



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO DE PESQUISAS ESPACIAIS

AUTORIZAÇÃO PARA PUBLICAÇÃO
AUTHORIZATION FOR PUBLICATION

AUTORES AUTHORS	PALAVRAS CHAVES/KEY WORDS		AUTORIZADA POR/AUTHORIZED BY	
	LINGÜÍSTICA LINGUAGEM NATURAL INTELIGÊNCIA ARTIFICIAL		Eduz Gylvan M. Filho Diretor Meteorologia	
AUTOR RESPONSÁVEL RESPONSIBLE AUTHOR		DISTRIBUIÇÃO/DISTRIBUTION		REVISADA POR/REVISED BY
Carlos A. Oliveira		☐ INTERNA / INTERNAL ☒ EXTERNA / EXTERNAL ☐ RESTRITA / RESTRICTED		Francisco E. C. Viola
CDU/UDC			DATA/DATE	
681.3.019			Maio, 1990	
TÍTULO/TITLE	PUBLICAÇÃO Nº PUBLICATION NO		ORIGEM ORIGIN	
	INPE-5064-PRE/1590		MET	
LINGÜÍSTICA TEXTUAL E REPRESENTAÇÃO DO CONHECIMENTO: UM ENFOQUE		PROJETO PROJECT		
		CAPL		
		Nº DE PAG. NO OF PAGES	ULTIMA PAG. LAST PAGE	
		25	18	
		VERSÃO VERSION	Nº DE MAPAS NO OF MAPS	
AUTORES/AUTHORSHIP	Carlos A. de Oliveira			
RESUMO - NOTAS / ABSTRACT - NOTES				
Procede-se a uma aproximação entre Processamento de Linguagem Natural e Lingüística Textual, através do estudo de problemas de representação de conhecimento.				
OBSERVAÇÕES/REMARKS				
Apresentado no Congresso de Lingüística do Grupo de Estudos Lingüísticos (GEL)/89, em Lorena, SP.				



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO DE PESQUISAS ESPACIAIS

PROPOSTA PARA
PUBLICAÇÃO

- ☐ DISSERTAÇÃO
☐ TESE
☐ RELATÓRIO
☒ OUTROS

TÍTULO

Linguística textual e representação de conhecimento:
um enfoque

AUTOR(ES)

Carlos Alberto de Oliveira

ORIENTADOR

CO-ORIENTADOR

IDENTIFICAÇÃO

DISS. OU TESE

LIMITE

DEFESA

CURSO

ORGÃO

DIVULGAÇÃO

☒ EXTERNA ☐ INTERNA ☐ RESTRITA

EVENTO/MEIO

☒ CONGRESSO ☐ REVISTA ☐ OUTROS

NOME DO REVISOR

Francisco E. C. Viola

RECEBIDO

DEVOLVIDO

ASSINATURA

REV. TÉCNICA

NOME DO RESPONSÁVEL

Francisco E. C. Viola

APROVADO

DATA

ASSINATURA

☐ SIM
☐ NÃO

— / — / —

APROVAÇÃO

Nº

PRIOR.

RECEBIDO

NOME DO REVISOR

REV. LINGUAGEM

PÁG.

DEVOLVIDO

ASSINATURA

OS AUTORES DEVEM MENCIONAR NO VERSO INSTRU-
ÇÕES ESPECÍFICAS, ANEXANDO NORMAS, SE HOUVER

RECEBIDO

DEVOLVIDO

NOME DA DATILÓGRAFA

DATILOGRAFIA

Nº DA PUBLICAÇÃO:

PÁG.:

CÓPIAS:

Nº DISCO:

LOCAL:

AUTORIZO A PUBLICAÇÃO

☐ SIM

☐ NÃO

— / — / —

DIRETOR

OBSERVAÇÕES E NOTAS

OBS: * para ser submetido ao ^{congresso de linguística do} Grupo de Estudos
Linguísticos (GEL) / 89 (resumo já aceito),

RESUMO: procede-se a uma aproximação entre procedimento
de linguagem natural e linguística textual,
através dos estudos de problemas de
representação de conhecimento.

ABSTRACT

This work discusses a possible link between Text Linguistics and Natural Language Processing: the knowledge representation problems studies.

SUMÁRIO

	<u>Pág.</u>
1 INTRODUÇÃO	1
1.2 - Considerações sobre o Processo de Tradução	2
2 CONSIDERAÇÕES SOBRE REPRESENTAÇÃO DE CONHECIMENTO .	4
2.1 - Considerações Iniciais	4
2.2 - Propriedades, Tipos e Esquemas de Representação..	7
3 CONSIDERAÇÕES GERAIS SOBRE LINGÜÍSTICA TEXTUAL E PLN	12
4 À GUIA DE CONCLUSÃO	14
REFERÊNCIAS BIBLIOGRÁFICAS	17

LISTA DE FIGURAS

	<u>Pág.</u>
1 - A relação sígnica e o processo de codificação/deco dificação	3
2 - Possíveis relações significativas de "palavras"den tro do mesmo e em diferentes domínios	3
3 - A tradução inter-códigos.....	4
4 - Algoritmo para identificação de classe morfológi ca de lexias	5
5 - Sugestão de representação para conhecimento morfo lógico	6
6 - Dois exemplos de rede semântica	8
7 - Uma representação DC para frase "João comprou pão (do padeiro)".....	9
8 - Um "script" elementar e simplificado para uma via gem de Ônibus	10
9 - Um "frame" elementar para a descrição de uma pes soa	11
10 - Definitivação: algumas regras de produção	14
11 - Sugestão para uma (de)codificação textual	15

LINGÜÍSTICA TEXTUAL E REPRESENTAÇÃO DE CONHECIMENTO: UM
ENFOQUE

Carlos Alberto de Oliveira
Instituto de Pesquisas Espaciais - INPE
Diretoria de Meteorologia - Núcleo de Computação
Linha de pesquisa: Inteligência Artificial
(Processamento de Linguagem Natural)
12200 - São José dos Campos - SP

1. INTRODUÇÃO

Os trabalhos na área de Processamento de Linguagem Natural (PLN) [1] têm como objetivo geral tornar os computadores aptos para a comunicação em Linguagem Natural (LN), isto é, torná-los capazes de compreender [2] e/ou produzir lexias, frases isoladas e/ou textos em dada(s) língua(s). Se, historicamente, em PLN, a Língua foi tomada como um conjunto de lexias e depois como um rol de frases, atualmente, na fase cognitiva, iniciando-se a fase dos problemas de representação do conhecimento, ela é tomada como um texto.

Esta fase abre um horizonte não explorado, onde o papel do usuário começa a ser levado em conta, no que concerne, por exemplo, à sua intencionalidade e à manei

1 No domínio da Lingüística, PLN é sucedâneo de Lingüística Computacional.

2 "Diz-se que um enunciado é compreendido quando a resposta do interlocutor na comunicação instaurada pelo locutor está de acordo com o que esse último dela espera, quer a resposta seja um enunciado ou um ato." (Dubois et alii, 1978, p. 128).

ra como, possivelmente, ele "apreende, armazena" e "manipula" conhecimento. Enquanto nas fases que antecederam a esta, dicionários e gramáticas aprioristicamente determinados se prestavam ao objetivo de "compreender" palavras e/ou frases isoladas, nesta fase atual do PLN, tendo-se a inter e a multidisciplinariedade como características principais, exige-se uma integração entre diversos modelos (lingüísticos, psico-cognitivos, lógicos, dentre outros) para a consecução do objetivo de aprender, compreender e/ou produzir LN.

1.2 - CONSIDERAÇÕES SOBRE O PROCESSO DE TRADUÇÃO

Os seres humanos comunicam-se através de sistemas de signos e, em especial, do sistema de signos lingüísticos. Este, o signo lingüístico, conforme Couto (1983), é uma realidade bifacial, isto é, consta de uma Expressão (dado concreto, sensorial) e de um Conteúdo (dado abstrato, conceitual), os quais estão em relação. Segue-se, pois, que a própria relação é um dado, ou seja, a significação. Através dessa relação é que se pode fazer o trânsito Expressão-Conteúdo, ou melhor, pode-se codificar/decodificar o signo, isto é, traduzi-lo (Figura 1).

Há que se acrescentar, então, que o processo de tradução (codificação/decodificação) da relação que une Expressão e Conteúdo pode ser intra-código, isto é, as relações (significações) valem dentro do código no qual o signo está inserto: dentro de cada domínio (código), um elemento pode encetar diferentes relações com os demais elementos desse mesmo domínio. Vide, por exemplo, a tradução que a lexia começasse necessita para ser compreendida pelos (sub)domínios morfológico e sintático (Figura 2A).

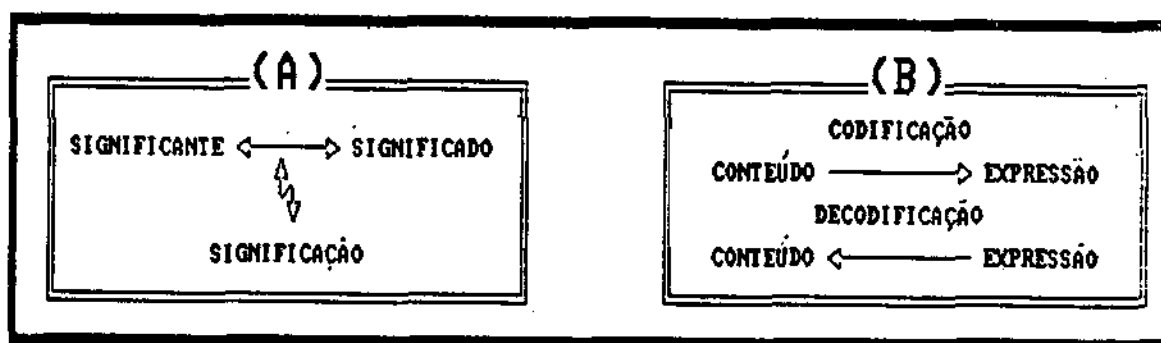


Fig. 1 - A relação sênica e o processo de codificação/decodificação.

-
- (A) COMEÇASSEM --) Lexema + VT + DMT + DNP
 ---) Classe Morfológica + PN + Tempo
- (B) SOL (----- antonímia -----) CHUVA
- (C) SOL ----) particularização_de -----) CORPO CELESTE
 LUA ----) particularização_de -----) CORPO CELESTE
- (D) SOL (----- sinonímia -----) DIA
 LUA (----- sinonímia -----) NOITE
 SOL (----- antonímia -----) LUA
-

Fig. 2 - Possíveis relações significativas de "palavras" dentro do mesmo e em diferentes domínios.

Na Figura 2B, para o elemento Sol: no domínio dos fenômenos meteorológicos, por exemplo, pode estabelecer uma relação de antonímia com Chuva (~ fazer sol(-)chover); no domínio de astronomia, por exemplo (Figura 2C), a relação pode se estabelecer de

maneira indireta com Lua, ou seja, ambos são particularizações de uma classe de objetos; no domínio de divisão do espaço de tempo 24 horas, por exemplo (Figura 2D), as relações estabelecidas são de antonímia com Lua e de sinonímia com Dia. Pode-se notar que, conforme mudam-se os domínios, mudam-se também as relações que dão significação a dada lexia. Esta é a tradução inter-códigos, devendo-se estabelecer, dentro de um conjunto de domínios, um conjunto de relações que delimitem e permitam o trânsito entre tais domínios (Figura 3).

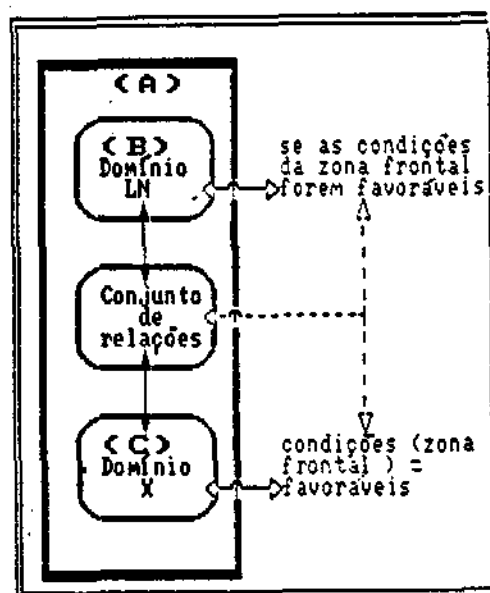


Fig 3 - A tradução inter-códigos.

2. CONSIDERAÇÕES SOBRE REPRESENTAÇÃO DE CONHECIMENTO.

2.1 CONSIDERAÇÕES INICIAIS

Tome-se por exemplo, em PLN, um algoritmo bem elementar que se presta à análise morfológica, conforme Figura 4.

```

SEGMENTAÇÃO(Provável_lexema,Resto_da_cadeia):-
    TIRAR_LETRA(Resto_da_cadeia,Letra,Nova_cadeia),
    INSERIR_LETRA(Provável_lexema,Letra,Novo_lexema),
    Se VERIFICAR_RESTO(Nova_cadeia,Dados) tiver sucesso
    Então VERIFICAR_LEXEMA(Novo_lexema,Dados)
    Caso contrário SEGMENTAÇÃO(Novo_lexema,Nova_cadeia).

```

Fig. 4 - Algoritmo para identificação de classe morfológica de lexias.

Pela Figura 4, um Provável_lexema (vazio) e o Resto_da_cadeia (lexia) entram como argumento da função SEGMENTAÇÃO. É segmentada a primeira letra do Resto_da_cadeia, retornando-se a Letra e a Nova_cadeia (sem a letra retirada). A seguir, insere-se a Letra no fim da cadeia do Provável_lexema, gerando-se um Novo_lexema. Se a Nova_cadeia existir como conhecimento, dela são retirados os Dados pertinentes para classificar a lexia em análise, verificando-se a seguir se o Novo_lexema já existe como conhecimento. Se não existir como conhecimento a Nova_cadeia, chama-se recursivamente a função SEGMENTAÇÃO com seus argumentos atualizados. Pode-se notar que as funções VERIFICAR_RESTO e VERIFICAR_LEXEMA devem manipular conhecimentos não incluídos no algoritmo.

Esses conhecimentos deverão estar então representados dentro de um formalismo qualquer que seja inteligível pela função em questão, sendo que o procedimento para manipulação de outros tipos de conhecimento é análogo. Assim, na Figura 5A, poder-se-ia representar cadeias, tais como, (começ)assem, onde ter-se-iam a VT (a), a DMT (sse) e a DNP (m), tirando-se daí então

a provável classe morfológica da lexia, bem como, suas categorias gramaticais.

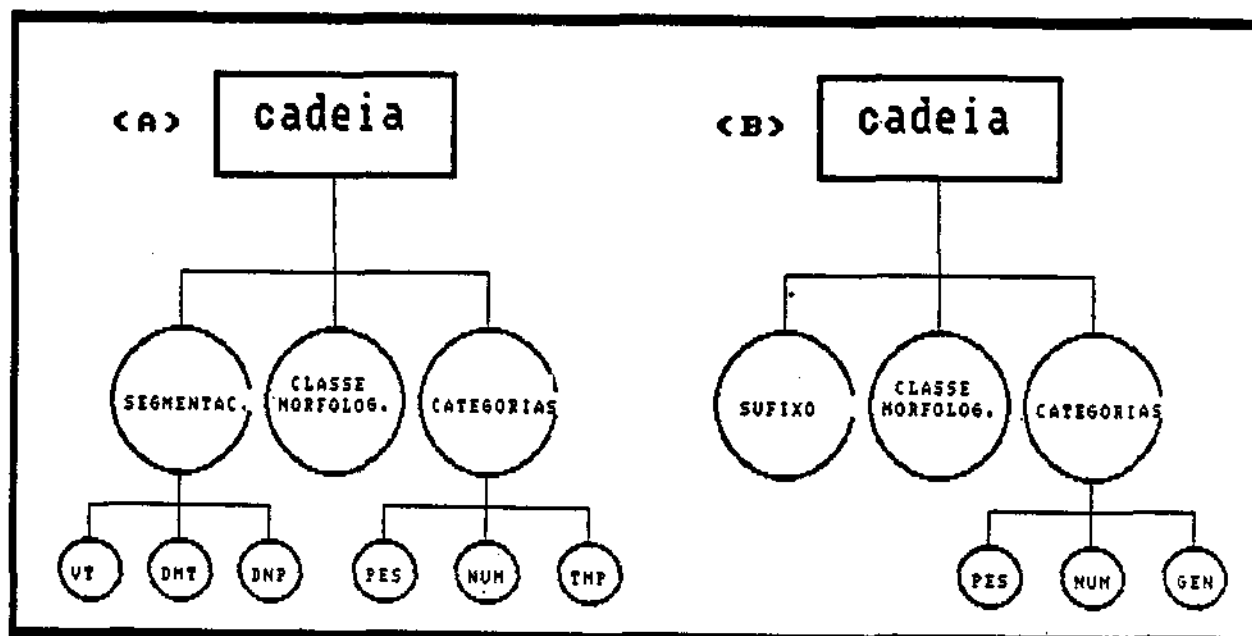


Fig. 5 - Sugestão de representação para conhecimento morfológico.

De modo análogo, para a cadeia (front)al, na Figura 5B, ter-se-ia a caracterização de um possível adjetivo. Deste modo, por este algoritmo, quase todas as lexias da língua portuguesa poderiam ser classificadas. Torna-se necessário somente que o conhecimento seja representado num formalismo qualquer. É evidente que, não somente para processos computacionais, tais formalismos são necessários, como, por exemplo, as árvores chomskyanas (Chomsky, 1965). Vários outros formalismos se prestam para representar conhecimento e, dentre eles, o mais comum e mais especial: o por frases em LN.

2.2 - PROPRIEDADES, TIPOS E ESQUEMAS DE REPRESENTAÇÃO

Para resolver problemas muito complexos (como os sobre LN) é necessário um volume grande de conhecimento e certos mecanismos para manipulá-los, visando criar soluções para novos problemas. Enfatize-se que duas entidades estão em jogo nesta discussão: as coisas que queremos representar (o conhecimento) e a representação desse conhecimento dentro de algum formalismo para que efetivamente possa ser manipulado. Para tanto, existem modelos de representação estruturada de conhecimento que visam a permitir o uso de mecanismos de inferência mais específicos e poderosos.

As propriedades que devem possuir um sistema de representação são, conforme Rich (1985, p. 201-202): a capacidade de representar todos os tipos de conhecimento que forem necessários naquele domínio; a capacidade de manipular as estruturas representacionais de modo a derivar novas estruturas correspondentes ao conhecimento novo inferido do antigo; a capacidade de incorporar, dentro da estrutura do conhecimento, informação adicional que possa ser utilizada para direcionar os mecanismos de inferência nas direções mais promissoras; a capacidade de adquirir novas informações com facilidade.

Dois tipos de representação de conhecimento se apresentam: a declarativa e a procedural. A vantagem da primeira diz respeito ao fato de que cada conhecimento só necessita ser representado uma vez e ao fato de que há facilidade de se acrescentar novos conhecimentos. A vantagem da segunda forma diz respeito à facilidade de representar conhecimento de como fazer, à facilidade de representar conhecimento que não se enquadra bem na forma declarativa simples (Rich, 1985, p.202).

Dentre as do tipo declarativo situam-se redes semânticas, dependência conceitual (Schank, 1975), "scripts" (Schank and Abelson, 1977) e "frames".

Projetadas primeiramente para representar o significado das palavras inglesas, as redes semânticas (Figura 6) são representadas por nós ligados um ao outro por arcos rotulados representativos de relações entre eles.

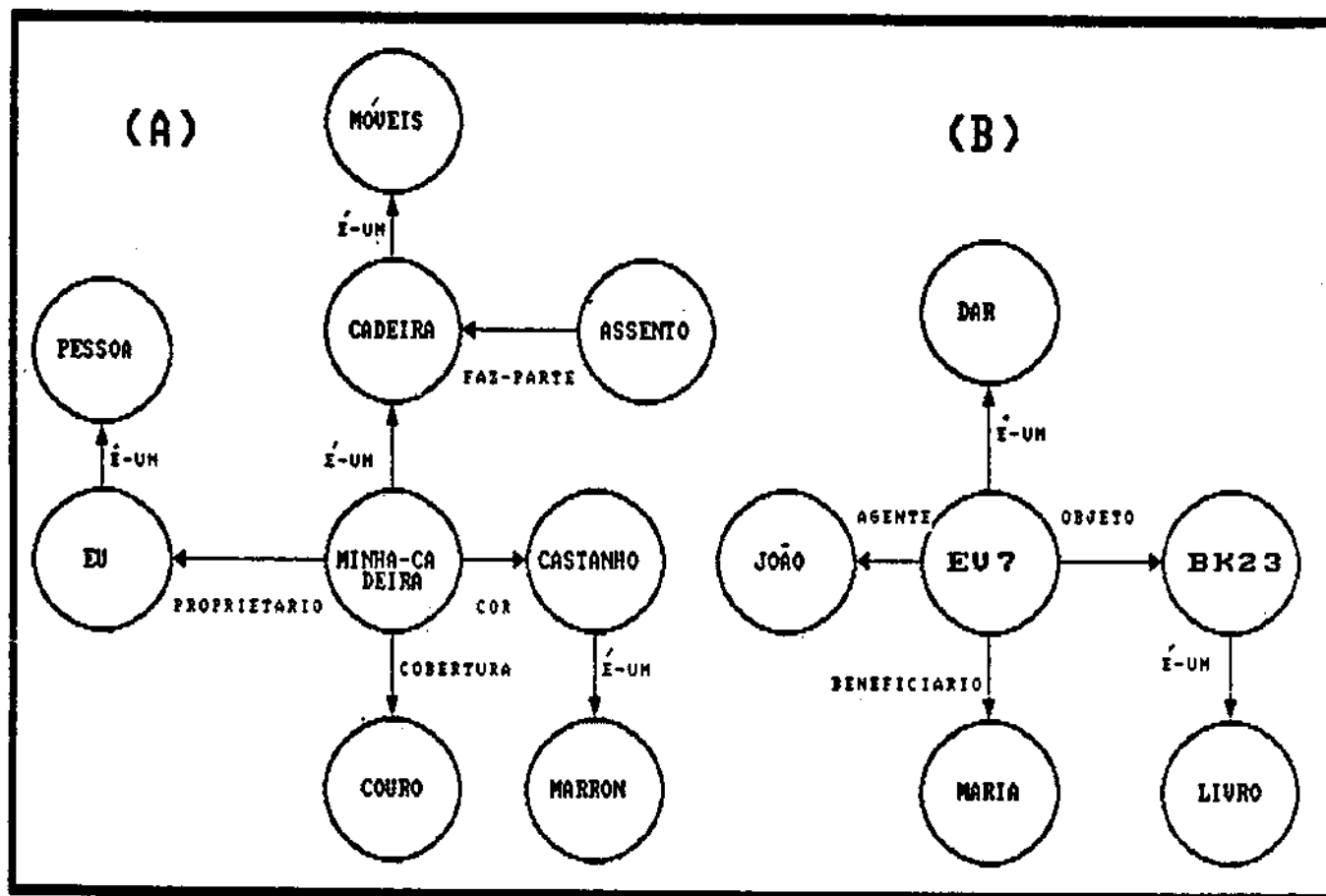


Fig 6 - Dois exemplos de rede semântica.

FONTE: Rich (1985), p. 253 e 255

A dependência conceitual (DC), como na Figura 7, é a representação de uma frase em primitivos conceituais que, combinados, podem formar significados de

palavras em qualquer língua em particular (conforme diz seu criador, Schank).

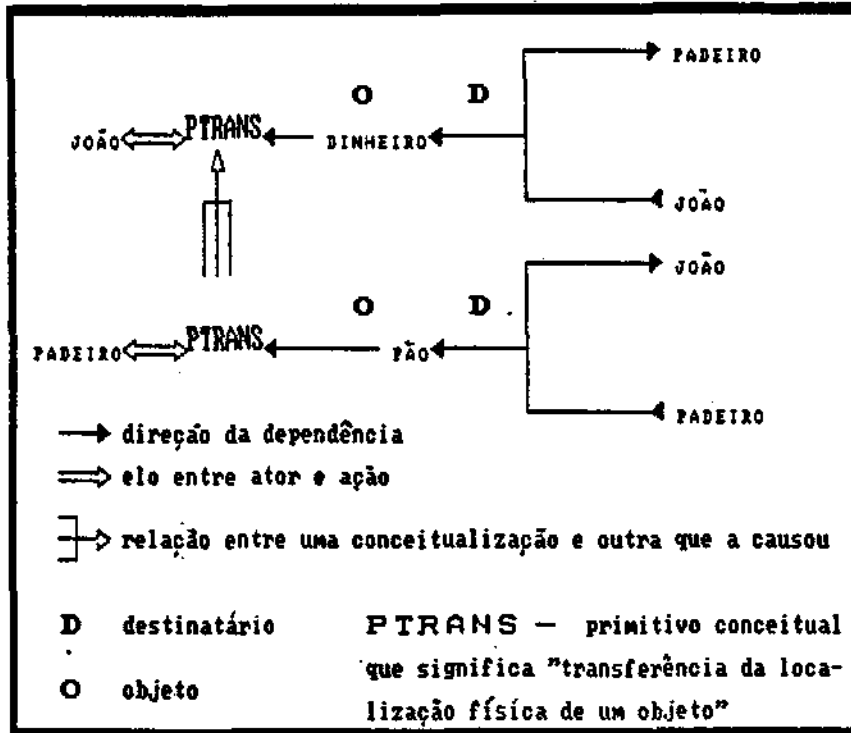


Fig. 7 - Uma representação DC para frase "João comprou pão (do padeiro)".

Notar na Figura 7 que, uma vez representada a frase, caracterizar-se-iam também relações do tipo vender-comprar/dar-receber pela própria força do formalismo em discussão.

O "script" é uma estrutura que descreve uma seqüência estereotipada de eventos num contexto particular, conforme Figura 8. Nesta, o primitivo MBUILD significa a construção de nova informação a partir da antiga (decidir, por exemplo).

Já o "frame", introduzido por Minsky para auxiliar a representação na área de visão (computacional), descreve uma classe de objetos. Compreende um rol de "slots" que descrevem aspectos de objetos, conforme Figura 9. "Frames" relacionados podem se agrupar num sistema.

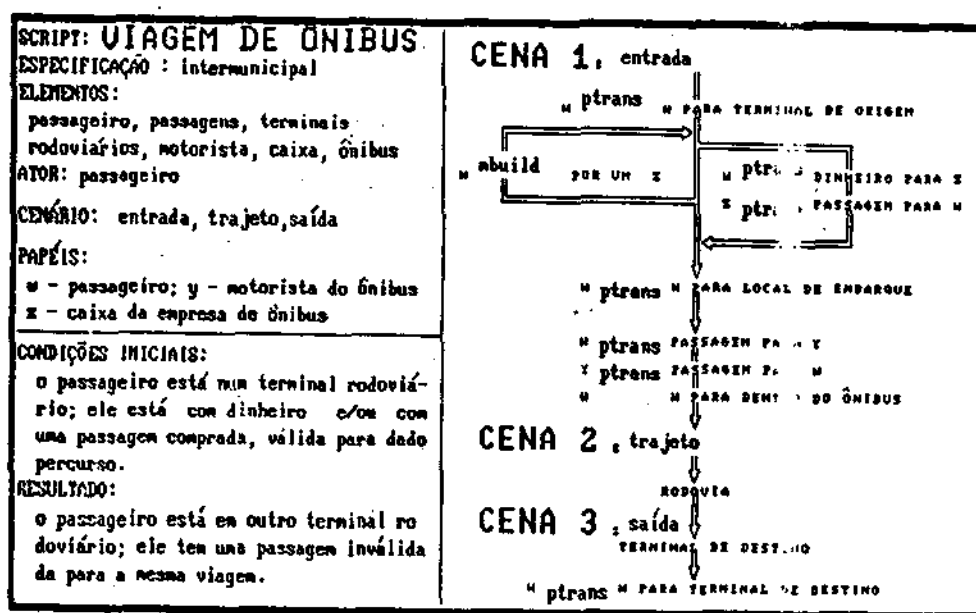


Fig. 8 - Um "script" elementar e simplificado para uma viagem de ônibus.

A representação do tipo procedural foi introduzida por Winograd (1976). Nesse esquema de representação, além de conhecimento do tipo declarativo, inserem-se informações sobre como proceder para atingir certas metas. Na Figura 9B, por exemplo, o "slot" IDADE contém o

"slot" PARA_ESTABELECEER que diz como obter a idade atual da pessoa descrita.

```

JOÃO      (É UM PESSOA)
           (NASCIMENTO      (ANO 1946)
                               (MÊS NOVEMBRO)
                               (DIA 13))
           (IDADE NIL))
           (IDADE (PARA_ESTABELECEER ( SUBTRAIR AGORA (ANO
                                                    DE NASCIMENTO))))

```

Fig. 9 - Um "frame" elementar para a descrição de uma pessoa.

FONTE: Rich (1985), p. 241

Dentre os esquemas [3] de representação de conhecimento sobressaem-se os "frames", os "scripts" e os modelos de regras. Como "frames" e "scripts" já foram apresentados, falar-se-á então dos modelos de regras: são utilizados para descrever recursos comuns partilhados entre um conjunto de regras de um sistema de produção (Davis, 1982). Uma regra de produção compreende um conjunto de premissas e um de conclusões. Dada premissa/conclusão é formalizada num tripla que contém os elementos: atributo ("attribute") e objeto do domínio e valor associado ao atributo. Assim, para a premissa "se a infecção for bacteremia primária" ter-se-ia, por exemplo, a tripla

3 Definido como em Bartlett (apud Rich, 1985, p. 203): "Schema refers to an active organisation of past reactions, or of past experiences, which must always be supposed to be operating in any well-adapted organic response".

atributo-objeto-valor, nesta ordem, (NOME, INFECÇÃO, BACTEREMIA PRIMÁRIA).

3. CONSIDERAÇÕES GERAIS SOBRE LINGÜÍSTICA TEXTUAL E PLN

A Lingüística Textual (LT) define-se como "uma lingüística dos sentidos e processos cognitivos e não da organização pura e simples dos constituintes da frase (Marcuschi, 1983, p. 14). Há então um ponto de convergência entre ambas as áreas de conhecimento (PLN e LT) que passará ser explorado neste trabalho. Enfatize-se, com relação a LT e a PLN, os modelos psico-cognitivos (fatores de coerência) que se realizam por esquemas de representação de conhecimento. Koch (1985, p. 154) cita algumas dessas representações ("frames", esquemas, planos, "scripts"), muito embora, tomados ali algo diferente do apresentado neste trabalho. No entanto, a questão é apenas de adequação terminológica: por exemplo, esquema em Koch é um sistema de "frames" em PLN.

Será segundo Marchuschi (1983) que alguns princípios da LT serão adotados neste trabalho. Assim, "Tomando a capacidade potencial de uma expressão de transmitir conhecimentos ou conteúdos, dizemos que o sentido é a realização atual desse conhecimento no texto. O texto é uma atualização seletiva de significações potenciais para possibilitar um só sentido.(...) O sentido deve manter uma continuidade, caso contrário o texto é incompreensível. Esta continuidade de sentido forma a coerência do texto e se expressa em conceitos e relações." (Marcuschi, 1983, p. 46).

Desde que a LT é uma lingüística de sentidos e processos cognitivos, adotá-la implicaria que: a) tomando-se uma frase em LN, poder-se-ia traduzir os conhecimentos nelas insertos para dada representação de

conhecimento e verificar, por inferências próprias ou através de interação com o usuário (produtor de tais frases), se tais conhecimentos constituem um "tecido" coerente naquele domínio; b) deverá existir um mecanismo que possibilite o caminho inverso quando da produção de LN.

Ressalvados os objetivos específicos de cada uma das áreas de estudo (LT e PLN), pode-se perceber o que há de comum entre elas. Como exemplo, notar os elementos de uma Gramática de Casos presentes nas representações de conhecimento visualizadas nas Figuras 6B e 7. Notar, ainda, a similaridade e/ou proximidade de algumas questões levantadas em ambas as áreas: que modelo(s) cognitivo(s) adotar e em que nível de análise? como representar o conhecimento para que possa ser manipulado convenientemente? como representar os fatores centrados no usuário (intencionalidade, por exemplo)? a análise deve ser feita em etapas estanques ou de um modo integrado?

Mas, afinal, em que a LT se conecta com um estudo sobre representação de conhecimento? Um pequeno e elementar exemplo pode ser apresentado no que concerne a um fator de coesão: a definitivação. Para as regras apresentadas em Fávero (1985, p. 149), optou-se por tentar representá-las simplificadamente por regras de produção, visto que estas podem conter fatores de crença que auxiliam no processo inferencial. Assim, pela Figura 10, o uso de representação por regras de produção parece ser pertinente. O é na realidade? E nos demais casos? Por isso, inúmeros estudos ainda vêm sendo feitos procurando-se achar "o modelo" de representação que mais se adeque a dado tipo de conhecimento, como se pode ver em Kaplan (1981), Rosenschein (1981) e Woods (1981), dentre outros. Em LT, não com o mesmo objetivo específico, mas sob a mesma ótica, cite-se o trabalho de Koch (1984).

(A)

Se houver uma seqüência de frases
 e se o elemento em análise for um referente
 e se esse referente for indefinido,
 então, com x % de certeza, deve-se aguardar um
 um referente definido .

(B)

Se o referente anterior for indefinido
 e se o referente em análise for também
 indefinido,
 então, com y % de certeza, indica-se o uso
 coletivo de uma entidade
 e, com z % de certeza, fala-se em tese
 sobre algo.

Fig. 10 - Definitivação: algumas regras de produção.

4. À GUISA DE CONCLUSÃO

A representação de conhecimento é um ponto de estrangulamento para a consecução dos objetivos tanto da LT como do PLN. Neste, mercê da distância astronômica que existe entre o grau de refinamento dos modelos e a rústica simplicidade das ferramentas computacionais, a emulação de um diálogo em LN está ainda circunscrita à elocubrações teóricas, sendo somente a simulação desse diálogo uma realidade. A LT, obviamente, não está sujeita ao cerceamento imposto ao PLN, mas, parece, carece ainda avançar nos formalismo de representação de conhecimento.

Visando, pois, um "casamento" entre as duas áreas, no que concerne a uma tentativa de solução parta

dirigidas pelo Domínio [4] (Coesão), como na Figura 10, e Regras dirigidas pelo Modelo Cognitivo (Coerência); c) daí para uma Representação Geral, válida para qualquer domínio, exigir-se-iam Regras de Tradução Inter-códigos e Esquemas de Representação de Conhecimento, estes definidos pelo(s) modelo(s) cognitivo(s) adotado(s). Para a produção (decodificação), o raciocínio seria da Representação Geral para a Intermediária e daí para o Texto.

Perceba-se que o Usuário foi propositalmente deixado de lado. Isto porque não se tem uma posição clara sobre onde e como seu conhecimento deve atuar: sobre o Domínio (Coesão) ? sobre o Modelo Cognitivo (Coerência) ? sobre ambos ? sobre as Regras de Tradução ? Ademais, por onde os usuários se compreenderiam: pelo Modelo Cognitivo subjacente ao Domínio ? pelo Domínio que determinou o Modelo ? por ambos ?

Outras questões poderiam ser levantadas aqui como, por exemplo, sobre o tipo de representação a ser usada em cada passo da Figura 11. Porém estas questões e outras propostas ficam para um próximo trabalho, enriquecidas, pressume-se, pela interação com os interessados neste assunto.

4 Este domínios poderiam ser, por exemplo, os tipos de texto.

REFERÊNCIAS BIBLIOGRÁFICAS

- CHOMSKY, N. Aspects of the theory of syntax. Cambridge, MIT Press, 1965
- COUTO, H. H. do. Uma introdução à semiótica. Rio de Janeiro, Presença, 1983
- DAVIS, R. Applications of meta-level knowledge to the construction, maintenance and use of large knowledge bases. In Knowledge-based systems in artificial intelligence. R. Davis & D.B. Lenat (eds.). McGraw-Hill, New York, 1982, p. 229-483
- DUBOIS, J. ; GIACOMO, M. ; GUESPIN, L. ; MARCELLESI, C. ; MARCELLESI, Jean-Baptiste; MEVEL, Jean-Pierre. Dicionário de Linguística. Tradução de Izidoro Blinkstein, Frederico Pessoa de Barros, Gesuina Domenica Ferreti, John Robert Schmitz, Leonor Scliar Cabral, Maria Elizabeth Leuba Salum e Valter Khedi. São Paulo, Cultrix, 1978
- FÁVERO, L. L. Mesa redonda: linguística textual. Grupo de Estudos Linguísticos, 10 (1): 146-152, 1985
- KAPLAN, S. J. Appropriate responses to inappropriate question. In: JOSHI, A. K. ; WEBBER, B. L.; SAG, I. A. (eds.). Elements of discourse understanding. New York, Cambridge University Press, 1981, p. 127-144
- KOCH, I. V. G. Argumentação e linguagem. S.Paulo, Cortez, 1984
- _____. Mesa redonda: linguística textual. Grupo de Estudos Linguísticos, 10 (1): 153-161, 1985

- MARCUSCHI, L. A. Linguística de Texto: o que é e como se faz. Dissertação de Mestrado em Linguística. Recife, Universidade Federal de Pernambuco, 1983 (Série Debates, 1)
- RICH, E. Artificial intelligence. 3rd. ed., London, McGraw-Hill, 1983, p. 135-246
- ROSENSCHEIN, S. J. Abstract theories of discourse and the formal specification of programs that converse. In: JOSHI, A. K. ; WEBBER, B. L.; SAG, I. A. (eds.). Elements of discourse understanding. New York, Cambridge University Press, 1981, p. 251-265
- SCHANK, R. C. Conceptual information processing. Amsterdam, Elsevier North-Holland, 1975
- SCHANK, R. C.; ABELSON, R. P. Scripts, plans, goals and understanding. Lawrence Erlbaum, 1977
- WINOGRAD, T. Understanding natural language. New York, Academic Press, 1976
- WOODS, A. W. Procedural semantics as a theory of meaning. In: JOSHI, A. K. ; WEBBER, B. L.; SAG, I. A. (eds.). Elements of discourse understanding. New York, Cambridge University Press, 1981, p. 300-334