Web. Este tipo de estrutura mostrou-se confiável, de fácil implementação e com poucas exigências computacionais para seu funcionamento;

- Apache *Web Server* [APA 99], como servidor HTTP, por ser utilizado por muitos servidores WWW, ser confiável, de fácil configuração e também exigir poucos recursos computacionais, além de estar disponível para a maioria dos sistemas operacionais utilizados;
- WAR FTP *Daemon* [WAR 99], como servidor FTP, por considerar-se mais simples a utilização de um *software* com estas características a um SGBD, a exemplo do que propôs Souza [SOU 94]. Considerou-se que a interação síncrona entre os usuários e o mediador da reunião dá-se através de uma ferramenta que suporte um *chat*, e a votação através de páginas Web, que devem retornar uma pequena quantidade de dados que podem ser manipulados de forma assíncrona pelo mediador, entre a geração de um ou outro resultado, não necessitando de uma estrutura mais complexa do que a que se propôs. Observa-se que tal estratégia somente torna-se necessária quando o servidor HTTP ou o Sistema ESCOP encontram-se em diferentes computadores. Quando estes dois últimos encontram-se na mesma instalação, portanto, o servidor de FTP é dispensável;
- Linguagem Delphi [OSI 97, LOY 96], para a construção dos módulos III, IV, V, VI, Filtro I e II, que podem ser observados no DFD (Diagrama de Fluxo de Dados) da figura 4.26. Determinou-se a utilização desta linguagem por permitir um desenvolvimento muito rápido, orientação a objetos e *interface* gráfica para o ambiente Windows. A intenção inicial era a de desenvolver todo o sistema utilizando a linguagem C através de *scripts* CGI. Abandonou-se esta premissa ao perceber-se várias incompatibilidades na execução de programas em C via *scripts* CGI no ambiente Windows, quando da utilização do servidor HTTP Apache. Entre estas, observou-se que as *system calls* em C, não eram corretamente executadas, apesar da correta configuração do servidor HTTP e do próprio programa, o que acarretou sérios problemas na implementação; em uma versão do *software* para ambiente Linux, os problemas deram-se também com as *system calls*, que exigiam o *userid* de *root* ao sistema, o que representa uma séria falha de segurança para a proposta; como solução a este problema, optou-se então por desenvolver os módulos anteriormente citados em uma linguagem com suporte ao ambiente Windows e, para a interação entre o mediador e os usuários, optou-se por uma comunicação via páginas Web, com *scripts* HTML;
- *Scripts* HTML [SAV 97], por serem mais fáceis e de rápida construção se comparados a outras propostas para a geração de páginas Web.

A implementação e os testes da ferramenta ESCOP foram feitos em um computador Pentium 200MHz, com 64 MB de memória RAM, 4GB de disco rígido e com um sistema operacional Windows 98 (mas também podem ser utilizados em sistemas Windows 95 e NT), o mesmo utilizado para fins de testes da ferramenta *Eurekha*. O tempo de processamento é limitado ao tamanho do arquivo de *log* e da quantidade de frases a serem agrupadas, além da quantidade de classes de *StopWords* que são selecionadas para a ferramenta de agrupamento. Também influi no desempenho o limite entre o tamanho mínimo e máximo de *clusters* a serem inspecionados.

5.3 Descrição Funcional dos Módulos

Neste sub-item , explicita-se, em descrição SDL, o funcionamento de cada um dos módulos apresentados na figura 5.1 e que foram comentados no sub-item 5.1.

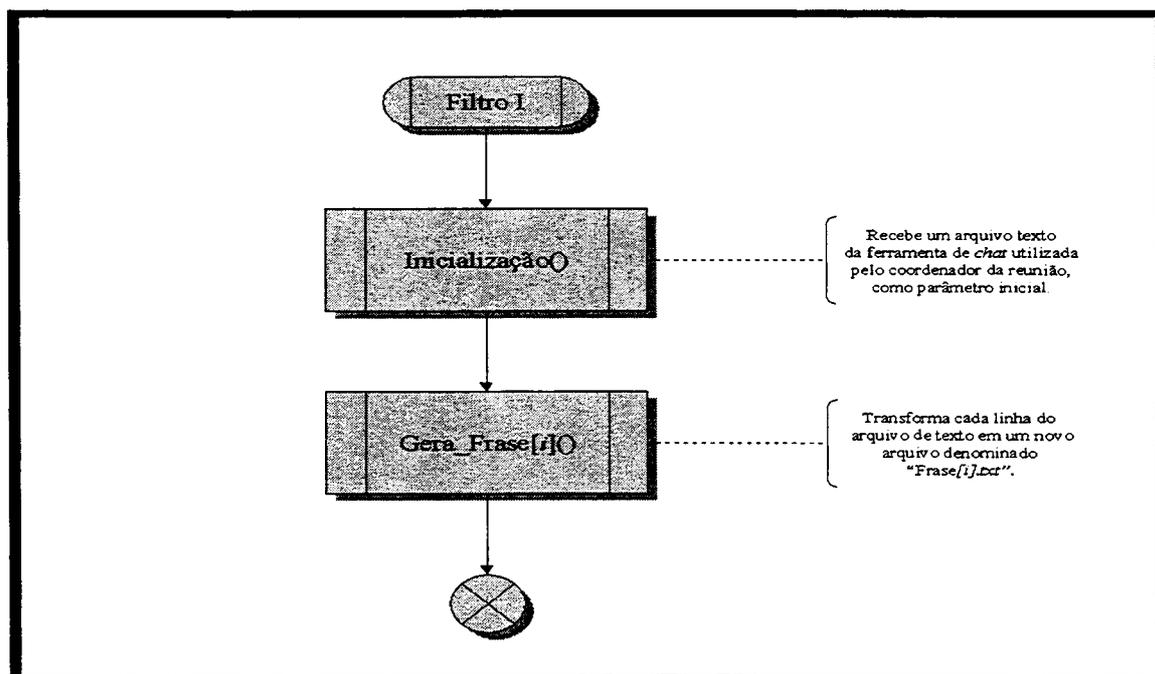


FIGURA 5.2 - O módulo Filtro I

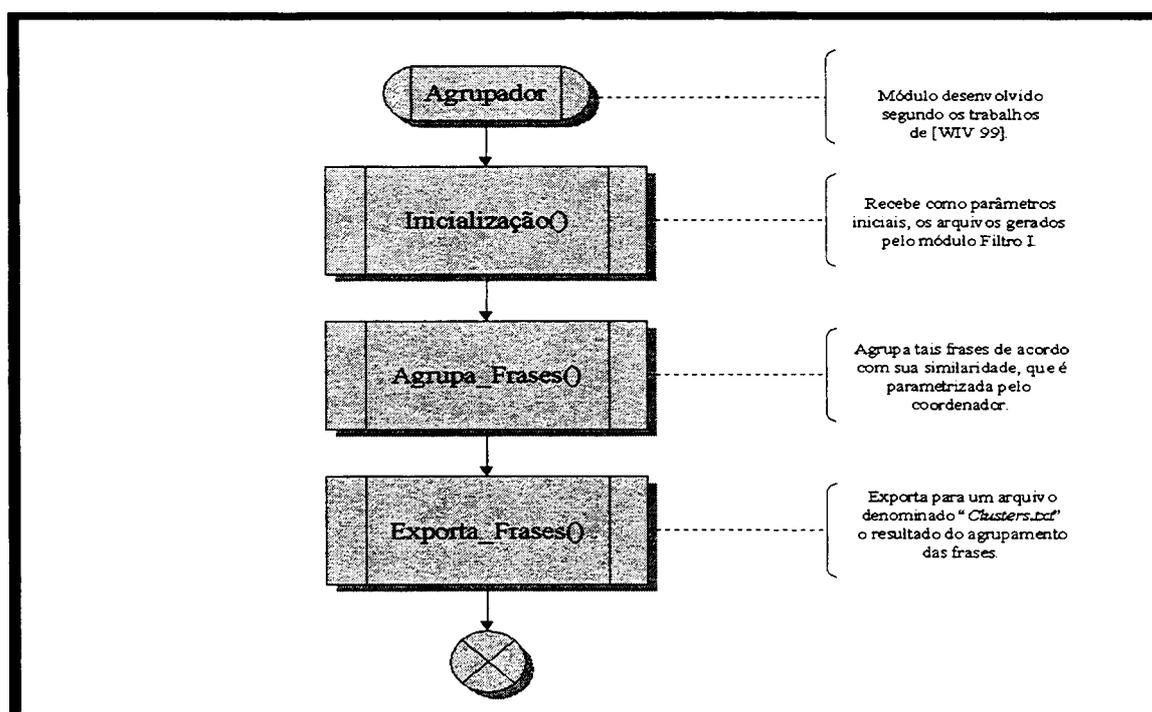


FIGURA 5.3 - O módulo Agrupador

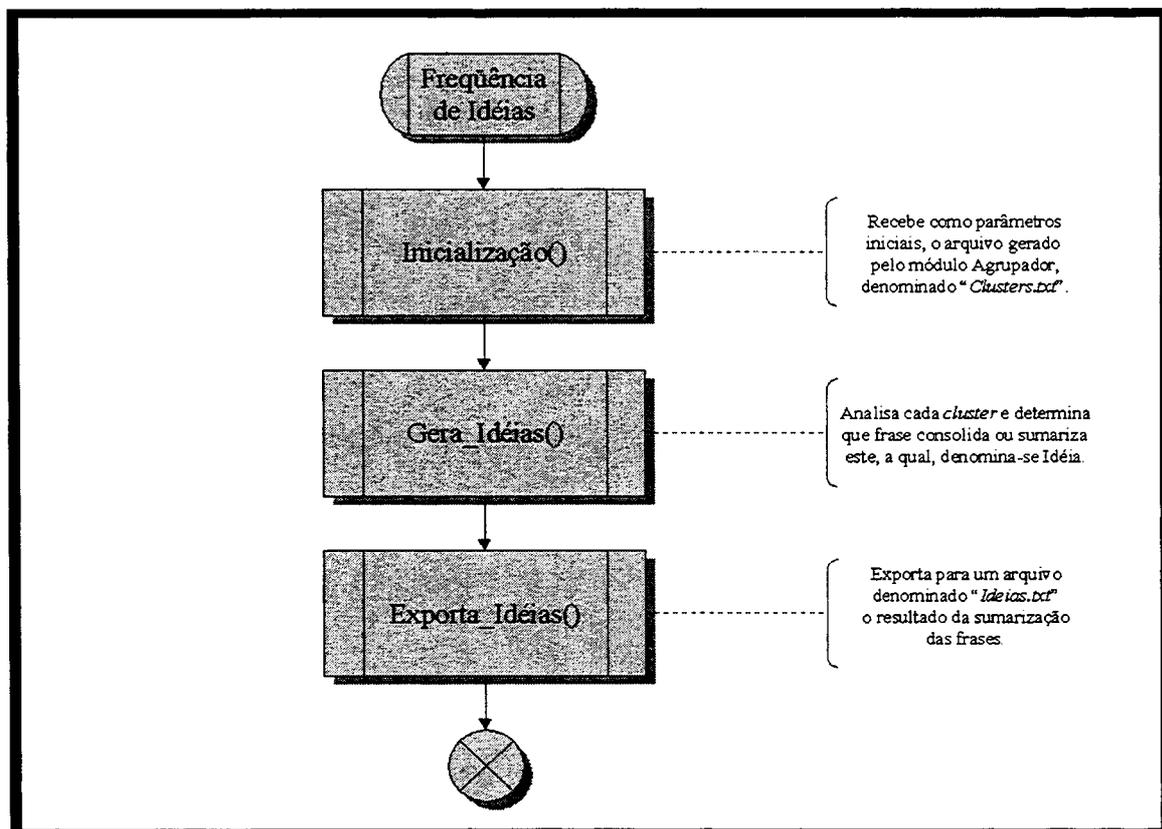


FIGURA 5.4 - O módulo de Frequência de Idéias

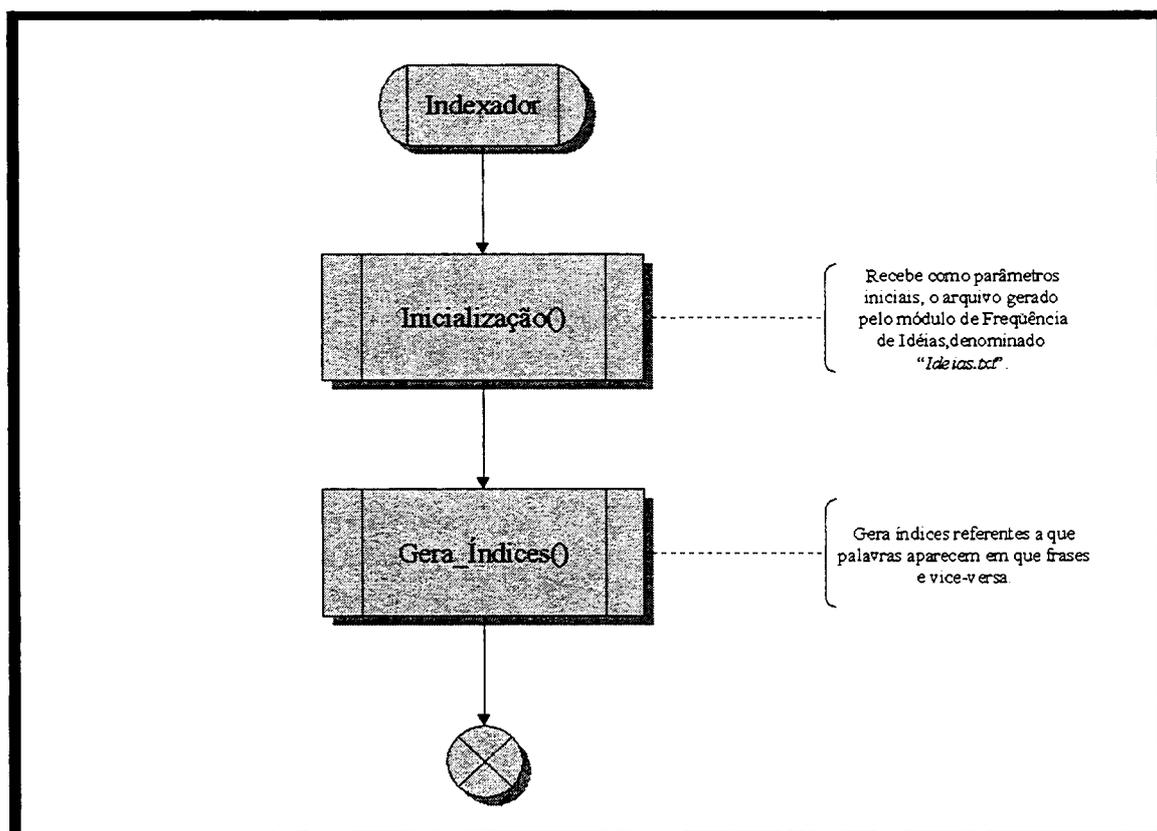


FIGURA 5.5 - O módulo Indexador

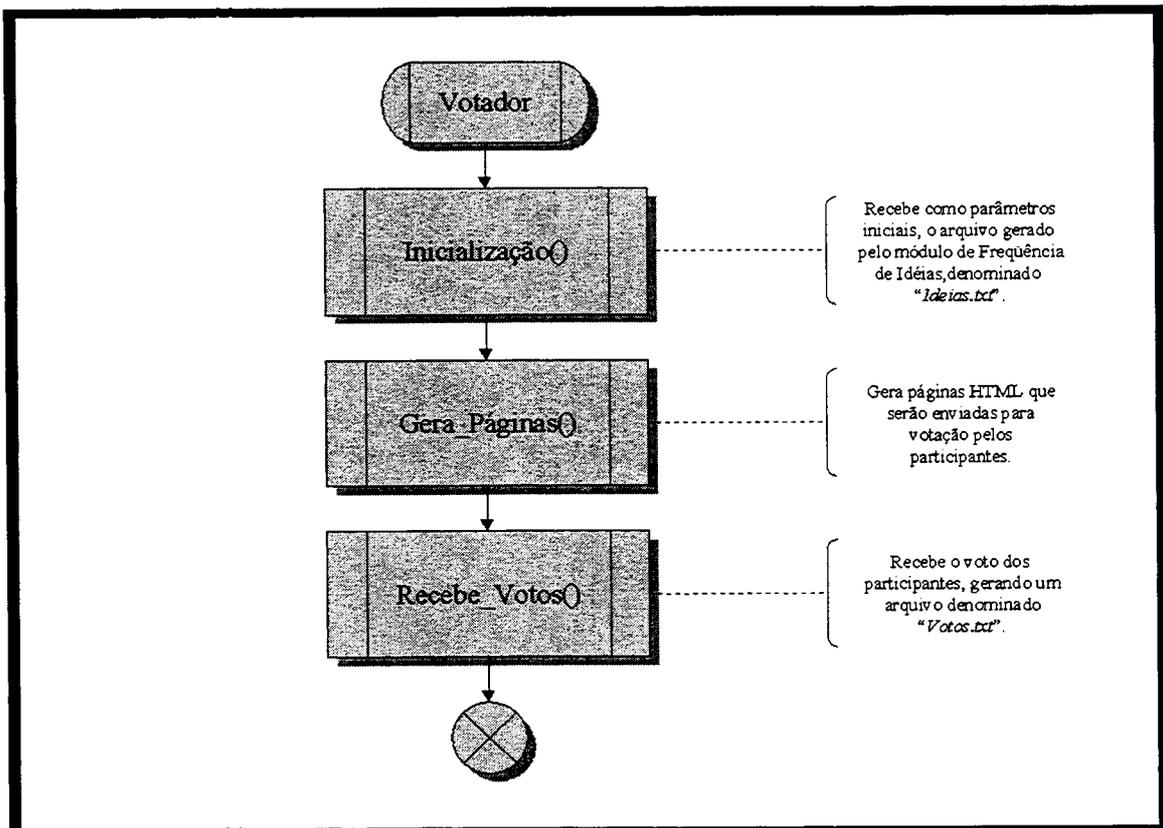


FIGURA 5.6 - O módulo Votador

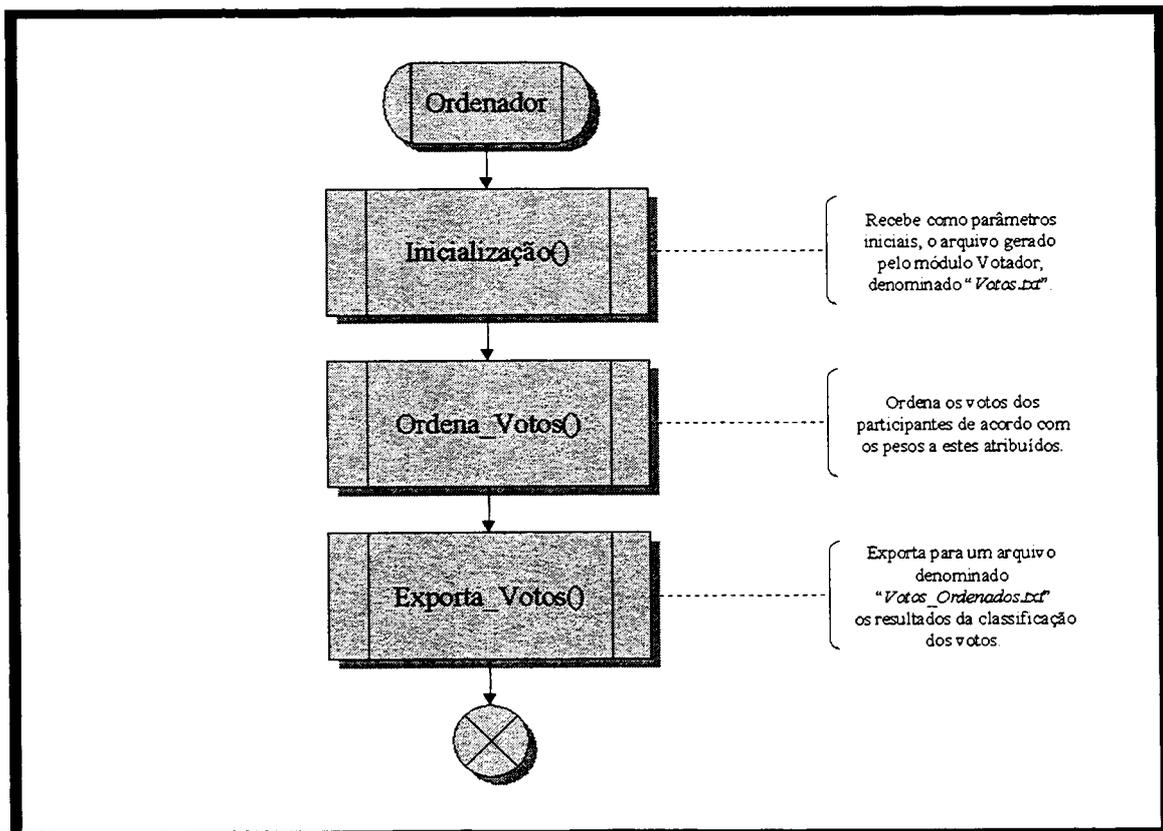


FIGURA 5.7 - O módulo Ordenador

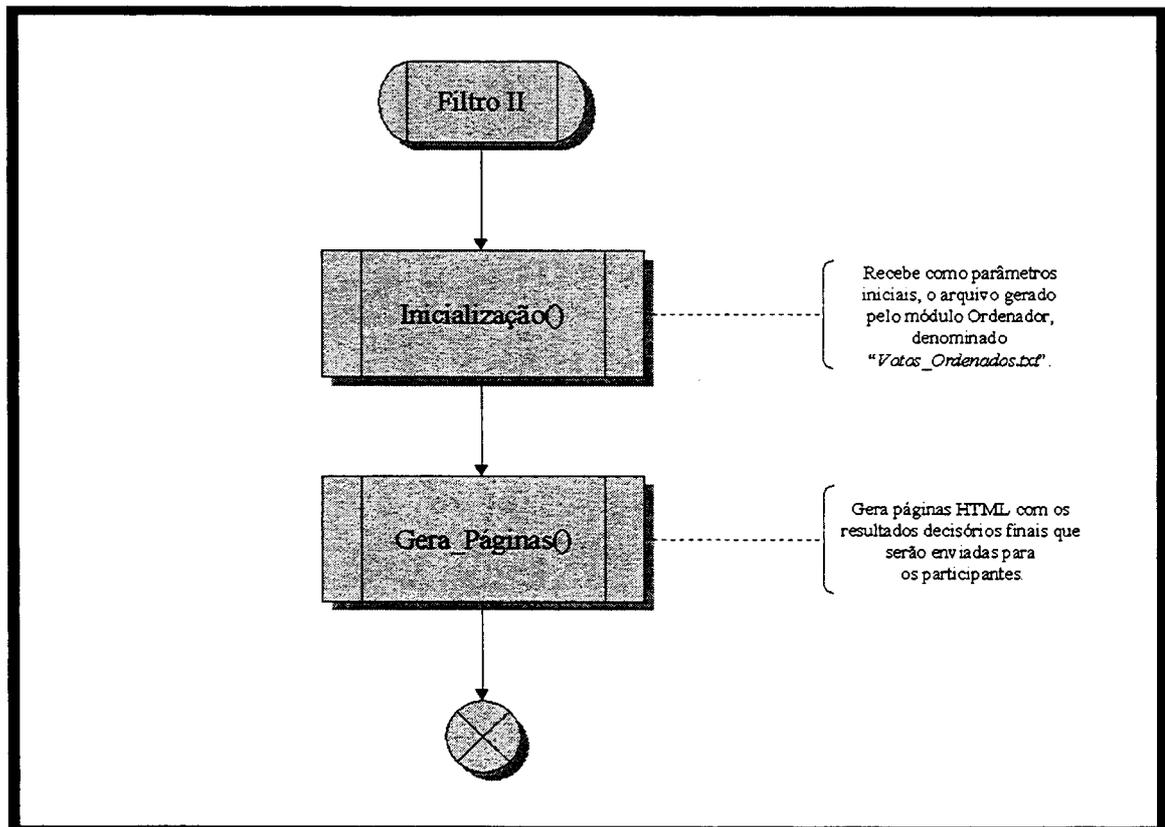


FIGURA 5.8 - O módulo Filtro II

5.4 Descrição da Interface

A interface da ferramenta é constituída por opções, que, ao serem selecionadas pelo usuário disparam um ou mais dos processos definidos na figura 5.1. A tela principal da ferramenta pode ser observada na figura 5.9. Como percebe-se, o sistema deve ser executado no mesmo computador e ambiente onde estiver instalada a ferramenta *Eureka*, já que esta última funciona como um dos módulos do protótipo proposto. No mesmo equipamento devem estar disponíveis um servidor HTTP (recomenda-se o referido em [APA 99]), um servidor de FTP (recomenda-se o referido em [WAR 99]) e um Web browser (utilizou-se o *Netscape Navigator*). A utilização de um servidor FTP não é absolutamente necessária em um primeiro momento e para esta versão do Sistema ESCOP, a menos que o servidor HTTP esteja em uma máquina diferente deste Sistema. A configuração destes sistemas (HTTP e FTP) foge do escopo deste estudo, mas a documentação completa pode ser obtida nas fontes de referência citadas.

Deve-se observar que nas janelas do Sistema desenvolvido, os botões de função são disponibilizados à direita destas, em ordem descendente a que devem ser utilizados. A única exceção a esta regra se dá no módulo "Frequência de Idéias", onde os botões estão distribuídos horizontalmente, com a ordem de utilização dando-se da esquerda para a direita.

A figura 5.10 mostra que, ao selecionar-se qualquer opção da tela principal, têm-se a opção "Informações" seguida da opção "Executar Módulo". A primeira fornece

A opção "Filtro I" pode ser observada na figura 5.11.

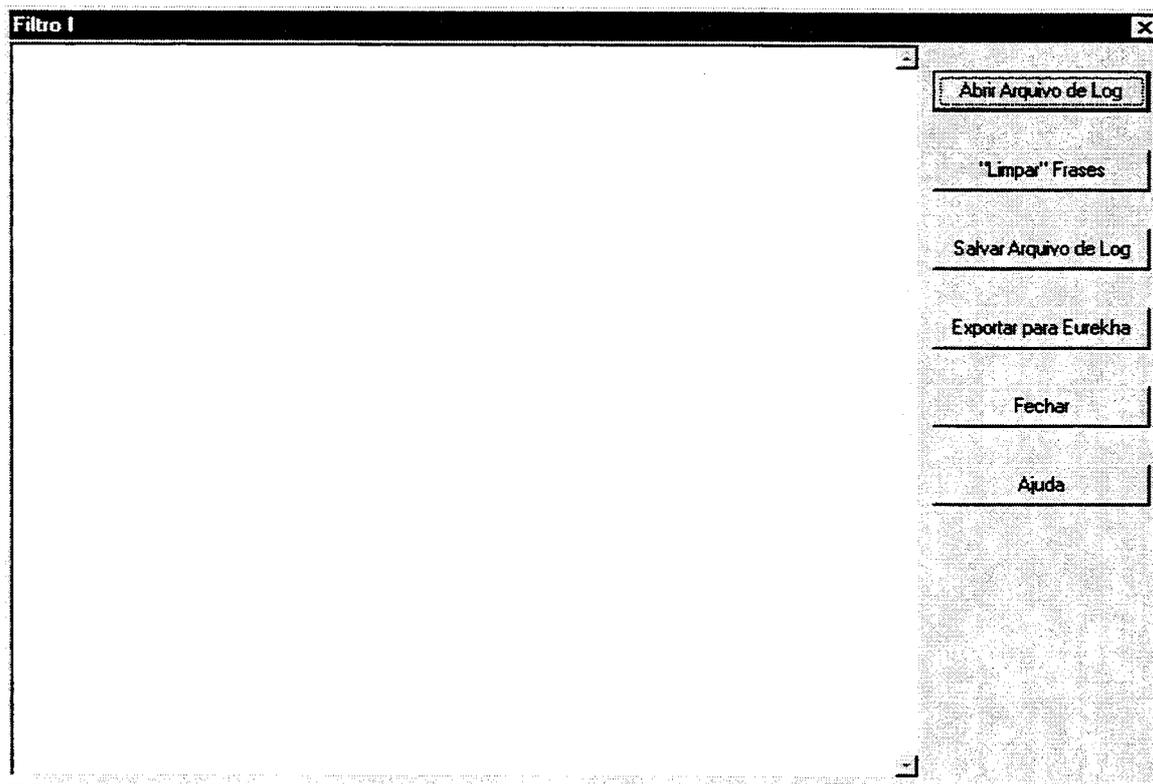


FIGURA 5.11 - O módulo "Filtro I"

O funcionamento do módulo ilustrado pela figura 5.11 é o seguinte (estas mesmas informações podem ser obtidas pressionando-se o botão "Ajuda"):

- O módulo Filtro I é o primeiro procedimento que deve ser efetuado pelo usuário do sistema ESCOP;
- Abrir Arquivo de *Log*: Esta opção permite abrir um arquivo texto e modificá-lo de acordo com as necessidades do usuário. Para fins de agrupamento, deve-se estar atento que cada linha de tal arquivo texto corresponde a um documento que será agrupado. Documentos muito extensos, devem ser utilizados com precaução e, se possível, evitados, já que não foram efetuados testes que assegurem seu perfeito funcionamento nestas condições;
- "Limpar Frases": Esta opção deve ser utilizada somente em casos especiais. Como caso especial, considera-se o arquivo de *log* de um programa de *chat* que, antes do início do conteúdo da frase do usuário, grava o nome deste, seguido do caractere ":". Este botão serve para eliminar o nome do usuário (bem como o caractere ":") deste arquivo de *log*. Após utilizar esta opção, não é necessário pressionar o botão "Salvar Arquivo de *Log*";
- Salvar Arquivo de *Log*: Esta opção permite salvar um arquivo texto;
- Exportar para *Eureka*: Esta opção irá abrir um arquivo texto e transformar cada linha deste em um novo arquivo denominado "*Frase[i].txt*", que será armazenado no diretório "EXPORT", o qual está localizado junto à instalação do software ESCOP. É neste diretório que o software *Eureka* irá encontrar os

arquivos para *clusterização*, que vem a ser o segundo passo. Além destes procedimentos, esta opção realiza ainda uma "normalização de termos", convertendo todas as letras do texto para maiúsculas (UPcase) e eliminando a acentuação gráfica de todas as palavras, para melhorar o processo de agrupamento. É importante observar que cada documento é delimitado por um "ENTER" ao seu final, ou seja, por convenção, utilizou-se este caractere como separador de documentos;

- Observações: Se a quantidade de frases indicadas após o processo de exportação for diferente daquele indicado pelo arquivo visualizado no diretório EXPORT, basta inserir um "ENTER" ao final do arquivo de frases antes do processo de exportação.

A opção "Agrupador" pode ser observada na figura 5.12.

Documento	Tamanho	Palavras	Ocorrências	Tipo	Origem
FRASE1.TXT	0.06 Kb	12	12	ASCII	C:\ESCOP\ESCOP\EXPORT\
FRASE10.TXT	0.06 Kb	9	9	ASCII	C:\ESCOP\ESCOP\EXPORT\
FRASE11.TXT	0.02 Kb	2	2	ASCII	C:\ESCOP\ESCOP\EXPORT\
FRASE12.TXT	0.09 Kb	14	14	ASCII	C:\ESCOP\ESCOP\EXPORT\
FRASE13.TXT	0.07 Kb	10	10	ASCII	C:\ESCOP\ESCOP\EXPORT\
FRASE14.TXT	0.05 Kb	8	8	ASCII	C:\ESCOP\ESCOP\EXPORT\
FRASE15.TXT	0.12 Kb	18	18	ASCII	C:\ESCOP\ESCOP\EXPORT\
FRASE16.TXT	0.15 Kb	19	23	ASCII	C:\ESCOP\ESCOP\EXPORT\
FRASE17.TXT	0.12 Kb	17	17	ASCII	C:\ESCOP\ESCOP\EXPORT\
FRASE18.TXT	0.04 Kb	8	8	ASCII	C:\ESCOP\ESCOP\EXPORT\
FRASE19.TXT	0.06 Kb	6	7	ASCII	C:\ESCOP\ESCOP\EXPORT\
FRASE2.TXT	0.03 Kb	6	7	ASCII	C:\ESCOP\ESCOP\EXPORT\
FRASE20.TXT	0.07 Kb	11	11	ASCII	C:\ESCOP\ESCOP\EXPORT\
FRASE21.TXT	0.02 Kb	4	4	ASCII	C:\ESCOP\ESCOP\EXPORT\
FRASE22.TXT	0.05 Kb	6	6	ASCII	C:\ESCOP\ESCOP\EXPORT\
FRASE23.TXT	0.09 Kb	16	16	ASCII	C:\ESCOP\ESCOP\EXPORT\
FRASE24.TXT	0.03 Kb	6	6	ASCII	C:\ESCOP\ESCOP\EXPORT\
FRASE25.TXT	0.07 Kb	9	9	ASCII	C:\ESCOP\ESCOP\EXPORT\
FRASE26.TXT	0.03 Kb	6	6	ASCII	C:\ESCOP\ESCOP\EXPORT\
FRASE27.TXT	0.11 Kb	16	17	ASCII	C:\ESCOP\ESCOP\EXPORT\

FIGURA 5.12 - O módulo "Agrupador"

O funcionamento do módulo ilustrado pela figura 5.12 é o seguinte (estas mesmas informações podem ser obtidas pressionando-se a opção "Ajuda" na tela principal na opção "Agrupador", referente a este módulo):

- O módulo Agrupador é o segundo procedimento que deve ser efetuado pelo usuário do sistema ESCOP;
- Adicionar documento local (*Ins*): Esta é a primeira opção que deve ser considerada neste módulo. Permite adicionar os documentos texto que serão

agrupados. A janela de seleção permite que se escolha por um diretório específico na árvore de diretórios, cujos arquivos são selecionados através do botão "Adicionar >>", e retirados da seleção através do botão "<< Remover". Observa-se que somente os arquivos que estiverem na janela da direita "Documentos Selecionados" realmente farão parte do processo de agrupamento. Tais documentos encontram-se no diretório "EXPORT" na sub-árvore de diretórios do sistema ESCOP;

- **Selecionar *StopWords*:** Esta opção permite selecionar as *StopWords* que serão utilizadas para fins de agrupamento. Quanto maior a quantidade de classes e de objetos destas classes de *StopWords*, maior será a qualidade dos *clusters* gerados;

- **Identificar Relacionamentos:** O botão "Identificar relacionamentos" inicia o procedimento de construção da matriz de similaridades. Este processo somente é interrompido quando estiver terminado. É este o procedimento mais demorado da ferramenta de agrupamento, já que pode consumir um tempo de processamento entre 10 minutos e 30 minutos, de acordo com as informações de entrada. Deve-se salientar que, até esta opção, ainda não selecionou-se nenhum tipo de algoritmo de agrupamento. Este procedimento também identifica as palavras relevantes de cada frase;

- **Visualizar Grupos:** É o processo responsável pela identificação dos grupos gerados, através de um algoritmo especificado pelo usuário. Observa-se que o algoritmo "*Cliques*" apresentou melhor desempenho nos testes efetuados, mas dispõem-se das outras 3 opções, segundo a opção "Algoritmo" disponibilizada pela ferramenta;

- **Coefficiente de Similaridade:** A barra deslizante intitulada "Coeficiente de Similaridade" é a responsável pelo nível que o usuário irá requerer para a similaridade entre os documentos. Quanto mais próximo de 0%, maior será o tamanho dos grupos gerados, entretanto, menor será a similaridade entre os documentos que compõem tais grupos. Ao contrário, quanto mais próximo de 100%, menores serão os tamanhos dos *clusters* gerados e menor a quantidade de documentos destes. De acordo com os experimentos práticos, a posição da barra de rolagem que formou grupos mais coesos ficou entre 15% e 25%. Ao selecionar-se valores fora destes limites, os grupos gerados não forneceram informações significativas. Entretanto, isto foi o que observou-se em casos específicos. O usuário deve estar familiarizado com a ferramenta para então poder escolher entre um ou outro algoritmo e também definir o grau de similaridade mais adequado para seu caso particular;

- **Reagrupar:** Deve ser utilizado sempre que modificar-se o algoritmo de agrupamento ou o valor do coeficiente de similaridade. A opção "Opções" deve estar desabilitada para eliminar-se os *clusters* unitários, que, geralmente, compõe-se de informações irrelevantes;

- **Ferramentas:** O último passo a ser executado com este software é a opção "Ferramentas", onde deve-se escolher a opção "Exportar resultado do agrupamento...", que será responsável por gerar um arquivo texto chamado "*clusters.txt*" no diretório corrente da ferramenta *Eureka*, arquivo este que será utilizado pelo sistema ESCOP.

A opção "Frequência de Idéias" pode ser observada na figura 5.13.

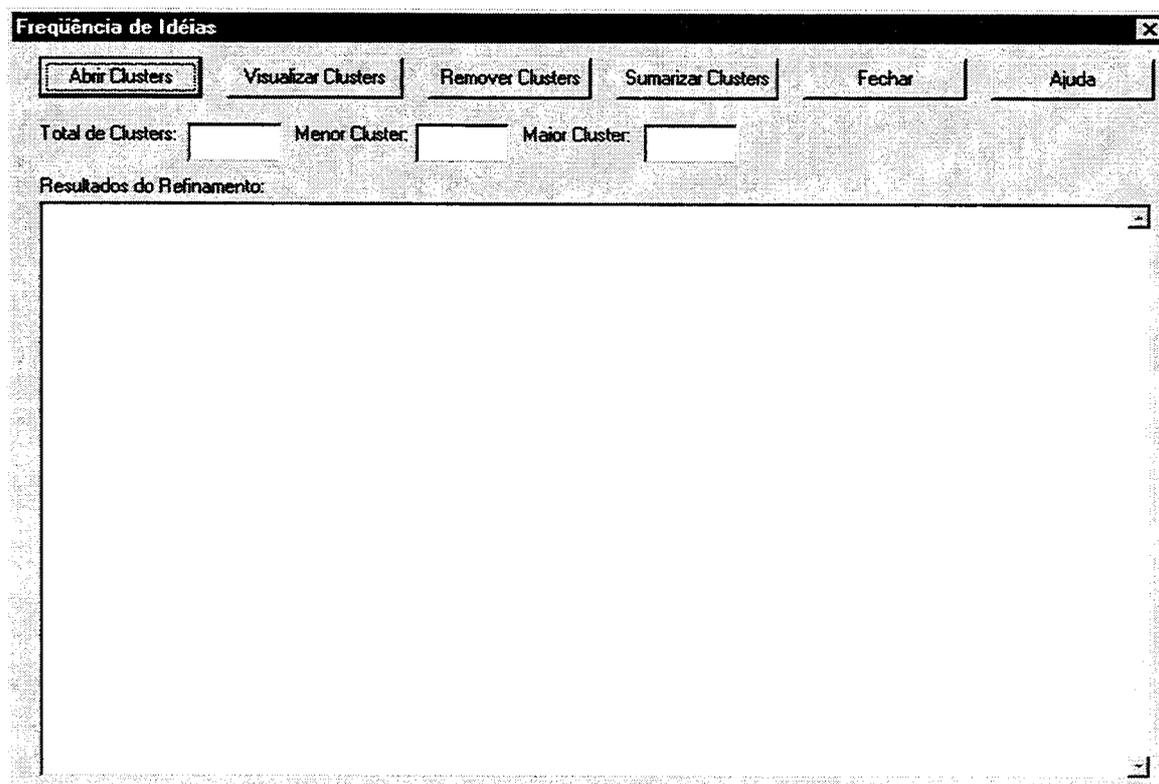


FIGURA 5.13 - O módulo "Frequência de Idéias"

O funcionamento do módulo ilustrado pela figura 5.13 é o seguinte (estas mesmas informações podem ser obtidas pressionando-se o botão "Ajuda"):

- O módulo Frequência de Idéias é o terceiro procedimento que deve ser efetuado pelo usuário do sistema ESCOP;
- *Abrir Clusters*: Esta opção permite abrir o arquivo texto "*clusters.txt*", gerado pela ferramenta *Eureka*. Após selecionar esta botão, basta que o usuário pressione duas vezes o botão do *mouse* sobre o nome do arquivo indicado anteriormente. Feito isto, os campos "Total de *Clusters*", "Menor *Cluster*" e "Maior *Cluster*" serão preenchidos, de acordo com as informações contidas no arquivo de *clusters*. Dependendo do tamanho do arquivo gerado, este botão pode demorar alguns segundos para apresentar a resposta na tela;
- *Visualizar Clusters*: Esta opção permite visualizar o conjunto de *clusters* gerados, através de um refinamento indicado pelo usuário, que pode, através desta possibilidade, concentrar-se em observar somente determinados grupos de documentos;
- *Remover Clusters*: Durante uma conversa livre em um "*Chat*", geralmente os participantes formam frases fora do contexto do que se está discutindo, frases estas que serão agrupadas durante o respectivo processo de agrupamento, independentes de seu sentido e escopo em relação ao que se está tratando. Para resolver este problema, o botão "*Remover Clusters*" permite que o usuário do sistema possa eliminar manualmente aqueles *clusters* que não formam informações consistentes para o processo de sumarização ou mesmo que já tenha passado por este processo;

- **Sumarizar Clusters:** É esta opção que permite ao usuário sumarizar, ou seja, extrair as informações realmente relevantes dos *clusters* gerados no processo de agrupamento. Se o usuário não estiver contente com o conteúdo de alguma das idéias geradas por este processo, ele pode optar por eliminar tal *cluster*. É este botão que gera o arquivo "*ideias.txt*" que será utilizado pelo módulo votador;
- **Observações:** Na utilização prática deste módulo, observou-se que as principais opções utilizadas (definidas pelos respectivos botões), são a primeira, a terceira e a quarta opções. Isto é um fato que depende da forma de utilização e da ambientação com o sistema, mas pode servir de dica para usuários mais experientes do sistema ESCOP.

A opção "Votador" pode ser observada na figura 5.14.

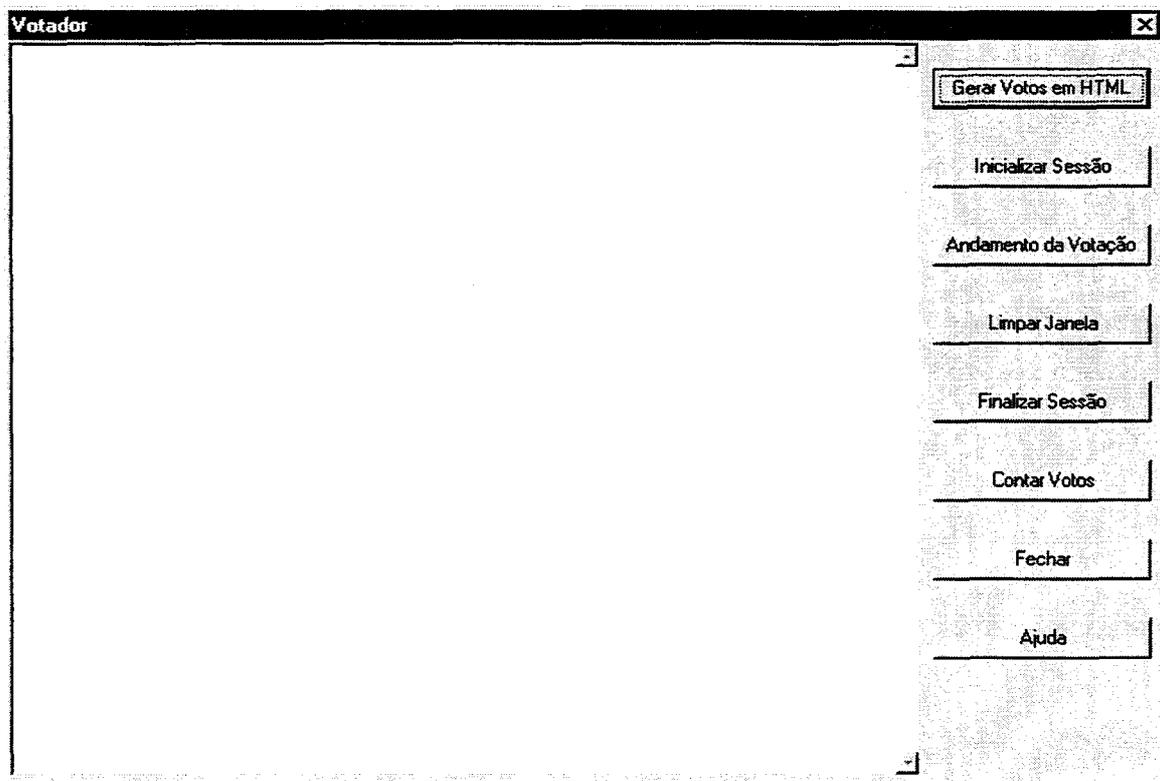


FIGURA 5.14 - O módulo "Votador"

O funcionamento do módulo ilustrado pela figura 5.14 é o seguinte (estas mesmas informações podem ser obtidas pressionando-se o botão "Ajuda"):

- O módulo Votador é o quarto procedimento que deve ser efetuado pelo usuário do sistema ESCOP;
- **Gerar Votos em HTML:** Esta opção permite criar o arquivo HTML, a partir de "*ideias.txt*", que deve ser distribuído aos votantes do sistema para que estes possam atribuir pesos às idéias sumarizadas pelo coordenador no terceiro módulo (Frequência de Idéias). Para isto, basta escolher um nome para o arquivo HTML e salvá-lo do diretório "*public_html*" da máquina onde o sistema ESCOP estiver instalado. Após este procedimento, o coordenador da reunião deve informar o endereço de tal página aos votantes para que estes possam então dar seus votos;

- **Iniciar Sessão:** No sistema ESCOP, somente uma sessão de votos pode ser trabalhada por vez. Desta forma, os dados de votação de uma sessão sempre ficam armazenados no diretório "VOTOS", presente junto a instalação do Sistema. Quando utiliza-se este botão, uma mensagem é dada ao usuário, dizendo que os resultados anteriores serão apagados (justamente para evitar a mistura de votos antigos com os novos votos). Se o usuário desejar, pode, manualmente (utilizando o *Windows Explorer*) guardar o conteúdo de tal diretório para fins de documentação. Sempre que quiser saber os resultados de uma reunião passada, basta transferir tais arquivos para o diretório "VOTOS" (tendo o cuidado de preservar os dados ali contidos) e então pressionar o botão "Contar Votos", quando então será mostrada na janela do módulo os resultados de votações anteriores. Explica-se que este é um procedimento que deve ser utilizado somente para a visualização de reuniões anteriores, já que no procedimento normal, ao pressionar-se o botão "Contar Votos", são mostrados os resultados da sessão atual;
- **Andamento da Votação:** Durante uma sessão de votação, o coordenador pode achar necessário acompanhar os votos dos usuário, para saber quem já votou, em que votou e quem ainda falta votar. É o que permite este botão. É importante observar que ele somente pode ser utilizado quando uma sessão estiver aberta. Se esta estiver fechada, é mais interessante utilizar o botão "Contar Votos";
- **Limpar Janela:** Em qualquer momento do processo de votação, este botão pode ser pressionado para limpar a janela de resultados do módulo;
- **Contar Votos:** Este botão é o responsável pela contabilização dos votos enviados pelos usuários. Seu funcionamento não requer nenhuma interação por parte do coordenador. Ao final do processamento dos dados da votação, exibe um relatório com os dados desta. Deve-se observar que este botão não permite nenhum processamento enquanto uma sessão estiver em aberto, já que, somente pode-se contar votos de uma sessão já encerrada;
- **Observações:** É importante salientar que os votos não enviados até o encerramento de uma sessão não poderão, de forma alguma, ser computados pelo sistema, já que cada vez que inicia-se uma sessão, os votos da sessão anterior são eliminados. A cópia dos resultados de sessões anteriores, como é citado acima, não permite a continuidade de uma sessão. Portanto, o coordenador deve ter o máximo cuidado ao encerrar uma sessão, observando sempre se todos os usuários já votaram, pois, caso contrário, alguns poderão ter sua oportunidade de voto suprimida.

A opção "Ordenador" pode ser observada na figura 5.15.

O funcionamento do módulo ilustrado pela figura 5.15 é o seguinte (estas mesmas informações podem ser obtidas pressionando-se o botão "Ajuda"):

- **O módulo Ordenador** é o quinto procedimento que deve ser efetuado pelo usuário do sistema ESCOP;
- **Editar Resumo Votação:** Esta opção permite abrir e editar o arquivo "*resumo.txt*" criado pelo módulo Votador. Todas as alterações feitas nesta janela serão mostradas na página HTML que será gerada a partir do arquivo "*votos_ordenados.txt*";

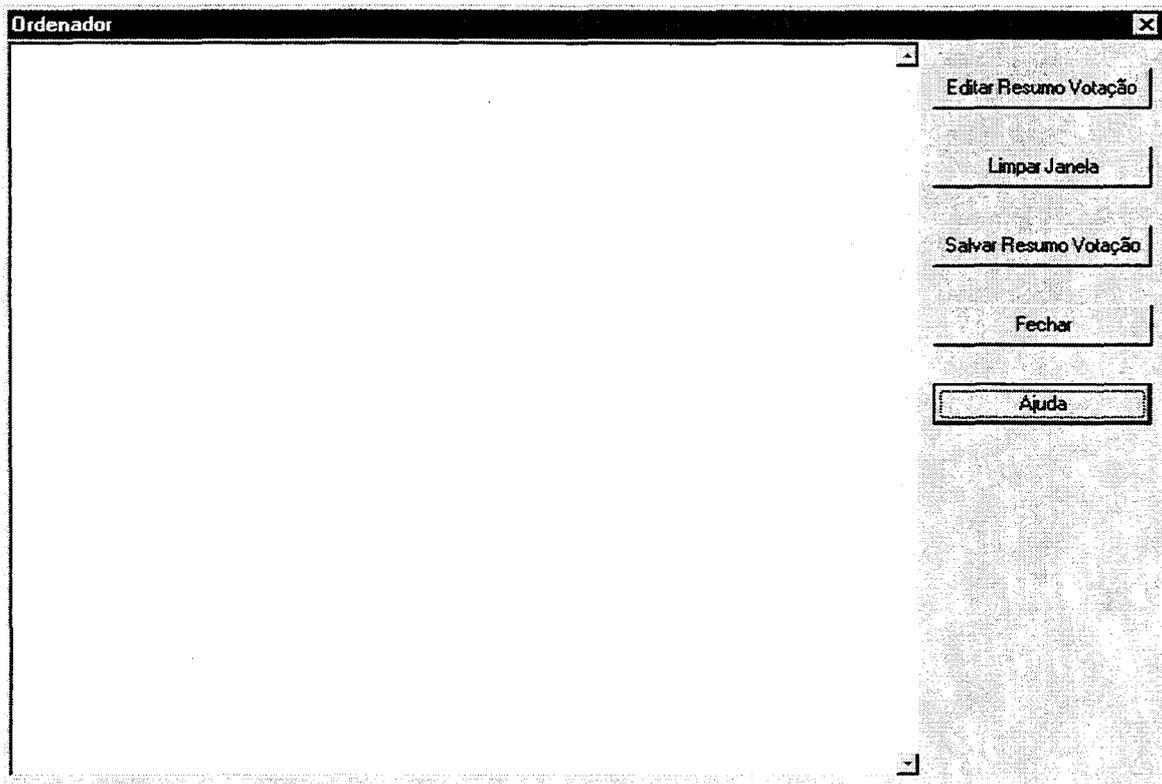


FIGURA 5.15 - O módulo "Ordenador"

- Limpar Janela: Em qualquer momento deste processo, este botão pode ser pressionado para limpar a janela de resultados do módulo;
- Salvar Resumo Votação: Salva os dados constantes na janela de edição para o arquivo "*votos_ordenados.txt*". Se esta janela estiver vazia, o arquivo salvo não terá conteúdo algum;
- Observações: É importante salientar que tudo o que aparecer na janela de edição deste módulo irá constar na página HTML final que será disponibilizada aos votantes, contanto que se pressione o botão "Salvar Resumo Votação".

A opção "Filtro II" pode ser observada na figura 5.16.

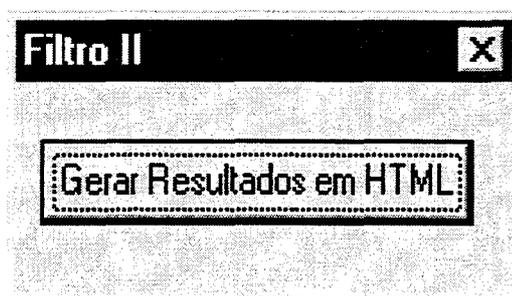


FIGURA 5.16 - O módulo "Filtro II"

O funcionamento do módulo ilustrado pela figura 5.16, por ser muito simplificado, não possui uma opção de ajuda, já que a única opção disponível aqui é o botão "Gerar Resultados em HTML". Este procedimento gera a página com o sumário

da sessão, através da manipulação dos dados pelo módulo "Ordenador", conforme mencionado anteriormente.

5.5 Considerações Gerais sobre a Implementação

Para o funcionamento da ferramenta ESCOP, são necessários, no mínimo, os seguintes requisitos de *hardware* e *software*:

- Microsoft Windows 95, com Winsock 2 instalado;
- *Web Server* APACHE, versão para Windows;
- Um computador PC, com pelo menos 32 MB de memória RAM.

Os testes realizados com o protótipo do Sistema ESCOP foram efetuados, tendo instalados no mesmo computador, a ferramenta *Eurekha* (o módulo "Agrupador" do sistema) e o *software* APACHE. Para uma situação diferente desta, devem ser feitas alterações no código fonte do Sistema, para permitir que este trabalhe em um computador diferente daquele onde está o servidor HTTP. Para esta situação, indica-se a utilização de um *software* servidor de FTP, conforme foi mencionado anteriormente. Uma observação importante é que o Sistema ESCOP sempre deverá ser executado no mesmo computador da ferramenta *Eurekha*.

Com este trabalho conseguiu-se avaliar estudos e ferramentas que permitem realizar automaticamente o agrupamento de idéias com foco ou temas comuns para a organização de tópicos discutidos no decorrer de um processo decisório em grupo, bem como algumas metodologias de apoio ao processo de tomada de decisão em grupo, situações estas comuns na Educação a Distância e no trabalho cooperativo.

Através da implementação do protótipo foi possível a confirmação das suposições levantadas no início deste trabalho. Dentre as principais vantagens observadas durante a fase de testes, pode-se citar:

- Como não foi projetado para ser utilizado por um usuário comum, e sim pelo coordenador de um processo decisório, o mediador precisa ter uma cópia do ESCOP instalada em sua máquina. O conteúdo das discussões torna-se disponível aos participantes de uma reunião através das páginas HTML geradas pelo Sistema;
- Devido ao vasto alcance da Internet e independência de plataforma do serviço WWW, o Sistema pode ter seus dados compartilhados por qualquer pessoa, independentemente da plataforma de *hardware* ou *software* por esta utilizado;
- A comunicação entre usuários votantes e o coordenador de uma reunião dá-se através do protocolo HTTP, o que torna os detalhes de comunicação transparente aos usuários do Sistema;
- Sendo desenvolvido em ambiente Windows, o Sistema torna possível sua utilização em Windows 95/98/NT e também em futuras atualizações destes sistemas, tendo seu nível de segurança determinado pelas restrições impostas por estes últimos.

Algumas dificuldades foram encontradas durante a implementação do ESCOP, as quais, podem servir de referência para futuros trabalhos:

- A implementação do sistema em uma linguagem visual (DELPHI) proporcionou um alto nível de interação com o usuário, através de uma *interface* clara e de fácil aprendizado. Entretanto, isto limitou o Sistema a ser executado em somente uma plataforma de *software* (Windows), o que considera-se pouco viável para um sistema deste tipo. Nos testes iniciais, optou-se por desenvolver toda a *interface* e o processamento através de *scripts* CGI, que, conforme justificativas anteriores neste trabalho, mostraram-se inviáveis. A independência de plataforma pode vir a ser obtida através da reformulação do sistema para código JAVA, por exemplo. Isto também deve ser efetuado com o módulo de agrupamento (*Eurekha*), que foi utilizado para estes trabalhos como uma ferramenta de código fechado;

- O desenvolvimento de uma aplicação de caráter local permitiu a manutenção de um contexto na utilização da aplicação por parte de seu operador. Tal possibilidade deve ser observada quando de uma proposta para utilização distribuída;

O protótipo desenvolvido possibilitou a implementação integral de todas as funcionalidades previstas no modelo proposto. Trabalhos futuros poderão aprimorar cada uma das funcionalidades do modelo, como por exemplo, o desenvolvimento de um módulo agrupador que possibilite um trabalho tendo este código fonte disponível, para possíveis melhorias nos algoritmos de agrupamento.

6 Conclusões

Neste capítulo procura-se apresentar as contribuições oferecidas pelo estudo realizado, assim como busca-se discutir as aplicações que podem ser dadas ao modelo apresentado, suas restrições e sugestões para que este possa ser melhorado.

Dentro de seus objetivos iniciais, este trabalho deu continuidade à proposta iniciada por Souza [SOU 94], além de estudar e comparar novos tipos de ferramentas que possuíssem características e facilidades similares, como foi o caso das ferramentas WebSaber, TANGO *Interactive*, Interloq e SAACI. Ainda dentro deste escopo, no início dos estudos relativos à construção de um modelo inicial, imaginou-se importante a análise da ferramenta Lotus Notes.

Sendo assim, primeiramente objetivou-se o projeto e a possível implementação de um sistema capaz de automatizar a votação e a sumarização de idéias com foco ou temas comuns, utilizando-se de uma *interface* WWW para o usuário final, um processo de ligação entre esta e um SGBD relacional, finalizando-se com a construção e/ou utilização de um módulo de agrupamento de documentos.

Nas primeiras etapas do desenvolvimento desta pesquisa, ainda durante o levantamento bibliográfico complementar, notou-se que a implementação de uma ferramenta para agrupamento de documentos seria inviável, devido a complexidade envolvida em tal tipo de processamento e ao pouco tempo disponível para o sua programação.

Ainda nesta fase, optou-se por descartar a análise da ferramenta *Lotus Notes*, primeiramente pelo alto nível de complexidade requerido para o entendimento específico de seu funcionamento, como também pela dificuldade em obter-se uma cópia de demonstração para testes, o que, aliado ao valor extremamente elevado do *software*, foi considerado como sendo uma das principais dificuldades técnicas para a sua utilização.

A partir deste ponto, durante o estudo das características da ferramenta SADGV2, proposta por Souza [SOU 94], notou-se o alto nível de complexidade envolvido em sua construção, já que foram necessários a implementação de rotinas de comunicação através de chamadas RPC, por exemplo, e a construção e validação do módulo de agrupamento proposto por Freitas [FRE 93], também bastante complexo e que, considerou-se, de difícil utilização e manutenção, porém de muita utilidade para a proposta por estes apresentada.

Após estas considerações, optou-se por repensar o tratamento que seria dado ao novo modelo, pelas dificuldades encontradas, primeiramente aquelas inerentes à complexidade e à falta de tempo para a implementação de um módulo de agrupamento; a impossibilidade de reconstruir, disponibilizar e utilizar a ferramenta SADGV2; e, finalmente, a necessidade de que a nova proposta pudesse ser utilizada em condições reais, com funcionamento simples, confiável e que não exigisse grandes recursos computacionais (o uso de estações de trabalho SUN, por exemplo), como no caso da proposta de Souza [SOU 94].

Seguiram-se os trabalhos em busca de um módulo de tratamento de linguagem natural, já que, ainda considerava-se necessária a sua utilização. Para fins de testes, utilizou-se a ferramenta desenvolvida por Magalhães [MAG 94], que comportou-se

estritamente dentro de suas propostas, que eram a de fazer uma análise sintática de uma frase qualquer dada ao sistema. Entretanto, percebeu-se que, limitada ao tamanho de seu dicionário, tal proposta apresentou-se pouco prática para a utilização em uma situação real, pelo motivo de que, ao encontrar uma nova palavra, o sistema não possuía a capacidade de "aprendê-la", como fazia a proposta de Freitas [FRE 93], cuja recuperação não foi possível, pelos mesmos motivos encontrados nos trabalhos de Souza [SOU 94], ou seja, somente recuperou-se fragmentos de código dos sistemas.

Como conclusões parciais, percebeu-se que o trabalho de Souza [SOU 94] era muito complexo e o de Magalhães [MAG 94] pouco prático para as necessidades da nova proposta.

A solução encontrada foi então procurar por uma nova ferramenta que fosse de fácil utilização e aprendizagem, que pudesse ser utilizada em condições reais, práticas e que disponibilizasse farta documentação.

Foi nesta busca que optou-se pelo trabalho desenvolvido por Wives [WIV 99], já que tal proposta encaixou-se perfeitamente dentro das suposições levantadas para a implementação de um módulo de agrupamento de documentos, principalmente pela fundamentação teórica explicitada neste trabalho, que defende teoricamente que não é necessário o um módulo de processamento de linguagem natural para o agrupamento de documentos.

Com estes conceitos, optou-se por descartar a utilização dos trabalhos de Magalhães [MAG 94]. Salienta-se que os testes práticos tornaram possível comprovar tais condições teóricas, neste caso específico de utilização para o Sistema ESCOP.

Em seu funcionamento básico, a proposta de Wives [WIV 99], denominada *Eurekha*, recebe um grupo de documentos e agrupa-os de acordo com sua similaridade, através de parâmetros definidos pelo usuário.

Convém acrescentar que, durante o procedimento anterior ao processo de agrupamento, utilizou-se de uma normalização mínima de termos das frases que seriam submetidas ao módulo de agrupamento.

Tomou-se esta atitude devido aos testes práticos realizados mostrarem que os dados normalizados (retirada de acentos e conversão das letras das palavras para maiúsculas - ou minúsculas) simplificam a tarefa do módulo agrupador, gerando melhores resultados finais.

Baseando-se nestas características, tornou-se premente a disponibilização de uma *interface* de comunicação entre vários usuários dispersos geograficamente, de fácil utilização, baixo custo e com possibilidade de geração de documentação de encontros. Determinou-se que tal solução poderia ser suportada por uma ferramenta de *chat*.

Para evitar a implementação de um módulo específico de *software* para tratamento de comunicação via *chat*, optou-se por utilizar aqueles mais conhecidos aos usuários, contanto que tais ferramentas disponibilizassem um arquivo de *log* com todas as mensagens trocadas entre estes durante uma sessão de discussões.

Resolvidos os problemas de comunicação entre os usuários e de agrupamento das frases similares criadas por estes, concentrou-se esforços na tarefa referente a sumarização destas mensagens, ou seja, na geração das idéias de tais grupos de frases similares. Convencionou-se que uma idéia é uma frase que contém todas as características de determinado grupo e, portanto, pode resumir todo o seu conteúdo.

Para a solução deste procedimento, buscou-se referências em Salton e Loh [SAL 83, LOH 99]. O primeiro indicava algoritmos de recuperação de informações e processamento de linguagem natural, enquanto que o segundo apresentava estudos de vários métodos de tratamento de informações bem como técnicas para sumarização.

Estudadas tais referências e comparadas as várias propostas apresentadas, imaginou-se utilizar um método computacional híbrido, simples e rápido que permitisse a obtenção do sumário de um *cluster* de frases.

Tal algoritmo baseia-se na premissa de que, dado um conjunto de frases similares e sendo retirada destas frases todas as suas *StopWords* (cujo conceito é definido por Wives [WIV 99] em seus trabalhos), somente sobram a estas frases as palavras ditas de conteúdo, ou seja, os substantivos e os verbos, por exemplo.

Com estas informações, testou-se na prática, a escolha da frase que contivesse a maior quantidade de palavras de conteúdo de um dado *cluster*, que seria então considerada a idéia consolidadora de um *cluster* de frases. Os resultados atenderam com eficiência aos objetivos esperados.

Mostrou-se que o algoritmo é muito rápido e confiável, sendo que também constatou-se, durante os testes, um fato inusitado para uma dada situação específica.

Tal situação, implementada somente com propósito de testes de integridade das informações geradas pelo módulo de agrupamento e sumarização, constou da entrada de um documento textual qualquer, com cada parágrafo deste texto inserido no sistema como se fosse uma frase única.

Durante o processo de agrupamento, o coeficiente de similaridade entre cada documento foi configurado para ser igual a zero, ou seja, forçou-se a geração de um único *cluster*, composto por todos os parágrafos daquele texto.

Como resultado final, após a aplicação do algoritmo de sumarização neste *cluster*, obteve-se um parágrafo que pode ser considerado como o mais importante daquele texto dado como entrada, ou seja, é o resumo daquela fonte de informação.

Considera-se esta situação atípica, porque, na maior parte dos casos, o sistema deverá tratar *logs* de ferramentas de *chat*, com frases curtas e não grandes documentos de texto.

Concluiu-se, neste ponto, que o sistema ESCOP implementado atendeu a todas as expectativas a que foi inicialmente submetido, além de prover a capacidade de trabalhar com textos extensos, característica suportada pela ferramenta *Eurekha* mas não projetada, inicialmente, para ser suportada pela nova proposta.

Resumidamente, as principais dificuldades encontradas para a modelagem do novo Sistema deram-se quando das referências iniciais ao trabalho de Souza [SOU 94].

Tal sistema, apesar de apresentar várias similaridades com o novo modelo, escondia falhas que não foram observadas de início, mas que sugeriram mudanças importantes para o novo trabalho.

Basicamente, portanto, tais dificuldades foram superadas com a utilização da ferramenta *Eurekha*, que atendeu a todas as exigências necessárias como módulo constituinte do sistema ESCOP.

6.1 Contribuições deste trabalho

Considera-se como contribuições as características que possam servir de futuras referências a novos trabalhos, que venham a comprovar determinadas teorias e que possam aproveitar os trabalhos de outros grupos de estudo e pesquisa.

Neste quesito, este estudo mostrou ser possível a utilização de uma solução para o agrupamento e recuperação de informações - a ferramenta *Eurekha* - utilizada para a organização de informações geradas em atividades de grupo. Tais atividades podem abranger o escopo de trabalhos em EAD - que é o objetivo principal do Sistema ESCOP - ou qualquer outro processo que envolva mais de um indivíduo para a solução de um problema.

Desta forma, considera-se importante que o Programa de Pós-Graduação em Computação da UFRGS mantenha sempre condições para o vínculo entre seus pesquisadores, como foi o caso deste trabalho.

Observa-se que, o ambiente de trabalho e as relações pessoais entre os envolvidos neste projeto contou positivamente para a conclusão do mesmo, conforme observa-se também nas referências obtidas dos pesquisadores da área de Processamento de Linguagem Natural.

Em relação às características de implementação do Sistema ESCOP, uma função crucial deste é a capacidade de sumarizar informações geradas de forma síncrona (uma discussão via *chat*) ou de forma assíncrona (uma lista de discussão por *email*). Esta peculiaridade oferecida pelo *software* possibilita sua utilização em várias situações, a critério de seu usuário.

Uma observação que foi considerada é que o Sistema ESCOP utiliza-se das técnicas de *Brainstorming* e *DELPHI* para a tomada das decisões em grupo, enquanto que o sistema de Souza [SOU 94] utiliza-se das técnicas de *Brainstorming* e Grupo Nominal.

Deste fato, observa-se que foi alcançado o objetivo de continuar, estender e ampliar as capacidades dos trabalhos de Souza [SOU 94].

O presente trabalho também foi submetido e aceito para apresentação em simpósios e congressos internacionais: [VIT 2000a] já apresentado e [TAR 2000 e VIT 2000b] a serem apresentados no primeiro semestre do corrente ano.

6.2 Sugestões e trabalhos futuros

Durante o desenvolvimento deste estudo, primou-se por cobrir todos os objetivos determinados, entretanto, alguns pontos deixaram de ser cobertos. Isso ocorre não só neste, mas em todo trabalho científico, pois nem todos os problemas podem ser resolvidos de uma única vez. Alguns destes problemas devem ser deixados de lado para que outros, mais marcantes, sejam resolvidos.

Conforme alguns comentários iniciais, feitos no item 5.5 deste trabalho, considera-se que a implementação do protótipo na linguagem DELPHI não é a solução ideal para o tipo de solução a que se propôs inicialmente, haja visto a forte dependência

de *hardware* (computadores PC) e *software* (ambiente Windows) requeridos por esta linguagem de programação.

Comentários a respeito de uma possível versão deste ambiente (DELPHI) para o sistema operacional Linux ainda não foram definidas no momento da apresentação desta dissertação. Mesmo neste caso, ainda resta a limitação de *hardware*. Como solução, aponta-se a implementação de uma versão do Sistema ESCOP em linguagem JAVA, que se propõem a ser independente de plataforma de *hardware* e *software*.

Os *scripts* CGI, que nesta versão foram desenvolvidos em linguagem C também poderiam ser portados para JAVA, pelos mesmos motivos descritos para a linguagem DELPHI.

O quesito de programação distribuída não foi implementado nesta primeira versão do *software*. Desta forma, sempre que for executado em um determinado computador, o Sistema ESCOP irá exigir a utilização local de um servidor HTTP, que pode sobrecarregar a máquina utilizada (se esta for um computador portátil, por exemplo), além de conferir-lhe possíveis falhas de segurança (se o servidor HTTP estiver mal configurado pelo usuário). Isto pode ser resolvido através de um remanejamento dos módulos de comunicação do Sistema desenvolvido.

O módulo "Agrupador", que utilizou-se da ferramenta *Eureka*, pode ser melhor explorado, com a determinação de algoritmos específicos para os casos tratados pelo Sistema ESCOP, algoritmos estes que podem ser otimizados especificamente para tais fins.

Um ponto que seria muito bem vindo é a implementação do módulo de agrupamento, haja visto que, apesar de toda a cooperação proporcionada pelo seu desenvolvedor (Wives [WTV 99]), muitas vezes percebeu-se as vantagens de se ter acesso a tal código fonte, que, por motivos pessoais, profissionais e fora do escopo deste estudo, não foi disponibilizado.

O módulo "Filtro I" pode também oferecer maiores opções para o tratamento de documentos que não estejam somente em formato texto.

Considera-se também oportuno, destacar as seguintes possibilidades para trabalhos futuros:

- A capacidade do "aprendizado" semântico do que se está tratando por um grupo de discussões;
- A capacidade de utilização de outras formas de comunicação entre pessoas, que não somente a escrita, mas a falada, por exemplo;
- A capacidade de extração de determinadas características dos atores de um grupo, por exemplo, aqueles que contam piadas; aqueles que mais contribuem para as discussões; e aqueles que participam mas não contribuem para as atividades do grupo;
- A capacidade de cruzamento de informações com um banco externo de "padrões" de idéias, para fins de comparação entre as idéias geradas pelo grupo e as que já foram geradas em outras oportunidades.

Finalmente, considera-se importante observar a utilização do Sistema ESCOP fora do ambiente de EAD, para verificar a sua versatilidade e praticidade de utilização, pois considera-se que sua implementação, conforme demonstrado pelos testes

efetuados, pode abranger um amplo leque de situações, sendo de grande utilidade para os mais variados fins de atividades em grupo.

Bibliografia

- [AIK 92] AIKEN, Milam; CARLISLE, Judith. Na automated idea consolidation tool for computer supported cooperative work. **Information & Management**, Amsterdam, v. 23, n. 6, p. 373-382, Dec. 1992.
- [APA 99] APACHE Web Server. **Apache 1.3 User's Guide**. Disponível por WWW em www.apache.org (nov. 1999).
- [BAN 92] BANNON, Liam J.; SCHMIDT, Kjeld. CSCW: Four characters in search of a context. In: BAECKER, R. M. (Ed.). **Readings in Groupware and Computer-Supported Cooperative Work: assisting human-human collaboration**. San Francisco: Morgan Kaufmann Publishers, Inc., 1992. 882p. p.50-56.
- [BAR 99] BARUA, Anitesh; CHELLAPPA, Ramnath; WHINSTON, Andrew B.. **Creating a collaboratory in cyberspace: theoretical foundation and an implementation**. Disponível por WWW em <http://cism.bus.utexas.edu/ram/papers/joc1/joc.html>. (nov. 1999).
- [BUI 87] BUI, Tung X. **Co-op a group decision support system for cooperative multiple criteria group decision make**. Germany: Springer-Verlag, 1987. 250p.
- [CAT 98] CATTANI, Airton; SILVA, Dinorá Fraga da. Representação do espaço: construir e reconstruir-se numa experiência em educação a distância. **Informática na Educação: teoria e prática**. Porto Alegre, v.1, n.1, p.7-24, outubro 1998.
- [CHE 96] CHEN, Hsinchun et al. **A concept space approach to addressing the vocabulary problem in scientific information retrieval: an experiment on the worm communit system**. Tucson, Arizona: Management Information System Department, University of Arizona, 1996. Disponível por WWW em <http://ai.bpa.arizona.edu/papers/wcs96/wcs96.html> (dez. 1999).
- [COR 91] CORREA, Heriberto Pintos. **DELAHP: uma metodologia de apoio à decisão para grupos aplicado na seleção de um sistema de telecomunicações**. Porto Alegre: CPGCC da UFRGS, 1991. 212p. Dissertação de mestrado.
- [CRU 98] CRUZ, Dulce Márcia; MORAES, Marialice de. **Tecnologias de comunicação e informação para o ensino à distância na integração universidade/empresa**. Disponível por WWW em <http://www.intelecto.net/ead/tecnol.htm> (nov. 1998).

- [CUN 85] CUNHA, Celso; CINTRA, Luis Filipe Lindley. **Nova gramática do português contemporâneo**. 2. ed. Rio de Janeiro: Nova Fronteira, 1985. 724 p.
- [DIE 96] DIETRICH, Elton. **Projeto de um Sistema de Suporte à Autoria Cooperativa de Hiperdocumentos**. Porto Alegre: CPGCC da UFRGS, 1996.
- [FRE 93] FREITAS, Sérgio Antônio Andrade de. **Deíticos e anáforas pronomiais em diálogos**. Porto Alegre: CPGCC da UFRGS, 1993. 112p. Dissertação de mestrado.
- [GRA 99] GRANVILLE, Lizandro Zambenedetti. **Tutorial SDL**. Disponível por WWW em <http://penta.ufrgs.br/redes296/sdl/inic.html> (Dez. 1999).
- [INT 99] INTERLOQ. **Software for cooperative thinking**. Disponível por WWW em <http://www.tarkvara.org/INTERLOQ/> (nov. 1999).
- [KAS 97] KASPARY, Adalberto José. **Português para profissionais: atuais e futuros**. Porto Alegre: Edita, 1997. 235 p.
- [LEI 98] LEITE, Lígia Silva; SILVA, Christina Marília Teixeira da. **A educação à distância capacitando professores: em busca de novos espaços para a aprendizagem**. Disponível por WWW em <http://www.intelecto.net/ead/ligia-cris.htm> (nov. 1998).
- [LEW 98] LEWIS, David. **Distributed course delivery for problem based learning**. Disponível por WWW em <http://edweb.sdsu.edu/clrit/home.htm> (nov. 1998).
- [LOH 99] LOH, Stanley. **Descoberta de conhecimento em textos: exame de qualificação**. Porto Alegre: CPGCC da UFRGS, 1999. (EQ-29).
- [LOY 96] LOYOLA, Afonso Celso Martins. **Delphi 2.0**. Rio de Janeiro: Infobook, 1996. 140 p.
- [LUC 89] LUCKESI, Cipriano Carlos. Democratização da educação: ensino a distância como alternativa. **Tecnologia Educacional**. Rio de Janeiro, v.18, n.89/90/91, p.9-12, jul./dez. 1989.
- [MAG 94] MAGALHÃES, Arthur Valcareggi. **Língua natural: classificação de orações**. Porto Alegre: CPGCC da UFRGS, 1994. 98p. Trabalho de diplomação.
- [NUN 98] NUNES, Ivônio Barros. **Noções de educação à distância**. Disponível por WWW em <http://www.intelecto.net/ead/ivonio1.htm> (nov. 1998).

- [OSI 97] OSIER, Dan. **Aprenda em 21 dias Delphi 2**. Rio de Janeiro: Campus, 1997. 840 p.
- [OTS 98] OTSUKA, Joice Lee. **SAACI - Sistema de apoio à aprendizagem colaborativa na internet**. Porto Alegre: CPGCC da UFRGS, 1998. Dissertação de mestrado.
- [PAD 99] PADILHA, Emiliano Gomes. Disponível por *email* em emiliano@inf.ufrgs.br (maio/ago. 1999).
- [PER 87] PEREIRA, Fernando; TWEED, Christopher. **C-Prolog user's manual**. Edinburg: Department of Architecture, University of Edinburgh, April 1987.
- [PER 94] PERRY, Greg. **Programação orientada para objeto com Turbo C++**. Rio de Janeiro: Berkeley, 1994. 752 p.
- [POD 98] PODGORNY, Marek. **TANGO product support**. Disponível por *email* em marek@npac.svr.edu (01 out. 1998).
- [RIB 98] RIBEIRO, Marcelo Blois. **CLEW: um ambiente de aprendizado cooperativo para a web**. Rio de Janeiro: PUC-RJ - Departamento de Informática, 1998. 116 p.
- [RIT 98] RITZEL, Marcelo Iserhardt. Aspectos básicos para se implementar ensino a distância utilizando sistemas multimídia. **Revista de Estudos**, Novo Hamburgo, v.21, n.1, p.41-64, jan./jun. 1998.
- [SAL 83] SALTON, G.; MCGILL, M. J. **Introduction to modern information retrieval**. New York: McGraw-Hill, 1993.
- [SAR 87] SARACCO, R.; TILANUS, P.A.J. CCITT SDL: overview of the language and its applications. **Computer Networks and ISDN Systems**, Amsterdam, v.13, n.2, p.65-74, 1987.
- [SAV 97] SAVOLA, Tom; WESTENBROEK, Alan; HECK, Joseph. **Usando HTML - O guia de referência mais completo**. Rio de Janeiro: Campus, 1997. 681 p.
- [SOU 94] SOUZA, Dávison Wisniewski de. **Um protótipo da ferramenta "issue analyser"**: projeto, implementação e validação. Porto Alegre: CPGCC da UFRGS, 1994. 176p. Dissertação de mestrado.
- [SOU 97] SOUZA, Paulo Nathanael Pereira de. Ensino a distância: significado e uso. **Cadernos Pedagógicos e Culturais**. Niterói, v.6, n.1/2, p.227-229, jan./dez. 1997.
- [TAN 98] TANGO Product Support. Disponível por WWW em <http://trurl.npac.svr.edu/tango> (set./out. 1998).

- [TAR 2000] TAROUCO, Liane Margarida Rockenbach; VIT, Antônio Rodrigo Delepiane de; GELLER, Marlise; HACK, Luciano. **Supporting Group Learning and Assessment Through the Internet**. Lisboa: TERENA Networking Conference 2000, May 22-25, 2000. Disponível por *email* em secretariat@terena.nl e warren@terena.nl (fev./maio 2000). Disponível por WWW em www.terena.nl/tnc2000/prog.html (fev. 2000).
- [TAR 98a] TAROUCO, Liane Margarida Rockenbach. **Teleducação - introdução**. Disponível por WWW em <http://penta2.ufrgs.br/edu/teleduc/teleduc1.htm> (abr. 1998).
- [TAR 98b] TAROUCO, Liane Margarida Rockenbach. **O que é educação à distância**. Disponível por WWW em http://penta.ufrgs.br/edu/edu1_1.html (dez. 1998).
- [VIT 2000a] VIT, Antônio Rodrigo Delepiane de; TAROUCO, Liane Margarida Rockenbach. Group learning through Internet. In: International Conference on Society for Information Technology & Teacher Education, SITE, 11., 2000, San Diego, CA. **Proceedings...** Charlottesville, VA: AACE, 2000. v.1, p. 143-147.
- [VIT 2000b] VIT, Antônio Rodrigo Delepiane de; TAROUCO, Liane Margarida Rockenbach. **Group activity handling**. Arlington: Education Technology 2000, July 24-26, 2000. Disponível por WWW em www.salt.org/ed_conf/ed_conf_info.htm (fev. 2000).
- [VIT 99a] VIT, Antônio Rodrigo Delepiane de. **ESCOP Estratégias de suporte à solução cooperativa de problemas: trabalho individual**. Porto Alegre: CPGCC da UFRGS, 1999. (TI - 767).
- [VIT 99b] VIT, Antônio Rodrigo Delepiane de; TAROUCO, Liane Margarida Rockenbach. Estratégia de Suporte à Solução Cooperativa de Problemas. In: Semana Acadêmica do PPGC da UFRGS, 4., 1999, Porto Alegre. **Anais...** Porto Alegre: PPGC da UFRGS, 1999. p. 339-442.
- [WAR 99] WAR FTP Daemon 1.70. Disponível por WWW em www.jgaa.com/tftpd.htm (nov. 1999).
- [WEB 98a] WEBOPEDIA. Disponível por WWW em <http://webopedia.internet.com>. (dez. 1998).
- [WEB 98b] WEBSABER. **Um ambiente para solução cooperativa de problemas**. Disponível por WWW em <http://mirror.kidlink.dk/brasil/acd/sabermas.htm> (dez. 1998).

- [WEI 93] WEISKAMP, Keith; HEINY, Loren; FLAMING, Bryan. **Programação orientada para objeto com Turbo C++**. São Paulo: Makron Books, 1993. 473 p.
- [WEI 97] WEINMAN, William E. **Manual de CGI**. São Paulo: Makron Books, 1997. 312 p.
- [WIE 90] WIELMAKER, Jan. **SWI-Prolog 1.6 reference manual**. Amsterdam: Department of Social Science Informatics, University of Amsterdam, 1990.
- [WIV 97] WIVES, Leandro K. **Um estudo sobre técnicas de recuperação de informações com ênfase em informações textuais: trabalho individual**. Porto Alegre: CPGCC da UFRGS, 1997. (TI - 672).
- [WIV 99] WIVES, Leandro Krug. **Um estudo sobre agrupamento de documentos textuais em processamento de informações não estruturadas usando técnicas de "clustering"**. Porto Alegre: CPGCC da UFRGS, 1999. 101p. Dissertação de mestrado.
- [YOU 89] YOUNG, Richard J. **Practical Prolog**. USA: Logicware International, 1989.



PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

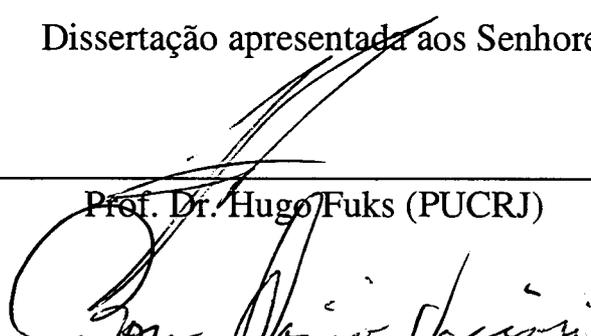
"ESCOPE: Estratégia de Suporte à Solução Cooperativa de Problemas"

por

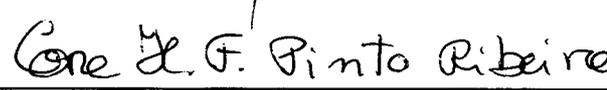
Antônio Rodrigo Delepiane de Vit

Dissertação apresentada aos Senhores:

Prof. Dr. Hugo Fuks (PUCRJ)



Profa. Dra. Rosa Maria Viccari



Profa. Dra. Cora Helena Francisconi Pinto Ribeiro

Vista e permitida a impressão.

Porto Alegre, ___/___/___.



Profa. Dra. Liane Margarida Rockenbach Tarouco,
Orientadora.



Prof.^a Carla Maria Dal Sasso Freitas
Coordenadora do Programa de Pós-Graduação
em Computação - PPGC
Instituto de Informática - UFRGS