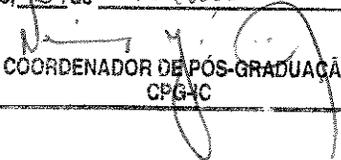


Este exemplar corresponde à redação final da
Tese/Dissertação devidamente corrigida e defendida
por: Norton Trevisan Roman
e aprovada pela Banca Examinadora.
Campinas, 27 de março de 2002

COORDENADOR DE PÓS-GRADUAÇÃO
CPG4C

Estudo de Diálogos Orientados à Tarefa
Usando a Teoria de Multiagentes

Norton Trevisan Roman

Dissertação de Mestrado

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

Estudo de Diálogos Orientados à Tarefa Usando a Teoria de Multiagentes

Norton Trevisan Roman¹

Agosto de 2001

Banca Examinadora:

- Profa. Dra. Ariadne Maria Brito Rizzoni Carvalho (Orientadora)
- Profa. Dra. Vera Lúcia Strube de Lima
Faculdade de Informática - PUC-RS
- Profa. Dra. Heloisa Vieira da Rocha
Instituto de Computação - UNICAMP
- Prof. Dr. Jacques Wainer (Suplente)
Instituto de Computação - UNICAMP

¹Com apoio da FAPESP (processo 99/00682-5).

UNIDADE	Be
Nº CHAMADA	T/UNICAMP
	R661e
V	SE
COMO	4838Z
	16-837102
RECIBO	R\$ 11,00
DATA	18/04/02
Nº CPD	

CM00166484-9

18 ID 236642

**FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO IMECC DA UNICAMP**

Roman, Norton Trevisan

R661e Estudo da teoria de multiagentes aplicada a diálogos orientados à tarefa / Norton Trevisan Roman – Campinas, [S.P. :s.n.], 2001.

Orientador : Ariadne Maria Brito Rizzoni Carvalho

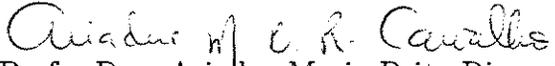
Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Computação.

1. Processamento de linguagem natural (Computação). 2. Sistemas multiagentes. 3. Inteligência artificial. I. Carvalho, Ariadne Maria Brito Rizzoni. II. Universidade Estadual de Campinas. Instituto de Computação. III. Título.

Estudo de Diálogos Orientados à Tarefa Usando a Teoria de Multiagentes

Este exemplar corresponde à redação final da
Dissertação devidamente corrigida e defendida
por Norton Trevisan Roman e aprovada pela
Banca Examinadora.

Campinas, 19 de Setembro de 2001.


Profa. Dra. Ariadne Maria Brito Rizzoni
Carvalho (Orientadora)

Dissertação apresentada ao Instituto de Com-
putação, UNICAMP, como requisito parcial para
a obtenção do título de Mestre em Ciência da
Computação.

08/09/2001

TERMO DE APROVAÇÃO

Tese defendida e aprovada em 19 de setembro de 2001, pela Banca Examinadora composta pelos Professores Doutores:

Vera Lúcia Strube de Lima

Profa. Dra. Vera Lúcia Strube de Lima
PUC - RS

Heloísa Vieira da Rocha

Profa. Dra. Heloísa Vieira da Rocha
IC - UNICAMP

Ariadne M. B. R. Cavalho

Profa. Dra. Ariadne Maria Brito Rizzoni Cavalho
IC - UNICAMP

Resumo

O processamento de diálogos em uma língua natural é um campo bastante vasto e não suficientemente explorado da Inteligência Artificial. Os inúmeros detalhes e particularidades existentes não só na própria língua, como também na estrutura de um diálogo em si impõem barreiras ao modelamento deste processamento. Dessa forma, para se modelar efetivamente um diálogo faz-se necessário o modelamento não só das regras gramaticais particulares da língua, mas também de fenômenos próprios de diálogos, como elipses, o uso de pronomes e a coerência entre as idéias expressas por cada fala presente.

Nessa dissertação é desenvolvido um sistema multiagentes que implementa o modelo para tratamento de diálogos originalmente desenvolvido por Grosz e Sidner. O objetivo deste sistema é testar a interação entre as estruturas que, de acordo com estas pesquisadoras, compõem a estrutura de um discurso, fazendo-o sob a luz do paradigma de multiagentes.

Abstract

Natural language dialogue processing is a very extensive and yet not enough explored field in Artificial Intelligence. The innumerous details and particularities, not only within the language itself, but also within the dialogue structure, present several obstacles to modeling the process. Therefore, in order to effectivally model a dialogue, it is necessary to model not only particular grammatical rules, but also the dialogue phenomena, wich involves ellipsis, the use of pronouns and the coherence between the ideas expressed by each utterance in the dialogue.

In this work a multiagent system that implements the dialogue modeling originally proposed by Grosz and Sidner is developed. The system's goal is to test the interaction between the structures that, according to these researchers, compose the discourse structure, using the multiagents paradigm.

Agradecimentos

Aos meus pais, por terem me criado insano o suficiente para não ter medo de ir em frente e por nunca terem me negado acesso ao conhecimento; à minha orientadora, por ter me apoiado em minha insanidade, e à FAPESP, pelo apoio financeiro.

Conteúdo

Resumo	v
Abstract	vi
Agradecimentos	vii
I Introdução	1
1.1 O Cenário de Estudo	2
1.2 A Contribuição deste Trabalho	3
1.3 O Cenário de Teste	5
1.4 Estrutura da Dissertação	5
II Fundamentos da Teoria	7
2.1 Definições Básicas	7
2.2 Modelo Teórico	8
2.2.1 Estrutura Lingüística	8
2.2.2 Estrutura Intencional	9
2.2.3 Estado Atencional	11
2.3 Interrupções	18
2.3.1 Interrupções Verdadeiras	18
2.3.2 Digressões	20
2.3.3 Flashbacks	20
2.4 Retornos Semânticos	22
III Sistemas Multiagentes e Colaboração	24
3.1 Introdução	24
3.2 Agentes	24
3.2.1 Características	25
3.3 Sistemas Multiagentes	26
3.4 Atividade Colaborativa	26

3.4.1	Características de Sistemas Colaborativos	27
3.4.2	A Atividade Colaborativa em um Discurso Orientado à Tarefa	29
3.4.3	Planos Compartilhados	30
3.4.4	Papel dos Planos Compartilhados no Processamento do Discurso	33
IV Modelamento do Sistema Desenvolvido		37
4.1	Visão Geral do Sistema	37
4.2	Estrutura Lingüística	39
4.2.1	Modelamento Proposto	40
4.2.2	Tipos de Ilhas da Língua Portuguesa	41
4.2.3	Características do modelo	43
4.3	A Estrutura Atencional	44
4.4	A Estrutura da Tarefa	45
4.5	A Estrutura Intencional	46
4.5.1	Sistema Desenvolvido por Lochbaum	46
V Implementação do Sistema		51
5.1	Visão Geral	51
5.1.1	A Forma Lógica	53
5.2	Estrutura Lingüística	54
5.2.1	O Agente EL1	59
5.2.2	O Agente EL2	61
5.2.3	O Agente EL3	62
5.2.4	Agente EL4	66
5.3	Estrutura da Tarefa	69
5.4	Estrutura Atencional	72
5.4.1	O Agente FI	73
5.4.2	O Agente FG	75
5.5	Estrutura Intencional	76
5.5.1	A Estrutura Intencional Desenvolvida	76
VI Exemplos do Sistema em Operação		79
6.1	Separação das Ilhas	79
6.2	Centros de Busca	80
6.3	Compreensão e Resposta	82
6.4	O Sistema em Operação	84
6.4.1	Diálogo I	84
6.4.2	Diálogo II	87

VI	Conclusão	92
7.1	Em Aberto	93
A	As Comunicações Entre os Agentes	95
	Bibliografia	102

Lista de Tabelas

A.1	Comunicação entre EL1 e EL.	95
A.2	Comunicação entre EL2 e EL.	95
A.3	As ações tomadas pelo EL, de acordo com a comunicação recebida.	96
A.4	Comunicação entre EL3 e outros agentes.	97
A.5	Comunicação entre EL4 e outros agentes.	98
A.6	Comunicação entre a ET e as outras estruturas.	98
A.7	Comunicação entre o FI e as outras estruturas e agentes.	99
A.8	Comunicação entre o FG e as outras estruturas e agentes.	100
A.9	Comunicação entre o EI e outros agentes.	101

Lista de Figuras

1.1	Diálogo entre o usuário (U) e o sistema (S).	5
2.1	Diálogo entre um especialista (E) e um aprendiz (A). Iniciado pelo especialista.	11
2.2	Um modelo simples da interação entre as três estruturas do discurso, mostrando como o estado atencional coordena as estruturas lingüística e intencional.	12
2.3	Pequeno exemplo da extração dos centros de uma fala.	15
2.4	Narrativa exemplificando as operações com os centros de busca.	16
2.5	Dois diálogos, um mostrando uma mudança constante de foco (à direita) e outro (esquerda) mantendo um único foco.	17
2.6	Exemplo de interrupção verdadeira.	19
2.7	Modelamento de um <i>flashback</i> usando uma pilha de espaços de foco auxiliar.	21
2.8	Modelamento de um <i>flashback</i> tratando-o como um segmento embutido normal.	22
2.9	Diálogo exemplificando o uso de um retorno semântico.	23
3.1	Planos colaborativos não são simplesmente a soma de planos individuais.	28
3.2	Descrição de um PCP.	32
3.3	Descrição de um PCC.	33
3.4	Processo de Incrementação do PC – Geração.	34
3.5	Processo de Incrementação do PC – Interpretação.	35
4.1	Esquema geral do sistema.	38
4.2	Gramática simples para o português.	39
4.3	Algoritmo para a tradução de uma oração em sua forma lógica.	42
4.4	Representações de receita e GrafoR.	47
4.5	Algoritmo de construção do GrafoR.	47
4.6	Esquema do sistema desenvolvido por Lochbaum.	48
4.7	Macros e seu processamento para as possíveis entradas do usuário.	49
5.1	Esquema geral do sistema.	52

5.2	Mensagens entre as estruturas ao processar a fala "mostre a porta da frente".	52
5.3	Esquema de formação da Estrutura Lingüística.	55
5.4	Processamento de uma fala dentro da EL.	57
5.5	Processamento de outra fala dentro da EL.	58
5.6	Processamento de duas falas dentro da EL.	60
5.7	Algoritmo para a separação de ilhas.	63
5.8	Algoritmo para verificação da compatibilidade entre duas palavras.	64
5.9	Algoritmo para codificação lógica da oração (já separada em ilhas).	67
5.10	Algoritmo para codificação lógica da oração (continuação).	68
5.11	Esquema geral da Estrutura Atencional.	72
6.1	Primeiro diálogo teste entre o Usuário e o Sistema.	83
6.2	Segundo diálogo entre o Usuário e o Sistema.	83

Capítulo I

Introdução

O tratamento computacional de um diálogo ainda é uma questão em aberto no campo da Inteligência Artificial. Isso porque um diálogo possui várias características de difícil modelamento.

O problema não se restringe única e exclusivamente à tradução de uma fala em língua natural para uma forma lógica, mas exige também que essa fala seja compreendida dentro de um contexto, e que uma resposta coerente com este mesmo contexto seja devolvida.

Então, além de fatores lingüísticos, como o uso correto da grafia e gramática, outros fatores, como atenção e intenções, devem ser levados em conta quando do processamento de um diálogo. Desse modo, para efetivamente compreender uma fala, um sistema deve saber não somente o que foi dito, mas também o contexto em que foi dito e porque foi dito.

O papel do foco da atenção dos participantes de um diálogo no processamento deste se torna bastante evidente ao se tratar de pronomes. Estes, por se referirem a elementos na maioria das vezes externos à oração, são profundamente dependentes do contexto ao qual a atenção dos participantes do diálogo está voltada. Naturalmente, mesmo frases sem pronomes também são dependentes do contexto em que foram ditas. Por exemplo, a frase "venha à feira de calçados de Franca", fora de um contexto, só chama para uma feira de calçados, não estando claro se a feira é em Franca, se os calçados são de Franca, ou ambos. Mesmo uma frase não ambígua, como "a porta está aberta", para ser entendida depende do conhecimento de que porta se está falando, ou seja, da porta que está no foco do diálogo.

Mas para compreender mesmo uma fala em um diálogo, resta ainda saber como esta se encaixa na lógica deste, ou seja, como a idéia que ela transmite se encaixa no conjunto de idéias já transmitidas pelas falas que a precederam, tornando o diálogo coerente.

Tal reconhecimento exige um mecanismo que possa não somente reconhecer a intenção que está por trás do ato de comunicar uma determinada fala, mas também que consiga

identificar as relações que esta intenção possui com as intenções das falas precedentes. Mais do que isto, deve também reconhecer qual a contribuição desta intenção para a intenção mais geral do diálogo, aquela que motivou o início deste e que norteia seu desenvolvimento.

Para poder modelar as características de um diálogo, Grosz e Sidner (1986) desenvolveram uma teoria acerca da estrutura deste (tratado pelas autoras também como discurso). Segundo esta teoria, um diálogo é composto por três estruturas independentes porém interagentes: uma *Estrutura Lingüística*, responsável pelo processamento e tratamento da estrutura das falas em um discurso; uma *Estrutura Atencional*, responsável pelo foco da atenção dos participantes deste discurso, ou seja, responsável pelo contexto em que o discurso se dá; e uma *Estrutura Intencional*, responsável pela detecção e compreensão da estrutura dos propósitos e intenções contidos nas falas que surgem quando do desenrolar do diálogo.

Além disto, a teoria também afirma que as falas em um diálogo são naturalmente agregadas em segmentos e sub-segmentos, cada um destes possuindo um propósito único, cuja função é auxiliar para que o propósito geral do diálogo seja atingido. Nesse contexto a intenção contida em cada fala do diálogo tem o papel de ajudar na definição e obtenção do propósito do segmento do discurso no qual ela está.

1.1 O Cenário de Estudo

O trabalho desenvolvido inicialmente por Grosz e Sidner (1985, 1986) apresenta um ponto de partida bastante importante, tratando não somente das definições básicas, mas também dando idéias de como se poderia fazer efetivamente um sistema que pusesse em prática a teoria proposta. No entanto, este trabalho ainda carece de definições e estudos mais detalhados acerca do funcionamento de cada estrutura.

A Estrutura Atencional consegue modelar satisfatoriamente o contexto geral em que o diálogo se dá, mas falha ao modelar o contexto mais específico, aquele existente entre pequenos grupos de frases, essencial para o referenciamento de pronomes e elipses.

Então, para preencher esta lacuna, neste trabalho de mestrado foi dada atenção a uma outra teoria, anterior à de Grosz e Sidner (mencionada anteriormente) – a Teoria de Centralização (Grosz 1978, Joshi 1979).

De acordo com esta teoria, cada fala em um discurso possui um centro de busca regressiva, que a conecta à sua predecessora, e um conjunto de entidades formando um centro de busca progressiva, que representa as possíveis conexões com a frase que a sucede, contendo os elementos que ficaram mais salientes, direta ou indiretamente, nesta fala. Tais centros teriam papéis decisivos na compreensão de como uma fala se conecta com as outras falas no discurso, sendo importantes fontes para busca de referências a pronomes e elipses,

e instrumentos úteis para a manutenção da coerência do diálogo.

Então, enquanto a teoria apresentada por Grosz e Sidner (1985, 1986) modela o foco global mantido pelos participantes do diálogo, a teoria de centralização busca modelar o foco imediato, existente entre uma fala e suas vizinhas.

A Estrutura Intencional, por sua vez, também apresenta no trabalho de Grosz e Sidner (1986) uma carência de estudo mais profundo, pois são apresentadas definições sobre intenções e relações entre estas, mas nada é dito sobre como alguém pode identificar as relações que surgem entre as intenções. O reconhecimento das intenções também se mostra bastante nebuloso, baseado apenas no conteúdo proposicional da fala.

Lochbaum (1994), em seu trabalho de doutorado, modelou esta estrutura usando uma outra teoria como base: a Teoria de Planos Compartilhados apresentada por Grosz e Sidner (1990). Esta teoria foi criada numa tentativa de modelar colaboração e, assim, permitir o modelamento da Estrutura Intencional do discurso. Ela se baseia na premissa de que diálogos são exemplos de comportamento colaborativo e que, dessa forma, ao modelar este comportamento, se estará ajudando a modelar o comportamento do próprio diálogo. A contribuição de Lochbaum neste contexto foi, então, de entender o diálogo como um exemplo de colaboração e transportar a Teoria de Planos Compartilhados para o contexto de diálogos.

1.2 A Contribuição deste Trabalho

Conforme visto acima, há duas teorias complementares que descrevem o funcionamento da Estrutura Atencional do discurso (Grosz e Sidner 1985, 1986 e Joshi 1979) e um trabalho que modela o funcionamento da Estrutura Intencional (Lochbaum, 1994).

A Estrutura Lingüística, porém, ainda não possui um modelamento bem definido. O que se sabe sobre ela é que sua função é a de processar as falas em um discurso, extraindo destas seu conteúdo proposicional. Para tal, ela pode ser auxiliada pelas outras estruturas e por fatores lingüísticos específicos, como marcas de discurso, que são frases que servem apenas para indicar contornos entre segmentos, ou que uma fala de um participante foi compreendida pelo outro.

Esse contexto motivou a execução desta pesquisa que buscou, dentre outras coisas, reunir as teorias e resultados já obtidos em um único trabalho para que, desta forma, pudesse ser feita uma avaliação não somente de cada teoria e modelo usados, mas também de sua integração.

Para tal foi desenvolvido um sistema computacional que aplica estas teorias (antes analisadas separadamente) em conjunto para o tratamento de diálogos orientados à tarefa, fazendo isso sob a luz do paradigma de multiagentes.

Esse tipo de diálogo foi escolhido por restringir a informação usada pelo sistema,

simplificando o seu desenvolvimento. Dessa forma, o vocabulário usado e as ações que devem ser executadas são limitadas pela tarefa que está sendo executada.

Além disso, a estrutura do diálogo também é simplificada uma vez que, como mencionam Grosz e Sidner (1986), os diálogos orientados à tarefa têm como característica o fato de que a intenção principal do participante que inicia o diálogo é que o outro participante execute alguma ação, sendo que o outro participante, por sua vez, adota a intenção de executar esta ação. Assim, um diálogo orientado à tarefa apresenta uma estrutura baseada em perguntas, respostas, ordens, pedidos e fornecimento de informação centradas na tarefa a ser executada ou em alguma sub-tarefa que a compõe.

O cenário de teste escolhido foi o de um sistema de segurança domiciliar que fosse capaz de interagir com o usuário em uma língua natural, recebendo deste ordens e perguntas, processando-as e devolvendo as respostas sob a forma de ações ou frases em língua natural. Mais detalhes sobre o cenário de teste serão vistos a seguir.

O uso do modelo de multiagentes foi motivado pelo fato deste ter como característica principal a interação entre módulos independentes, característica esta intrínseca à teoria desenvolvida por Grosz e Sidner (1986). Além disso, buscou-se obter um ganho em simplicidade no sistema, ou seja, houve uma expectativa de que o uso do modelo de multiagentes tornasse o projeto do sistema mais simples, sem comprometer seu funcionamento.

Devido ao fato de a Estrutura Lingüística ainda não possuir um modelamento definido, uma das contribuições deste trabalho foi fazer o modelamento, especificamente para a língua portuguesa, para a implementação desta estrutura e sua integração com as outras estruturas do diálogo. Por se tratar da língua portuguesa, foi feito um estudo do comportamento desta língua, assim como um modelamento computacional para este comportamento. Em adição, foi necessário o desenvolvimento de meios para que o modelo desenvolvido pudesse tirar proveito do trabalho das estruturas Atencional e Intencional.

Para dar uma maior generalização ao uso deste sistema foi desenvolvida uma nova estrutura – a Estrutura da Tarefa – responsável pela informação dependente do domínio que o sistema necessita para compreender o diálogo (como, por exemplo, o vocabulário utilizado). O objetivo por trás de sua criação foi permitir que o mesmo sistema pudesse ser usado em domínios diferentes, bastando para tal que esta estrutura fosse substituída por uma nova versão contendo a informação dependente do domínio necessária para o novo cenário.

Por fim, uma última contribuição deste trabalho foi a união destes três trabalhos complementares (estruturas do discurso, centros e o trabalho de Lochbaum) mais o modelamento para as estruturas Lingüística e da Tarefa aqui desenvolvido em um único sistema multiagentes que tratasse de diálogos orientados à tarefa, para que seu funcionamento pudesse ser testado, possibilitando assim que uma avaliação mais precisa de cada teoria e, principalmente, de sua complementaridade, sob a luz da teoria de multiagentes, pudesse

ser feita.

1.3 O Cenário de Teste

O cenário escolhido para testar o sistema, conforme mencionado, foi a interação entre o usuário e um sistema de segurança domiciliar.

O sistema seria, então, o responsável pelo monitoramento de todas as portas e janelas de uma residência. Para tal ele disporia de alarmes conectados a cada porta e janela, além de um circuito interno de câmeras que permitisse monitorar todas as entradas da residência, registrando imagens destas entradas vinte e quatro horas por dia.

Ao detectar que uma porta ou janela foi aberta, o sistema guardaria em um arquivo informação sobre qual porta ou janela foi aberta e o momento quando isto ocorreu.

Chegando em casa, o usuário poderia interrogar o sistema acerca do que ocorreu em sua ausência, podendo inclusive obter imagens de alguma porta ou janela no instante de sua preferência.

Este cenário usado nos testes não corresponde a um cenário real, tendo sido criado por meio de simulações apenas para testar o sistema.

A Figura 1.1 mostra um pequeno diálogo ilustrando a interação entre o sistema e o usuário. Nele o usuário deseja saber o que ocorreu durante sua ausência (no caso, após as 14 horas). O sistema, então, começa a narrar o ocorrido, sendo interrompido, vez ou outra, por ordens ou perguntas do usuário, como na terceira fala do exemplo.

Usuário	-	O que aconteceu após as 14:00?
Sistema	-	A porta da frente foi aberta às 14:30.
Usuário	-	Mostre.
Sistema	-	<imagem>
Usuário	-	Tudo bem, prossiga.

Figura 1.1: Diálogo entre o usuário (U) e o sistema (S).

Pode-se observar também, da Figura 1.1, que as falas usadas neste diálogo são todas dirigidas ao objetivo mantido pelo usuário de obter uma descrição detalhada dos fatos ocorridos em sua ausência. Isso cria um estilo de conversação mais simples e direto, facilitando o processamento de cada fala.

1.4 Estrutura da Dissertação

O restante da dissertação está assim dividido:

No Capítulo 2 é explicada a teoria de discurso desenvolvida por Grosz e Sidner (1986), bem como a Teoria de Centralização (Gordon 1993, Grosz 1978, 1981, 1995, 1998 e 1999a) e seu uso na Estrutura Atencional do discurso. Também é feito um resumo dos vários tipos de interrupções que a teoria prevê e como elas podem ser tratadas.

Como este trabalho dá ênfase ao uso de multiagentes, o Capítulo 3 traz definições básicas acerca destes, caracterizando agentes, sistemas multiagentes e comportamento colaborativo, fazendo, inclusive, um resumo da teoria de Planos Compartilhados e seu uso no processamento de um diálogo.

O Capítulo 4 trata do modelamento usado em cada estrutura do discurso (Grosz e Sidner, 1986), apresentando os modelos propostos para o tratamento da Estrutura Lingüística e da Tarefa, bem como os modelamentos existentes para as estruturas Atencional e Intencional.

A implementação do sistema e os algoritmos utilizados são apresentados no Capítulo 5. Nele também são encontradas descrições de cada agente que compõe o sistema.

O Capítulo 6 contém alguns testes criados para demonstrar as características e funcionamento do sistema e, por fim, o Capítulo 7 apresenta as conclusões resultantes deste trabalho e as questões deixadas em aberto.

Capítulo II

Fundamentos da Teoria

Esse capítulo é dedicado à descrição da teoria geral apresentada por Grosz e Sidner (1985, 1986) para o processamento de um discurso orientado à tarefa, teoria esta usada como base neste trabalho. Inicialmente a teoria é explicada conforme descrita por suas autoras. Em adição às idéias iniciais apresentadas, é introduzida a teoria de centralização e, ao final do capítulo são apresentados alguns tipos de interrupções e seu tratamento conforme a teoria.

2.1 Definições Básicas

Antes de tratar do processamento de discurso (ou diálogo, como poderá ser também utilizado aqui) é necessário definir discurso. Um discurso, segundo Grosz e Sidner(1986), é uma parte do comportamento da linguagem que tipicamente envolve múltiplas falas e múltiplos participantes, podendo ser produzido por um ou mais desses participantes, como oradores ou escritores, e podendo a audiência compreender um ou mais dos participantes também, como ouvintes ou leitores.

A ênfase desse trabalho está no processamento de diálogos orientados à tarefa, que são diálogos que se preocupam com a execução de uma tarefa; seu tópico é esta tarefa. A escolha desse tipo de diálogo não restringe significativamente a gama de discursos que podem ser tratados aqui, uma vez que, como afirmou Grosz (1985), discursos são geralmente de algum modo "orientados à tarefa", sendo que os tipos de "tarefa" executadas são bastante variados – algumas são físicas, outras mentais, outras lingüísticas.

Para que se possa fazer o processamento do discurso é necessário o conhecimento de como as falas deste se agregam em segmentos, das intenções expressas no discurso e das relações entre elas, e o acompanhamento do foco da atenção dos participantes, por meio de algum mecanismo (Grosz 1986).

Nesse âmbito, a teoria sobre a estrutura do discurso proposta por Grosz e Sidner

(1985, 1986, 1999a) é uma tentativa de suprir essas necessidades para o processamento do discurso. Segundo essa teoria, a estrutura de qualquer discurso é composta de três partes distintas, porém interagentes: a estrutura da seqüência de falas atual no discurso, a estrutura de intenções e o estado atencional. Cada uma delas será discutida mais detalhadamente na próxima seção.

Juntos, os três constituintes fornecem a informação necessária aos participantes conversacionais para determinar como uma fala individual se encaixa no resto do discurso, ou seja, por que foi dita e o que significa no contexto em que foi dita. Este contexto forma a base para certas expectativas acerca do que está por vir, tendo essas expectativas papel relevante na acomodação de novas falas no discurso (Grosz 1986).

2.2 Modelo Teórico

O modelo que aqui será apresentado foi originalmente proposto por Grosz e Sidner (1985, 1986). De acordo com esse modelo, conforme já mencionado, a estrutura de qualquer discurso é composta de três componentes distintos, porém interagentes: uma estrutura lingüística, que é a estrutura da seqüência de falas no discurso; uma estrutura intencional, que é a estrutura dos propósitos e intenções que ficam salientes no discurso à medida que este se desenrola; e um estado atencional, que é o estado do foco da atenção dos participantes do discurso. Cada uma delas será discutida separadamente nas próximas seções, sendo que seu modelamento computacional será apresentado no Capítulo 4.

2.2.1 Estrutura Lingüística

A Estrutura Lingüística é a estrutura da seqüência de falas que um discurso compreende. Não se trata da estrutura sintática de uma única sentença, mas da estrutura de uma seqüência de falas.

Falas – o dizer ou escrever de seqüências particulares de palavras – são os elementos básicos desta estrutura. Num discurso, as falas formam segmentos de discurso que, por sua vez, podem conter uma combinação de subsegmentos e falas. É possível, então, que duas falas consecutivas estejam em segmentos diferentes, bem como que duas falas não consecutivas estejam no mesmo segmento. Sobre os segmentos de discurso, Grosz (1985) afirma o seguinte:

Da mesma forma que palavras em uma única sentença formam frases constituintes, as falas em um discurso são naturalmente agregadas em segmentos de discurso. As falas em um segmento, como as palavras em uma frase, têm papéis particulares com respeito a esse segmento.

Sendo assim, podemos dizer que a Estrutura Lingüística consiste dos segmentos de discurso e das relações implícitas que podem aparecer entre eles.

Um ponto ainda nebuloso é onde está a fronteira de um segmento. Em suas pesquisas, Mann (1975) pediu para algumas pessoas dividirem em segmentos um conjunto de diálogos. As pessoas segmentaram os discursos aproximadamente do mesmo modo, sendo que as discordâncias surgiam somente com respeito às falas situadas nos limites dos segmentos. Então podemos concluir que reconhecemos a segmentação, embora não tenhamos certeza de seus contornos exatos.

Dentre os principais indicadores de fronteiras de segmentos estão as expressões lingüísticas (como, por exemplo, "em primeiro lugar", "e finalmente" etc). Outros marcadores seriam entonação, mudança na tensão etc. Naturalmente, ainda a melhor "pista" é a semântica, pois uma variação no assunto discutido indica uma mudança de segmento.

De qualquer forma, a segmentação do discurso é um fator importante para sua compreensão, pois afeta a interpretação das expressões lingüísticas situadas neste segmento. Ela restringe o uso de expressões referenciais ao delinear certos pontos nos quais há uma mudança significativa em quais entidades são discutidas.

Nesse sentido, de grande ajuda na segmentação do discurso são as ações da Estrutura Intencional, responsável pelo reconhecimento e tratamento das intenções por trás de cada segmento de discurso, e do Estado Atencional, que é o responsável pelo tratamento do foco da atenção dos participantes do discurso. Com seu auxílio, podemos identificar essas mudanças significativas, o que indica a presença de um novo segmento.

2.2.2 Estrutura Intencional

A Estrutura Intencional é a estrutura das intenções presentes no discurso à medida que este se desenvolve. Ela captura os propósitos relevantes ao discurso, expressos em cada um dos segmentos lingüísticos, bem como as relações entre eles. Seus elementos básicos são as intenções e um pequeno número de relações entre elas.

Embora os participantes de um discurso possam ter várias razões para participar dele, podemos distinguir uma delas como sendo fundamental para o discurso. A essa intenção Grosz e Sidner deram o nome de propósito do discurso (PD), que é a razão para o engajamento dos participantes conversacionais em um discurso particular. O PD nos dá tanto a razão pela qual um discurso está sendo feito quanto a razão pela qual um conteúdo particular está sendo transmitido.

Seguindo essa linha de raciocínio, Grosz (1985) também definiu o propósito do segmento de discurso (PSD) como o propósito que está por trás da criação deste segmento de discurso, a intenção que leva à iniciação deste segmento. Dessa forma, o propósito do segmento de discurso nos diz como este segmento contribui para que o propósito do

discurso seja atingido.

Tanto o propósito do segmento de discurso quanto o propósito do discurso podem ser distingüidos de outras intenções pelo fato de que se pretende que eles sejam reconhecidos. Contudo, a motivação principal do participante que inicia o discurso em começá-lo pode ser uma intenção diferente do propósito do discurso, uma intenção pessoal, cujo reconhecimento não é desejado. Naturalmente, esse reconhecimento do propósito do discurso ou do propósito do segmento de discurso é essencial para a obtenção dos efeitos desejados no discurso. Sem ele, os participantes do discurso não saberão do que estão falando nem de por que estão falando.

Mas propósitos ou intenções isoladas não são suficientes para que se possa compreender um discurso, visto que em um discurso as idéias não estão isoladas, mas sim inter-relacionadas de algum modo. A teoria que aqui é seguida (Grosz e Sidner 1985, 1986) identifica duas relações estruturais que têm um papel importante na estrutura do discurso. São elas: dominância e pré-satisfação.

Se uma ação que satisfaz uma intenção (PSD_1) pretende satisfazer parte de outra (PSD_2), então PSD_1 contribui para PSD_2 , ou, da mesma forma, PSD_2 domina PSD_1 (ou PSD_2 DOM PSD_1). Já se uma intenção (PSD_1) deve ser satisfeita antes de outra (PSD_2), dizemos que PSD_1 pré-satisfaz PSD_2 (ou PSD_1 SP PSD_2).

A relação de dominância invoca uma ordenação parcial dos propósitos do segmento de discurso. Essa ordenação é chamada hierarquia de dominância e apresenta, dentre outras coisas, um registro completo das intenções no nível do discurso e suas relações; é, portanto, uma listagem de todas as intenções surgidas no discurso juntamente com as relações entre elas.

A Figura 2.1 mostra um exemplo de discurso com suas intenções e relações entre estas. As intenções referentes às falas que indicam que uma tarefa foi executada (falas 4 e 8) foram omitidas por simplicidade. Este diálogo foi retirado (e reduzido) de Grosz e Sidner (1986), e tem como objetivo o conserto de um compressor de ar.

Nesse exemplo podemos ver as relações de dominância e pré-satisfação, com os PSDs para cada segmento indicados à direita. Assim, o segmento iniciado na fala 2 tem seu propósito do segmento de discurso (receber uma receita – o conhecimento de como fazer uma ação – para a remoção da hélice) dominado pelo PSD ao qual a fala 1 pertence (remover a hélice).

Essa relação surge pelo fato de que, ao ensinarmos alguém a executar uma ação (PSD_2), estaremos satisfazendo parte da intenção de que esse alguém execute a ação (PSD_1).

Nessa mesma linha de raciocínio vemos que o PSD_2 domina os PSD_3 , PSD_4 e PSD_7 , pois esses são passos que compõem a receita e, naturalmente, cada passo ensinado contribui para a satisfação da intenção de ensinar a receita.

As relações de pré-satisfação são observadas entre os PSD_3 e PSD_4 e PSD_4 e PSD_7 .

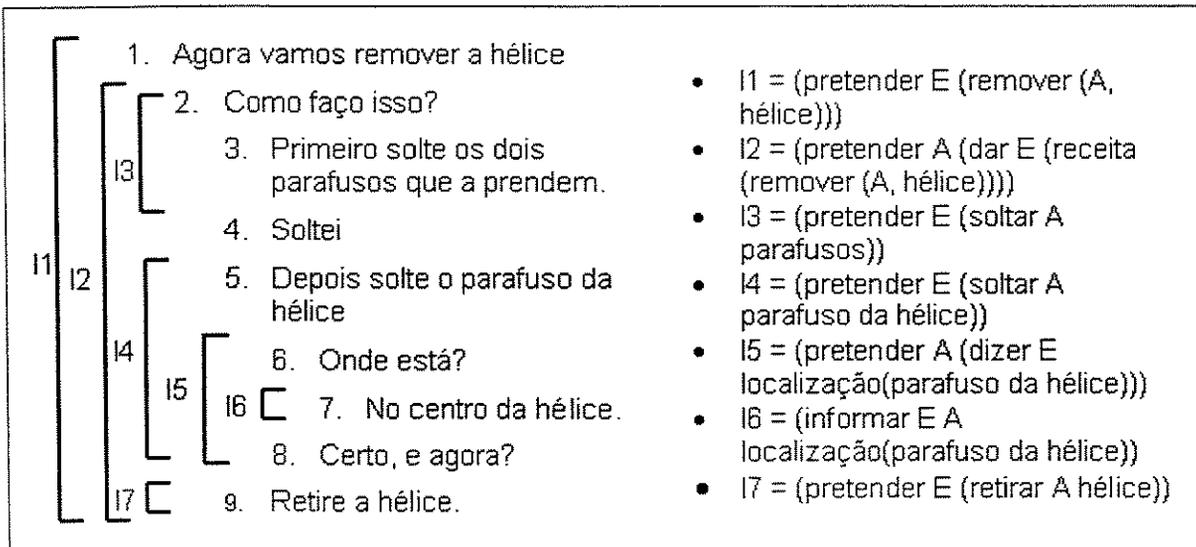


Figura 2.1: Diálogo entre um especialista (E) e um aprendiz (A). Iniciado pelo especialista.

Isso se deve ao fato de que essa receita mostra passos que claramente devem ser executados na ordem dada. Esse fato fica mais claro através do uso de palavras-pista como "primeiro" (frase 3) e "depois" (frase 5), além do pedido explícito feito pelo aprendiz para que o novo passo seja comunicado (frase 8).

Vemos, então, que para podermos determinar as relações entre duas intenções, devemos lançar mão não somente de conhecimento sobre as ações, mas também de pistas lingüísticas que a própria interatividade e necessidade de fazer-se entender em um diálogo apresentam.

2.2.3 Estado Atencional

O Estado Atencional é uma abstração do foco de atenção dos participantes do discurso à medida que este se desenvolve. Ele serve para resumir a informação de falas anteriores, crucial para o processamento das falas subseqüentes, através do registro dos objetos, propriedades e relações que estão salientes em cada ponto do discurso e, por isso, é inerentemente dinâmico.

Dessa forma o Estado Atencional fornece os meios para usar a informação nas duas outras estruturas, tanto na geração como na interpretação de falas individuais.

Seu modelamento é feito por meio de um conjunto de espaços de foco alocados em uma pilha. Ao conjunto de espaços de foco disponíveis em um dado momento damos o nome de estrutura focal. Um exemplo genérico do modelo é apresentado na Figura 2.2 (Grosz e Sidner, 1986).

O funcionamento da Pilha de Espaços de Foco pode ser descrito da seguinte maneira:

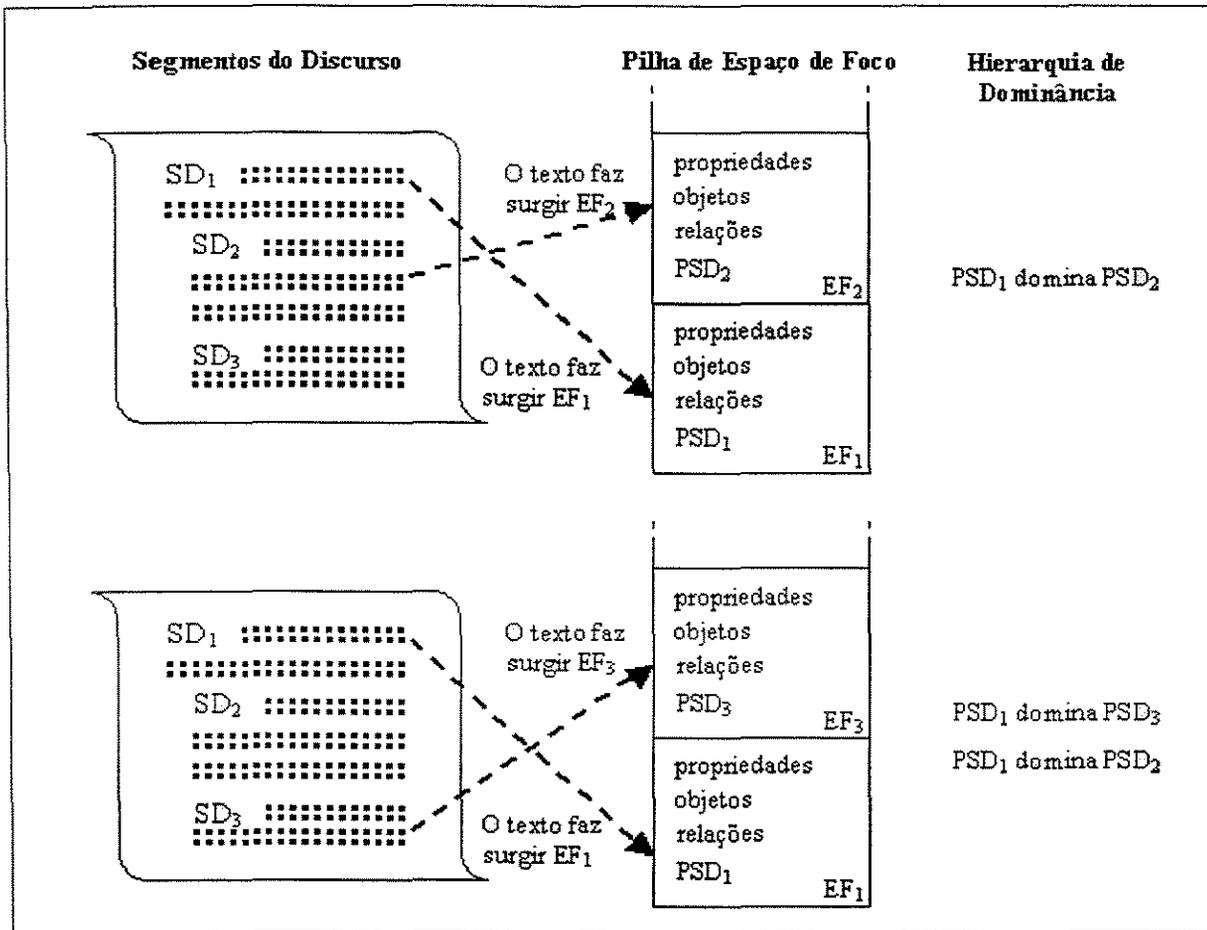


Figura 2.2: Um modelo simples da interação entre as três estruturas do discurso, mostrando como o estado atencional coordena as estruturas lingüística e intencional.

- A cada segmento de discurso é associado um espaço de foco contendo as entidades que estão salientes no segmento, porque foram explicitamente mencionadas, ou porque se tornaram salientes no processo de produção ou compreensão das falas no segmento. Além dessas entidades, o PSD também é incluído, pois os participantes conversacionais (PCs) estão focados não somente no que estão falando, mas também em por que estão falando.
- O espaço do topo corresponde ao segmento de discurso atual. Espaços mais abaixo na pilha correspondem a segmentos de discurso que envolvem (embutem) o atual. Esses segmentos foram iniciados mas não completados.
- Quanto mais alto um espaço de foco estiver na pilha, mais acessível será a informação nele contida. Nesse caso, a informação contida em um espaço no meio da pilha estará menos acessível que a informação contida nos espaços superiores, mas mais acessível

que a contida nos inferiores. Então, na parte superior da Figura 2.2 vemos que as informações no espaço de foco EF_1 são acessíveis a partir de EF_2 , mas menos salientes que as informações no próprio EF_2 , ou seja, uma busca por uma entidade primeiro seria efetuada em EF_2 , e então, caso não fosse satisfeita, seria efetuada em EF_1 .

- Um empilhamento ocorre quando o PSD para um novo segmento contribui para o PSD do segmento imediatamente precedente. Se ele contribuir para um PSD mais alto na hierarquia (e mais baixo na pilha), alguns espaços devem ser desempilhados (o suficiente para que se chegue ao espaço cujo PSD domina o PSD do espaço a ser empilhado) e, então, este espaço é empilhado. A parte superior da Figura 2.2 mostra que o segmento de discurso SD_1 já foi processado, e agora está sendo processado o segmento SD_2 . Ao chegar no processamento de SD_3 (parte inferior da Figura 2.2), vemos que seu PSD (PSD_3) é dominado somente pelo PSD do segmento SD_1 (PSD_1); então o espaço de foco para o segmento 2 (EF_2) é desempilhado e o espaço de foco para o segmento 3 (EF_3) é empilhado em seu lugar.

No exemplo da Figura 2.2 a hierarquia de dominância mostra somente relações de dominância, embora ela possa incluir relações de pré-satisfação. Foi incluída somente a relação de dominância porque, para ilustrar o funcionamento da Pilha de Espaços de Foco, a relação de pré-satisfação não tem nenhuma função, uma vez que a Pilha de Espaços de Foco somente apresenta espaços que possuem uma relação de dominância entre si (hierárquica também, já que os de baixo dominam os de cima).

Vale notar que ao final do discurso a pilha estará vazia, enquanto que a Estrutura Intencional terá sido totalmente construída se o discurso foi encerrado.

A grande importância do Estado Atencional está no fato dele restringir a interpretação de expressões referenciais diretamente. Se uma referência não puder ser encontrada no espaço atual basta buscá-la nos espaços acima na hierarquia, sendo que nenhuma referência fora da Pilha de Espaços de Foco será considerada. No caso dessa referência ser a descrição de um objeto, ela é tida como se referindo à entidade mais próxima na pilha que satisfaz a descrição, ou seja, à entidade que corresponde à descrição no mais alto espaço da Pilha de Espaços de Foco no qual houver uma entidade assim.

A estrutura de espaços de foco limita a informação que deve ser considerada no reconhecimento do propósito do segmento de discurso, bem como na identificação de referências a certas classes de sintagmas nominais definidos. Ela restringe a gama de PSDs candidatos à dominância (ou pré-satisfação) do PSD do segmento corrente, pois serão candidatos apenas os segmentos que estiverem na Pilha de Espaços de Foco.

Justamente para limitar a informação a ser considerada é que a noção de focalização foi introduzida na teoria. Focalização, segundo Grosz (1978 e 1981), é o processo no qual

se engajam os participantes em um diálogo de concentrar a atenção em, ou colocar em evidência, um subconjunto de sua realidade. É um processo ativo, pois à medida que o diálogo progride os participantes mudam sua atenção para novas entidades ou perspectivas sobre entidades previamente enfatizadas pelo diálogo. Dessa forma o que é dito influencia a focalização, e o que é focalizado influencia o que é dito.

De acordo com essa teoria, há dois níveis de focalização no discurso: o foco global e o foco imediato. Enquanto o foco global afeta a produção e interpretação de descrições definidas, o foco imediato afeta a produção e interpretação de pronomes (Gordon 1993).

Basicamente, o foco global é o conjunto de entidades relevantes ao discurso como um todo. Ele modela as propriedades do Estado Atencional no nível intersegmental (Grosz 1998) e, por sua vez, é modelado pela Pilha de Espaços de Foco. Já o foco imediato se relaciona com a identificação da entidade à qual uma fala individual mais centralmente diz respeito. Ele modela as mudanças intrasegmentais no Estado Atencional. Para modelar esse componente foi proposta a teoria da centralização.

Segundo a Teoria de Centralização (Gordon 1993, Grosz 1978, 1981, 1995, 1998 e 1999a, Joshi 1979), centros de uma fala são aquelas entidades que servem para ligar essa fala a outras no segmento de discurso que a contém. É uma fala que possui centros, e não uma sentença isolada. Uma mesma sentença dita em diferentes situações pode ter diferentes centros. Então, centros são objetos semânticos, não palavras, frases ou formas sintáticas (Grosz 1995).

Um dos pontos principais dessa teoria é a sustentação de que a cada fala (F_n) em um segmento de discurso (SD) é atribuído um conjunto de centros de busca progressiva, $C_p(F_n, SD)$ ou, para simplificar, somente $C_p(F_n)$, e, a cada fala que não a fala inicial do segmento é atribuído um único centro de busca regressiva $C_r(F_n, SD)$, simplificando, $C_r(F_n)$.

O $C_p(F_n)$ fornece as potenciais conexões com falas subseqüentes, determinando como a sentença será ligada ao discurso que a sucede. É, então, um conjunto de entidades que estão salientes na fala. Os elementos desse conjunto são ordenados segundo vários fatores, dentre eles, proeminência, papel gramatical, posição superficial e realização por pronome. Segundo Grosz (1995) e Gordon (1993), o fato de uma entidade estar no papel do sujeito gramatical contribui por si só à probabilidade desta entidade ser o mais alto membro do C_p . Um ponto interessante a ser notado é que o $C_p(F_n)$ depende somente das expressões que constituem essa fala; seus elementos não são restringidos por fatores de qualquer fala prévia no segmento.

Já o $C_r(F_n)$ serve para ligar a fala F_n ao discurso precedente, em particular, a F_{n-1} . Ele se conecta a um dos $C_p(F_{n-1})$ e assim auxilia na coerência da seqüência de falas no segmento de discurso. Dessa forma, o C_r determina como a sentença será incorporada no discurso precedente. Para ser mais exato, uma das regras de centralização proposta

originalmente estipula que o elemento de maior classificação do $C_p(F_{n-1})$ que é realizado em F_n é o $C_r(F_n)$ (Gordon 1993). Em seus experimentos, Gordon (1993) notou que "o sujeito de uma fala será seu C_r se isso for possível", não importando sua posição inicial, embora ele tenha notado também que a posição superficial é importante na ordenação dos outros membros do C_p . Ou seja, a ordem superficial afeta a ordenação dos C_p s, mas não a identificação do C_r (Grosz, 1995).

Uma vez que o fato de um determinado elemento ser sujeito contribui para que este seja o mais alto membro do C_p , e dado que o maior elemento do $C_p(F_{n-1})$ é o $C_r(F_n)$ (se realizado em F_n), a conclusão de que o sujeito de uma fala F_n tende a ser o $C_r(F_{n+1})$ é imediata. Deve ser ressaltado que nem sempre o sujeito será o C_r da própria fala, pois isso pode levar a incoerências em certos casos, como nos casos de mudança de centro (conforme será visto mais adiante).

A Figura 2.3 mostra um exemplo da extração dos centros de uma fala (extraído de Gordon, 1993). Nela, por simplicidade, os membros do C_p foram ordenados de acordo com sua posição superficial (com exceção do sujeito, que sempre foi tido como sendo o mais alto membro na hierarquia do C_p). Note que à primeira frase não foi atribuído um C_r , uma vez que ela é inicial no segmento.

Um outro ponto interessante do exemplo na Figura 2.3 é o uso dos pronomes "ela" e "a" na fala 2, que foram associados a "Susan" e "Betsy", respectivamente.

1. Susan deu a Betsy um hamster.

$$C_p = \{\text{Susan, Betsy, hamster}\}$$

2. Ela lembrou-a que esses hamsters são bastante tímidos.

$$C_r = \text{Susan}; C_p = \{\text{Susan, Betsy, hamsters}\}$$

Figura 2.3: Pequeno exemplo da extração dos centros de uma fala.

Em um diálogo geralmente a entidade focalizada muda vez ou outra, causando uma transição dos centros de busca. A teoria define três tipos de relações de transição entre pares de falas (Gordon 1993, Grosz 1995):

- Continuação do centro: $C_r(F_{n+1}) = C_r(F_n)$, sendo que esta entidade é o elemento de classificação mais alta do $C_p(F_{n+1})$. Então $C_r(F_{n+1})$ é o mais provável candidato a $C_r(F_{n+2})$.
- Retenção do centro: $C_r(F_{n+1}) = C_r(F_n)$, não sendo esta entidade o elemento de classificação mais alta do $C_p(F_{n+1})$, ou seja, é improvável que seja o $C_r(F_{n+2})$.
- Mudança do centro: $C_r(F_{n+1}) \neq (C_r(F_n))$.

A Figura 2.4 (Grosz, 1995) ilustra essas três relações com um diálogo exemplo, onde o C_r é primeiro continuado, depois retido e finalmente mudado.

- | |
|--|
| <p>a. John está tendo bastante trabalho para planejar suas férias.
 $C_r = \text{John}; C_p = \{\text{John}\}$</p> <p>b. Ele não consegue achar ninguém que tome conta de tudo.
 $C_r = \text{John}; C_p = \{\text{John}\}$</p> <p>c. Ele chamou Mike ontem para ajudar a planejar.
 $C_r = \text{John}; C_p = \{\text{John}, \text{Mike}\}$ (Continua)</p> <p>d. Mike o perturbou um bocado recentemente.
 $C_r = \text{John}; C_p = \{\text{Mike}, \text{John}\}$ (Retém)</p> <p>e. Ele chamou John às 5 da manhã na sexta passada.
 $C_r = \text{Mike}; C_p = \{\text{Mike}, \text{John}\}$ (muda)</p> |
|--|

Figura 2.4: Narrativa exemplificando as operações com os centros de busca.

Além dessas relações, também são definidas duas regras com relação à movimentação dos centros e sua realização. A primeira regra diz respeito à realização dos centros por pronomes e é assim enunciada:

Se qualquer elemento do $C_p(F_n)$ for referenciado por um pronome em F_{n+1} , então o $C_r(F_{n+1})$ deve ser referenciado por pronome também.

Essa regra estipula que nenhum elemento em uma fala pode ser referenciado por um pronome a menos que o C_r da fala seja referenciado por pronome também (Grosz 1995). O argumento intuitivo para essa regra é que pronomes são um mecanismo lingüístico para indicar continuidade e coerência (Gordon 1993) e, uma vez que o C_r também traz coerência com relação à fala anterior, é improvável que outra entidade seja uma candidata mais forte a ser referenciada por pronome.

Ainda segundo Gordon, entidades altamente focalizadas tendem a ser referenciadas por pronomes, elipses etc, enquanto que as menos focalizadas tendem a ser realizadas com descrições definidas. As entidades focalizadas na fala F_n pertencem ao $C_p(F_n)$, de onde concluímos que os elementos mais baixos do C_p não podem ser pronominalizados a menos que os mais altos também o sejam. Os experimentos de Gordon também levaram à conclusão que o C_r de uma frase é preferencialmente referenciado por pronome pois, caso contrário, observa-se um aumento nas inferências necessárias para compreender o diálogo. A lógica dessa conclusão é que, uma vez que o C_r é o elemento central da frase, ele está altamente focalizado, tendendo, assim, a ser referenciado por pronome.

Isso nos dá uma visão de como se dá o referenciamento de pronomes nessa teoria de discurso. Se um pronome é encontrado, ele é preferencialmente atribuído ao C_r da fala. Se isso não for possível, tentamos atribuí-lo aos elementos do C_p da fala anterior (de acordo com a hierarquia) e, se novamente não for possível, buscamos nos outros espaços de foco da Pilha de Espaços de Foco por uma possível referência.

A segunda regra apresentada pela teoria diz respeito à coerência de um diálogo. Ela é assim enunciada:

Seqüências de continuação são preferidas a seqüências de retenção que, por sua vez, são preferidas a seqüências de deslocamento.

Em outras palavras, transições de continuação são mais coerentes que transições de retenção, e transições de retenção são mais coerentes que transições de mudança.

Então, se houver a necessidade de uma mudança de centro, que ela ocorra por meio de retenções para que, assim, a transição para o novo centro seja suave. Essa regra é especialmente útil na geração de falas. Na compreensão não se deve esperar que isso sempre ocorra uma vez que, em diálogos, também é muito comum ocorrerem falhas na coerência, o que faz com que um dos participantes interrompa o discurso e, usando um sub-diálogo, busque trazer de volta a coerência entre as idéias.

A Figura 2.5 (Grosz, 1995) mostra duas narrativas, uma claramente centrada em "João", e outra que fica mudando o foco de atenção entre várias entidades diferentes.

• João foi à sua loja musical favorita comprar um piano.	• João foi à sua loja musical favorita comprar um piano.
• Ele freqüentou a loja por muitos anos.	• Era uma loja que João havia freqüentado por anos.
• Ele estava feliz por finalmente poder comprar um piano.	• Ele estava feliz por finalmente poder comprar um piano.
• Ele chegou bem na hora da loja fechar.	• Ela estava fechando bem na hora em que ele chegou.

Figura 2.5: Dois diálogos, um mostrando uma mudança constante de foco (à direita) e outro (esquerda) mantendo um único foco.

Podemos notar que o segundo texto é mais difícil de se entender, embora passe a mesma mensagem que o primeiro. Essa mudança constante de foco deixa mais incoerente o segundo discurso, embora não necessariamente ininteligível, o que está de acordo com a segunda regra apresentada.

Então, podemos dizer que a estrutura atencional é o repositório central para a informação contextual necessária para processar as falas em um dado ponto no discurso. Ela contém os objetos, propriedades e relações mais salientes nesse ponto e também contém ligações com as partes da Estrutura Lingüística e Atencional que são relevantes.

A próxima seção apresentará o tratamento de interrupções segundo a teoria de Grosz e Sidner (1985, 1986), de grande importância para o tratamento de diálogos. Além de aparecerem freqüentemente em diálogos, as interrupções são um importante teste para qualquer teoria de processamento de discurso, por significarem uma mudança no fluxo do discurso.

2.3 Interrupções

Basicamente, interrupções são pedaços de discurso que quebram o fluxo do discurso precedente (Grosz e Sidner 1986) sendo que, após sua conclusão, o foco do discurso volta ao segmento interrompido. Esse retorno é, na verdade, um efeito do progresso normal de uma conversação.

Para poder explicar melhor as interrupções, Grosz e Sidner (1985, 1986) apresentaram duas definições para interrupções: a definição forte, usada para interrupções verdadeiras e digressões; e a definição fraca, usada para *flashbacks*.

Definição forte: uma interrupção é um segmento de discurso cujo PSD não é dominado nem pré-satisfeito pelo PSD de nenhum segmento precedente.

Definição fraca: uma interrupção é um segmento de discurso cujo PSD não é dominado nem pré-satisfeito pelo PSD do segmento imediatamente precedente.

A teoria de Grosz e Sidner, conforme mencionado, distingue três tipos de interrupções, verdadeiras, digressões e *flashbacks*, embora afirme que isso não elimina os possíveis tipos de interrupções que podem surgir num discurso. Cada um deles será tratado com mais detalhes a seguir.

2.3.1 Interrupções Verdadeiras

As interrupções verdadeiras são interrupções que seguem a definição forte. A Figura 2.6 (Grosz e Sidner, 1986) mostra um exemplo de interrupção verdadeira e o seu modelamento usando o Estado Atencional.

Nesse exemplo, como em toda interrupção verdadeira, há duas estruturas de PD/PSD para as falas na seqüência, uma para D_1 e outra para D_2 . Uma vez que D_2 muda totalmente a atenção para um novo propósito (podendo mudar também os ouvintes pretendidos), o narrador não pode usar expressões referenciais em D_2 que dependam da acessibilidade de entidades em D_1 . Então o contorno entre D_1 e D_2 deve ser impenetrável (o que, na Figura

2.6 é mostrado como uma linha interseccionada por várias linhas diagonais) de modo a prevenir que entidades nos espaços mais baixos estejam disponíveis para os espaços acima desse contorno.

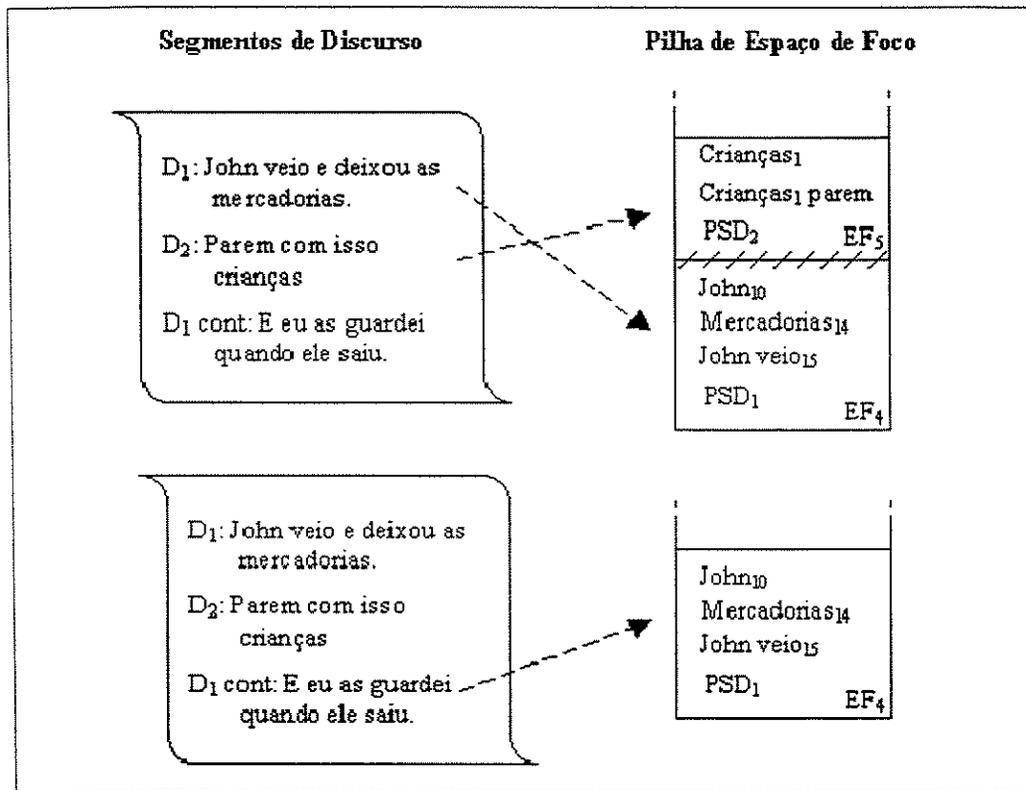


Figura 2.6: Exemplo de interrupção verdadeira.

Quando a interrupção terminar, esse "lacre" é retirado e o discurso prosseguirá normalmente.

Vários fatores indicam o término da interrupção, dentre eles a mudança de público alvo (o retorno ao público original), pistas lingüísticas (como "agora voltando a...") etc. Naturalmente, um modo de ver que uma interrupção terminou é usando a Estrutura Intencional. Se, de repente, um novo segmento é introduzido e seu PSD não tem relação alguma com o PSD do segmento anterior pertencente à interrupção (ou seja, este segmento age como uma interrupção da interrupção), basta verificar se ele apresenta alguma relação com o segmento anteriormente interrompido. Se isso acontecer, é um sinal de que a interrupção acabou, devendo seus espaços de foco serem desempilhados e o contorno impenetrável retirado para que se possa empilhar esse novo espaço de foco.

2.3.2 Digressões

Uma digressão é definida como uma interrupção forte (ou seja, que segue a definição forte) que contém uma referência a alguma entidade que está saliente tanto na interrupção quanto no segmento interrompido. Dessa forma, uma entidade qualquer do segmento interrompido permanece saliente no discurso enquanto que o propósito do discurso muda. Normalmente as digressões começam com frases do tipo "falando de...", "isso me lembra..." etc, indicando que haverá uma mudança de assunto mas ainda centrado na entidade que foi por último discutida.

Para processar esse tipo de interrupção agimos como nas interrupções verdadeiras. Uma vez que o PSD para a digressão forma a base de uma estrutura intencional separada (como nas interrupções verdadeiras), um novo espaço de foco é empilhado (como na Figura 2.6, apresentada anteriormente), contendo pelo menos uma entidade do espaço de foco do segmento interrompido. Naturalmente sua ligação com o segmento interrompido também é inviabilizada, como na interrupção verdadeira.

Seu término também é sinalizado praticamente da mesma forma que a interrupção verdadeira mas, neste caso, ele é menos brusco, geralmente feito por meio de uma fala explícita como "agora voltando a..." ou "mas, de qualquer modo...".

2.3.3 Flashbacks

Um *flashback* é um segmento cujo PSD pré-satisfaz o PSD do segmento interrompido e é dominado pelo PSD de algum outro segmento da Pilha de Espaços de Foco. Ele ocorre quando esquecemos de mencionar algo e temos que fazer uma interrupção para fornecer a informação necessária que está faltando. O *flashback* segue a definição fraca de interrupção e seu PSD mantém alguma relação com o propósito do discurso. Tipicamente, esses segmentos contêm PSDs adicionais que devem ser satisfeitos antes do PSD atual (do segmento interrompido).

Seus indicadores lingüísticos geralmente incluem um comentário sobre algo estar errado (como "opa, esqueci...").

A Figura 2.7 mostra um exemplo retirado de Grosz e Sidner (1986) de como um *flashback* pode ser modelado. Nesse exemplo podemos ver (como mostra a pilha principal) que a conversa iniciou sobre um tema A, passando depois a B, C e finalmente chegando em "Bill". Mas nesse ponto um problema surgiu: quem estava falando esqueceu que, para falar de "Bill", teria de falar sobre ABC. Considerando-se o PSD de ABC (PSD_{ABC}) como sendo dominado pelo PSD de B (PSD_B), o que se deve fazer é retirar os espaços de foco indo de EF_{Bill} até EF_C da pilha e colocá-los em uma pilha auxiliar, empilhando o espaço de foco para ABC na pilha principal, acima de EF_B .

Então o que fazemos é empilhar o espaço de foco para o *flashback* após um número

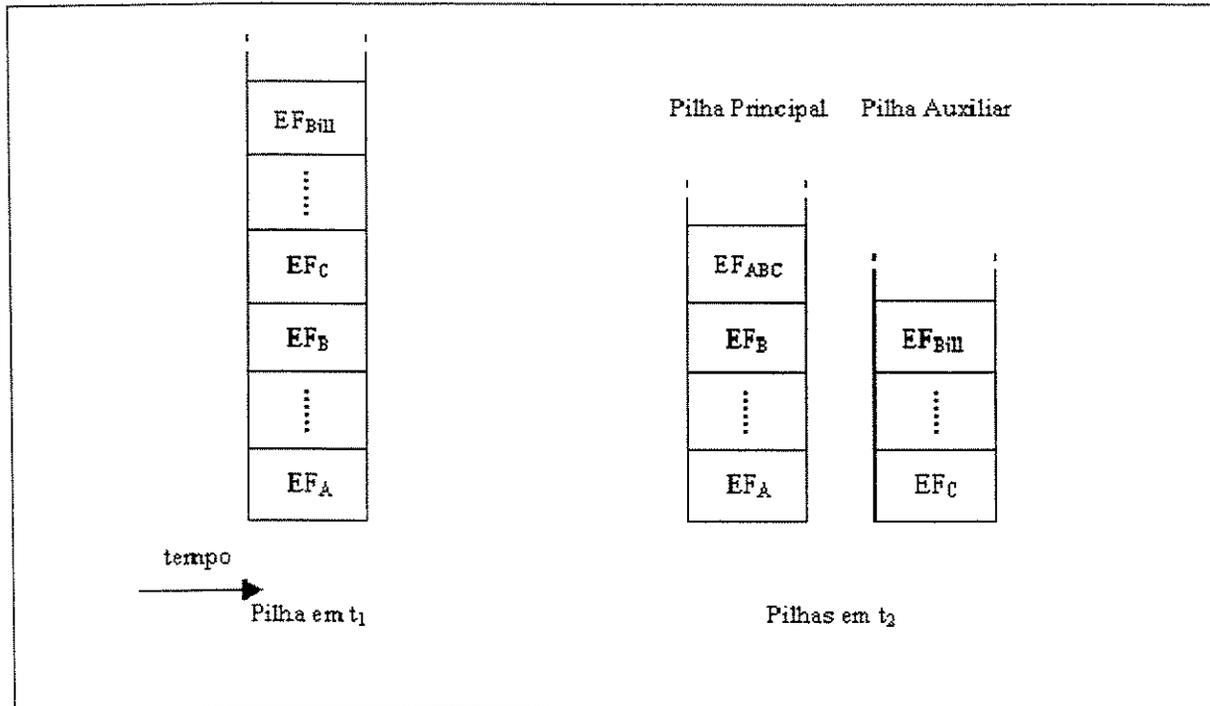


Figura 2.7: Modelamento de um *flashback* usando uma pilha de espaços de foco auxiliar.

apropriado de espaços de foco, incluindo o espaço interrompido, ter sido retirado da pilha principal e empilhado na pilha auxiliar. É importante notar que a pilha auxiliar fica inacessível a partir da principal, uma vez que estas são estruturas diferentes.

Após uma indicação explícita de retorno (como "agora voltando a...") os espaços na pilha principal decorrentes do *flashback* são desempilhados e os espaços na pilha auxiliar são devolvidos à pilha principal.

As relações que existem entre o PSD_{ABC}, o PSD_{Bill}, o PSD_B e o PD são as seguintes:

PD DOM PSD_{Bill}
 PD DOM PSD_{ABC}
 PD DOM PSD_B
 PSD_{ABC} SP PSD_{Bill}
 PSD_B DOM PSD_{Bill}
 PSD_B DOM PSD_{ABC}

Com base nessas relações um outro modelo torna-se possível: o segmento para o *flashback* poderia ser tratado como um segmento embutido normal, como na Figura 2.8. Neste modelo, o espaço de foco para ABC (EF_{ABC}) é empilhado sobre o espaço de foco para Bill (EF_{Bill}) e todas as entidades em ambos espaços de foco são normalmente acessíveis para referência.

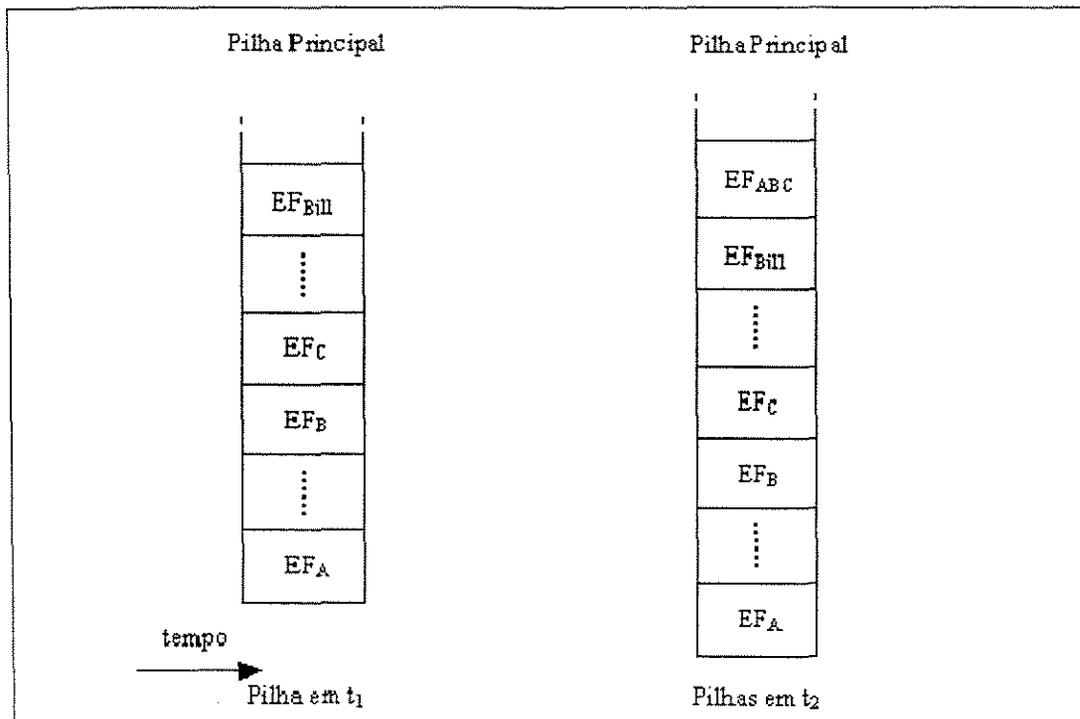


Figura 2.8: Modelamento de um *flashback* tratando-o como um segmento embutido normal.

A maior diferença entre esses dois modelos é que o segundo permite que as entidades relevantes para o material interrompido sejam acessíveis durante a interrupção enquanto que no primeiro elas não o são.

A decisão sobre qual modelo é melhor, devendo portanto ser utilizado quando do tratamento de um discurso ainda depende de um estudo mais detalhado deste fenômeno lingüístico.

2.4 Retornos Semânticos

Um retorno semântico ocorre quando entidades que estiveram salientes durante um discurso passado são reintroduzidas explicitamente. Portanto, o retorno semântico não interrompe nada, apenas "apela para a memória" para poder se referir a alguém, alguma coisa ou algum fato de alguma conversa passada.

A Figura 2.9 mostra um exemplo onde duas pessoas (Júlio e Marcos) conversam sobre fazer compras. A fala 2 representa um retorno semântico onde Marcos, ao buscar em sua memória algo que ajude Júlio, lembra de determinada camisa que Júlio havia comentado que queria e trouxe essa entidade, trazendo implicitamente outras que a acompanhavam

no momento da conversa original, para a discussão atual.

1. Júlio: Estou com vontade de ir ao shopping. Preciso mesmo de camisas novas.
2. Marcos: Camisas? Ei, lembra aquela camisa muito bonita que você me falou que queria no mês passado mas estava muito cara?
3. Júlio: Claro.
4. Marcos: Está com 50% de desconto na loja X do shopping.

Figura 2.9: Diálogo exemplificando o uso de um retorno semântico.

Esse tipo de ação (trazer entidades salientes de um discurso passado para o discurso atual) envolve o empilhamento de um novo espaço de foco contendo, entre outras coisas, representações das entidades reintroduzidas. De forma alguma trata-se de um retorno a um espaço prévio na Pilha de Espaços de Foco, uma vez que o espaço do discurso anterior não mais faz parte da pilha.

No Capítulo IV será visto em maiores detalhes o modelamento computacional da teoria que foi apresentada aqui nesse capítulo. Nele veremos como cada estrutura foi modelada, que novas estruturas foram criadas e por que.

Nesta dissertação somente a interrupção verdadeira foi tratada. As demais interrupções e a digressão não foram modeladas no sistema, uma vez que demandavam um certo tempo para que fossem melhor estudadas e posteriormente implementadas, tempo este que estava além dos limites dessa dissertação.

Capítulo III

Sistemas Multiagentes e Colaboração

Nesse capítulo serão apresentadas algumas características dos sistemas multiagentes. O capítulo começará definindo um único agente, passando então à definição de um sistema multiagentes e explicando características básicas de um comportamento colaborativo e apresentando justificativas de por que modelar um diálogo sob essa perspectiva. Ao final do capítulo é feito um resumo sobre Planos Compartilhados e seu papel no processamento de um discurso.

3.1 Introdução

Uma vez que o sistema desenvolvido neste projeto é um sistema multiagentes, antes de passar à descrição do modelamento e implementação do sistema faz-se necessária a introdução de conceitos básicos acerca de agentes, sistemas multiagentes e a ação conjunta dos agentes, conforme será apresentado a seguir.

3.2 Agentes

Um agente, de acordo com Russell e Norvig (1995), é qualquer coisa que possa perceber seu ambiente por meio de sensores e agir sobre este ambiente por meio de executores. Alternativamente, é um sistema que tenta preencher um conjunto de objetivos num ambiente complexo e dinâmico (Maes, 1995). Se esse sistema consegue operar com completa autonomia, decidir por si só como relacionar os dados obtidos com ações de modo que seus objetivos sejam atingidos com sucesso, ele é chamado de agente autônomo. Da mesma forma, se ele for capaz de melhorar com o tempo, de se adaptar, melhorando suas estruturas internas, é chamado de adaptativo.

Os agentes projetados no sistema descrito nessa dissertação não são agentes adaptativos. Com relação à sua autonomia, eles a possuem em um determinado grau, mas com

certeza não são totalmente autônomos, uma vez que para processar as informações por eles obtidas necessitam de informações adicionais providas por outros agentes.

3.2.1 Características

Muitas arquiteturas para agentes propostas têm características em comum (Maes, 1995):

- **Módulos Orientados à Tarefa:** Na Inteligência Artificial tradicional, os módulos se dividem de acordo com sua função: percepção, execução, planejamento etc, e são desenvolvidos independentemente. Agentes são conjuntos de módulos de competência, cada um responsável por uma pequena competência orientada à tarefa, comunicando-se uns com os outros por mensagens simples em vez de ter uma representação comum.
- **Soluções Específicas à Tarefa:** Tradicionalmente, os diferentes componentes do sistema são modelados o mais geral e independente do domínio possível. A esperança é que os mesmos componentes funcionais possam ser usados em diferentes domínios. Em multiagentes, não há um módulo geral ou independente da tarefa. Cada agente é responsável por toda a representação, computação e execução necessárias para sua função particular.
- **O Papel das Representações é Pouco Enfatizado:** Tradicionalmente, um agente tem um modelo interno completo, isto é, uma cópia perfeita do ambiente dentro do sistema ao qual pode se referir para resolver seus problemas. Já em multiagentes, não há uma representação central compartilhada pelos módulos; cada módulo representa localmente o que precisa para obter sua competência. As representações podem ser até inconsistentes com relação às de outro agente.
- **Estrutura de Controle Descentralizada:** A Inteligência Artificial tradicional adota uma organização seqüencial dos diferentes módulos dentro do sistema. Já as arquiteturas baseadas em agentes são altamente distribuídas e descentralizadas. Todos os módulos componentes operam em paralelo. Contudo, algum método simples de arbitragem é incluído para coordenar todo o processo.
- **A Atividade Direcionada ao Objetivo é uma Propriedade Emergente:** Em domínios baseados em agentes, a atividade é modelada como uma propriedade emergente da interação entre os módulos de competência e entre os módulos e o ambiente. Não há um processo de "pensar" como na inteligência artificial tradicional.

3.3 Sistemas Multiagentes

Um sistema multiagentes é um sistema em que, como o nome já diz, vários agentes trabalham em conjunto. Baseia-se na premissa de que as estruturas internas que controlam um agente não precisam ser complexas para produzir um comportamento resultante complexo (Maes, 1995).

Tal premissa tem origem na observação do comportamento de sociedades de insetos. Sobre isso, Wilson (1971) observa:

O inseto social individual, em comparação com o inseto individual celibatário, apresenta modelos de comportamento que não são nem excepcionalmente engenhosos nem excepcionalmente complexos. As qualidades notáveis da vida social são fenômenos de massa que emergem da interação destes modelos individuais simples por meio de comunicação.

Dessa forma, um sistema multiagentes tem como característica a cooperação e comunicação entre módulos; assim, o sucesso de um grupo de agentes pode ser maior – individualmente e coletivamente – se eles comunicarem suas crenças e objetivos uns aos outros (Russell 1995).

O uso de multiagentes para modelar um sistema de processamento de uma língua natural é decorrente dessas características dos sistemas multiagentes e da própria teoria usada como base e apresentada no Capítulo 2, a qual tem pelo menos uma grande característica em comum com os sistemas multiagentes: é composta por módulos distintos porém interagentes. Além disso, tais módulos são parte de uma grande estrutura, que também pode ser considerada como um grande agente, que interage com o usuário do programa.

Nesse contexto, o sistema aqui desenvolvido pode ser visto como um grande agente que deve participar de um diálogo com um agente humano. Mas tal agente (como o humano), por sua vez, é composto de vários agentes "órgãos", por assim dizer, independentes e com funções diferentes mas que, agindo conjuntamente, garantem o funcionamento do agente maior.

A seguir será discutida colaboração, sua definição e importância não só em sistemas multiagentes, mas também no próprio diálogo orientado à tarefa.

3.4 Atividade Colaborativa

A definição de colaboração vem do próprio termo em latim *colaborare*, que significa trabalhar junto a, trabalhar em conjunto. Ou seja, a colaboração para a obtenção de um objetivo nada mais é que trabalhar em conjunto em prol deste objetivo.

Nessa mesma linha de raciocínio, um sistema colaborativo é um sistema que busca trabalhar em conjunto com outros sistemas, com o usuário ou seu ambiente, não servindo de escravo, mas sim de parceiro na obtenção de um determinado objetivo, ou seja, a relação mestre-servo, que é atualmente típica de interações humano-computador, não é apropriada em sistemas colaborativos (Grosz, 1994).

3.4.1 Características de Sistemas Colaborativos

Sistemas colaborativos possuem inúmeras características, dentre as quais podemos citar (Grosz 1994):

- **Comprometimento com o Resultado:** Em um sistema colaborativo, cada membro deve estar comprometido com o resultado final da tarefa executada. Ou seja, ele deve fazer todo o possível para que a execução dessa tarefa tenha sucesso.
- **Comprometimento com a Atividade Conjunta:** Além de se comprometerem com o resultado, os participantes de uma atividade colaborativa devem ter um comprometimento mútuo de auxílio, para que o objetivo seja atingido com sucesso. Ou seja, não basta querer executar a ação, deve-se também permitir que os outros membros façam sua parte. Dessa forma, por exemplo, um participante não usaria uma ferramenta que outro iria precisar a menos que fosse realmente necessário. Para tal, cada membro deve, de alguma forma, ter crenças sobre a capacidade, conhecimento e planos dos outros membros.
- **Planos Colaborativos não são Simplesmente a Soma dos Planos Individuais:** Se dois membros de um sistema colaborativo colaboram para a execução de uma tarefa, cada um deles forma dentro de si planos de como executar essa tarefa. Ou seja, forma planos individuais. Mas planos individuais não são suficientes para executar a tarefa, eles precisam também de algum plano compartilhado pelos membros para mutuamente executar a tarefa, algo que garanta o sucesso na execução dessa tarefa caso cada membro execute sua parte. Por exemplo, imagine que duas pessoas (Carlos e Pedro) têm como tarefa fazer uma torre de blocos vermelhos e verdes intercalados (exemplo adaptado de Grosz 1994 e mostrado na Figura 3.1). Carlos possui somente blocos vermelhos e Pedro possui somente verdes. O plano de Carlos é empilhar seus blocos vermelhos intercalados com algum outro tipo de bloco, que ele não possui. O plano de Pedro é idêntico ao de Carlos, salvo que os blocos são verdes. Agora, basta olhar a Figura 3.1 para notar que a soma dos dois planos não constitui o plano original, sequer ajuda no sucesso da execução da tarefa. Para que os blocos possam ser empilhados, os participantes devem antes criar um plano conjunto de execução, por exemplo, determinar quem coloca o primeiro

bloco e fazer com que cada um, ao colocar seu bloco espere que o segundo coloque o seu também, para então prosseguir com seu plano individual. Tal plano conjunto é necessário para nortear as ações dos participantes e garantir que a tarefa que era objetivo da interação seja executada com sucesso.

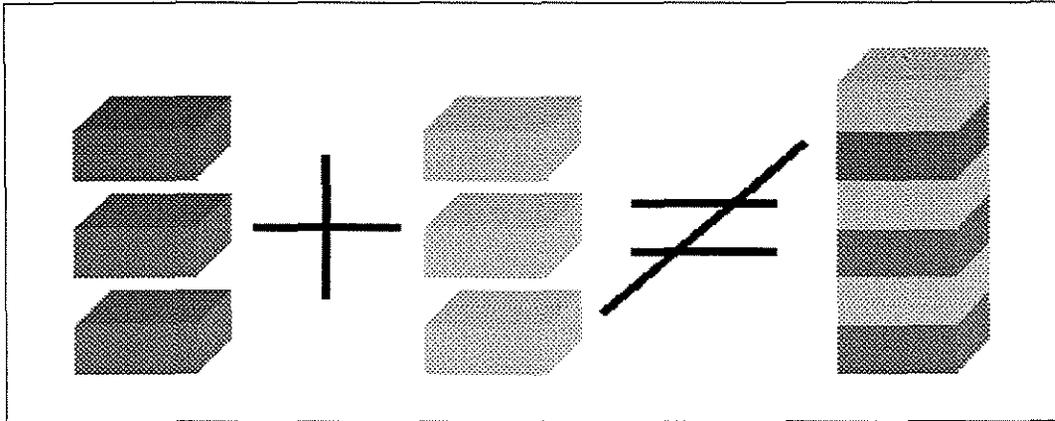


Figura 3.1: Planos colaborativos não são simplesmente a soma de planos individuais.

Dessa forma, vemos que em uma atividade colaborativa o grupo precisa formar um plano conjunto de como executar a tarefa. Para tal ele deve concordar sobre quais são as ações subsidiárias que compõem essa tarefa e quem executará cada ação subsidiária. Também deve ter confiança que o subgrupo designado para executar uma ação subsidiária pode e deseja fazer tal ação e, então, deve se comprometer com o sucesso desse subgrupo, auxiliando-o se necessário. Como afirmam Russel e Norvig (1995), em domínios multiagentes é importante para um agente raciocinar acerca dos "processos mentais" dos outros agentes, pois conhecendo as crenças dos outros, um agente pode decidir a melhor maneira de efetuar uma tarefa.

Naturalmente, além disso, cada elemento do subgrupo deve possuir um plano compartilhado com os companheiros do subgrupo, nos moldes do plano que possui com o resto do sistema, para a execução da ação subsidiária a eles incumbida. E cada elemento, além desse plano compartilhado, deve possuir planos individuais de como executar sua parte, o que exige que ele conheça um modo de executar sua parte, que possua a habilidade necessária para fazê-lo e que, naturalmente, deseje executá-la. Assim, em um sistema colaborativo é importante que os participantes concordem com o plano de domínio que está sendo construído, com as ações tomadas para desenvolvê-lo e com as crenças que norteiam essa construção (Carberry, 1993).

Uma outra característica da atividade colaborativa é a comunicação. Sem ela não há como o grupo chegar a um consenso sobre como executar uma tarefa ou resolver os possíveis conflitos.

A existência de conflitos também é uma característica interessante de sistemas colaborativos. Quando engajados num processo de planejamento conjunto, cada agente envolvido contribui com diferentes conhecimentos sobre a tarefa que podem afetar a validade do plano resultante (Chu-carroll 1995a); dessa forma, gerenciamento de conflitos, comunicação e negociação são importantes componentes da atividade colaborativa multiagente (Chu-Carroll, 1995b).

Esse gerenciamento de conflitos se mostra importante no processamento de uma língua natural para a geração de respostas ao usuário, uma vez que, para que uma resposta do sistema pareça colaborativa, o sistema deve ser capaz de detectar esses conflitos assim que eles surgem e começar a negociação com o outro agente (humano) para resolver os conflitos relevantes (Chu-Carroll, 1996). Além disso, em vez de rejeitar indiscriminadamente propostas que o sistema não tenha razões suficientes para aceitar, ele deve compartilhar seu conhecimento com o usuário e solicitar dele informações relevantes (sob a forma de um sub-diálogo) de modo a fazer com que ambos os agentes (sistema e usuário) efetivamente reavaliem a proposta e cheguem à melhor decisão (Chu-Carroll, 1995c).

Essas características são o que diferencia a colaboração de uma simples interação. Como afirma Grosz (1999b):

Colaboração se distingue de outros tipos de interações multiagentes (como cooperação, coordenação e competição) pela existência de um objetivo em comum e pelo comprometimento dos agentes com este objetivo.

Embora colaborar implique interagir, uma pura e simples interação não é colaborativa por não possuir as características citadas acima. Assim, por exemplo, o trânsito em uma grande cidade é altamente interativo, mas nada colaborativo. As vantagens da colaboração, em oposição à mera interação, podem ser vistas ao considerarmos o que aconteceria se algo desse errado no sistema (Grosz, 1994). Ou seja, os problemas que surgem podem ser atacados em conjunto, se um dos participantes não puder fazê-lo sozinho e, caso não haja meios de resolvê-los, tal fato pode ser comunicado aos outros participantes, economizando tempo e energia.

3.4.2 A Atividade Colaborativa em um Discurso Orientado à Tarefa

Conforme afirmam Grosz e Sidner (1990):

Discursos são fundamentalmente exemplos de comportamento colaborativo. Os participantes em um discurso trabalham juntos para satisfazer várias das suas necessidades conjuntas.

Especificamente, em diálogos de consulta orientados à tarefa (como é o caso do sistema desenvolvido), o usuário e o sistema constroem juntos um plano para atingir o objetivo do usuário (Chu-Carroll, 1993) e trabalham nele. Ou seja, eles colaboram para que o objetivo do usuário seja atingido com sucesso. Dessa forma, o modelamento de um sistema para interagir com o usuário na forma de um diálogo em uma língua natural deve passar pelo modelamento da atividade colaborativa.

O sistema aqui desenvolvido possui todas essas características. Mais adiante será explicado detalhadamente o modelamento do sistema. Nele, a colaboração com o usuário existe na medida que, ao ser iniciado o diálogo, o sistema cria um plano parcial para a execução da tarefa que serve de objetivo a esse diálogo. À medida que o diálogo se desenrola, o sistema vai completando esse plano, até que a tarefa principal do diálogo seja executada. Para tal, ele entende cada frase como tendo o objetivo de aproximar mais os agentes da satisfação do propósito principal do discurso, ou seja, a frase é entendida de acordo com sua participação na obtenção do resultado final. Esse plano é o modelo de colaboração entre o sistema e o usuário, pois, ao identificar o propósito que está por trás do diálogo, o sistema começa a trabalhar junto com o usuário para atingir tal propósito (propósito este dado pelo usuário). Dessa forma o sistema auxilia o usuário em sua tarefa ao assumir tal tarefa como sua também e considerar o usuário um parceiro.

3.4.3 Planos Compartilhados

A teoria de Planos Compartilhados, originalmente desenvolvida por Grosz e Sidner (1990), foi criada numa tentativa de fornecer um tratamento adequado do comportamento colaborativo exibido em diálogos (Grosz, 1996), sendo assim um modelo formal de planejamento colaborativo, um modelo de colaboração.

O modelo de Planos Compartilhados adota uma visão de planos baseada em estados mentais: os agentes têm planos quando têm um conjunto particular de intenções e crenças (Grosz, 1996).

Um plano, segundo Lochbaum (1994), é a coleção de atitudes mentais que um agente, ou grupo de agentes deve ter para agir com sucesso. Nesse contexto, os agentes envolvidos em um plano comum para executar alguma ação devem ter crenças individuais e mútuas sobre como esta ação e suas sub-ações constituintes devem ser executadas (Grosz, 1999b). Essas crenças são especificadas em termos de receitas. Uma receita para executar alguma ação α é um conjunto de ações e restrições tais que a execução destas ações sob tais restrições constituem a execução de α (Grosz, 1999b), ou seja, as receitas representam o que os agentes sabem quando conhecem um modo de executar uma ação (Lochbaum, 1994).

Ter um plano não é meramente saber como executar uma ação, mas também ter

a intenção de executar as ações agregadas a esta (Grosz, 1999c). A teoria de Planos Compartilhados distingue entre dois tipos de planos; parciais e completos, que podem ser individuais ou compartilhados.

Planos completos são aqueles nos quais todas as crenças e intenções requisitadas foram estabelecidas (Grosz, 1999c). Já um plano parcial, é aquele no qual somente parte dessas crenças e intenções está estabelecida. Embora planos completos sejam o verdadeiro objetivo do planejamento, planos parciais são os que melhor caracterizam o estado mental dos participantes em uma atividade conjunta (Grosz, 1999c). Isso porque o processo de planejamento colaborativo é um processo de refinamento; a descrição de um plano parcial é modificada no curso do planejamento pelos múltiplos agentes envolvidos na colaboração (Grosz, 1996), ou seja, na maior parte do tempo, os participantes de uma atividade conhecem apenas parte do plano para a execução dessa atividade. À medida que a interação se desenvolve, novos fatores surgem para completar tal plano.

Um plano individual é um plano, conforme o nome diz, formado por um único agente, não necessitando, portanto ser conhecido por outros agentes. De acordo com Lochbaum (1994), para um agente G ter um plano individual para um ato α ele deve satisfazer quatro requisitos:

1. G deve saber como executar α , isto é, deve ter uma receita para o ato;
2. G deve crer que pode executar os sub-atos da receita de α ;
3. G deve pretender executar os sub-atos; e
4. G deve ter um plano individual subsidiário para cada um dos sub-atos.

Um plano compartilhado – plano construído por grupos de agentes colaborativos (Grosz, 1996) – difere de um plano individual por requerer que o conjunto de agentes tenha uma crença mútua dos requisitos acima descritos (Lochbaum, 1994). Além disso, os planos subsidiários de um plano compartilhado podem ser individuais ou compartilhados, dependendo se são formados por um único agente dentro do grupo ou por um subgrupo de agentes. O grupo de agentes deve também crer mutuamente que o agente de cada sub-ato tem um plano para este, mas somente o agente executor precisa ter crenças específicas sobre os detalhes deste plano. Dessa forma, segundo Grosz (1999b), são alguns dos elementos essenciais dos Planos Compartilhados:

- Comprometimento com a atividade conjunta;
- Comprometimento com as ações dos outros agentes em serviço da atividade conjunta;

- Comprometimento com um processo conjunto para a seleção de uma receita para a atividade conjunta; e
- Comprometimento com um processo conjunto para adicionar agentes ou subgrupos às ações constituintes.

Com base nisso, é possível então redefinir plano parcial e completo: um agente tem um plano completo (individual ou compartilhado) quando cumpriu todos os requisitos acima descritos. Quando um agente satisfaz somente um subconjunto deles, ele possui um plano parcial (Lochbaum, 1994).

Então, dada esta caracterização de planos compartilhados completos e parciais, podemos dizer que um conjunto de agentes tem um Plano Compartilhado para α se eles têm um Plano Compartilhado Completo para α , ou têm um Plano Compartilhado Parcial para α e um Plano Compartilhado para completá-lo (Lochbaum, 1993). As Figuras 3.2 e 3.3 (Grosz, 1996) apresentam as definições de Plano Compartilhado Parcial (PCP) e de Plano Compartilhado Completo (PCC) e ilustram as características aqui descritas.

PCP($P, GR, \alpha, T_p, T_a, C_a$)

0. O grupo GR tem crença mútua que todos os membros do grupo estão comprometidos com o sucesso do grupo em executar α .
1. O grupo GR mutuamente crê que há uma receita para α , mas essa receita pode ser parcial; ou seja, eles podem somente ter identificado algumas das sub-ações que devem ser executadas. Eles têm um PCC para completar sua receita parcial.

Para cada sub-ação β_i na receita parcial, um de (2) a (4) vale:

2. Caso núcleo:

- (a) **Sub-ação monoagente:** Um membro do grupo G_k pretende fazer a sub-ação, pode ter somente um plano parcial para tal.
- (b) **Sub-ação multiagente:** Um subgrupo GR_k tem um plano compartilhado para fazer a sub-ação, este plano pode ser parcial.

3. Caso de contrato:

- (a) **Sub-ação monoagente:** O grupo decidiu subcontratar um agente externo G_e para fazer a sub-ação. Pode ter somente um plano parcial (individual ou compartilhado) para executar a contratação.
- (b) **Sub-ação multiagente:** O grupo decidiu subcontratar um grupo externo GR_e para executar a sub-ação. Pode ter somente um plano parcial (individual ou compartilhado) para executar a contratação.

4. **Caso não conciliado:** GR não deliberou acerca da sub-ação; nenhuma decisão foi tomada sobre que agente (ou agentes) irá executá-la.

Figura 3.2: Descrição de um PCP.

<p>PCC(P, GR, α, T_p, T_a, C_a)</p> <ol style="list-style-type: none"> 0. O grupo GR tem crença mútua que todos os membros do grupo estão comprometidos com o sucesso do grupo em executar α. 1. O grupo GR tem crença mútua dos atos (β_i) que ele precisa executar para obter α e das suas restrições (ρ_j). <p>Para cada β_i ou (2) ou (3) vale:</p> <ol style="list-style-type: none"> 2. Caso núcleo: <ol style="list-style-type: none"> (a) A sub-ação β_i é uma ação monoagente: Algum membro do grupo executará a sub-ação. (b) A sub-ação β_i é uma ação multiagente: Algum subgrupo fará a sub-ação. 3. Caso de contrato: <ol style="list-style-type: none"> (a) A sub-ação β_i é uma ação monoagente: O grupo fará com que outro agente, G_e, execute a ação. (b) A sub-ação β_i é uma ação multiagente: O grupo fará com que outro grupo de agentes execute a ação.

Figura 3.3: Descrição de um PCC.

Para obter as características acima citadas, Grosz e Sidner definem uma coleção de operadores lógicos que buscam modelar formalmente tais características. Tal coleção não será mostrada aqui por estar fora do escopo deste trabalho, mas pode ser encontrada descrita em detalhes em Grosz (1996, 1999b e 1999c).

3.4.4 Papel dos Planos Compartilhados no Processamento do Discurso

De acordo com Lochbaum (1994), com relação ao modelo dos Planos Compartilhados, os agentes engajados em um discurso são tidos como colaborando na obtenção de algum objetivo. As falas são, então, entendidas em termos da sua contribuição à colaboração dos agentes.

Até o ponto no qual todas as atitudes mentais necessárias para o sucesso da colaboração foram estabelecidas, os agentes têm um Plano Compartilhado Parcial. Este plano serve para delinear a informação que os agentes devem considerar ao interpretar as falas dos outros agentes e ao determinar o que eles próprios devem dizer ou fazer. Dessa forma, as falas dos participantes do discurso são entendidas como incrementando o Plano Compartilhado Parcial que representa o estado da sua colaboração. O processo de incrementação do plano é mostrado nas Figuras 3.4 e 3.5 (Lochbaum, 1994), que assumem que dois agentes G_1 e G_2 estão colaborando em um ato α e modelam o raciocínio de G_1 nesse sentido.

A Figura 3.4 apresenta o processo de geração de novas falas, mostrando como o Plano Compartilhado Parcial serve para restringir a gama de informações que o agente G_1 deve considerar ao formular suas falas. No curso de sua colaboração, os agentes devem decidir sobre uma receita para α (item 1 da figura), devem dividir a execução dos sub-atos entre eles (itens 2 e 3), e devem formar planos individuais ou compartilhados para esses atos (itens 2a e 3a). Estes requisitos formam a base para as falas de G_1 e são mantidos em sua agenda. A agenda de G_1 indica as crenças e intenções requeridas para que os agentes tenham um Plano Compartilhado completo para α , mas que estão ausentes do plano parcial corrente. Com base em sua agenda, G_1 escolhe um item para o qual direcionar sua atenção, decide o que ele quer dizer sobre esse item, e diz (item 2b). No item 3, G_1 espera que G_2 responda se concorda ou não com a informação comunicada por G_1 . Então G_1 atualiza o Plano Compartilhado Parcial dos agentes para refletir a informação que G_2 comunicou.

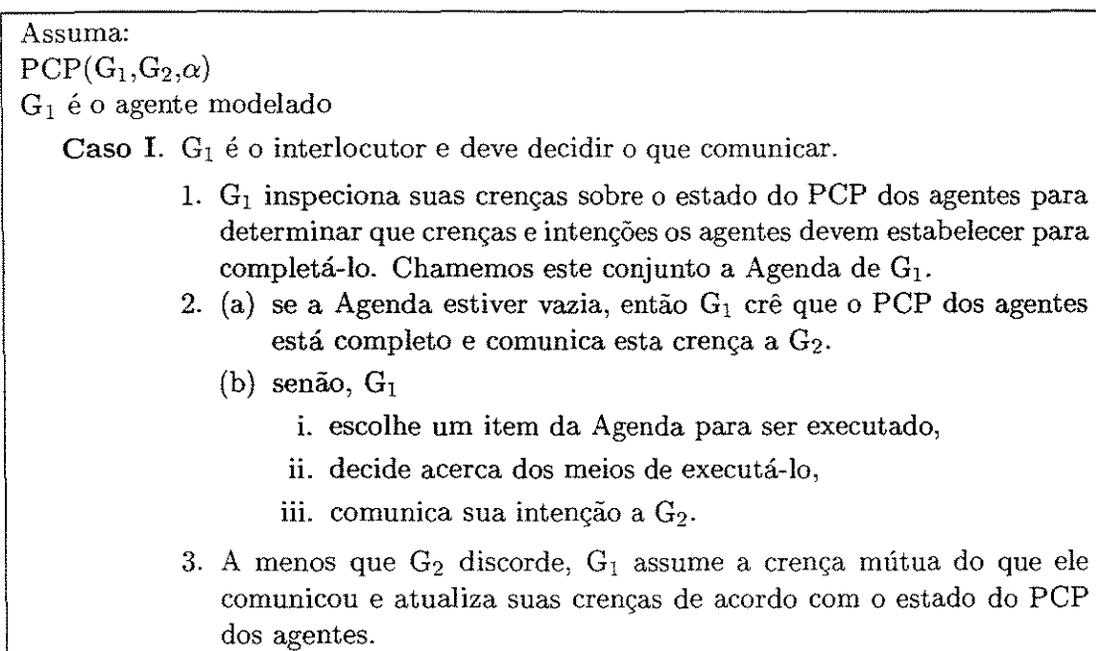


Figura 3.4: Processo de Incrementação do PC – Geração.

Na Figura 3.5 é mostrado o processo de interpretação de falas. Neste processo, para entender uma fala de G_2 , G_1 deve determinar a contribuição desta para o Plano Compartilhado Parcial dos agentes.

A figura assume que G_2 recém comunicou uma fala a G_1 contendo o conteúdo proposicional *prop*. Assim, *prop* pode ser interpretado de três modos:

1. Pode indicar o início de um Plano Compartilhado subsidiário (caso (a) do passo 5);

2. Assinalar o completamento do Plano Compartilhado atual (caso (b)); ou
3. Contribuir para esse plano (caso (c)).

Após determinar a contribuição de *prop* para o Plano Compartilhado Parcial dos agentes, G_1 deve atualizar esse plano, para refletir a fala de G_2 (passo 6a). Caso G_1 não consiga determinar tal contribuição, ele avisa G_2 deste fato (passo 6b).

Assuma:
 $PCP(G_1, G_2, \alpha)$
 G_1 é o agente modelado

Caso II. G_1 é o ouvinte e deve interpretar a fala de G_2 .
 Seja *Prop* a proposição comunicada pela fala de G_2 .

4. Como resultado da comunicação, G_1 assume crença mútua que G_2 acredita em *Prop*.
5. G_1 deve então determinar a relação de *Prop* com o contexto do Plano Compartilhado atual:
 - (a) Se G_1 crê que *Prop* indica o início de um Plano Compartilhado subsidiário para um ato β , então G_1 irá:
 - i. Atribuir a G_2 uma intenção que ambos os agentes formem um PCC para β ,
 - ii. Determinar se ele também quer adotar tal intenção.
 - (b) Se G_1 crê que *Prop* indica o completamento do Plano Compartilhado atual, então G_1 irá:
 - i. Atribuir a G_2 uma crença que ambos os agentes formaram um PCC para α ,
 - ii. Determinar se ele também crê que o Plano Compartilhado atual dos agentes está completo.
 - (c) De qualquer outra maneira, G_1 irá:
 - i. Atribuir a G_2 uma crença de que *Prop* é relevante para o Plano Compartilhado atual dos agentes,
 - ii. Determinar se ele também crê ser este o caso.
6. (a) Se o passo (5) tiver sucesso, então G_1 assinalará sua concordância e assumirá crença mútua da relação inferida em (5a), (5b), ou (5c), como for apropriado.
- (b) Senão, G_1 irá interrogar G_2 e/ou comunicar sua discordância.

Figura 3.5: Processo de Incrementação do PC – Interpretação.

A grande contribuição do trabalho de Lochbaum (1994) foi o modelamento da Estrutura Intencional do discurso através do uso de Planos Compartilhados. Tal modelamento

é feito com base no comportamento da Estrutura Intencional descrito no Capítulo 2. Conforme já mencionado, a Estrutura Intencional é composta por intenções e relações entre estas, sendo que a cada segmento de discurso é associado um propósito (Propósito do Segmento de discurso - PSD) que contribui para a determinação do Propósito do Discurso (PD). As relações que surgem entre os segmentos de discurso são, conforme já visto, de pré-satisfação e dominância.

Dessa forma, cada segmento de discurso é modelado através de um Plano Compartilhado. O propósito do segmento é tido como sendo a intenção que os participantes do discurso formem esse plano, intenção esta mantida pelo agente que inicia o segmento. Assim, para verificar se uma fala inicia um segmento, completa o atual ou contribui para este, basta ver a contribuição do conteúdo proposicional desta fala para o Plano Compartilhado correspondente a este segmento de discurso (passo 5 da Figura 3.5).

No próximo capítulo serão apresentados maiores detalhes sobre o trabalho realizado por Lochbaum no modelamento da Estrutura Intencional, além de ser mostrado o modelamento teórico das outras estruturas.

Capítulo IV

Modelamento do Sistema Desenvolvido

Neste capítulo será detalhado o modelamento teórico do sistema computacional desenvolvido. Inicialmente será dada uma visão geral do sistema e de suas partes principais e, então, cada uma de suas partes (correspondentes às estruturas definidas no capítulo 2 mais a Estrutura da Tarefa, que será explicada mais adiante) terá seu modelamento discutido mais detalhadamente.

4.1 Visão Geral do Sistema

O sistema desenvolvido é composto por quatro módulos independentes, porém interagentes (Figura 4.1):

- *Estrutura da Tarefa*: responsável pelo gerenciamento do dicionário e da biblioteca de receitas (planos pré-definidos de como executar uma determinada tarefa), servindo como um banco de dados do sistema.
- *Estrutura Atencional*: responsável pelo foco da atenção dos participantes do diálogo, armazenando informação sobre as entidades e intenções mais salientes durante a conversação.
- *Estrutura Lingüística*: responsável pela tradução das frases do usuário em língua natural para uma forma lógica que o sistema seja capaz de processar.
- *Estrutura Intencional*: é um mapa das intenções que estão por trás de cada fala no discurso, bem como das relações entre elas. É a responsável pela compreensão e manipulação dessas intenções.

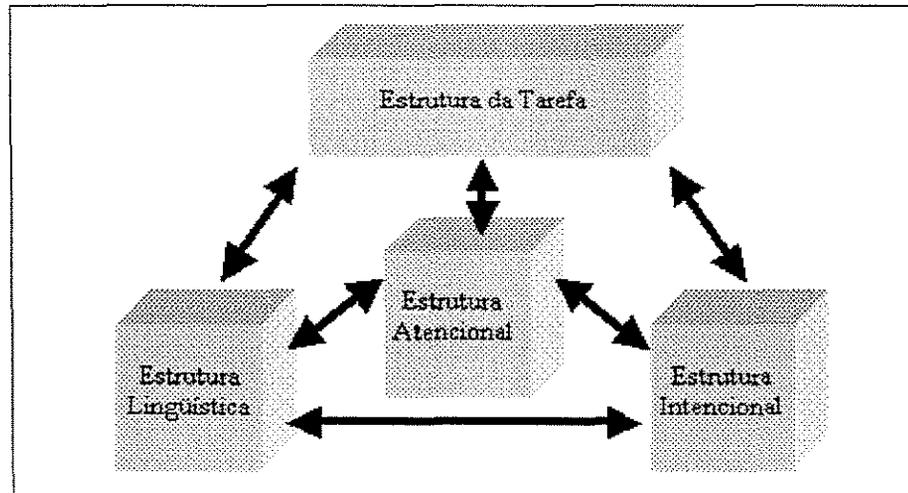


Figura 4.1: Esquema geral do sistema.

Em linhas gerais o sistema funciona da seguinte maneira (a figura 5.2 do próximo capítulo apresenta um exemplo mais detalhado deste comportamento):

- O usuário digita uma nova fala do discurso.
- Essa fala é analisada pela Estrutura Lingüística, que a transforma em uma frase lógica. Para tal, a Estrutura Lingüística acessa a Estrutura da Tarefa para saber o significado das palavras, e a Estrutura Atencional, para saber quais entidades são candidatas ao referenciamento de pronomes e elipses. Nesse ponto, a Estrutura Atencional pode vir a acessar a Estrutura da Tarefa também, para auxiliar nesse referenciamento. Ao final do processamento da fala do usuário, a Estrutura Lingüística atualiza os centros de busca na Estrutura Atencional, enviando as entidades salientes na fala recém processada.
- A seguir, a frase lógica completa é enviada à Estrutura Intencional. Esta, por sua vez, com o auxílio da Estrutura da Tarefa e da Estrutura Atencional, verifica qual o papel dessa fala com relação às suas predecessoras no discurso. Ao determinar tal relação, a Estrutura Intencional atualiza a Pilha de Espaços de Foco na Estrutura Atencional (o que pode acontecer por meio de desempilhamentos e empilhamentos de espaços de foco, ou de simples atualização do propósito do segmento de discurso) e dá a resposta apropriada ao usuário.

Dessa forma, ainda que em linhas gerais, pode-se verificar o funcionamento do sistema. Naturalmente, seu processamento é muito mais complexo. Nessa descrição foi omitido praticamente todo o processamento interno de cada estrutura, enfatizando-se como a

interação entre as estruturas afeta o processamento final do diálogo. Os detalhes da construção e funcionamento de cada estrutura serão apresentados no Capítulo 5.

4.2 Estrutura Lingüística

Conforme mencionado, essa estrutura é a responsável pela tradução efetiva da frase do usuário em língua natural para uma forma lógica capaz de ser processada pelo sistema. Para tal, ela necessita de um analisador sintático, além de mecanismos que permitam sua interação com as outras estruturas do sistema.

A interação da Estrutura Lingüística com as outras estruturas já foi mostrada na Seção 4.1. Agora a atenção será voltada para o modelamento interno da Estrutura Lingüística, mais especificamente, de seu analisador sintático.

Uma das aspirações desse sistema é que ele deixe o usuário o mais livre possível para formular sua fala durante o diálogo. Para tal não é aceitável que o sistema possua somente um conjunto de "formas válidas" de se dizer algo; ele deve ser o mais flexível possível.

Essa foi uma das razões pelas quais não foi usada uma gramática para comandar as ações do analisador sintático. Por exemplo, na Figura 4.2 é mostrada uma gramática simples para o português. É visível que ela fica presa à ordem sintagma nominal (SN), seguido de um sintagma verbal (SV). Mas tal ordem nem sempre é respeitada. Da mesma forma que alguém pode dizer "João vai ao parque", também pode dizer "Vai ao parque João". Essa é uma frase perfeitamente válida em português, não ambígua, embora rara. Mas as possibilidades não terminam por aí; ainda restam outras quatro possibilidades de dizer a mesma coisa, todas válidas: "João ao parque vai", "Vai João ao parque", "Ao parque João vai" e "Ao parque vai João".

$S \rightarrow SN SV$
$SN \rightarrow \epsilon \mid \text{nome-próprio} \mid \text{substantivo}$
$SV \rightarrow SV SN \mid \text{verbo} \mid SV SP \mid SV \text{advérbio}$
$SP \rightarrow \text{preposição SN}$

Figura 4.2: Gramática simples para o português.

O trabalho que daria prever cada possibilidade com regras em uma gramática é inimaginável. Isso porque a gramática apresenta a característica inata de estar vinculada a uma determinada ordem.

As dificuldades encontradas para criar uma gramática abrangente para a língua portuguesa levaram a um novo modo de abordar o problema, conforme será descrito a seguir.

4.2.1 Modelamento Proposto

O modelamento que aqui será apresentado tem origem em uma característica da língua latina: sua quase total independência de ordem. Em latim, devido às declinações características de cada caso, os elementos de uma frase podem ser permutados quase que livremente. Só não são totalmente livres porque, com o passar dos anos, algumas declinações assumiram terminações iguais (como genitivo e dativo singular e nominativo plural da primeira declinação, que de um final "a" passaram para "æ") causando uma certa ambigüidade.

Mas o ponto chave da questão é que as sub-idéias que formam uma frase (cada elemento semântico) recebem uma marca para indicar seu papel na idéia maior que a frase expressa. Então, dado que o português se origina do latim, pode-se pensar que algumas dessas características podem ainda estar presentes nele.

Dessa forma, o que o modelo proposto busca é determinar essa sub-idéias em português, postulando que, como em latim, elas também podem ser permutadas ao longo de uma oração, não com tanta liberdade quanto o latim, mas ainda assim com considerável liberdade.

Tomemos a frase "João vai ao parque", citada acima. Conforme visto, ela tem seis permutações que são válidas em português. Essas são permutações de basicamente três elementos: "João" (nome), "vai" (verbo) e "ao parque" (objeto indireto). Então, de acordo com o modelo, esses elementos seriam as três ilhas que compõem a oração.

Por ilha deve ser entendido todo conjunto de palavras com ordem fixa, mas que pode "navegar" dentro da frase com um determinado grau de liberdade, sem alterar o sentido da mesma. Por navegar entende-se permutar de lugar com outro desses conjuntos de palavras.

Uma exceção é a ilha que serve de objeto direto. Essa ilha é composta basicamente por substantivo, adjetivo e artigo. Embora o artigo seja fixo, o substantivo e o adjetivo permitem uma certa permutação de suas posições, gerando essa exceção.

O grau de liberdade das ilhas em uma frase não é total. Em frases como "João viu Pedro", se invertermos e falarmos "Pedro viu João", o sentido muda; mais especificamente, a entidade que ocupa a posição de sujeito muda. Nesse caso, as ilhas seriam "João", "viu" e "Pedro". Essa ambigüidade surge porque, em português, o nominativo e o acusativo possuem a mesma forma. Agora, se em vez de "Pedro" usarmos a palavra "bandeira", aí a ambigüidade some: "João viu a bandeira", "João a bandeira viu", "A bandeira viu João", "A bandeira João viu", "Viu a bandeira João" e "Viu João a bandeira". Mas essa ambigüidade somente sumiu porque quando alguém escuta essa frase, está usando uma informação semântica a mais: bandeira é objeto inanimado, logo, não vê.

A partir dessas idéias iniciais, foi feito um estudo da língua portuguesa. Como resultado, chegou-se à conclusão de que o português possui quatro tipos de ilhas, apresentadas

a seguir.

4.2.2 Tipos de Ilhas da Língua Portuguesa

A definição dos tipos de ilhas é feita a partir do verbo. Em português duas estruturas são consideradas básicas: o verbo e o substantivo. As outras estruturas são satélites destas ou servem para unir orações, determinar o papel destas na oração, etc.

Nesse trabalho, a estrutura considerada principal é o verbo, porque ele determina a ação, a idéia central. As outras estruturas são vistas como complementares a ele, servindo para responder às questões "quem/que?", "quando?", "como?", "onde?" e "com que/quem?". Então as ilhas seriam as seguintes:

- *Objeto direto*: composta por artigo, substantivo (ou nome próprio), adjetivo e numeral – os satélites do substantivo (além do próprio substantivo). Naturalmente a ordem não necessariamente é essa (com exceção do artigo, que sempre vem no início), pois tanto podemos ter "as duas grandes casas" quanto "as duas casas grandes". Além dessas entidades, alguns pronomes podem fazer parte dessa ilha, como os pessoais, de tratamento, demonstrativos etc. Pronomes são casos à parte, pois sua função é fazer a ligação com alguma entidade já citada ou que vai ser citada.
- *Objeto indireto*: composta por preposição seguida de um objeto direto. Essa ilha existe para poder responder às questões supracitadas, e também pelo fato de que não há como separar a preposição do que se segue dela.
- *Verbo*: é a ilha principal. Um dos pontos básicos do enfoque proposto é que cada oração tem uma só ação (não necessariamente um só verbo). Se houver mais de um verbo, então eles estão se complementando, como na frase "a janela está sendo aberta".
- *Advérbio*: uma ilha não tão livre, mas ainda assim uma ilha. Se na oração houver um só verbo, então o advérbio é livre para "navegar" por ela. Já se houver mais de um verbo, ele deve estar mais próximo do verbo que ele complementa.

O modelo adotado assume que com essas quatro ilhas é possível processar uma frase em língua natural, transformando-a em lógica.

O algoritmo para tal processamento é mostrado, em linhas gerais, na Figura 4.3, sendo que detalhes da implementação deste serão apresentados no Capítulo 5.

Inicialmente, determina-se a ação a partir dos verbos na frase, e então, de acordo com essa ação, busca-se o complemento exigido pelo verbo principal. Cada complemento é representado por uma ilha. Dessa forma, quando uma ilha necessária para a complementação do verbo é encontrada, ela é transformada em lógica e adicionada à ação que o verbo

Entrada: fala do usuário em língua natural.

1. Separa as orações que constituem a frase.
2. Para cada oração:
 - (a) Classifica cada palavra da oração conforme sua classe gramatical (substantivo, adjetivo, verbo etc).
 - (b) Separa a oração nas ilhas que a formam.
 - (c) Analisa a oração da seguinte maneira:
 - i. Busca a ilha verbo.
 - ii. Se houver somente uma, então traduz sua ação correspondente a este verbo para lógica.
 - iii. Se houver duas ilhas verbo, então uma delas é de ligação (no indicativo) e a outra corresponde à ação que deverá ser traduzida para lógica, sendo que este segundo verbo está ou no particípio ou no gerúndio.
 - iv. Se houver três ilhas verbo, então duas delas são verbos de ligação e uma é a ação a ser codificada. Nesse caso, as possibilidades são as seguintes:
 - Verbo de ligação no indicativo + verbo de ligação no gerúndio + verbo de ação no particípio, como em "está sendo aberta".
 - Verbo de ligação no indicativo + verbo de ligação no particípio + verbo de ação no particípio, como "tem sido feito".Em qualquer um destes casos, a frase lógica contém a tradução do verbo de ação. Os demais verbos servem para definir a voz (ativa ou passiva) e o tempo da ação.
 - v. Verbos no infinitivo funcionam como objeto direto (são tratados como ilha objeto direto) ou mesmo sujeito, como em "fazer isso é difícil". Verbos no gerúndio que acompanhem verbos de ação, significam uma ação conjunta, como em "trabalhe sorrindo", que equivale a "trabalhe enquanto sorri". Dessa forma, monta-se a lógica da ação.
 - vi. Busca-se, então, um candidato a sujeito, que é uma ilha do tipo objeto direto que concorde com o verbo. Ela preferencialmente estará à esquerda do verbo e próxima a ele. Ao encontrar traduz para lógica essa ilha.
 - vii. Conforme o tipo do verbo, busca-se o complemento exigido (objeto direto ou indireto). Ao encontrá-lo, traduz para lógica (da mesma forma que o sujeito) e o adiciona à lógica final.
 - viii. Finalmente busca algum outro complemento, traduz este, e adiciona à frase lógica final.

Figura 4.3: Algoritmo para a tradução de uma oração em sua forma lógica.

(ou o conjunto de verbos) representa. Por fim, outros complementos, como advérbios, são buscados para complementar a ação e, se encontrados, também são traduzidos em lógicas e incorporados à representação lógica da ação.

Por exemplo, na análise da frase "João viu Pedro hoje", a primeira entidade buscada é a ilha verbo. Nesse caso ela é única, representando então a ação "ver()". Tal ação exige objeto direto e sujeito. Então busca-se a primeira ilha objeto direto. Ao encontrá-la, sua posição é verificada e, por estar à esquerda do verbo, é tomada como candidata preferencial a sujeito (caso não haja nenhuma à esquerda do verbo, a primeira que for encontrada será tida como sujeito). Nesse ponto essa ilha (no caso "João") é traduzida para lógica e incorporada à representação lógica da ação, que fica sendo "ver(JOÃO)".

No passo seguinte, uma nova ilha objeto direto é procurada, mas desta vez para servir de objeto direto. Tal ilha é encontrada ("Pedro") e, após transformada em lógica, incorporada à representação da ação: "ver(JOÃO,PEDRO)".

Por fim, outros complementos são buscados. O único encontrado é uma ilha advérbio, "hoje", que é um advérbio de tempo, o que responderia à pergunta "quando?". Tal ilha é traduzida para lógica e incorporada à representação da ação.

No final desse procedimento, a representação lógica da frase "João viu Pedro hoje" é "ver(JOÃO,PEDRO,Hoje())". Vale notar que "Hoje()" também pode receber argumentos (como um horário, ou intervalo de tempo).

Dessa forma foi obtida uma representação lógica de uma oração sem a necessidade do uso de uma gramática computacional, apenas separando a oração nas ilhas que a compõem e vendo como cada ilha contribui para a formação dessa representação lógica. Um ponto importante nisso é que, com exceção das ilhas "João" e "Pedro", todas as outras poderiam estar em qualquer ordem e, ainda assim, a lógica seria a mesma. Assim, "Hoje João viu Pedro", "João viu hoje Pedro", "João hoje viu Pedro", "Viu João hoje Pedro" resultariam na mesma representação lógica, "ver(JOÃO,PEDRO,Hoje())", desde que a ordem de "João" e "Pedro" não fosse mudada, o que mudaria as entidades nos papéis de sujeito e objeto direto.

Esse modelo, por tratar apenas de orações simples (que contêm somente uma ação), deixa de lado as orações múltiplas, isto é, frases compostas por várias orações unidas por um pronome relativo ou conjunção.

A seguir serão enumeradas as características do modelo que o tornam útil para o processamento de uma língua natural.

4.2.3 Características do modelo

Três características desse modelo o tornam bastante útil:

1. *Maior liberdade com relação à ordem:* conforme mostrado, esse sistema analisa várias entradas semanticamente iguais (porém com suas ilhas permutadas) da mesma forma e sem custo adicional de programação, gerando a mesma representação lógica. Isso é possível devido à sua forma de raciocínio não ser tão dependente da seqüência das palavras, como em uma gramática.
2. *Sem árvores de derivação:* num analisador sintático que trabalhe com uma gramática é comum, durante a análise de uma sentença, a geração de uma quantidade grande de árvores de derivação. Esse modelo, por estar integrado às outras estruturas, não gera mais do que uma possibilidade semântica, uma frase lógica. Essa frase no contexto do sistema é enviada à estrutura apropriada e, somente se considerada ilógica (de acordo com o contexto do diálogo), é descartada e nova possibilidade é procurada. Então, o modelo, por ter sido criado para trabalhar em conjunto com as outras estruturas do sistema, não se preocupa em gerar o maior número possível de interpretações, poupando bastante tempo de processamento.
3. *Mudança de ponto de vista:* nesse sistema, o foco do raciocínio é a apresentação de novas informações que, ao final da oração, a completam, diferente do foco usado por um analisador sintático baseado em gramática, que exige que a informação buscada esteja na oração e esteja na ordem prevista. Nesse modelo, nenhuma exigência é feita (excetuando-se as impostas pela gramática portuguesa, como concordância, uso correto de preposições etc) porque, em diálogos, haverá falas que têm tão somente a função de completar alguma fala anterior, ou seja, serão incompletas e não estarão erradas.

Por essas razões o sistema desenvolvido usa o modelo aqui descrito para o processamento das falas do usuário. No próximo capítulo será mostrado o desenvolvimento do protótipo e no Capítulo VI os resultados obtidos da integração desse modelo (usado na Estrutura Lingüística) com as outras estruturas do sistema.

4.3 A Estrutura Atencional

O Estado Atencional, conforme visto no Capítulo 2, é o estado do foco da atenção dos participantes do discurso, guardando as entidades que estão salientes em cada segmento de discurso, bem como o propósito do segmento de discurso.

Ele também apresenta dois focos distintos com relação ao discurso: o foco global e o imediato.

O foco global contém as informações e intenções relevantes ao discurso como um todo, trabalhando intersegmentalmente. O foco imediato, por sua vez, trata da identificação da

entidade à qual uma fala individual mais centralmente diz respeito, ocorrendo intrasegmentalmente e entre pares de falas seqüenciais em um diálogo.

O modelamento do foco global se dá por meio de uma Pilha de Espaços de Foco, pilha esta que, por ter sido caracterizada e ter seu comportamento descrito no Capítulo 2, não será aqui discutida.

O foco imediato, conforme já visto, é composto pelos centros de busca, que são as entidades responsáveis pela conexão de uma fala com suas vizinhas (a anterior e posterior). Nesse sistema ele é modelado usando-se uma pilha de três elementos, correspondentes às três últimas falas do diálogo. Esse número foi escolhido porque, uma vez que os centros de busca são entidades que servem para ligar uma determinada fala à fala anterior ou posterior, não há porque guardar mais do que três falas consecutivas do diálogo. De fato, dois elementos seriam suficientes para tal função, mas como sempre há a possibilidade de uma oração ser rejeitada pela Estrutura Intencional, toma-se três elementos para poder restaurar os dois elementos anteriores caso a terceira fala seja rejeitada.

Quando esta pilha está cheia e mais uma fala deve ser empilhada, a frase da base da pilha é retirada desta, sendo suas entidades mais relevantes enviadas à Pilha de Espaços de Foco.

4.4 A Estrutura da Tarefa

A Estrutura da Tarefa funciona como um banco de dados do sistema, servindo de provedor de informações para as outras três estruturas.

É composta por duas bibliotecas distintas: o dicionário e o livro de receitas. O dicionário é uma biblioteca que contém as palavras reconhecidas pelo sistema, bem como sua classificação sintática.

O livro de receitas, por sua vez, e como o próprio nome sugere, guarda as receitas que poderão ser usadas pelo sistema para entender as falas do usuário ou gerar suas próprias falas e ações.

Dessa forma, quando uma estrutura precisa saber a classificação sintática de alguma palavra (como o fazem as estruturas lingüística e atencional), ela consulta a Estrutura da Tarefa, e quando uma estrutura precisa de uma receita para uma ação (como a Estrutura Intencional), também busca tal informação na Estrutura da Tarefa. No Capítulo 5 esta interação será mostrada com maiores detalhes.

Essa estrutura foi projetada assim para permitir uma maior mobilidade do sistema. Uma vez que ela é a responsável por toda a informação dependente do domínio que o sistema precisa, para mudar de domínio bastaria mudar essa estrutura (ou substituí-la). Mas isso, por si só, não garante que ela seja totalmente independente do domínio. Há que se considerar outras restrições, como os meios físicos de executar as ações, e o contexto de

um diálogo. Desse modo, pode-se afirmar que essa estrutura torna mais fácil a mudança de domínio no sistema, desde que nenhuma função física nova seja incluída (como, por exemplo, a adição de uma habilidade física que o sistema não tenha, como um telefone para discar) e que a interação se dê por meio de um diálogo orientado à tarefa.

Além disso, a separação desse banco de dados em uma estrutura única serve também para facilitar a inclusão de novas informações no sistema, ou seja, o aumento do banco de dados, quer seja com palavras novas, quer seja com receitas novas.

4.5 A Estrutura Intencional

A Estrutura Intencional é a responsável no sistema pela compreensão e manipulação das intenções que estão por trás de cada fala do diálogo, bem como das que representam os segmentos de discurso. Seu papel é, a cada fala, descobrir como esta contribui para que o propósito do discurso seja atingindo, conforme será visto a seguir.

Seu modelamento computacional foi feito por Lochbaum (1994), utilizando a teoria de planos compartilhados desenvolvida por Grosz e Sidner (1990) e apresentada no Capítulo III. A seguir será apresentado o modelamento desenvolvido por Lochbaum e utilizado nesse trabalho.

4.5.1 Sistema Desenvolvido por Lochbaum

O sistema de Lochbaum apresenta duas estruturas de dados principais: um grafo de receitas e uma pilha de intenções e endereços.

Um grafo de receitas, é uma representação dos atos que os participantes do discurso planejam executar para atingir seus objetivos, ou seja, uma representação dinâmica de receitas. A Figura 4.4 mostra a representação gráfica de uma receita e do grafo de receitas (GrafoR). Vale notar que uma receita é um GrafoR com apenas um nível de decomposição, ou seja, enquanto que uma receita representa informação sobre a execução abstrata de uma ação, um GrafoR representa informação mais especializada por incluir parâmetros, agentes e tempos, bem como níveis múltiplos de decomposição (Lochbaum, 1994).

Também de acordo com Lochbaum, a construção de um GrafoR corresponde ao raciocínio que um agente executa para determinar se a execução de um ato particular faz ou não sentido dadas as crenças do agente sobre as receitas e o estado de seus planos individuais e compartilhados. O algoritmo para implementar tal construção é dado na Figura 4.5 e é baseado na suposição de que os agentes G_1 e G_2 estão colaborando para executar um ato α , modelando o raciocínio de G_1 .

Ao receber uma fala de G_2 com conteúdo proposicional β , G_1 determina se β contribui diretamente para α , verificando se β pertence a alguma receita para α . Este procedimento

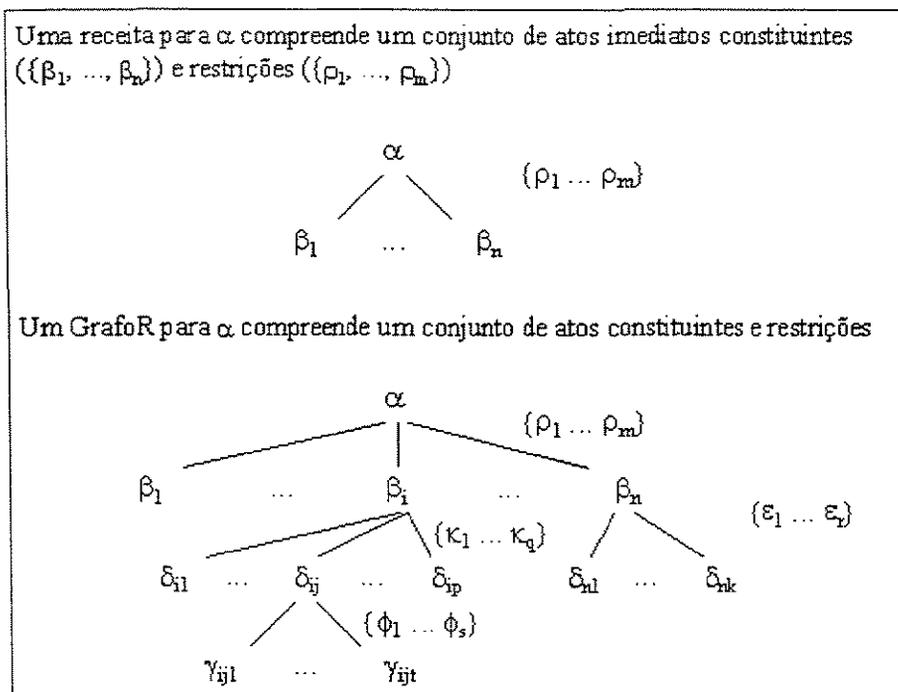


Figura 4.4: Representações de receita e GrafoR.

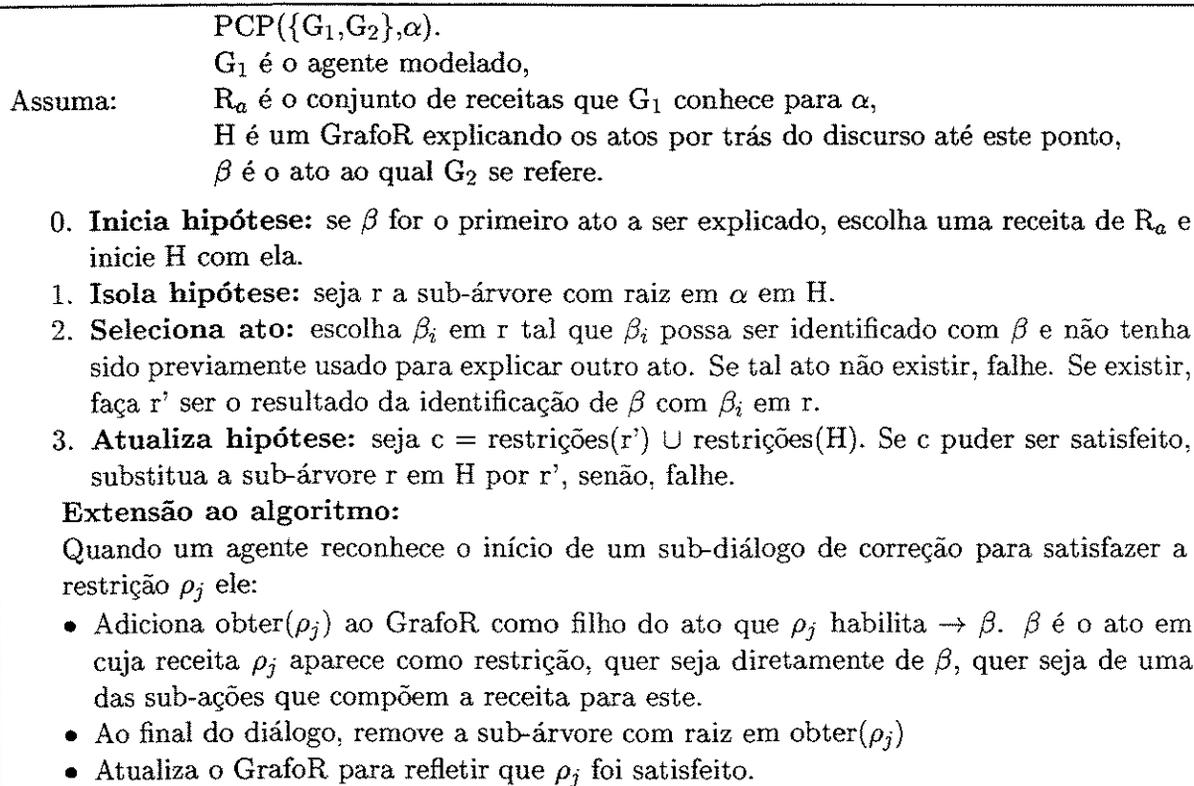


Figura 4.5: Algoritmo de construção do GrafoR.

é mostrado nos passos 1 e 2 do algoritmo. O passo 3 verifica se esta explicação de β é consistente com o resto do GrafoR, testando as restrições já existentes no grafo e as que β trouxe consigo para ver se podem ser satisfeitas pelo sistema.

A pilha de intenções e endereços, por sua vez, executa a função de uma estrutura atencional simplificada, sendo capaz somente de manter registro das intenções em foco e seus respectivos endereços no GrafoR.

Dessa forma, o sistema desenvolvido por Lochbaum é composto por quatro módulos (Figura 4.6):

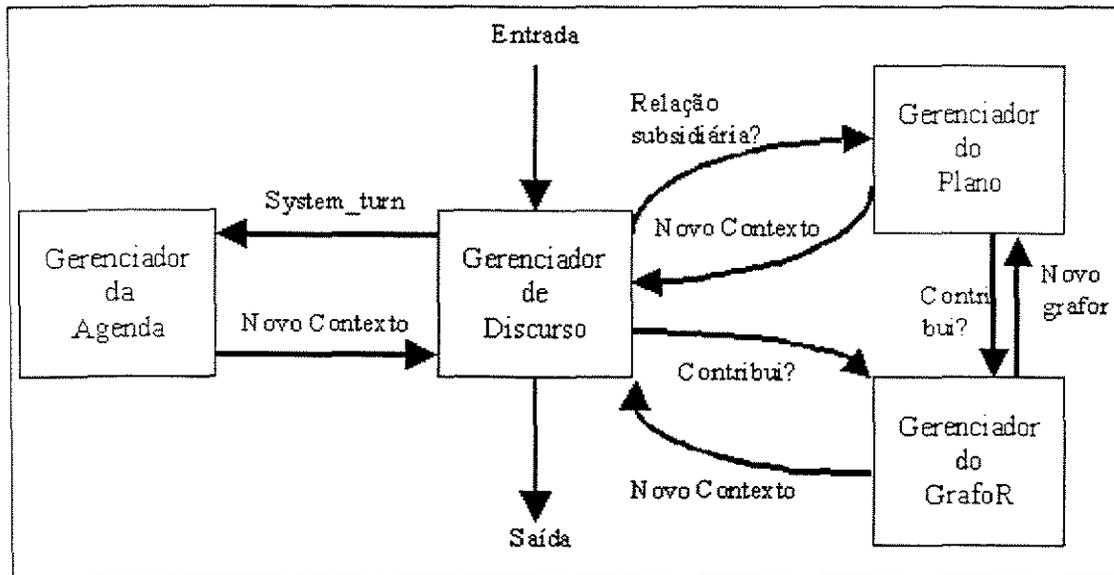


Figura 4.6: Esquema do sistema desenvolvido por Lochbaum.

- **Gerenciador de Discurso:** recebe a entrada do usuário e a envia ao componente apropriado. Conforme o formato da entrada, ele a envia a um dos componentes abaixo. A entrada do usuário segue uma coleção de 7 macros, descritas na Figura 4.7.
- **Gerenciador do Plano:** determina a existência de relações subsidiárias entre planos. Para tal recebe um ato do Gerenciador de Discurso (na forma "deseja(G, β, C)" e "deseja_p(G, β, C)") e determina, usando o Gerenciador do GrafoR, se este ato é subsidiário ao plano correntemente em foco. Se for, ele empilha o ato na pilha de intenções, indicando que este novo ato é o que está em foco.
- **Gerenciador do GrafoR:** implementa o algoritmo de construção do GrafoR. Recebe um ato β sob as restrições C do gerenciador de discurso (na forma "sugere($G,$

$\{\beta_i\}, C)$ ", "comunica($G, \{\beta_i\}, C)$ " e "comunica_p($G, P, C)$ ") e, usando o GrafoR, verifica se β contribui para o plano atual, verificando se β pertence a alguma receita para este. Caso contribua, atualiza o GrafoR.

- **Gerenciador da Agenda:** é o responsável pelas atividades que o sistema deve executar para atingir o propósito do discurso. Gera um conjunto de tarefas que o sistema deve executar no atual contexto e então as mostra ao usuário, para que este selecione uma das tarefas para o sistema executar.

deseja(G, β, C): G deseja o ato β sob as restrições C . C opcional.

Envia β e o contexto atual (GrafoR e pilha de intenções) ao Gerenciador do Plano, para determinar se uma relação subsidiária existe entre β e o plano atual, dado C .

deseja_p(G, P, C): G deseja a proposição P sob as restrições C . C opcional.

Envia obter(P) e o contexto atual ao Gerenciador do Plano, para ver se existe uma relação subsidiária entre obter(P) e o plano em foco.

sugere($G, \{\beta_i\}, C$): G sugere que os atos $\{\beta_i\}$ sejam executados sob as restrições C . C opcional.

Envia cada β_i e o contexto ao Gerenciador do GrafoR para determinar se uma relação de contribuição existe entre os β_i e o plano atual.

comunica($G, \{\beta_i\}, C$): G comunica que os atos $\{\beta_i\}$ devem ser executados sob as restrições C . C opcional.

Envia a entrada e o contexto ao Gerenciador do GrafoR para determinar se uma relação de contribuição existe entre comunica($G, \{\beta_i\}, C$) e o plano atual. Se não puder ser explicado, tenta explicar os $\{\beta_i\}$.

comunica_p(G, P, C): G comunica que a proposição P é verdadeira sob as restrições C . C opcional.

Envia a entrada e o contexto ao Gerenciador do GrafoR. Se comunica_p(G, P, C) não puder ser explicado, tenta explicar P como uma restrição dos atos no grafo.

ok: o usuário assinala sua crença de que o plano atual está completo.

Sonda a agenda e o GrafoR para ver se o plano atual está realmente completo.

system.turn: é a vez do sistema fazer ou dizer algo.

Passa o contexto ao Gerenciador da Agenda para determinar as tarefas que o sistema pode executar.

Figura 4.7: Macros e seu processamento para as possíveis entradas do usuário.

As entradas do usuário, (como mencionado) são feitas em uma lógica proposicional, seguindo uma lista pequena de proposições, ou macros (Figura 4.7). Esta entrada é

analisada pelo gerenciador de discurso, que a envia ao componente apropriado, para que seja determinada a contribuição desta fala para o propósito do discurso.

Após esta ser processada, um novo contexto é devolvido ao usuário, ou seja, a pilha de intenções e o GrafoR são atualizados e devolvidos ao usuário para inspeção. Nas ações em que cabe ao sistema tomar a iniciativa, este consulta um "oráculo", que é o próprio usuário, para decidir o que fazer. Então, quando é a vez do sistema agir, quer seja tomando iniciativa em algo ou executando uma ordem, o usuário deve entrar a macro "system_turn" e fazer as vezes de oráculo, escolhendo ações e receitas para o sistema poder, então, atualizar novamente a pilha de intenções e o GrafoR e devolvê-los ao usuário, para que este entenda que o sistema executou a tarefa.

Caso o sistema não consiga explicar a entrada do usuário, ou que não consiga satisfazer alguma restrição, ele devolve o contexto (pilha de intenções e GrafoR) ao usuário juntamente com uma explicação para a falha.

A seguir, no próximo capítulo, será explicada em detalhes a implementação do sistema, mostrando como foi feita a programação de cada uma das estruturas que o compõem.

Capítulo V

Implementação do Sistema

No texto que segue será apresentada, em detalhes, a implementação do sistema desenvolvido para testar o modelo proposto, descrevendo, para cada estrutura, suas estruturas de dados e os agentes desenvolvidos, bem como a comunicação entre eles.

5.1 Visão Geral

Como já mencionado anteriormente, para testar o modelamento das estruturas da Tarefa e Lingüística (apresentado no Capítulo 4), bem como o comportamento das quatro estruturas quando agindo em conjunto, foi desenvolvido um sistema computacional multiagentes.

O sistema foi totalmente escrito em C, sendo composto por quatro módulos independentes, porém interagentes (conforme visto no capítulo 4): Estrutura da Tarefa, Estrutura Atencional, Estrutura Lingüística e Estrutura Intencional. Além desses módulos, foi desenvolvido um agente para servir de ambiente de comunicação - o agente COM - cuja função é controlar a comunicação entre essas quatro estruturas e o usuário (Figura 5.1).

Toda a comunicação entre os agentes e entre as estruturas é feita por meio de linhas de texto pré-formatadas. O formato de cada linha é " $\{A|Msg\}$ ", onde A é o agente ou estrutura de origem, se está enviando, ou destino, se está recebendo a mensagem. Por exemplo, para que o agente G_1 envie uma mensagem, Msg, para o agente G_2 , ele envia a mensagem " $\{G_2|Msg\}$ " ao agente COM; este, por sua vez, repassa a mensagem a G_2 , modificando o campo A de destinatário para remetente. Então a mensagem que chega em G_2 é " $\{G_1|Msg\}$ " e, deste modo, quando um agente recebe uma mensagem, ele sabe quem a enviou.

A Figura 5.2 contém um exemplo do funcionamento do sistema. Uma fala do usuário é "empacotada" conforme o modelo pré-definido acima descrito e enviada à Estrutura

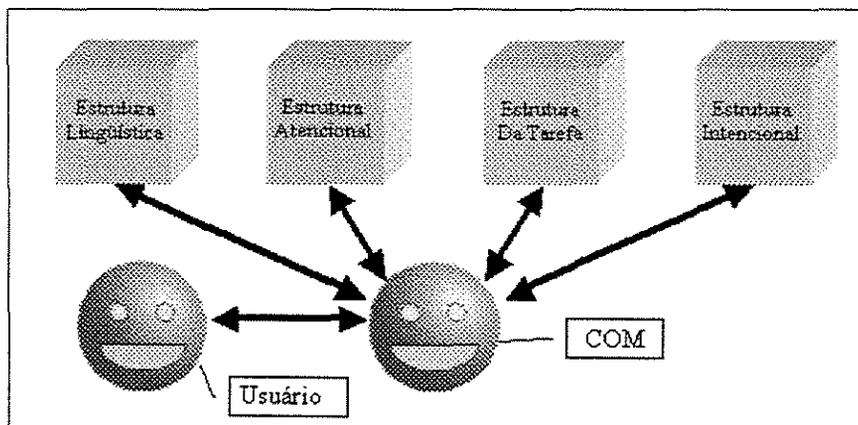


Figura 5.1: Esquema geral do sistema.

Lingüística (item 1).

```

1. O usuário enviou "{U|mostre a porta da frente}" ao EL
2. EL enviou "{L| <mostre>}" ao ET
3. ET enviou "{T|[v-m-p-3-s-d-mostrar]}" ao EL
4. EL enviou "{L| <a>}" ao ET
5. ET enviou "{T|[a-f-s][p-n-s][r-p-f-s-3-o]}" ao EL
6. EL enviou "{L| <porta>}" ao ET
7. ET enviou "{T|[s-f-s-PORTA]}" ao EL
8. EL enviou "{L| <da>}" ao ET
9. ET enviou "{T|[p-f-s]}" ao EL
10. EL enviou "{L| <frente>}" ao ET
11. ET enviou "{T|[s-f-s-FRENTE]}" ao EL
12. EL enviou "{L|ent:N@PORTA(? ,1,FRENTE(? ,1,?))|FRENTE(? ,1,?)}" ao EI
13. EI enviou "{A|F}" ao EA
14. EA enviou "{I|N}" ao EI
15. EI enviou "{T|R:mostrar(S,PORTA_FRENTE,ta)}" ao ET
16. ET enviou "{I|{det_posicao(S,PORTA_FRENTE)|ter_descr(PORTA_FRENTE)}
    {determinar_camera(S,C)}|{mostrar_imagem(S,C,ta)|ligada(C)}" ao EI
17. EI enviou "{U| <ação>}" ao usuário.

```

Figura 5.2: Mensagens entre as estruturas ao processar a fala "mostre a porta da frente".

A partir daí, as estruturas começam sua intercomunicação. A Estrutura Lingüística, por sua vez, se comunica com a Estrutura da Tarefa para saber a função sintática de cada palavra (itens 2 a 11). Conforme o conteúdo da fala do usuário, a Estrutura Lingüística pode ou não consultar a Estrutura Atencional para a resolução de pronomes e/ou elipses. Após terminar a tradução da fala do usuário para lógica, a Estrutura Lingüística envia à Estrutura Atencional as entidades mais salientes na fala analisada (item 12) e, então

envia à Estrutura Intencional o conteúdo proposicional desta fala (item 13).

A Estrutura Intencional, de posse da forma lógica, comunica-se com a Estrutura Atencional para saber que intenções estão correntemente em foco (item 13). Novamente, dependendo do conteúdo proposicional da fala do usuário, a Estrutura Intencional pode ou não consultar a Estrutura da Tarefa em busca de alguma receita para executar um ato qualquer (item 15 e 16) e, ao final do processamento, ela devolve uma ação ou uma resposta ao usuário.

Nesse protótipo, tanto a ação quanto a resposta vêm em forma escrita; a única diferença é que uma resposta vem sob a forma de uma frase em língua natural, enquanto que a ação tem uma representação única e independente da ação – "<ação>" – servindo apenas para avisar ao usuário que a ação desejada foi executada com sucesso.

5.1.1 A Forma Lógica

A codificação lógica de uma frase é feita com base em macros e funções lógicas que o sistema deve entender. Existem três tipos diferentes de codificação:

- **Ações:** codificadas como funções escritas com letras minúsculas e com campos definidos: ação(1,2,3,4,5,6,7). Os campos são:
 1. Quem: é o campo do executor da ação, o sujeito.
 2. O que: é o campo do complemento direto ou indireto, o alvo da ação.
 3. Quando: esse campo situa no tempo o alvo da ação.
 4. Como: indica modo ou instrumento.
 5. Onde: indica o local da ação.
 6. Enquanto: determina uma ação que deve ser executada paralelamente à ação principal.
 7. Tempo: é o tempo da ação. Ele difere do campo 3) por ser o instante de execução da ação e não o tempo do objeto direto ou indireto. Por exemplo: "Mostre o que aconteceu ontem às 14:00". Nessa frase, o campo 7 é o presente, enquanto que o campo 3 é "Ontem(14)". Dessa forma, a fala "eu quebrei violentamente a janela ontem às 14:00" é representada como "quebrar(U,JANELA(,),Ontem(14),Violento,,,r)". Uma vez que não foi dito onde a ação se deu nem o que o sujeito fazia (se efetivamente fazia algo) durante o ato de quebrar, os campos 5 e 6 foram deixados em branco.
- **Objetos e complementos:** são objetos físicos que preenchem, após codificados, os campos acima descritos, sendo representados por uma função com três campos e escrita em maiúscula: OBJETO(1,2,3). Seus campos são:

1. Os adjetivos do objeto, suas características. Estes são representados por uma lista de atributos escritos em maiúscula e unidos por "&". Por exemplo, "a casa branca" é representada como "CASA(BRANCA,1,)" e "a casa branca e velha" é representada como "CASA(BRANCA&VELHA,1,)"
 2. A quantidade de objetos, determinada por números. Caso não se saiba a quantidade exata, o campo é deixado em branco. Então, se deixarmos em branco o campo da quantidade na fórmula lógica acima – CASA(BRANCA&VELHA,,) – isso significaria "as casas brancas e velhas".
 3. O genitivo. Esse campo indica a quem o objeto pertence. É usado também para diferenciar objetos semelhantes; assim, "a porta da frente" se diferencia de "a porta dos fundos" pela forma lógica, que é "PORTA(,1,FRENTE(,1,))" para a primeira e "PORTA(,1,FUNDO(,,))" para a segunda. Ou seja, apesar de ambas serem PORTAS, sua forma lógica permite que o sistema diferencie as entidades através da inspeção de seus campos do genitivo. Esse tipo de codificação permite, também, o uso infinito do genitivo, ou seja, "a porta da frente da casa da mãe da amiga de..." e assim por diante, seria codificada como "PORTA(,1,CASA(,1,MÃE(,1,AMIGA(,,...)))" etc.
- **Complementos da Ação:** São representados principalmente pelos advérbios, e são codificados como macros (ou funções) em letra minúscula com a inicial maiúscula. A representação por uma função ou macro simples depende do complemento. Assim, complementos temporais (como advérbios de tempo) possuem um sub-complemento - o horário - enquanto que complementos de modo (advérbios de modo) não necessitam de sub-complementos. Então, "ontem", "ontem à meia noite" e "ontem após a meia-noite" são representados, respectivamente, por "Ontem()", "Ontem(0)" e "Ontem(>0)"; enquanto que "rapidamente" é representado pela macro "Rapido".

5.2 Estrutura Lingüística

A Estrutura Lingüística é a responsável pela tradução da fala do usuário para uma forma lógica que o sistema consiga processar. Para tal ela deve receber a frase do usuário, analisá-la e tentar obter seu conteúdo proposicional lógico.

Essa estrutura é composta por cinco agentes (Figura 5.3):

- EL: responsável pela comunicação entre os outros quatro agentes. O EL está para a Estrutura Lingüística assim como o COM está para o sistema como um todo.
- EL1: responsável pela separação da fala do usuário em sub-falas e pela análise da pontuação utilizada.

- EL2: responsável pela classificação sintática das palavras que compõem uma dessas sub-falas encontradas pelo agente EL1.
- EL3: de posse de uma das sub-falas já classificada pelo agente EL2, este agente responde pela separação dessa sub-fala em ilhas.
- EL4: responsável pela tradução efetiva da sub-fala em lógica. Após o agente EL3 separar as ilhas, ele envia um sinal ao agente EL4, que inicia uma conversação com o primeiro para formar a frase lógica.

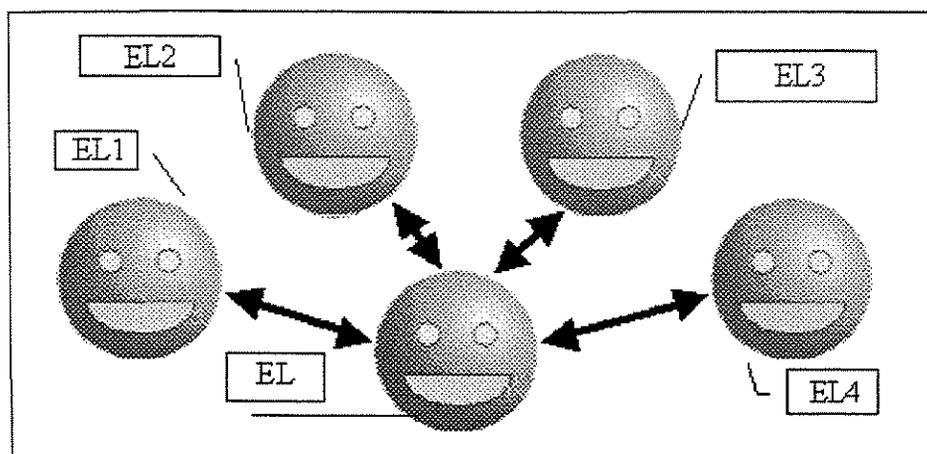


Figura 5.3: Esquema de formação da Estrutura Lingüística.

Então, ao chegar uma nova fala do usuário nessa estrutura, ela é imediatamente remetida ao agente EL1 que, após separá-la em sub-falas, remete a primeira delas ao agente EL2. Este, por sua vez, com a ajuda da Estrutura da Tarefa, classifica cada palavra da frase. Caso alguma palavra tenha mais de um significado, o EL2 faz permutações, criando frases com todos os significados possíveis. Em seguida ele envia uma dessas frases ao agente EL3, que busca e separa as ilhas dentro da frase. Caso esse agente encontre alguma falha com relação à gramática portuguesa, ele acusa tal erro ao EL2, que envia uma nova frase de sua lista de possíveis combinações sintáticas.

Após conseguir separar as ilhas de uma frase, o agente envia uma mensagem de prontidão ao agente EL4. Este, por sua vez, inicia um diálogo com o agente EL3 pedindo que determinadas ilhas sejam enviadas, caso elas existam, ou perguntando se há alguns elementos específicos na fala.

Terminada essa conversa o EL4 já terá, a esta altura, obtido uma tradução para a frase que o EL3 recebeu. Caso ele não consiga obter tal tradução, ele assinala um erro ao EL2 que, por sua vez, envia uma nova frase ao EL3, recomeçando o ciclo. Se todas as possibilidades do agente EL2 se esgotarem e nenhuma frase lógica tiver sido criada, então

é enviado um erro ao usuário, dizendo que a frase dele não pode ser compreendida. Dessa forma, quase todos os erros são detectados somente quando o agente EL4 processa a fala, uma vez que incoerências geralmente são detectadas ao tentarmos dar um sentido lógico a uma frase.

Para ilustrar o processo acima descrito, na Figura 5.4 é mostrado o processamento da fala "mostre a porta da frente", retirado de um dos arquivos de monitoramento das comunicações entre os agentes do sistema.

Conforme explicado, a fala é enviada ao agente EL1 (item 1). Como ela é composta de somente uma oração, tudo que EL1 tem a fazer é classificar a pontuação da frase (tal classificação será descrita mais adiante, quando for detalhado o funcionamento do agente EL1) e enviá-la ao EL2 (item 2).

De posse da fala enviada por EL1, o agente EL2 começa a classificar sintaticamente cada palavra, usando, para tal, a Estrutura da Tarefa (itens 3 até 12). Como algumas palavras têm mais de uma classificação sintática, EL2 faz uma lista das possíveis combinações e envia uma delas a EL3 (item 13).

O EL3, por sua vez, separa as ilhas encontradas, enviando um sinal de prontidão a EL4 (item 14). Começa então o diálogo entre eles. EL4 primeiro pede os *flags* da frase, que são a classificação dada pelo EL1 (item 15). Após recebê-los (item 16), ele pede um verbo (item 17) e, assim que recebe este verbo (item 18), passa a pedir os possíveis complementos do verbo (como advérbios, pronomes, objetos diretos e objetos indiretos). No caso da Figura 5.4, o verbo recebido é "mostrar" (item 18) em seu modo imperativo.

De posse deste verbo, EL4 busca saber se há algum advérbio de negação que possa negar a ação de mostrar (item 19), ao que o EL3 responde que não há (item 20). EL4, então, busca uma ilha contendo "que" para verificar a existência deste pronome relativo (item 21). Novamente o EL3 afirma não haver tal entidade (item 22). EL4, então, busca um pronome pessoal na terceira pessoa (item 23) e um verbo no infinitivo (item 25), numa tentativa de achar um objeto direto para o verbo. Como esta busca resultou em nada, EL4 finalmente pede uma ilha Objeto Direto (OD) (item 27) ao EL3. Tal ilha existe, e "a porta da frente" é enviada ao EL4.

Para finalizar, EL4 busca por um advérbio qualquer e um objeto indireto (OI), tentando encontrar algum complemento extra para o verbo (como, por exemplo, um advérbio de modo, de tempo, ou alguma ilha com preposição que indique como a ação foi executada). Uma vez que nesta frase eles não existem, o EL3 envia apenas negativas ao EL4. Por fim, no item 33, EL4 pergunta se há mais alguma ilha não enviada, ao que EL3 responde que não (item 34). Então EL4 abastece a Estrutura Atencional com as entidades mais salientes na frase (item 35) e envia o resultado lógico à Estrutura Intencional.

Da mesma forma, na Figura 5.5 é mostrado o processamento da fala "a porta está aberta". A diferença deste diálogo para o anterior é que a primeira análise sintática

Fala: "A porta está aberta".

1. COM enviou "{d|mostre a porta da frente}" a EL1
2. EL1 enviou "{d| <mostre a porta da frente><c>}" a EL2
3. EL2 enviou "{T| <mostre>}" a COM
4. COM enviou "{T|[v-m-p-3-s-d-mostrar]}" a EL2
5. EL2 enviou "{T| <a>}" a COM
6. COM enviou "{T|[a-f-s][p-n-s][r-p-f-s-3-o]}" a EL2
7. EL2 enviou "{T| <porta>}" a COM
8. COM enviou "{T|[s-f-s-PORTA]}" a EL2
9. EL2 enviou "{T| <da>}" a COM
10. COM enviou "{T|[p-f-s]}" a EL2
11. EL2 enviou "{T| <frente>}" a COM
12. COM enviou "{T|[s-f-s-FRENTE]}" a EL2
13. EL2 enviou "{d| <[mostre|v-m-p-3-s-d-mostrar][a|a-f-s][porta|s-f-s-PORTA][da|p-f-s][frente|s-f-s-FRENTE]><c>}" a EL3
14. EL3 enviou "{P|}" a EL4
15. EL4 enviou "{D|flags}" a EL3
16. EL3 enviou "{d|c}" a EL4
17. EL4 enviou "{D|v}" a EL3
18. EL3 enviou "{d|[mostre|v-m-p-3-s-d-mostrar]#1}" a EL4
19. EL4 enviou "{D|d-n}" a EL3
20. EL3 enviou "{d|N}" a EL4
21. EL4 enviou "{D|[que]}" a EL3
22. EL3 enviou "{d|N}" a EL4
23. EL4 enviou "{D|r-p-*-3-*}" a EL3
24. EL3 enviou "{d|N}" a EL4
25. EL4 enviou "{D|v-n}" a EL3
26. EL3 enviou "{d|N}" a EL4
27. EL4 enviou "{D|OD}" a EL3
28. EL3 enviou "{d|[a|a-f-s][porta|s-f-s-PORTA][da|p-f-s][frente|s-f-s-FRENTE]#2}" a EL4
29. EL4 enviou "{D|d}" a EL3
30. EL3 enviou "{d|N}" a EL4
31. EL4 enviou "{D|OI}" a EL3
32. EL3 enviou "{d|N}" a EL4
33. EL4 enviou "{D|@}" a EL3
34. EL3 enviou "{d|@}" a EL4
35. EL4 enviou "{A|ent:N@PORTA(? ,1,FRENTE(? ,1,?))|FRENTE(? ,1,?)}" a COM
36. EL4 enviou "{I|Ordem(U,S,mostrar(S,PORTA(? ,1,FRENTE(? ,1,?)),?, ?, ?, ?, p))}" a COM

Figura 5.4: Processamento de uma fala dentro da EL.

Fala: "A porta está aberta".

1. COM enviou "{d|A porta está aberta}" a EL1
2. EL1 enviou "{d| <a porta está aberta><c>}" a EL2
3. EL2 enviou "{T| <a>}" a COM
4. COM enviou "{T|[a-f-s][p-n-s][r-p-f-s-3-o]}" a EL2
5. EL2 enviou "{T| <porta>}" a COM
6. COM enviou "{T|[s-f-s-PORTA]}" a EL2
7. EL2 enviou "{T| <está>}" a COM
8. COM enviou "{T|[v-i-p-3-s-l-estar][v-i-p-3-s-d-estar]}" a EL2
9. EL2 enviou "{T| <aberta>}" a COM
10. COM enviou "{T|[v-p-f-s-N-d-abrir][j-f-s-ABERTO]}" a EL2
11. EL2 enviou "{d| <[a|a-f-s][porta|s-f-s-PORTA][está|v-i-p-3-s-l-estar][aberta|v-p-f-s-N-d-abrir]><c>}" a EL3
12. EL3 enviou "{P}" a EL4
13. EL4 enviou "{D|flags}" a EL3
14. EL3 enviou "{d|c}" a EL4
15. EL4 enviou "{D|v}" a EL3
16. EL3 enviou "{d|[está|v-i-p-3-s-l-estar]#2}" a EL4
17. EL4 enviou "{D|v}" a EL3
18. EL3 enviou "{d|[aberta|v-p-f-s-N-d-abrir]#3}" a EL4
19. EL4 enviou "{D}" a EL2
20. EL2 enviou "{d| <[a|p-n-s][porta|s-f-s-PORTA][está|v-i-p-3-s-l-estar][aberta|v-p-f-s-N-d-abrir]><c>}" a EL3
21. EL3 enviou "{P}" a EL4
22. EL4 enviou "{D|flags}" a EL3
23. EL3 enviou "{d|c}" a EL4
24. EL4 enviou "{D|v}" a EL3
25. EL3 enviou "{d|[está|v-i-p-3-s-l-estar]#3}" a EL4
26. EL4 enviou "{D|v}" a EL3
27. EL3 enviou "{d|[aberta|v-p-f-s-N-d-abrir]#4}" a EL4
28. EL4 enviou "{D}" a EL2
29. EL2 enviou "{d| <[a|r-p-f-s-3-o][porta|s-f-s-PORTA][está|v-i-p-3-s-l-estar][aberta|v-p-f-s-N-d-abrir]><c>}" a EL3
30. EL4 enviou "{D}" a EL2
31. EL2 enviou "{d| <[a|a-f-s][porta|s-f-s-PORTA][está|v-i-p-3-s-d-estar][aberta|v-p-f-s-N-d-abrir]><c>}" a EL3
32. EL3 enviou "{D}" a EL2
33. EL2 enviou "{d| <[a|p-n-s][porta|s-f-s-PORTA][está|v-i-p-3-s-d-estar][aberta|v-p-f-s-N-d-abrir]><c>}" a EL3
34. EL4 enviou "{D}" a EL2
- ...
35. EL2 enviou "{d| <[a|a-f-s][porta|s-f-s-PORTA][está|v-i-p-3-s-d-estar][aberta|j-f-s-ABERTO]><c>}" a EL3
- ...
36. EL4 enviou "{A|ent:PORTA(? ,1,?)@PORTA(? ,1,?)}" a COM
37. EL4 enviou "{I|comunicar(U,estar(PORTA(? ,1,?) ,?(ABERTO,1,?) ,?,?,??,?,?,p)}" a COM p))}" a COM

Figura 5.5: Processamento de outra fala dentro da EL.

enviada pelo EL2 ao EL3 é rejeitada pelo EL4 (item 19), forçando o EL2 a enviar nova fala ao EL3 para análise (como foi descrito anteriormente). Esse ciclo é repetido até que o EL2 envie uma formação sintática que tenha lógica. Por problemas de espaço, algumas dessas tentativas são omitidas, tendo sido toda a comunicação entre os agentes EL3 e EL4 também omitida.

A primeira tentativa é rejeitada pelo agente EL4 por se tratar de uma fala com somente um verbo, sendo que este é de ligação, o que não é permitido, além de "a" ter sido classificado como preposição. Dessa forma o diálogo entre os agentes continua até que uma formação sintática aceitável seja enviada ao EL3 (item 35).

A Figura 5.6 apresenta o diálogo interno entre os agentes da Estrutura Lingüística e as estruturas do sistema quando do processamento seguido de duas frases: "a porta está aberta" e "o vento a abriu" (pelo fato de "a porta está aberta" ter sido a fala ilustrada na Figura 5.5 ela não foi repetida aqui). Esse exemplo ilustra a comunicação entre as estruturas para o referenciamento de um pronome (itens 33 e 34, que mostram o momento em que o agente EL4 pede à Estrutura Atencional uma referência ao pronome). Novamente, tentativas infrutíferas foram eliminadas do diálogo por questão de simplicidade.

A comunicação entre os agentes da Estrutura Lingüística segue as mesmas regras que a comunicação entre as estruturas (por meio de mensagens do tipo "{A|Mensagem}"), fazendo o agente EL o mesmo papel do agente COM. Uma diferença do agente EL é que o campo "A" da mensagem não contém o agente de destino, mas sim uma outra codificação, que dá, inclusive, pistas do conteúdo da mensagem. Com base nisso e no conhecimento de qual agente enviou a mensagem, o agente EL decide para quem repassá-la. A Tabela A.3 do Apêndice A mostra as ações tomadas pelo sistema baseadas na determinação do campo "A" da mensagem, ou seja, do agente de destino, e no conhecimento do agente remetente.

A seguir será detalhado o funcionamento de cada agente (com exceção do EL, que foi descrito aqui), incluindo sua função, suas estruturas de dados e comunicação com outras estruturas e agentes. Detalhes da codificação dos dados, assim como da classificação das palavras no dicionário, serão mostrados mais adiante, quando for apresentada a Estrutura da Tarefa.

5.2.1 O Agente EL1

Este agente, como já mencionado, é o responsável pela classificação da frase do usuário conforme a pontuação, e pela separação desta frase de entrada em sub-frases, de acordo com a necessidade.

Sua entrada é a frase digitada pelo usuário e transmitida ao EL1 pelo agente COM. Ao receber essa frase, EL1 verifica se ela é composta por mais de uma oração, separadas

Falas: "A porta está aberta" e "O vento a abriu".

1. COM enviou "{d|o vento a abriu}" a EL1
2. EL1 enviou "{d| <o vento a abriu><c>}" a EL2
3. EL2 enviou "{T| <o>}" a COM
4. COM enviou "{T|[a-m-s][r-p-m-s-3-o]}" a EL2
5. EL2 enviou "{T| <vento>}" a COM
6. COM enviou "{T|[s-m-s-VENTO]}" a EL2
7. EL2 enviou "{T| <a>}" a COM
8. COM enviou "{T|[a-f-s][p-n-s][r-p-f-s-3-o]}" a EL2
9. EL2 enviou "{T| <abriu>}" a COM
10. COM enviou "{T|[v-i-r-3-s-d-abrir]}" a EL2
11. EL2 enviou "{d| <[o|a-m-s][vento|s-m-s-VENTO][a|a-f-s][abriu|v-i-r-3-s-d-abrir]><c>}" a EL3
12. EL3 enviou "{D|}" a EL2
13. EL2 enviou "{d| <[o|a-m-s][vento|s-m-s-VENTO][a|r-p-f-s-3-o][abriu|v-i-r-3-s-d-abrir]><c>}" a EL3
14. EL3 enviou "{P|}" a EL4
15. EL4 enviou "{D|flags}" a EL3
16. EL3 enviou "{d|c}" a EL4
17. EL4 enviou "{D|v}" a EL3
18. EL3 enviou "{d|[abriu|v-i-r-3-s-d-abrir]#3}" a EL4
19. EL4 enviou "{D|d-n}" a EL3
20. EL3 enviou "{d|N}" a EL4
21. EL4 enviou "{D|[você]}" a EL3
22. EL3 enviou "{d|N}" a EL4
23. EL4 enviou "{D|[ele]}" a EL3
24. EL3 enviou "{d|N}" a EL4
25. EL4 enviou "{D|[ela]}" a EL3
26. EL3 enviou "{d|N}" a EL4
27. EL4 enviou "{D|OD}" a EL3
28. EL3 enviou "{d|[o|a-m-s][vento|s-m-s-VENTO]#1}" a EL4
29. EL4 enviou "{D|[que]}" a EL3
30. EL3 enviou "{d|N}" a EL4
31. EL4 enviou "{D|r-p-*-3-*}" a EL3
32. EL3 enviou "{d|[a|r-p-f-s-3-o]#2}" a EL4
33. EL4 enviou "{A|P:s-f-s}" a COM
34. COM enviou "{A|PORTA(? , 1, ?)}" a EL4
35. EL4 enviou "{D|d}" a EL3
36. EL3 enviou "{d|N}" a EL4
37. EL4 enviou "{D|OI}" a EL3
38. EL3 enviou "{d|N}" a EL4
39. EL4 enviou "{D|@}" a EL3
40. EL3 enviou "{d|@}" a EL4
41. EL4 enviou "{A|ent:VENTO(? , 1, ?)@VENTO(? , 1, ?)|PORTA(? , 1, ?)}" a COM
42. EL4 enviou "{I|comunicar(U,abrir(VENTO(? , 1, ?),PORTA(? , 1, ?),?, ?, ?, r), ?, ?, ?, p)}" a COM

Figura 5.6: Processamento de duas falas dentro da EL.

por sinais de pontuação. Se for, ele separa a frase nas sub-frases que a compõem. Assim, uma frase de entrada do tipo "Certo. Agora mostre a porta da frente." seria separada em duas sub-frases guardadas em uma lista ligada contendo "Certo" e "Agora mostre a porta da frente.", nessa ordem. Isso é necessário porque este tipo de frase realmente apresenta duas orações. A primeira, muito provavelmente, indica o término de um segmento, ou simplesmente indica que o usuário compreendeu a mensagem recebida, enquanto que a segunda é uma nova ordem. A Estrutura Intencional as processa de um modo diferente; então elas não podem chegar até ela como sendo uma sentença única. Por essa razão a frase é separada e enviada como duas orações distintas para a Estrutura Intencional.

Além de separar a frase em orações, este agente transforma todas as letras das palavras em minúsculas e também classifica essas orações conforme a pontuação final de cada uma. As regras para tal são simples: se houver um ponto final, a oração recebe o status "c", indicando que é um comentário ou que não há como dizer pela pontuação do que se trata; se houver um ponto de interrogação, é uma pergunta, e recebe o status "p"; e, por fim, se houver um ponto de exclamação, recebe o status "o", indicando que é ordem ou uma interjeição.

Então, uma entrada do tipo "Mostre a porta da frente" geraria uma saída do tipo "<mostre a porta da frente><c>"; uma entrada do tipo "O que aconteceu após o meio-dia?" geraria uma saída do tipo "<o que aconteceu após o meio-dia><p>".

Um outro status não mencionado é um adicional "s", que indica que a oração tem uma seqüência, ou seja, que a frase de entrada foi quebrada em sub-frases e após esta oração há outra armazenada. Dessa forma, a entrada "Certo. Agora mostre a porta da frente." é armazenada em uma lista de dois elementos: "<certo><cs>" e "<agora mostre a porta da frente><c>", que são enviados nessa ordem aos outros agentes.

Naturalmente, as orações que compõem essa lista não são enviadas juntas aos outros agentes. A primeira é enviada e, então, se assim for pedido, a segunda é enviada. Esse pedido é feito geralmente pela Estrutura Intencional, e é tomado como uma nova entrada do usuário.

A comunicação entre este agente e os outros é bem simples, dado que ele não interage tanto. A Tabela A.1 do Apêndice A mostra as possíveis comunicações recebidas e enviadas e as ações que o agente executa em cada caso.

5.2.2 O Agente EL2

EL2 é o agente responsável pela classificação sintática das palavras que compõem a oração. Ele recebe a oração de entrada enviada pelo agente EL1 e, a cada palavra, consulta a Estrutura da Tarefa para saber qual a classificação sintática desta. Ao receber esta classificação, o agente EL2 verifica quantas classificações diferentes há e, após to-

da a frase ter sido verificada, ele cria orações contendo todas as possíveis combinações, de acordo com a classificação sintática das palavras, empilhando-as para consulta futura. Por exemplo, em uma entrada do tipo "O menino caiu", as palavras "caiu" e "menino" têm apenas uma classificação; "o", por sua vez, possui duas: pode ser artigo ou pronome pessoal oblíquo. Nesse caso o agente criaria uma pilha de dois elementos (as duas possíveis combinações), contendo "o(artigo) menino(substantivo) caiu(verbo)" e "o(pronome) menino(substantivo) caiu(verbo)". Naturalmente a segunda será descartada quando o agente EL4 for criar a frase lógica, mas o trabalho do EL2 é fornecer todas as possíveis combinações, não importando quão esdrúxulas elas sejam. Por se tratar de uma pilha, as interpretações já enviadas são desempilhadas, de modo a evitar que o agente envie uma interpretação sintática que já foi processada e descartada.

Tal política foi adotada para garantir que todas as possibilidades de interpretação de cada frase fossem levadas em consideração. Assim, se a frase não fosse entendida, poder-se-ia tão somente considerar uma outra combinação sintática. Ou seja, caso uma frase ambígua fosse rejeitada pela Estrutura Intencional, bastaria pedir à Estrutura Lingüística que considerasse uma outra interpretação da mesma, sem aborrecer o usuário com uma mensagem de erro. Tal mensagem de erro somente seria enviada caso todas as possibilidades tivessem sido consideradas e, ainda assim, a lógica da frase tivesse sido rejeitada pela Estrutura Intencional.

Com relação aos *flags* da oração recebidos do agente EL1, eles são apenas repassados na frase, sem ter utilidade alguma para este agente (embora sua utilidade seja mostrada mais adiante). A comunicação entre o EL2 e o resto do sistema é mostrada na Tabela A.2 do Apêndice A, nos moldes da Tabela A.1.

5.2.3 O Agente EL3

EL3 é o agente responsável pela separação da oração do usuário em ilhas. Ele recebe a entrada enviada pelo agente EL2 e a analisa, separando as ilhas nela existentes. Durante a análise, há a possibilidade do EL3 detectar a presença de uma elipse do substantivo. Nesse caso, ele se comunica com a Estrutura Atencional em busca de uma referência para essa elipse.

Uma elipse é detectada quando uma ilha Objeto Direto é encontrada sem o substantivo, ou seja, quando há artigo e algum outro elemento, mas falta um substantivo (como em "a <porta> da frente" ou "o <carro> azul", onde "porta" e "carro" foram omitidos da sentença). Dessa forma, uma ilha Objeto Direto que, após totalmente formada apresenta essa característica, é uma forte candidata à existência de elipse (de outra forma seria um erro).

A separação das ilhas segue o algoritmo descrito na Figura 5.7. Este algoritmo se

baseia na análise das palavras que compõem a frase, tomadas duas a duas. Basicamente, um par de palavras é retirado da frase e sua compatibilidade é determinada (a Figura 5.8 apresenta o algoritmo da função que determina essa compatibilidade). Se elas forem compatíveis, então são agrupadas na mesma ilha; se não, são consideradas pertencentes a ilhas diferentes. Uma ilha, após definida, é adicionada a uma lista ligada de ilhas, guardada pelo EL3.

Separação das ilhas:

1. Retira o primeiro par de palavras da frase, atribui essas palavras a P_1 e P_2 , respectivamente. Cria nova ilha, adicionando P_1 a ela.
2. Se P_1 for verbo, a ilha está completa. Inclui essa ilha na lista ligada de ilhas, devolve P_2 à frase e volta a 1.
3. Se P_1 e P_2 forem compatíveis, ou seja, se podem fazer parte da mesma ilha, então concatena P_2 à ilha atual. Faz $P_1 = P_2$. Retira da frase de entrada nova palavra, atribuindo-a a P_2 . Volta a 3. Se a frase de entrada foi completamente lida, fecha a ilha e a inclui na lista de ilhas, pulando o passo 4.
4. Se não forem compatíveis, então:
 - (a) Se a incompatibilidade for devida a uma elipse do substantivo, então pede uma referência a ela à Estrutura Atencional. Ao receber a referência, devolve o atual P_2 à frase de entrada e faz $P_2 = \text{referência}$. Volta a 3 (note que em 3, o próximo elemento a ser obtido é o antigo P_2).
 - (b) Se não for devida à elipse, então fecha a ilha com P_1 , adicionando-a à lista de ilhas. Cria uma nova ilha, adicionando P_2 a ela. Faz $P_1 = P_2$, retira nova palavra da frase, atribuindo-a a P_2 e volta a 2. Se a frase de entrada foi completamente lida, fecha a ilha e a inclui na lista de ilhas.

Figura 5.7: Algoritmo para a separação de ilhas.

Por exemplo, para identificar as ilhas da frase "mostre a porta da frente agora", o EL3 retira da frase as palavras "mostre" e "a", criando uma nova ilha e adicionando "mostre" a ela (passo 1 da Figura 5.7). Como "mostre" é um verbo, EL3 inclui esta ilha na lista de ilhas, devolvendo "a" à frase original, que contém, então, "a porta da frente agora" (passo 2). Mais uma vez ele retira as duas primeiras palavras da frase, "a" e "porta", criando uma nova ilha, adicionando "a" a ela e verificando a compatibilidade entre essas duas palavras (passo 3). Como "a" é artigo e "porta" substantivo, segundo o passo 1 do algoritmo para verificação da compatibilidade (Figura 5.8), eles são compatíveis. Desse modo, seguindo o passo 3 do algoritmo da Figura 5.7, "porta" é adicionada à ilha, juntamente com "a".

Compatibilidade – entrada: par de palavras, P_1 e P_2 Caso P_1 for:

1. Artigo:

(a) Se P_2 for adjetivo, substantivo ou numeral, verifica se gênero e número concordam, se sim, então P_1 e P_2 são compatíveis, senão, não.

(b) Se P_2 for pronome possessivo, ou for "próprio" ou "mesmo", verifica se gênero e número concordam, se sim, então P_1 e P_2 são compatíveis, senão, não.

(c) Se P_2 for verbo no infinitivo, então P_1 e P_2 são compatíveis, senão, não.

2. Numeral ou adjetivo: se P_2 for adjetivo ou substantivo, verifica se gênero e número concordam, se sim, então P_1 e P_2 são compatíveis, senão, não.

3. Substantivo: se P_2 for adjetivo, numeral, pronome possessivo, ou preposição que indique posse ou função (como em "porta da frente" e "vara de pescar") verifica se gênero e número concordam, se sim, então P_1 e P_2 são compatíveis, senão, não.

4. Preposição: se P_2 for adjetivo, substantivo, numeral ou artigo, verifica se gênero e número concordam, se sim, então P_1 e P_2 são compatíveis, senão, não.

5. Pronome:

(a) Se pessoal ou de tratamento, é ilha. Incompatível com qualquer outra coisa.

(b) Se possessivo ou demonstrativo e P_2 for substantivo, adjetivo ou numeral, então verifica se gênero e número concordam, se sim, então P_1 e P_2 são compatíveis, senão, não.

(c) Se indefinido e P_2 for substantivo, então verifica se gênero e número concordam, se sim, então P_1 e P_2 são compatíveis, senão, não.

6. Qualquer outro é ilha por si só. Incompatível com qualquer outra coisa.

Figura 5.8: Algoritmo para verificação da compatibilidade entre duas palavras.

Até este momento, o EL3 possui duas ilhas, uma completamente determinada - "mostre" - e outra em aberto - "a porta".

Ele, então, e ainda seguindo o passo 3, retira mais uma palavra da frase, "da", verificando a compatibilidade entre "porta" e "da". O passo 3 da Figura 5.8 afirma que há tal compatibilidade. Dessa forma, o EL3 adiciona "da" à ilha contendo "a porta", retirando a palavra seguinte da frase ("frente").

A compatibilidade entre "da" e "frente" é garantida pelo passo 4 da Figura 5.8, fazendo com que, mais uma vez, o EL3 adicione "frente" à ilha em formação e verifique a próxima palavra, "agora".

O algoritmo da Figura 5.8 afirma que "frente" e "agora" não são compatíveis (passo 3). Então o EL3 segue ao passo 4 do algoritmo da Figura 5.7. Naturalmente, a incompatibilidade não se deve a uma elipse do substantivo, uma vez que "frente" é um substantivo. Então a ilha sendo processada é considerada completa e incluída na lista de ilhas.

Ainda seguindo o passo 4, o EL3 cria uma nova ilha, adicionando "agora" a ela. Como a frase de entrada acabou, esta ilha é considerada completa e adicionada à lista de ilhas.

Desse modo, ao final da execução do algoritmo da Figura 5.7, o EL3 possuirá uma lista com três ilhas: "mostre", "a porta da frente" e "agora", que são as ilhas presentes na frase "mostre a porta da frente agora".

Ao receber uma nova entrada do agente EL2, o EL3, então, esvazia a lista de ilhas, uma vez que elas pertencem a uma fala anterior. Essa lista somente não estará vazia caso tenha sido detectado algum erro com relação à oração anterior (como uma frase mal-formada que foi rejeitada pelo EL4 antes mesmo de todas as ilhas terem sido enviadas a ele). Além dessa lista de ilhas a serem enviadas, o agente também mantém uma lista das que foram enviadas, para o caso de haver a necessidade de recuperar alguma delas.

A definição da compatibilidade das palavras advém da definição dos tipos de ilhas e de um estudo da língua portuguesa. Tomando como exemplo a ilha Objeto Direto, pode-se dizer que ela tem cinco elementos: artigo, substantivo, adjetivo, numeral e um genitivo (indicador de posse, como, por exemplo "a porta da frente"). Então, o que essa função faz é definir esses cinco elementos como compatíveis, ou seja, são elementos que podem vir em seqüência (o algoritmo também trata da exigência de que o artigo, quando existir, seja o primeiro elemento da ilha), e agrupá-los em uma única ilha. Qualquer outro elemento, como verbo, advérbio ou preposição (que não o "de") seria tido como incompatível, não sendo incluído corretamente nessa ilha.

Uma vez feito esse trabalho de separação das ilhas, o agente envia um sinal ao agente EL4 indicando que terminou a separação. Este, por sua vez, começa a inquirir o agente EL3 acerca das ilhas encontradas. O formato da mensagem para o envio de uma ilha é "{d|ilha#pos}", onde pos é a posição da ilha na frase (a ilha 2 é a segunda que foi encontrada na frase, por exemplo). Esse número é usado pelo EL4 para determinar, por

exemplo, qual verbo um advérbio de negação está modificando.

Um quadro contendo a codificação de todas as comunicações entre o agente EL3 e outros agentes é apresentado na Tabela A.4 do Apêndice A.

5.2.4 Agente EL4

Este é o agente responsável pela montagem da frase lógica a partir das ilhas separadas pelo agente EL3, conforme mostrado anteriormente, nas Figuras 5.3 e 5.4. EL4 recebe um sinal do EL3 indicando que este está pronto para enviar as ilhas e, então, começa a pedi-las, uma a uma, montando, assim, a frase lógica correspondente à oração do usuário. Caso durante o processamento o agente EL4 encontre uma incoerência, ele sinaliza um erro ao agente EL, que envia ao agente EL2 um pedido para que este remeta uma nova combinação sintática ao EL3, que separará as ilhas dessa nova combinação e, novamente, enviará o sinal de prontidão ao EL4 (como demonstram os passos 12, 22 e 23 da Figura 5.5).

Ao receber o sinal de prontidão, EL4 pede que EL3 lhe diga quais são os *flags* da frase (passo 13). De posse desses *flags*, ele os analisa da seguinte maneira:

- Se o *flag* for "o", inicie a lógica com "Ordem(U,S,)"
- Se for "p", inicie a lógica com "Questão(U,S,)", e
- Se for "c", deixe em branco a lógica.

Em seguida, EL4 busca o verbo e seus complementos (passos 17 a 32 da Figura 5.4), adicionando-os a essa frase lógica inicial. Ao terminar a busca pelos complementos principais do verbo, EL4 ainda pede a EL3 que envie alguma ilha que tenha sobrado (passo 33). Caso haja tal ilha (ou tais ilhas), EL4 a traduz para lógica, adicionando-a à frase lógica inicial. Conforme já dito, se ele encontrar alguma incoerência na frase de entrada, ou seja, se não puder criar uma frase lógica correspondente, EL4 sinaliza um erro ao agente EL, que, por sua vez, envia um pedido para que o EL2 apresente uma nova interpretação sintática ao EL3.

Tendo terminado a definição da lógica correspondente à oração de entrada, EL4 então envia à Estrutura Atencional as entidades mais salientes nessa oração (passo 35), formadas pelos substantivos e nomes presentes nesta, bem como seus complementos, traduzidos para lógica.

Por fim, a frase lógica final é enviada à Estrutura Intencional para que possa ser analisado seu papel dentro do propósito do discurso (passo 36).

As Figuras 5.9 e 5.10 ilustram o algoritmo usado para buscar pelo verbo da frase e seus complementos, ou seja, para montar a correspondente lógica da oração do usuário.

Este algoritmo implementa a parte 2c da Figura 4.3, apresentada no Capítulo 4. Nele, o verbo é obtido fazendo-se um pedido explícito ao EL3 (passo 1).

Montagem da Lógica:

1. Busca uma ilha verbo.
2. Se não houver:
 - (a) Busca uma ilha OD. Se encontrar, inclui na frase lógica "OD(*ilha*)".
 - (b) Busca advérbio. Se encontrar, cria o verbo indeterminado com este advérbio preenchendo o campo devido. Por exemplo, se for "Hoje", preenche o campo "quando", incluindo a função "-(---,Hoje(),---,---)" na frase lógica.
 - (c) Busca ilha OI. Se encontrar, inclui na frase lógica "OI(*ilha*)".
3. Se Houver, então, se o verbo for:
 - (a) Infinitivo. Se de ligação:
 - i. Pede novo verbo. Se de ligação, então, se gerúndio ou particípio, pede um novo verbo. Se não achar ou achar e for de ligação acusa erro. Se achar, não for verbo de ligação, e for particípio ou gerúndio, então busca um sujeito e complementos e monta a lógica.
 - ii. Senão, se particípio ou gerúndio, busca um sujeito, complementos e monta a lógica.
 - (b) Senão, se infinitivo, busca um sujeito e complementos e monta a lógica.
 - (c) Indicativo: se seu *flag* for c, adiciona "comunicar(U,S,-)" à frase lógica. Então trata igual ao infinitivo
 - (d) Imperativo: se seu *flag* for c, adiciona "ordem(U,S,-)" à frase lógica. Se for de ligação, acusa erro. Senão, inclui o verbo na frase lógica, busca um sujeito e complementos e os adiciona a esta.

Figura 5.9: Algoritmo para codificação lógica da oração (já separada em ilhas).

O EL3, por sua vez, envia a ilha verbo ao EL4, caso ela exista. Se não houver verbo, o sistema busca somente os complementos, considerando que estes completam um verbo indeterminado, que foi apresentado numa fala anterior do diálogo (passo 2).

De posse do verbo, o sistema, então, o analisa, buscando se necessário outros verbos que o auxiliem (passo 3). Ao conseguir determinar completamente a ação, o sistema então busca uma entidade que possa servir de sujeito para esta ação e outros complementos, conforme ilustra o passo 2 da Figura 5.9.

Para encontrar o sujeito, EL4 pede uma ilha Objeto Direto ao EL3 e, ao recebê-la,

Montagem da Lógica: (continuação)

3. *Se houver verbo, então, se ele for:*

(e) **Particípio.** Se de ligação:

i. Busca um novo verbo. Se de ligação:

A. Se for indicativo, então busca um novo verbo. Se este for de ligação acusa erro. Senão, e se for particípio, então busca um sujeito e complementos e monta a lógica.

B. Senão, acusa erro.

ii. Senão, se for particípio, então pede um novo verbo e, se este for de ação e no indicativo, busca um sujeito e complementos, montando a lógica. Caso não seja particípio, então acusa erro.

(f) **Senão,** pede um novo verbo e, se não for de ligação, acusa erro. Se for de ligação:

i. Se indicativo, pede um novo verbo, se diferente de gerúndio de ligação, devolve, senão, busca um sujeito, os complementos e monta a lógica.

ii. Se gerúndio de ligação, pede novo verbo e, se este for de ligação no indicativo, busca um sujeito, complementos e monta a lógica. Se não for de ligação, acusa erro.

(g) **Gerúndio.** Se de ligação:

i. Pede novo verbo. Se este for de ligação, então, se indicativo ou particípio, pede verbo e, se este for de ligação, acusa erro. Senão, se for particípio, então busca um sujeito, os complementos e monta a lógica.

ii. Senão, pede um novo verbo e, se este for de ligação no indicativo, busca o sujeito, complementos e monta a lógica, senão acusa erro.

(h) **Senão,** completa o campo "enquanto" do verbo.

Figura 5.10: Algoritmo para codificação lógica da oração (continuação).

verifica se esta concorda ou não com o verbo da ação (no caso de haver um único verbo), ou com o verbo que estiver no indicativo (caso haja mais de um verbo). Isso garante que a primeira ilha Objeto Direto seja sempre que possível tomada como sujeito.

Dessa forma, para traduzir a frase "mostre a porta da frente agora", EL4 espera que o EL3 avise que separou as ilhas e, então, busca os *flags* da frase. Como estes são "c", EL4 não inclui nada na frase lógica, por não saber se trata-se de comentário ou ordem. Em seguida, ele busca o verbo da frase, no caso, "mostre".

Como o verbo está no modo imperativo, não é verbo de ligação e o *flag* da frase é "c", EL4 abastece a frase lógica com "ordem(U,S)" (conforme o passo 3(d) da Figura 5.9). EL4, então, traduz o verbo para sua forma lógica (descrita na Seção 5.1.1 deste capítulo) e o inclui na frase, obtendo "ordem(U,S,mostrar(S,,,,,p))". Vale notar que, pelo fato do verbo estar no imperativo e vir do usuário, EL4 assume que o sistema é o sujeito da ação.

Após isso, o agente busca os possíveis complementos da frase, encontrando a ilha Objeto Direto "a porta da frente" e a ilha Advérbio "agora". Como o papel de sujeito já foi preenchido e "mostrar" pede um objeto direto, esta ilha é codificada e incluída na frase lógica: "ordem(U, S,mostrar(S,PORTA(1,FRENTE(1,)),,,,,p))". Da mesma forma, "agora", por ser advérbio de tempo, é codificado e incluído no campo "quando" da ação: "ordem(U, S, mostrar(S,PORTA(1,FRENTE(1,)),Agora(,,,,,p))". Como não há mais elementos na frase de entrada, esta é a forma lógica correspondente de "mostre a porta da frente agora".

A comunicação desse agente com outras estruturas e outros agentes é descrita na Tabela A.5 do Apêndice A. Por ser grande parte dessa comunicação feita com o agente EL3 (e descrita na Tabela A.4), só é apresentada a comunicação com os outros agentes.

5.3 Estrutura da Tarefa

Conforme mencionado no Capítulo 4, esta é a estrutura responsável pelo armazenamento da informação dependente do domínio utilizada pelas outras estruturas. Seu funcionamento é bastante simples, sendo por isso composta por apenas um agente.

A função deste agente é atender os pedidos feitos pelas demais estruturas, buscando as respostas nos arquivos de dados correspondentes.

Ao receber uma entrada ele analisa sua origem e o pedido feito. Caso esta venha da Estrutura Intencional, pedindo uma receita para um determinado ato, o agente ET (responsável pelas ações da Estrutura da Tarefa) busca no arquivo de receitas uma receita para a ação desejada, remetendo-a à Estrutura Intencional.

Se a entrada vem da Estrutura Lingüística ou da Atencional, então é um pedido para que se busque uma palavra no dicionário. A Estrutura Intencional consulta essa estrutura para obter a classificação sintática das palavras que compõem a frase do usuário, necessária

para a separação dessa frase em ilhas. Ela, então envia tal pedido ao ET (contendo também a palavra a ser procurada), que busca esta palavra no dicionário, enviando todas as possíveis classificações sintáticas à Estrutura da Tarefa.

Já a Estrutura Atencional consulta a Estrutura da Tarefa em busca da classificação sintática de uma macro (descrita a seguir). Ela necessita desta informação para poder fazer o referenciamento de pronomes e elipses. O agente EL, então, procura uma palavra no dicionário que contenha tal macro, devolvendo sua classificação sintática. Nesse caso, o ET retorna uma só classificação, uma vez que somente substantivos ou nomes próprios são procurados.

As macros existem no dicionário para que o sistema possa mapear várias palavras numa única representação. Por exemplo, o substantivo "casa" é representado pela macro "CASA"; assim, estando no plural ou no singular, essa será a representação do sistema para a entidade "casa", entidade esta que ele deve "conhecer".

Da mesma forma, um verbo é representado pelo seu infinitivo; assim, todos os tempos verbais ditos em todas as pessoas são mapeados para uma única forma (usada depois na frase lógica). As informações restantes julgadas importantes também estão presentes na codificação, sob a forma de sinais (informações estas apresentadas a seguir).

As palavras contidas no dicionário são codificadas, conforme sua classificação gramatical, da seguinte maneira:

- Cada classificação é uma lista de atributos, separados por "-".
- Um par palavra-classificação é codificado como <palavra>[classificação] e, caso haja mais de uma classificação sintática, então ela é concatenada a este par.
- Cada classificação tem o seguinte formato, conforme a função gramatical da palavra:
 - Verbos: v-modo-tempo-pessoa-número-transitividade-infinitivo;
 - Artigos: a-gênero-número;
 - Adjetivos: j-gênero-número-MACRO;
 - Numerais: n;
 - Interjeições: i;
 - Substantivos: s-gênero-número-MACRO;
 - Preposições: p-gênero-número;
 - Pronomes: r-tipo-gênero-número-pessoa-caso.
- Os sinais usados para representar cada campo da classificação são:

- Gênero: m (masculino), f (feminino) e n (neutro, usado quando a palavra serve tanto como masculina quanto feminina);
- Modo: usado nos verbos, são eles: i (indicativo), n (infinitivo), p (particípio), m (imperativo), g (gerúndio);
- Tempo: p (presente), r (passado) e f (futuro);
- Pessoa: as pessoas verbais, 1, 2, e 3;
- Número: s (singular), p (plural) e n (neutro, para palavras que independem do número);
- Transitividade: d (transitivo direto), i (transitivo indireto), n (intransitivo), l (verbo de ligação) e b (bitransitivo);
- Tipo: p (pessoal), t (tratamento), o (possessivo), d (demonstrativo), r (relativo), i (indefinido), n (interrogativo);
- Caso: r (reto), o (oblíquo) ou N (não possui);

Dessa forma, a palavra "casa", citando como exemplo, existe no dicionário codificada como "<casa>[s-f-s-CASA][v-i-p-3-s-b-casar]", ou seja, é tanto um substantivo feminino singular representado pela macro "CASA", quanto um verbo que pode ser transitivo direto ou indireto (ou ambos), dito na terceira pessoa do singular do presente do modo indicativo e representado pelo seu infinitivo, "casar".

A codificação do livro de receitas é bem mais simples, contendo apenas o par ação-receita, codificado como no dicionário: <ação>[receita₁]. . . [receita_n]. Pelo fato que uma receita pode ser composta por uma ou mais sub-ações, ela é codificada da seguinte maneira: {sub-ação₁|restrição₁}. . . {sub-ação_n|restrição_n}, onde restrição_i é o conjunto de restrições (separadas por ";"), para sub-ação_i.

Por exemplo, a ação "mostrar" é representada no livro de receitas como "<mostrar (X,Y,T)> [{det_posição(X,Y)|ter_descr(X,Y)} {determinar_câmera(X,C,Y)} | {mostrar_imagem(X,C,T)|ligada(C)}]". Ou seja, há somente uma receita conhecida pelo sistema para a ação de "mostrar", dizendo que para que alguém (X) mostre algo (Y) no instante T, X deve determinar primeiro a posição de Y (tendo como pré-requisito para esta ação que Y seja conhecido por X), depois determinar que câmera (C) consegue filmar Y e, então, finalmente, mostrar a imagem desta câmera no instante estabelecido (desde que esta esteja ligada).

Na Tabela A.6 do Apêndice A é mostrada a comunicação entre essas estruturas as demais no sistema, mostrando a codificação usada.

5.4 Estrutura Atencional

A Estrutura Atencional é a responsável, conforme visto, pelo foco da atenção dos participantes. Ela mantém registro das entidades e intenções mais importantes no discurso, sendo, por isso, valiosa para o referenciamento de pronomes e elipses e para a própria compreensão do discurso pela Estrutura Intencional.

No sistema aqui desenvolvido, ela é composta por três agentes (Figura 5.11):

- EA: este agente está para a Estrutura Atencional assim como o COM está para o sistema. Ele é o responsável pela organização e transmissão da comunicação entre os agentes.
- FG: o FG é o agente responsável pelo foco geral, ou seja, pela manipulação da Pilha de Espaços de Foco.
- FI: responsável pela manipulação dos centros, este agente cuida do foco imediato do discurso.

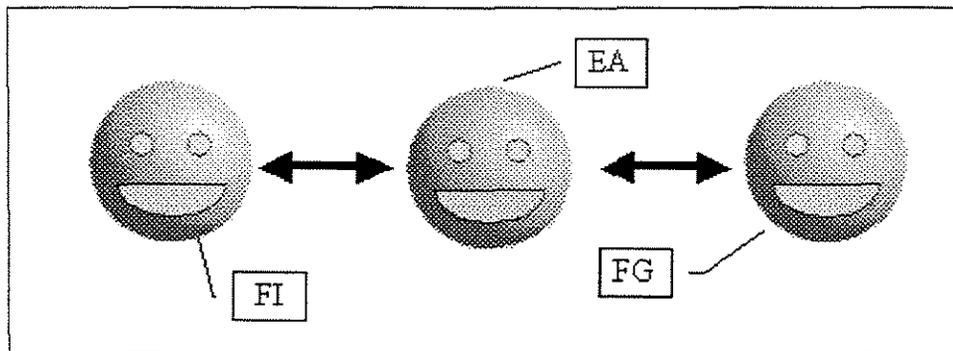


Figura 5.11: Esquema geral da Estrutura Atencional.

A codificação da comunicação entre os agentes segue o mesmo padrão do resto do sistema. Ao receber uma mensagem de fora da estrutura, o agente EA a analisa, de acordo com seu conteúdo e origem, e a envia ao agente responsável.

Se esta comunicação for proveniente da Estrutura Lingüística, então ela pode ser um pedido ou abastecimento de dados. No caso de pedido, o agente EA repassa a mensagem ao agente FI, pois tal pedido é para o referenciamento ou de uma elipse ou de algum pronome. Se for uma ordem, novamente ela é repassada ao agente FI, uma vez que se trata do abastecimento dos centros de busca.

Quando a comunicação vem da Estrutura Intencional, então o agente EA deve analisar seu conteúdo. Se for um pedido de centro de busca (usado em ordens do tipo "mostre" para determinar o alvo da ação), então este é repassado ao agente FI. Já se for um pedido para

que se verifique que ação está em foco atualmente, o pedido é enviado ao agente FG, que é o responsável pela Pilha de Espaços de Foco. Por fim, se a comunicação recebida se tratar de uma ordem para que um espaço de foco seja empilhado ou desempilhado, então esta ordem é repassada ao agente FI, para que este possa enviar as entidades presentes nos centros de busca para o agente FG e então repassar a ordem de empilhamento ou desempilhamento a este. Caso a ordem trate de um desempilhamento, não há a necessidade de passar as entidades presentes nos centros de busca ao FG, uma vez que serão perdidas. Nesse caso o agente FI apenas descarta tais entidades, repassando a ordem de desempilhamento ao FG.

Cada agente será visto com mais detalhes a seguir. O funcionamento e integração das Pilhas de Espaços de Foco e de Centros será mostrada na Seção 6.2 do Capítulo 6, quando serão apresentados os testes com o sistema.

5.4.1 O Agente FI

Este agente é o responsável pelo foco imediato do sistema, ou seja, ele é quem carrega as informações mais relevantes contidas no nível intrasegmental do diálogo.

Basicamente, ele implementa o funcionamento do mecanismo de centros de busca explicado no Capítulo 2. Para tal ele possui uma pilha de três elementos, conforme mencionado no Capítulo 4, que guarda os centros de busca das três últimas falas do diálogo.

Ao receber da Estrutura Lingüística as entidades mais salientes em uma fala (em forma lógica), o agente FI as classifica da seguinte maneira:

- Se esta fala possui algum membro resultante de referenciamento de elipse ou pronome, então este será seu C_r . Se houver mais de um elemento assim, o C_r será o que estiver no papel de sujeito ou, caso não haja sujeito, o primeiro na ordem superficial da fala.
- Caso não haja um membro assim, o C_r é deixado em branco (pode se tratar da primeira frase do diálogo).
- A entidade servindo de sujeito é então colocada como primeiro elemento do C_p , sendo as outras entidades incluídas neste na ordem em que estão (que corresponde à ordem superficial na oração), dando preferência ao elemento que está no papel de C_r , se houver.
- Caso não haja um Sujeito, a entidade presente no C_r assume a primeira posição do C_p , sendo seguida pelas demais.

Dessa forma, após ter classificado os membros mais salientes da oração, o FI está pronto para fazer o referenciamento posterior de entidades. Exemplos dessa classificação podem ser encontrados na Seção 6.2 do Capítulo 6.

Para fazer a referência a um pronome ou elipse, o FI dispõe apenas da classificação sintática do membro procurado (por exemplo, substantivo feminino singular, conforme o passo 33 da figura 5.6). Então ele começa a busca no C_r da fala anterior (seguindo a regra que seqüências de continuação são preferidas a seqüências de retenção, apresentada no Capítulo 2). Se houver tal entidade, ele verifica se esta é candidata a servir de referência para este pronome ou elipse. Para tal, ele envia a macro da entidade (que é sua representação lógica) à Estrutura da Tarefa, que busca em seu dicionário, retornando a classificação sintática desta. Caso a entidade seja compatível com o que é procurado, então o sistema analisa se o pedido é de elipse ou pronome. Se for de pronome, ele envia a entidade lógica como um todo; se for de elipse, envia somente sua classificação sintática. Dessa forma, uma entidade do tipo "CASA(BRANCA,1,)" será enviada como "CASA(BRANCA,1,)" em resposta a um pedido de referência de pronome, e como "s-f-s-CASA" para um pedido de referenciamento de elipse.

Caso o C_r não seja compatível ou esteja vazio, cada elemento do C_p é testado em ordem e, se nenhum candidato compatível for encontrado, a busca é repetida para cada par de centros na Pilha de Centros e, se novamente essa busca for infrutífera, então um pedido é enviado ao FG e a busca é feita na Pilha de Espaços de Foco. A busca na pilha de focos é feita somente para pronomes, pois elipses são entendidas como sendo mais locais e, portanto, tratadas somente na Pilha de Centros.

Cada vez que o sistema é abastecido com as entidades mais salientes em uma fala, ele, após classificar estas entidades, empilha-as na Pilha de Centros. Caso a pilha esteja cheia (contenha já três elementos), o C_p de seu último membro é retirado desta e enviado ao FG, para ser incluído no espaço de foco correspondente. O C_r não é enviado por já fazer parte do C_p desta ou da fala anterior. A Seção 6.2 do próximo capítulo apresentará um exemplo do referenciamento de pronomes e da manipulação da Pilha de Centros acima descritos.

Se, após ter sido abastecido com os centros de uma sentença, o sistema recebe um aviso da Estrutura Intencional dizendo que a oração foi rejeitada, então estes centros são desempilhados, restando na Pilha de Centros o contexto anterior ao processamento desta fala.

Na Tabela A.7 do Apêndice A está listada a comunicação deste agente com os demais no sistema, responsável pelo processamento acima descrito.

5.4.2 O Agente FG

Responsável pelo foco global do sistema, este agente trata da manipulação dos espaços de foco em sua pilha, carregando, além das entidades mais salientes no segmento de discurso ao qual o espaço de foco pertence, a intenção que corresponde ao propósito deste segmento de discurso.

Para tal, ele possui uma pilha, a Pilha de Espaços de Foco, conforme descrita no Capítulo 2. Cada espaço de foco nessa pilha é uma estrutura de quatro elementos:

- O propósito do segmento de discurso, representado por uma intenção abastecida pela Estrutura Intencional.
- O endereço dessa intenção no grafo de receitas da Estrutura Intencional.
- Uma pilha dos elementos mais salientes no segmento de discurso, em ordem de aparição, conforme enviados pelo agente FI.
- Um campo indicando se este segmento faz parte de uma interrupção ou não.

Este agente age mais como um coadjuvante do agente FI, sendo sua comunicação com outros agentes pequena. Pedidos de empilhamento, desempilhamento e referência para pronomes e elipses, além de abastecimento de novas entidades, são todos feitos pelo FI, cabendo à Estrutura Intencional apenas fazer o pedido direto de qual intenção está atualmente em foco (e se faz parte de uma interrupção) e à Estrutura da Tarefa tão somente responder o que lhe é pedido pelo FG.

Ao receber um pedido da Estrutura Intencional para que a intenção em foco seja remetida, o FG não somente envia tal intenção, como também seu endereço no grafo de receitas e uma indicação sobre ser esta intenção parte ou não de uma interrupção (seguindo o formato apresentado na Tabela A.8 do Apêndice A). Todas estas informações são usadas pela Estrutura Intencional na interpretação e geração de falas.

O restante do trabalho deste agente é atuar conforme as ordens recebidas. Se o agente FI envia uma ordem de empilhamento, contendo uma nova intenção, seu endereço no grafo de receitas e uma indicação de ser ou não interrupção, o FG empilha tal espaço, preenchendo os campos devidos na estrutura acima descrita. Da mesma forma, se a mensagem enviada trata do abastecimento de novas entidades que se tornaram salientes no diálogo, o FG as empilha, tomando o cuidado de não repetir entidades. Se uma entidade for citada novamente, ela é retirada de seu lugar original na pilha e recolocada em seu topo, ou seja, sua prioridade em uma busca é aumentada. O empilhamento de entidades na Pilha de Espaços de Foco (PEF) pode ser visto na Seção 6.2 dos testes do sistema no Capítulo 6.

A comunicação com a Estrutura da Tarefa se dá quando o FI envia um pedido ao FG para o referenciamento de determinada entidade. Neste caso, o FG busca na Pilha de Espaços de Foco tal entidade (tomando o cuidado de não cruzar limites entre interrupção e segmento interrompido). Para isso ele envia uma mensagem à Estrutura da Tarefa contendo, para cada segmento permitido e cada elemento da pilha de elementos mais salientes no segmento, um pedido de remessa da classificação sintática deste elemento. Ao encontrar o elemento buscado (cuja classificação sintática combine com a classificação da entidade buscada), ele o remete ao FI, conforme pedido.

A Tabela A.8 do Apêndice A descreve todas as comunicações recebidas e enviadas por este agente, mostrando o que foi acima discutido.

5.5 Estrutura Intencional

A Estrutura Intencional é a responsável pela compreensão propriamente dita das falas em um diálogo. A cada nova fala, sua função é determinar como esta contribui para que seja atingido o propósito do diálogo.

Conforme dito no Capítulo 4, seu modelamento foi feito por Lochbaum (1994), usando a teoria de Planos Compartilhados de Grosz e Sidner (1990). Por essa razão, o sistema aqui desenvolvido utiliza o sistema desenvolvido por Lochbaum para implementar essa estrutura, sistema este (de Lochbaum) que implementa o algoritmo para o processo de incrementação do plano compartilhado mostrado nas figuras 4.6 e 4.7.

5.5.1 A Estrutura Intencional Desenvolvida

O módulo projetado para fazer o papel da Estrutura Intencional no sistema desenvolvido, como já dito, é uma adaptação do sistema desenvolvido por Lochbaum e, justamente por isso, é composto por um único agente, o EL, que é o responsável por toda a Estrutura Intencional.

O agente EL recebe uma entrada da Estrutura Lingüística numa forma lógica determinada e, então, deve traduzir esta forma lógica em uma das proposições de entrada do sistema desenvolvido por Lochbaum. Para tal ele faz um mapeamento simples entre uma forma lógica e outra. De posse dessa nova proposição lógica, ele então a analisa da mesma forma que o sistema de Lochbaum faria.

Devido à natureza das duas aplicações (o sistema desenvolvido por Lochbaum e este) ser diferente, há um conjunto de diferenças entre eles:

- No sistema desenvolvido não há mais o "oráculo", como no sistema original de Lochbaum. O sistema sempre assume um papel ativo, ou seja, ele sempre presume

que, após uma fala do usuário, cabe a ele fazer algo. Para decidir o que fazer, ele se baseia na forma da entrada. Se for uma pergunta, ele tenta respondê-la e, se for uma ordem, tenta executá-la. Da mesma forma, ao detectar uma restrição não satisfeita durante a execução de uma tarefa, o sistema tenta satisfazer esta restrição sem a ajuda do usuário, só recorrendo a este em último caso.

- O sistema desenvolvido usa a Pilha de Espaços de Foco da Estrutura Atencional para guardar as intenções que estão em foco e seu endereço no GrafoR correspondente, seguindo a teoria de Grosz e Sidner (1985, 1986), apresentada no Capítulo 2. No sistema de Lochbaum uma pilha de intenções faz o papel da Pilha de Espaços de Foco para esta estrutura.
- No sistema desenvolvido a saída para o usuário é feita em língua natural, e não através da apresentação do GrafoR e da pilha de intenções, como no sistema de Lochbaum. Justamente por isso, quando o sistema dá uma resposta para o usuário, ele também envia uma mensagem à Estrutura Atencional fornecendo as entidades mais salientes nesta fala, para que essa estrutura possa atualizar sua Pilha de Centros. Essa resposta é gerada por uma função pré-definida no sistema, não tendo nenhuma característica mais elaborada. A proposição lógica é traduzida diretamente para palavras, sem o uso da Estrutura Atencional para auxiliar na coerência (como sugerido no Capítulo 2).
- Em vez de acusar erro caso não possa explicar como uma fala do usuário se ajusta ao plano em foco (como faria o sistema de Lochbaum), este sistema assume ser esta fala parte de uma interrupção, o que não é tratado no sistema de Lochbaum. Então ele envia uma mensagem à Estrutura Atencional, contendo o conteúdo proposicional desta fala e indicando que é uma interrupção, para que esta possa tomar as medidas cabíveis. Esta reação do sistema é essencial para que se possa tratar de uma interrupção pois, do contrário, quando esta ocorresse, seria tida como se fosse um erro do usuário, uma vez que a intenção nela contida não tem relação com o Propósito do Segmento de Discurso que está em foco.
- Para não perder o contexto do segmento interrompido, quando uma interrupção é detectada ela é armazenada em um segundo GrafoR, GR2, que existe somente para guardar o contexto dentro das interrupções.

Além dessas diferenças, uma outra que surge é a presença de comunicação com as outras estruturas, conforme mostra a Tabela A.9 do Apêndice A. Nesta tabela pode ser visto que o sistema também faz consultas à Estrutura Atencional em busca de entidades que estejam em foco, e não somente em busca de uma intenção ou para atualizar entidades. Esse tipo de busca se faz necessário para resolver diálogos do tipo:

S – A porta da frente foi aberta.

U – Mostre.

Assim, para saber o que mostrar, o sistema deve pedir à Estrutura Atencional qual entidade está em foco no momento. A Estrutura Atencional, por sua vez, busca em sua Pilha de Centros o C_r da última fala do diálogo e, caso este não exista, busca o elemento de classificação mais alta do C_p , retornando-o à Estrutura Intencional para que esta possa completar a frase de entrada do usuário.

No próximo capítulo serão mostrados exemplos do sistema em operação, onde poderá ser visto o modo que o sistema reage (e age) de acordo com as falas do usuário.

Capítulo VI

Exemplos do Sistema em Operação

Neste capítulo serão apresentados alguns testes feitos com o sistema desenvolvido. Para executar estes testes foram elaborados dois diálogos, mostrados nas Figuras 6.1 e 6.2, além de dois outros testes (não feitos por meio de diálogos), visando mostrar o referenciamento de pronomes, através do monitoramento dos centros de busca, e o tratamento e separação das ilhas que compõem a entrada do usuário.

6.1 Separação das Ilhas

O ponto principal do mecanismo de separação e classificação das ilhas descrito neste capítulo é a idéia de que, sem nenhuma programação adicional, uma mesma sentença, dita de formas diferentes (trocando-se a ordem das suas ilhas), resultaria na mesma frase lógica, ou seja, se o sistema aceita uma frase com seus componentes em uma determinada ordem, então ele deve aceitar todas as outras ordens (dadas as restrições mencionadas).

Para testar este mecanismo, a estrutura intencional foi separada do sistema (para evitar que o sistema tentasse entender as frases usadas no teste). Então as seguintes frases foram testadas:

- A porta dos fundos agora mostre.
- A porta dos fundos mostre agora.
- Agora mostre a porta dos fundos.
- Agora a porta dos fundos mostre.
- Mostre agora a porta dos fundos.

- Mostre a porta dos fundos agora.

Estas frases correspondem às seis possíveis permutações das ilhas "a porta dos fundos", "agora" e "mostre".

Do processamento de cada uma destas frases resultou uma única frase lógica: $Ordem(U,S,mostrar(S,PORTA(1,FUNDO(,)),Agora(,.,.,p))$. Ou seja, o sistema traduziu todas essas frases para esta única forma lógica.

É claro que, se a frase de entrada for ambígua, há uma grande chance de que o sistema a interprete de modos diferentes caso suas ilhas sejam permutadas. Por exemplo, a frase "a porta foi aberta" pode significar que alguém abriu a porta ou que a porta esteve aberta alguma vez no passado. Pior do que isso, se a ordem for mudada e obtivermos "a porta aberta foi", "aberta" pode ser interpretada como adjetivo, e não verbo no particípio. Essa última combinação pode ser bastante incomum e o uso do particípio ser claro para um humano. Mas vale lembrar que ela deve ser entendida no contexto de um diálogo e, dependendo deste contexto, a interpretação de que "<a porta aberta> <foi> <algo>" é válida.

6.2 Centros de Busca

Para testar o funcionamento dos centros de busca, essenciais ao referenciamento de pronomes, foi escolhida uma série de quatro frases que ilustram as mudanças de centro, conforme explicadas no Capítulo 2.

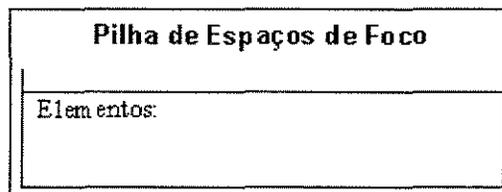
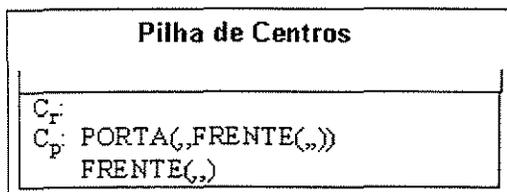
Cada frase foi entrada na ordem mostrada abaixo e as pilhas de Centros e de Espaços de Foco foram monitoradas para verificar seu comportamento. Além disso, sua tradução lógica (conforme enviada à Estrutura Intencional) também é apresentada.

Novamente, como no teste com as ilhas, a estrutura intencional foi retirada do sistema para facilitar o trabalho e a coleta de dados para estes testes.

As frases testadas foram:

1. A porta da frente foi aberta.

Lógica: $comunicar(U,abrir(p,PORTA(1,FRENTE(1,)),.,.,r),.,.,p)$



Nessa entrada vemos que a Pilha de Espaços de Foco está vazia, enquanto que a Pilha de Centros contém os elementos mais salientes na oração. Na Pilha de Espaços de Foco, os itens referentes à ação e seu endereço no GrafoR foram omitidos por simplicidade (além do que, a Estrutura Intencional foi separada do sistema). Com relação à Pilha de Centros, vale notar que, por ser a primeira sentença, esta não possui um C_r . Além disso, está em foco não somente a entidade "porta da frente", mas também a entidade "frente", que é considerada implicitamente em foco.

2 Ela abriu agora.

Lógica: comunicar(U,abrir(PORTA(1,FRENTE(1)),D,Agora(),,,,r),,,,p)

Pilha de Centros	
C_r	PORTA(,FRENTE(,))
C_p	PORTA(,FRENTE(,)) FRENTE(,)
C_r	PORTA(,FRENTE(,))
C_p	PORTA(,FRENTE(,)) FRENTE(,)

Pilha de Espaços de Foco
Elementos:

Com essa frase, já pode ser visto o referenciamento de um pronome - ela. Para fazer tal referenciamento, o sistema buscou o C_r da fala anterior e, como esta não possuía um, usou o elemento de classificação mais alta do C_p que pudesse servir de referência (no caso, buscou um substantivo feminino singular). Então, pelo fato dessa entidade ser o sujeito da oração 2 e ter sido referenciada por um pronome, ela foi incluída no C_r da segunda oração. Como ambas se referem à mesma entidade, seus C_p s são idênticos.

A Pilha de Espaços de Foco continua sem elementos porque ainda não foram ditas mais de três frases no discurso.

3 O vento a abriu.

Lógica: comunicar(U,abrir(VENTO(1),PORTA(1,FRENTE(1)),,,,r),,,,p)

Pilha de Centros	
C_r	PORTA(,FRENTE(,))
C_p	VENTO(1) PORTA(,FRENTE(,)) FRENTE(,)
C_r	PORTA(,FRENTE(,))
C_p	PORTA(,FRENTE(,)) FRENTE(,)
C_r	PORTA(,FRENTE(,))
C_p	PORTA(,FRENTE(,)) FRENTE(,)

Pilha de Espaços de Foco
Elementos:

Novamente, por ter sido representada por um pronome na fala 3, a entidade "porta da frente" foi atribuída ao C_r desta fala. A diferença entre esta fala e a fala 2 é que aqui a entidade que foi atribuída ao C_r da fala não é o elemento de classificação mais alta de seu C_p . "Vento" é a entidade de mais alta classificação por ser o sujeito da fala 3.

4 Ele está forte hoje.

Lógica: comunicar(U,estar(VENTO(,1,),D,,Forte,,,P),,,,,P)

Pilha de Centros	
C_r :	VENTO(,1,)
C_p :	VENTO(,1,)
C_r :	PORTA(,FRENTE(,))
C_p :	VENTO(,1, PORTA(,FRENTE(,)) FRENTE(,))
C_r :	PORTA(,FRENTE(,))
C_p :	PORTA(,FRENTE(,)) FRENTE(,)

Pilha de Espaços de Foco	
Elementos:	
	PORTA(,1,FRENTE(,1,))
	FRENTE(,1,)

Aqui, finalmente a Pilha de Espaços de Foco recebeu a frase que ocupava a última posição da Pilha de Centros. Por ser o sujeito da fala 4 e também ter sido referenciado por pronome, a entidade "vento" assume a posição de elemento de mais alta classificação no C_p e de C_r da fala 4. Este exemplo ilustra o funcionamento dos centros de uma fala no decorrer do diálogo, conforme descrito no Capítulo 2. As quatro sentenças apresentam uma mudança de centro, feita por meio de uma continuação (falas 1 e 2), uma retenção (fala 3) e, finalmente uma mudança (fala 4). Vale lembrar que as mudanças de centro não precisam ser tão coerentes assim, uma vez que o sistema, ao tentar achar a referência de um pronome, leva em conta informações como gênero e número. Dessa forma, a referência a um pronome como "ele", por exemplo, será tida como sendo o primeiro elemento da Pilha de Centros que represente um substantivo masculino singular.

6.3 Compreensão e Resposta

Para testar o funcionamento da Estrutura Intencional e a integração de todas as estruturas quando juntas, foram criados dois diálogos que deveriam ser "entendidos" pelo sistema. Estes diálogos estão apresentados nas Figuras 6.1 e 6.2.

O primeiro exemplo (Figura 6.1) é um diálogo simples, composto apenas por uma intenção básica (o usuário querendo saber os fatos ocorridos após as 14 horas) e interrupções geradas pelo próprio usuário com o intuito de obter mais informações sobre o fato em foco ou sobre outros fatos, como nas falas 3, 5 e 7. Após ter sua curiosidade satisfeita,

o usuário indica o desejo de que o sistema retome a ação original (fala 9), o que o sistema faz com sucesso (fala 10).

O segundo diálogo (Figura 6.2) apresenta um exemplo mais elaborado do modo como o sistema funciona. A intenção por trás deste diálogo é obtida da fala 1, e é o desejo do usuário que o sistema mostre a ele uma determinada (ou indeterminada) porta. Diálogos desta natureza são geralmente sub-segmentos de diálogos maiores, mas aqui este sub-segmento é tomado como o diálogo todo para fins de testar o sistema.

1. O que aconteceu após as 14 horas? (Usuário)
2. A porta dos fundos foi aberta às 14:30. (Sistema)
3. Mostre. (Usuário)
4. <ação> (Sistema)
5. Mostre a porta da frente agora. (Usuário)
6. <ação> (Sistema)
7. Alguma vez ela foi aberta? (Usuário)
8. Não. (Sistema)
9. OK. Continue. (Usuário)
10. Às 15:00 a janela dos fundos foi aberta. (Sistema)

Figura 6.1: Primeiro diálogo teste entre o Usuário e o Sistema.

1. Mostre a porta. (Usuário)
2. Qual? (Sistema)
3. A da frente. (Usuário)
4. Não posso. A câmera está desligada. Você terá de ligá-la. (Sistema)
5. Como? (Usuário)
6. Vá até a sala, ela está em cima da lareira. Aperte o botão "liga". (Sistema)
7. OK. Feito. (Usuário)
8. <ação> (Sistema)

Figura 6.2: Segundo diálogo entre o Usuário e o Sistema.

Ao receber a ordem do usuário o sistema verifica que a informação dada por este é insuficiente, pedindo então mais detalhes (fala 2). Ao receber esta nova informação o sistema conclui que esta é suficiente para a execução da ordem, e tenta executá-la. Nesse ponto, um problema ocorre e o sistema deve pedir a ajuda do usuário para sua solução (fala 4). O usuário, por sua vez, não sabe como pode ajudar o sistema, buscando que este lhe diga o que fazer (falas 5 e 6). Sabendo o que fazer, o usuário resolve o problema, avisando, então, o sistema (fala 7), que finalmente executa a tarefa sem problemas (fala 8).

6.4 O Sistema em Operação

A seguir, serão mostrados os resultados do processamento dos dois diálogos-teste pelo sistema. Após cada fala do usuário é mostrado o estado da Pilha de Espaços de Foco e do GrafoR usado (GR, caso a fala não faça parte de uma interrupção, e GR₂, caso faça), além de ser dada uma pequena explicação de como se deu o processamento.

Por questão de simplicidade, somente serão apresentadas as intenções presentes na Pilha de Espaços de Foco, deixando de fora a lista de entidades mais salientes na fala. Além disso, a Pilha de Centros não é mostrada por ter sido seu funcionamento demonstrado anteriormente.

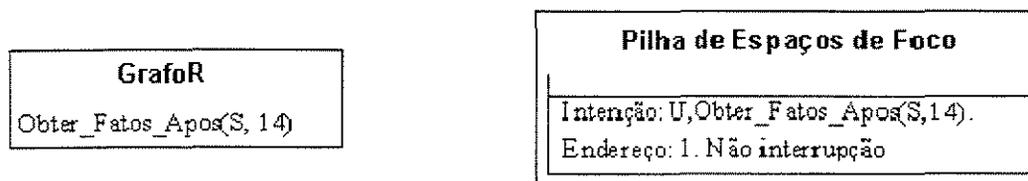
As proposições lógicas usadas neste processamento possuem o formato determinado por Lochbaum (1994), obtido após ter havido um mapeamento da forma lógica do sistema (gerada pela Estrutura Lingüística) para a lógica proposta por Lochbaum (1994).

6.4.1 Diálogo I

O primeiro diálogo-teste (Figura 6.1) é iniciado pelo usuário com a frase:

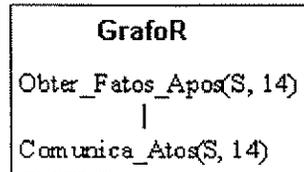
- *O que aconteceu após as 14 horas?*

O sistema, então, transforma esta frase na proposição *deseja(U,Obter_Fatos_Apos(S,14))*. Como esta é a frase inicial do diálogo, ela é diretamente empilhada na pilha de espaços de foco (PEF) e adicionada ao GrafoR principal – GR – que, naturalmente, estará vazio.



O item "endereço" na Pilha de Espaços de Foco é o endereço da intenção no GrafoR (1 é a raiz). Então de posse deste endereço e sabendo tratar-se de interrupção ou não, pode-se encontrar a intenção no GrafoR apropriado.

Como trata-se de um desejo do usuário, o sistema tenta executá-lo. Para tal ele busca uma receita para *Obter_Fatos_Após(X,Y)* e coloca no GrafoR, instanciando as variáveis necessárias:



Após incluir a receita, o sistema a executa. Comunica_Atos é uma função básica do sistema, e consiste deste narrar os fatos acontecidos após o horário determinado, um a um, verificando se o usuário diz algo. O primeiro fato é lido, e repassado ao usuário :

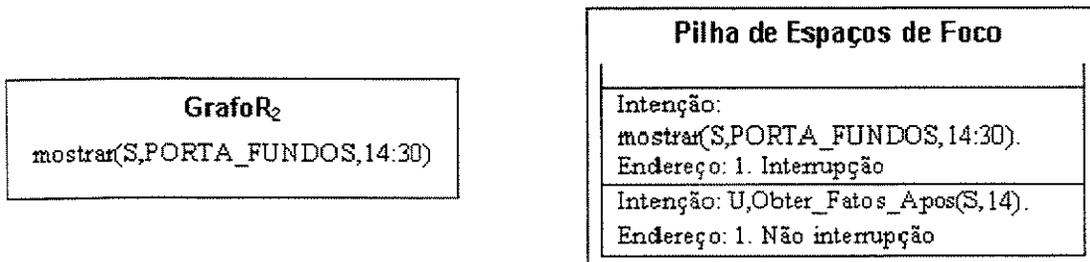
- *A porta dos fundos foi aberta às 14:30.*

Nesse ponto, as entidades PORTA(1,FUNDO(,)) e FUNDO(,) são enviadas à Estrutura Atencional, para fazer parte da Pilha de Centros.

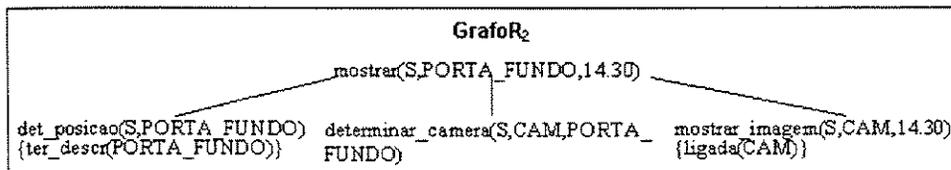
O usuário então gera uma interrupção;

- *Mostre.*

Para explicar esta frase, o sistema pega a proposição de entrada, *deseja(U,mostrar(S, X, t))*, e verifica se *mostrar(S, X, t)* faz parte das receitas para *Obter_Fatos_Apos*. Como não faz parte, ele supõe que se trata de uma interrupção, empilhado-a na PEF e incluindo-a no GrafoR para interrupções (GR₂). Antes de executar tal empilhamento, o sistema busca a entidade que está em foco para poder fazer o referenciamento de X, efetivamente atualizando a PEF e GR₂ com *mostrar(S,PORTA_FUNDOS,14:30)*.



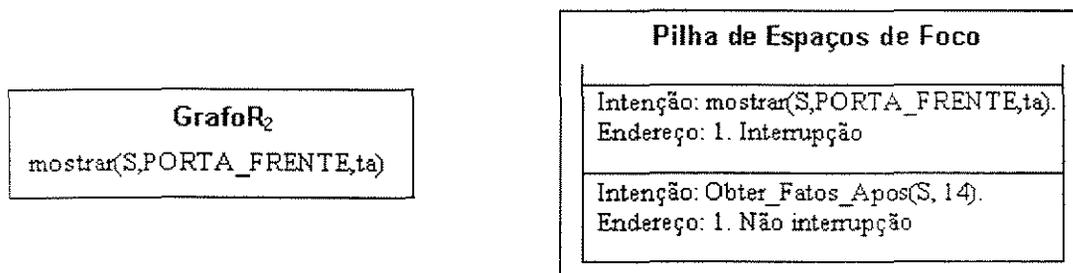
Após empilhar a intenção, o sistema busca uma receita para esta, atualizando o grafo de interrupções e, então, começa a executá-la. Como nenhuma restrição foi violada, o sistema, em resposta, mostra a imagem.



Em sua próxima fala, o usuário faz nova interrupção:

- *Mostre a porta da frente agora.*

Novamente o sistema tenta explicar a entrada, *deseja(U,mostrar(S,PORTA_FRENTE,ta))*, buscando se esta existe na receita da função em foco. Para tal ele verifica na PEF quem está em foco e busca no grafo se esta função está completa, neste caso, a função *mostrar(S,X,t)* está completa no grafo e o sistema então limpa GR₂ e desempilha esta intenção da PEF. Novamente verifica a função em foco, *Obter_Fatos_Apos(S,14)*, e, como esta não está completa em GR, verifica se *mostrar(S,PORTA_FRENTE, ta)* pertence à sua receita. Uma vez que não pertence, ele trata novamente como interrupção, atualizando a PEF e GR₂.



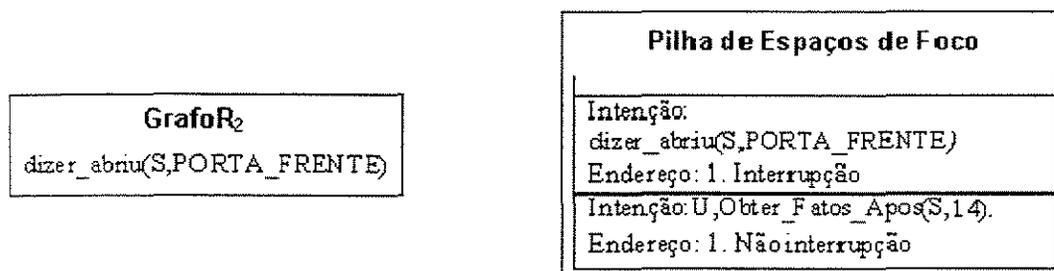
Seguindo os passos anteriormente descritos, o sistema busca uma receita para a ação, incluindo-a no GR₂ e executando-a, mostrando uma imagem ao usuário.

Em sua próxima frase, o usuário busca mais informações sobre a entidade em foco:

- *Alguma vez ela foi aberta?*

Mais uma vez o sistema tenta explicar a entrada do usuário, *deseja(U,dizer_abriu(S,PORTA_FRENTE))*. Para tal, ele verifica a ação que está em foco e, ao ver que ela foi completada no GR₂, a retira deste, desempilhando-a também da PEF. Em seguida, o sistema verifica a ação que assumiu o foco após o desempilhamento e busca saber, no GR, se *dizer_abriu(S, PORTA_FRENTE)* pertence à sua receita.

Uma vez que não pertence, novamente o sistema a trata como interrupção, incluindo a ação no GR₂ e empilhando-a na PEF.



O sistema, então, como antes, tenta executar o pedido a ele feito. Uma vez que dizer *abriu* é uma função básica do sistema, ele a executa diretamente, dando o resultado, isto é, respondendo que não.

O usuário, então, resolve continuar a ação interrompida, no caso, saber o que aconteceu após as 14 horas:

- *OK. Continue.*

Essa frase é entendida pelo sistema como o envio de um "ok" por parte do usuário (Figura 4.11), indicando que ele crê que a interrupção acabou. Nesse ponto o sistema verifica, usando a PEF e o GR₂, se a interrupção foi realmente terminada, ou seja, se as ações que a compõem foram executadas.

Ao verificar que isso aconteceu ele as retira do GR₂ e da PEF, busca a ação que agora está em foco (*Obter_Fatos_Apos(S, 14)*) e a executa, continuando a narrar os fatos acontecidos após as 14 horas.

Dessa forma o sistema consegue tratar de diálogos que sejam seqüências de ordens, tratando não somente do referenciamento de pronomes, como na quarta fala do usuário, mas também da manutenção do foco da atenção dos participantes, o que pode ser visto na segunda e quinta falas do usuário.

6.4.2 Diálogo II

O objetivo deste diálogo é mostrar o tratamento de sub-diálogos que surgem para completar uma determinada informação que falta (chamados por Lochbaum de sub-diálogos de pré-condição de conhecimento¹), ou para corrigir problemas relacionados à tarefa executada (sub-diálogos de correção²).

Ao digitar sua primeira frase

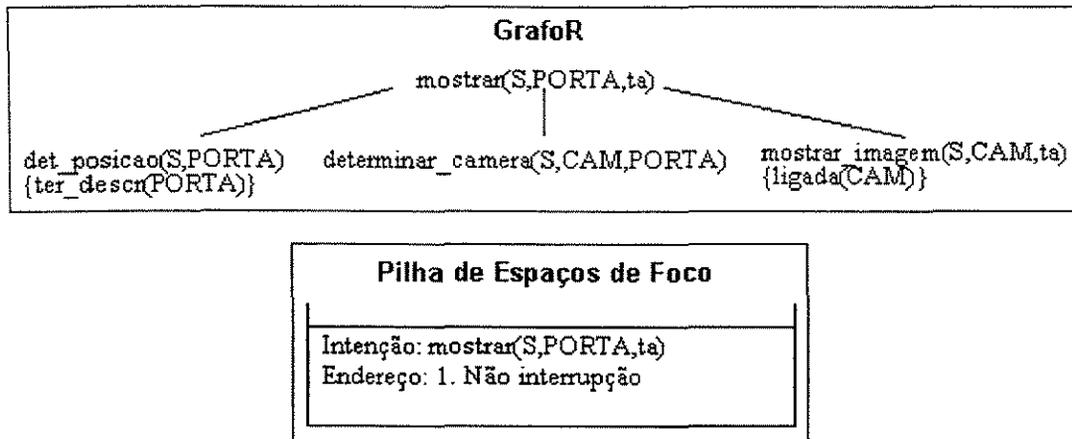
- *Mostre a porta*

o usuário passa ao sistema a proposição *deseja(U, mostrar(S, PORTA, ta))*, onde "ta" é o tempo atual do sistema. O sistema então busca uma receita para o ato de mostrar e, ao recebê-la, inclui a mesma no RG e empilha na PEF (ambos inicialmente vazios).

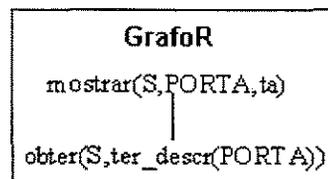
Em seguida, por se tratar de um desejo do usuário, o sistema tenta executar a receita. Para tal, ele verifica, para cada ato na receita, se suas restrições podem ser satisfeitas. Neste caso ele falha em satisfazer *ter_descr(PORTA)*, uma vez que PORTA não é uma entidade reconhecida.

¹Knowledge precondition subdialogue

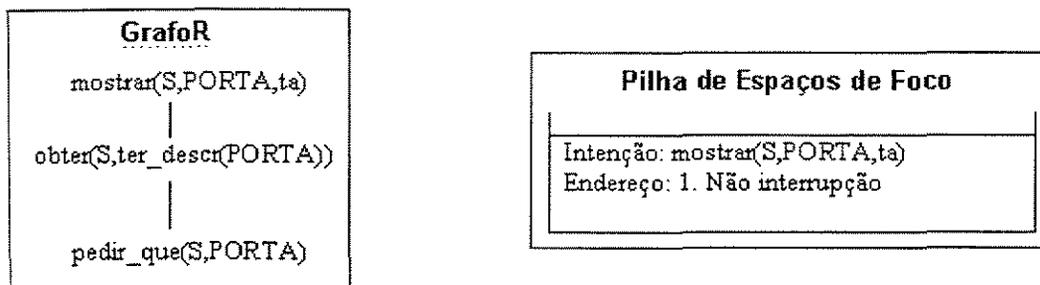
²Correction subdialogue



O sistema então retira do GR esta receita, incluindo a ação de *obter*($S, ter_descr(PORTA)$), ou seja, a receita para executar a ação *mostrar*($S, PORTA, ta$) passa pela satisfação desta restrição.



O próximo passo seguido pelo sistema é tentar executar esta nova ação. Para tal ele busca uma receita para a ação "obter" e a inclui no GR. Vale notar que nada foi empilhado na PEF, uma vez que o foco ainda está na ação de mostrar.

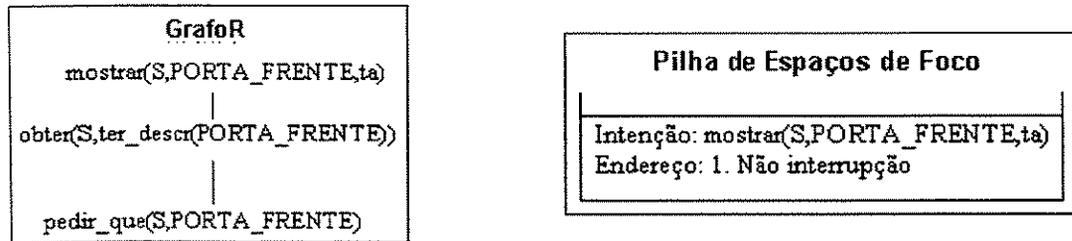


Então, finalmente, o sistema executa a ação *pedir_que*, pedindo ao usuário a que porta este se refere. O usuário então dá a resposta:

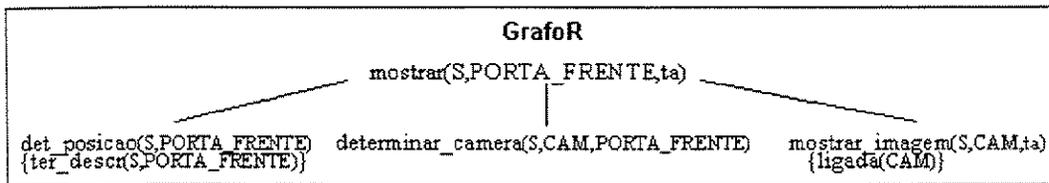
- *A da frente.*

Esta frase é entendida como *comunica*($U, PORTA_FRENTE$) (tendo sido a elipse resolvida quando da formação da frase lógica).

De posse da proposição, o sistema busca na PEF a ação que está em foco e, inspecionando o GR identifica a ação que ainda está incompleta. Ao encontrar esta ação o sistema substitui a entidade antiga pela nova, atualizando tudo no GR e na PEF.

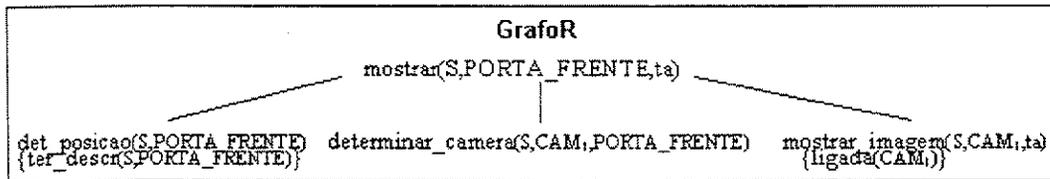


Então, o sistema desempilha a sub-árvore com raiz em *obter(S, ter_descr (PORTA_FRENTE))*, buscando novamente uma receita para o ato de mostrar e incluindo-a novamente no GR.



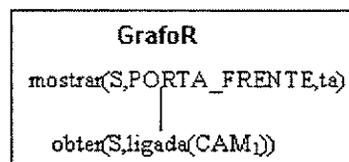
Mais uma vez o sistema executa a receita e, desta vez, PORTA_FRENTE é uma entidade conhecida.

Quando da execução de *determinar_camera(S, CAM, PORTA_FRENTE)* o sistema verifica que a câmera a ser usada é a CAM_1 , atualizando então o grafo para refletir este fato.

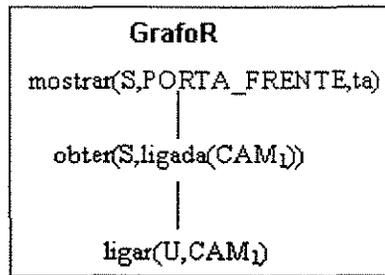


Antes da execução de *mostrar_imagem(S, CAM, ta)*, o sistema, ao testar suas restrições, verifica que *ligada(CAM₁)* não foi satisfeita, ou seja, a câmera 1 não está ligada.

Como antes, para resolver este problema o sistema retira do GR a receita de *mostrar_imagem(S, CAM₁, ta)*, adicionando *obter(ligada(CAM₁))*.



Em seguida, busca uma receita para $obter(S, ligada(CAM_1))$, incluindo-a no grafo.

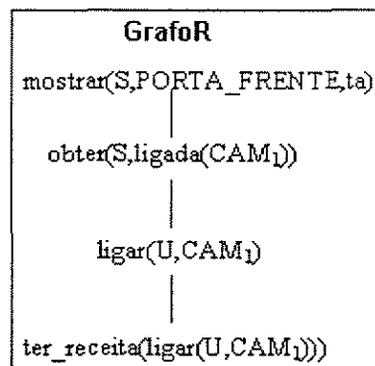


O sistema, então, verifica que não pode executar a ação, visto que *ligar* deve ser executada pelo usuário. Então ele avisa este fato ao usuário, dizendo "Não posso. A câmera está desligada. Você terá que ligá-la", e atualizando os centros de busca com a entidade CAM_1 (que, na representação usada pela Estrutura Lingüística, é $CAMERA(1, PORTA(1, Frente(1,)))$).

Em resposta, o usuário pede uma receita de como executar tal ato.

- Como?

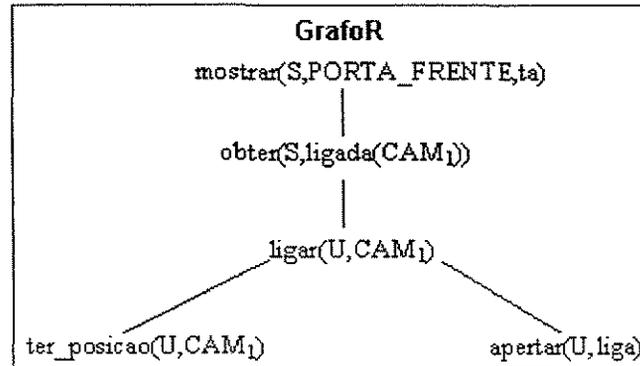
O que é entendido como $deseja_p(U, ter_receita(X))$. Para instanciar a variável X, o sistema busca na PEF a ação que está correntemente em foco e, ao encontrá-la, busca sua receita no GR. De posse da receita, o sistema busca (em profundidade) a primeira ação incompleta e instancia X com esta ação. A proposição lógica torna-se então $deseja_p(U, ter_receita(ligar(U, CAM_1)))$. $Ter_receita(ligar(U, CAM_1))$ é então incluída no grafo como filha da ação $ligar(U, CAM_1)$.



Finalmente, o sistema executa $ter_receita(ligar(U, CAM_1))$, buscando uma receita para $ligar(U, CAM_1)$ e a repassando ao usuário:

- Vá até a sala, ela está em cima da lareira. Aperte o botão "ligar".

Para repassar a receita ao usuário, o sistema primeiro a incluiu no grafo (no lugar de $ter_receita(ligar(U, CAM_1))$) e a executa. Como cada ação da receita deveria ser feita pelo usuário (pelo fato de $ligar$ ser executada também por este), o sistema se limita a narrar cada ato da receita, atualizando os centros de busca conforme a necessidade.



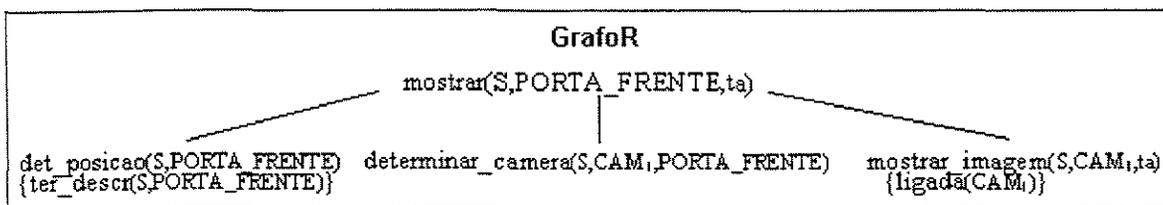
$Ter_posicao(U, CAM_1)$ é uma ação básica entendida como sendo "o usuário deve ter a posição de CAM_1 ". Por isso o sistema determina a posição desta e a passa ao usuário.

O usuário, então, executa sua parte e dá uma resposta indicando isto:

- OK. Feito.

Esta fala é entendida como um "ok" (como no diálogo I). O sistema, então, busca a ação que está em foco, verifica a última que ainda não foi executada, $ligar(U, CAM_1)$, e, como é uma ação para o usuário executar, a toma como tendo sido executada. Ou seja, o "ok" do usuário é entendido como um "fiz minha parte, prossiga". Dado que esta é a única ação na receita de $obter(S, ligada(CAM_1))$, esta também é tida como terminada.

Por se tratar de uma ação do tipo $obter(restrição)$, esta é retirada do grafo e a receita para $mostrar(S, PORTA_FRENTE)$ é reintroduzida e novamente executada.



Desta vez todas as restrições foram satisfeitas e o sistema pode finalmente mostrar a imagem requisitada pelo usuário.

Este exemplo mostrou como se dá o tratamento de sub-diálogos de correção, como o iniciado na segunda fala da Figura 6.2, e de sub-diálogos de pré-condição de conhecimento, como o iniciado na quarta fala desta mesma figura. Além disso, serviu para demonstrar o referenciamento de elipses (segunda fala do usuário) e a manipulação do foco do sistema que, neste caso, nunca se distanciou do ato de mostrar uma determinada porta.

Capítulo VII

Conclusão

Neste trabalho foi desenvolvido um sistema para o processamento computacional de diálogos orientados à tarefa em uma língua natural. Tal sistema teve como objetivo estudar e implementar a teoria de discurso desenvolvida por Grosz e Sidner (1985, 1986), a Teoria de Centralização apresentada por Grosz (1978) e o modelamento para a Estrutura Intencional introduzido por Lochbaum (1994), buscando também o estudo e implementação da integração entre estas teorias.

Além disso, o sistema também teve por objetivo implementar o modelamento para a Estrutura Lingüística desenvolvido neste trabalho, comparando seu comportamento com o esperado segundo o modelo teórico.

Os diálogos usados para testar o sistema no Capítulo 5 demonstram este comportamento, nos dando uma dimensão das características que o tornam atraente.

A atuação do sistema no referenciamento de pronomes e elipses, na própria extração do conteúdo proposicional das frases e na compreensão destas é bastante animador com relação a estudos futuros.

Deve ser ressaltado que, pelo fato de ter sido somente desenvolvido com de multiagentes, não há como dizer se o sistema funcionaria melhor com ou sem eles, uma vez que não há como comparar o sistema aqui desenvolvido com um criado usando-se outro paradigma.

No entanto, o que se pode dizer é que o uso do enfoque multiagentes quando do planejamento do sistema (especialmente da Estrutura Lingüística) permitiu que este pudesse ser visualizado com mais clareza, uma vez que, ao distribuir as competências entre seus agentes, foi possível separar cada um dos sub-problemas que surgiam a partir do problema geral e tratá-los isoladamente, sem haver uma preocupação com o resto do sistema.

Com relação à velocidade de processamento, houve sim uma perda pelo uso de multiagentes, uma vez que eles precisam se comunicar e gastam um certo tempo nisso. Em contrapartida, houve um ganho enorme em mobilidade, isto é, este sistema não precisa es-

tar rodando numa única máquina. Ele pode estar espalhado em várias máquinas, através de uma rede. De fato, cada agente pertencente ao sistema poderia ser rodado isoladamente em uma máquina em separado, bastando para tal que suas mensagens fossem redirecionadas através da rede para a máquina que contém o agente que serve de "central de comunicações".

Além disso, a teoria de planos compartilhados usada por Lochbaum para modelar a Estrutura Intencional foi criada para modelar colaboração entre sistemas multiagentes, donde se pode concluir que o uso destes para modelar este sistema é bastante aceitável.

7.1 Em Aberto

As principais questões que ficaram em aberto neste trabalho são:

- Um estudo mais a fundo do referenciamento de pronomes e elipses: mais especificamente, como tratar uma fala onde o pronome faz referência a uma entidade não mencionada, mas a ser mencionada, como na frase "eu a vi, Ana, esta manhã". Além disso, a única elipse detectada pelo sistema é a elipse de substantivo. Outros tipos de elipse precisam ser melhor estudados.
- O dicionário precisa conter informação semântica para cada palavra: isso tornaria possível a resolução de ambigüidades. Por exemplo, como está, o sistema representa a frase "a bandeira viu João" literalmente, ou seja, que a entidade bandeira realmente viu a entidade João. Se alguma informação semântica fosse utilizada (como por exemplo, se o objeto é animado ou inanimado), essa interpretação seria desconsiderada.
- O referenciamento de elipses é simplista: em um diálogo do tipo:
 - Mostre a casa branca.
 - Qual?
 - A da esquina.

O sistema vai entender como "a casa da esquina", e não "a casa branca da esquina". Novamente, um estudo mais aprofundado de elipses deve ser levado adiante.

- A geração de respostas deve ser melhor trabalhada: no momento, a geração é feita pelas funções básicas da receita que têm que responder algo ao usuário. Para tal, essas funções simplesmente traduzem a proposição lógica em palavras, sem fazer uso da Estrutura Atencional, o que permitiria a inclusão de, por exemplo, pronomes na resposta, tornando-a mais coerente e natural.

- Sinais de pontuação devem ser estudados: até agora, apenas pontos finais, de exclamação e interrogação são levados em conta. É necessário um estudo do uso da vírgula, por exemplo, e de como ela pode ser tratada. Além disso, apenas orações simples são entendidas. Se algo mais complexo, repleto de sinais de pontuação, for usado como entrada, o sistema não conseguirá interpretar corretamente.
- A interpretação correta de pronomes, conjunções e locuções: pronomes do tipo relativo não foram tratados a contento, justamente pela falta de uma melhor caracterização de seu uso. Tal caracterização também vale para conjunções e para a detecção de locuções adverbiais e adjetivas.
- Uma melhor integração das lógicas deve ser feita: a tradução da lógica usada pela Estrutura Lingüística para a forma lógica usada por Lochbaum é feita por mapeamento direto. É necessário o desenvolvimento de funções que permitam esta tradução de modo mais flexível e aceitável.
- O uso de marcas de discurso: marcas de discurso não são efetivamente entendidas pelo sistema (com exceção do "OK" e "Certo" ou "Feito"), um estudo melhor de que tipo de marcas de discurso existem no português se faz necessário para melhorar a compreensão do discurso pelo sistema.
- Teste do sistema em domínios diferentes: por escassez de tempo, não foi possível testar este sistema em outros domínios, ou seja, mudar os arquivos da estrutura da tarefa e ver como o sistema se comporta em um novo ambiente.
- Incluir o abastecimento de receitas em língua natural: seria interessante se o sistema pudesse ser abastecido pelo usuário com novas receitas, apresentadas em língua natural. Assim, o usuário descreveria uma nova função em termos das ações básicas executadas pelo sistema e este, por sua vez, abasteceria o seu livro de receitas, como se fosse uma espécie de programação de uma sub-rotina em língua natural.
- São estas ilhas suficientes? Finalmente, é necessário um amadurecimento maior do modelo de ilhas, para saber se as que foram utilizadas são suficientes, se está faltando ou sobrando alguma. Essa conclusão somente poderá vir de um estudo aprofundado da língua portuguesa, a ser efetuado futuramente.

Apêndice A

As Comunicações Entre os Agentes

Entrada	Saída	Ações
{d frase do usuário}	{d <sub-frase><flags>}	É a entrada do usuário, processa e envia a saída definida.
{D }	{d <sub-frase><flags>}	Pedido de nova sub-frase, envia ao EL a resposta.
	{E }	Erro. Ocorre caso tenha havido um pedido de sub-frase e ela não exista.
	{e }	Erro de sistema.

Tabela A.1: Comunicação entre EL1 e EL.

Entrada	Saída	Ações
{d <frase do EL1> <flags>}	{d <[pal ₁ c ₁]...[pal _n c _n]> <flags>}	Recebe a entrada de EL1 e gera a saída. Na saída, pal _i e c _i representam uma palavra da frase e sua classificação.
{D }	{d <[pal ₁ c ₁]...[pal _n c _n]> <flags>}	Pedido de nova interpretação, envia ao EL a resposta.
	{E }	Erro. Ocorre caso tenha havido um pedido de nova interpretação e ela não exista.
	{e }	Erro de sistema.

Tabela A.2: Comunicação entre EL2 e EL.

Remetente	Campo "A"	Ação
Usuário	D	É uma nova fala, envia ao EL1.
Estrutura da Tarefa	Não importa	Repassa ao EL2 a mensagem.
Estrutura Atencional	Não importa	É resposta a algum pedido, repassa a quem enviou tal pedido.
EL1	E	Erro, envia mensagem ao usuário dizendo que a frase não foi compreendida.
EL1	D	Dados, sub-fala que deve ser enviada ao EL2.
EL1	E	Erro de sistema no processamento de algum dado. Reinicia.
EL2	E	Erro, envia mensagem ao usuário dizendo que a frase não foi compreendida.
EL2	D	Dados. Envia ao EL3.
EL2	T	Pedido para Estrutura da Tarefa, envia ao COM.
EL2	E	Erro de sistema no processamento de algum dado. Reinicia.
EL3	E	Erro, envia um {D } ao EL2 pedindo nova fala.
EL3	P	Prontidão, remete ao EL4.
EL3	D	Dado, repassa ao EL4.
EL3	A	Pedido ao Estado Atencional, remete ao COM.
EL3	E	Erro de sistema no processamento de algum dado. Reinicia.
EL4	E	Erro, envia um {D } ao EL2 pedindo nova fala.
EL4	D	Dado para a Estrutura Intencional. Substitui o d por um I e envia ao COM.
EL4	D	Pedido, envia ao EL3.
EL4	O	Ordem, envia ao EL3.
EL4	A	Dados para a Estrutura Atencional, envia ao COM.
EL4	E	Erro de sistema no processamento de algum dado. Reinicia.

Tabela A.3: As ações tomadas pelo EL, de acordo com a comunicação recebida.

Entrada	Saída	Ações
{d <[pal ₁ c ₁]...[pal _n c _n]>><flags>	{P }	Recebe a entrada do EL2 e, após processá-la, envia o sinal de prontidão ao EL4.
{D @}	{d @} ou {d \$}	Recebe um pedido do EL4 querendo saber se há alguma ilha sobrando. Se não houver envia um @, se houver, envia \$.
{D flags}	{d flags da frase}	O EL4 pede os flags da frase.
	{d N}	Enviado quando o agente não encontrou a resposta a algum pedido a ele feito.
{D dev:ilha}		Devolve a ilha à lista de não enviadas.
{D desc:ilha}		Inclui a ilha na lista de ilhas enviadas, retirando da lista de não enviadas.
{D [palavra]}	{d ilha#pos}	Devolve a primeira ilha contendo "palavra".
{D classificação}	{d ilha#pos}	Devolve a primeira ilha que contenha uma palavra com essa classificação sintática. A classificação pode ser parcial: assim, um "a-f" retorna a primeira ilha contendo um artigo feminino, não importando o gênero.
{D OD}	{d ilha#pos}	Retorna a primeira ilha Objeto Direto encontrada.
{D OI}	{d ilha#pos}	Retorna a primeira ilha Objeto Indireto encontrada.
{D #pos}	{d ilha#pos}	Retorna a ilha existente na posição pos.
	{A E:clas}	Busca na EA a referência a uma entidade de classificação sintática clas para resolver uma elipse.
{A [pal clas]}		Mensagem da EA dando a palavra e sua classificação sintática completa encontrada para referenciar a elipse.
{A N}		Mensagem da EA, avisando que nenhuma referência foi encontrada.
	{E }	Erro. O agente COM transforma esse erro num {D } e envia ao EL2, pedindo nova interpretação.
	{e }	Erro de sistema.

Tabela A.4: Comunicação entre EL3 e outros agentes.

Entrada	Saída	Ações
	{d frase lógica}	Envia à EI a correspondente lógica da oração. O agente EL transforma o "d" num "A" antes de enviar ao COM.
	{A ent:E ₁ @E ₂ ... E _n }	Envia à EA as entidades mais destacadas na oração, E ₁ é o sujeito, e as demais E _i são as entidades conforme ordem superficial na oração. Caso não haja sujeito, E ₁ é omitido.
	{A P:clas} ou {A N}	Busca na EA a referência a uma entidade de classificação sintática clas para resolver um pronome. Um "N" é enviado se nada for encontrado.
{A entidade}		Mensagem da EA dando a entidade lógica encontrada para referenciar o pronome.
{A N}		Mensagem da EA, avisando que nenhuma referência foi encontrada.
	{E }	Erro. O agente COM transforma esse erro num {D } e envia ao EL2, pedindo nova interpretação.
	{e }	Erro de sistema.

Tabela A.5: Comunicação entre EL4 e outros agentes.

Entrada	Saída	Ações
{A M:macro}	{A classificação} ou {A N}	Recebe da EA o pedido de uma macro; após buscar no dicionário envia a classificação da palavra correspondente a esta, ou N se não achar.
{L <palavra>}	{L [c ₁][c ₂]...[c _n]} ou {L N}	A EL busca uma palavra. O agente envia as possíveis classificações sintáticas dessa palavra à EL, ou N se não achar.
{I ação}	{I receita} ou {I N}	Busca no livro de receitas uma receita para a ação pedida pela EI. Se encontrar envia no formato {s ₁ r ₁ }...{s _n r _n }, onde s _i é uma sub-ação pertencente à receita e r _i suas restrições. Caso não encontre receita, envia um N.
	{e }	Erro de sistema.

Tabela A.6: Comunicação entre a ET e as outras estruturas.

Entrada	Saída	Ações
{L P:clas}	{L entidade} ou {L N}	Recebe da EL o pedido de um pronome com classificação sintática clas. Busca e se encontrar devolve a entidade, senão envia um N.
{L E:clas}	{L clas. compl.} ou {L N}	A EL busca uma elipse com classificação sintática clas. O FI busca a referência e devolve sua classificação sintática completa ou N, se não encontrar.
{L ent:E ₁ @E ₂ ... E _n }		A estrutura Lingüística envia as entidades mais salientes na frase. E ₁ é o sujeito, as demais vêm na mesma ordem que estão na oração.
	{T M:macro}	Envia à ET o pedido para que devolva a classificação sintática da palavra correspondente a essa macro.
{T classificação}		Contém a classificação sintática da palavra pesquisada pelo FI. É resposta à ação acima tomada.
{d E ₁ ... E _n }		Envia para o FG as entidades salientes de uma fala, conforme presentes no C _p desta. É usada quando a pilha de centros já possui três elementos e um novo deve ser empilhado, ou quando uma ordem de empilhar vem da EI, o que força o envio de toda a pilha de centros para a pilha de espaços de foco.
{I E:ação end int}		Ordem de empilhamento vinda da EI. O FI deve enviar ao FG a pilha de centros e repassar a ele a ordem.
{I D:ação}		Ordem de desempilhamento vinda da EI. O FI deve esvaziar a pilha de centros e repassar a ordem ao FG.
{I EF}	{I entidade} ou {I N}	O EI busca a entidade em foco.
	{e}	Erro de sistema.

Tabela A.7: Comunicação entre o FI e as outras estruturas e agentes.

Entrada	Saída	Ações
{i P:clas}	{i entidade} ou {i N}	Recebe do FI um pedido para o referenciamento de um pronome. Busca tal referência (consultando a ET) e devolve a entidade lógica, ou N se não encontrar.
{d centro progressivo}		O FI enviou uma ordem de abastecimento das entidades relevantes em uma fala.
	{T M:macro}	Envia à ET o pedido para que devolva a classificação sintática da palavra correspondente a esta macro.
{T classificação}		Contém a classificação sintática da palavra pesquisada pelo FG. É a resposta à ação acima tomada.
{I F}	{I ação end int} ou {I N}	A EI está buscando a ação que está em foco no momento. O FG retorna a ação, seu endereço no grafo e um indicativo de ser ou não interrupção, ou N senão houver tal ação.
{i E:ação end int}		Ordem de empilhamento de ação, vinda do FI
{i D:ação}		Desempilha ação, caso ela exista na pilha.
	{e}	Erro de sistema.

Tabela A.8: Comunicação entre o FG e as outras estruturas e agentes.

Entrada	Saída	Ações
	{A F}	Pede ao EA a ação que está atualmente em foco na pilha de espaços de foco.
{A ação end int}		A EA envia uma ação, seu endereço no GrafoR e um indicativo desta ser ou não parte de uma interrupção.
	{A ação end int}	O EI envia ao EA para que este abasteça a pilha de espaços de foco.
	{A E:ação end int}	O EI envia ao EA uma ordem de empilhamento de ação.
	{A EF}	Pede ao EA a entidade que está em foco.
{A entidade}		A EA retorna a entidade em foco.
	{T R:ato}	Pede à ET uma receita para ato.
{T {B ₁ r ₁ }...{B _n r _n }}		A ET envia as sub-ações (B _i) que formam a receita para o ato consultado e suas restrições (r _i).
{A N}		Indicador que a resposta ao pedido feito não foi encontrada.
	{A D:ação}	Envia ao EA um ordem para que ação seja desempilhada.
	{T r:ato}	Pede ao ET as restrições de ato. A resposta vem sob a forma {T {B r}} ou {T N}, caso ato não tenha restrições.
	{e }	Erro de sistema.

Tabela A.9: Comunicação entre o EI e outros agentes.

Bibliografia

- [1] Carberry (1993), Sandra; Chu, Jennifer; Lambert, Lynn. *User-System Conflict in Task-Oriented Consultation*. Em Proceedings of the 1993 IJCAI Workshop on Computational Models of Conflict Management. pp. 41-53. 1993.
- [2] Chu-Carroll (1993), Jennifer. *Responding to User Queries in a Collaborative Environment*. Em Proceedings ACL-93. pp. 280-282. 1993.
- [3] Chu-Carroll (1995a), Jennifer; Carberry, Sandra. *Conflict Detection and Resolution in Collaborative Planning*. Em Proceedings of the IJCAI-95 Workshop on Agent Theories, Architectures, and Languages. pp. 67-79. 1995.
- [4] Chu-Carroll (1995b), Jennifer; Carberry, Sandra. *Communication for Conflict Resolution in Multi-Agent Collaborative Planning*. Em Proceedings of the First International Conference on Multiagent Systems. pp. 49-56. 1995.
- [5] Chu-Carroll (1995c), Jennifer; Carberry, Sandra. *Generating Information-Sharing Subdialogues in Expert-User Consultation*. Em Proceedings for the 14th International Joint Conference on Artificial Intelligence. pp. 1243-1250. 1995.
- [6] Chu-Carroll (1996), Jennifer. *Response Generation in Collaborative Dialogue Interactions*. Em Proceedings of the Twente Workshops on Language Technology: Dialogue Management in Natural Language Processing. pp. 177-186. 1996.
- [7] Gordon (1993), Peter; Grosz, Barbara; Gillion, Laura. *Pronouns, Names, and the Centering of Attention in discourse*. Em Cognitive Science, 3(17). pp. 311-347. 1993.
- [8] Grosz, (1978), Barbara. *Focusing in Dialog*. Em Proceedings of the Workshop on Theoretical Issues in Natural Language Processing-2. 1978.
- [9] Grosz (1981), Barbara. *Focusing and Description in Natural Language Dialogues*. Em Elements of Discourse Understanding. A. K. Joshi, I. Sag e B. Webber (eds.). Cambridge University Press. Cambridge, Inglaterra. 1981.

- [10] Grosz (1985), Barbara; Sidner, Candace. *Discourse Structure and the Proper Treatment of Interruptions*. Em Proceedings of the 9th International Joint Conference on Artificial Intelligence. 1985.
- [11] Grosz (1986), Barbara; Sidner, Candace. *Attention, Intentions, and the Structure of Discourse*. Em Computational Linguistics, 12(3). pp 175-204. 1986.
- [12] Grosz (1990), Barbara; Sidner, Candace. *Plans for Discourse*. Em Intentions in Communication. P. R. Cohen, J. L. Morgan e M. E. Pollack (eds.). MIT Press, Cambridge, MA. pp 417-444. 1990.
- [13] Grosz (1994), Barbara. *Collaborative Systems*. Em AAAI-94 Presidential Address, 2(17). pp. 67-85. 1994.
- [14] Grosz (1995), Barbara; Joshi, Aravind; Weinstein, Scott. *Centering: A framework for Modeling the Local Coherence of Discourse*. Em Computational Linguistics, 2(21). Pp. 203-225. 1995.
- [15] Grosz (1996), Barbara; Kraus, Sarit. *Collaborative Plans for Complex Group Action*. Em Artificial Intelligence, 86(2). Pp. 269-357. 1996.
- [16] Grosz (1998), Barbara; Ziv, Yael. *Centering, Global Focus, and Right Dislocation*. Em Centering in Discourse. M. Walker, A. Joshi e E. Prince (eds.). Oxford University Press. Pp. 293-307. 1998.
- [17] Grosz (1999a), Barbara; Sidner, Candace. *Lost Intuitions and Forgotten Intentions*. Em Centering in Discourse. M. Walker, A. Joshi e E. Prince (eds.). Oxford University Press. Pp. 39-51. 1999.
- [18] Grosz (1999b), Barbara ; Kraus, Sarit. *Planning and Acting Together*. Em AI Magazine, 4(20). Edição especial em Planejamento Contínuo Distribuído. 1999.
- [19] Grosz (1999c), Barbara; Kraus, Sarit. *The Evolution of Shared Plans*. Em Foundations and Theories of Rational Agencies. A. Rao e M. Wooldridge (eds.). Pp. 227-262. 1999.
- [20] Joshi (1979), Aravind; Kuhn, Steve. *Centering Logic: The Role of Entity Centered Sentence Representation in Natural Language Inferencing*. Em Proceedings of the International Joint Conference on Artificial Intelligence. pp. 435-439. 1979.
- [21] Lochbaum (1993), Karen. *A Collaborative Planning Approach to Discourse Understanding*. Technical Report TR-20-93, Universidade de Harvard. Cambridge, MA. 1993.

- [22] Lochbaum (1994), Karen. *Using Collaborative Plans to Model the Intentional Structure of Discourse*. Technical Report TR-25-94. Universidade de Harvard. Tese de Doutorado. Cambridge, MA 1994.
- [23] Maes (1995), Pattie. *Modeling Adaptive Autonomous Agents*. Em *Artificial Life, An Overview*. C. Langton (ed.). MIT Press. Cambridge, MA. 1995.
- [24] Mann (1975), William; Moore, James; Levin, James; Carlisle, James. *Observation Methods for Human Dialogue*. Technical Report RR/75/33. Information Science Institute. Marina del Rey. CA. 1975.
- [25] Russell (1995), Stuart; Norvig, Peter. *Artificial Intelligence: A Modern Approach*. Prentice Hall Series in Artificial Intelligence. Prentice Hall. 1995.
- [26] Wilson (1971), Edward. *The Insect Societies*. Harvard University Press. Cambridge, MA. 1971.