

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA
DA COMPUTAÇÃO**

Gilberto Corrêa de Souza

**Proposta de sistema de diálogo textual
independente de aplicação**

Dissertação submetida à Universidade Federal de Santa Catarina como parte dos requisitos para a obtenção do grau de Mestre em Ciência da Computação.

Prof. Ilson Wilmar Rodrigues Filho, Dr.

Florianópolis, Agosto de 2004

Proposta de sistema de diálogo textual independente de aplicação

Gilberto Corrêa de Souza

Esta Dissertação foi julgada adequada para a obtenção do título de Mestre em Ciência da Computação, área de concentração Sistemas de Conhecimento e aprovada em sua forma final pelo Programa de Pós-Graduação em Ciência da Computação.

Prof. Raul Sidnei Wazlawick, Dr.

Banca Examinadora

Prof. Ilson Wilmar Rodrigues Filho, Dr.

Prof. Fernando Álvaro Ostuni Gauthier, Dr.

Prof. Mauro Roisenberg, Dr.

Prof. Marco Antonio Esteves da Rocha, Dr.

A Deus e a minha família.

Agradecimentos

Aos meus pais e irmão que me incentivam sempre a dar passos seguros e acurados no tortuoso e maravilhoso caminho que leva ao horizonte da vida. Ao professor João Bosco da Mota Alves pelo incentivo e oportunidade de conhecimento de um mundo novo.

Ao professor Ilson Wilmar Rodrigues Filho por ter me acolhido e orientado neste caminho cheio de surpresas.

Aos professores: Mauro Roisenberg, Marco Rocha, Fernando Gauthier e João Dovicchi por explicações e contribuições importantes para o desenvolvimento deste trabalho.

Aos amigos e colegas por tornar a vida mais divertida e proveitosa em especial a Ana Castro, Francisco Santos, Andrea Miranda, Gilberto Medeiros, Paula Renata e Lisiane Oliveira.

Sumário

Sumário	v
Lista de Figuras	viii
Resumo	x
Abstract	xi
1 Introdução	1
1.1 Apresentação	1
1.2 Referências iniciais	2
1.3 Problemas em aberto	4
1.4 Objetivos	9
1.4.1 Objetivo geral	9
1.4.2 Objetivos específicos	10
1.5 Metodologia	10
1.6 Organização da dissertação	12
2 Reconhecimento de linguagem natural	13
2.1 Níveis Lingüísticos	15
2.1.1 Nível Fonológico	15
2.1.2 Nível Gramatical ou Sintático	17
2.1.3 Nível Semântico	18
2.2 Processamento de Linguagem Natural	19
2.2.1 Reconhecimento e Geração de Fala	20
2.2.2 Análise Léxico-morfológica	21

2.2.3	Análise Sintática	23
2.2.4	Análise Semântica	25
2.2.5	Análise Pragmática	28
2.3	Aplicações do processamento de linguagem natural	29
3	Sistemas de diálogo	31
3.1	Definição	31
3.2	Abordagens para a codificação de diálogo	31
3.2.1	Abordagens Estruturadas	32
3.2.2	Abordagens baseadas em planos	33
3.2.3	Abordagens baseadas em colaboração	36
3.2.4	Abordagem baseada em interação racional	38
3.3	Fontes de conhecimento	40
3.3.1	Modelo de diálogo	40
3.3.2	Modelo de discurso	41
3.3.3	Modelo de tarefa	42
3.3.4	Modelo de domínio	44
3.3.5	Modelo de usuário	45
3.3.6	Relações entre modelos	48
3.4	Componentes de um sistema de diálogo	49
3.4.1	Entendimento de linguagem	51
3.4.2	Gerente de diálogo	54
3.4.3	Comunicação com sistema externo	56
3.4.4	Geração de linguagem natural	59
3.5	Controle de diálogo	65
3.6	Conclusão	68
4	Trabalhos correlatos	69
4.1	Sistemas de diálogo	69
4.1.1	TRAINS	69
4.1.2	TRIPS (<i>The Rochester Interactive Planning System</i> / Sistema de Planejamento Interativo da Universidade de Rochester)	73

4.1.3	Verbmobil	76
4.1.4	SUNDIAL – <i>Speech UNderstanding in DIAlogue</i>	79
4.1.5	GALAXY	80
5	Sistema proposto	83
5.1	Arquitetura	86
5.1.1	Entendimento de linguagem	87
5.1.2	Gerência de diálogo	107
5.1.3	Comunicação com sistema externo	120
5.1.4	Geração de linguagem	125
5.2	Conclusão	126
6	Considerações	127
6.1	Considerações Finais	127
6.2	Trabalhos Futuros	130
	Referências Bibliográficas	132
	Apêndice A	141
	Exemplo de sistema	141

Lista de Figuras

3.1	Arquitetura de um sistema de diálogo genérico	49
3.2	Arquitetura detalhada de um sistema de diálogo genérico	50
4.1	Sistema TRAINS-95	71
4.2	Sistema TRIPS	74
4.3	Sistema Verbmobil	77
4.4	Arquitetura do SUNDIAL	80
4.5	Arquitetura do GALAXY-II	82
5.1	Arquitetura do sistema de diálogo	87
5.2	Um conjunto de regras sintáticas	90
5.3	Árvore sintática	91
5.4	Um conjunto de regras LFG	92
5.5	Entradas lexicais	93
5.6	Estrutura-c	94
5.7	Instanciação da estrutura-c	95
5.8	Estrutura-c após instanciação	96
5.9	Significado de alguns tipos de equação funcional	97
5.10	Descrição funcional	98
5.11	Estrutura-f	99
5.12	Mapeamento de estrutura-f para estrutura semântica	103
5.13	Atribuição de significados à estrutura semântica	104
5.14	Estrutura-f para a frase “Maria contratou Silvia”	105
5.15	Processo de dedução	106

A.1 Estrutura de um sistema de diálogo genérico 143

Resumo

Este trabalho propõe uma abordagem para o desenvolvimento de um sistema de diálogo. Trata-se de um sistema computacional, uma interface para humanos interagirem de forma mais natural com aplicações que fornecem serviços diversos, tais como: previsão de tempo, consulta de reservas, dentre outras. Para se entender a base de funcionamento desse tipo de sistema, são apresentados os conceitos do processamento de linguagem natural, sistema de diálogo genérico e alguns exemplos de sistemas de diálogo desenvolvidos, explicitando suas características e aplicações. Com base nos conceitos e experiências comprovadas de sistemas anteriores, apresenta-se a especificação de um modelo de diálogo que possui dois componentes-chaves, a saber: um componente capaz de gerenciar o diálogo por meio de princípios racionais e cooperativos e outro, cujo uso da noção computacional de ontologia toma lugar na comunicação entre usuários e aplicações fornecedoras de serviços. Ao final deste trabalho a proposta de um sistema de diálogo é apresentada, entretanto sua implementação é apenas sugerida por um exemplo, visto que o desenvolvimento de um sistema de diálogo completo estava fora dos objetivos do presente trabalho.

Palavras-chave: Processamento de Linguagem Natural, Sistema de Diálogo, Ontologia.

Abstract

This work presents an approach to the development of a dialogue system. It proposes the use of a computational system—an interface allowing human beings to interact more naturally with applications that perform a variety of tasks, such as giving weather forecasts and checking hotel reservations, among others. In order to provide some understanding of the functional basis of this kind of system, the concepts of processing of natural language and generic dialogue systems will be presented, along with some examples of dialogue systems already developed, their characteristics and their applications. With the support of concepts and evidence derived from experience with previous systems, the specifications for a model of dialogue will be presented. That model shall feature two key components, namely: one component that is capable of managing dialogue through rational, cooperating principles; and one in which the computational notion of ontology will have a place in the communication between the users and the applications performing tasks. By the end of the work the proposal of a dialogue system is presented; however, its implementation will only be suggested through an example, since the development of a complete dialogue system was not one of the goals of the present work.

Capítulo 1

Introdução

1.1 Apresentação

A comunicação entre pessoas e computadores é um sonho antigo dos pesquisadores da área de inteligência artificial. Essa comunicação é a mesma que ocorre entre duas ou mais pessoas. Um marco à perseguição desse sonho, ocorreu quando Allan Turing em [1] propôs uma questão: “Máquinas podem pensar?”, e um jogo: “Jogo da imitação”. Esses elementos originaram o chamado Teste de Turing, cujo objetivo consiste em verificar se um sistema computacional tem a capacidade de dialogar com um humano de modo que o mesmo não perceba que está dialogando com uma máquina. Contudo, a idéia de Turing era medir a inteligência do sistema em termos de produção de linguagem.

O teste de Turing pode verificar a fluência de um sistema computacional em uma linguagem natural¹. A fluência ou processamento em linguagem natural consiste em uma aplicação das pesquisas em inteligência artificial. Seu objetivo principal é capacitar um sistema a entender e gerar fala ou escrita de uma linguagem natural, tão bem quanto um humano.

¹Uma linguagem pode ser natural, quando se refere a uma língua falada e/ou escrita por um grupo de pessoas; ou artificial quando um grupo de pessoas cria uma linguagem específica para determinadas comunicações.

A busca a possibilitar um computador dialogar como o homem possui a mesma intensidade motivadora a de Turing quando este fez o questionamento sobre a inteligência dos computadores. Desse modo, buscar o aperfeiçoamento de sistemas conversacionais (estes serão tratados a partir deste ponto por sistemas de diálogo) é um dos objetivos das pesquisas em inteligência artificial na tentativa de proporcionar uma simulação do comportamento inteligente do homem.

A pesquisa em sistemas de diálogo evoluiu de forma semelhante às demais áreas da inteligência artificial. Cada nova teoria e tecnologia possibilitaram o desenvolvimento de sistemas capazes de entender e gerar os sons da fala humana, interpretar e responder determinadas formas de perguntas, além de possibilitar uma busca por informações em bancos de dados de um modo mais próximo a que um ser humano a faria junto a outro ser humano.

Além da inteligência artificial, outras áreas do conhecimento serviram de base para o desenvolvimento dos sistemas de diálogo. Áreas tais como: lingüística, responsável pela descrição da estrutura da linguagem natural humana; psicologia, no sentido de entender os processos mentais que levam a produção e interpretação da linguagem; e filosofia, que tem seu papel apresentado por Searle da seguinte forma: [2] p. 4. “A filosofia da linguagem é uma tentativa de fornecer descrições iluminatórias filosóficas a certas características da linguagem”; apenas citando algumas delas. Este trabalho considerará as três primeiras áreas como pilares ao desenvolvimento de sistemas de diálogo.

1.2 Referências iniciais

Em alguns dos primeiros textos sobre Inteligência Artificial e Processamento de Linguagem Natural, pode-se interpretar o desejo humano de fazer com que computadores possam conversar com seus criadores.

As pesquisas voltadas para esse fim podem ser divididas em duas linhas: uma, onde

se utiliza modelos teóricos desenvolvidos por estudos nas diversas áreas contidas na área de Inteligência Artificial; outra onde, se busca um entendimento mais amplo dos fatores lingüísticos envolvidos em conversações humanas [3].

Alguns dos primeiros programas que simulavam certos comportamentos humanos durante uma conversa foram:

- PARRY² foi desenvolvido pelo psiquiatra Kenneth Colby e se prestava a simular o comportamento de uma pessoa com esquizofrenia paranóica. Sua implementação contava com um modelo comportamental simples baseado em conceitos, conceitualizações e crenças sobre essas conceitualizações de um doente esquizofrênico paranóico. Sua estratégia de conversação era mais elaborada do que a empregado no ELIZA;
- ELIZA³ foi desenvolvido por Joseph Weizenbaum e tentava imitar o comportamento de um psiquiatra rogeriano e se tornou um dos mais famosos programas de computador. Ele se baseava simplesmente no reconhecimento de padrões e substituição de palavras-chave em sentenças pré-determinadas. Com essa estratégia simples o programa poderia ser confundido com uma pessoa real;
- Outro programa é o SHRDLU⁴ que foi desenvolvido por Terry Winograd e permitia a um usuário instruí-lo por meio de termos em inglês sobre a movimentação de objetos dentro de um ambiente de blocos. Seu grande diferencial era o domínio que facilitou a escolha de um conjunto de palavras relativamente pequeno e ao mesmo tempo suficiente para lidar com várias formas de sentenças. Também utilizava a idéia de memória para armazenar termos anteriormente proferidos e dessa forma tornar o diálogo mais convincente.

Os programas apresentados acima visavam simular o comportamento humano em

²Kenneth Mark Colby, Sylvia Weber, Franklin Dennis Hilf. Artificial Paranoia. *Artificial Intelligence*, 2(1): 1–25 (1971)

³Joseph Weizenbaum. ELIZA – A Computer Program for the Study of Natural Language Communication between Man and machine. *CACM*, Vol. 10, 1967.

⁴Terry Winograd. Procedures as a Representation for Data in a Computer Program for Understanding Natural Language. MIT AI Technical Report n° 235, Fevereiro de 1971

conversao. Existem outros programas com o mesmo objetivo, alguns deles sero descritos com mais detalhes no decorrer deste trabalho.

Barr, [4], apresenta uma introduo sobre os primeiros sistemas computacionais que pretendiam entender e simular linguagem natural.

1.3 Problemas em aberto

Apesar da evoluo das pesquisas para o desenvolvimento de melhores sistemas de dilogo, alguns problemas ainda permanecem sem solues satisfatrias. Um dos motivos, contudo no o mais importante, para eles continuarem sem solues  a natureza dinmica da linguagem natural, isto , ela est sempre se transformando, seja por meio de adio ou eliminao de palavras, seja por atribuies de novos sentidos a palavras existentes.

Cabe ressaltar que mesmo uma linguagem no dinmica apresenta dificuldades semelhantes as dinmicas, tais como: ambigüidade, co-referncia, entre outras. Isto , alm da questo dinmica de uma linguagem existem outras caractersticas das lguas humanas que podem e de fato dificultam o seu processamento computacional.

Isso expressa a necessidade de busca por mtodos mais flexveis e adaptativos no processo de resoluo de problemas na rea de sistemas de dilogo: uma vez que o ser humano possui a capacidade de domnio lingüstico de uma linguagem, enquanto que os sistemas atuais no a possuem no mesmo nvel.

Alm da busca por novos mtodos existem outros desafios a serem compreendidos para o processamento computacional de linguagens naturais. Allen em [5] indica quatro desafios ainda pendentes para a conversao prxima da ideal entre homem e computador. Outros obstculos como a ambigüidade, por exemplo, so ainda um problema srio para o processamento de linguagem, contudo os desafios apresentados a seguir podem ser estudados de forma mais pontual. Em outras palavras, apesar dos

problemas a serem apresentados, existem outros com mais ou menos importância para o processamento de linguagem natural, todavia os seguintes serão objeto de estudo deste trabalho.

- Complexidade da linguagem associada ao domínio de conhecimento da tarefa associada ao sistema de diálogo.
- Processos envolvidos na comunicação do sistema de diálogo com uma aplicação ou módulo externo.
- Reconhecimento das intenções como fator fundamental para o processo de entendimento lingüístico⁵.
- Possibilitar iniciativa mista no sistema⁶.

Primeiro, um sistema de diálogo em algum momento de seu desenvolvimento necessitará da definição de seu propósito, ou seja, a que atividade ele se presta (tarefa do sistema). Esta dependerá do domínio a que o sistema estará associado, por exemplo, um sistema de diálogo para contadores terá de levar em consideração alguma atividade que um contador queira executar em seu domínio.

Dentro do domínio da atividade ou tarefa existem certas informações lingüísticas que são específicas ao domínio, por exemplo, palavras comuns ou não à uma língua que são utilizadas de forma a espelhar os significados característicos do domínio em questão, ou seja, um léxico próprio e expressões com significados e usos particulares em um dado domínio.

⁵O reconhecimento apresentado aqui diz respeito a tentativa de reconhecer as intenções que levam a produção de enunciações lingüísticas, isto antes de seu processamento computacional, uma vez que a determinação das intenções antes mesmo da proferição de novas enunciações pode acarretar em problemas de interpretação (circularidade no processamento). Por este motivo, este trabalho não tem por objetivo propor uma reconhecimento de intenções como forma de processamento computacional.

⁶A iniciativa em um diálogo pode ser entendida como o conjunto de vezes que uma pessoa inicia ou introduz novos tópicos ou idéias. Assim, em um diálogo pode-se ter a iniciativa da pessoa que fala, neste contexto iniciativa ou do usuário ou do sistema, e iniciativa mista, onde tanto usuário quanto sistema poderão tomar a iniciativa, mas não para durante todo o diálogo.

Assim, o sistema além de necessitar de conhecimentos que o façam compreender a tarefa do usuário, terá ainda de possuir informações sobre o uso específico de termos relativos ao domínio em que a tarefa do usuário está associada.

A dificuldade levantada por este ponto reside no fato de que sistemas de diálogo realizam tratamento lingüístico geral sobre as emissões lingüísticas do usuário, ou seja, eles consideram padrões genéricos que uma linguagem apresenta, os quais são expressos por regras gramaticais simples que estabelecem significados gerais e abrangentes aos termos que podem ser utilizados em cada emissão lingüística, tanto do usuário quanto do próprio sistema.

Isto quer dizer que quando uma pessoa fala algo ela, a princípio, segue regras de construção, as quais possibilitaram uma interpretação segura da fala. Todavia, não existem regras gerais para se tratar construções específicas à cada domínio de conhecimento, sendo necessário uma codificação de regras específicas. Logo, para cada novo domínio que o sistema tenha que trabalhar (nova tarefa), ele necessitaria interpretar um léxico específico para o novo domínio.

Esta capacidade de adequação a cada novo léxico de tarefa ainda não foi alcançada de forma satisfatória pelos sistemas de diálogo atuais. Ressalta-se ainda a dificuldade de tratamento lingüístico geral, isto porque o conhecimento independente de domínio também mostra obstáculos tão representativos quanto os apresentados por domínios específicos.

Segundo ponto, fora a dificuldade lingüística apresentada anteriormente, um sistema de diálogo teria ainda que lidar ou se adaptar a tarefas diferentes. Idéia semelhante à de que o ser humano pode realizar tarefas em domínios diferentes com a mesma desenvoltura. Então a separação dos conhecimentos específicos a cada domínio possibilitaria a um sistema lidar com tarefas diferentes. Algo semelhante a uma técnica de solução de problemas de dividir para conquistar⁷.

⁷Dividir para conquistar é uma abordagem algorítmica bastante difundida na área de inteligência

Logo a comunicação com uma fonte de conhecimento externa que conteria as informações necessárias para a realização de uma tarefa, possibilitaria ao sistema trocar de fonte de conhecimento e dessa forma trocar de domínio. Contudo, para a troca de domínios, o sistema teria de possuir um mecanismo genérico para o tratamento de tarefas diferentes. Esse mecanismo teria que assegurar uma separação satisfatória da dependência de conhecimento específico de uma tarefa do conhecimento que pode ser generalizado a um conjunto de tarefas semelhantes.

Em relação ao terceiro ponto, é adotada a crença de que a comunicação escrita ou falada é governada pelas intenções de seus participantes. Dessa forma, o reconhecimento das intenções do usuário é um fator importante para o reconhecimento e geração de falas em um diálogo.

A dificuldade encontrada ao se adotar essa crença é causada pela necessidade do sistema reconhecer corretamente as intenções do usuário. Pois, a descoberta das intenções necessita de raciocínio sobre uma tarefa e da interpretação que leve em consideração um contexto em particular. Allen, [5], diz que as técnicas que se apóiam somente na especificação da forma da linguagem serão limitadas por não relevarem a complexidade da linguagem. Complexidade que pode ser superada com a consideração do contexto e do conteúdo no tratamento da especificação de modelos para uma representação formal da linguagem.

A adoção da representação do processo comunicativo baseado na intencionalidade pode ser motivo de debates, contudo, seu uso vem da tentativa de expressar a idéia de que as palavras transmitem ou exercem uma ação, o que foi observado por Austin e aprimorado por Searle para a expressão de uma teoria de atos ou ações de fala. A

artificial. Ela consiste em dividir recursivamente ou não um problema em dois ou mais sub-problemas (semelhantes) até que estes possam ser resolvidos sem necessidade de novas divisões. Assim cada solução desses sub-problemas será combinada para gerar a solução final para o problema inicial. A vantagem dessa abordagem é a diminuição das dificuldades de solução de um problema complexo. Desvantagem é o custo que a recursão atribui a velocidade de resolução (diminuição).

partir dessa tentativa de expressão surgiram teorias que tentavam formalizar o processo comunicativo segundo os estudos desses dois filósofos. Um trabalho marcante sobre a formalização da teoria de atos de fala, e mais amplamente para uma teoria da comunicação foi o desenvolvido por Cohen e Levesque [6].

A adoção de uma representação formal para a questão da intencionalidade é motivo para debates entre os pesquisadores da área de processamento de linguagem natural. Isto sobre o ponto de vista dos pesquisadores da lingüística e da filosofia, onde considera-se não se considera a noção da representação formal um meio adequado para a resolução do problema da intencionalidade. Contudo, apesar de seus argumentos este trabalho seguirá a direção da formalização do problema, ressaltando que o modelo teórico a ser adotado neste trabalho de fato não explica de forma satisfatória as características desejáveis de uma possível solução para a produção intencional de fala.

Apesar de resultados expressivos terem sido apresentados, alguns problemas ainda permaneceram sem solução aparente. Tais problemas eram originados pela noção de escolha correta de atitudes mentais, uma delas a intenção, mais apropriadas para a composição de um núcleo que expressa-se os elementos mínimos da comunicação. E pelo próprio formalismo adotado, uma vez que certas idealizações não condiziam com fenômenos reais da comunicação humana.

Por fim, chega-se ao problema de manutenção de um diálogo coerente e natural, ou seja, algo semelhante a uma conversa entre duas pessoas. Para isso ocorrer o sistema necessita considerar alguns requisitos que lhe permitam entender seu papel no decorrer de um diálogo. Em um diálogo um participante pode ser alguém que dite o desenrolar do diálogo, comandando as direções das falas, ele pode ser também alguém que segue as direções que vão se estabelecendo durante o diálogo, ou ele pode ser ainda aquele que tanto estabelece quanto segue algumas direções que vão se estabelecendo durante um diálogo.

No primeiro caso diz-se que o participante teria a iniciativa das ações em um diálogo, no segundo quem a teria seria o outro participante e no terceiro ambos possuiriam em graus equivalentes ou não a iniciativa sobre o diálogo. Em diálogos entre sistema de diálogo e usuário o sistema pode iniciar o diálogo e manter a iniciativa ou esperar pela iniciação do usuário. Nessas duas situações, o sistema possui uma estrutura simples de controle de diálogo, algo que ele pode compreender. Com essa compreensão superficial de um diálogo, o sistema poderia simular certos comportamentos humanos em diálogo dessa natureza.

Todavia, em diálogo onde a iniciativa é mista a arquitetura do sistema de diálogo tem de estar preparada para lidar com falas que podem ou não seguir um fluxo constante de raciocínio. Para lidar com esse problema os sistemas consideram alguns fatores que podem ser observados em diálogos entre humanos dessa natureza.

Esses fatores dizem respeito à definição de fim da emissão do usuário, tratamento de informações extras, as quais não são necessárias diretamente ao diálogo, mas que atribuiriam naturalidade e dinâmica ao diálogo, e controle do andamento da tarefa, considerando que o sistema estaria tentando alcançar um objetivo comum ao usuário (cooperação). Existem outros fatores que estariam ligados a manutenção de diálogo e que acrescentariam mais dificuldades as tentativas de produção de sistemas de diálogo como iniciativa mista.

1.4 Objetivos

Tendo em vista o objeto de estudo desta dissertação, os objetivos deste trabalho desdobram-se em um objetivo geral e três específicos, quais sejam:

1.4.1 Objetivo geral

Propor uma abordagem que apresentará algumas soluções, as quais podem ser expressas no desenvolvimento de um sistema de diálogo escrito. Este trabalho considerará

as dificuldades apontadas anteriormente como ponto de partida para a tentativa de busca de soluções para algumas das mesmas.

Em relação ao desenvolvimento do sistema, quer se dizer que neste estágio somente soluções que possam ser geradas e utilizadas sem a necessidade de estudos mais aprofundados serão consideradas, ou seja, apesar do desejo de apresentar algo concreto este trabalho tentará apontar soluções sobretudo teóricas.

1.4.2 Objetivos específicos

Os objetivos específicos deste trabalho são:

- Apresentar um modelo de sistema de diálogo escrito simples para uma língua específica;
- Especificar uma descrição parcial do processo de comunicação utilizando o formalismo para interação racional cooperativa definido por Sadek;
- Descrever a comunicação entre sistema de diálogo e aplicação externa por meio de uma linguagem ontológica pode levar a independência de domínio.

1.5 Metodologia

Como primeiro passo esta pesquisa leva em consideração informações sobre o desenvolvimento e pesquisas na área de sistemas de diálogo, também se apropriando de idéias afins de áreas que contribuam para soluções ainda não exploradas.

Um levantamento preliminar de trabalhos correlatos indica o que se tem feito e o que é oferecido como teoria para o começo do levantamento de viabilidade da constru-

ção de um sistema de diálogo, uma pesquisa bibliográfica seria necessária, contudo a praticidade que se tentará adicionar a pesquisa iniciada por este trabalho considerará todas as possíveis ocorrências disponibilizadas pela literatura dessa área. De posse de material introdutório, parte-se para a definição de metas de pesquisa e desenvolvimento.

O passo seguinte é o estabelecimento de referenciais teóricos, os quais auxiliarão no entendimento da área de sistemas de diálogo e no esclarecimento de questões sobre possíveis técnicas a serem utilizadas no desenvolvimento desses sistemas. Portanto, para se estabelecer o lugar destinado aos sistemas de diálogo em um plano mais abrangente da literatura pesquisada, faz-se necessário uma introdução ao processamento de linguagem natural, área na qual eles estão inseridos.

Após essa introdução ao processamento de linguagem natural, pode-se iniciar a apresentação do que é um sistema de diálogo e de suas características funcionais. Note-se que essa introdução não tem o objetivo de ser completa, todavia apenas ser uma forma de introduzir o assunto e tentar esclarecer o tema principal deste trabalho.

Uma vez que se tem conhecimentos sobre o processamento de linguagem natural, sistema de diálogo e referências concretas advindas de sistemas correlatos, então se parte para a definição do objetivo principal desta pesquisa, ou seja, comentar as idéias e os conhecimentos adquiridos durante o levantamento de material teórico para então apontar soluções próprias que visem encaminhar uma pesquisa para a apresentação de uma proposta de um sistema de diálogo escrito.

Um exemplo incompleto da proposta para tal sistema será mostrado em anexo ao texto. Uma vez que o objetivo do presente trabalho não é apresentar um exemplo funcional de um sistema de diálogo, mais sim fornecer uma base simples para um futuro desenvolvimento do mesmo.

1.6 Organização da dissertação

No próximo capítulo será apresentada uma visão sobre a área de processamento de linguagem natural, sua divisão básica e algumas aplicações.

O capítulo seguinte fornecerá uma contextualização e introdução à área de sistemas de diálogo, classificações, componentes e tipos de conhecimento necessários para se modelar o diálogo entre homem e computador.

Alguns trabalhos na área de sistemas de diálogo falado ou escrito serão apresentados no capítulo posterior.

A apresentação de uma arquitetura de sistema de diálogo escrito com seus componentes e teorias fundamentais, a qual é a semente desta pesquisa, será apresentada no capítulo após os trabalhos correlatos.

Considerações sobre esta proposta e possíveis desdobramentos futuros decorrentes das primeiras experiências adquiridas durante o processo de construção deste texto e da amostra prática de um sistema de diálogo serão comentadas em capítulo final.

Em anexo é apresentado um exemplo simples e não funcional do que se pretende caso a presente proposta fosse de fato implementada.

Capítulo 2

Reconhecimento de linguagem natural

Para a tentativa de reconhecimento e interpretação das informações que estão contidas nas emissões proferidas em linguagem natural, sejam elas faladas ou escritas, é necessário se fazer um estudo sobre a natureza dessas emissões. Isso é feito através do estudo da comunicação, seja por convenções (língua) ou por uso de seus usuários (fala), onde se entenderão os mecanismos de uma linguagem.

A lingüística é o estudo de uma linguagem, tendo como objetivo básico determinar a sua natureza por meio do estudo de sua estrutura e funcionamento [7, 8]. Uma linguagem, sobretudo a humana, pode ser descrita segundo Saussure, em Barthes [9], por um princípio, o princípio da dupla articulação. Onde na primeira articulação estão as palavras, possuidoras de sentido, e na segunda articulação estão os sons (fonemas), os quais participam da formação das palavras, porém não possuem sentido associado.

Sendo assim, a linguagem pode ser entendida como a separação dessas duas articulações. Logo, dada uma língua qualquer, a partir de um conjunto finito e reduzido de sons, seria possível produzir uma quantidade diversa e numerosa de palavras. A união disso evidenciaria a linguagem por trás dessa língua e por conseqüência, expressa pela fala.

A análise ou descrição lingüística de uma linguagem constitui-se da distinção feita por Saussure entre língua e fala¹, onde língua representa um conjunto ou sistema finito de regras e convenções (norma) estabelecido independentemente de seu uso individual. Fala seria o conjunto infinito baseado nessas regras, isto é, o uso de fato da norma em casos particulares [9]. Um exemplo disto seria a seguinte analogia: um livro qualquer pode ser entendido como a língua dentro do mundo literário, a linguagem.

O conjunto de regras é dividido em duas partes: gramática e vocabulário, onde a gramática procura descrever as unidades significativas (palavras) e suas relações de combinação (frases) [10, 8, 7]. E o vocabulário estabelece de modo semelhante uma descrição para as unidades distintivas (sons). Assim os componentes da língua fornecem os meios para limitar e ao mesmo tempo esclarecer os seus falantes sobre a combinação dessas unidades (significativas e distintivas) na fala.

Um estudo mais detalhado da estrutura de uma linguagem considera normalmente três aspectos distintos: o significado das palavras, a associação de sons a esses significados² e a combinação de palavras para a formação de frases [8], que aqui corresponderão respectivamente aos níveis lingüísticos: semântico, fonológico e gramatical [7]. É possível ainda mencionar outro nível, o pragmático, que considera as relações entre as combinações de palavras e frases, as quais formará uma estrutura mais complexa denominada discurso.

Outra forma de estudo da linguagem considera sua divisão em dois planos: plano dos significantes e plano dos significados. Em ambos existem explicações lingüísticas (forma) e não lingüísticas (substância) sobre a natureza das unidades da linguagem. Onde no primeiro plano podem-se unir esses elementos (forma e substância) para ex-

¹Língua, neste trabalho, se refere a *langue* e fala a *parole*, ambas traduções dos termos originalmente utilizados por Saussure.

²A associação de sons ao significados das palavras não é a única forma de se ver o que representa o nível fonológico, outra forma é considerar o processo de delimitação dos sons específicos de uma língua qualquer.

plicar as regras que descrevem a combinação de signos³, regras sintáticas e, ainda, a natureza articulatória da produção de sons, fonética. Já outro plano, o dos significados, está unido os mesmos tipos de elementos, porém sua explicação esclarecerá os aspectos referentes ao significado das unidades da linguagem [9].

2.1 Níveis Lingüísticos

2.1.1 Nível Fonológico

O estudo de uma linguagem pode relacionar ou estabelecer as correspondências entre as unidades significadas (palavras) e unidades distintivas (sons ou representações fonéticas), onde as palavras expressam sentido, seja ele conceitual ou psicológico, que representa a divisão de significado (conceito) e significante (meios para representação do significado) [9]. As representações fonéticas determinam a forma de pronúncia dessas palavras [8].

Essas correspondências são estudadas por duas linhas: na primeira, fonética, os estudos são referentes aos sons da fala, divididos no estudo da produção dos vários sons da fala humana levando em consideração o aparelho fonador humano (pronúncia das palavras) [11, 7], nas características físicas do sons da fala e da forma de percepção destes sons. Esta linha pode ser entendida como aquela que estuda a substância no plano dos significantes. Na segunda, fonologia, tem-se o objetivo de determinar a natureza das representações fonológicas de uma língua, ou seja, sua forma. Por meio da definição da função dos elementos sonoros, fonemas, na formação de diferentes significados (signos) a partir do plano dos significantes.

³Um signo é uma unidade de significado que se refere a outra unidade significativa. Para Saussure “The linguistic sign unites not a thing and a name, but a concept and a sound-image. The latter is not the material of sound, the impression that it makes on our senses: the sound-image is sensory, and if I happen to call it ‘material,’ it is only in that sense, and by way of opposing it to other terms of the association, the concept, which is generally more abstract.”(SAUSSURE, p 98, 1992. *Cours de linguistique générale*. Versão em inglês.).

Segundo Langacker, [12], essas representações fonológicas podem ser entendidas sob três aspectos: o do som, o da palavra e o do enunciado. Quanto ao do som, a fonologia tem como elemento de estudo os fonemas, unidades distintivas, de uma língua, que são representações simbólicas atribuídas a sons específicos produzidos na fala. Esses símbolos são representados por sinais gráficos denominados letras. Ou ainda, a partir de uma mesma substância pode-se atribuir duas formas diferentes, uma fônica e outra gráfica [9].

Uma língua apresenta um conjunto finito de fonemas que podem ser produzidos pelo aparelho fonador durante a fala. Nem sempre o número de fonemas corresponde ao número de letras que fazem parte de uma palavra. Por exemplo, na palavra *demonstrar* existem 10 letras e 9 fonemas, então, seja L for um conjunto de letras de uma palavra, para demonstrar $L = \{d, e, m, o, n, s, t, r, a, r\}$, e seja F um conjunto de fonemas de uma palavra, então $F = \{/d/, /m/, /o/, /n/, /s/, /t/, /r/, /a/, /r/\}$.

Em relação ao segundo aspecto alguns conjuntos de sons específicos representados por meio de fonemas quando combinados atribuem ou modificam o sentido destes conjuntos [9], por exemplo, a troca do fonema $/s/$ pelo fonema $/z/$ causa a modificação de sentido (caça \rightarrow casa). Esses conjuntos são então representados por várias letras e a reunião dessas letras permite interpretações gráficas desses conjuntos, as quais são denominadas palavras.

O último aspecto considera que as representações fonológicas podem ser utilizadas com algumas variações, as quais são caracterizadas por certos fenômenos da produção de fala: enunciado, entonação e pausa. Esses aspectos podem ser entendidos pela divisão de Hjelmslev da língua em três planos, no primeiro se define a forma pura, como a definição fonológica de uma letra, no segundo está a relação da forma material, como pelo uso oficial ou característico falado da mesma letra e por fim no último nível, o hábito, no qual a letra pode ser pronunciada de forma regional. Por exemplo, o ‘r’ definido por uma gramática, o ‘r’ pronunciado em um país (capital) e o ‘r’ pronunciado

por uma região do país (interior) [9].

Existem regras que servem a classificação e agrupamento dos diversos fonemas de uma língua. Essas regras definem unidades denominadas vogais (classificação de acordo com pronúncias de fonemas) e sílabas (conjunto de fonemas).

2.1.2 Nível Gramatical ou Sintático

O estudo dos sons e suas relações com as formas das palavras foi a primeira parte do estudo da linguagem. O prosseguimento desse estudo envolve as relações lógicas que existem entre as palavras.

Uma forma de iniciar esta continuação de estudo é admitir e definir elementos significativos mínimos que compõem as palavras, ou seja, sua parte significante. Uma vez que, mesmo uma palavra pode ter uma parte significante e outra significativa. Assim, pode-se estudar a estrutura e regras que definem a criação das palavras. Onde os elementos que compõem a porção significante de uma palavra são os morfemas (radical, afixos, desinências, tema, vogais e consoantes), que representam assim, as unidades significativas derivadas dos fonemas. Este estudo é denominado morfologia.

As palavras podem ser agrupadas em certas categorias, de acordo com as suas partes significantes, denominadas categorias gramaticais, as quais identificam qualidades comuns a um grupo de palavras (verbos, substantivos, adjetivos, etc.) em relação aos elementos que as formam. Essas categorias podem ser agrupadas dentro de outros conjuntos, os sintagmas. Sintagmas representam a combinação de signos (produto da união de significante e significado, isto é, significação) ou porque não de palavras [9]. Dessa forma os sintagmas poderiam expressar conjuntos de categorias gramaticais, por exemplo: sujeito, predicado nominal, predicado verbal, predicado adverbial, etc.

As relações lógicas e combinatórias estabelecidas entre palavras e expressas por regras gramaticais são denominadas regras sintáticas. Estas regras permitem o reco-

nhecimento de certos fatores que determinam a estrutura de uma frase (concordância⁴, interpretação⁵, etc.) [11]. Tais regras são usadas em um processo que determinará as categorias gramaticais de cada palavra em uma frase, tal processo é denominado etiquetagem por Vieira, [11].

Portanto, uma frase pode ser construída pela combinação de unidades significativas (palavras), onde suas partes significantes estão ligadas sintaticamente no plano dos significantes por meio de regras. As quais restringem as relações lógicas que permitem a união com suas partes significativas para a definição final do sentido expresso. Logo a reunião de várias unidades significativas por meio de um processo de significação (união entre significante e significado) permite a posterior definição ou combinação daquelas para a formação de frases [9].

2.1.3 Nível Semântico

De posse das informações estruturais (sintáticas) das frases resta buscar seu entendimento semântico. Isso pode ser alcançado pela interpretação dos significados das palavras que a compõem e de como elas estão dispostas, isto é, a estrutura da frase.

A interpretação semântica tenta expressar o conteúdo da comunicação contido em frases. Dessa forma, para se entender o que uma pessoa quer transmitir é necessário entender e conhecer os significados ou conceitos psicológicos das palavras e como estas são agrupadas. Ou seja, para se estabelecer o entendimento das palavras que uma pessoa fala é, primeiro, necessário o conhecimento das unidades distintivas da língua e entender sua combinação como forma de estabelecimento de sentido nas palavras, unidades significativas.

⁴Mecanismo de adaptação das terminações de palavras para a harmonização de uma frase. Ex: **Regra:** Um verbo concorda com o sujeito mesmo que este esteja deslocado. – “*Faltaram*, naquele dia, quatro *funcionários*.”

⁵Possibilidade de que combinações entre elementos em uma frase poder gerar interpretações variadas (ambigüidade), em decorrência da estrutura da frase. Ex: “Uma mulher viu um rapaz com óculos.”

Segundo, saber que as palavras possuem duas partes, uma significante e outra significada. Onde na parte significante existem as noções de formação da palavra e na parte significada está o que estabelece sua ligação com um conceito, significado. Esse conceito foi adquirido e modificado pela cultura lingüística de uma população. Assim, para Saussure tal conceito estabelecido pelo significado liga uma palavra a sua imagem psicológica [9]. Um exemplo disso seria o significado da palavra mesa não tem a ver com sua forma real, isto é, a palavra mesa não determina a forma física de uma mesa, assim é necessária uma associação entre a palavra mesa e o conceito psicológico do objeto físico mesa.

Portanto a construção da interpretação semântica pode ser obtida pela categorização de palavras, as quais expressam categorias do conhecimento. E dessa forma, a combinação dos conceitos expressos por classes de conhecimento (categorias do conhecimento) permitirá a construção da comunicação.

2.2 Processamento de Linguagem Natural

Com as noções de lingüísticas apresentadas parte-se para sua utilização em uma área denominada lingüística computacional. Uma forma de introduzir o objeto de estudo dessa área (processamento), é considerando a premissa de que a gramática de uma linguagem representa a sua descrição lingüística [12]. Assim, o processamento de uma linguagem por meio de computadores espelha as áreas do estudo lingüístico (fonologia, morfologia, sintaxe, semântica e pragmática) [11] apresentadas informalmente como níveis lingüísticos. Portanto, lingüística computacional é uma área de estudo que entre outras áreas afins une elementos da lingüística e da computação. Onde uma de suas aplicações ou área subjacente é o processamento de linguagem natural.

O processamento de uma linguagem natural pode seguir o caminho do estudo de expressões faladas, textos escritos e a noção de uso da linguagem por falantes nativos

[12], para gerar teorias que expliquem o processamento humano da linguagem e dessa forma atribuir a sistemas computacionais a capacidade humana de entendimento e geração de linguagem natural.

Porém existem obstáculos impostos pela própria linguagem ao seu processamento completo por meio de computadores. Um deles é a natureza biológica ainda não revelada da linguagem, pois não se sabe como ela funciona no ser humano⁶. Outro seria a natureza incomensurável de uma linguagem, onde não é possível determinar uma gramática que represente todo o universo de sentenças produzíveis por uma linguagem.

A manipulação que a linguagem sofre pelos seus usuários é também um obstáculo, pois as adaptações estilísticas, sociais e econômicas acrescentam novas palavras, modificam significados, escrita ou pronúncia de palavras antigas.

Contudo, as pesquisas tentam contornar tais adversidades, porém apenas o entendimento do funcionamento da linguagem no cérebro humano poderá indicar a melhor forma de se realizar o processamento de linguagem natural por meio de computadores.

As etapas usuais para o processamento de linguagem natural serão descritas sucintamente a seguir.

2.2.1 Reconhecimento e Geração de Fala

O reconhecimento e a geração de fala são etapas do processamento de linguagem natural responsáveis respectivamente: pelo processo que começa com a aquisição da fala humana, passando pela interpretação e codificação dos sinais obtidos na aquisição em um modelo capaz de suprir as necessidades computacionais de processamento de uma língua. A partir de um modelo fonológico baseado nas características de um língua, pode-se tentar uma tradução de uma codificação interna para sinais que se

⁶Os estudos em processamento de linguagem natural são baseados em noções de como entendemos a linguagem, não existindo ainda uma representação científica para tal fenômeno.

assemelhem a fala humana, ou seja, a geração pode realizar total ou parcialmente o papel inverso do reconhecimento de fala.

Tanto o reconhecimento quanto a geração de fala encontram problemas práticos para utilização em computadores. As teorias atuais ainda não conseguiram propor um modelo que se ajuste a qualquer tipo de fala e língua. Pois, cada pessoa possui uma forma característica de executar a linguagem. Ainda, os idiomas possuem características peculiares e algumas semelhanças entre si. Estas são decorrentes da formação histórica de uma língua.

O reconhecimento de fala ocorre por meio da captação de sinais sonoros e sua posterior codificação baseada em teorias voltadas para o estudo dos sons (acústica). Esses sinais são então convertidos para forma de fonemas.

Com relação à geração de fala, os modelos utilizados atualmente levam em consideração fatores como: entonação, sotaque e pronúncia [11], os quais são de difícil reprodução e utilização por computadores.

2.2.2 Análise Léxico-morfológica

A segunda etapa no processamento é a análise dos dados disponibilizados pelo reconhecimento de fala. É necessária uma tradução dos fonemas obtidos anteriormente em letras e palavras. Dessa forma, pode-se começar a utilizar as regras que definem as combinações válidas dos elementos componentes das palavras.

Esses elementos são os morfemas que representam combinações de fonemas [8]. Essas combinações permitem definir unidades significativas mínimas, e que são objeto de estudo da morfologia [7]. Na morfologia são estudados os processos combinatórios dos morfemas, ou seja, o estudo das manifestações fonéticas dos morfemas [8].

A morfologia pode ser dividida em morfologia lexical e morfologia flexional. Na pri-

meira estudam-se as derivações que contribuem para a formação de novos morfemas. Na segunda estudam-se os processos de flexão que indicam as categorias gramaticais em que os morfemas serão agrupados, no formato de palavras.

Um dos objetivos do estudo morfológico é a criação de dicionários ou léxicos. Eles são estruturas que contêm palavras (itens lexicais) e suas informações, tais como: categoria gramatical, suas possíveis descrições semânticas e variáveis morfossintáticas como gênero, número, grau, pessoa, modo, etc. [11]. As palavras podem ser armazenadas conjuntamente com suas informações gramaticais ou por meio de morfemas básicos que podem depois gerar palavras completas.

A recuperação das informações em um dicionário pode ser então feita de acordo com a maneira com que as palavras foram armazenadas. Quando as palavras são armazenadas diretamente no dicionário a recuperação é simples e direta. Entretanto, quando são utilizados morfemas são necessárias técnicas complexas para a formação das palavras a partir dos morfemas.

Uma vez que se têm informações sobre as palavras armazenadas em dicionário, parte-se para uma análise das palavras captadas na etapa anterior. A abordagem mais comum para a análise morfológica é considerar a interpretação de uma sentença (frase) de cada vez, no caso de existirem mais de uma sentença.

O processo de análise morfológica consiste da separação de sentenças em palavras, de palavras em morfemas, e da associação das palavras, por meio dos morfemas, às suas informações gramaticais presentes no dicionário [11]. As palavras possuem definições diferentes que dependem de um contexto em particular para uma interpretação de seu significado. Por isso, dependendo da sentença as suas palavras podem ser associadas a mais de uma categoria gramatical.

Como o contexto não pode ser determinado apenas a partir de uma sentença, a ambigüidade das categorias gramaticais atribuídas às palavras tem de ser retirada por

técnicas específicas para tal fim. Além da ambigüidade outros problemas estão associados à análise morfológica. Alguns deles são: a possibilidade de atribuição errada de morfemas durante o processo de construção de palavras e variações nas regras de escrita das palavras (ortografia) e falta de clareza na decomposição das palavras [11, 10].

Isso acontece, pois as partes componentes das palavras (afixos, radicais, etc.) podem não se combinar de forma bem definida por regras morfológicas ou ortográficas. Grande parte desses problemas ocorre por existirem muitas exceções a essas regras. Por exemplo, existe uma regra que diz: “utiliza-se a letra z no caso de se ter o sufixo *-izar* formando um verbo”, assim a palavra ‘hospital’ na sua forma de verbo é escrita como: ‘hospitalizar’. Porém, no caso da palavra ‘análise’ o verbo resultante é ‘analisar’ e não ‘anализar’ como a regra prevê.

2.2.3 Análise Sintática

Após a aquisição de informações léxico-morfológicas parte-se para a análise da organização das palavras em frases. Um dos objetivos dessa análise é determinar se um conjunto ou arranjo seqüenciado de palavras em uma frase é uma construção válida ou não em uma dada linguagem [11]. Considerando um outro ponto de vista, o objetivo da análise sintática é fazer com que certas representações semânticas sejam transmitidas através de seqüências de palavras [7].

Esses objetivos mostram que as palavras e sua organização possuem informações relevantes para a interpretação do significado geral de um conjunto de frases. Isto não quer dizer que de posse de informações sintáticas possa-se inferir o significado total de um texto, são necessárias reflexões semânticas e pragmáticas para tal fim.

Para a análise sintática são necessárias regras que definam restrições às construções ou combinações de seqüências de palavras em uma frase de uma determinada língua. Ainda, essas regras têm como função mapear informações semânticas contidas nas palavras de um vocabulário de linguagem para formas derivadas de palavras definidas

pelas regras sintáticas [8].

As regras sintáticas são agrupadas em uma gramática específica para uma linguagem, e o número de frases que podem ser definidas por essas regras é finito. Contudo o conjunto de frases de uma linguagem é infinito, dessa forma uma gramática não poderia conter todas as regras necessárias para a definição de todas as frases da uma linguagem [13]. Então seguindo esse paradigma, um conjunto limitado de regras que permitam a definição do maior número de frases será suficiente para a análise sintática.

Utiliza-se uma padronização formal para a codificação das regras sintáticas contidas em uma gramática. Essa padronização permite a reutilização de regras entre padrões formais, adaptação de novas formas de codificação de regras entre outras vantagens, segundo Zaenen, [10].

Esses padrões formais utilizam uma descrição complexa das unidades gramaticais (morfemas, palavras, frases) na forma de pares do tipo atributo-valor denominados termos característicos. A representação desses termos varia em cada padrão, podendo ser na codificação das regras de estrutura de frase ou ainda na incorporação de informações das estruturas de frases nos próprios termos [10].

Alguns desses padrões são: gramáticas de constituintes imediatos (*Phrase Structure Grammar - PSG*), gramáticas de estrutura de frases generalizadas (*Generalized Phrase Structure Grammar - GPSG*), gramática de estrutura de frase direcionada (*Head-driven Phrase Structure Grammar - HPSG*), gramática de cláusulas definidas (*Definite Clause Grammar - DCG*), gramáticas de unificação funcional, gramática léxico funcional (*Lexical Functional Grammar - LFG*), gramática semântica, gramáticas de caso, redes de transição, gramáticas de árvores adjuntas, e PATR-II, entre outros [10, 11].

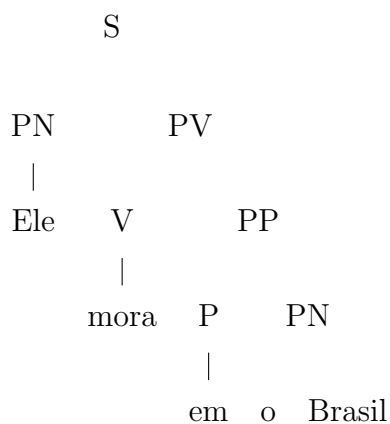
Apesar da padronização da análise sintática, uma frase pode apresentar mais de uma representação sintática, isto por causa da ambigüidade dos significados das as-

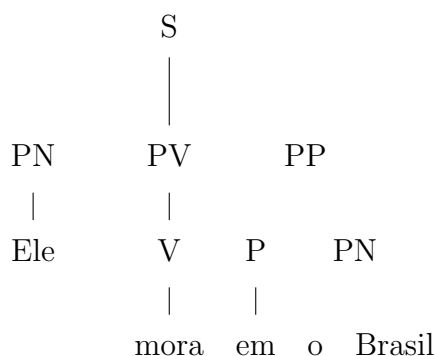
sociações que as palavras podem ter. Mais uma vez, lembrando que na análise morfológica foi necessário o uso de técnicas de retirada de ambigüidade, será necessário o uso de alguma técnica de retirada de ambigüidade, na análise sintática.

2.2.4 Análise Semântica

Nesta etapa têm-se informações sobre a estrutura de uma frase, quais são seus elementos constituintes, a ordem deles e quais relações eles mantêm entre si, ou seja, as palavras de uma frase estão associadas as suas respectivas categorias gramaticais e os conjuntos de palavras estão definidos sintaticamente.

Contudo, devido à natureza dinâmica da linguagem apenas a análise sintática não é suficiente para se ter um entendimento completo de uma frase. Isto porque no processo de análise sintática, certas frases podem gerar construções sintáticas diferentes (árvores sintáticas), e não existem regras sintáticas que possam definir apenas uma construção sintática daquela frase. Isto caracteriza uma parte ambígua da natureza da linguagem. Um exemplo, a analisar sintática da frase “Ele mora no Brasil.”pode gerar as seguintes árvores sintáticas:





Para a resolução deste tipo de ambigüidade utilizam-se técnicas de retirada de ambigüidade. Uma delas poderia ser a verificação de qual das árvores, geradas na análise sintática, é mais usual na linguagem corrente, isto é, qual delas é mais utilizada. Existem, ainda, outras formas, além da mencionada anteriormente, não se esgotando, assim, as opções e maneiras de se retirar a ambigüidade utilizando conhecimentos diversos, tais como redes neurais, abordagens probabilísticas, entre outras.

Desse modo, apesar das contribuições de Chomsky para o estudo sintático, este trabalho considerará que a análise semântica é necessária e complementar as análises anteriores. Assim, por mais independente que uma regra sintática seja de seu conteúdo significativo, ela sozinha não poderá fornecer meios de interpretação significativa, mesmo para a retirada de ambigüidade.

Porém, as construções sintáticas são possuidoras de pistas semânticas que podem ser utilizadas para o entendimento do contexto em que elas podem vir a serem interpretadas. Dessa forma as estruturas sintáticas (árvores sintáticas) podem ser complementadas por outros tipos de regras que levem em consideração o significado não apenas sintático dos conceitos apresentados na frase, mais também sua extensão semântica.

Por exemplo, considere uma regra sintática com a seguinte definição: uma frase é formada por um nome ou substantivo em segunda pessoa, de gênero masculino e número plural. Por um verbo conjugado para acordar com a segunda pessoa de número plural, e um complemento verbal concordando em gênero e número. Nesta regra carac-

terísticas tais como: gênero, número e pessoa, são destacadas como parte integrante da regra, isto é, a regra define explicitamente qual a possível configuração de nome, verbo e complemento a que ela faz distinção, isto difere da definição simplificada de regra sintática, onde apenas nome, verbo e complemento são necessários para a sua definição.

A análise semântica considera a estrutura fornecida pela análise sintática como base para uma representação e interpretação do significado de uma frase [11]. De acordo com Borba, [7] a análise semântica pode ser dividida em: análise semântica lexical, que estuda a organização do léxico (dicionário) e tem como premissa o princípio de que este não é apenas uma lista de palavras, mas sim um depósito de conhecimentos pontuais representados por palavras; e análise semântica frasal, que estuda a organização da significação de uma sentença.

Essa divisão permite uma divisão da análise semântica em duas partes, na primeira os aspectos lexicais podem ajudar na identificação de significados pontuais. Na segunda os significados pontuais, expressos pelas palavras, aplicados à organização sintática de uma frase podem determinar o significado da frase. Isto também pode auxiliar no tratamento das ambigüidades encontradas tanto na análise sintática quanto nas demais análises que fazem parte do processamento de linguagem natural.

Ao final da análise semântica pode-se considerar que a linguagem natural foi traduzida para uma linguagem formal mais adequada ao tratamento computacional. Esse tratamento pode ser feito por teorias com base lógica, por exemplo, a lógica proposicional. Com a representação formal da linguagem é possível extrair informações sobre o significado de uma sentença.

Em adição, outras teorias e abordagens podem e são empregadas para o tratamento computacional mencionado até então neste capítulo. Além da base lógica utilizam-se abordagens baseadas em regras ou casos, onde o conhecimento necessário para as análises lingüísticas é obtido por casos encontrados no uso das línguas e codificados em formato de regras genéricas, as quais então podem ser adotadas por sistemas compu-

tacionais.

Outra linha de abordagens são aquelas baseadas em corpus, nelas informações lingüísticas são obtidas por meio de técnicas de probabilísticas, porém não se limitando a elas, ou seja, outras técnicas podem ser empregadas para a extração de informações relevantes para o estudo de uma língua e no caso deste trabalho, para a extração de regras que possam ser utilizadas para direcionar as diversas análises lingüísticas.

Um outro meio de se abordar o estudo lingüístico é por meio do uso de técnicas conexionistas. Tais técnicas surgiram dos estudos da área de inteligência computacional e neurobiologia, tendo como objetivo simular o funcionamento das redes neurais humanas em computadores, as chamadas redes neuronais artificiais. Esta abordagem difere das demais por não possuir um conjunto de regras definido, ao invés, como em humanos elas tentam aprender através de exemplos as regras necessárias para a realização das diversas análises lingüísticas.

2.2.5 Análise Pragmática

Como visto anteriormente a semântica e sua análise tratam do estudo ou representação do significado contido em uma sentença. Porém existem casos onde o significado não pode ser determinado sem se levar em consideração o contexto em que eles acontecem. Para casos como esses se necessita de outro mecanismo ou conjunto de teorias que tentem prover as informações que possam explicar o sentido de uma sentença considerando o contexto em que ela se encontra ou faz parte, contexto esse determinado por outras sentenças e conhecimento extralingüísticos (culturais e situacionais).

A pragmática se ocupa do estudo do significado de sentenças levando em consideração o contexto de seu uso. Para tal, ela busca meios de interpretar o contexto não apenas com informações de uma sentença ou da anterior, ela então leva em consideração o discurso (conjunto de sentenças) como um todo, para descrever um contexto geral e por fim interpretar o sentido expresso por uma sentença componente desse

discurso.

No estudo pragmático é necessária uma maneira de representação formal das relações de antecedência que ocorrem entre os elementos de uma sentença e dentro de um discurso. Os fenômenos lingüísticos (as relações anafóricas, a co-referência, a determinação de antecedentes, o foco ou tema do discurso, as elipses, entre outros [11]) ajudam a compor os elementos utilizados no estudo pragmático.

2.3 Aplicações do processamento de linguagem natural

Adotando a posição de Allen, [14], a pesquisa em processamento de linguagem natural pode ser dividida em duas classes: aplicações baseadas em texto e aplicações baseadas em diálogos.

Na primeira classe estão as aplicações que lidam com tarefas relacionadas a textos (livros, manuais, relatórios, etc.). Algumas das tarefas podem ser:

- Procura de informações relevantes em um banco de dados de documentos (uma biblioteca);
- Extração de informações a partir de documentos ou mensagens;
- Tradução de documentos;
- Sumarização de textos.

A segunda classe de aplicações envolve a comunicação homem-computador (escrita ou falada). Algumas de suas aplicações são:

- Sistema de perguntas e respostas;

- Sistemas de resolução cooperativa de problemas;
- Sistemas tutores de ensino;
- Controle de fala sobre equipamentos eletrônicos;
- Serviços ao consumidor disponibilizados por telefone (informações bancárias);
- Recuperação de informações em banco de dados.

Existem ainda outras aplicações para o processamento de linguagem natural, muitos deles envolvendo aplicações na área de inteligência artificial.

As aplicações da segunda classe serão destacadas neste trabalho, principalmente os sistemas que permitam o diálogo entre homens e computadores.

Capítulo 3

Sistemas de diálogo

3.1 Definição

São sistemas computacionais que possibilitam comunicação entre homem e computador por meio de diálogo falado ou escrito. Esses sistemas têm como funções: servir de interface entre aplicações (banco de dados, sistemas especialistas) ou serviços (consultas através de telefone ou Internet) e o homem, e, assistência na resolução e planejamento colaborativos de tarefas (itinerário de trens, projeto de circuitos eletrônicos). Alguns exemplos de tais sistemas são apresentados no capítulo 4.

Os pontos fundamentais de um sistema de diálogo, que serão explanados neste capítulo, são: abordagens de codificação de diálogo, fontes de conhecimento, componentes do sistema e a forma de controle de diálogo.

3.2 Abordagens para a codificação de diálogo

Os sistemas de diálogo necessitam formalizar sua comunicação, no caso o diálogo. Para este fim são usados modelos computacionais que tentam descrever as características gerais de um diálogo, as quais são baseadas em estudos empíricos, cognitivos, psicológicos e lingüísticos de diálogos humanos.

A seguir serão apresentados os tipos de abordagens, os quais podem ser considerados como uma forma de classificação de sistemas de diálogo. Outras classificações são fornecidas por Flychteriksson em [15], relacionada a aplicação do sistema, e por McTear em [16], relacionado ao método de controle de diálogo utilizado no sistema.

3.2.1 Abordagens Estruturadas

Também conhecidas como abordagens baseadas em gramáticas de diálogo ou em diagramas de estados finitos. Elas consideram que existe uma estrutura regular de diálogo que pode ser representada por uma gramática [17], ou seja, o diálogo é uma seqüência de estados pré-determinados que obedecem a uma gramática de estados finitos [16].

Assim, a estrutura de um diálogo pode ser descrita a partir da observação de seqüências de sentenças, por meio de um conjunto de regras finitas (regras de diálogo), as quais compõem uma gramática. Isso permite definir que a coerência de um diálogo está na observação e identificação das seqüências subseqüentes do diálogo e não na explicação do que está sendo observado [17]. Este ponto final é o caracterizador dos sistemas que adotam esta classe de abordagens [18].

Uma forma de se utilizar essas abordagens é relacionar a estrutura do diálogo com a estrutura do contexto, o que é apontado como uma desvantagem por Sadek [17], pois, fixando-se em um contexto específico perde-se a naturalidade contextual da comunicação humana. Esse relacionamento ocorre pela definição de conjuntos de regras de diálogo capazes de determinar rigidamente todas as possibilidades de sentenças, isto é, todas as sentenças que podem ser utilizadas no diálogo são definidas previamente por esse conjunto de regras.

Essa rigidez no diálogo permite ao sistema conhecer ou prever as seqüências de sentenças que o usuário poderá utilizar, dessa forma as informações requeridas pelo usuário serão fornecidas sem processamentos complexos. A vantagem de limitar a en-

trada do usuário é poder manter um vocabulário e uma gramática restritos, além de diminuir a complexidade do processamento de reconhecimento da entrada lingüística e de dispensar algumas das fontes de conhecimento que um sistema de diálogo pode necessitar para manter um diálogo coerente com o usuário. Tais fontes serão discutidas adiante na seção 3.3.

Uma desvantagem está na própria restrição da entrada, pois, o usuário fica impossibilitado de se expressar de forma natural, tendo que seguir as opções apresentadas pelo sistema. Isto acarreta em baixo grau de satisfação por parte do usuário [17]. Também impossibilita ao usuário tomar o controle do diálogo, pois, na comunicação humana a tomada de controle possibilita entre outras funções a correção de sentenças mal expressas e o esclarecimento de sentenças mal entendidas.

Outras desvantagens são feitas por Churcher e Uszkoreit [19, 18]: a falta de flexibilidade da estrutura de diálogo, limitação de conhecimentos, específico a um domínio particular, impossibilidade de tratamento de sentenças com mais de uma função comunicativa simultânea, impraticabilidade de adoção de referências diferentes ao contexto corrente do diálogo, não consideração de sentenças com sentidos indiretos (aquelas onde a interpretação não é literal).

Apesar das desvantagens apontadas, em [20] Dahalback apresenta um estudo que alega que para casos particulares, interfaces em linguagem natural para sistemas computacionais, é possível utilizar abordagens estruturadas. Estabelecido ainda que essa abordagem não segue modelos cognitivos da linguagem humana, mas que simula características pontuais da mesma.

3.2.2 Abordagens baseadas em planos

Estas abordagens consideram que a fala humana não é meramente uma troca de palavras. A fala é constituída de ações comunicativas, representadas por atos de fala [2], que são executadas com a finalidade de alcançarem-se objetivos pré-determinados

[21, 17, 14, 22].

Assim, por meio da modelagem das ações comunicativas, no formato de plano (conjunto de ações que levam a um ou mais objetivos), pode-se tentar reproduzir o objetivo da comunicação humana, que é tentar modificar o estado mental do interlocutor por meio da mudança de seu estado mental (crenças, intenções, etc.) [17].

Então quando uma pessoa fala, ela executa ações que fazem parte de um plano cuja referência está no objetivo dessa pessoa querer comunicar algum tipo de informação a outra pessoa. A pessoa que percebe essa fala considera que a fala possui objetivos e planos que expressam o estado mental de quem fala, e por meio da interpretação dos efeitos causados pelas ações da fala mudará seu próprio estado mental para reconstruir o plano de quem fala e poder tomar as ações que lhe forem convenientes.

Os sistemas que adotam a abordagem de planos necessitam de mecanismos para o reconhecimento e o planejamento de objetivos. Para isso, eles utilizam regras de inferência, definição de ações, modelos de estados mentais do usuário e um conjunto de expectativas apropriadas aos objetivos e ações em determinado contexto [18], cabe ressaltar que esta subseção não tem a intenção de prover todo o conhecimento referente à pesquisa de reconhecimento de plano.

Um exemplo de sistema de diálogo que utiliza planejamento como forma de controle de diálogo são sistemas conhecidos como sistemas de perguntas e respostas, que são sistemas compostos de vários quadros¹, onde as perguntas do sistema e respostas do usuário serão usadas para cumprir um plano do usuário.

A estrutura desses sistemas é uma rede ou árvore hierárquica de quadros que re-

¹Quadros são estruturas de dados que representam uma situação estereotipada. Essa estrutura pode ser entendida como uma rede composta de nós que representam os elementos da situação e de relações que representam as relações desses elementos dentro da situação representada, segundo Minsky em [23].

presenta a estrutura do contexto. Assim, o fluxo do diálogo entre sistema e usuário se dá por meio de perguntas e respostas, cujo objetivo é reconhecer o plano do usuário por meio de suas respostas. Isso acontece pelo preenchimento sucessivo de quadros inter-relacionados, de acordo com as informações extraídas das respostas do usuário.

O tratamento lingüístico das respostas do usuário tem sua complexidade diminuída pela utilização dos quadros, pois, o sistema pode condicionar seu processamento a uma busca ou casamento de conceitos extraídos das respostas do usuário, o que permite um tratamento mais adequado a erros na entrada do usuário [16].

As perguntas do sistema têm por objetivo guiar o preenchimento dos quadros, o que fornecerá a resolução do plano que executará o objetivo do usuário. Entretanto, o fluxo do diálogo não é pré-definido, ou seja, as respostas do usuário não pertencem a um conjunto finito de sentenças, como nos sistemas que utilizam abordagem estruturada. O sistema tem de levar em consideração o contexto atual da interação para gerar perguntas que mantenham o rumo de execução do plano do usuário.

Para facilitar a execução dos objetivos do plano do usuário, as perguntas que o sistema fará podem ser acrescidas de condições, as quais proverão um grau de relevância a cada pergunta. Isso possibilita a escolha da pergunta mais adequada a determinado ponto do diálogo (planejamento de plano do sistema) [16].

O controle das etapas do planejamento atribui um grau maior de complexidade a esses sistemas. Pois, no planejamento é considerado que um plano tem de alcançar um objetivo final, contudo, para diminuir a complexidade dessa tarefa considera-se divisão do objetivo final em sub-objetivos. Tais objetivos podem ser controlados de forma isolada, o que permite o processo de reconhecimento do plano geral mais simples. Contudo determinar quais os objetivos ainda estão pendentes é um problema de manutenção de plano. Para resolver este problema é necessária a adoção de uma fonte de conhecimento adicional ao contexto, o modelo de tarefa (a ser apresentado em 3.3.3).

A vantagem destas abordagens é permitir a resolução de algumas das desvantagens das abordagens anteriores, tais como o problema de sentenças com mais de uma função comunicativa simultânea, flexibilidade da estrutura de diálogo e o problema das sentenças com sentidos indiretos. Ainda, a iniciativa do diálogo pode ser atribuída ao usuário.

A desvantagem mais aparente é a dificuldade de precisar-se qual o plano atual do usuário, isto ocorre devido à: ambigüidade inerente ao processo de mapeamento da entrada lingüística e a sua função comunicativa e a formalização teórica das ações comunicativas.

Para Sadek [17] os motivos dessa dificuldade estão: na fraqueza da formulação lógica usada para representar as atitudes mentais, cuja interpretação ainda é intuitiva e não formal, a dificuldade encontrada por técnicas de inteligência artificial em caracterizar o efeito de uma ação comunicativa como uma função do contexto mental onde ela está sendo executada e a sub-estimação do problema de estabelecer os tipos de ações e suas especificações segundo as atitudes mentais. Outras desvantagens são comentadas por Churcher e Uszkoreit [19, 18].

3.2.3 Abordagens baseadas em colaboração

Tais abordagens partem do princípio de que o processo comunicativo e interativo entre os participantes de um diálogo é um processo colaborativo no qual os participantes trabalham juntos para alcançar um entendimento mútuo sobre este diálogo. Dessa forma, o diálogo é considerado uma atividade mútua onde os participantes possuem certos compromissos, tais como confirmação e verificação, entre outros tipos, necessários a sua compreensão [19].

Diferente das abordagens apresentadas anteriormente, aqui a estrutura do diálogo é deixada de lado e a forma de modelagem do diálogo recai sobre a descoberta das motivações dos participantes e também dos mecanismos do próprio diálogo, os quais servirão de base para a colaboração mútua entre os participantes.

Um sistema de diálogo colaborativo usa um modelo que descreve as crenças sobre o diálogo de ambos participantes, onde um deles faz uma proposta que pode ser aceita ou rejeitada pelo outro participante. Quando uma proposta é aceita ela é compartilhada como uma crença mútua entre ambos os participantes. Os objetivos do usuário são normalmente adotados pelo sistema e propostas feitas com relação a estes objetivos são trabalhadas cooperativamente para os cumprir [19].

Os participantes do diálogo, usuário e sistema, são considerados como agentes inteligentes, termo provido pela área de inteligência artificial para a modelagem de uma entidade autônoma em relação a um ambiente qualquer. Assim, cada agente possuirá um conjunto de crenças que ao serem confrontadas com as do outro agente fornecerão o entendimento mútuo de ambos sobre o diálogo. Churcher em [19] apresenta alguns modelos de sistemas colaborativos baseados em agentes.

Sistemas colaborativos permitem um fluxo de diálogo mais natural que o permitido pelas outras abordagens, pois, não limita o usuário quanto à forma de se expressar. Isso ocorre nas abordagens estruturadas e baseadas em plano por existir a necessidade do conhecimento prévio do contexto do diálogo. Contudo, em abordagens colaborativas o contexto não é mais tão importante para o progresso do diálogo, uma vez que nessas abordagens utiliza-se a noção de elementos comuns de um diálogo e não conhecimentos específicos sobre a estrutura do diálogo, os quais necessitam de conhecimento específico de contexto.

Essa vantagem permite ainda estabelecer um diálogo mais próximo do natural, pois, considerando a adição de alguns poucos elementos dependentes de contexto, crenças e intenções, ao modelo genérico adotado pelo sistema colaborativo, este pode satisfazer melhor ao usuário. Outra vantagem é que algumas de suas abordagens podem simular certas características da comunicação humana que não podem ser representadas pelas abordagens anteriores [19].

As desvantagens são: a alta complexidade das metodologias utilizadas na construção de sistemas colaborativos e a necessidade de modelar todas as crenças, em um contexto específico, dos participantes do diálogo. Esta última colocação se torna uma desvantagem quando a inferência dessas crenças ocasionalmente sai do escopo do contexto estabelecido para o diálogo, o que provoca a determinação incompleta e inconsistente do contexto [19].

3.2.4 Abordagem baseada em interação racional

Esta abordagem considera que a comunicação é um caso especial de comportamento inteligente, e segundo Sadek [17] ela faz uso de elementos das abordagens baseadas em planos e das baseadas em cooperação. Tais como a noção de intenção na comunicação fornecida por Cohen em [24], e de alguns dos princípios de interação cooperativa.

A noção básica considera que os participantes de um diálogo tendem a possuir um comportamento cooperativo governado por princípios racionais e cooperativos. Com esses princípios podem-se descrever as ações que revelam as intenções formadoras de um plano, no qual consideram-se ainda as crenças e objetivos referentes aos conhecimentos específicos a um domínio.

Assim, em um diálogo podem-se extrair alguns aspectos que caracterizam um comportamento racional, sendo o principal deles a cooperação. Esses aspectos são apresentados como princípios primitivos que descrevem o comportamento racional de um participante de um diálogo. Por meio desses princípios a estrutura do diálogo pode ser inferida, o que dispensa a manutenção de um modelo de diálogo explícito.

Os princípios racionais, que descrevem o comportamento racional, refletem as crenças e planos decorrentes sobre um dado domínio, além das ações que expressam as intenções dos participantes racionais. As intenções dos participantes são definidas pelas suas crenças sobre um domínio, objetivos e comprometimentos para o entendimento mútuo dos participantes da situação do diálogo [16].

Esses princípios são formalizados logicamente por axiomas, os quais possuem como elementos principais às atitudes mentais (são as relações mantidas entre participantes de um diálogo, objetos ou situação atual). Para Sadek [17] elas são: crença que é uma situação sobre um modelo de domínio em que um participante acredita ser verdadeira, incerteza que é uma forma de se representar uma crença não verdadeira sobre uma situação do domínio e intenção que está relacionada a ações futuras motivadas pelos princípios racionais e cooperativos.

Os princípios cooperativos expressam as motivações dos participantes de agirem cooperativamente, por meio de respostas cooperativas. Essas respostas manifestam então um comportamento cooperativo que pode ou não ser seguido por um participante do diálogo. E sua produção depende diretamente do contexto ou situação em que o diálogo está [17].

Para a produção de comportamento cooperativo é necessária a análise da comunicação e do domínio. Isso ocorre pela identificação das intenções presentes na comunicação (diálogo). Essa identificação pode ser feita por técnicas de reconhecimento de plano sobre um dado domínio, o qual requer um modelo de domínio que especifique as ações possíveis dentro do domínio [17].

Porém, apesar do uso de princípios cooperativos ainda existe a necessidade de um esquema de planejamento dependente do domínio. Uma solução apresentada por Sadek em [17] seria a de adicionar funções que expressem informações referentes ao domínio ao conjunto de axiomas lógicos que representam os princípios racionais e cooperativos. Com essa medida, mantém-se a base de princípios e acrescentam-se informações necessárias para o planejamento adequado de ações comportamentais racionais e cooperativas em um domínio específico.

Vantagens dessas abordagens são a explicação do porque das informações extras em respostas cooperativas, tratamento de contextos de interação diferentes e possibilidade

de explicação de fenômenos complexos da comunicação, como a negociação [16].

3.3 Fontes de conhecimento

As fontes de conhecimento utilizadas pelos sistemas de diálogo para manter um diálogo coerente, realizar a tarefa do sistema, recuperar-se de erros, entre outras utilidades são, segundo Flycht-Eriksson [15]: diálogo, discurso, modelo de usuário, tarefa e domínio.

Essas fontes são representadas por modelos que descrevem suas informações características em um formato que o sistema de diálogo possa trabalhar. Logo cada fonte possuirá um modelo correspondente.

Uma análise da funcionalidade do sistema de diálogo e de seus possíveis comportamentos em diálogo determinarão quais das fontes de conhecimento necessárias serão necessárias para a construção do sistema [15].

3.3.1 Modelo de diálogo

As informações contidas em um modelo de diálogo estão relacionadas com a representação das relações entre os constituintes de um diálogo [15]. Cohen em [18] considera a existência de duas linhas de pesquisa sobre a modelagem de diálogo. Na primeira busca-se uma formulação que explique teoricamente um diálogo. Na outra o objetivo é desenvolver algoritmos que permitam a um computador participar de um diálogo cooperativo.

Ambas tem de descrever como um diálogo pode ser construído e a representação de um diálogo resultante da aplicação de uma dessas teorias. Assim, as informações sobre a construção de um diálogo consideram as relações dos seus constituintes [15].

Apesar de não existir ainda um consenso sobre qual dessas teorias é a mais adequada para modelar um diálogo, duas categorias de abordagens são apontadas por Cohen em [18]: gramática de diálogo e modelos de diálogo baseados em plano. Elas são semelhantes a algumas das abordagens apresentadas em 3.2.

A primeira baseia-se na observação da existência de regularidades seqüenciais em diálogos, denominadas pares adjacentes (perguntas são seguidas de respostas, por exemplo). E por essa observação foram estabelecidas regras que representassem essa natureza seqüencial, algo semelhante às regras gramaticais utilizadas para descrever sentenças. Assim como essas regras são usadas para analisar a estrutura de uma sentença, aquelas podem ser utilizadas para analisar a estrutura de um diálogo. As gramáticas de diálogo como teoria não são práticas, pois, não explicam o fenômeno que elas tentam descrever, segundo Cohen em [18].

Em modelos de diálogo baseados em plano, como apresentado anteriormente, considera-se que as sentenças representam ações comunicativas que são descritas por planos que representam meios de se alcançar certos objetivos. Essas ações são creditadas como parte integrante dos planos de um dos participantes do diálogo, sendo assim, a função do outro participante é identificar, a partir das ações, o plano do outro participante, isto segundo Cohen em [18].

Um outro componente para o modelo de diálogo é o histórico de diálogo. Nele ficam armazenadas as informações sobre o estado corrente do diálogo, além de armazenar o que foi dito e do que se está falando até o momento atual [15]. Essas informações podem ser usadas para o controle de diálogo e definição de referências pronominais.

3.3.2 Modelo de discurso

Discurso é um conjunto de elementos que possuem algum tipo de relacionamento. Esses elementos podem ser falas, o que formará um diálogo, ou um conjunto de sentenças, o que forma um texto. As abordagens para a modelagem de discurso seguem

quatro teorias de discurso segundo Scott e Kamp em [18]:

- Teoria de coerência de discurso baseada em um conjunto de relações de coerência entre os elementos do discurso.
- Teoria que considera a divisão do discurso em três partes que lidam com diferentes aspectos do discurso: foco de atenção, a estrutura intencional da fala e a estrutura de seqüência de falas.
- A Teoria da Estrutura Retórica de discurso.
- Organização hierárquica do discurso por meio de esquemas fixos que garantem a coerência e permitem a geração de conteúdo.

Um formalismo para a modelagem de discurso é a Teoria de Representação de Discurso [25]. Onde se considera que as sentenças ou falas possuem referencias que mantêm a coerência do discurso, tais como anáforas, elipses e outras formas de coesão textual.

Scott e Kamp em [18] fornecem uma lista não exaustiva de teorias computacionais que podem ser utilizadas para a construção do modelo de discurso.

3.3.3 Modelo de tarefa

A tarefa do usuário pode ser representada por uma hierarquia de objetivo e sub-objetivos, cuja representação expressará o possível plano (combinação relacionada de ações para se alcançar objetivos) do usuário. As ações desse plano podem ser construídas a partir do reconhecimento e emprego de conjuntos de atos de diálogo [26]. Também, a tarefa do sistema pode ser representada por uma hierarquia, porém, uma hierarquia de sub-tarefas que capturem a estrutura geral da tarefa do usuário [15].

Para construir um modelo de tarefa do usuário, um sistema de diálogo tem de considerar o efeito que a tarefa possuirá sobre o diálogo, pois, o estabelecimento do contexto da tarefa afetará o contexto do diálogo, e como foi visto anteriormente o

conhecimento do contexto influencia nas decisões de controle de diálogo. As tarefas podem delimitar o diálogo de três formas [26]:

1. Determinam quais objetivos são relevantes. Isso permite a execução de uma sub-tarefa apenas quando os seus atributos necessários para execução se tornarem relevantes para a tarefa atual.
2. Delimita a ordem das trocas de atos de diálogo que utilizam sub-tarefas específicas.
3. Quando duas sub-tarefas ocorrem simultaneamente, no caso de sistemas multimodais². Elas serão sincronizadas para permitir a continuidade do diálogo.

Uma vez que a tarefa está definida, resta decidir como o sistema agregará a tarefa do usuário, por meio do modelo da tarefa, a sua estrutura interna. Essa definição conta com a relação entre o modelo de tarefa e o modelo de diálogo adotado pelo sistema. Flycht-Eriksson em [15] aponta três categorias de relacionamento entre esses modelos:

- Diálogos de tarefa – onde o papel do sistema é guiar o usuário na execução de sua tarefa, o que necessitará de um modelo da tarefa do usuário.
- Diálogo planejado – onde o sistema ajudará no planejamento da tarefa do usuário, e como no caso anterior necessitará de um modelo da tarefa do usuário.
- Diálogos parametrizados – o sistema fornecerá informações ao usuário sem possuir conhecimento de qual é a sua tarefa, ou seja, o sistema não possui um modelo da tarefa do usuário.

No primeiro caso o sistema de diálogo tentará modelar um diálogo com o usuário de forma que a tarefa seja realizada com sucesso. O modelo de tarefa é conhecido por ambos os participantes do diálogo. Isso permite o ajuste do modelo de diálogo ao contexto específico da tarefa do usuário.

²Sistema de diálogo que trabalham com mais de um modo de interação: fala, escrita, imagens, etc.

Na segunda categoria o modelo de tarefa não necessita estar ligado explicitamente ao modelo de diálogo, pois, o modelo de diálogo obedece a uma estrutura de planejamento que não necessita fixar seu funcionamento a um contexto em particular, logo em princípio o sistema poderia trocar de tarefa se o modelo de uma nova tarefa for compatível com um tipo de modelo de tarefa que o modelo de diálogo planejado possa entender.

No último caso o sistema possui um modelo de diálogo baseado em princípios específicos sobre a forma de diálogo, o que permite a troca do modelo de tarefa, sem a necessidade de ajustes no modelo de diálogo, devido a troca de contexto.

Portanto, o modelo de tarefa separado do modelo de diálogo permitiria um diálogo mais natural com o usuário e tornaria o sistema de diálogo mais eficiente, pois possibilitará a adição de novas tarefas e modificação nas existentes, de forma simples [15].

3.3.4 Modelo de domínio

É a representação do conhecimento de mundo que é referente à uma aplicação ou área que o sistema de diálogo serve como interface entre o usuário e a aplicação ou área. Este conhecimento é usado para auxiliar no mapeamento das informações lingüísticas provenientes do usuário aos conceitos de um dado domínio. As informações provenientes do modelo de domínio, logo, auxiliam na interpretação semântica e pragmática das entradas do usuário e na manutenção de um diálogo mais natural e intuitivo.

Essas informações podem ser divididas em três modelos que fazem parte do modelo de domínio: modelo de informação ou conceitual, modelo de interação e modelo de tarefa [26].

O modelo de informação ou conceitual não é o modelo de domínio, é apenas uma parte dele. Pois, o modelo de domínio representa uma formalização estrutural do mundo, composto por sub-conjuntos de conhecimentos gerais do mundo. Já no modelo

de informação estão representados apenas os relacionamentos conceituais estabelecidos entre os objetos de um domínio [15].

Esses relacionamentos são agrupados em estruturas denominadas ontologias, as quais armazenam as relações entre os conceitos de objetos em um domínio em particular. Para Hulstijn [26], os conceitos representam possíveis tópicos de assunto em um diálogo, onde cada tópico apresentará objetos que contêm as informações referentes ao que se está em foco no diálogo corrente.

Assim, a estrutura de um diálogo poderia ser representada por uma ontologia, que pode ser entendida como um espaço hierárquico de tópicos conectados por meio de atributos e conceitos relacionados (sub-tópicos). As conexões representam o espaço hierárquico e o encadeamento dos tópicos indica a coerência de um diálogo.

Ao se utilizar a estrutura construída pelo modelo de informação e o conhecimento do contexto específico do diálogo, delimitado pelo modelo de tarefa, é possível descrever o diálogo com a utilização de atos de diálogo, cuja função é determinar segmentos de diálogo, os quais podem ser agrupados em tópicos.

Essa descrição permite descrever um diálogo por meio de uma árvore hierárquica, que conterá as informações de cada interação entre usuário e sistema. Cada interação terá de ser agrupada em um ramo que represente adequadamente a relação de coerência estabelecida pelo ramo como um todo. As relações de coerência são determinadas por conjuntos de atos de diálogo e por meio delas pode-se estabelecer o modelo de interação para o sistema de diálogo [26].

3.3.5 Modelo de usuário

Segundo Kobsa,[27], modelo de usuário é um conjunto de suposições levantadas pelo sistema sobre alguns aspectos da interação com o usuário que sejam relevantes para um domínio em particular. Tais aspectos são: objetivos do usuário, planos do

usuário para alcançar esses objetivos e conhecimento (crenças) do usuário sobre um domínio específico.

Esses aspectos são importantes para o controle do diálogo, uma vez que a partir deles o sistema identifica fatores da comunicação que o ajudem a dialogar de forma mais flexível ou natural e também cooperativa, fornecendo meios para o sistema levar em consideração: quais informações são necessárias para os objetivos e planos do usuário, o que o usuário sabe ou não sobre determinado ponto do diálogo, o que evitará respostas sem sentido ou redundantes e as crenças do usuário sobre do domínio [27].

A questão de flexibilidade do diálogo pode ser entendida quando se considerar o tratamento de mal entendidos e erros de reconhecimento de emissões do usuário. Muitos desses erros e mal entendidos são causados pela falta de conhecimento do usuário por parte do sistema. Se o sistema não sabe quais as características do usuário que interage com ele, então técnicas de recuperação de erros terão de ser empregadas para corrigir ações futuras do sistema. Entretanto, se o sistema possuir algum tipo de conhecimento sobre o usuário então ele poderá fazer suposições sobre como interagir em um diálogo. Essas suposições pode ser utilizadas para evitar mal entendidos [28].

As suposições podem ser identificadas segundo Kobsa [27] como: padrões a serem obtidos pelas contribuições do usuário e do sistema. Onde o primeiro tipo agrupa informações que são extraídas de crenças e objetivos específicos a um domínio, e que podem definir características comuns a qualquer usuário desse domínio, ou seja, o sistema considera que qualquer usuário de um domínio terá inicialmente um conjunto semelhante de crenças e objetivos. Essas suposições se referem ao conhecimento geral do usuário sobre o domínio, suas crenças sobre o mesmo e seus objetivos, outro conjunto de suposições diz respeito às crenças do usuário sobre as crenças e objetivos do sistema.

As suposições obtidas das contribuições do usuário são mais confiáveis que as apresentadas anteriormente, pois são fornecidas pelos próprios usuários durante sua interação com o sistema. As formas de obtenção dessas suposições são: questionamentos

diretos sobre crenças e objetivos, utilização da forma sintática para a inferência da suposição e pistas lingüísticas fornecidas por certas palavras, também usadas para inferência. No caso das contribuições do sistema, existe a noção de que o sistema planejou contribuições futuras para coletar informações sobre o usuário e dessa forma atualizar seu modelo de usuário [27].

A forma de agrupar essas informações e suposições em um modelo de usuário pode seguir dois tipos de abordagens: abordagem estereotipada e formação de histórico de interação. Outras formas podem ser utilizadas para este fim, porém estas duas são mais simples e se ajustam ao propósito de criar um modelo de usuário para uso em sistema de diálogo.

Os estereótipos são elementos que armazenam as informações características a um grupo de usuários. Quando um sistema os utiliza, algumas de suas características marcantes são comparadas as obtidas pela primeira interação de um usuário com o sistema, isto visa a identificação ou classificação desse usuário em um estereótipo. Assim, o sistema poderá preencher um modelo de usuário individual com as informações obtidas da interação e das obtidas pela classificação [29]. Logo, o sistema pode obter algumas suposições sobre usuário sem a necessidade de uma interação completa, o que se bem feito pode melhorar a qualidade e coerência do diálogo.

Uma segunda abordagem considera a gravação de grupos de dados extraídos do comportamento do usuário na interação com o sistema, onde cada grupo é associado a um conjunto de usuário ou parte específica de um plano para então generalizar os dados obtidos e descobrir padrões de interação. Assim quando novos usuários interagirem com o sistema, padrões de interação coletados anteriormente serão usados para classificá-los e então guiar sua interação. Outra abordagem é atribuir novos usuários a grupos pré-existentes por meio de algumas de suas informações características, e desse modo generalizar seu comportamento na interação para padrões de cada grupo de usuário mantido pelo sistema [29].

A utilização do modelo de usuário por um sistema de diálogo dependerá da abordagem de controle de diálogo e do modelo de interação. Assim, para sistemas baseados em estados finitos e variações simples (sistemas do tipo pergunta-resposta) um modelo de diálogo pode não ser necessário, uma vez que a estrutura do diálogo já está definida e a natureza do domínio pode não necessitar de informações adicionais de usuários para melhorar a interação deles com o sistema.

Entretanto para sistemas de pergunta e resposta, baseados em quadros e sistemas baseados em agentes que adotem a noção de diálogo cooperativo o uso de um modelo de usuário que utilize qualquer tipo de abordagem, pode ajudar no controle de diálogo. Isso devido a um modelo de usuário possuir informações relevantes ao diálogo [15].

3.3.6 Relações entre modelos

Os modelos apresentados possuem algumas relações referentes ao tipo de informação armazenada e a forma de utilização em determinado tipo de sistema de diálogo.

Por estes modelos apresentarem semelhanças em determinado tipo de sistema de diálogo, Flycht-Eriksson em [15] discute a possibilidade de junção de tais modelos a partir do estudo das relações dos mesmos. Essa discussão será mostrada resumidamente a seguir, pois podem existir outros fatores que sejam relevantes para a interligação dos modelos apresentados.

3.3.6.1 Modelo de usuário e modelo de discurso

Uma possível junção de informações entre modelos seria considerar a metodologia cognitiva sobre memória de longa e curta duração para efetuar o controle de diálogo. Sendo assim o modelo de usuário seria o armazenador de conhecimentos de longa duração, que seriam usados para decidir sobre possíveis formatos para resposta ao usuário. E o modelo de discurso, neste caso um histórico de diálogo, representaria as últimas informações ou último diálogo com o usuário (memória de curta duração) que então

seria utilizado para a interpretação sensível ao contexto e para a geração de linguagem.

3.3.6.2 Modelo de diálogo, modelo de tarefa e modelo de domínio

Uma forma de união de modelos é a incorporação dos modelos de tarefa e domínio ao modelo de diálogo. Isso é possível pelo sistema depender de informações de um domínio para execução de uma tarefa. Pois, a interligação entre os modelos de domínio e tarefa, em sistemas orientados a execução de tarefa, ocorre devido à execução da tarefa depender do domínio. Assim os modelos não necessitarão de separação na arquitetura do sistema de diálogo. O modelo de tarefa pode ser incorporado ao de domínio e o inverso também pode ocorrer.

3.4 Componentes de um sistema de diálogo

Na figura 3.1 podem-se destacar os seguintes componentes de um sistema de diálogo: reconhecimento de fala/texto, entendimento de linguagem, gerente de diálogo, aplicação externa, geração de linguagem e saída falada/textual.

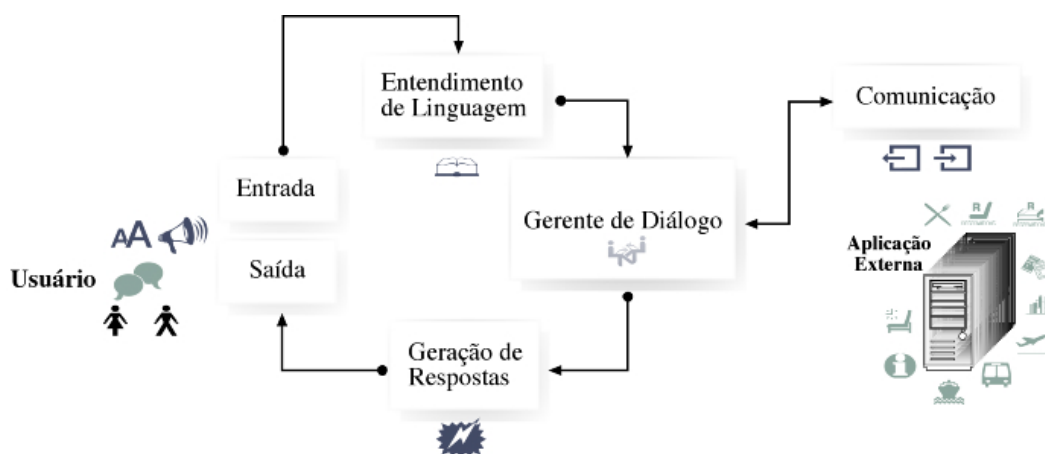


Figura 3.1: Arquitetura de um sistema de diálogo genérico

O fluxo de informações começa pelo reconhecimento da contribuição do usuário no

diálogo, a qual é normalmente convertida em texto no caso de contribuição falada, em seguida ocorrem etapas de análises lingüísticas para a extração de seu significado, o qual será convertido em um formato que represente suas características e que seja adequado ao tratamento do componente Gerente de Diálogo.

O Gerente de Diálogo tem como função principal controlar a interação entre sistema e usuário, para isso ele utiliza o formato representativo da contribuição do usuário para então acessar uma aplicação externa ao sistema de diálogo. As informações colhidas da aplicação externa (formato lógico) fornecerão meio para o sistema poder gerar sua contribuição, normalmente textuais. Ela é então processada pelo componente de geração de linguagem do sistema de diálogo e finalmente repassada ao usuário. Essa saída dependerá do tipo de sistema de diálogo, podendo ser textual, falada ou visual.

Na figura 3.2 é apresentada uma visão mais detalhada dos componentes, além da ligação entre as fontes de conhecimento e os componentes do sistema de diálogo.

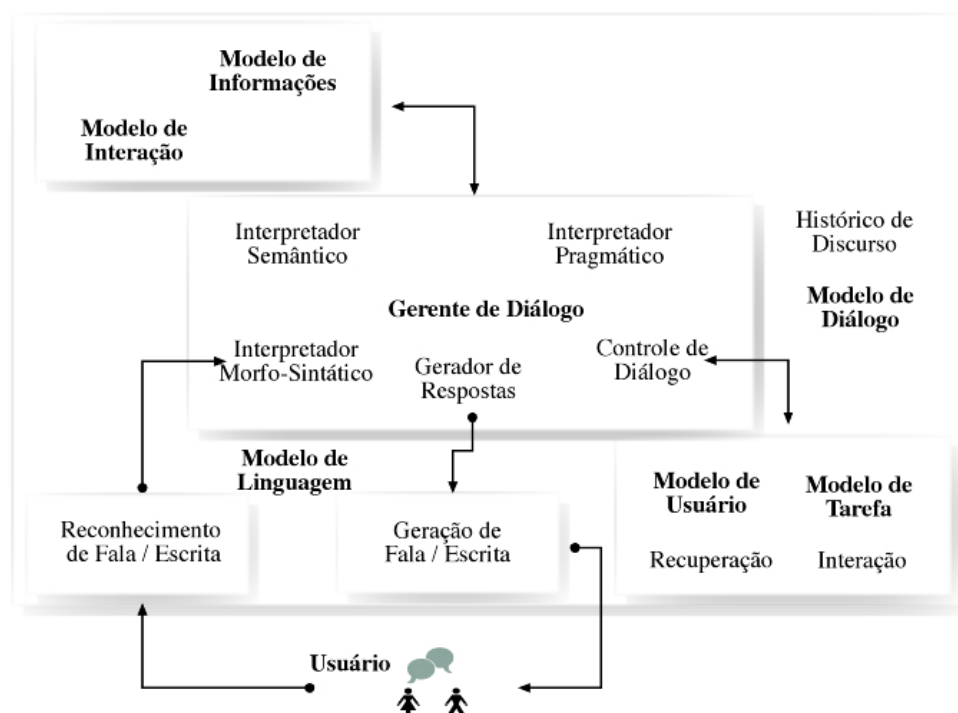


Figura 3.2: Arquitetura detalhada de um sistema de diálogo genérico

Os pontos novos são a definição de um modelo de linguagem que contém informações (dicionário e gramática) sobre a linguagem que o sistema interage com o usuário. Tal modelo é utilizado pelos componentes de Reconhecimento e Geração de linguagem. O detalhamento do componente de Gerência de Diálogo é outro ponto de detalhamento apresentado pela figura 3.2. Seus novos constituintes são:

- Modelo de conversação – contém as estratégias de interação e recuperação de erros, além dos modelos de tarefa e de diálogo.
- Interpretação semântica – contém os modelos de domínio e de interpretação baseada em atos de fala ou outra abordagem que expresse as intenções do usuário.
- Histórico de interação – contém os modelos de discurso e de usuário.

Os componentes de Reconhecimento de fala/texto e de Saída falada/textual não serão discutidos neste trabalho, uma geral é apresentada por McTear em [16]. Os demais componentes serão apresentados em mais detalhes a seguir.

3.4.1 Entendimento de linguagem

A função principal deste componente é analisar a saída do componente de reconhecimento de fala/texto e gerar uma representação que expresse seu significado e que possa ser manipulada pelo controle de diálogo. Essa análise segue os moldes da teoria apresentada no capítulo sobre entendimento de linguagem, dessa forma, para a obtenção dessa representação significativa as análises sintática, semântica e pragmática podem ser utilizadas de maneira consecutiva ou não.

A análise adotada por este componente pode utilizar informações sintáticas, semânticas e pragmáticas sem uma análise sintática anterior ou ainda sintática e semântica sem a necessidade da análise pragmática. A decisão de quantas e quais análises utilizar dependerá da abordagem de diálogo escolhida para o sistema de diálogo e a consideração sobre o aumento de complexidade geral da análise em decorrência das possíveis

falhas inerentes a cada etapa de análise adicional.

Um fator importante, ainda sobre a análise geral, é qual língua será adotada pelo sistema. Isso porque uma língua possui aspectos especiais (morfológicos, sintáticos e semânticos) que influenciam as etapas de análise. Para tentar padronizar ou amenizar os efeitos destes aspectos sobre os demais componentes do sistema, monta-se um modelo de linguagem.

Um modelo de linguagem é uma representação de uma linguagem, dessa forma ele necessita representar as palavras e as regras que definem as relações entre elas. As regras adotadas por um modelo de linguagem são as regras gramaticais de uma língua. As palavras são armazenadas ou organizadas em uma estrutura denominada dicionário, a qual é responsável pela codificação precisa dos significados das palavras.

Este modelo pode ser comparado, mas não tomado como, uma gramática de uma língua. Pois, uma gramática contém as regras que definem a formação e características das palavras, e este modelo além de regras semelhantes, possui ainda um conjunto único de palavras válidas e usuais de uma língua, ou seja, as palavras em si e não somente as regras para sua formação. Cabe ressaltar que o dicionário adotado pelo modelo de linguagem teria de passar por um tratamento de retirada de ambigüidade circular³ para melhorar seu uso por componentes do sistema de diálogo.

Com a adoção de um modelo de linguagem este componente poderá lidar com mais de uma língua, desde que ela seja codificada de forma semelhante à língua corrente do sistema. Logo, além de permitir uma padronização de codificação dos aspectos dependentes de linguagem um modelo de linguagem pode também possibilitar a adoção de mais de uma língua em um sistema de diálogo.

A definição da representação significativa é um outro fator importante para a aná-

³Ambigüidade circular é considerada aqui como a definição circular do significado de uma palavra, por exemplo: suar = transpirar; transpirar = suar.

lise geral e ainda mais para a continuidade de transporte das informações adquiridas do usuário para o componente de gerente de diálogo. Essa representação é feita utilizando-se algum formalismo lógico, apesar de outras formas para essa representação serem possíveis. O formalismo mais comum para se representar o significado semântico de uma sentença é a lógica de cálculo de predicados de primeira ordem. Outras formas mais expressivas para a representação lógica são a lógica intencional de Montague ou a teoria de representação de discurso [16].

O último fator constituinte deste componente é a integração com o componente de reconhecimento de fala/texto. Ela pode ser feita serialmente, onde sua saída completa é repassada assim que ele finalize seu processamento ou paralelamente, onde saídas parciais são repassadas a cada etapa de reconhecimento para o componente de entendimento de linguagem. Se entrada for falada o componente de reconhecimento envia como saída um conjunto de possíveis textos reconhecidos, junto com um tipo de classificação que indique o grau de validade dessa entrada. Na entrada escrita o texto já é a saída do componente.

Existem duas formas de se realizar a integração, na primeira são enviadas ao componente de entendimento hipóteses de textos classificados e na segunda são atribuídas probabilidades às palavras do texto reconhecido. De posse dessas informações esse componente poderá verificar a validade gramatical da construção enviada pelo componente de reconhecimento [16].

Erros de reconhecimento da entrada do usuário, decorrentes do processo de reconhecimento e de entendimento, decorrentes de uma má interpretação dos textos enviados pelo componente de reconhecimento prejudicam a integração entre esses dois componentes. Para solucionar ou abrandar erros dessa natureza o componente de entendimento terá de possuir estratégias com a finalidade correção ou prevenção dos mesmos.

3.4.2 Gerente de diálogo

Sua função principal é controlar o fluxo do diálogo. Para isso o sistema precisa: determinar se ele possui informações suficientes sobre o usuário para iniciar a comunicação com o componente externo, comunicar-se efetivamente com o componente externo e repassar ao usuário, em linguagem natural, os dados obtidos junto ao componente externo.

A determinação de informações corretas sobre o usuário é um fator importante para sua comunicação com o componente externo, pois, se as informações adquiridas junto ao usuário não forem suficientes para ele fazer uma consulta ao componente externo o processo de comunicação falhará.

O fluxo das informações entre os componentes do sistema de diálogo é fator importante para a determinação dessas informações. Para lidar com esse fator o componente de gerente de diálogo pode ser estruturado com dois tipos de arquitetura: serial ou distribuído. Esses dois são diferenciados pela forma de tratamento de entradas não usuais, aquelas que advêm de erros de processamento ou outro tipo de erro.

Em uma arquitetura serial o erro de interpretação será propagado de componente a componente até este ser detectado pelo gerente de diálogo, que fará a entrada ser processada novamente para que ele verifique novamente se o novo processo de interpretação está correto ou não.

Em uma arquitetura distribuída ou *blackboard* as informações de sub-processamento de cada componente podem ser requisitadas em qualquer ponto de seu processamento global por qualquer componente do sistema de diálogo, assim tanto o gerente de diálogo quanto os outros componentes que fazem parte do processamento global da entrada do usuário poderão efetuar correções sem sobrecarregar o processamento. Isso permite um controle de erros mais dinâmico e aprimorado, uma vez que possíveis erros, que seriam repassados a outros componentes, podem ser corrigidos imediatamente com a volta da informação errada ao componente que a gerou.

Além do tipo de arquitetura do sistema, o componente gerente de diálogo tem que possuir estratégias para o tratamento de entradas mal formadas, mal interpretadas ou insuficientes para o processo de comunicação com o componente externo e estratégias para a confirmação das informações obtidas [16].

Uma forma simples de lidar com erros é informá-los ao usuário assim que eles forem detectados, porém esta solução acaba tirando a naturalidade do diálogo, pois o sistema apenas informa que houve um erro de processamento, o que interrompe o fluxo normal do diálogo e coloca no usuário uma obrigação de tentar corrigir o erro cometido pelo sistema ou por ele mesmo. Uma maneira diferente e menos intrusiva seria possuir estratégias de recuperação implícitas baseadas nos conhecimentos disponibilizados pelos componentes de reconhecimento e entendimento de linguagem, além das fontes de conhecimento [16]. Assim, o sistema poderia usá-las para tentar corrigir ou requisitar ao usuário uma correção para o erro, algo natural mesmo em diálogos entre pessoas.

Com um modelo de discurso integrado ao sistema de diálogo o componente gerente de diálogo pode obter informações sobre o contexto do diálogo, o que permite a interpretação de itens dependentes de contexto, tais como itens referenciados por pronomes e por alguns fenômenos lingüísticos, característicos de discurso, que são dependentes de contexto local, anáforas e elipses. Essa interpretação recebe o nome de análise pragmática e por meio dela a determinação das informações do usuário podem ser mais bem esclarecidas, uma vez que se sabe o contexto em que elas estão.

Outra estratégia de correção de entradas é a pré-determinação das emissões seguintes do usuário, assim se o sistema conseguir prever as próximas emissões ele poderá diminuir ou eliminar o processamento necessário para interpretação de uma emissão, pois ele já sabe qual será a próxima. Essa predição pode ser feita com base no modelo de diálogo, no histórico de discurso e com um modelo de usuário.

Com o modelo de diálogo o sistema sabe em que ponto o diálogo está, com o

histórico de discurso ele saberá quais foram as emissões anteriores e com o modelo de usuário (crenças, objetivos, preferências e planos) ele poderá determinar possíveis planos que expressam os objetivos do usuário, assim tentar prever quais seriam as futuras emissões do usuário.

As estratégias apresentadas anteriormente tentam tratar de entradas mal interpretadas ou incompletas sem a intervenção do usuário, contudo nem sempre o sistema conseguirá corrigir sozinho esse tipo de erro. Então estratégias de verificação e confirmação são necessárias. A verificação é necessária para lidar com prováveis erros de interpretação e entendimento que o sistema de diálogo terá de lidar, já a confirmação é mais comum em sistemas de diálogo falado, onde fatores tecnológicos podem influenciar a interpretação das emissões.

A verificação pode ser implícita ou explícita. Na primeira o sistema adiciona na sua próxima pergunta uma repetição do que ele entendeu da emissão anterior do usuário, isso permite ao usuário corrigir o que não foi entendido corretamente, caso não o faça, a verificação será confirmada implicitamente. Na segunda o sistema pergunta explicitamente ao usuário para requerer confirmação da sua entrada.

A confirmação pode ser obtida por uma resposta sim ou não. Esta verificação pode ter a desvantagem de se a pergunta de confirmação for mal formulada ou estiver fora de contexto local o usuário será forçado a corrigir a confirmação e a pergunta de confirmação também. Uma dificuldade desse tipo de verificação é a grande possibilidade de respostas que o usuário pode fornecer, o que dificultará as futuras confirmações do sistema.

3.4.3 Comunicação com sistema externo

Um sistema de diálogo normalmente realiza alguma tarefa que necessita de informações provenientes de uma ou mais fontes externas, tais fontes podem ser compreendidas como qualquer tipo de aplicação externa que possa ser associada a um sistema

de diálogo. Algumas dessas fontes podem ser classificadas dentro dos seguintes grupos: bancos de dados, sistemas especialistas e bases de conhecimento.

3.4.3.1 Comunicação com banco de dados

Esta comunicação permite ao usuário acessar informações armazenadas em um banco de dados. A comunicação em si não representa um motivo para muitos comentários quando se trata de sistemas de diálogo [16], pois existem pesquisas específicas sobre interface em linguagem natural para banco de dados.

Para realizar a integração entre um sistema de diálogo e um banco de dados, além do componente de comunicação pode-se necessitar de um componente a mais, chamado gerente de informação, o qual é responsável pelo acesso direto ao banco de dados e também pela interligação com o componente de comunicação externa ou diretamente com o componente gerente de diálogo.

Um aspecto importante para a comunicação com o banco de dados é o mapeamento entre o seu vocabulário e o do diálogo, isso porque se o mapeamento do diálogo não puder ser realizado diretamente pelo vocabulário do banco, então a resposta produzida com as informações colhidas da consulta poderá ser incompleta ou mesmo ambígua.

Alguns dos problemas relacionados a comunicação baseada em banco de dados podem ser encontrados: no mapeamento, decorrentes de mau entendimentos ou interpretações por parte do usuário sobre estrutura que representa o domínio no banco de dados, na forma de tratamento pelo componente de comunicação da situação de não encontro de informações requeridas junto ao banco de dados e na produção das informações advindas do banco de dados ao usuário.

3.4.3.2 Comunicação com base de conhecimento

Sistemas de diálogo que tem como aplicação a resolução de problemas necessitam de conhecimentos fornecidos por uma base de conhecimento, que é uma estrutura complexa que armazena conhecimentos específicos sobre um domínio. Ela é diferente de um banco de dados por não seguir uma regra de armazenamento e recuperação baseada em álgebra e teoria de conjuntos, podendo utilizar teorias diversas tanto para armazenamento quanto para recuperação.

Para a integração com essa base utilizam-se componentes comuns em sistemas baseados em conhecimentos (processador de domínio, módulo de raciocínio geral e módulo de conhecimento geral) [16]. Esses novos componentes serão acrescentados ao componente de comunicação externa ou diretamente ao gerente de diálogo como sub-componentes.

Considerando a adição ao componente de comunicação externa, o processador de domínio será responsável pela recomendação ao componente de comunicação e posteriormente ao gerente de diálogo dos passos necessários a resolução da tarefa do sistema, uma vez que ele é um componente dependente de domínio e que contém toda a informação da aplicação necessária para a recomendação. Essas informações são compartilhadas com o módulo de raciocínio, que as utiliza como fonte de seus mecanismos genéricos de resolução, os quais são independentes do domínio das informações. Já o módulo de conhecimento geral possui conhecimentos relevantes (lingüísticos e de interação gerais em diálogo) para a manutenção de diálogos.

3.4.3.3 Comunicação com um sistema especialista

Um sistema especialista é um sistema que tenta simular o conhecimento de um especialista humano em uma determinada área do conhecimento geral. Assim as crenças e regras utilizadas pelo especialista para resolver um problema são repassadas para um sistema que as usará de forma semelhante.

Um sistema de diálogo que forneça uma interface em linguagem natural para um sistema especialista necessitará do acréscimo de um sub-componente, chamado de componente de raciocínio baseado em planos, ao componente de comunicação externa. O qual será responsável pelo processo de reconhecimento de plano e elaboração de planos para resolução de tarefa.

Assim o sistema de diálogo passará a lógica de interação do componente gerente de diálogo para esse novo componente, restando a ele a tarefa de identificar nas contribuições do usuário os elementos (crenças) que sejam necessários para a comunicação com o sistema especialista.

Essas crenças serão utilizadas pelo componente de raciocínio baseado em planos para montar um plano cujas ações serão repassadas ao sistema especialista. As informações colhidas junto a ele serão incorporadas a este plano e com isso a resolução da tarefa do usuário será determinada pela resolução do sistema especialista e pela completude deste plano. Este plano tem de estar de acordo com a estrutura de resolução do sistema especialista para a comunicação funcionar.

3.4.4 Geração de linguagem natural

A geração de linguagem é um processo de transformação de uma estrutura significativa (lógica) em um conjunto de sentenças de uma linguagem natural. Essa estrutura é gerada ou fornecida pelo Gerente de Diálogo e pode ser originada das informações obtidas junto ao componente de Comunicação Externa ou seguindo a abordagem de diálogo adotada pelo sistema.

O processo de geração é dividido em três etapas, segundo Theune [30], onde cada uma possui desdobramentos importantes.

- Planejamento do texto – primeiro, as informações obtidas de outros componentes são classificadas de sorte que seu conteúdo seja normalizado. Segundo, após a

normalização as informações são ordenadas para representar uma estrutura da saída do processo de geração.

- Microplanejamento – a primeira tarefa dessa etapa consiste da escolha das palavras que melhor representem os conceitos apresentados na estrutura de saída, levando em consideração um dado contexto. Na segunda ocorre uma decisão de verificar dentro da estrutura de saída quais informações podem ser agrupadas para formarem uma sentença coesa. Por fim, informações (pronomes, descrições literais, etc.) que possam descrever entidades de um contexto, serão utilizadas para criar novas frases.
- Realização de superfície – nesta etapa existem dois tipos de realizações, a lingüística e a de estrutura. Na primeira estão relacionados os aspectos gramaticais (regras sintáticas e morfológicas) envolvidos na criação de sentenças gramaticais. E na segunda ocorre a conversão das sentenças em algum formato de saída utilizado por um sistema de diálogo (arquivo texto, informações para um componente de sintetização, conversão em imagens, etc.).

Na primeira etapa as informações são independentes de linguagem, na segunda tanto informação dependente de linguagem quanto dependente de contextos serão necessárias para a definição de sentenças. E na terceira etapa, mais especificamente na realização lingüística, as informações são dependentes de linguagem, pois neste caso trata-se da apresentação formal e gramatical da mensagem de saída [30].

Em sistemas de diálogo, a geração de linguagem não necessita de todas as etapas apresentadas anteriormente. Apenas as etapas de microplanejamento e realização de superfície, uma vez que as informações a serem repassadas ao usuário já foram trabalhadas pelo Gerente de Diálogo, o que dispensa a utilização da etapa de planejamento de texto, pelo menos como uma parte conjunta do processo de geração de linguagem.

Theune [30] considera que os sistemas de diálogo focados a apenas um domínio não necessitam de todas as etapas da geração de linguagem, pois, alguns desses necessitam tão somente de formação de sentenças simples (sistema de perguntas e respostas). O

que envolve a determinação da estrutura coesa e gramatical da sentença, duas partes das etapas da geração de linguagem. E para isso, o uso de certas técnicas para a manipulação de palavras é o suficiente para produzir saídas em linguagem natural.

Algumas dessas técnicas são apresentadas por McTear [16]:

- Texto inalterável – utiliza partes das informações advindas do Gerente de Diálogo para a construção da mensagem.
- Dados disponíveis diretamente – uso de informações obtidas diretamente, por meio do componente de Comunicação Externa (banco de dados ou base de conhecimento), para a construção da estrutura da sentença de saída.
- Dados computáveis – utilização de informações derivadas de algum processo de racionalização adotado pelo sistema de diálogo para a determinação das palavras a serem integradas na estrutura de saída.
- Dados não disponíveis – inclusão de informações complementares, ou seja, informações que não estão presentes diretamente nos dados resultantes, mas que possuem relevância para a resposta ao usuário.

Contudo para sistemas mais complexos tais técnicas não são suficientes, pois elas não produzem sentenças flexíveis e dependentes de contexto, algo que se assemelha a que os humanos fazem quando dialogam. Segundo Theune [30] alguns pontos relevantes da geração de linguagem em diálogos por humanos são os seguintes: estilo de diálogo, relação entre falante e ouvinte, estados cognitivo, entre outros.

Por tais motivos Theune [30] considera quatro tipos de abordagens como mais condizentes com a geração complexa de diálogo. Elas são: análise reversa, realização baseada em gramática, realização baseada em moldes⁴ e geração completa.

⁴Termo em inglês *Template*

3.4.4.1 Análise reversa

A análise de linguagem transforma uma sentença em uma representação que expresse seu significado. A geração de linguagem realiza o caminho inverso, a partir de uma representação ela atribui os componentes lingüísticos que expressem seu significado. Partindo dessa idéia poderia ser utilizada a mesma gramática de interpretação sintática para a geração sintática e a mesma representação significativa de contextos, a qual é produzida pelo Gerente de Diálogo e fornecida, via usuário pelo componente de Entendimento de Linguagem para a geração de linguagem natural.

Esse tipo de abordagem para geração trás algumas vantagens: simplicidade de implementação, pois somente uma gramática seria utilizada no sistema, uso de um padrão para a representação significativa de conceitos, aceitação psicológica, uma vez que o sistema entenderia o que gera e interpreta e possibilidade de utilização de componentes diferentes para a sintetização de fala [30].

Entretanto, a gramática utilizada para o entendimento de linguagem é desenvolvida especialmente para esta tarefa, utilizando adaptações que permitam o entendimento de sentenças que não sejam totalmente gramaticais. Isso invalidaria o uso de tal gramática para a geração, pois nela as sentenças produzidas têm de estar de acordo com a gramaticalidade da linguagem adotada pelo sistema.

Um outro fator negativo é a ambigüidade que a representação significativa pode apresentar quando na geração de linguagem. Isso é explicado pela própria ambigüidade de uma linguagem natural, onde uma palavra pode expressar mais de um conceito, e uma reunião de conceitos podem determinar mais de uma sentença que expresse o mesmo significado dependente de contexto.

Sistemas de diálogo que possuam um domínio específico, logo um vocabulário fechado, podem se valer dessa abordagem. Pois, eles podem determinar um número finito de sentenças a serem produzidas e ainda possuir poucas palavras a serem agregadas a algumas estruturas de saída em linguagem natural. Todavia, para estes tipos

de sistema esta abordagem apresenta uma desvantagem marcante, a falta de um mecanismo de produção de sentenças que leve em consideração fatores não significativos, ou proposicionais, tais como estilo de linguagem, estrutura de informações, etc. Portanto, uma representação significativa será sempre mapeada para uma mesma sentença [30].

3.4.4.2 Realização baseada em gramática

Utiliza uma gramática especializada para a geração de uma linguagem natural, o que evita muitos dos problemas de uso de gramática apresentados anteriormente. Contudo, essa abordagem ainda apresenta algumas desvantagens, porém todas de natureza prática e não teórica.

O tipo de representação significativa usada como entrada necessita de certas informações lingüísticas que não estão disponíveis em representações significativas que utilizam lógica para sua expressão. A resolução desse problema é utilizar técnicas específicas para a etapa de microplanejamento, onde elas adicionam informações em falta (tempo, número, grau, gênero, etc.) na representação significativa fornecida pelo componente de Gerência de Diálogo.

Porém, essa medida causa uma outra desvantagem prática ao uso de realização baseada em gramática, a lentidão do processo de geração de linguagem, o que pode atrapalhar o diálogo entre usuário e sistema [30].

3.4.4.3 Realização baseada em moldes

Outra forma de realização lingüística é utilizar um conjunto de estruturas semi-preenchidas de sentenças pré-definidas ou regras que expressam como produzir realizações lingüísticas [31] denominados moldes, para produzir sentenças em linguagem natural. Isso permite o uso de menos recursos gramaticais para a geração de sentenças, pois os detalhes lingüísticos da etapa de microplanejamento são codificados na estrutura dos moldes.

Sendo assim, a representação significativa, mesmo a que utiliza algum tipo de formalismo lógico, pode ser utilizada sem maiores adaptações no processo de geração de linguagem, uma vez que os recursos lingüísticos necessários para possíveis adaptações estarão presentes nos moldes.

Apesar da vantagem de praticidade de implementação, esta abordagem apresenta a desvantagem de necessitar de codificação antecipada dos possíveis grupos de sentenças a serem gerados por este componente. Também como na abordagem anterior ainda é necessário que o componente Gerente de Diálogo forneça informações que seriam tratadas pela etapa de planejamento de texto do processo de geração de linguagem.

Outra desvantagem é a necessidade de adaptação ou mesmo reconstrução dos moldes de um sistema de diálogo. Isso ocorre por conseqüência da determinação de moldes ser dependente de informações específicas de domínio e linguagem [30].

3.4.4.4 Geração completa

Para sistemas de diálogo simples o componente de geração de linguagem utiliza apenas uma ou no máximo duas etapas do processo de geração de linguagem, realização e microplanejamento. Todavia, para diálogos mais complexos o uso de todas as etapas pode ser necessário. Assim, com a utilização de todas as etapas de geração de linguagem o componente de geração poderia receber informações que não contivessem aspectos lingüísticos e ainda assim gerar sentenças gramaticais.

Contudo o uso de todas as etapas do processo de geração em um componente de geração de linguagem natural acarreta em uma cobertura maior do que a necessária para um sistema de diálogo, mesmo os que necessitem lidar com diálogos complexos, pois a união de todas as etapas do processo de geração de linguagem normalmente gera sentenças adequadas a um texto e não a um diálogo [30].

Então para o uso adequado de um componente de geração de linguagem, adaptações serão necessárias para que a saída textual deste componente se integre com os demais componentes do sistema e assim permite a geração de sentenças para um diálogo. Isto pode ser alcançado pela adaptação de uma de suas etapas à estrutura de histórico de diálogo, mantida pelo componente de Gerência de Dialogo, o que adequaria a saída textual ao tipo de coesão encontrada em diálogos.

3.5 Controle de diálogo

Um diálogo consiste de elementos que se referem a uma tarefa de base, tarefas comunicativas e a aspectos de controle da comunicação, tais como monitorar a atenção, tomada de vez, reconhecimento e entendimento, recuperação de falhas de comunicação, entre outras. O controle de dialogo é um termo que se refere a esses aspectos [32].

Uma tarefa de base refere-se a um objetivo que uma pessoa possui. Dependendo do tipo de objetivo uma das formas de alcançá-la será por meio da comunicação. Na comunicação existem alguns fatores que representam objetivos comunicativos, tais fatores podem ser utilizados para o controle da comunicação. Assim, por meio de objetivos comunicativos podem-se alcançar objetivos referentes a uma tarefa de base.

Esses objetivos comunicativos podem ser entendidos como as unidades segmentadas de uma tarefa, apresentadas por Allwood [33, 34]. Onde uma tarefa de base pode ser dividida em sub-tarefas de menor complexidade. Com essa divisão a comunicação pode ser entendida como seqüências de atos comunicativos.

O controle de diálogo pode ser obtido por quatro tipos de abordagens sobre o formato da comunicação via diálogo. Elas são: medidas de confiança no reconhecimento de fala [35], a distribuição da iniciativa do diálogo [16], contexto do diálogo [32] e diferença entre custo de confirmação e transferência de informação do sistema [36].

Na primeira abordagem o sistema aceita a entrada do usuário caso ela contenha um elevado grau de confiança, se ela contiver um pequeno grau de confiança o sistema tentará recuperar o conteúdo da entrada utilizando uma estratégia de confirmação. Tal estratégia pode ser rejeitar a entrada e sempre pedir confirmação direta ou indireta sobre o conteúdo da mesma.

O formato de medição do grau de confiança dependerá da estratégia adotada pelo sistema de diálogo, entretanto ela terá de considerar a relevância da informação aceita pelo sistema, o número de troca de vezes entre sistema e usuário. E mais, o valor da relevância da informação pode ser aumentado pelo uso de confirmação direta enquanto que o número de troca de vez é diminuído pela adoção de confirmação indireta. Niimi em [35] comprova a última afirmação.

Uma vez é o direito de um participante falar. Uma vez difere de uma emissão, por nela não ser necessário emitir nada para que um participante a tenha. É por um conjunto de normas que se determina o direito de fala, tais normas podem expressar: quem tem a direito a vez, sobre qual tópico a vez se refere, a maneira de se falar, o tempo para se ter a vez e duração que uma vez pode ter [33, 34].

Em relação à distribuição da iniciativa em diálogo, o controle de diálogo é feito pela manutenção de qual dos participantes do diálogo possui a vez de fala e mantém o ritmo de tomada de vez. Existem três tipos iniciativa: iniciativa do usuário, onde ele controla a desenrolar do diálogo fazendo perguntas e esperando respostas do sistema, que apenas espera por perguntas do usuário e não tem interesse em tomar o controle do diálogo. O segundo tipo é a iniciativa do sistema, onde o sistema tentará conduzir um diálogo com o usuário. Por fim a iniciativa mista, onde ambos podem ter a iniciativa de controlar o diálogo, algo semelhante a uma conversa entre humanos.

Segundo Allwood [33, 34] a importância da manutenção de iniciativa é motivada por requisitos que possibilitam um diálogo racional, cooperativo e motivacionado, além da necessidade de se ter convenções que suportem tais requisitos. Também, outra mo-

tivação são restrições físicas e psicológicas do ser humano, que não consegue processar e enviar muitas mensagens ao mesmo tempo.

O controle de iniciativa é importante para a determinação das ações que um sistema de diálogo pode executar, assim considerando os tipos de abordagens para representação de diálogo apresentadas na seção 3.2, pode-se identificar o tipo de iniciativa suportada que elas permitem a um sistema de diálogo. Para sistemas de diálogo baseados em estados finitos a iniciativa é do sistema, em sistemas baseados em quadros é possível se ter iniciativa tanto do usuário quanto do sistema, e em sistemas baseados em agentes e em princípios racionais a iniciativa poderá ser mista [16].

Um diálogo pode ser visto como uma troca de ações, onde a linguagem é uma forma de executar ações que modifiquem um contexto. E mais, dentro de um diálogo existem unidades funcionais que permitem um participante modificar o contexto, essas unidades não correspondem diretamente às falas, pois, estas são multifuncionais por natureza. Essa visão de diálogo como um meio de modificar o contexto é apresentada por Bunt em [32] com o nome de teoria de interpretação dinâmica.

Contexto é considerado aqui como o conjunto de fatores que são relevantes para a compreensão do comportamento comunicativo. Alguns desses fatores são: a natureza e estado corrente da tarefa de base que iniciou o diálogo, o que foi dito anteriormente, as relações sociais entre os participantes e o conhecimento que um participante tem do outro. Tais fatores podem ser agrupados em cinco categorias de contexto: lingüístico, semântico, físico, social e cognitivo [32].

As unidades funcionais que modificam o contexto podem ser distinguidas de acordo com sua função. Algumas são orientadas a tarefa de base e outras específicas para o controle do diálogo, ou seja, enquanto que certas funções são responsáveis pela modificação do contexto semântico, o que representa as informações a serem transmitidas, outras são responsáveis pela modificação das demais categorias de contexto sem alterar o contexto semântico. Cada aspecto relevante para o controle de diálogo poderá

ser representado por uma unidade funcional que expressa uma forma particular de modificação de alguma classe de contexto [32].

A última abordagem, apresentada por Dohsaka em [36], considera que em diálogos é comum o sistema não reconhecer a fala do usuário. Para corrigir este problema o sistema tenta confirmar a partir do reconhecimento de partes da fala do usuário o reconhecimento completo do que foi dito por ele, isto leva ao chamado diálogo de confirmação. Quanto menor o tamanho desse tipo de diálogo melhor para a comunicação entre usuário e sistema.

Contudo, existem casos onde ocorrem erros de reconhecimento durante um diálogo de confirmação, o que levará a uma cadeia de novas confirmações e em conseqüência ao aumento do diálogo. Um desses casos é o de o sistema tentar confirmar informações que estão fora do seu limite de conhecimento, ou seja, o sistema não faz uma interpretação real do que está tentando confirmar.

Assim, para casos onde o sistema entende o que recebe do usuário, mas seus conhecimentos não são suficientes para confirmar certas entradas mal entendidas, Dohsaka em [36] propõem uma forma de controle de diálogo de confirmação que considere a avaliação de dois valores que expressam como determinar dentro da entrada do usuário quais porções serão utilizadas para gerar de respostas para confirmação (custo de confirmação e custo de transferência de informação).

3.6 Conclusão

Para o desenvolvimento e pesquisa de um sistema de diálogo genérico, seja ele falado ou escrito, pode-se seguir a maioria do que foi apresentado neste capítulo e a seguir serão apresentadas as bases para o início de uma pesquisa que visa o desenvolvimento de um sistema de diálogo genérico textual independente de domínio.

Capítulo 4

Trabalhos correlatos

No capítulo anterior foram mostrados alguns dos conceitos e características que definem um sistema de diálogo genérico. Existem vários sistemas de diálogo apontados pela literatura desta área, contudo os apresentados neste capítulo servirão de substrato para as idéias que serão expostas no capítulo seguinte.

Os sistemas a seguir não seguem uma ordem temporal e nem de importância. Alguns deles serão mais detalhados outros apenas mencionados, isso porque existem diversos sistemas de diálogo e a suas descrições completas não são parte dos objetivos deste trabalho.

4.1 Sistemas de diálogo

4.1.1 TRAINS

Este projeto teve início em novembro de 1994 [37] e tinha como objetivo principal provar a viabilidade de se modelar e posteriormente utilizar sistemas de diálogo em domínios orientados à tarefa, aqueles onde os conhecimento necessário à resolução de um problema é dependente de um domínio em particular.

A partir do aprimoramento de técnicas extraídas de teorias bem fundamentadas,

mas que ainda não tinham sido empregadas na prática [37], poderia-se desenvolver um sistemas de diálogo robusto o bastante para interagir com humanos quase de forma natural.

Para alcançar esse objetivo os projetistas do sistema de diálogo se decidiram a usar uma tarefa simples cujo domínio pudesse ser bem determinado. Desse modo, foi escolhido um domínio dos sistemas viários de trens e como tarefa o planejamento do itinerário das viagens de trens entre cidades nos Estados Unidos.

O projeto tinha ainda os seguintes objetivos [38]:

- Fazer um estudo de diálogos homem-homem.
- Construir vários protótipos que pudessem ser utilizados por usuários não treinados.
- Uso dos resultados obtidos dos protótipos para desenvolver uma plataforma para entendimento de linguagem natural e planejamento colaborativo de uma tarefa específica.

A figura 4.1, extraída de [37], apresenta a arquitetura TRAINS-95, uma explicação sobre os módulos que compõem esta arquitetura será fornecida a seguir. Houve modificações nesta arquitetura, porém elas não mudaram a base da arquitetura, verifique em [39].

Para o reconhecimento de voz foi utilizado o *Sphinx-II system* e para a sintetização ou geração de fala o *TRUETALK system*. Um módulo denominado Interface possibilita a comunicação (feita com a utilização de uma linguagem interna) com outros módulos do sistema, o controle das informações a serem apresentadas ao usuário e o controle da interação via teclado, que é outra possibilidade de interação do usuário com o sistema.

O módulo de Reconhecimento de Fala tem como saída dados estatísticos referentes à entrada falada do usuário, porém podem existir erros nesse reconhecimento e dos dados podem não representar corretamente o que foi dito pelo usuário, para resolver

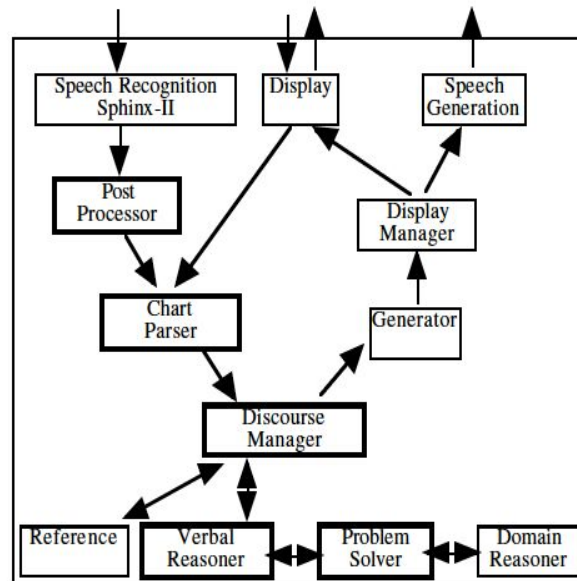


Figura 4.1: Sistema TRAINS-95

este possível problema o sistema utiliza algumas das técnicas usadas na área de tradução automática. Essa correção é feita no módulo chamado Pós-Processador.

As entradas do usuário são repassadas ao módulo denominado Etiquetador, que ao receber dados fornecidos pela interação falada, via Pós-Processador, e da interação por teclado, via Interface, gera um conjunto de atos de fala que serão interpretados no módulo chamado de Gerente de Discurso.

Esse módulo possui sub-componentes que lidam com fatores: lingüísticos e referentes à tarefa do sistema. Esses sub-componentes são: Referência, Planejador, Resolução e Domínio. O primeiro lida com fatores lingüísticos e os demais com fatores referentes à tarefa do sistema.

O sub-componente Planejador é responsável pela determinação de um itinerário mais adequado às entradas fornecidas pelo usuário. Isso é feito pela interação com o usuário, onde o sistema conhece as rotas e o usuário fornece os parâmetros (nomes de cidades ou estações) para estabelecer qual a rota que melhor lhe agrada, dentre as que

o sistema apresenta em cada interação. Dessa forma a tarefa (estabelecer um itinerário de viagem) é cumprida com contribuições de ambas as partes.

O processo de geração de resposta do sistema ao usuário tem início quando o módulo denominado Gerador recebe do Gerente de Discurso um conjunto de atos de fala que correspondam à resposta do sistema. Por meio desse conjunto o Gerador constrói uma sentença que será enviada aos módulos de Geração de Fala e Interface. Isso porque o sistema gera saída tanto auditiva quanto visual.

A construção de uma sentença é baseada na utilização de padrões simples de respostas, onde cada um deles é associado a diferentes formas dos atos de falas, os quais descrevem classes de objetos específicos ao domínio da tarefa.

Após a rápida apresentação da arquitetura do sistema TRAINS, pode-se classificá-lo como um sistema de diálogo baseado em planos de iniciativa mista, apontar algumas limitações, muitas referentes as técnicas adotadas para o processamento de sinal auditivo e pelas técnicas utilizadas para manipulação de atos de fala, pois o erro na identificação da fala causa o aumento na duração da conversa entre usuário e sistema; e a escolha errada de atos de fala para representar a entrada do usuário (sua intenção) pode causar erros no prosseguimento da resolução da tarefa.

A principal limitação apontada por Allen em [37] está no seu módulo de planejamento de itinerários, que necessita de interação com o usuário para a completa resolução da tarefa de programação do itinerário de uma viagem de trem.

Porém a simplificação da modelagem da tarefa escolhida demonstrou que um dos objetivos do sistema, provar a viabilidade de conversação e não a melhor forma de se modelar uma tarefa, era possível [37].

4.1.2 TRIPS (*The Rochester Interactive Planning System* / Sistema de Planejamento Interativo da Universidade de Rochester)

A continuação das pesquisas do projeto TRAINS forneceram meios para a criação de um novo projeto denominado de TRIPS. Nesse novo projeto o objetivo tomou a direção da colaboração como forma de melhoria no planejamento de tarefas complexas. A colaboração ocorre por meio de diálogo em linguagem natural e interface gráfica [40].

Segundo Ferguson [41] as principais diferenças entre o projeto TRIPS e o TRAINS são as seguintes. Outras diferenças são descritas em [42, 40].

- As tarefas do TRIPS são mais complexas, reflexo da modelagem mais detalhada e da escolha de domínios mais complexos. No TRAINS o domínio da tarefa era o de planejamento de rotas de trens entre cidades.
- Os tipos de planos suportados pelo TRIPS são mais complexos do que os do TRAINS, tais como: manipulação de recursos baseada em restrições, racionalização de ações temporais e coordenação de ações de múltiplos agentes.
- Adoção de modelos de colaboração para resolução de problemas mais complexos.

A arquitetura do sistema TRIPS é apresentada na figura 4.2, extraída de [43], e sua descrição será apresentada a seguir.

Cada componente da arquitetura do sistema é representado por um agente computacional responsável pelo cumprimento de um objeto específico. Esses agentes interagem entre si por meio de troca de mensagens, as quais permitem a delimitação de fluxos de informações em áreas de processamento. Na arquitetura do sistema existem três áreas identificáveis: interpretação, comportamento e geração [43].

Uma interação entre usuário e sistema começa pelo processamento de informações auditivas ou via teclado pelo componente de Análise. Atos de fala convencionais são

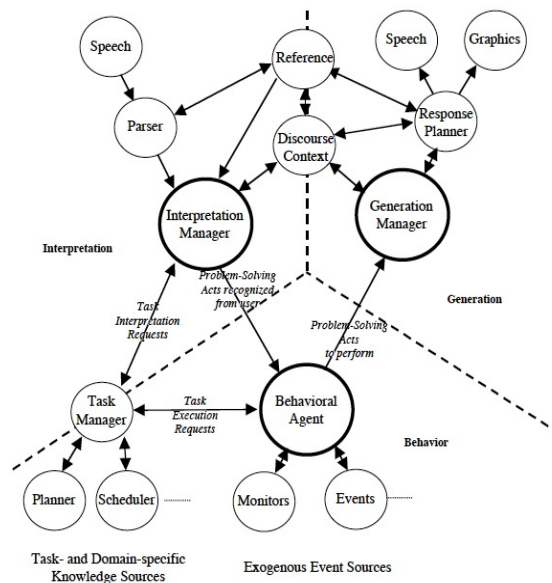


Figura 4.2: Sistema TRIPS

gerados como resultado do processamento desse componente. Esses atos são utilizados para manter referências de elementos introduzidos anteriormente no decorrer de um diálogo, isso é responsabilidade do componente denominado Referências.

Esses atos são convertidos para atos de fala específicos referentes à interpretação do componente Gerente de Interpretação do estado de resolução da tarefa do sistema. O processo de interpretação necessita de informações sobre as intenções do usuário (Análise), das referências de diálogo (Referências), do estado da resolução da tarefa (Gerência de Tarefa) e de informações sobre o contexto da tarefa ou diálogo (Contexto de Discurso).

Esses processos realizados são parte da área de interpretação. A continuação da interação do usuário ocorre na área de comportamento, onde o componente chamado de Agente Comportamental é responsável pelas ações do sistema e pela comunicação com sistemas externos.

As informações advindas do Gerente de Interpretação são refinadas e utilizadas

para decidir que ações o sistema tomará. Essa decisão é baseada nos planos mantidos pelo Gerente de Tarefa e pelos atos de fala quem expressão as intenções do usuário, obrigações e objetivos do sistema, além das mudanças no estado de resolução da tarefa mantidas pelo Gerente de Tarefa [43].

As ações do sistema que envolve interação com o usuário (clarificação, notificação de problemas, resposta ao usuário, etc.) são então repassadas ao Gerente de Geração, que planeja o conteúdo da fala do sistema e as saídas utilizadas para comunicá-las ao usuário. Este planejamento decorre de obrigações de discurso¹ (Contexto de Discurso) e das informações do Gerente de Geração [43].

No componente de Planejamento de Resposta as informações de ações do sistema, dos componentes de Referências e Contexto de Discurso são utilizadas para montar a resposta que o sistema fornecerá ao usuário. O uso dos dois últimos componentes é necessário para garantir a manutenção da coerência do diálogo. Por fim, a resposta montada é processada pelos componentes de Sinterização de Voz e Geração de Gráficos.

As informações necessárias à geração de respostas manipuladas pelos componentes do sistema definem a área de geração.

O projeto TRIPS, segundo seus autores, tem como objetivo ser um sistema de diálogo genérico, capaz de fornecer suporte a tarefas baseadas em plano em domínios diferentes. Para isto algumas versões do sistema estão sendo testadas [40].

Houve uma evolução desde do projeto original (TRAINS), onde várias técnicas foram empregadas para solucionar os problemas encontrados no desenvolvimento do primeiro projeto. Contudo ainda restam problemas a serem resolvidos e a adequação da noção de planejamento de identificação de intenções baseada em crenças para a colaboração utilizando a troca de recursos permite ao sistema melhores resultados no

¹São comportamentos sociais esperados durante uma conversa, por exemplo ao se cumprimentar alguém espera-se um cumprimento em resposta.

diálogo com o usuário, além de possibilitar a flexibilidade no tratamento de tarefas complexas.

4.1.3 Verbmobil

Este sistema advém de um projeto que tem como objetivo auxiliar a comunicação entre dois falantes humanos de línguas nativas diferentes, um tendo o alemão e o outro o japonês como língua nativa, com a utilização de uma terceira língua, o inglês. Isso permite aos participantes se comunicarem sem a necessidade de conhecerem a língua do outro, bastando apenas conhecimento de uma língua em comum [44].

O sistema se vale do não conhecimento total da língua em comum, pelo menos não o que um deles teria se fosse sua própria língua nativa. Dessa forma, cada participante pode emitir falas em sua língua nativa e o sistema se encarregará de traduzi-las para o inglês. Logo, eles poderão verificar se o sistema gerou a tradução mais adequada para sua fala e no caso de negativo corrigi-la se necessário. Pode-se entender este sistema como tendo a mesma função que um interprete humano tem.

O início deste projeto ocorreu em 1993 e foi gerenciado pela *German Federal Ministry for Education, Science, Research and Technology* que o dividiu em duas fases distintas: Fase 1 de 1993 até 1996 e Fase 2 de 1997 até 2000 [45].

Na fase 1 foi desenvolvido o protótipo *Verbmobil Demonstrador* que traduzia do alemão e japonês para o inglês no domínio de agendamento de encontros de negócios e possuía uma arquitetura baseada em multi-agentes. Essa tradução era feita pela junção de várias técnicas de tradicionais de tradução (baseada em exemplos, baseada em atos de diálogo e baseada em estatística).

Na fase 2 o sistema já era capaz de traduzir bidirecionalmente do alemão para o inglês e do japonês para os japoneses diálogos com emissões espontâneas, ou seja, falas não previstas pelos projetistas podiam ser traduzidas.

Os objetivos da segunda fase eram [45]: a adaptação rápida a novos domínios de discurso, tradução de fala espontânea em várias línguas, oferecer assistência em traduções em aplicações internacionais, ser acessível via celular e a assistência de tradução não se limitaria a diálogos entre dois participantes somente.

A arquitetura do projeto Verbmobil é apresentada na figura 4.3, extraída de [44].

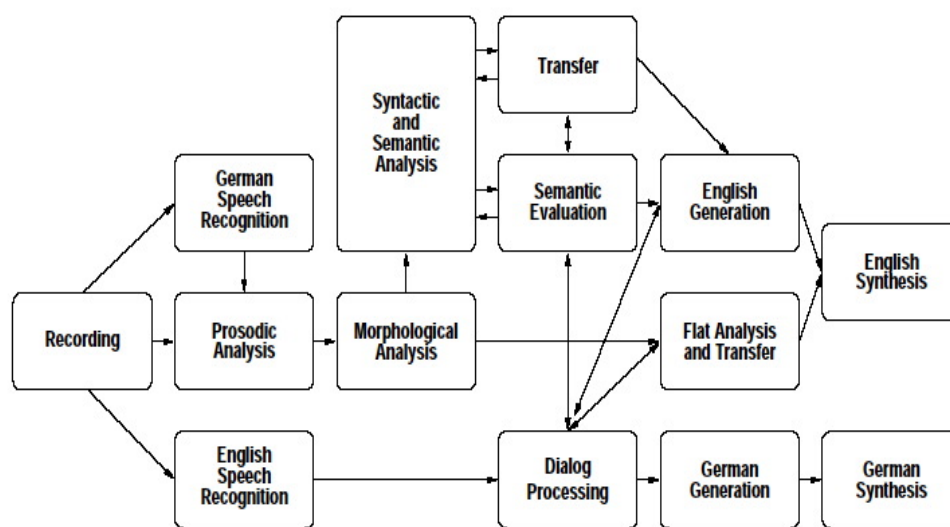


Figura 4.3: Sistema Verbmobil

O fluxo de informações começa pela aquisição do sinal da fala por meio do componente denominado Gravação. Nele são realizados ainda o controle dos dispositivos de gravação do sinal e o tratamento de erros de captação desse sinal.

As informações adquiridas anteriormente são repassadas ao componente chamado de Reconhecedor que as processa utilizando modelos estatísticos da linguagem captada (alemão, japonês e inglês) e gera um conjunto de probabilidades para as palavras que foram reconhecidas no sinal captado. O reconhecimento continua com a avaliação dessas probabilidades junto a dicionários com informações ortográficas e fonéticas. E para finalizar é feita a descoberta de palavras-chaves dentre as palavras restantes do

processo de reconhecimento. Logo o resultado final será as palavras que o sistema reconheceu do sinal de entrada.

Após o reconhecimento, o componente de Análise Semântica e Sintática realiza uma análise mais detalhada das palavras identificadas. No caso de ser necessário uma obtenção de mais informações lexicais o componente de Análise Semântica e Sintática fará uma requisição ao componente de Avaliação Semântica.

A tradução é responsabilidade do componente de Transferência que realiza o mapeamento da estrutura semântica da língua origem, adquirida no processo anterior, para a do inglês. Após, o resultado do mapeamento é repassado para o componente Gerador de Inglês que forma frases gramaticalmente corretas a partir desse resultado.

O processo de mapeamento pode gerar traduções alternativas, assim é necessário um componente que faça a escolha da mais adequada ao contexto do diálogo (componente de Seleção). Essa escolha é baseada no histórico de atos de diálogo mantido pelo componente de Processamento de Diálogo, que é responsável ainda pela iniciativa e clarificação do sistema.

Então a frase selecionada será repassada ao componente Sintetizador de Inglês, responsável pela emissão de sinais auditivos. Contudo, pode haver uma falha na tradução e faz-se necessário o uso de técnicas de clarificação ou correção, tarefa do componente de Processamento de Diálogo. Ele então força o sistema a se comunicar com o usuário por meio de sua língua nativa (alemão) e não mais o inglês. Para isso ele requisita o auxílio dos componentes Gerador de Alemão e Sintetizador de Alemão.

Um dos problemas encontrados em sistemas de tradução é a sua velocidade de tradução, que é demorado por processamentos internos do sistema, como o reconhecimento de fala e as análises de linguagem. Para amenizar isso este sistema utiliza um componente denominado Análise Superficial, que realiza a análise sintática com um conjunto limitado de regras usuais e com isso consegue maior velocidade do que um

analisador sintático comum que utilizada um conjunto com várias regras.

As limitações deste sistema recaem sobre as técnicas adotadas, isto é, o uso de técnicas probabilísticas permite um bom controle sobre o diálogo e resultados satisfatórios para a tarefa de tradução, entretanto ela necessita de muitos dados para se fazer confiável e problemas como dados esparsos ainda não foram superados por completo.

4.1.4 SUNDIAL – *Speech UNderstanding in DIAlogue*

O projeto SUNDIAL tinha como objetivo construir um sistema computacional capaz de manter um diálogo cooperativo com usuários por meio de linha telefônica [46, 47, 48]. Consiste de um conjunto de sistemas em quatro línguas voltados para os domínios de reservas de passagens para aviões (Francês e Inglês), itinerários de trens (Alemão) e banco de dados de um hotel (Italiano).

Essa interação em línguas diferentes é possibilitada pela arquitetura (figura 4.4, extraída de [47]) distribuída adota no projeto. Que lhe permite distribuir três funções distintas: interpretação das emissões do usuário, geração de emissões do sistema e gerenciamento do diálogo; em cinco módulos em cada sistema [48].

Os módulos responsáveis pela interpretação lingüística são: módulo de aquisição acústica que recebe os sinais acústicos e retorna palavras ou fonemas e o módulo de processamento lingüístico que atribui a um conjunto de palavras retornadas seu respectivo significado sintático e semântico.

No gerenciamento de diálogo o módulo de gerência de diálogo atribui a cada representação lingüística uma interpretação baseada no contexto do diálogo, dessa forma ele poderá decidir como continuar o diálogo. A geração de emissões fica a cargo do módulo de geração de mensagem, que produz uma representação lingüística adequada a mensagem que o sistema quer passar ao usuário e o módulo de sintetização de fala, que produz essas representações no formato acústico para ter como saída o telefone.

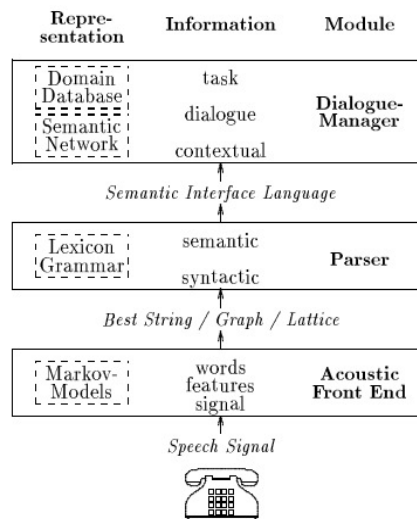


Figura 4.4: Arquitetura do SUNDIAL

O módulo de gerenciamento de diálogo atribuí aos sistemas do projeto SUNDIAL flexibilidade em relação à troca de tarefa e de língua adotadas pelo sistema e ainda sobre o problema de cooperatividade em diálogo (As emissões do sistema têm de percebidas pelo usuário como naturais, coerentes e úteis dentro do contexto do diálogo). Assim a estratégia utilizada foi distribuir o processo de controle de interação e cooperação em sub-módulos que repassam ao usuário a decisão de aceitar ou não certas direções que o diálogo poderia tomar.

As respostas dos sistemas ainda não se aproximam de um sistema de respostas em tempo real, ficando na casa das dezenas de segundos as melhores respostas [48].

4.1.5 GALAXY

É um de sistema de diálogo falado que permite a obtenção de informações por meio de conversação entre homem e computador via por telefone [49]. A sua primeira arquitetura considerava a distribuição de informações por meio de servidores de informações a diversos clientes. Com objetivo inicial de fornecer informações sobre o domínio de viagens (guia de cidade, reservas de vôos e previsão de tempo).

Esses servidores continham conhecimento necessário para lidar com perguntas específicas sobre o domínio. Isso os permitia interpretar perguntas dos clientes e fornecer respostas mais adequadas a situação, algo semelhante a um módulo de gerência de discurso de um sistema de diálogo convencional [49].

A arquitetura do GALAXY é composta por sistemas independentes que unidos fornecem ao usuário as informações requeridas de forma falada e/ou visual. Os componentes de base são chamados de servidores e são os seguintes: servidor de reconhecimento de fala (SUMMIT), servidor de entendimento de linguagem (TINA), servidor de geração de linguagem e sintetização de fala (GENESIS).

As fontes de informações sobre domínio específico são fornecidas pelos seguintes sistemas: agendamento de vôos (PEGASUS), previsão de tempo (JUPITER), guia de cidade (VOYAGER), guia de restaurante (DINEX), classificados de automóveis (WHEELS) e informações sobre páginas na Internet (WebGALAXY).

Para permitir o acesso dos usuários aos serviços disponibilizados pela arquitetura GALAXY e durante estudos sobre o componente de diálogo, em 1996 se iniciou uma reestruturação da arquitetura, figura 4.5 extraída de [50].

Na nova arquitetura as modificações mais significativas ocorreram no re-direcionamento das informações dos servidores e todos os sistemas de domínios para o novo componente denominado Concentrador (centro da figura), e a separação das funcionalidades da interface com o usuário para servidores separados (som, textos e gráficos), o que permite ao Concentrador diminuir suas responsabilidades [50].

Em [50] GALAXY-II é considerada uma arquitetura melhorada, contudo em [51] parte-se para o entendimento dessa arquitetura como um meio de estudo da interação com usuário, com bancos de dados e com sistemas de tarefas complexas específicas. Essa visão permite considerar uma mudança de paradigma, passando do cliente—servidor para o componente de comunicação intra-agentes. Outra modificação,

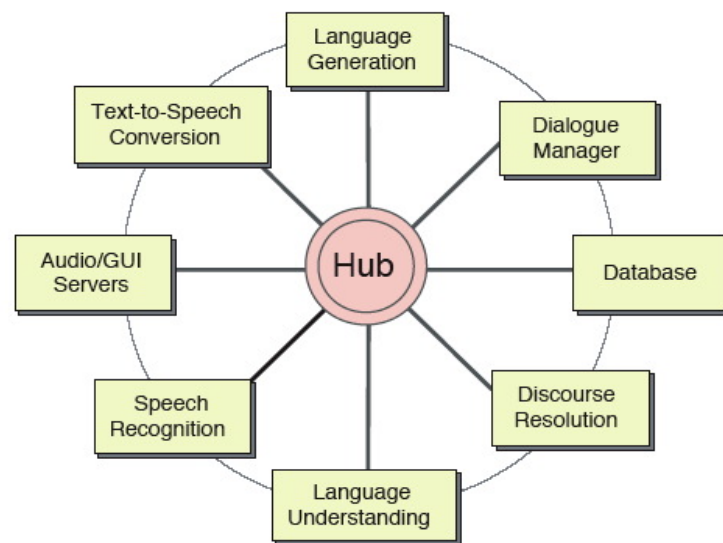


Figura 4.5: Arquitetura do GALAXY-II

considerada em [52], seria adaptar os sistemas de domínio para lidarem com outras línguas, o que poderia ser feito sem problemas uma vez que o Concentrador lida com informações sem o conhecimento fixo de uma língua.

Capítulo 5

Sistema proposto

Neste capítulo é apresentado o objetivo principal deste trabalho, que é a definição de uma estrutura genérica para um sistema de diálogo escrito (em Português) independente de domínio (aplicação), que atua como meio para a comunicação escrita entre uma pessoa e um programa de computador e que forneça algum tipo de serviço ou informação relevante para a pessoa. Para isso o sistema de diálogo necessita adotar uma abordagem que lhe permita representar tanto a estrutura de comunicação entre ele e o usuário, quanto o canal de informações entre ele e um programa de computador que forneça um serviço ou informações para o usuário.

No capítulo 3 algumas abordagens para a representação de comunicação foram comentadas sucintamente, neste capítulo será adotada uma delas para representar a comunicação entre usuário e sistema de diálogo. A abordagem escolhida é a baseada em agentes de interação racional [53, 6, 24, 54, 55, 17, 56, 57, 58, 59], que considera a comunicação como uma forma de comportamento inteligente, o qual é guiado por princípios racionais e cooperativos. Tais princípios são usados para delimitar a execução de ações, representadas pelas emissões dos participantes, dentro do processo de comunicação. Dessa forma as ações tomadas pelos participantes visam alcançar os objetivos não comunicativos, segundo Bunt [32], os quais levaram os participantes a entrarem em um diálogo.

Uma vez definida uma abordagem para sustentar teoricamente o processo comuni-

cativo do sistema de diálogo, resta definir sua estrutura. Essa definição será exposta no decorrer deste capítulo por meio da explicação da arquitetura do sistema de diálogo e de seus componentes: componente de entendimento de linguagem, componente de gerência de diálogo, componente de comunicação externa e componente de geração de linguagem.

A organização deste capítulo apresenta a estrutura e o processamento de linguagem do sistema de diálogo. Na arquitetura fundamental os componentes são agrupados em conjuntos que representam determinadas tarefas que juntas visam obter um funcionamento homogêneo do sistema como um todo. O primeiro componente do sistema é o de entendimento de linguagem.

Este componente é responsável pela análise lingüística das contribuições do usuário dentro de um diálogo textual com o sistema e posterior tradução da mesma em um formato mais adequado aos processos envolvidos no processamento das informações obtidas do usuário. Para isso, será utilizado o formalismo sintático denominado Gramática Léxico Funcional, o qual permite, além da representação sintática, uma codificação semi-semântica de uma sentença.

A interpretação semântica que é uma tarefa complexa que envolve conhecimentos referentes ao contexto do diálogo e do domínio em que este se apóia, utilizando esse formalismo sintático possibilita a escolha de uma linguagem de ligação entre a estrutura sintática e sua estrutura semântica conseqüente. Isso será realizado com o uso da lógica linear, a qual durante o processo de interpretação converterá a estrutura sintática em uma representação semântica lógica que será utilizada pelo sistema de diálogo para o reconhecimento de atos de fala referentes às contribuições do usuário, o que será feito pelo componente de gerência de diálogo.

Esse componente é o principal concentrador de informações do sistema, pois, ele recebe as informações do usuário e coordena os demais componentes para que o processamento dessas informações produza os resultados esperados pelo usuário.

Como mencionado anteriormente o sistema de diálogo adotará a abordagem de diálogo baseada em interação racional, a qual será mais bem explicada no decorrer do texto, além de seus componentes principais: os princípios racionais e cooperativos. Ainda, um ponto a ser destacado é a utilização de uma variação da lógica linear, a lógica linear temporal, como meio de contornar alguns detalhes referentes a utilização dessa abordagem para a representação do diálogo entre usuário e sistema.

As informações representadas em formato lógico são então processadas de acordo com um modelo de interação que faz parte de um sub-componente da gerência de diálogo, o controle de diálogo. Esse modelo organiza a forma como o sistema agirá no diálogo. Em conjunto com este sub-componente existem outros que servirão no auxílio de coordenação do processamento ou fluxo de informações dentro do sistema. Eles são o modelo de usuário, onde as informações sobre as contribuições do usuário serão armazenadas e servirá de base para os processos envolvidos na interação racional, proposta pelo sistema como abordagem de diálogo, e uma estratégia de recuperação e confirmação, responsável pela manutenção de um diálogo coerente e mais próximo do natural.

A função do sistema de diálogo é servir de intermediário para usuário e alguma aplicação que lhe forneça informações relevantes. Para isso o sistema pode ou não adotar a tarefa do usuário, ou seja, ele pode escolher se agirá como o usuário quando na comunicação direta com a aplicação ou se apenas auxiliará o usuário no processo de comunicação com a aplicação. A abordagem adotada pelo sistema a ser apresentado neste trabalho tentará auxiliar o usuário sem, no entanto ficar dependente da tarefa do usuário. Isto quer dizer que caso o sistema seja utilizado como intermediário para outro tipo de aplicação, ele poderá realizar sua função sem a necessidade de modificações internas em sua arquitetura.

Para se alcançar tal objetivo este sistema utilizará uma linguagem ontológica para realizar a comunicação com a aplicação externa. Com essa linguagem o sistema po-

derá auxiliar os usuários em diferentes aplicações sem a necessidade de modificações internas, necessárias para a adaptação às novas aplicações. Por um lado essa decisão de utilizar uma linguagem como protocolo de comunicação entre sistema e aplicação tem a vantagem de permitir a adaptação do sistema a outras aplicações, por outro lado ainda será necessário um processo de adaptação da aplicação à linguagem protocolar adotada pelo sistema.

A independência de aplicação é uma característica que permitirá ao sistema de diálogo realizar sua função sem a preocupação com detalhes referentes à aplicação. Porém, informações dependentes de aplicação ainda serão necessárias para o sistema executar sua função, pois, o contexto de uma aplicação pode não ser o mesmo que de outra, assim o vocabulário utilizado para uma não necessariamente será o mesmo utilizado em outras aplicações.

Essa questão se reflete no último componente do sistema de diálogo, o de Geração de Linguagem. Uma vez que uma aplicação possui um vocabulário específico, que pode ser considerado como um subconjunto do vocabulário de uma língua, as contribuições do sistema terão que pertencer a ele. Assim, para tentar se adequar ao vocabulário da aplicação e do usuário, o método de geração de linguagem será o que utiliza um misto de análise reserva e de realização baseada em gramática, capítulo 3.

5.1 Arquitetura

A especificação da estrutura por meio da arquitetura interna deste sistema de diálogo, figura 5.1, é caracterizada pelo agrupamento de algumas das fontes de conhecimento em módulos componentes da estrutura, mostradas no capítulo 3. As fontes utilizadas, e os detalhes sobre cada componente do sistema serão apresentados nas subseções seguintes.

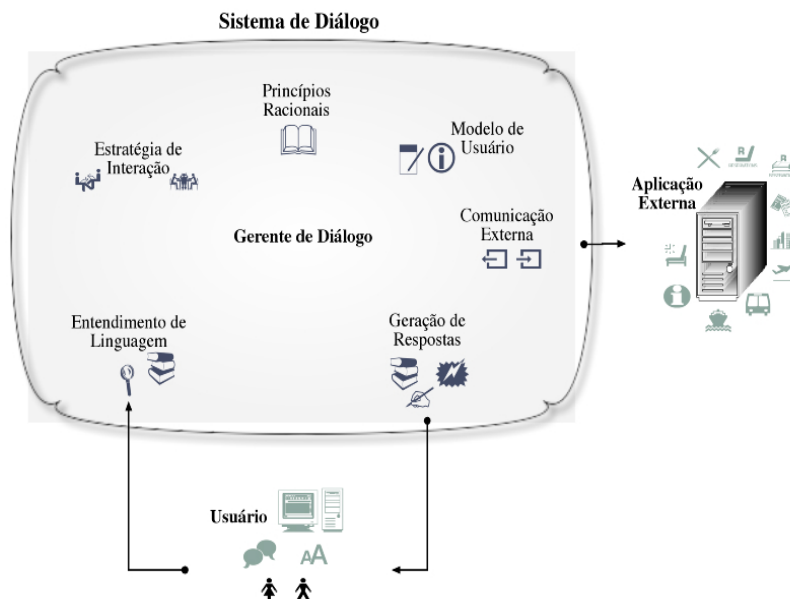


Figura 5.1: Arquitetura do sistema de diálogo

5.1.1 Entendimento de linguagem

Este componente do sistema é responsável pela aquisição do texto de entrada do usuário e posterior conversão do mesmo em um formato adequado ao tratamento computacional adotado pelo sistema de diálogo, o qual será apresentado no decorrer deste capítulo.

Antes de mostrar os processos internos desta parte do sistema de diálogo é bem vinda uma revisão dos mesmos. Para isto confira o capítulo 2.

Em linhas gerais, a aquisição é feita pelo recebimento do texto de entrada seguida da chamada do componente de análise sintática, o qual com a ajuda de um dicionário¹ e de um conjunto de regras gramaticais de uma língua, no caso deste trabalho o Português, montará árvores sintáticas para cada sentença do texto de entrada.

Após a construção dessa floresta sintática, um processo de interpretação semântica

¹Dicionário representa o significado de palavras por meio de palavras. Sua utilização se limita a princípio a consultas na análise sintática e na geração de respostas.

será iniciado. A interpretação semântica traduzirá cada uma das árvores em expressões lógicas, as quais serão repassadas aos próximos componentes do sistema.

Em resumo, os processos internos desta parte do sistema de diálogo são três: aquisição textual, análise sintática e codificação semântica em formato lógico. Cada um desses processos será explicado em maiores detalhes a seguir.

O processo de aquisição textual considera que a forma de interação do sistema com o usuário é feita por meio de sentenças textuais. Assim sua codificação dependerá apenas da linguagem adotada para a programação do sistema e também da língua. Isto é, a forma como o sistema receberá as entradas do usuário dependerá de como ele é codificado, o estilo dos elementos de interface (caixa de texto, campo textual, entre outros), e os detalhes que envolvem uma linguagem, por exemplos os textos produzidos em linguagens indo-européias como o Português e Francês necessita de codificação apropriada para lidar com os sinais de pontuação.

Um ponto importante da aquisição textual é o reconhecimento das palavras adquiridas, esta tarefa é facilitada pela não necessidade de reconhecimento de fala, porém ainda resta o reconhecimento da grafia correta das palavras, uma vez que a escrita incorreta de uma palavra pode modificar o sentido de uma contribuição do usuário, por exemplo, em um contexto de um hotel a pergunta “Quanto *quatos* vocês têm?” soa errada devido a falta de uma letra.

O segundo processo é o de análise sintática, onde as informações textuais adquiridas pelo processo anterior serão tratadas sintaticamente para se começar a interpretação do que foi dito pelo usuário. Para isso será utilizado um formalismo sintático, que é uma maneira de codificar o processo de análise sintática (extração das informações gramaticais de uma sentença). O formalismo que será adotado neste trabalho é denominado gramática léxico funcional ou LFG, da contração do nome em inglês.

LFG é uma teoria lingüística que atribui dois níveis ou estruturas de representação

sintática a uma sentença: uma estrutura constituinte (estrutura-c) que é representada por uma árvore sintática baseada em regras gramaticais livres de contexto, e uma estrutura funcional (estrutura-f) que é uma matriz hierárquica com elementos na forma atributo-valor, na qual se representam as relações gramaticais presentes dentro da sentença [60].

Assim para a representação sintática de uma sentença é necessário um processo que especificará quais informações lexicais (as categorias gramaticais a que as palavras podem pertencer) e quais restrições serão acrescentadas a representação sintática final de uma sentença. Para a definição da estrutura-c é necessário um conjunto de regras gramaticais livres de contexto, onde as anotações funcionais dessas regras fornecem uma descrição formal restritiva para construção de uma estrutura-f, a qual deverá conter apenas as informações sintáticas necessárias para representar uma sentença [60], ou seja, por meio de regras gramaticais que incorporam informações lexicais pode-se obter a estrutura-c de uma sentença. Por meio dessas informações lexicais pode-se determinar o arranjo mais adequado para a representação sintática da sentença, estrutura-f.

Para a análise sintática é necessário um conjunto de regras que estabeleçam os relacionamentos válidos entre os componentes de sentenças, e o qual indicará as sentenças que fazem parte de uma linguagem. Esse conjunto é denominado gramática e pode ser classificado de acordo com o trabalho de Noam Chomsky [13]. Onde um conjunto de regras gramaticais podiam ser assemelhadas a regras de produção, ou seja, regras do tipo ‘se tem letras então é um livro’. Isto trás a implicação de que por meio de um conjunto de regras e de elementos gramaticais, poderia-se inferir as combinações que produzissem sentenças válidas em uma linguagem. Isso quer dizer que por meio de um conjunto de regras que especificasse fenômenos sintáticos seria possível realizar uma análise sintática.

Uma gramática livre de contexto é aquela onde suas regras indicam a necessidade de um elemento que termine a análise sintática. Esse elemento é representado por uma entrada lexical, ou seja, a informação de uma palavra. Assim a análise sintática

ocorrerá pela substituição de todos os elementos do lado direito das regras aplicáveis até se encontrar algum elemento lexical. A final da aplicação de todas as regras e uso de todos os itens lexicais disponíveis então é montada uma árvore sintática, que representa o produto da análise sintática para uma sentença.

S	→	PN	PV
PN	→	Det	N
PV	→	V	PN

Figura 5.2: Um conjunto de regras sintáticas

Na figura 5.2 é apresentada uma gramática simples onde: S representa sentença, PN representa predicado nominal, PV significa predicado verbal, Det determinante ou artigo, V verbo e N nome ou substantivo. Adotando a frase: “Um livro cria mundos.” como exemplo e esse conjunto de regras, o processo de análise sintática ocorrerá pela substituição dos elementos lexicais segundo as regras gramaticais:

S = Um livro cria mundos

PN = Um livro

Det = Um

N = livro

PV = cria mundos

V = cria

PN = mundos

PN = mundos

Det = \emptyset

N = mundos

O resultado da análise sintática é apresentado na figura 5.3:

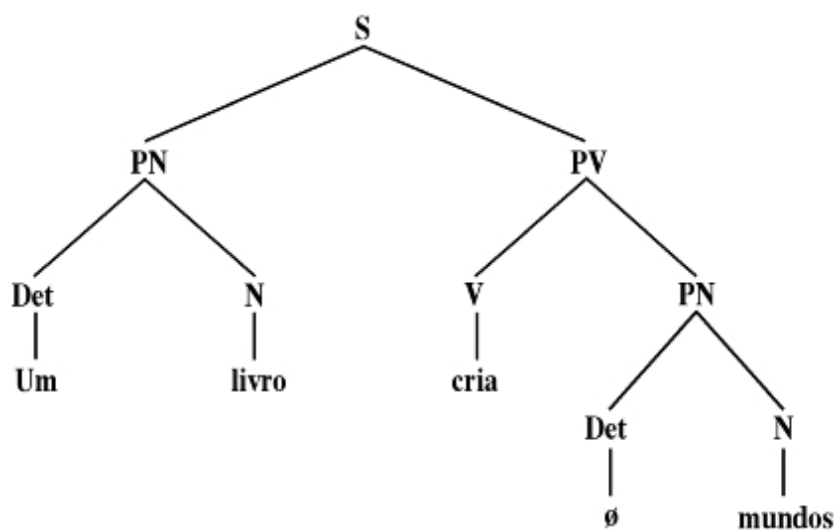


Figura 5.3: Árvore sintática

Apesar da simplicidade desse processo o resultado da análise sintática quando utilizando este tipo de gramática tende a produzir mais de uma representação ou árvore sintática para uma única sentença. Isso ocorre devido a natureza ambígua da linguagem e de sua classificação gramatical.

Como mencionado anteriormente em LFG acrescenta-se as regras sintáticas informações lexicais (estrutura-c), as quais podem ajudar a diminuir a geração ambígua de árvores sintáticas. Outra contribuição de LFG são princípios que tentam estabelecer uma certa unicidade a representação sintática final, estrutura-f. Assim os componentes necessários para a representação de uma sentença em LFG são: regras gramaticais, entradas lexicais, estrutura-c e uma estrutura-f.

As regras utilizadas em LFG apresentam expressões denominadas esquemas funcionais, os quais estão associados com os símbolos gramaticais destas regras (S, PN, PV, N, Det, etc.) [61]. Na figura 5.4 é mostrado um conjunto de regras em LFG.

Os esquemas funcionais são: $(\uparrow\text{Subs}) = \downarrow$, $(\uparrow\text{Obj}) = \downarrow$ e $\uparrow = \downarrow$. Onde em $(\uparrow\text{Subs}) = \downarrow$

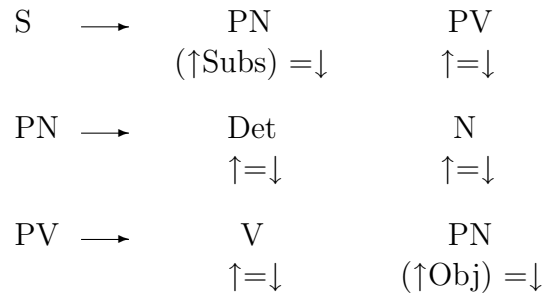


Figura 5.4: Um conjunto de regras LFG

a seta para cima indica que a expressão *Subs* será atribuída ao símbolo imediatamente acima do símbolo em que a seta foi posicionada. Assim S receberá a informação de que PN é um Substantivo. A seta para baixo diz que a atribuição será no próprio símbolo, então PN receberá também a informação de que ele é um Substantivo. Explicação semelhante a esta pode ser dada a $(\uparrow\text{Obj}) = \downarrow$. Já o esquema funcional $\uparrow = \downarrow$ indica que a característica do símbolo será compartilhada com o símbolo imediatamente acima [62].

As entradas lexicais possuem informações sobre as características dos itens lexicais, tais como sua estrutura de argumentos e suas funções gramaticais associadas aos argumentos [62]. Essas características são representadas por esquemas funcionais. As entradas lexicais para a sentença utilizada anteriormente podem ser vistas na figura 5.5.

A estrutura-c (figura 5.6) é semelhante às árvores sintáticas obtidas por outros formalismos gramaticais como resultado da análise sintática. A diferença está na adição dos esquemas funcionais junto aos símbolos e itens lexicais. Assim seu processo de construção é semelhante ao apresentado anteriormente.

A instanciação, segundo Wescoat [61], é o processo de transformação dos esquemas funcionais em equações funcionais, as quais são usadas na descrição formal de uma estrutura-f.

Em LFG cada nó da árvore sintática ou estrutura-c possui uma estrutura-f associada, e cada estrutura-f possuirá um identificador único chamado de meta-variável,

		Representação do item lexical
		Categoria sintática
		Lista de esquemas funcionais
Um	Det	(↑PRED) = 'um' (↑Num) = singular (↑Gen) = masculino
Livro	N	(↑PRED) = 'livro' (↑Num) = singular (↑Gen) = masculino
Cria	V	(↑PRED) = criar <(↑Subs)(↑Obj)> (↑Subs Num) = singular (↑Subs Pess) = 3
Mundos	N	(↑PRED) = 'mundos' (↑Num) = plural (↑Gen) = masculino

Figura 5.5: Entradas lexicais

figura 5.7, extraída de [61] (NP quer dizer PN e SUBJ é equivalente a SUBS). As setas tomam um valor referencial durante o processo instanciação.

Então, a seta para cima faz referência ao nó acima do qual ela aparece, logo ela será substituída pela variável que representa a estrutura-f do nó imediatamente acima. A seta para baixo será substituída pela estrutura-f referente ao nó em que ela se encontra.

Em resumo o processo de instanciação é o processo de substituição das setas pelas variáveis que nomeiam estruturas-f as quais elas se referem, isto é, de esquemas funcionais para equações funcionais. A figura 5.8 representa a estrutura-c depois da instanciação.

Após o processo de instanciação a próxima etapa na análise sintática será a construção da estrutura-f. Desse modo, o conjunto de todas as equações funcionais obtidas durante o processo de instanciação da estrutura-c serão transformadas em um descrição funcional da estrutura-c [61]. A partir desta descrição é que a estrutura-f será construída.

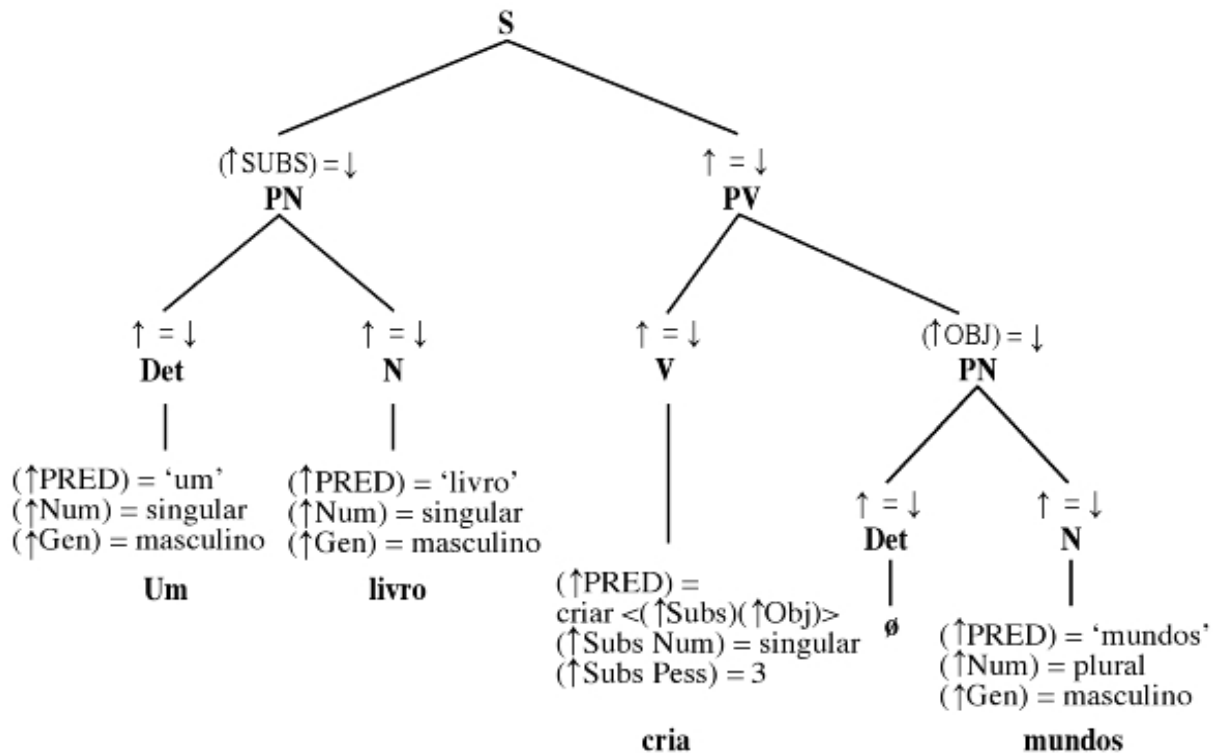


Figura 5.6: Estrutura-c

Na figura 5.9 apresenta-se o formato de uma equação funcional, onde f_p representa a estrutura-f associada à equação, *atrib* representa um atributo e *val* representa o valor deste atributo. E na figura 5.10 está listada a descrição funcional da estrutura-c.

A estrutura-f para a sentença exemplar é apresentada na figura 5.11. A leitura de uma estrutura-f é a seguinte, na coluna à esquerda estão representados os atributos e na coluna da direita estão os valores. Cada linha dessa matriz conterá uma equação funcional.

As entradas lexicais ou formas semânticas possuem: uma expressão que será interpretada na análise semântica, estruturas de argumentos de predicado, que são listas de argumentos de entidades semânticas [61].

Essa associação entre funções gramaticais e argumentos semânticos (valores dos

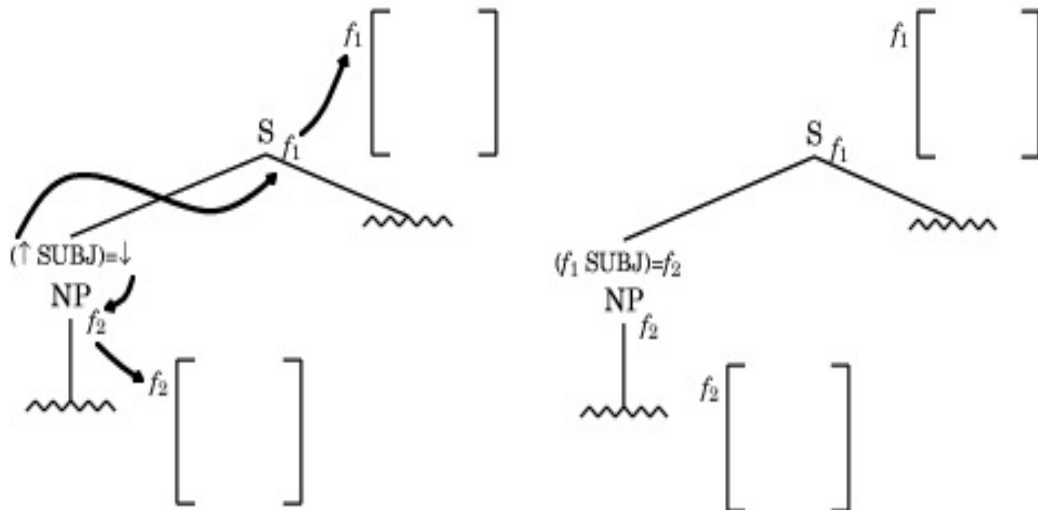


Figura 5.7: Instanciação da estrutura-c

atributos PRED) é estabelecida pela nomeação das funções gramaticais dentro de uma posição de argumento, assim a função será governada pelo argumento [61].

Uma estrutura-f tem de obedecer a algumas condições que garantam a sua boa formação e conseqüentemente a gramaticalidade da sentença a ser representada por ela. Elas são consistência, completude e coerência.

Consistência ou unidade funcional requer que cada atributo da matriz tenha apenas um único valor. Assim, por exemplo, um item lexical terá apenas um valor para o atributo de gênero [62].

Completude diz que as funções gramaticais presentes na estrutura de argumento de um predicado, ou seja, sub-categorizada pelo predicado, terão de ser representadas na estrutura-f [61]. Esta condição garante que certos elementos sejam obrigatórios. Isto é importante em línguas que exijam um sujeito presente em suas frases, como o Inglês e o Francês.

Coerência requer que toda entrada lexical tenha uma função gramatical na estrutura de argumentos de um predicado definida por algum argumento de predicado

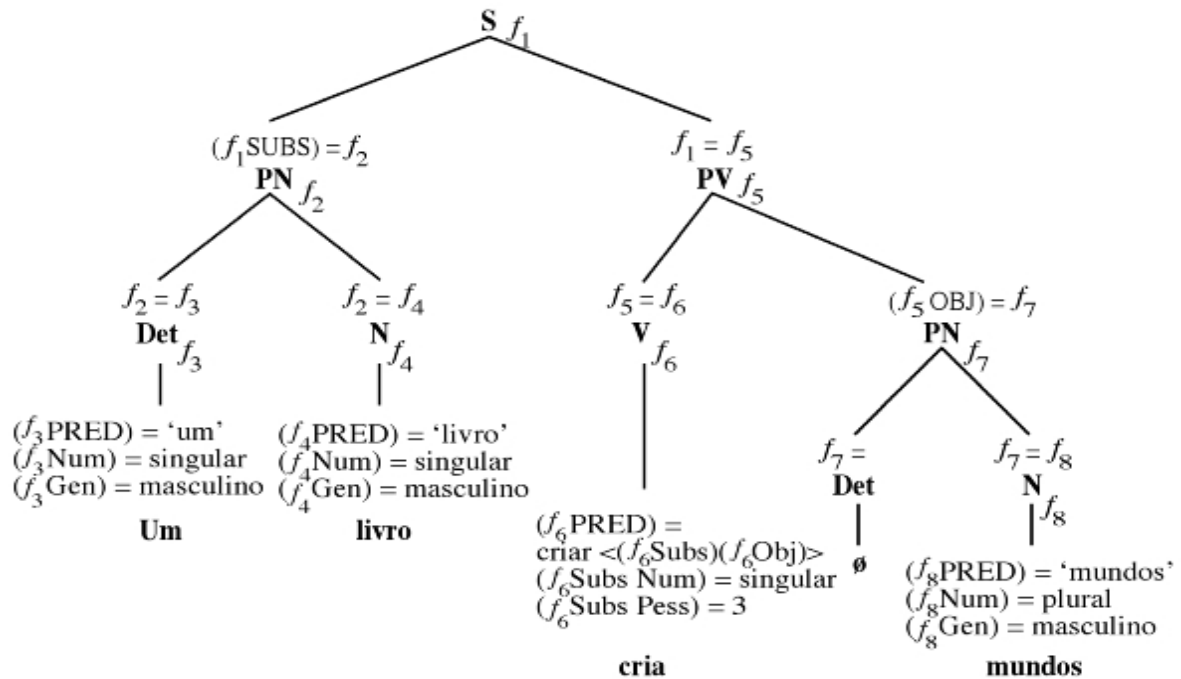


Figura 5.8: Estrutura-c após instanciação

presente na estrutura-f [62]. Um ponto negativo dessa condição é que caso um elemento não seja descrito por nenhum argumento de predicado ele se tornará inútil dentro da estrutura-f [61].

Após a definição da estrutura-f e de suas condições de boa formação, um ponto importante a ser apresentado é a questão de ordenamento interno dos itens lexicais de uma sentença, isto é importante em casos onde o ordenamento é importante, como para a resolução de anáforas².

A estrutura-f não apresenta ordenamento interno definido, isto permite representar livremente itens lexicais. Para atribuir algum ordenamento à estrutura-f faz-se necessário um ordenamento na estrutura-c, pois o mapeamento entre as estruturas induzirá um certo ordenamento à estrutura-f [60].

²Anáfora é a co-referência de uma expressão com seu antecedente, o qual fornecerá as informações necessárias para a interpretação desta expressão, ou seja, é uma expressão que se refere a um antecedente. Por exemplo, em “*Um menino travesso* quebrou uma janela; *Ele* foi castigado pela mãe.”, o pronome *Ele* se refere ao sujeito *Um menino travesso*.

$$(f_p \text{ ATRIB}) = \text{VAL} \qquad f_p \left[\text{ ATRIB} \quad \text{ VAL} \right]$$

Na estrutura-f f_p representa uma linha
 ATRIB é um atributo
 VAL é o valor deste atributo

$$((f_p \text{ ATRIB1}) \text{ ATRIB2}) = \text{VAL} \qquad f_p \left[\text{ ATRIB1} \left[\text{ ATRIB2} \quad \text{ VAL} \right] \right]$$

Na estrutura-f $(f_p \text{ ATRIB1})$ representa uma linha
 ATRIB2 é um atributo
 VAL é o valor deste atributo

$$f_p = f_q \qquad \begin{matrix} f_p \\ f_q \end{matrix} \left[\text{ ATRIB} \quad \text{ VAL} \right]$$

Figura 5.9: Significado de alguns tipos de equação funcional

O último processo é a tradução das informações sintáticas, obtidas no processo anterior, em um formato lógico que abrigue informações semânticas.

Este processo não é denominado de análise semântica devido a não intenção de se interpretar semanticamente as informações sintáticas, contudo querem-se apenas traduzir as mesmas em um formato que permitirá um tratamento com uma taxa mínima de ambigüidades nas próximas partes do sistema.

Uma vantagem da utilização de uma representação lógica, resultado da tradução semântica, é a possibilidade de uma posterior análise semântica mais simplificada, pois ela permite a utilização de premissas lógicas para a inferência de informações, que no caso deste trabalho será a inferência por deduções.

A desvantagem da escolha de uma lógica para o processamento interno do sistema é uma possível falta de expressibilidade, isso ocorre devido à lógica ser uma linguagem rígida quanto a sua estrutura e organização de seus componentes, o que não ocorre em uma linguagem natural como o Português, onde se pode expressar qualquer tipo de informação referente aos vários componentes da língua natural (verbos, substantivos,

$$\begin{aligned}
 (f_1 \text{SUBS}) &= f_2 \\
 f_2 &= f_3 \\
 (f_3 \text{PRED}) &= \text{'um'} \\
 (f_3 \text{Num}) &= \text{singular} \\
 (f_3 \text{Gen}) &= \text{masculino} \\
 f_2 &= f_4 \\
 (f_4 \text{PRED}) &= \text{'livro'} \\
 (f_4 \text{Num}) &= \text{singular} \\
 (f_4 \text{Gen}) &= \text{masculino} \\
 f_1 &= f_5 \\
 f_5 &= f_6 \\
 (f_6 \text{PRED}) &= \text{criar} \langle (f_6 \text{Subs})(f_6 \text{Obj}) \rangle \\
 (f_6 \text{Subs Num}) &= \text{singular} \\
 (f_6 \text{Subs Pess}) &= 3 \\
 (f_5 \text{OBJ}) &= f_7 \\
 f_7 &= f_8 \\
 (f_8 \text{PRED}) &= \text{'mundos'} \\
 (f_8 \text{Num}) &= \text{plural} \\
 (f_8 \text{Gen}) &= \text{masculino}
 \end{aligned}$$

Figura 5.10: Descrição funcional

adjetivos, etc.).

Por ter-se adotado LFG como forma de representar as informações sintáticas é possível escolher entre duas opções para a tradução semântica.

Uma delas seria a utilização da Teoria de Representação de Discurso ou TRD, descrita por Kamp e Uwe [25]. Onde a interpretação semântica é dividida em duas partes. Na primeira é feita a representação semântica das sentenças de um discurso por meio de estruturas de representação de discurso ou ERD e na segunda a interpretação dessas estruturas [25].

Nesta teoria a interpretação semântica leva em consideração não apenas o significado de uma sentença por vez, mas o conjunto de sentenças que formam um discurso. Dessa forma, a interpretação de cada sentença afetaria a interpretação das posteriores. Isto permite lidar com mudanças de contexto dentro de um discurso.

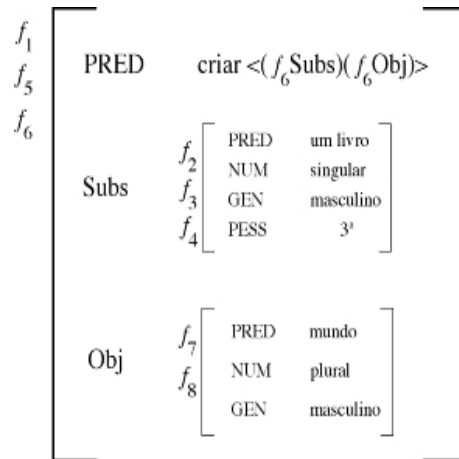


Figura 5.11: Estrutura-f

O contexto atribui um significado específico a cada sentença, ou seja, a interpretação de cada sentença dependerá de um contexto em particular. O conjunto de sentenças que compõem um discurso pode pertencer a um mesmo contexto, pois tendem a manter a coerência do discurso.

Logo, a interpretação semântica proposta pela TRD leva em consideração as mudanças que a interpretação de uma nova sentença causará no contexto estabelecido durante o processo de interpretação, com isto gerando uma atualização no contexto do discurso e como consequência uma interpretação mais completa do discurso.

Os elementos da TRD são: a ERD e as referências de discurso ou RD. RD's são partes constituintes das ERD's que servem para representar entidades e variáveis de contexto, as quais são úteis para o tratamento de referência pronominal [63].

As ERD's são formadas por dois elementos, um conjunto de referência de discurso, denominado de universo, e um conjunto de condições. O universo engloba todas as entidades que são referenciadas no discurso. As condições expressam restrições, relações ou propriedades dessas entidades, no discurso [63].

A construção das ERD's e o processo de interpretação fornecido pela TRD não

serão apresentados neste trabalho, pois eles estão fora do escopo do mesmo. Kamp em [63] realiza uma possível descrição do processo completo de interpretação.

A outra opção seria utilizar uma linguagem que permita representação (ligação) do significado semântico de uma sentença por meio da atribuição correta dos significados semânticos aos seus itens lexicais constituintes.

Os métodos convencionais utilizados para se fazer essa ligação consideravam uma certa ordenação dos constituintes, o que não é o caso da LFG, onde os constituintes representados na estrutura-f não possuem ordenação definida [64, 65].

A partir de Dalrymple [65] foi considerado que para ser possível a atribuição de significado aos itens lexicais das estruturas-f seria necessário à utilização de uma linguagem lógica de ligação, a linguagem escolhida foi a lógica linear, apresentada por Jean-Yves Girard em 1987, [66].

Essa ligação do significado de uma sentença com sua estrutura sintática ocorre, então, pela consideração dos itens lexicais de uma sentença à premissas lógicas que serão usadas para derivar logicamente o significado da mesma. Contudo, os itens lexicais terão de sofrer algumas alterações para que possam ser usados como premissas lógicas.

Tais alterações podem ser demonstradas ao considerar-se a continuação do exemplo anterior sobre a análise de uma sentença e posterior conversão para um formato de estrutura-f. Assim considerando a frase “Um livro cria mundos.” e seus itens lexicais apresentados na figura 5.5, além de sua estrutura-f, figura 5.11. Tem-se o seguinte conjunto de itens lexicais:

$$\begin{aligned} \text{um livro} &= (\text{subs}, \text{'um livro'}) \\ \text{mundos} &= (\text{obj}, \text{'mundos'}) \\ \text{criar} &= !(X, Y.(\text{subs}, X) \otimes (\text{obj}, Y) \multimap (f, \text{criar}(X, Y))) \end{aligned}$$

Onde $(\text{subs}, \text{'um livro'})$ representa um conjunto para descrição do formato lógico que uma entrada lexical pode vir a ter, onde ‘um livro’ representa uma constante e

‘subs’ um termo associativo correspondente ao tipo sintático do item lexical. Ainda, X e Y representam variáveis que podem ser instanciadas por constantes encontradas em premissas derivadas de itens lexicais.

$$\begin{aligned} X &= \text{‘um livro’} \\ Y &= \text{‘mundos’} \\ &(\text{subs, ‘um livro’}) \otimes (\text{obj, ‘mundos’}) \multimap (f, \text{criar(‘um livro’, ‘mundos’)}) \end{aligned}$$

Então para a frase “Um livro cria mundos.” após a análise sintática e posterior conversão lógica tem-se como resultado a expressão lógica: $\text{criar(‘um livro’, ‘mundos’)}$. A qual corresponde ao significado pontual da frase apresentada.

A lógica linear é uma extensão da lógica clássica e intuicionista [67]. Seu caráter extensivo é dividido em parte pela inclusão de novos conectivos lógicos [66] e em outra a teoremas que comprovam sua força e expressividade.

Com alguns desses conectivos é possível lidar com atualização de situações, onde situações são casos onde uma expressão lógica que represente um evento por meio de uma implicação causal, não apresenta uma reação ou esta pode ser desconsiderada [67].

O tratamento de atualização de situações se encaixa dentro do processo de representação semântica dinâmica, onde a representação não é baseada apenas em uma sentença, mas sim no conjunto de sentenças que definem um evento qualquer.

Os conectivos introduzidos pela lógica linear são os seguintes:

- Multiplicativos (\otimes, \wp, \multimap) correspondentes bilineares de: ‘e’, ‘ou’ e ‘implicação’.
- Aditivos ($\oplus, \&$) versões lineares de: ‘e’ e ‘ou’ da lógica clássica.
- Exponenciais ($!, ?$) que tem semelhança com os conectivos modais: \Box e \Diamond .

Uma descrição breve de cada um é a seguinte:

\otimes – conjunção simultânea – onde $A \otimes B$ quer dizer: A e B podem ser executados ou consumidos ao mesmo tempo.

$\&$ – conjunção alternativa ou escolha interna – onde $A \& B$ quer dizer: tem-se a opção de escolher entre A e B, porém sabendo que está disponível exatamente um de cada.

\oplus – disjunção ou escolha externa – onde $A \oplus B$ é semelhante a $A \vee B$ na lógica clássica e quer dizer: tem-se A e B podem ser consumidos exatamente uma vez, contudo pode-se escolher qual deles será consumido.

\wp – disjunção – onde $A \wp B$ representa a operação dual em relação a $A \otimes B$, assim na primeira operação A e B terão de ser consumidos, independente da ordem de seu consumo. Isto passa a idéia de compartilhamento de recursos.

\multimap – implicação linear – onde $A \multimap B$ quer dizer: utilizando o recurso A será obtido o recurso B, com o consumo de A durante o processo.

! – claro – onde $A!$ quer dizer: existe um suprimento ilimitado de recursos do tipo A.

? – porque não – onde $A?$ quer dizer: existe a possibilidade do final dos recursos do tipo A.

Um exemplo simples da expressividade da lógica linear e ainda da diferença em relação a lógica clássica. Supondo que A represente uma nota de dinheiro e B um copo de suco, ainda, com uma nota de dinheiro possa-se comprar um um suco ($A \rightarrow C$).

Em lógica clássica a seguinte dedução estaria correta: Tendo uma nota e sabendo que com ela pode-se comprar um suco então compra-se um suco ($A, A \rightarrow C$, então C, porém A ainda é verdadeiro), ou seja, apesar de se ter gastado uma nota, ainda esta a disposição outra nota.

Em lógica linear o mesmo exemplo ficaria assim: Tendo uma nota e sabendo que com ela pode-se comprar um suco então compra-se um suco ($A, A \multimap C$, então C), e A é consumida durante o processo de implicação, ou seja, não resta mais notas.

Como pôde ser visto a lógica linear pode expressar melhor certos fenômenos reais, os quais a lógica clássica não pode.

Sendo assim, o uso de lógica linear, como linguagem de ligação, permite considerar que as sentenças de um diálogo ou discurso e seus itens lexicais contribuem exatamente em um único momento para a definição do significado semântico de uma sentença, pois a contribuição semântica é dada pelas diversas ocorrências de informações que são geradas e usadas apenas uma vez neste processo [64].

Todas as demais ocorrências que não forem utilizadas na representação de uma sentença serão utilizadas em outros passos do processo de interpretação semântica. Isto garante que todas as informações sobre o evento serão utilizadas para fornecer o significado semântico para um conjunto de sentenças.

Para a representação semântica da estrutura- f obtida a partir de uma sentença é necessário uma função σ que realize o mapeamento de uma estrutura- f (f) para uma estrutura semântica (f_σ) correspondente [64], figura 5.12.



Figura 5.12: Mapeamento de estrutura- f para estrutura semântica

O significado pode ser associado a estrutura semântica de duas formas: diretamente, figura 5.13a, ou no caso de verbos, tem-se que especificar como o significado será associado com a estrutura semântica, figura 5.13b [64].

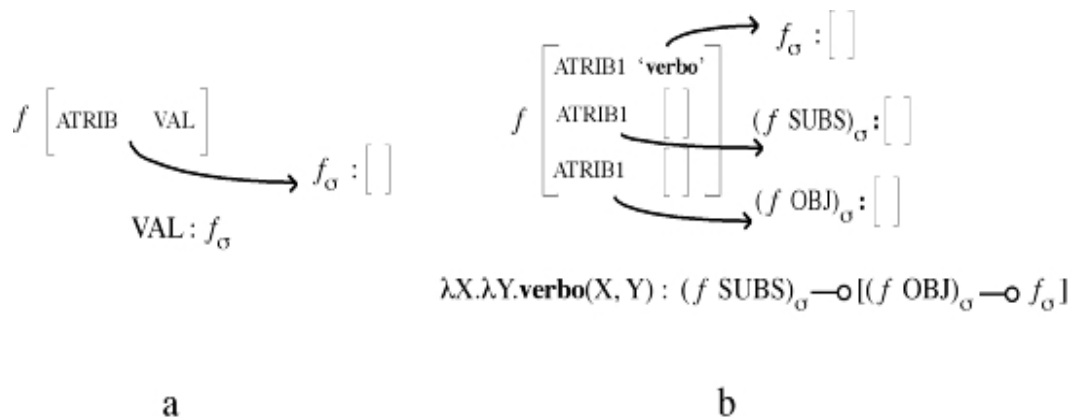


Figura 5.13: Atribuição de significados à estrutura semântica

A leitura da primeira fórmula diz que o valor da estrutura- f será mapeado diretamente para o significado da estrutura semântica. Na segunda fórmula, para verbos, o verbo conterá dois elementos (X e Y) com existência única, ou seja, existirá apenas um de cada. Por exemplo, a leitura da primeira parte desta fórmula para o verbo apontar seria X aponta para Y ($\text{apontar}(X, Y)$). A segunda parte da fórmula para verbos diz que se existir um significado para o substantivo associado ao verbo então avaliasse em seguida. Se o significado do objeto associado ao verbo for encontrado então será, por fim, produzida a representação do verbo na estrutura semântica.

Em se considerando a frase: “Maria contratou Silvia” a conversão de sua estrutura- f para a estrutura que represente a sua significação semântica, ocorrerá primeiramente pela definição da estrutura- f . Assim, seguindo o que foi exposto anteriormente a estrutura- f para essa frase é apresentada na figura 5.14. A estrutura está mais simplificada apenas para efeito de exemplo. As letras s e o representam as estruturas contidas na estrutura- f final f .

O passo seguinte será determinar quais são os elementos construtores de significado ($[\text{maria}]$, $[\text{silvia}]$ e $[\text{contratou}]$) da estrutura semântica para que seja possível montar o conjunto de premissas lógicas que servirão ao mapeamento entre as estruturas [64].

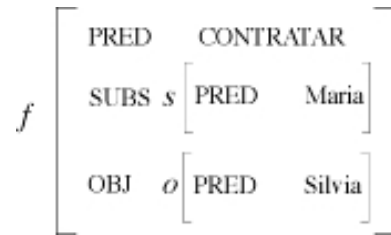


Figura 5.14: Estrutura-f para a frase “Maria contratou Silvia”

Construtores:

[Maria] Maria : s_{σ}

[Silvia] Silvia : o_{σ}

[contratou] $\lambda X.\lambda Y.\text{contratar}(X, Y) : (f \text{ SUBS})_{\sigma} \multimap [(f \text{ OBJ})_{\sigma} \multimap f_{\sigma}]$

O que pode ser reescrito, considerando que $(f \text{ SUBS}) = s$ e $(f \text{ OBJ}) = o$, como:

Construtores:

[Maria] Maria : s_{σ}

[Silvia] Silvia : o_{σ}

[contratou] $\lambda X.\lambda Y.\text{contratar}(X, Y) : s_{\sigma} \multimap [o_{\sigma} \multimap f_{\sigma}]$

Após a definição desses construtores, parte-se para o processo de mapeamento entre a estrutura-f e estrutura semântica. Os construtores obtidos depois da construção da estrutura-f serão usados como premissas em um processo de dedução lógica. Logo as premissas podem ser agrupadas da seguinte forma:

[maria] \otimes [silvia] \otimes [contratou]

O processo dedutivo avança com a utilização da operação de *modus ponens*³ sobre as premissas disponíveis, figura 5.15. Assim

Portanto por meio de um processo dedutivo pôde-se extrair a partir de de uma frase sua representação semântica em formato lógico.

³Essa operação considera que se em tendo-se uma premissa verdadeira que diga: se ter sede então beber água; e outra premissa que seja verdade e diga: tem-se sede, então por esta operação pode-se deduzir ou afirmar: beber água.

$$\begin{array}{c}
\frac{\lambda X.\lambda Y.\mathbf{contratar}(X, Y) : s_{\sigma} \multimap [o_{\sigma} \multimap f_{\sigma}] \quad \text{Maria} : s_{\sigma}}{\lambda Y.\mathbf{contratar}(\text{Maria}, Y) : o_{\sigma} \multimap f_{\sigma} \quad \text{Silvia} : o_{\sigma}} \\
\hline
\mathbf{contratar}(\text{Maria}, \text{Silvia}) : f_{\sigma}
\end{array}$$

Figura 5.15: Processo de dedução

Além das abordagens apresentadas anteriormente, existem outras formas de se prover meios para a interpretação semântica a partir de LFG, essas formas podem ser reunidas em três grupos de abordagens para a interpretação, onde dois desses grupos foram apresentados anteriormente.

No primeiro, estão as abordagens baseadas em mapeamento da estrutura-f para uma forma quasi-lógica. No segundo, linguagens de representação de significado dinâmico e, no último, abordagens para gerenciamento de contexto por lógica linear [64].

O uso da lógica linear como meio de interpretação semântica visa utilizar suas características de interpretação de contexto e atualização de contexto sem a necessidade uma linguagem de representação de significado. Sendo para isso necessário apenas algumas modificações, sugeridas por Genabith em [68] para que ela possa lidar com aspectos particulares da atualização de contexto.

Então a lógica linear permite o tratamento de interpretação e atualização de contexto apenas com a utilização de derivações da linguagem de ligação entre estrutura-f e estrutura semântica de uma sentença [68], o que pode simplificar o processo de interpretação semântica e ainda permitir o tratamento de alguns fenômenos da linguagem natural, como elipses e anáforas.

5.1.2 Gerência de diálogo

Este componente tem a função de coordenar o fluxo de informações entre os demais componentes do sistema de diálogo. Para isso, ele possui algumas subdivisões: controle de diálogo, modelo de interação racional, estratégia de recuperação, modelo de usuário e comunicação com aplicação externa.

O controle de diálogo foi incorporado ao componente de gerência de diálogo, porque sua função é indispensável para a manutenção do fluxo de informações entre os componentes do sistema e por estabelecer uma organização sobre as informações que estão sendo requisitadas e sobre aquelas que estão disponíveis para o processo de comunicação com o componente externo.

No modelo de interação, será apresentada a contribuição deste trabalho e um dos objetivos do mesmo. Levando em consideração as teorias sobre interação racional será apresentada uma alternativa que visa solucionar em um primeiro momento o problema da falta de consideração sobre os limites de recursos disponíveis para o tratamento lógico de formalismos antigos.

A estratégia de recuperação está fora do controle de diálogo apenas para ressaltar seu papel dentro do processo de manutenção de um diálogo coerente com o usuário.

O modelo de usuário está ligado ao modelo de interação, no fornecimento de informações sobre o estado mental do usuário.

A última interligação entre este componente e uma função do sistema de diálogo é a comunicação com a aplicação externa. Ela se dará por meio de uma linguagem padronizada que será especificada por uma ontologia. Essa ontologia servirá como um repositório de conhecimentos mútuos entre o sistema de diálogo e a aplicação externa.

5.1.2.1 Controle de diálogo

A estratégia para o controle de diálogo considera algumas idealizações quanto ao processo de tratamento lingüístico, comunicação com o usuário e aplicação externa. Isso é necessário para evitar possíveis problemas envolvendo tais processos.

Sendo assim, as informações advindas do componente de entendimento de linguagem estarão em um formato lógico que expresse corretamente a semântica da emissão do usuário. O processo de reconhecimento das ações do usuário sobre a representação lógica de sua emissão será considerado como não possuindo erros, uma vez que o não reconhecimento adequado das ações pode levar a falha no processo de racionalização.

Com essas idealizações iniciais parte-se para a definição da estrutura de diálogo a ser mantida durante a interação racional entre usuário e sistema. Essa estrutura considerará que a comunicação (diálogo) está dividida em seqüências de segmentos, denominados ações comunicativas [33, 34]. Tais ações podem ser agrupadas em:

- Abertura de tópico ou assunto
- Manutenção de tópico
- Mudança de tópico
- Fechamento de tópico ou assunto

Essas ações podem ser compreendidas como ações comunicativas que expressam a vontade ou intenção de um participante de alcançar um objetivo. Assim tais ações serão tratadas como atos de fala [2].

Os atos de fala são as ações que um falante executa em uma emissão⁴ [57]. Isto se considerar a comunicação como um processo de execução de ações. Existem três tipos de ações que podem ser executadas na comunicação, os quais correspondem a três tipos de atos [14]:

⁴No caso deste trabalho uma emissão é uma entrada textual, podendo ser uma palavra, uma frase ou mesmo um conjunto de frases.

- Ato locutionário: é o ato de emitir uma seqüência de palavras.
- Ato ilocutionário: é o ato que o falante executa quando diz as palavras, tendo um sentido proposicional e intencional.
- Ato perlocutionário: é o ato que produz um efeito particular no ouvinte, por meio da emissão das palavras.

Esses atos foram agrupados por Austin segundo suas observações em estudos sobre verbos de uma língua, a língua inglesa, pois a teoria sobre os atos de fala foi desenvolvida para esta língua. Searle realizou melhoramentos nesta teoria de Austin sobre atos ilocutionários ao expressar a necessidade de condições necessárias e suficientes para a execução destes atos [2].

A partir dessas condições Searle apresenta as seguintes classes de atos de fala [14, 57]

:

- Classe representativa – onde o falante está compromissado como verdade do que é emitido – atos: informar, negar, afirmar e confirmar.
- Classe comissionada – onde o falante se compromete com uma ação futura – atos: prometer e assegurar.
- Classe de diretrizes – onde o falante tenta influenciar as intenções e comportamentos do ouvinte para que ele execute alguma ação, atos: comandar, convidar, perguntar, pedir e requisitar.
- Classe expressiva – onde o falante expressa seu estado psicológico – atos: desculpar, agradecer, parabenizar e receber.
- Classe declarativa – onde o falante, por meio do ato de fala, executa uma convenção social – atos: nomear, demitir-se, batizar, etc.

As classes expressivas e declarativas não serão consideradas no decorrer deste trabalho, pois como apontado por Pauchet em [69], os atos da classe expressiva podem ser considerados como casos particulares de atos da classe representativa, uma vez que

a descrição do estado psicológico do falante é comparada a descrição de qualquer outro ato que o falante execute. Os atos da classe declarativa não são considerados válidos, pois sua execução apenas transmite a sua essência e não causa os efeitos esperados, por exemplo, ao declarar uma situação uma pessoa não estará mudando o estado da situação por meio dessa declaração.

Bunt [32, 70, 71, 72, 73, 74] e Traum [75, 76, 77, 78, 57] estendem a noção de atos de fala como simples ações executáveis por emissões para a idéia de que determinados tipos de atos podem ser agrupados de tal forma que possam representar algumas características da comunicação. Para Traum esses agrupamentos são chamados de atos conversacionais e para Bunt são denominados de atos de diálogo, contudo ambos possuem semelhança de definição, são tipos de atos de fala que se prestam ao controle de diálogo, diferindo apenas no uso, para Traum existem quatro níveis de ação que permitem a manutenção e coerência de um diálogo. Já Bunt considera dois tipos de atos de diálogo, os que controlam o diálogo e outros que expressam o seu conteúdo.

Neste trabalho os atos de fala serão utilizados apenas para expressar o conteúdo semântico contido nas emissões do usuário e do sistema. Os demais tipos de atos de fala serão aproveitados em continuação do mesmo.

Decisão semelhante será tomada em relação ao gerenciamento de troca de vezes. Apesar da comunicação ser um processo voltado para muitos participantes, onde cada um pode contribuir sempre que necessário, este sistema considerará a iniciativa fixa como forma de controle de vezes. Entretanto, a abordagem adotada pode permitir a adoção de iniciativa mista, uma vez que ela considera além de princípios de ações racionais, também princípios de comportamento cooperativo.

Após a definição de alguns fatores referentes ao controle de diálogo, resta definir o fluxo de informações a serem gerenciadas por este componente. As informações são processadas da seguinte forma:

1. As informações advindas do componente de entendimento de linguagem são com-

paradas ao estado mental do sistema, o qual é composto por uma rede de crenças sobre fatos obtidos junto à aplicação externa. Essa rede é definida a partir da linguagem ontológica utilizada para a comunicação com a aplicação externa.

2. Após a verificação das informações junto ao modelo mental do sistema, as informações do usuário são convertidas para o formato de atos de fala, algo que é possível pelo formalismo lógico adotado neste trabalho. O resultado da conversão e a informação original são utilizados para atualizar o modelo de usuário, que mantém o suposto estado mental do usuário.
3. Interpretado o que o usuário emitiu, o sistema tentará fornecer informações que sejam adequadas à emissão do usuário, por exemplo, se o usuário emitiu uma pergunta então o sistema tentará respondê-la de acordo com informações extraídas do seu estado mental e do modelo de usuário atualizado. Dependendo do domínio da aplicação externa e da emissão do usuário o processamento do sistema será regido pelos princípios racionais e cooperativos de seu modelo de interação.
4. Se o sistema não possuir informações suficientes em seu estado mental e modelo de usuário para realizar seu processo racional, então ele requisitará mais informações junto ao componente de comunicação externa. Feito isto, o sistema atualizará seu estado mental.
5. A geração de emissões do sistema será realizada pelo componente de geração de linguagem no momento em que algum ponto do processo de racionalização necessitar, tanto durante sua execução quando no seu final.

O processo de racionalização e a definição do estado mental do sistema serão apresentados a seguir.

5.1.2.2 Modelo de interação racional

Este modelo será utilizado para fornecer meios de o sistema de diálogo interagir com o usuário. Isso será possível pela adoção de uma teoria para agentes racionais que considera o diálogo como uma interação racional, uma vez que um agente, nos termos de IA, é uma entidade autônoma que produz ações que modificam o mundo ao seu redor. Uma vez que a teoria dos atos de fala considera a fala como um meio para a execução de uma ação então a adoção de uma teoria para modelagem de agentes pode ser adequada à modelagem de diálogo.

A base para a teoria adota por este modelo é aquela que considera a noção de que pela definição de atitudes intencionais pode-se especificar o comportamento de um agente. A forma de se realizar esta especificação será por meio de formalismos que utilizam algum tipo de lógica para essa especificação. Wooldridge em [58] apresenta algumas tentativas de se alcançar este objetivo.

As tentativas que serão consideradas neste trabalho são: a teoria de Cohen e Levesque [24] que considera as atitudes mentais: crença, objetivo e intenções que são derivadas das duas primeiras e a teoria de Sadek [54, 55, 17, 79] que utiliza princípios racionais e cooperativos para representar as ações de atos comunicativos. Esta será adotada como ponto de partida para o modelo de interação racional deste componente. Ainda, outra tentativa é a lógica BDI desenvolvida por Rao em [56], que utiliza as três atitudes mentais: crenças, desejos e intenções e foi baseada na teoria de sistema BDI desenvolvidos por Bratman [21], que não será discutida aqui por seguir outra linha de raciocínio.

A teoria de Cohen e Levesque teve início com o trabalho de Cohen e Perrault [53], onde uma metodologia para a adoção de atos de fala em uma teoria baseada em plano era fundamentada pela definição de atos de fala como operadores lógicos em um esquema de planejamento. Esse esquema considerava que os resultados das ações expressos por atos de fala afetam o estado mental de ambos os participantes de uma interação racional. As atitudes mentais consideradas por esta metodologia eram duas:

crenças e querer que modelavam os desejos dos participantes da interação.

A seqüência desse trabalho Cohen [6] promoveu uma formalização mais detalhada do papel dos atos de fala em relação a uma teoria formal para a representação de ação racional. O ponto fundamental era a demonstração de que os atos ilocutionários não necessitam ser primitivos, ou seja, eles não são dependentes da satisfação de certas condições [2] para terem seus efeitos representados na teoria de comunicação por meio de ações. Assim alguns atos podem ser derivados a partir de princípios básicos de ação racional.

Com a base teórica desenvolvida anteriormente, em [24] Cohen e Levesque apresentaram um formalismo que utilizava as noções de interação racional e da teoria de atos de fala desenvolvida até então. Sua teoria avançou para a idéia de que as intenções são o elemento que mantém o balanço racional entre crenças, objetivos e ações do estado mental de um agente. Assim, do ponto de vista lógico eles desenvolveram um formalismo para a interação racional de agentes com motivações intencionais, para isso usaram uma lógica modal dinâmica que utiliza como operadores crença, objetivos e dois outros para representar ações temporais. Do ponto de vista lingüístico o formalismo lógico desenvolvido permitia a representação dos atos comunicativos não como recursos para agentes, mas sim modelos descritivos de ações [57].

O trabalho de Cohen e Levesque teve papel fundamental para o desenvolvimento de outras teorias que tentam contornar problemas da representação de interação racional [59, 58]. Pois, conseguiu expressar claramente o papel semântico de sua teoria lógica no cenário da interação racional para agentes [57]. Todavia, algumas desvantagens ainda persistem, uma delas é o problema da onisciência lógica (um problema onde um agente é considerado um racionalizador completo que possui todas crenças e todos os conhecimentos sobre um domínio, algo que não é real, pois o um agente real possui recursos limitados) e outra está relacionada à manutenção das crenças dos agentes, isto é, um agente durante o processo de racionalização poderá abandonar suas crenças antigas caso ele decida que seus objetivos ligados a estas crenças não são mais válidos

ou que sejam inalcançáveis.

A teoria de Sadek [54, 17] sobre comunicação por meio de interação racional se apóia em duas noções: na primeira o processo de diálogo é expresso por meio de um formalismo lógico de teoria de interação racional, baseado somente em princípios de comportamento racional que descrevem atos comunicativos como ações que podem ser planejadas, definindo então um processo totalmente dedutivo de planejamento comunicativo. Na segunda essa teoria lógica pode ser usada representar diferentes aspectos do comportamento racional, principalmente o cooperativo.

A teoria lógica formalizava, primeiramente no trabalho de Sadek [54], as atitudes mentais (crenças e intenções) e ações ou eventos em uma linguagem modal lógica baseada em um modelo lógico denominado KD45 [58]. Baseado no formalismo desenvolvido por Cohen e Levesque, Sadek considerava as ações como eventos e o processo de planejamento racional como a ocorrência de seqüências de ações codificadas em expressões de ações que passariam por dois tipos de escolhas, uma determinística e outra não determinística.

O componente principal do modelo de ação comunicativa são os atos comunicativos [54, 17], os quais representam dentro do processo de planejamento de ações, as razões pelas quais um emissor seleciona determinadas ações e a relevância das condições que têm de serem satisfeitas para essas ações serem planejadas. Para uma ação seus efeitos são denominados efeitos racionais e as condições para sua execução são denominados de pré-condições de viabilidade.

Esses efeitos são derivados da tentativa de descrever os atos comunicativos como derivações dos atos de fala. Sendo assim, o efeito racional corresponderia ao efeito causado pela procura de um ato pelo próprio emissor, as pré-condições de viabilidade corresponderiam ao efeito indireto que um ato tem sobre o receptor que tenta manter a persistência de suas pré-condições e por fim o efeito intencional que corresponderia a intenção que um receptor tem de manifestar em comprometimento da execução do

ato [79].

Em [54] Sadek apresenta um exemplo desse formalismo. Já em [17] tal formalismo é expandido para lidar com outras atitudes mentais, além de fornecer elementos para a codificação de princípios racionais e cooperativos.

Apesar das contribuições das teorias apresentadas, ainda persistem alguns tipos de problemas: onisciência lógica [58], atualização de crenças [79], definição de quais atitudes mentais se ajustam melhor a uma teoria lógica que represente atos de fala e a dificuldade de representação de certos tipos de execuções temporais de ações pelo uso de alguma variação de lógica dinâmica [57]. Dois deles tentarão ser contornados pelo modelo de interação a ser definido a seguir. O primeiro problema é a não consideração da limitação de recursos para o processo de interação racional e o segundo é a questão da não representatividade adequada de ações no decorrer do tempo. Ambos tentarão ser solucionados com a adoção da lógica linear temporal [80] como mecanismo de representação lógica do modelo de interação racional.

Com relação ao problema da onisciência lógica o modelo de interação racional a ser utilizado neste componente tentará fornecer uma solução simples. Tal solução se baseia no princípio de que o problema surgiu pela adoção de uma lógica que não considera os recursos disponíveis para o processo de dedução ou inferência, assim ao se adotar uma lógica que leve em consideração esse fator então o problema estaria resolvido.

A forma de prover essa solução será pela adaptação da teoria de Sadek quanto a interação racional cooperativa. A escolha por essa teoria tem duas motivações: a primeira diz respeito a evolução do tratamento da teoria de interação racional, ao acrescentar a noção de cooperação, e a segunda diz respeito ao formalismo que permite deduções por meio de seus próprios elementos formadores. Esse processo de racionalização será semelhante ao adotado por Sadek em [17], todavia, algumas modificações terão de ser realizadas.

Tais modificações recaem na diferença de tratamento entre as lógicas adotadas por Sadek e por este modelo. No formalismo de Sadek se utiliza uma lógica dinâmica que é baseada em lógica modal de primeira ordem (KD45), a qual considera as crenças como fatores principais do processo de dedução. Já neste modelo será usada a lógica linear temporal, que considera premissas como recursos a serem consumidos durante o processo de dedução. Sadrzadeh em [81] mostra uma forma de conversão da lógica modal para a lógica linear, o que viabilizará a conversão dos princípios racionais e cooperativos para a nova lógica.

Para a realização dessa conversão, além de adaptações nos princípios, serão utilizados apenas quatro atos de fala para representar o modelo de ação de Sadek: perguntaSN, pergunta, respostaSN e resposta. Neste trabalho será considerado que esses atos podem ser responsáveis pela representação de um diálogo simples. Sabe-se que com a adição de outros tipos atos a representação ficaria mais próxima da ideal, contudo para efeito de comprovação de adaptação apenas esses atos serão utilizados.

Um ato de fala desde [53] pode ser entendido como uma ação que pode causar certos efeitos e que possui algumas condições para sua execução. Essa noção considera que uma pessoa tem intenção alcançar ou causar algum efeito sobre outra pessoa quando ela executa uma ação, representada por um ato de fala.

Para Cohen e Levesque [53] o ato pergunta é definido por dois atos complementares, um pergunta e outro convencer, os dois atos tem a seguinte expressão: um falante que deseja perguntar a um ouvinte sobre uma informação p e obter tal informação.

pergunta(falante, ouvinte, p)

condição: falante ACREDITA p

falante ACREDITA falante DESEJA inf

efeito: ouvinte ACREDITA falante ACREDITA p

convencer(agente1, agente2, p)

condição: agente1 ACREDITA agente2 ACREDITA p

efeito: agente1 ACREDITA p

O ato pergunta foi desenvolvido nos trabalhos seguintes de Cohen e mesmo por Sadek, que em [54] define o ato perguntaSN como a ação de um agente tem que realizar para alcançar os efeitos proporcionados pelo ato.

<agente1, perguntaSN(agente2, proposição)>
 condições: ACREDITA(agente1, SABE(agente2, proposição)) e
 NÃO SABE(agente1, proposição) e
 ACREDITA(agente2, NÃO SABE(agente1, proposição))
 efeito: SABE(agente1, proposição)

A definição dos atos a serem usados neste modelo não apresentará diferença da definida por Sadek, apenas será adequada a notação apropriada à lógica linear temporal.

Assim, as informações codificadas logicamente pelo componente de entendimento de linguagem terão de ser identificadas para um dos quatro atos de fala a serem considerados por este modelo. O reconhecimento dos atos pode ser feito pela comparação dos elementos da representação com os elementos componentes do modelo mental de crenças do sistema de diálogo, o qual será baseado nos conhecimentos extraídos da fonte externa.

O modelo mental do sistema será formado por crenças que serão consideradas como recursos que o sistema possui para serem consumidos durante o processo de inferência do modelo de interação racional. Além das crenças do sistema existiram crenças que estarão ligadas ao modelo de usuário. Tais crenças podem ser consideradas como de conhecimento mútuo, e podem ser consideradas como crenças que são consumidas durante um diálogo.

A duração de uma crença dependerá de quão consumíveis elas forem, isto é, crenças estabelecidas apenas durante o processo de inferência podem ser consideradas como crenças voláteis. Logo crenças de usuário e sistema serão consideradas como de duração ou uso ilimitado, enquanto que as crenças mútuas serão consumidas durante o

processo de inferência. A questão de crenças mútuas ainda permanece sem solução, tanto em relação a sua definição quanto a validade de seu uso [57, 6]. Contudo neste trabalho elas serão consideradas como recursos que serão consumidos apenas uma vez durante o processo de inferência.

Dessa forma, as crenças a serem utilizadas no processo de racionalização serão as crenças que o sistema mantém e que são obtidas junto a fonte externa, que representa o domínio, as crenças do sistema sobre o usuário (modelo de usuário) e um conjunto pequeno de crenças de mútuo conhecimento entre sistema e usuário.

Em resumo, este modelo pode ser entendido como um conjunto composto por um provedor de teoremas e um sistema especialista para um dado domínio, o qual pode ser substituído por outros sistemas referentes a outros domínios. Assim, os princípios racionais e cooperativos podem ser comparados ao núcleo do provedor de teoremas, enquanto que a estrutura representacional dos relacionamentos de conceitos contidos na linguagem ontológica são comparados ao sistema especialista. Logo este modelo fornecerá os meios devidos para uma interação racional do sistema com o usuário levando em consideração um domínio em comum.

5.1.2.3 Estratégia de recuperação e clarificação

A estratégia a ser adotada pelo sistema de diálogo para lidar com erros e emissões mal interpretadas será dividida em duas etapas, uma para a recuperação do foco do diálogo mantido pelo usuário e sistema e outra que informará o usuário as possíveis dúvidas do sistema sobre determinado ponto do diálogo.

Na primeira etapa (recuperação) é realizado o tratamento de entradas mal formadas e também de erros no processo de representação das entradas em formas lógicas. A abordagem escolhida será informar ao usuário o erro e requisitar uma nova entrada, mesmo que o erro não tenha sido causado por ele. Esta não é a melhor forma de resolver tais problemas, porém a definição de uma abordagem para recuperação não é

um dos objetivos deste trabalho, além de acreditar-se que com a utilização do modelo de interação possam-se diminuir as chances de tais erros acontecerem.

A segunda etapa (clarificação) consiste de verificação ou confirmação das informações adquiridas e interpretadas. Estes processos são necessários para garantir que o que foi interpretado pelo sistema seja realmente o que o usuário teve intenção de passar com sua entrada, uma vez que uma emissão pode falhar em sua interpretação. Para isso existem duas maneiras de se obter confirmações junto ao usuário [16]: explícita ou implícita. Ambas foram discutidas no capítulo 3.

A abordagem utilizada neste trabalho será a de confirmação explícita, ou seja, para cada não entendimento do sistema ele fará uma consulta ao usuário. Esta abordagem é mais robusta, contudo ela pode tirar a naturalidade do diálogo entre sistema e usuário, por exemplo, em situações em que o sistema não entenda uma seqüência de entradas do usuário ele insistirá na correção de cada uma, ao invés ele poderia tentar contornar a situação.

5.1.2.4 Modelo de usuário

Neste trabalho o modelo de usuário funcionará como um repositório para as atitudes mentais do usuário, e ainda como um histórico do diálogo entre um usuário e o sistema. Assim, a observação das atitudes mentais do usuário ao longo do tempo pode descrever as interações do usuário com o sistema no decorrer de um diálogo, o histórico de atitudes mentais do usuário.

Esse histórico pode ser utilizado para a construção de um padrão de comportamento do usuário ou ainda para a definição de um conjunto de crenças e intenções do usuário, repositório de atitudes mentais. Este repositório pode ser usado na determinação de quais ações tem maiores chances de serem executadas pelo usuário na interação racional com o sistema.

5.1.3 Comunicação com sistema externo

Este componente visa adquirir as informações armazenadas ou processadas por uma fonte de dados externa ao sistema de diálogo. Essa aquisição será feita por meio de uma linguagem cuja gramática é definida por uma ontologia (linguagem ontológica⁵), termo que será explicado mais adiante. Assim, tanto a fonte externa quanto o sistema possuem um comum canal para comunicação.

Da parte da fonte externa é necessária ou não uma adaptação em seu mecanismo de comunicação externa, pois para se comunicar com o sistema de diálogo ela precisará conhecer a linguagem adotada pelo sistema. No sistema de diálogo as informações providas pela fonte externa serão convertidas em um formato lógico, o mesmo utilizado no modelo de interação racional. Isto permitirá ao componente de gerência de diálogo ter acesso a informações sobre o contexto de um domínio.

Contudo, apesar das adequações na fonte externa, ainda restaria uma outra modificação, esta no próprio componente de comunicação, mais especificamente no mecanismo que interpreta a linguagem ontológica. Isto pode ser atribuído a adequação dos conceitos representados na ontologia ao formalismo lógico usado no componente de gerência de diálogo. Todavia, esta nova adaptação pode não afetar o processo de comunicação de modo geral com a fonte externa, uma vez que apenas alguns pontos da ontologia teriam de ser modificados para acomodar as restrições lógicas necessárias ao tratamento por outras partes do sistema. Então um complemento ou passo extra, no processo de definição de uma ontologia, seria necessário para a adequação final da ontologia ao sistema de diálogo.

O primeiro passo para a comunicação entre sistema e fonte externa será a definição

⁵O termo linguagem ontológica sugere a noção da existência ou definição de uma linguagem artificial com características que possam ser utilizadas de forma única para a determinação da comunicação, isto é, ela permitirá uma comunicação mais direta do seguinte ponto de vista hipotético: em analogia a uma linguagem humana, as palavras são substituídas por conceitos ontológicos e dessa maneira os mesmos usos que elas admitem e suportam podem ser aplicados a eles.

do que é uma ontologia. O termo ontologia vem da filosofia e segundo Bueno [82] é o estudo ou ciência do ser enquanto ser, em si mesmo e independente do modo como este se manifesta. Contudo, na área de estudos em Inteligência Computacional este termo recebe outras definições providas por vários autores, e que possuem enunciados semelhantes entre si.

Gruber em [83] disse que “uma ontologia é uma especificação explícita de uma conceitualização”. E que conceitualização é uma visão abstrata do mundo feita sobre um ponto de vista específico.

Em [84] Uschold define uma ontologia como um termo utilizado para expressar um entendimento comum de um domínio em particular, representando uma visão de mundo de um grupo ou sociedade. Tal visão é determinada por grupos de conceitos, suas definições e seus inter-relacionamentos. Uma ontologia possui, ainda, um vocabulário de significados e definições dos conceitos em sua estrutura.

Para Noy [85] uma ontologia é uma descrição formal de um domínio de discurso (classes ou conceitos), propriedades de cada conceito descrevendo suas características e atributos (conectores⁶), restrições e/ou operações nos conectores (facetadas).

Após tais definições uma ontologia será entendida como uma rede de classificações comuns a um grupo de pessoas em uma área de conhecimento em particular e que pode ser utilizada para além de formalizar e unificar as várias classificações, compartilhar conhecimento comum sobre um ponto de vista mais geral sobre uma área de conhecimento.

Estabelecido o que é uma ontologia, parte-se para seu uso. Neste trabalho uma ontologia será utilizada para permitir o compartilhamento e armazenamento de conhecimentos específicos de um domínio ou contexto, os quais estão disponíveis em uma fonte externa ao sistema de diálogo.

⁶Este termo é uma tradução de *slot*, o qual é mais utilizado na literatura.

Porém, a construção de uma ou mais ontologias para representar os conhecimentos de uma fonte externa seria custoso, pois, dependendo do tipo de fonte (base de dados, base de conhecimento e sistema especialista) as informações seriam diversas e invalidariam a idéia de consulta, uma vez que se estaria transportando o conhecimento para dentro do sistema de diálogo.

Para contornar este problema a opção será adotar uma linguagem cuja gramática seja baseada em uma ontologia comum tanto ao sistema de diálogo e quanto a fonte externa. Assim pelo uso dessa linguagem poderia se estabelecer uma comunicação entre o sistema de diálogo e a fonte externa. Para se alcançar este objetivo um primeiro passo seria a definição da ontologia e o segundo passo seria a definição da linguagem de comunicação.

Para a construção de uma ontologia é necessária a definição do seu escopo e domínio. Uma forma de se alcançar isto seria se basear na formulação de perguntas gerais sobre o domínio, as quais serão respondidas com o conhecimento da fonte externa. A abrangência do tipo de conhecimento contido nessas perguntas estabelecerá o escopo da ontologia. Algo como a definição de conhecimento concentrado em apenas determinado assunto de uma área de conhecimento.

Essas perguntas são chamadas de perguntas de competência por Noy [85], e servirão para: limitar o conhecimento contido na ontologia e o nível de detalhamento das informações específicas do domínio.

Após a definição das perguntas, parte-se para a determinação de quais termos são mais importantes ou mais utilizados no domínio. Esses termos compõem o vocabulário característico do domínio (jargão do domínio). Onde os termos definem um ou mais conceitos referentes a uma área de conhecimento, e os conceitos são representações formais de conhecimentos específicos do domínio, os quais são caracterizados por um conjunto de propriedades, que estão presentes na sua definição formal e na definição

dos seus relacionamentos.

Porém termos apesar de relevantes para a definição do vocabulário da ontologia, podem apresentar a característica ambígua de representar mais de um conceito, isto tem de ser contornado, pois as definições de conceitos têm de ser aceitas e entendidas de forma única por várias pessoas do mesmo domínio.

Com a definição dos termos em conceitos únicos, parte-se para sua classificação. Essa classificação é produzida pelo agrupamento de conceitos que apresentem características comuns entre si. Realizada a classificação dos conceitos únicos, é então definida uma estrutura que represente formalmente esta classificação.

Essa estrutura pode ser caracterizada como um arranjo hierárquico dos conceitos. Noção semelhante a da teoria de representação orientada a objetos. Onde objetos com características comuns são agrupados em arranjos hierárquicos de classes de objetos e os seus relacionamentos são apresentados por métodos ou ações próprias de suas classes.

Essa estrutura hierárquica possibilitará o relacionamento direcionado entre conceitos de uma classe, ou seja, conceitos de uma mesma classe podem ser agrupados de forma a representar níveis de informações. Ainda, ela permitirá que novas informações sejam obtidas por meio de agrupamentos de conceitos que levem em consideração diversas classes e não somente o ramo hierárquico a que eles pertencem.

A classificação de conceitos podem ser definida por meio de termos. Termos estes que representam informações mais genéricas sobre determinado ramo de conhecimento do domínio. Logo, termos que representem conceitos gerais nomearão classes, termos que representem conceitos específicos serão agrupados em relação às classes de conceitos que melhor determinem a especificação dos conceitos. Os termos escolhidos para nomear classes fazem referência a conceitos únicos, pois não pode haver ambigüidade na definição das classes.

Definida a estrutura formal de representação e a classificação dos conceitos, parte-se para a definição das propriedades e relacionamentos dos conceitos. Mais uma vez, apropriando-se das noções de orientação a objetos, propriedades representam características marcantes do conjunto de objetos de uma classe. Neste caso as propriedades caracterizarão os conceitos que fazem parte de uma classe em particular.

Possíveis termos que não tenham sido aproveitados como conceitos únicos, serão considerados como propriedades das classes de conceitos a que mais se assemelhem. As propriedades podem ser definidas a partir de necessidades de representação dos conceitos e não apenas por termos restantes do processo de classificação.

As relações entre conceitos e termos podem ser tratadas mais formalmente por meio aparatos lógicos, normalmente utilizam-se axiomas⁷ lógicos, os quais apresentam condições que serão responsáveis pela recuperação de informação na ontologia.

Com um conjunto de axiomas podem-se expressar os relacionamentos e as condições destes relacionamentos entre as classes da ontologia. Por fim, a ontologia estará definida para a fonte externa. Resta então definir a linguagem para a comunicação entre sistema de diálogo e fontes externas.

Em seguida, parte-se para a definição da linguagem de comunicação entre sistema de diálogo e fonte externa. Essa linguagem será composta por grupos de conceitos presentes na ontologia do domínio, assim a própria ontologia será a linguagem de comunicação, pois, uma vez estabelecido uma estrutura para a ontologia, esta estrutura será utilizada como meio de comunicação.

As informações contidas na ontologia serão convertidas para o mesmo formato lógico usado pelo modelo de interação, dessa forma o sistema poderá realizar sua função de ser o intermediário entre usuário e uma fonte externa sem ficar dependente da co-

⁷Axioma lógico é uma proposição evidente, ou seja, uma apresentação de um fato verdadeiro e diretamente provável logicamente.

dificação da utilizada pela fonte externa, pois, ele usará uma ontologia que garantirá a comunicação entre ele e outros tipos de fontes externas de conhecimento.

Com o uso desse mecanismo de comunicação o sistema não precisará armazenar todo o conhecimento da fonte externa, e ainda permanecerá independente de domínio, pois o que determina o conhecimento de um domínio é o conhecimento da existência de uma ontologia que sirva de linguagem de comunicação com uma fonte externa ao sistema que forneça tais informações.

Resta ainda um detalhe, como se realiza a comunicação de forma prática com um dos tipos de fonte externa? Para responder a esta pergunta é necessário considerar a realização de alterações na fonte externa ou a criação de um módulo, externo ao sistema de diálogo, que funcione como meio prático entre sistema e fonte externa.

5.1.4 Geração de linguagem

Esta parte é responsável pelas saídas do sistema. A estratégia utilizada para a produção de respostas e de perguntas com fins de esclarecimento será o processo inverso ao utilizado no componente de entendimento de linguagem, ou seja, a partir de uma representação lógica este componente montaria uma sentença que expresse de certo modo o significado que se pretende passar ao usuário.

O formato lógico usado será o mesmo usado pelo componente de entendimento de linguagem, logo sua função se resumirá apenas a tradução das representações lógicas em sentenças. Entretanto essa tradução utilizará o mesmo dicionário, mas não a mesma gramática utilizada no outro componente. Isso para evitar alguns problemas apontados em 3.4.4.1, e ainda assim permitir ao sistema produzir de forma simples as sentenças que expressem suas ações em um diálogo.

5.2 Conclusão

Este capítulo apresentou a definição de um sistema de diálogo escrito, que incorpora algumas novidades em relação aos trabalhos correlatos apresentados no capítulo 4, modelo de interação racional baseado em lógica linear temporal e utilização de uma ontologia como meio de comunicação entre sistema de diálogo e uma fonte externa.

Um fator importante a ser ressaltado é a tentativa de separação de funções entre os componentes do sistema, pois o objetivo deste trabalho era fornecer uma arquitetura simples para um sistema de diálogo textual. Isso permitirá futuras substituições das abordagens adotadas e também utilização de abordagens mais complexas em alguns deles.

Capítulo 6

Considerações

6.1 Considerações Finais

Este trabalho apresentou as bases para o início de uma pesquisa que tentará mostrar a viabilidade de algumas soluções para o desenvolvimento prático de um sistema de diálogo escrito. Isso foi possível pelo levantamento de conceitos envolvidos na construção de um sistema de diálogo genérico em pesquisas e estudo anteriores, o que permitiu a definição de um modelo de sistema de diálogo escrito simples para uma língua específica que adote a noção de independência de aplicação.

Os pontos-chaves desse modelo são: primeiro, a utilização de uma linguagem lógica responsável pela codificação das informações externas ao sistema e ainda pela formalização de um conjunto de regras de inferência, baseadas em um formalismo lógico para a comunicação racional e cooperativa. Segundo, a especificação de uma linguagem ontológica para a comunicação entre sistema de diálogo e aplicação externa que fornece algum tipo de serviço (previsão de tempo, consulta a banco, reservas, etc.).

Isto em princípio pode levar a uma comunicação livre do sistema de diálogo, isto é, sua arquitetura não fica atrelada a um conjunto de aplicações relacionadas. Bastando a leitura de uma nova definição ontológica para que o sistema possa se comunicar com uma nova aplicação.

As idéias apresentadas no trabalho ainda não foram verificadas em um protótipo totalmente funcional que representasse o modelo por completo, contudo alguns pontos puderam ser observados em relação ao desenvolvimento dos componentes chaves da arquitetura do sistema proposto. Considerando o componente de entendimento de linguagem, a descrição formal lógica apresentada pode fornecer incompatibilidade com os dados necessários para que o componente de gerência de diálogo possa desenvolver seu papel. Esse ponto pode ser considerado como um trabalho futuro e que aparentemente não será motivo de grandes problemas.

A utilização do formalismo lógico definido por Sadek em dois de seus artigos, [54, 17], para o componente de interação racional e cooperativa mostrou-se trabalhosa e ainda incompleta. A utilização de algumas noções encontradas em [86] amenizou as dificuldades do processo de codificação dos princípios em um formato lógico adotado pelo sistema.

Em relação ao componente de comunicação com a aplicação externa, o desenvolvimento de uma ontologia seguindo os passos apresentados no capítulo 5, mostrou-se simples, contudo algumas adaptações foram necessárias para a adequação ao formato lógico adotado pelo sistema. O processo de adaptação e mesmo o de construção de uma ontologia não foram apresentados, pois a definição de uma ontologia estava fora do escopo deste trabalho. Entretanto, a falta de um exemplo é considerada como um ponto fraco deste trabalho, o que não invalida a apresentação das etapas para a construção de uma ontologia apresentada no capítulo 5.

A seguir serão mostradas algumas observações sobre as dificuldades envolvidas no desenvolvimento de um sistema de diálogo. Isso com o intuito de caracterizar a delimitação do objetivo principal deste trabalho, que era iniciar uma pesquisa e não o desenvolvimento completo de um sistema de diálogo funcional.

Observações sobre os problemas em abertos na área de sistema de diálogo e processamento de linguagem natural que podem servir de incentivo para o desenvolvimento

desta pesquisa e melhoramento dos pontos fracos de um sistema de diálogo:

- O processo comunicativo não é apenas caracterizado por princípios de cooperação e racionalidade, existem outros fatores que podem e influenciam seu funcionamento. Tais fatores ainda não foram detectados por completo, apesar de estudos em diversas áreas do conhecimento (psicologia, linguística, neurologia, etc.).
- Certos fatores comportamentais como obrigações, normas sociais, confiança, cooperação complexa teriam de ser levados em consideração, além dos fatores (princípios) de interação racional e inteligente para que um sistema possa conversar, ou seja, o projeto ou a capacidade de dialogar de um sistema de diálogo teria de levar em consideração muitos dos fenômenos conversacionais (fatores) encontrados na fala e escrita.
- O conhecimento é um fator que também influencia o processo comunicativo. Mais relevante do que ter grandes quantidade e diversidade de conhecimentos é ter a capacidade de aprender e utilizar conhecimentos novos para representar o encadeamento de pensamentos que forma a base de um processo comunicativo natural.
- A capacidade humana de amenizar erros de reconhecimento de fala ainda é um objetivo a ser alcançado. Não que meios para se contornar erros não existam, mas uma metodologia que represente a capacidade completa ainda está por vir.
- Ainda persiste o debate sobre qual o melhor formalismo para representar o processo comunicativo, e as soluções apresentadas até então não são capazes de representá-lo por completo, contudo certas características podem ser simuladas parcialmente. Isto porque, ainda não foi decifrado o processo cognitivo de linguagem e tais formalismos, em sua grande maioria, são baseados em estudos

observacionais.

Contudo nem tudo são problemas, existem soluções parciais para os problemas apresentados acima e no decorrer deste trabalho algumas delas foram apresentadas e novas sugeridas. Tais soluções ainda não podem ser comparadas à capacidade humana de dialogar, todavia, elas apresentam um grande potencial para simular determinados comportamentos humanos que podem ser importantes para comunicação entre humanos e computadores.

6.2 Trabalhos Futuros

Comprovação ou validação das idéias apontadas neste trabalho por meio da adição dos componentes não destacados no último capítulo. Uma vez que seu desenvolvimento permitiria experiências com usuário reais, o que seria benéfico para o desenvolvimento do sistema como um todo.

Utilização da lógica linear temporal para representar melhor as características dinâmicas da comunicação. Pois, apenas a lógica linear não poderia, a princípio, representar a natureza temporal da comunicação humana, e mesmo de diálogo entre humanos e computadores.

Verificação completa da proposição de que ao se utilizar uma lógica linear pode-se eliminar o problema de onisciência lógica, apresentado pelo modelo representacional de conhecimento (lógica KD45), no qual Sadek apóia seu formalismo lógico de princípios racionais e cooperativos.

Definição de estratégias para o entendimento ou reconhecimento de linguagem, o que poderia ser feito pela consideração do estado corrente do diálogo, ou seja, com a utilização mais detalhada das informações armazenadas no modelo de usuário e histórico de diálogo. A adoção de teorias que lidem com fala ao invés de escrita pode ser

outro tipo de trabalho futuro.

Estudo mais aprofundado de conceitos relacionados à lógica, para que seja definido um formalismo lógico próprio, o qual poderá lidar com diálogos cooperativos de forma mais objetiva e consistente.

Continuação desta pesquisa, por meio do desenvolvimento de um protótipo funcional que possua todos os componentes propostos no capítulo 5 e posterior estudo prático com usuários. Além do melhoramento ou inclusão de teorias para cada parte do sistema de diálogo.

Referências Bibliográficas

- [1] TURING, A. M. *Computing Machinery and Intelligence*. 1950.
- [2] SEARLE, J. R. *Speech Acts An Essay in the Philosophy of Language*. [S.l.]: Cambridge University Press, 1978.
- [3] WILKS, Y.; CATIZONE, R. Human-computer conversation. In: *Encyclopedia of Microcomputers*. New York: Marcel Dekker, Inc., 2000. Disponível em: <<http://citeseer.ist.psu.edu/316623.html>>.
- [4] BARR, A. Natural language understanding. In: AMERICAN ASSOCIATION FOR ARTIFICIAL INTELLIGENCE. 1980. v. 1, p. 5–10. Disponível em: <<http://www.aaai.org/Library/Magazine/Vol01/01-01/Papers/AIMag01-01-002.pdf>>.
- [5] ALLEN, J. et al. Towards conversational human-computer interaction. *AI Magazine*, v. 4, p. 27–38, 2001. Disponível em: <<http://www.cs.rochester.edu/research/cisd/pubs/2001/allen-et-al-aimag2001.pdf>>.
- [6] COHEN, P. R.; LEVESQUE, H. J. Speech acts and rationality. In: *Proceedings of the 23th ACL*. Chicago, IL: [s.n.], 1985. p. 49–60.
- [7] BORBA, F. S. *Introdução aos estudos lingüísticos*. 12. ed. Campinas-SP: Editora Pontes, 1998.
- [8] LANGACKER, R. W. *A Linguagem e sua Estrutura*. 4. ed. Petrópolis-RJ: Editora Vozes, 1980.

- [9] BARTHES, R. *Elementos de semiologia*. 5. ed. São Paulo-SP: Editora Cultrix, 1964.
- [10] ZAENEN, A. Language analysis and understanding. In: _____. [S.l.: s.n.], 1995. (Survey of State of the Art in Human Language Technology), cap. 3.
- [11] VIERA, R.; LIMA, V. L. S. de. *Linguística computacional: princípios e aplicações*. IX Escola de Informática da SBC-Sul, Passo Fundo, Maringá, São José, p. 27–58, 2001.
- [12] LANGACKER, R. W. *Fundamentals of Linguistic Analysis*. 1. ed. California-USA: Harcourt Brace Jovanovich, Inc, 1972.
- [13] CHOMSKY, N. *Estruturas Sintáticas*. Lisboa-Portugal: Edições 70, 1980.
- [14] ALLEN, J. *Natural Language Understanding*. 2. ed. Redwood City-CA: The Benjamin/Cummings Publishing Company, Inc., 1995.
- [15] FLYCHT-ERIKSSON, A. A survey of knowledge sources in dialogue systems. In: ALEXANDERSSON, J. (Ed.). *Proceedings of the IJCAI-99 Workshop on Knowledge and Reasoning in Practical Dialogue Systems*. Murray Hill, New Jersey: International Joint Conference on Artificial Intelligence, 1999. p. 41–48. Disponível em: <<http://citeseer.ist.psu.edu/328028.html>>.
- [16] MCTEAR, M. F. Spoken dialogue technology: Enabling the conversational user interface. *ACM Computing Surveys*, ACM Press, v. 34, n. 1, p. 90–169, 2002.
- [17] SADEK, D. Design considerations on dialogue systems: from theory to technology - the case of artimis. In: *Proceedings of IDS'99 (ETR Workshop on Interactive dialogue for multimedia systems*. Kloster Irsee, Germany: [s.n.], 1999. p. 173–187.
- [18] USZKOREIT, H. Discourse and dialogue. In: _____. *Survey of State of the Art in Human Language Technology*. [S.l.: s.n.], 1995. cap. 6.
- [19] CHURCHER, G. E.; ATWELL, E. S.; SOUTER, C. *Dialogue Management Systems: a Survey and Overview*. 1997.

- [20] DAHLBÄCK, N.; JÖNSSON, A. An empirically based computationally tractable dialogue model. In: *cogsci-92*. Bloomington, Indiana: [s.n.], 1992.
- [21] BRATMAN, M. E.; ISRAEL, D. J.; POLLACK, M. E. Plans and resource-bounded practical reasoning. *Computational Intelligence*, v. 4, p. 349–355, 1988.
- [22] POLLACK, M. Plans as complex mental attitudes. In: *Intentions in Communication*. Cambridge, Massachusetts: MIT Press, 1990. p. 77–103. Disponível em: <<http://citeseer.ist.psu.edu/pollack90plans.html>>.
- [23] MINSKY, M. *A Framework for Representing Knowledge*. [S.l.], 1974. Disponível em: <<http://www.ai.mit.edu/people/minsky/papers/Frames/frames.html>>.
- [24] COHEN, P. R.; LEVESQUE, H. J. Intention is choice with commitment. *Artificial Intelligence*, Elsevier Science Publishers Ltd., v. 42, n. 2-3, p. 213–261, 1990.
- [25] KAMP, H.; REYLE, U. *From Discourse to Logic Part 1*. [S.l.]: Klumer Academic Publishers, 1993.
- [26] HULSTIJN, J. Modelling usability development methods for dialogue systems. In: *Natural Language Engineering*. [S.l.]: Cambridge University Press, 2000. v. 1, n. 1, p. 1–16.
- [27] KOBSA, A. User modeling in dialog systems: Potentials and hazards. v. 4, p. 214–240, 1990.
- [28] Veldhuijzen van Zanten, G. E. User modelling in adaptive dialogue management. In: *Proceedings of the 6th European Conference on Speech Communication and Technology, Eurospeech-99*. [s.n.], 1999. p. 1183–1186. Disponível em: <<http://citeseer.ist.psu.edu/556910.html>>.
- [29] KOBSA, A. Supporting user interfaces for all through user modeling. In: *Proceedings HCI International-95*. Yokohama, Japan: [s.n.], 1995. p. 155–157.

- [30] THEUNE, M. *Natural Language Generation for dialogue: system survey*. [S.l.], May 6 2003.
- [31] MCROY, S. W.; CHANNARUKUL, S.; ALI, S. S. A natural language generation component for dialog systems. In: *Working Notes of the AAAI'99 Workshop on Mixed-Initiative Intelligence*. Orlando, FL: [s.n.], 1999. Disponível em: <<http://citeseer.ist.psu.edu/mcroy99natural.html>>.
- [32] BUNT, H. C. Context and dialogue control. *THINK Quarterly*, v. 3, p. 19–31, May 1994. Disponível em: <<http://citeseer.ist.psu.edu/bunt94context.html>>.
- [33] ALLWOOD, J. Reasons for management in spoken dialog. In: BEUN, R.; BAKER, M.; REINER, M. (Ed.). *Dialogue and Instruction*. Springer-Verlag: [s.n.], 1995. p. 241–250.
- [34] ALLWOOD, J. The structure of dialog. In: TAYLOR, M.; BOUWHUIS, D.; NEAL, F. (Ed.). *The Structure of Multimodal Dialogue II*. Amsterdam, Benjamins: [s.n.], 2001. p. 3–24.
- [35] NIIMI, Y.; KOBAYASHI, Y. A dialog control strategy based on the reliability of speech recognition. In: *International Conference on Spoken Language Processing, ICSLP'96*. [S.l.: s.n.], 1996. p. 534–537.
- [36] DOHSAKA, K. et al. An efficient dialogue control method under system's limited knowledge. In: *Proceedings of the 6th International Conference on Spoken Language Processing (ICSLP-2000)*. [S.l.: s.n.], 2000. p. 739–742.
- [37] ALLEN, J. F. et al. A robust system for natural spoken dialogue. In: *Proceedings of the 34th Annual Meeting of the ACL (ACL'96)*. [S.l.: s.n.], 1996. p. 62–70.
- [38] THE TRAINS Project: Natural Spoken Dialogue and Interactive Planning. Acessado em: 09/05/2004. Disponível em: <<http://www.cs.rochester.edu/research/trains/>>.
- [39] ALLEN, J.; BLAYLOCK, N.; FERGUSON, G. *A Problem Solving Model for Collaborative Agents*. 2002. In First International Joint Conference on Autonomous

Agents and Multiagent Systems, Bologna, Italy, July 15-19 2002, To appear.
Disponível em: <<http://citeseer.ist.psu.edu/allen02problem.html>>.

[40] TRIPS: The Rochester Interactive Planning System. Acessado em: 07/05/2004.
Disponível em: <<http://www.cs.rochester.edu/research/cisd/projects/trips/>>.

[41] FERGUSON, G.; ALLEN, J. TRIPS: An intelligent integrated problem-solving assistant. In: *Fifteenth National Conference on Artificial Intelligence (AAAI-98)*. Madison: [s.n.], 1998. p. 567–573.

[42] A Comparison of TRAINS and TRIPS. Acessado em: 07/05/2004. Disponível em: <<http://www.cs.rochester.edu/research/cisd/projects/trains/tripscompare.html>>.

[43] ALLEN, J.; FERGUSON, G.; STENT, A. An architecture for more realistic conversational systems. In: *Proceedings of the 6th international conference on Intelligent user interfaces (IUI-01)*. Santa Fe-NM: [s.n.], 2001. p. 1–8.

[44] BUD, T.; SCHWINN, J. Verbmobil: The evolution of a complex large speech-to-speech translation system. In: *Proceedings of ICSLP 96*. Philadelphia, PA: [s.n.], 1996. p. 2371–2374.

[45] THE Verbmobil Project. Acessado em: 05/05/2004. Disponível em: <<http://verbmobil.dfki.de/verbmobil/VM2.info.us.html>>.

[46] MCGLASHAN, S. et al. Dialogue management for telephone information systems. In: *Proceedings of the Third Conference on Applied Natural Language Processing*. Trento, Itália: [s.n.], 1992. p. 245–246.

[47] ECKERT, W.; MCGLASHAN, S. Managing spoken dialogues for information services. In: ESCA. *Proceedings of the EUROSPEECH'93*. Berlin, 1993. p. 1653–1656.

[48] MCGLASHAN, S. et al. *Dialogue Management for Telephone Information Systems*. May 1996.

- [49] GODDEAU, D. et al. Galaxy: A human-language interface to on-line travel information. In: *Proceedings of the 1994 International Conference on Spoken Language Processing (ICSLP-94)*. Yokohama, Japan: [s.n.], 1994. p. 707–710.
- [50] SENEFF, S. et al. GALAXY-II: A reference architecture for conversational system development. In: *The 5th International Conference on Spoken Language Processing – (ICSLP'98)*. Sydney, Australia: [s.n.], 1998.
- [51] SENEFF, S.; LAU, R.; POLIFRONI, J. Organization, communication, and control in the GALAXY-II conversational system. In: ESCA. *Proceedings of EUROSPEECH'99*. Budapest, Hungary, 1999.
- [52] POLIFRONI, J.; SENEFF, S. Galaxy-II as an architecture for spoken dialogue evaluation. In: *Proceedings of Second International Conference on Language Resources and Evaluation (LREC)*. Athens, Greece: [s.n.], 2000.
- [53] COHEN, P. R.; PERRAULT, C. R. Elements of a plan-based theory of speech acts. *Cognitive Science*, ABLEX, v. 3, n. 3, p. 177–212, 1979.
- [54] SADEK, M. *Communication Theory = Rationality Principles + Communicative Act Models*. 1994. Disponível em: <<http://citeseer.ist.psu.edu/sadek94communication.html>>.
- [55] SADEK, M. et al. Effective human-computer cooperative spoken dialogue: The ags demonstrator. In: *Proceedings of ICSLP'96*. Philadelphia, USA: [s.n.], 1996.
- [56] RAO, A.; GEORGEFF, M. Modeling rational agents within a BDI architecture. In: ALLEN, J.; FIKES, R.; SANDWALL, E. (Ed.). *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning*. San Mateo, CA: Morgan Kaufmann, 1991. p. 473–484.
- [57] TRAUM, D. R. Speech acts for dialogue agents. In: WOOLDRIDGE, M. J.; ANAND, R. (Ed.). *Foundations And Theories Of Rational Agents*. [S.l.]: Kluwer Academic Publishers, 1999. p. 169–201.
- [58] WOOLDRIDGE, M. J.; JENNINGS, N. Knowledge engineering review. *Intelligent Agents: Theory and Practice*, v. 10, n. 2, p. 155–152, 1995.

- [59] HOEK, W. van der; WOOLDRIDGE, M. Towards a logic of rational agency. *Logic Journal of the IGPL*, v. 11, n. 2, p. 133–157, 2003.
- [60] KAPLAN, R. M. The formal architecture of lexical-functional grammar. In: DALRYMPLE, M. et al. (Ed.). *Formal Issues in Lexical-Functional Grammar*. Stanford, CA: CSLI Publications, 1995.
- [61] WESCOAT, M. T. Practical instructions for working with the formalism of lexical functional grammar. MS, Xerox PARC. 1989.
- [62] NEIDLE, C. Lexical functional grammar. In: *Encyclopedia of Language and Linguistics*. New York: Pergamon Press, 1994. p. 2147–2153.
- [63] KAMP, H.; GENABITH, J. van; REYLE, U. *DRAFT: Discourse Representation Theory*. 2003.
- [64] DALRYMPLE, M. et al. Linear logic for meaning assembly. In: *Proceedings of the Workshop on Computational Logic for Natural Language Processing*. Edinburgh, UK: [s.n.], 1995. Versão revisada em: 2002.
- [65] DALRYMPLE, M.; LAMPING, J.; SARASWAT, V. Lfg semantics via constraints. In: *Proceedings of the 6th Meeting of the European Association for Computational Linguistics*. Utrecht, The Netherlands: [s.n.], 1993. p. 97–105. Disponível em: <<http://citeseer.ist.psu.edu/181032.html>>.
- [66] GIRARD, J.-Y. Linear logic. *Theoretical Computer Science*, v. 50, p. 1–102, 1987.
- [67] GIRARD, J.-Y. Linear logic: Its syntax and semantics. In: GIRARD, J.-Y.; LAFONT, Y.; REGNIER, L. (Ed.). *Advances in Linear Logic*. [S.l.]: Cambridge University Press, 1995. p. 1–42.
- [68] GENABITH, J. van; CROUCH, R. *On Comparing Dynamic and Underspecified Semantics for LFG*. 1997. Disponível em: <<http://citeseer.ist.psu.edu/vangenabith97comparing.html>>.

- [69] PAUCHET, A.; CHAIGNAUD, N.; FALLAH-SEGHRUCHNI, A. E. Modelling planning and interaction in the framework of human cooperative problem solving. In: *The Second International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'03)*. Melbourne, Australia: [s.n.], 2003. Disponível em: <http://www.traclabs.com/~cmartin/hmas/wkshp_2003/papers/Pauchet.pdf>.
- [70] BUNT, H. C. Dialogue control functions and interaction design. In: BEUN, R.; BAKER, M.; REINER, M. (Ed.). *Dialogue and Instruction*. Springer-Verlag: [s.n.], 1995. p. 197–214. Disponível em: <<http://citeseer.ist.psu.edu/bunt95dialogue.html>>.
- [71] BUNT, H. C. Interaction management functions and context representation requirements. In: LUPERFOY, S.; NIJHOLT, A.; ZANTEN, G. van (Ed.). *Proceedings of the Eleventh Twente Workshop on Language Technology*. Enschede, NL, 1996. p. 187–198. Disponível em: <<http://citeseer.ist.psu.edu/bunt96interaction.html>>.
- [72] BUNT, H. C. Dynamic interpretation and dialogue theory. In: TAYLOR, M. M.; NÉEL, F.; BOUWHUIS, D. G. (Ed.). *The Structure of Multimodal Dialogue, Volume 2*. Amsterdam: John Benjamins, 1996. Disponível em: <<http://citeseer.ist.psu.edu/bunt97dynamic.html>>.
- [73] BUNT, H. C. *Belief Contexts in Human-Computer Dialogue*. August 1996. Disponível em: <<http://citeseer.ist.psu.edu/bunt96belief.html>>.
- [74] BUNT, H. C. Dialogue pragmatics and context specification. In: BUNT, H. C.; BLACK, W. (Ed.). *Abduction, Belief and Context in Dialogue: Studies in Computational Pragmatics*. Amsterdam, Benjamins: [s.n.], 2000. p. 81–150.
- [75] TRAUM, D. R.; ALLEN, J. F. A speech acts approach to grounding in conversation. In: *Proceedings 2nd International Conference on Spoken Language Processing (ICSLP-92)*. [S.l.: s.n.], 1992. p. 137–140.
- [76] TRAUM, D. R.; HINKELMAN, E. Conversation acts in task-oriented spoken dialogue. *Computational Intelligence*, v. 8, n. 3, p. 579–599, 1992.

- [77] POESIO, M.; TRAUM, D. R. Conversational actions and discourse situations. *Computational Intelligence*, v. 13, n. 3, p. 309–347, 1997.
- [78] TRAUM, D. R. A reactive-deliberative model of dialogue agency. In: MÜLLER, J. P.; WOOLDRIDGE, M. J.; JENNINGS, N. R. (Ed.). *Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages – ATAL'96: Intelligent Agents III*. Berlin: Springer Verlag, 1997. p. 157–172.
- [79] LONGIN, D.; SADEK, D. Dialogue et dynamique des croyances. In: *Le temps, l'espace et l'évolutif en Sciences du Traitement de l'Information*. [S.l.]: Cépaduès-Éditions, 2000. (Actes des 5èmes journées nationales du PRC-GDR I3 (Information, Interaction, Intelligence), Document et évolution).
- [80] HIRAI, T. *Temporal Linear Logic and Its Applications*. Tese (Doutorado) — The Graduate School of Science and Technology of Kobe University, September 2000.
- [81] SADRZADEH, M. Modal linear logic in higher order logic, an experiment in coq. In: BASIN, D.; BURKHART, W. (Ed.). *Emerging Trends Proceedings of Theorem Proving in Higher Order Logics (TPHOLS)*. ARACNE, Roma: [s.n.], 2003.
- [82] BUENO, F. S. *Minidicionário da Língua portuguesa*. 3. ed. São Paulo-SP: Editora Lisa, 1988.
- [83] GRUBER, T. R. A translation approach to portable ontology specifications. *Knowledge Acquisition*, v. 5, n. 2, p. 199–220, 1993.
- [84] USCHOLD, M.; GRÜNINGER, M. Ontologies: Principles, methods and applications. v. 11, n. 2, p. 93–155, June 1996. Disponível em: <<http://citeseer.ist.psu.edu/uschold96ontologie.html>>.
- [85] NOY, N. F.; MCGUINNESS, D. L. *Ontology Development 101: A Guide to Creating Your First Ontology*. [S.l.], 2001. Disponível em: <ftp://ftp.ksl.stanford.edu/pub/KSL_Reports/KSL-01-05.pdf.gz>.
- [86] HERZIG, A.; LONGIN, D. A logic of intention with cooperation principles and with assertive speech acts as communication primitives. In: *Proceedings of the*

first international joint conference on Autonomous agents and multiagent systems.
Bologna, Italy: ACM Press, 2002. p. 920–927.

Apêndice A

Exemplo de sistema

Neste apêndice será apresentado um exemplo de como poderia funcionar o sistema de diálogo proposta. Isto porque uma implementação completa de tal sistema acarretaria em grandes dificuldades técnicas, uma vez que cada componente do sistema apresenta uma peculiaridade singular em relação ao processamento de linguagem natural, tais como todas as etapas de processamento, desde a entrada de uma frase até sua respectiva representação em formato lógico, para mais detalhes confira o capítulo 2. Assim, o exemplo apresentado aqui tentará descrever um possível cenário de uso do sistema.

O cenário do exemplo será o seguinte: o sistema age como interface para um programa de reservas de passagens permitindo a usuário desta aplicação consultar, por meio de perguntas escritas, a disponibilidade de viagens para os destinos, fornecidos por uma base de dados administrada pelo próprio programa.

Relembrando a estrutura proposta (figura A.1). O fluxo de funcionamento do sistema de diálogo seria o seguinte: após a entrada textual (Entrada) o texto é repassado ao componente de Entendimento de Linguagem. Nele, ocorrerão as análises léxico-morfológica, sintática e semântica. Na primeira as palavras serão verificadas quanto sua ortografia e formação lexical, ou seja, será feita a verificação se uma palavra: está presente no dicionário do sistema, escrita corretamente e se possível extrair algumas informações lexicais, a serem utilizadas na análise sintática.

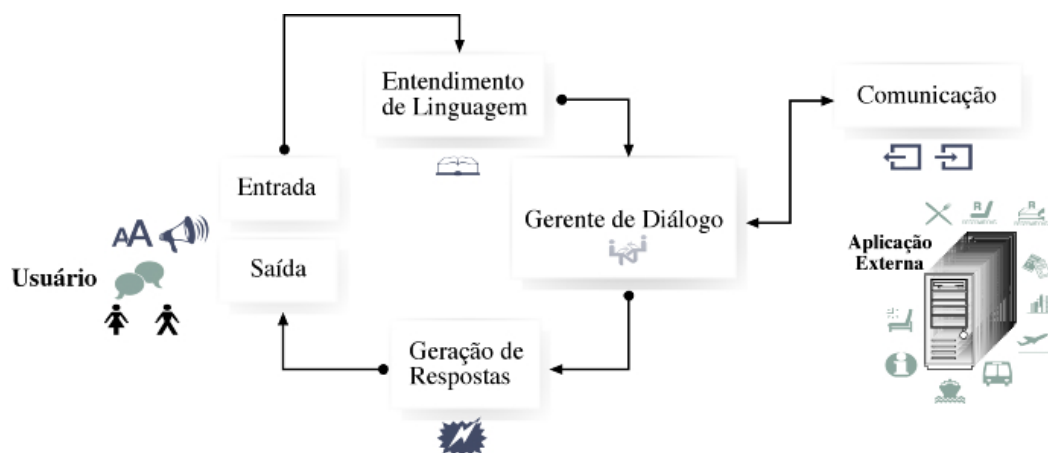


Figura A.1: Estrutura de um sistema de diálogo genérico

Na análise sintática¹, recomenda-se o uso de um formalismo gramatical (apresentados no capítulo 2), no caso deste trabalho o LFG². O processo de análise sintática fará a verificação da concordância de uma frase com as regras gramaticais de uma dada língua. Computacionalmente este processo resulta na produção de árvores sintáticas que expressam as relações existentes entre as palavras de uma frase. O uso do formalismo LFG permite a inserção de algumas informações sintáticas que serão úteis durante a análise semântica.

Por fim, a análise semântica fornecerá uma possível interpretação para o significado da frase inserida no sistema pelo usuário. Neste processo a árvore sintática fornecida pela análise sintática será transformada em uma sentença lógica, a qual fornecerá o significado semântico da frase analisada. Essa transformação é realizada por uma teoria de “cola-semântica”, onde com uma árvore sintática e uma estrutura ou conjunto de regras de ligação obtém-se uma sentença lógica.

Terminado a fase de análises de uma frase o passo seguinte do fluxo de processa-

¹Uma nota faz-se necessária, o processo de análise sintática pode produzir mais de uma árvore sintática para uma frase, isto ocorre devido a uma possível ambigüidade inerente da frase em questão e não apenas ao processo de produção de árvores sintáticas.

²Um exemplo do uso de LFG foi fornecido no capítulo 5.

mento será a descoberta ou enquadramento da sentença lógica a um ato de fala. Nesta etapa as informações extraídas da sentença lógica auxiliaram o sistema, mais precisamente o componente de gerência de diálogo, a classificar a frase. Essa classificação obedece à noção de atos de fala apresentada no capítulo 5.

Tal classificação ocorre pela análise de determinados predicados lógicos, os quais podem ser qualificados em categorias de atos de fala. Então após uma sentença lógica ser produzida, seu predicado principal é classificado de acordo com atos de fala e dessa forma tem-se a identificação do ato de fala associado à frase de entrada.

De posse da informação semântica, em formato lógico, e do ato de fala correspondente à frase de entrada, o passo posterior será a consulta à fonte de conhecimentos a que o sistema de diálogo participa como interprete. Para o caso de um sistema do tipo pergunta-resposta, categorias de atos de fala mais importantes seriam aquelas onde se poderia identificar o desejo de obtenção de informações e de notificação de desejos.

O sistema de diálogo, em seu componente de comunicação, possui uma cópia da estrutura da ontologia correspondente às informações da aplicação externa, neste exemplo um programa de verificação de destinos para viagens. Essa cópia fornecer ao componente de comunicação o meio necessário para a comunicação com a aplicação externa.

Em outras palavras, a partir da estrutura ontológica de um domínio de conhecimento, o sistema de diálogo pode verificar se a informação semântica contida na frase analisada pertence ao domínio da aplicação e ainda se essa informação necessitará de complementos adicionais, isto é, informações que estejam ligadas por dependências onde, por exemplo, ao se perguntar sobre dia de uma reserva espera-se outra pergunta sobre seu horário e validade, neste caso as informações possuem um certo nível de dependência complementar.

Se a estrutura ontológica do domínio deste exemplo fosse hierárquica, um possível

seguimento seria o seguinte, em notação XML³:

```

...
<Reserva>
    <Reserva_Viagem/>
    <Reserva_Simples>
        <Reserva_Alojamento/>
        <Reserva_Transporte>
            <Reserva_Vôo/>
        </Reserva_Transporte>
    </Reserva_Simples>
</Reserva>
...

```

O retorno de informação obtida junto à aplicação externa dependerá do processo de comunicação entre o sistema de diálogo e a aplicação externa. Sendo assim, considerando que não existam erros de comunicação, o componente de comunicação do sistema de diálogo fará uma consulta à aplicação externa utilizando para isto a linguagem ontológica proposta no capítulo 5. Então, após o processo comunicativo entre os dois programas e ainda considerando que a houve sucesso na busca, o componente de comunicação repassa ao componente de gerência de diálogo as informações requeridas.

A informação obtida será então comparada com a informação semântica e com o ato de fala, utilizados para iniciar a consulta com a aplicação externa. Caso tal comparação seja satisfatória, então o componente de gerência de diálogo adicionará ao histórico de diálogo: a informação semântica de entrada, o ato de fala derivado e a resposta obtida junto à aplicação externa. Porém, no caso do processo comparativo não resultar em dados satisfatórios, o componente de gerência tem de utilizar-se de alguma estratégia de correção de erros, algumas delas foram apresentadas no capítulo 3.

Prosseguindo com o processamento, a informação obtida será então codificada em

³*eXtensible Markup Language* (XML) é um formato de texto derivado do SGML (ISO 8879) que permite a codificação de uma grande variedade de dados. Fonte: *World Wide Web Consortium* (W3C) <<http://www.w3.org>>.

formato lógico para então ser enquadrada em uma das categorias de atos de fala disponíveis no sistema de diálogo. Assim, o componente de gerência de diálogo pode acionar o componente de geração de resposta para então gerar a resposta do sistema ao usuário.

No componente de geração as fases de processamento de geração textual seguirão o caminho inverso ao adotado pelo componente de entendimento de linguagem, a saber, a partir de uma sentença em formato lógico e de um ato de fala, para denotar a intenção da fala do sistema, este componente irá gerar uma árvore sintática e a partir dela, por meio de consulta em dicionário e gramática da língua de saída do sistema, gerará uma frase. Esse processo, apesar de não é o mais sofisticado, pois não elimina o aspecto ambigüidade da produção de fala, garante geração de respostas simples (capítulo 3).

Este exemplo tentou fornecer uma visão de como seria o funcionamento de um sistema de diálogo que se adota a proposta deste trabalho. Ressalta-se ainda que apesar dele não ter apresentado detalhes suficientes para demonstrar o processamento real de um sistema de diálogo, sua importância permanece inalterada, pois ele mostra a idéia principal de funcionamento de um sistema de diálogo textual.