

Utilização de um Ambiente Inteligente para Interfacear um Sistema de Processamento de Imagens.

JOSÉ DEMISIO SIMÕES DA SILVA¹
CARLOS ALBERTO DE OLIVEIRA¹
JOÃO MARQUES DE CARVALHO²

¹INPE - Instituto Nacional de Pesquisas Espaciais
Caixa Postal 515
12201-900 São José dos Campos, SP, Brasil

²UFPB - Universidade Federal da Paraíba
Av. Aprígio Veloso, 882
58100 Campina Grande, PB, Brasil

Abstract. This paper describes the use of an intelligent environment as an interface for an Image Processing System for educational purposes. The Image Processing System treated is of low complexity, and deals with basic tasks of Image Processing. The environment itself uses Artificial Intelligence techniques, such as frame-based knowledge representation within a blackboard approach and Natural Language Processing. The result is an environment adequate for interfacing general purposes systems, considering the knowledge involved in the tasks to be performed within each application domain. The user gets rid of computational details, such as parameters specifications and procedure sequence choices, while using a software package for Image Processing. A prototype of the system was developed for translating queries in the Portuguese language to an Image Processing language. The user specifies his queries using a free Natural Language within an application domain, then enhancing Man-Machine Interaction and the use of the interfaced system. The resulting knowledge of a session can be captured and stored in order to optimize future queries.

1 - Introdução

Os sistemas computacionais em geral objetivam atingir apenas as metas computacionais para as quais foram projetados. Assim, a interação homem-máquina (IHM) é dirigida para a realização das tarefas, e dificultada pelo fato dos sistemas não tentarem detectar a "intenção" do usuário na sua solicitação de ação. Com isso, sempre se executa ou não a tarefa desejada, e nenhuma outra chance de diálogo é dada ao usuário, caso este não domine a linguagem exigida para a

interação.

Este fato decorre da utilização de metodologias de aplicativos, utilitários e sistemas de "softwares" em geral, com abordagens direcionadas aos procedimentos ou aos dados. Portanto, a maioria dos sistemas são paramétricos exigindo um conhecimento detalhado dos parâmetros envolvidos nas várias ações possíveis.

Como consequência muitos usuários deixam de usar todo o potencial do "software" pela falta de conhecimento necessário, e pela necessidade de trabalho exaustivo para "dominar" o "software" e sua respectiva

Este trabalho faz parte do Projeto Temático de Equipe no. 91/3532-2

parametrização.

Uma maneira de resolver o problema é o projeto de interfaces que tornemos sistemas mais amigáveis, ou seja, que melhorem a interação homem/máquina. Muitos pesquisadores (Thompson, 1984; Schmitt, 1989; Oliveira, 1988; Ambler e Burnett, 1989) têm procurado meios de desenvolvimento de interfaces que facilitem a IHM. Para isto eles têm utilizado vários métodos que consideram a modelagem das características e as necessidades do usuário do sistema. Os sistemas objetivam permitir que o usuário possa executar suas tarefas com maior eficiência, uma vez que a parametrização e a seqüência de funções a ser usada, são fornecidas em alto nível de abstração.

Na maioria dos vezes os sistemas desenvolvidos incorporam interfaces que trazem vantagens ergonômicas, mas que não executam as tarefas automaticamente, do ponto de vista do usuário. Ou seja, durante toda a execução, o usuário é requisitado para interagir com a máquina, fornecendo parâmetros e escolhendo os procedimentos que devem ser adotados.

No geral os sistemas de Processamento de Imagens caracterizam-se como o tipo anteriormente mencionado, por exigirem conhecimento dos procedimentos e da sintaxe necessários na realização das tarefas. Isto torna os "softwares" difíceis de manipular pelos usuários de aplicação, por estes necessitarem de treinamentos exaustivos no sistema em uso, antes de operacionalizar seu uso na aplicação.

Como citado várias abordagens de desenvolvimento de interfaces que tornem o trabalho mais acessível, do ponto de vista da aplicação, têm sido propostos (Thompson, 1984; Schmitt, 1989; Oliveira, 1988; Ambler e Burnett, 1989).

Este trabalho apresenta o uso da

Linguagem Natural (LN) em uma abordagem cognitiva para conceber um ambiente inteligente para a interação homem-máquina. O ambiente traduz uma requisição do usuário em LN para a ação ou conjunto de ações mais adequado à realização da tarefa especificada. Neste processo a "intenção" do usuário é detectada através de interações com o mesmo por um sistema de menus, ícones ou perguntas, permitindo o mapeamento automático da solicitação do usuário no conjunto de ações que deve ser tomado. O ambiente utiliza-se do sistema lingüístico da Língua Portuguesa do Brasil (LPB), onde as solicitações do usuário são feitas em linguagem natural livre dentro de um domínio de aplicação restrito.

Da maneira como foi concebido, o ambiente pode mapear um conhecimento dentro de um domínio para uma linguagem de transformações como: Linguagem de Transformações Geométricas, Linguagem de Processamento Digital de Imagens como em (Batchelor, 1986), etc.

O mapeamento é a especificação dos procedimentos que devem ser adotados para a execução da tarefa desejada. Os parâmetros ainda são fornecidos pelo usuário, só que o ambiente, baseado no conhecimento armazenado do domínio, requisita de forma amigável, podendo inclusive explicitar a necessidade dos mesmos. Isso caracteriza um processo de treinamento do usuário, através do próprio ambiente, onde qualquer falha de especificação pode ser detectada antes da execução da tarefa.

O conhecimento resultante das interações com o(s) usuário(s) durante o processo de análise e interpretação, é "apreendido" pelo ambiente, estabelecendo um processo de aquisição do conhecimento (Oliveira, 1988). Esta facilidade permite que em novas sessões de uso do sistema, o usuário tenha seu trabalho racionalizado.

zado.

O uso da LN na concepção de ambientes inteligentes mantém a interação homem-máquina em um nível de abstração alto, permitindo que o usuário dedique-se às tarefas mais específicas, como a pesquisa de novos métodos de resolução de problemas.

Como exemplo, do uso da LN na interação homem-máquina, pode-se citar o trabalho de Schmitt (1989), que desenvolveu um sistema para gerar seqüências de transformações morfológicas encadeadas, no domínio da Morfologia Matemática, com interação em Língua Inglesa.

2 - Desenvolvimento

A experiência com sistemas de processamento de imagens mostra que a grande maioria dos pesquisadores restringem-se às tarefas básicas do processamento de imagens, envolvendo transformações simples. Este fato indica que alguns tipos de processamento são mais solicitados. Como exemplo tem-se a manipulação histográfica para realce de imagem, a filtragem passa-baixa para a eliminação de ruídos na imagem, classificação de padrões pelo método do paralelepípedo, o fatiamento de níveis de cinza para realce e/ou classificação de imagens, etc.

Alguns sistemas de processamento de imagens de pequeno e médio porte disponíveis no mercado, consideram estes fatos e generalizam alguns procedimentos para toda e qualquer imagem. Evidentemente que a generalização é suportada por resultados estatísticos resultantes de pesquisas, que traduzem a necessidade de certos tipos de procedimentos ou não. Ou seja, muitas tarefas de processamento de imagens são consideradas básicas em qualquer tratamento digital, resguardando o domínio de aplicação.

O desenvolvimento do ambiente apresentado neste trabalho, também considera este fato e assume que alguns processamentos são sempre requisitados no tratamento digital. Dentre estes tratamentos enquadram-se o realce por modificação histográfica, a eliminação de ruído por filtro passa-baixa, o realce de bordas por um filtro passa-alta e a classificação de padrões pelo método do paralelepípedo.

Uma outra consideração feita é que as configurações dos filtros utilizados não são complexas e os usuários utilizam apenas alguns tipos específicos, que estão disponíveis na literatura, e que a maioria dos sistemas comerciais trazem como padrões a serem utilizados. Isto permite que o usuário não entre no mérito de validar o uso ou não de determinado tipo de filtro.

O tipo de filtro mais comum é uma janela de 3 X 3, que convolui com janelas de dimensões iguais sobre toda a imagem. Como exemplo, considere os seguintes filtros passa-baixa e passa-alta utilizados:

passa-baixa	passa-alta
$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix}$	$\begin{vmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{vmatrix}$

A classificação de imagens é feita através do método do paralelepípedo. Não se procurou utilizar métodos de classificação mais complexos, que envolvem decisão estatística, por que o objetivo principal do trabalho, é contribuir na obtenção de uma interface homem / máquina melhorada. Assim, qualquer método pode ser escolhido. O método do paralelepípedo foi escolhido, por ser de baixa complexidade.

O sistema compõe-se de módulos que executam as tarefas de interpretação linguística e tradução no conjunto de ações a ser adotado.

O conhecimento do domínio de aplicação é representado utilizando "frames", que permite a representação de estereótipos, como por exemplo, o conceito de um objeto (como "cadeira") ou uma situação (como "entrar numa sala"). As várias fontes de conhecimento são interrelacionados dentro de uma abordagem de "Blackboard", que permite que várias fontes de conhecimento atuem no conhecimento armazenado para a resolução do problema, de forma cooperativa sob a coordenação de uma fonte específica.

Inicialmente desenvolveu-se um protótipo de um ambiente, para o domínio da Computação Gráfica, para realizar algumas tarefas sobre figuras geométricas regulares em duas dimensões (2D), como por exemplo translação, escalonamento, alteração de cor, etc.

No desenvolvimento do protótipo utilizou-se as linguagens C, na codificação das rotinas para apresentação de resultados gráficos, e PROLOG na construção de uma linguagem de manipulação de frames e na representação das regras de inferência necessárias à resolução dos problemas propostos.

Utiliza-se o conceito de primitivas de ação que compõem qualquer conjunto de ação que deve ser tomado para executar uma tarefa. Assim, as solicitações de ação escritas em sentenças da língua são traduzidas para um conjunto de primitivas de ação, segundo a representação por Dependência Conceitual (Schank, 1975).

Para o caso do domínio da Computação Gráfica, foram concebidas as seguintes primitivas:

IDENT - Identifica se o objeto explicitado na solicitação de ação pertence ao domínio em questão e se ele está presente no contexto gráfico corrente considerado.

LOC - Uma vez identificado o objeto algumas transformações podem requisitar seus atributos, o que é realizado por esta primitiva.

ALT - Realiza qualquer alteração em qualquer atributo do objeto.

Os objetos são representados utilizando o esquema de representação por frames, que permite estruturar hierarquicamente o conhecimento. A figura 1 abaixo exemplifica a representação por frames utilizada para o domínio da computação gráfica considerado no desenvolvimento do protótipo.

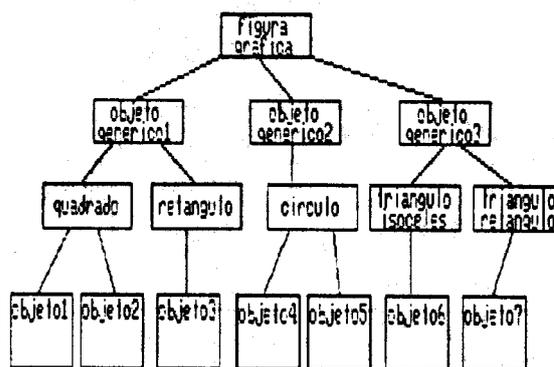


Figura 1 - Estrutura de frames para figuras geométricas.

Conforme se muda o domínio de aplicação, podem mudar também as regras de inferência, as primitivas e o conhecimento necessário do domínio. Assim, o usuário deve reespecificar os arquivos onde encontram-se suas descrições, em função do domínio de aplicação.

O protótipo desenvolvido foi adaptado para o domínio do Processamento de Imagens, considerando as questões sobre os "softwares" de processamento de imagens comentadas na introdução. O conhecimento ainda é representado através de frames. Na figura 2 tem-se um exemplo de representação para um catálogo de imagens. O conjunto de primitivas é composto por:

CLEAN - Primitiva para apagar a tela.

DISP - Exibe a imagem especificada na tela.

HFIL - Executa a filtragem passa-alta na imagem especificada.

LFIL - Executa a filtragem passa-baixa na imagem especificada.

CLASS - Executa a classificação da imagem especificada.

"DESLOCAR" - [IDENT, LOC, ALT [atributo : posição]].

A classificação anterior pode variar com o domínio de aplicação, entretanto um mesmo verbo pode ser utilizado para executar tarefas com semânticas diferentes em domínios distintos. Para o caso do protótipo desenvolvido, utilizou-se a capacidade gráfica do equipamento SITIM para demonstração do ambiente. Para isto classificou-se algumas ações básicas de carga e tratamento envolvendo filtragem de imagens. As seguintes ações foram classificadas como segue:

"CARREGAR" - [IDENT, DISP [atributo : nome_imagem]]

"SUAVIZAR" - [IDENT, LFIL [atributo : nome_imagem]]

"AGUDIZAR" - [IDENT, HFIL [atributo : nome_imagem]]

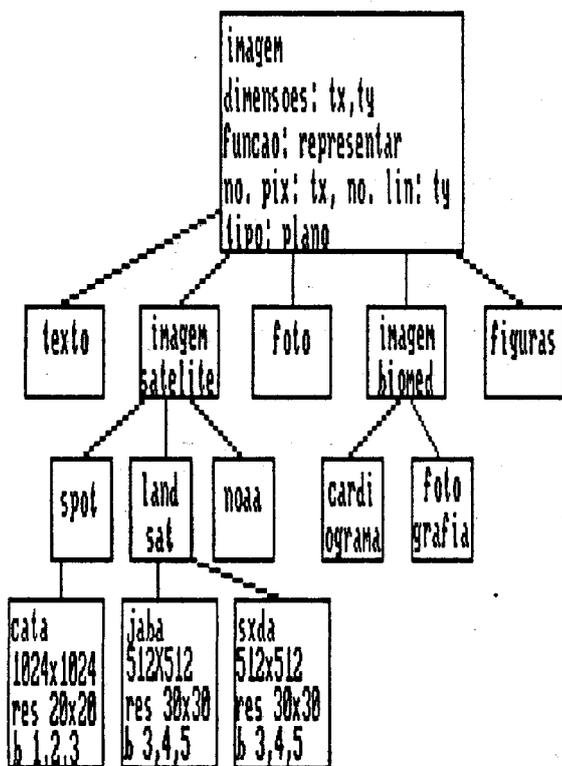


Figura 2 - Estrutura de frames para suportar as operações com imagens.

O mapeamento da solicitação de ação no conjunto de primitivas adequado, é feito a partir da interpretação da intencionalidade do usuário, resultante da análise lingüística. Este mapeamento é feito pela classificação das ações em função das primitivas. A seguir tem-se alguns exemplos de classificação para o domínio da Computação Gráfica.

"LOCALIZAR" - [IDENT, LOC]

O tratamento da Linguagem Natural, no domínio restrito, é executado pelo sistema IDEAL de (Oliveira, 1990). Este sistema constitui-se numa ferramenta para construção de interfaces em LN para sistemas especialistas. Assim, o sistema IDEAL, através das interações com o usuário, fornece o universo lingüístico suficiente para compreender, interpretar e traduzir frases em LN para uma representação de conhecimento. Em outras palavras, a aquisição do conhecimento léxico-semântico-pragmático das solicitações dos usuários é executada pelo sistema IDEAL, adaptado para os domínios em questão.

O ambiente como concebido pode ser usado para permitir interações com sistemas de Processamento Digital de Imagens, Processamento Digital de Sinais, Computação Gráfica, etc, onde as especificações de ação e o fornecimento de parâmetros são feitos

através de diálogos. A seguir tem-se exemplos de possíveis interações com o ambiente nos domínios do Processamento Digital de Imagens e da Computação Gráfica, respectivamente:

Exemplo 1:

```
ENTRADA > carregar a imagem
AMBIENTE > Especifique o nome da imagem.
ENTRADA > sao_jose
```

Mapeamento resultante:

```
carregar : [ IDENT, DISP ]
objeto : [ imagem ]
atributo : [nome: sao_jose]
```

```
ENTRADA > suavizar a imagem sao_jose
AMBIENTE > Especifique o canal.
ENTRADA > 0
```

Mapeamento resultante:

```
suavizar : [ IDENT, LFIL ]
objeto : [ imagem ]
atributo : [nome : sao_jose,
canal:0 ]
```

Exemplo 2:

```
ENTRADA > mudar o quadrado
AMBIENTE > Entre com a posição de origem.
ENTRADA > 10,20
AMBIENTE > Entre com a posição destino.
ENTRADA > 40,10
```

Mapeamento resultante:

```
mudar : [ IDENT, LOC, ALT ]
objeto : [ quadrado ]
atributos [ posicao_atual :
[10,20], posicao_destino :
[40,10]]
```

Exemplo 3:

```
ENTRADA > mostrar a imagem cata
AMBIENTE > "mostrar" denota ação sobre um objeto ?
ENTRADA > sim
```

AMBIENTE > A ação ainda não foi classificada. Você deseja classificá-la?

ENTRADA > sim

AMBIENTE > Você quer classificá-lo entre as ações primitivas ou derivadas?

ENTRADA > derivadas.

AMBIENTE > (exibe relação de ações classificadas como derivadas) Qual é então a ação sinônimo de "mostrar"?

ENTRADA > carregar.

No primeiro exemplo a medida que o usuário interage com o sistema, esse especifica o que deve ser fornecido, e pode inclusive esclarecer as dúvidas com relação à necessidade dos parâmetros. No segundo, o usuário é requisitado a fornecer os parâmetros necessários para a conclusão da tarefa.

No terceiro exemplo a requisição de ação do usuário não pode ser mapeada por que o sistema não possui a representação conceitual do verbo especificado. Assim, o sistema passa para uma fase pragmática requisitando a classificação da ação solicitada. Após o usuário fornecer o sinônimo da ação solicitada, o sistema classifica segundo a representação conceitual existente para a ação sinônimo. Assim, a representação da ação "mostrar" com a mesma semântica de "carregar" resulta no mapeamento do exemplo 1.

Os exemplos mostram alguns diálogos possíveis. Outras possibilidades necessitam o mapeamento de um maior número de ações do processamento de imagens, e a definição de um conjunto de primitivas mais completo que permita o mapeamento das ações.

O sistema de processamento de imagens idealizado para exemplificação do ambiente, utiliza a capacidade gráfica do equipamento SITIM, que permite a concepção de rotinas para realizar as tarefas básicas de processamento de imagens, comentadas na introdução. O "software" desenvolvido

resume-se a um conjunto de procedimentos que realizam tarefas simples como carregar imagens no visualizador, filtrar imagens, remover imagens da tela, apagar o canal de visualização, fatiar a imagem por níveis de cinza, etc.

O ambiente como um todo está dentro de uma estrutura de Blackboard, onde várias fontes de conhecimento atuam e são coordenadas por uma fonte, que ativa seqüencialmente os vários módulos de sistema do ambiente para a solução do problema.

3 - Conclusões

O uso da Linguagem Natural na concepção de sistemas inteligentes mantém a IHM em um nível de abstração alto, o que permite que os usuários finais dediquem-se às tarefas mais específicas, como por exemplo, a pesquisa por novos métodos de resolução de problemas, ao invés de preocupar-se com a especificação e a utilidade de parâmetros no uso de pacotes de "software" complexos. Portanto a IHM torna-se mais eficiente, uma vez que o usuário não precisa ter conhecimentos específicos do "software" que utiliza. Além disso, as tarefas são executadas quase que automaticamente do ponto de vista do usuário, permitindo que seu esforço na resolução de um problema seja racionalizado.

Muitos pesquisadores, cientistas, analistas, projetistas, etc, são céticos em relação ao uso da LN em sistemas computacionais. Isto ocorre porque o conhecimento envolvido no processamento de LN é grande e dinâmico, ou seja, as regras podem mudar constantemente de acordo com os costumes, cultura, etc, de uma certa região, o que implica na necessidade de se estabelecer novas regras e de se armazenar novos dados. Um outro motivo, está relacionado à representação de um mesmo sentimento que pode variar de um usuário para outro e de-

pender da aplicação. Isso implica que os ambientes precisam ser flexíveis no tocante à representação e dispor de capacidade suficiente para o armazenamento adequado. O problema de ambiguidade inerente à LN ainda não está resolvido, o que impede que ela seja usada de forma sistemática em qualquer aplicação. Porém, pode-se usá-la para domínios de aplicação restritos para tratar de situações típicas do domínio, o que implica em uma LN livre mas para domínio restrito, ou seja, não são todas as situações lingüísticas que podem ser tratadas.

Apesar das restrições de uso, ambientes em LN podem ser usados para permitir uma melhor IHM, em casos onde o sistema de "software" envolve a especificação de parâmetros e procedimentos, que muitas vezes o usuário especialista no domínio de aplicação, não está acostumado. Assim, pode-se pensar em um sistema de Processamento Digital de Imagens onde as especificações de ação e o fornecimento de parâmetros são feitos através de diálogos.

O uso de ambientes como o apresentado neste trabalho pode significar um uso mais eficiente de sistemas de processamento de imagens, e o desenvolvimento de novos "softwares" com capacidade e flexibilidade para suportar a programação pelo próprio usuário final. Utilizando-se técnicas de Inteligência Artificial, que permitam a execução das tarefas no nível da representação do conhecimento, possivelmente os sistemas de processamento tornem-se menos complexos.

Um outro aspecto positivo de "softwares" com ambientes, como o apresentado neste trabalho, é a possibilidade de treinamento de novos usuários em sistemas específicos, sem a necessidade de treinamento exaustivo no pacote propriamente dito.

Entretanto, existem muitos proble-

mas que ainda devem ser solucionados. O mais significativo é o problema da representação de imagens. A representação atual como matrizes de pontos e linhas não oferece grande expressividade na representação. Muitos pesquisadores trabalham em busca de formalismos de representação de imagens, que permitam que estas sejam tratadas em um alto nível de abstração.

O protótipo do ambiente desenvolvido realiza algumas funções consideradas básicas, como exposto na introdução. Para se tratar de um maior número de casos é necessário inserir novas primitivas de ação (regras) e representar um maior número de possíveis ocorrências lingüísticas.

Referências

- AMBLER, A. L. ; BURNETT, M. M..** Influence of Visual Technology on the Evolution of Language Environments. Computer, vol. 22, No. 10, October 1989, pp. 9 - 22.
- ARARIBOIA, G..** Inteligência Artificial: Um Curso Prático, Livros Técnicos e Científicos Editora Ltda, Rio de Janeiro 1988.
- BATCHELOR, B. G..** Merging the AUTOVIEW Image Processing Language with PROLOG. Image and Vision Computing, vol 4, No. 4, November 1986.
- BORBA, F. S. ; IGNÁCIO, S. E..** Critérios para Identificação dos Verbos de Ação e de Processo. X Anais de Seminários do Grupo de estudos Lingüísticos do Estado de São Paulo, Bauru 1985, pp. 01 - 03.
- BORBA, F. S. ; DEZOTTI, J. D..** Critérios para Identificação dos Verbos de Ação - Processo. X Anais de Seminários do Grupo de Estudos Lingüísticos do Estado de São Paulo, Bauru 1985, pp. 04 - 05.
- BORBA, F. S. ; NEVES, M. H. M. ; DALL'AGLIO, M. M..** Critérios para Identificação dos Verbos de Estado. X Anais de Seminários do Grupo de Estudos Lingüísticos do Estado de São Paulo, Bauru, 1985, pp. 06 - 10.
- FIKES, R.; KEHLER, T..** The Role of Frame-Based Representation in Reasoning, Communications of the ACM, September 1985, Vol 28, No. 9, pp. 904-920.
- HAYES-ROTH, B..** A Blackboard Architecture for Control, Artificial Intelligence, Vol. 26, No. 3, July 1985, pp. 260 - 261.
- NIE, G. L. ; MORIZET MAHOUEAUX, P. ; GAILLARD, P..** Application of the AI techniques to signal processing: A knowledge-based-system for digital filters synthesis. Signal processing, vol 23, No. 2, May 1991, pp. 121 - 136.
- NII, H. P..** Blackboard Systems: The Blackboard Model of Problem Solving and the Evolution of Blackboard Architectures, The AI Magazine, Summer 1986, Vol. 7, No. 2, pp 38 - 53.
- OLIVEIRA, C. A..** Sobre Interfaces em Linguagem Natural. INPE-4681 PRE/1376, Agosto 1988.
- OLIVEIRA, C. A..** IDEAL, uma Interface Dialógica em Linguagem Natural para Sistemas Especialistas. Tese de Doutorado em Computação Aplicada, INPE-5151-TDL/424, Outubro 1990.
- PASSOS, E. L..** Inteligência Artificial e Sistemas Especialistas ao Alcance de Todos, Livros Técnicos e Científicos Editora Ltda, Rio de Janeiro, 1989.
- SCHANK, R. C..** Conceptual Information Processing, North-Holland, Holland Publishing Company, Amsterdam 1975.

SCHANK, R. ; ABELSON, R.. Scripts Plans Goals and Understanding: An Inquiry into Human Knowledge Structures. Lawrence Erlbaum Associates, Publishers, Hillsdale, New Jersey 1977.

SCHMITT, M.. Mathematical Morphology and Artificial Intelligence: An Automatic Programming System. Signal Processing, vol. 16, No. 4, April 1989, pp. 389 -401.

SILVA, J. D. S. da. Ambiente Inteligente para Conversão de Aplicações em Linguagem Natural para uma Linguagem de Transformações. Dissertação de Mestrado, Universidade Federal da Paraíba - UFPB, Agosto de 1992.

THOMPSON, C. W.. Recognizing Values in queries or commands in a natural language interface to database. CAIA, vol. 1, Proceedings, 1984, pp. 25 - 30.

WEXELBLAT, R. L.. On Interface Requirements for Expert Systems. AI Magazine, Fall 1989, pp. 66 - 78.

WINSTON, P. H.. The Psychology of Computer Vision, McGraw-Hill 1975.

WINSTON, P. H.. Artificial Intelligence, Addison-Wesley 1979.