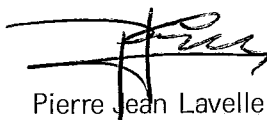


# LINGUAGENS DE CONSULTA A BASE DE DADOS

Celino Rojas

Tese submetida ao corpo docente da coordenação dos programas de pós-graduação de engenharia da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários para a obtenção do grau de mestre em ciências (M.S.C.)

Aprovada por



Pierre Jean Lavelle



Nelson Maculan Filho



José Lucas Mourão Rangel Netto



Daniel Alberto Menascé

Rio de Janeiro, RJ — Brasil  
JUNHO DE 1981

Rojas, Celino

Linguagens de Consulta a Bases de Dados (Rio de Janeiro) 1980

XII, 66 p. 29,7 cm (COPPE–UFRJ, M.Sc., Engenharia de Sistemas) 1980

Tese – Universidade Federal do Rio de Janeiro. Faculdade Engenharia

1. Linguagem de Consulta I. COPPE/UFRJ II. Título (série)

DEDICATÓRIA

*A minha esposa e filhas*

## RESUMO

A utilização de Sistemas Computerizados através de dispositivos interativos torna-se a cada dia um imperativo para a agilização dos processos decisórios e de controle. Sugere-se que na elaboração de Sistemas Computerizados interativos utilizem-se as técnicas das assim chamadas Linguagens de Consulta, como forma de facilitar a interação entre o usuário e o sistema.

Este trabalho apresenta as características necessárias para uma Linguagem de Consulta, implementável no todo ou em parte, em Sistemas Computerizados interativos.

**ABSTRACT**

The utilization of Computerized Systems by mean of interative devices has become more and more necessary to improve the decision and control process.

For implementation of interactive Computerized Systems, we suggest the aplication of the tecniques of the so called Query Languages to increase the user-system interaction.

This work presents the necessary features of the Query Languages to be implemented on such systems in partial or total form.

## **ÍNDICE**

### **PRÓLOGO**

#### **CAP. I INTRODUÇÃO ÀS LINGUAGENS DE CONSULTA**

- I.1 Níveis das Linguagens de Consulta a Bases de Dados
- I.2 Níveis de Utilização das Linguagens de Consulta
- I.3 Os Problemas do Diálogo
- I.4 Classificação dos Problemas do Diálogo
- I.5 Os Dispositivos de Execução do Diálogo
- I.6 Conclusão

#### **CAP. II CARACTERÍSTICAS COMUNS AOS VÁRIOS NÍVEIS DE DIÁLOGO**

- II.1 Início da Sessão
- II.2 Reconhecimento do Usuário
- II.3 Autenticação do Usuário e do Sistema
- II.4 Acesso a Base de Dados
- II.5 Informações Preliminares sobre o Sistema
- II.6 Escolha do Contexto Adequado ao Usuário
- II.7 Normas de Segurança e Recuperação
- II.8 Conclusão

#### **CAP. III DIÁLOGO DE ALTO NÍVEL**

- III.0 Introdução
- III.1 Preparação das Perguntas
- III.2 Interpretação das Perguntas
- III.3 Confirmação das Perguntas
- III.4 Recomposição das Perguntas
- III.5 Segmentação das Perguntas
- III.6 Informações Preliminares sobre as Perguntas
- III.7 Informações de Controle sobre o Processamento
- III.8 Preparação das Respostas e Entrega dos Resultados
- III.9 Treinamento do Usuário
- III.10 Conclusão

## **CAP. IV CONCLUSÕES E RECOMENDAÇÕES**

- IV.1 Implementações em Sistemas Atuais
- IV.2 Premissas sobre Implementações Futuras
- IV.3 Objetivos de uma Linguagem de Consulta
- IV.4 Conclusão

Bibliografia Seleccionada

## PRÓLOGO

A comunicação é a essência da natureza do homem. O surgimento da linguagem entre os humanos, foi o ponto de partida para sua diferenciação entre os demais animais. No aprimoramento contínuo desta linguagem diversas etapas foram vencidas, e sobre esta capacidade de vencer obstáculos mediante o uso da comunicação tem o homem depositado a sua confiança a tal ponto de não mais utilizar a linguagem apenas como instrumento de diálogo, de lazer, mas sim como forma definitiva de trabalho e ação.

A partir do momento em que a linguagem passa a ser considerada um instrumento de trabalho, surge a sua estilização e de certa forma uma fulga às veleidades da comunicação informal, e nas diversas categorias profissionais tem lugar a formação de verdadeiros dialetos técnicos aos quais apenas os iniciados conseguem interpretar plenamente, da mesma forma como no passado as tribos de cada região possuíam seu dialeto.

A comunidade de processamento de dados (PD) também possui a sua própria linguagem técnica, constituída dos inúmeros jargões da língua inglesa utilizados no ramo com significados exclusivos muitas vezes distanciados de seus significados originais. Entretanto, esta mesma comunidade de PD percebe, no momento, as barreiras que a especialização das classes e os dialetos técnicos próprios a cada uma delas vem criando à consecução plena dos objetivos sócio-econômicos das atividades de PD. Os usuários dos serviços de PD tem dificuldade em gerar as solicitações, pois a comunicação com os técnicos de PD é cada vez mais estreita devido ao desconhecimento de parte a parte dos termos empregados e dos conceitos que estes expressam. Esta barreira, fruto da estilização das linguagens vem a ser removida parcial ou totalmente com uma nova modalidade de oferecimento dos serviços de PD — a comunicação direta dos usuários com os sistemas sem interferência dos técnicos de PD.

A comunicação direta dos usuários com os sistemas exige no entanto que estes sistemas sejam capazes de “entender” as consultas dos usuários numa linguagem acessível a usuários de qualquer nível, pois se diferente fosse, teríamos novamente a necessidade de um técnico de PD especialista nesta linguagem para formular as consultas dos usuários.

Sugere-se pois um avanço no sentido das linguagens naturais, não no sentido completo que entendemos de difícil implementação, mas na direção de um modelo capaz de dotar o diálogo usuário-sistema dos contornos de um diálogo natural, eliminando-se ao menos parcialmente a rigidez que caracterizam os sistemas de aplicações atuais.

Acreditamos que num ambiente de comunicação mais ameno, obtido pela utilização das técnicas de linguagens de consulta, a produtividade dos usuários de sistemas tenda a crescer, bem como o interesse pela pesquisa seja renovado e ampliado.

No presente trabalho discutiremos a sistematização do “diálogo” com o objetivo de traçar as suas linhas mestras, detalhando as características mais importantes do processo e dando a conhecer exemplos de aplicação destas técnicas em sistemas de atualidade.



## CAP. I INTRODUÇÃO ÀS LINGUAGENS DE CONSULTA

### I.1 NÍVEIS DAS LINGUAGENS DE CONSULTA A BASES DE DADOS

O acesso às bases de dados é feito por linguagens que são processadas em vários níveis, segundo o distanciamento da base de dado real.

#### NÍVEL-3 Linguagens de Consulta

Consiste num interpretador de comandos de alto nível compostos de termos da linguagem natural.

#### NÍVEL-2 Analisador de Sintaxe

Consiste num mapeador da linguagem de consulta noutra linguagem de mais baixo nível.

#### NÍVEL-1 Analisador de Semântica

Consiste num analisador de relações e inferências da pergunta, cujo objetivo é evitar o acesso a base de dados (ABR74).

#### NÍVEL-0 Linguagem de Manipulação de Dados

Linguagem de acesso a base de dados que se utiliza das rotinas de acesso do sistema operacional.

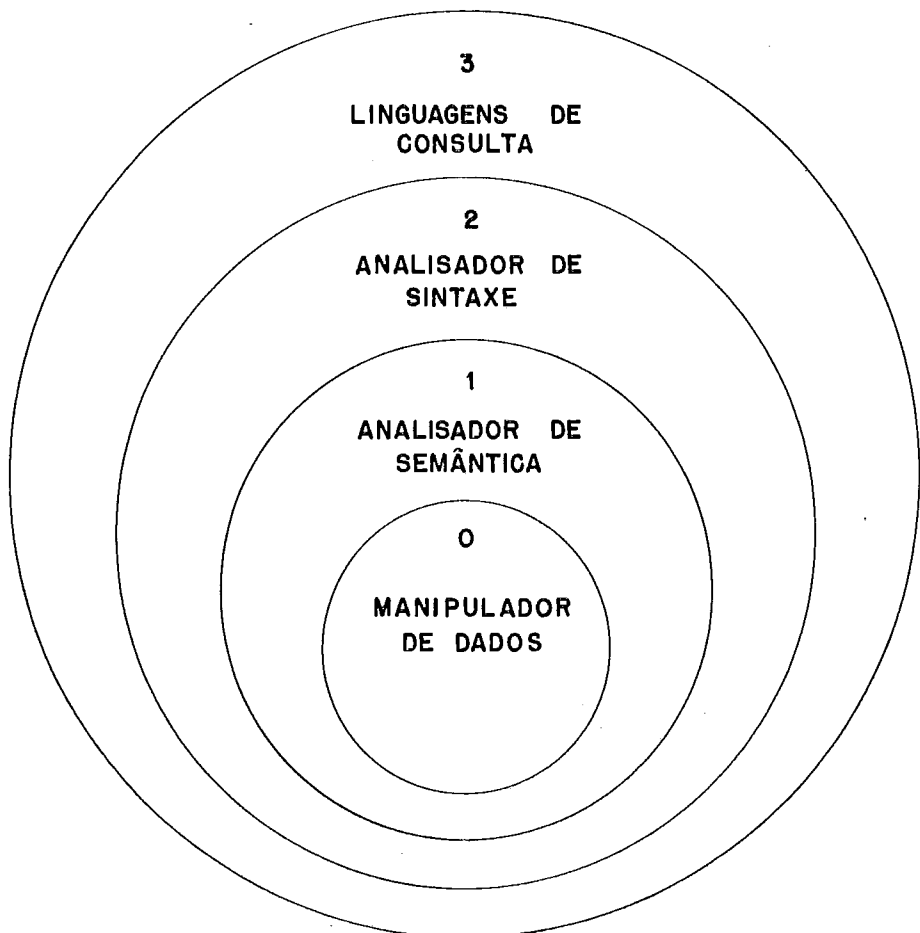


Fig. I.1

## I.2 NÍVEIS DE UTILIZAÇÃO DAS LINGUAGENS DE CONSULTA.

A existência de uma base de dados pressupõe a existência de uma organização que provenha os recursos necessários a sua realização. Além disso e pressuposto também a existência de um "software" capaz de prover ao usuário o acesso e gestão das informações contidas na base de dados. Neste contexto pode-se caracterizar o "espectro" do relacionamento dos usuários da organização (SPR77) com o sistema, para o acesso as informações da base de dados, pela figura abaixo.

# QUEM USA LINGUAGENS DE CONSULTA ?

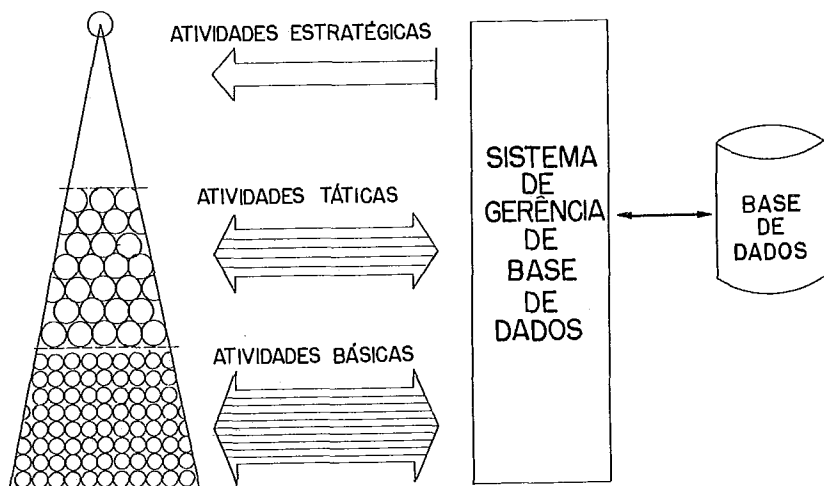


Fig. 1.2

Percebe-se assim que quanto maior o nível hierárquico do usuário na organização, maior será a necessidade que este possui de dispor de uma linguagem de consulta que lhe permita efetuar consultas com rapidez e precisão.

### **I.3 OS PROBLEMAS DO DIÁLOGO**

O diálogo se caracteriza por uma interação entre duas partes com interesse comuns, cujo processo é limitado por alguns problemas que podem ser classificados nos tipos abaixo:

#### **Entrada das Informações**

Ou o usuário não consegue fornecer as informações adequadamente, ou o sistema, por suas limitações, não está preparado para receber informações ao nível que o usuário necessita informar.

#### **Tratamento das Informações**

O sistema não é capaz de executar todos os processamentos solicitados pelo usuário, ou o faz em tempos que impedem sua utilização.

#### **Comunicação dos Resultados**

O sistema fornece os resultados obtidos sem a "apresentação" adequada ao fim desejado, ou sem a flexibilidade exigida pelo usuário.

### **I.4 CLASSIFICAÇÃO DOS PROBLEMAS DO DIÁLOGO**

Analisando-se os problemas do diálogo, percebe-se que eles são originários de 3 fontes básicas:

- Dispositivos de Interação
- Capacitação Técnica do Usuário
- Flexibilidade do Sistema

#### **I.4.1 PROBLEMAS ORIUNDOS DOS DISPOSITIVOS DE INTERAÇÃO**

Há uma evolução contínua na tecnologia dos terminais utilizados para o diálogo, entretanto a tendência é uma especialização para aplicações muito específicas, como por exemplo os terminais de entrada de dados (Data-Entry) e os terminais gráficos. Este fato dificulta a utilização dos terminais como instrumento de diálogo.

#### I.4.2 PROBLEMAS ORIUNDOS DA FALTA DE CAPACITAÇÃO TÉCNICA DO USUÁRIO

Em virtude da acirrada competição mercadológica as pessoas tendem a especializar-se em sua área técnica em detrimento da capacitação mínima requerida para utilização dos sistemas.

#### I.4.3 PROBLEMAS ORIUNDOS DA FALTA DE FLEXIBILIDADE DOS SISTEMAS

O sistemas de computação produzidos atualmente são herméticos e dirigidos a condições previamente estabelecidas, as quais pressupõe-se sejam do conhecimento do usuário, o que no entanto nem sempre é uma realidade permanente, dificultando assim sua utilização por usuários não treinados ou leigos.

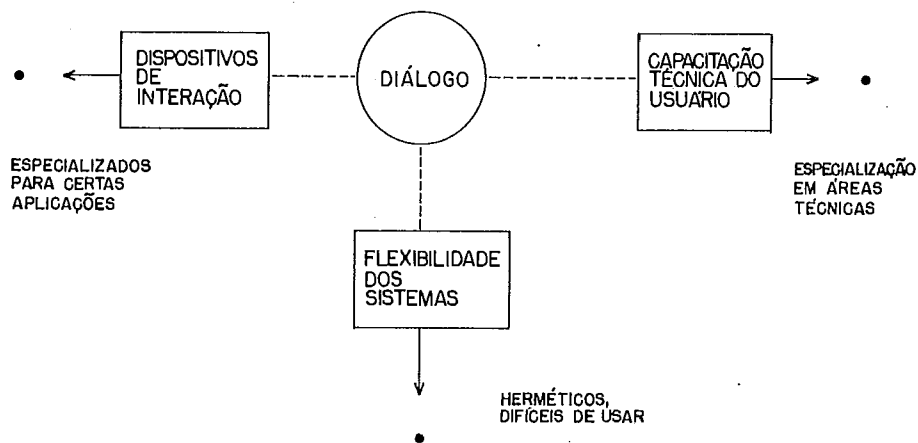


Fig. 1.4

## I.5 OS DISPOSITIVOS DE EXECUÇÃO DO DIÁLOGO

Os dispositivos mais comuns nos dias atuais são os terminais de vídeo, entretanto existe uma gama de equipamentos que são aplicados na utilização dos sistemas, os quais são mostrados na tabela abaixo (MAR73).

APLICAÇÃO	TIPOS DE TERMINAIS			
	TELEX	VÍDEO	GRÁFICO	ÁUDIO
ENTRADA DE DADOS		■	■	
RECUPERAÇÃO DE INFORMAÇÕES	■	■	■	■
ESTOQUE	■	■		
PASSAGENS	■	■		
FINANÇAS		■	■	
PROJETOS DE MÁQUINAS E EQUIPAMENTOS		■	■	
TREINAMENTO		■	■	
SEGURANÇA				■

Fig. I.5

## I.6 CONCLUSÃO

Pode-se ver, pela exposição dos fatos procedentes, que o diálogo está sendo prejudicado por diversos fatores. Entretanto, o fator que preponderantemente impede a difusão dos sistemas interativos dotados de linguagem de consulta atualmente é a má qualidade técnica dos sistemas. Desta forma, acreditamos que uma melhor avaliação das necessidades do usuário e a adoção das técnicas expostas nos capítulos que se seguem poderão contribuir para uma efetiva utilização das informações disponíveis, pelos usuários em nível de decisão.

A literatura técnica em PD não apresenta até a data uma discussão das técnicas aqui propostas de forma organizada e que efetivamente contribua para a adoção de técnicas de diálogo pelos Sistemas de Aplicação.

## CAP. II CARACTERÍSTICAS COMUNS AOS VÁRIOS NÍVEIS DO DIÁLOGO

### II.1 INÍCIO DA SESSÃO

O início da sessão é um procedimento que consiste na execução de alguns passos que independem do sistema de aplicação que será usado.

- Ligar o Terminal a Rede Elétrica
- Acionar o botão "Liga/Desliga" do Terminal
- Colocar-se como Usuário Ativo do Sistema Operacional
- Chamar o Sistema de Aplicação Desejado

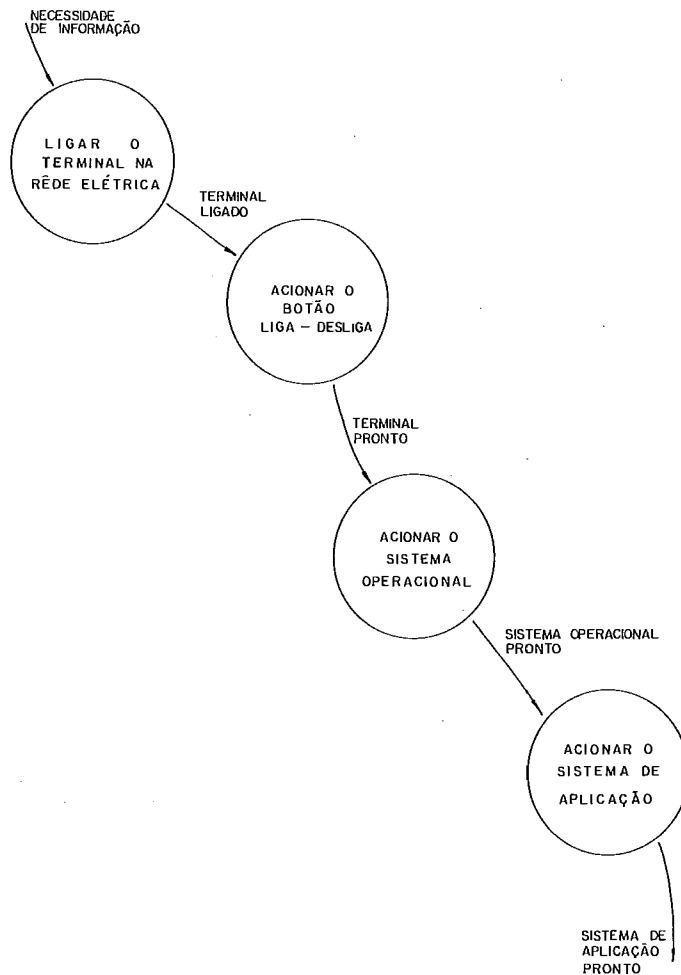


Fig. II.1

A partir do momento em que o sistema de aplicação está pronto, passa-se ao domínio do que consideramos o diálogo propriamente dito entre o usuário e o sistema.

## II.2 RECONHECIMENTO DO USUÁRIO

Dado que as informações armazenadas em uma base de dados possuem na maior parte das aplicações, um certo grau de sigilo, ou ainda porque o acesso a estas informações demandem um determinado processo de faturamento, é usual a identificação prévia de toda e qualquer pessoa ou programa (DAT77) que tente o acesso a base de dados.

A identificação se processa através do fornecimento pela pessoa ou programa, de senhas, cuja interpretação decidirá se o usuário pode ou não ter acesso a base de dados bem como o nível de utilização permitido.

A fim de proteger as senhas, é comum a utilização de métodos de ocultação das senhas no momento em que são fornecidas ao sistema, que variam de acordo com o equipamento utilizado. Quando se utilizam terminais de vídeo retira-se o "eco" dos caracteres da senha, ou seja as senhas não são impressas no vídeo quando digitadas pelo usuário, porém o sistema as recebe e executa as funções de validação. Quando são utilizados terminais tele-impressores, a região onde a senha será digitada é previamente preenchida pela sobreposição de caracteres tais como "X" e "W" de tal sorte que a senha uma vez impressa sobre esta "máscara" se torne ininteligível. A utilização por parte de outros programas se dá pelo uso de "Interfaces" que se incumbem da identificação segundo os mesmos mecanismos de validação das senhas.

Quando a senha fornecida não for localizada no arquivo de senhas, deve-se dar ao usuário a chance de digitar corretamente a senha por um número finito de vezes (3, 4 vezes) ao fim do qual o sistema deve deixar de atender a esse usuário.

Caso a senha fornecida tenha sido invalidada por decurso de prazo, deve-se informar claramente ao usuário e orientá-lo sobre os procedimentos para obtenção de nova senha. O uso de senhas com prazo de validade é útil nos ambientes de sistemas com fins comerciais, ou quando o nível de sigilo das informações o exigir.

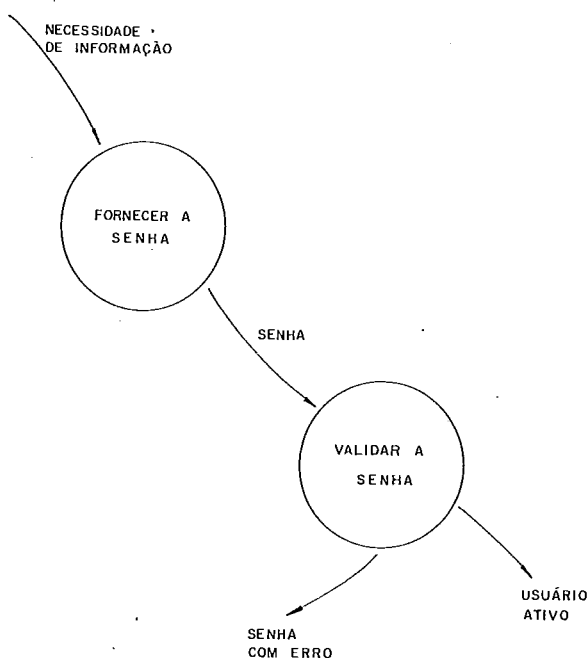


Fig. II.2

## II.3 AUTENTICAÇÃO DO USUÁRIO E DO SISTEMA

A fim de proteger a base de dados contra o acesso de pessoas desonestas e inescrupulosas que venham a ter conhecimento de senhas de outros usuários, é necessário que o sistema antes de permitir o acesso a base de dados proceda a uma autenticação do usuário (DAT77). A autenticação do sistema pelo usuário é necessária para assegurar a correta manipulação da base de dados selecionada, e a obtenção dos resultados esperados.

### II.3.1 AUTENTICAÇÃO DO USUÁRIO PELO SISTEMA

Consiste na solicitação ao usuário de determinadas informações, de composição variável, cujo conteúdo, seqüência ou forma de apresentação sejam conhecidas apenas pelo usuário verdadeiro. Algumas das formas de autenticação conhecidas são:

- **Cartão Magnético**  
Após o fornecimento da senha, um cartão magnético contendo um código de autenticação deve ser introduzido em um dos orifícios existentes para esse fim, cuja posição depende da senha fornecida.
- **Questionário Pessoal**  
Algumas perguntas de cunho pessoal devem ser respondidas pelo usuário logo após o fornecimento da senha, sem que no entanto estas respostas reflitam a situação real do usuário, e sim uma situação fictícia estabelecida previamente que pode ser alterada a cada fim de sessão.
- **Relatório da Última Utilização**  
O sistema solicita do usuário alguns detalhes da última utilização, e ao mesmo tempo fornece um relatório completo sobre a última sessão.
- **Algoritmos Especiais**  
O sistema fornece ao usuário um número aleatório e este tem que fornecer a resposta correta de acordo com um algoritmo simples que o usuário tenha anteriormente memorizado.

Exemplo: O Sistema Fornece — X, Y, Z  
O Usuário Fornece — Os dois algarismos após a vírgula do resultado da expressão abaixo:

$$S = \sqrt{(X + Y) * Z}$$



O algoritmo pode incluir a data, o número do terminal, a última resposta dada, ou seja, alguma informação gerada aleatoriamente.

- **Códigos de Senha Variáveis**

O sistema altera a senha do usuário a cada sessão, pela variação de algumas letras de sua senha básica, segundo uma lei de formação simples que o usuário deve conhecer, de forma que a cada sessão uma senha diferente deve ser fornecida para a autenticação do usuário.

Exemplo: 1) SEGREDO ... 2) SFGRFDP ... 3) SGGRGDQ ...

Neste exemplo as letras nas posições 2, 5 e 7 são trocadas pelas suas sucessoras na ordem alfabética a cada sessão.

- **Lista de Senhas Usadas Ciclicamente**

O usuário não dispõe de apenas uma senha, mas sim de uma série delas que devem ser usadas numa determinada ordem.

Exemplo: JACXYZ, JACZXY, ABD347, ACY324

Neste exemplo, após uma sessão em que tenha sido usada a senha JACXYZ — deverá ser utilizada a senha — JACZXY.

### II.3.2 AUTENTICAÇÃO DO SISTEMA PELO USUÁRIO

Consiste na verificação pelo usuário das características do sistema através das suas informações preliminares (veja item II.5) ou de testes específicos.

- **Relatório de Última Utilização**

O usuário solicita detalhes da última utilização, tais como a hora da sessão, sua duração, operações executadas, etc.

- **Algoritmos Especiais**

O usuário seleciona um algoritmo (ou função) de teste e fornece um valor inicial para que o sistema forneça a resposta. Este algoritmo poderá ser formulado pelo próprio usuário durante a implantação.

## II.4 ACESSO A BASE DE DADOS

O acesso a base de dados, disciplinado pelas senhas, se processa condicionado ao tipo de usuário e aos tipos de ações que lhes são permitidas.

### II.4.1 TIPOS DE USUÁRIOS DE BASES DE DADOS

Os usuários de uma base de dados podem ser classificados em função do nível de autorização ou restrição de acesso que lhes foi atribuída.

- **Usuário Principal**  
Normalmente conhecido por “Administrador da Base de Dados”, tem acesso irrestrito a Base de Dados podendo executar quaisquer ações (Veja item II.4.2)
- **Usuário Privilegiado**  
Tem acesso irrestrito a Base de Dados, porém tem limitada autorização para execução de determinadas ações de controle tais como Criação/Destruição, Geração de Cópias, etc.
- **Usuário Final**  
Tem acesso restrito a Base de Dados e não tem permissão para execução de ações de controle operacional sobre a Base de Dados.

### II.4.2 AÇÕES SOBRE A BASE DE DADOS

As ações constituem o atendimento a uma solicitação que o usuário expressa através da linguagem de consulta. Existe uma gama de ações que podem ser solicitadas pelo usuário, para serem executadas sobre uma base de dados.

- **Criação/Destruição**
  - Criação — Ação de definir a estrutura da Base de Dados e proceder a carga inicial de dados
  - Destruição — Ação de remover a Base de Dados do meio de gravação acessível a todos os usuários
- **Geração de Cópias**  
Ação de reproduzir a Base de Dados de um meio de gravação para outro sem perda ou alteração das informações.

- **Reorganização**  
Ação de reorganização dos diretórios de uma Base de Dados visando melhorar a sua performance de respostas.
- **Manutenção de Senhas**  
Ação de inclusão de novos usuários no sistema ou a manutenção das senhas por ações de alterações, exclusões, mudança da classificação do usuário, listagem das senhas, etc.
- **Relatórios de Controle**  
Funções do sistema que fornecem relatórios de análise ou de avaliação de performances ou ainda de estatísticas de utilização e faturamento.
- **Leitura de Campos**  
Ações de consulta aos campos de Base de Dados através dos comandos da linguagem de consulta.
- **Alterações de Campos**  
Ação de alteração do conteúdo dos campos da Base de Dados através dos comandos da linguagem de consulta.
- **Exclusões de Campos**  
Ação de remover campos de informação de uma determinada entidade que existe na Base de Dados, através de comandos da linguagem de consulta.

### II.4.3 NÍVEIS DE RESTRIÇÃO DE ACESSO

As rotinas de autorização de acesso são acionados para identificar o usuário e determinar a sua classificação no tocante a utilização das Bases de Dados, fixando para toda a sessão da consulta, a faixa de ação permitida, ou seja o nível de restrição (ou de autorização) de acesso.

- **Restrição de Acesso a Base de Dados**  
Este é o maior nível de restrição e se verifica pelo fornecimento de uma senha válida.
- **Restrição de Uso de Funções de Manipulação**  
Este é o nível em que se delimitam as funções do administrador de Bases de Dados e os outros usuários.

- Restrição de Acesso a Campos de Informação da Base de Dados

Este é o nível de separação entre usuários privilegiados e não privilegiados. Os primeiros podem ter acesso a todos os campos enquanto que os últimos não têm essa autorização.

- Restrição de Ação Sobre Campos de Informação

Este é o nível de maior controle das restrições, e se destina a permitir a manipulação controlada das Bases de Dados por pessoas pouco especializadas ou com funções bem definidas na organização. Podem ser sumarizadas em 3 tipos:

- Leitura de determinados campos

Alguns campos de informação podem ter o acesso protegido, tais como:

Salários de funcionários

Preços de novos produtos

Códigos militares

Restringe-se pois a leitura destes campos da parte das senhas não autorizadas.

A restrição de leitura implica automaticamente na restrição de alterações e exclusões, não sendo no entanto a recíproca necessariamente verdadeira.

- Alterações de Determinados Campos

A ação de alteração pode ser restringida em todos os campos de informação, ou apenas naqueles de maior importância.

Como em algumas aplicações a ação de alteração é freqüente e envolve o manuseio de grandes volumes de dados, é comum atribuir esta tarefa a técnicos especializados, com autorização exclusiva de executar alterações, sem permitir-lhes a leitura dos campos da Base de Dados.

A restrição de alteração pode ainda ser dirigida, permitindo-se ao usuário que altere os valores dos campos somente se os novos valores preencherem em algumas condições, relacionadas com os antigos valores, geralmente da forma:

$$VN > VA * P$$

$$VN < VA * P$$

$$VN = F (VA)$$

onde

VN – Valor Novo

VA – Valor Antigo

P – Percentual

F (VA) – Função Algébrica do Valor Antigo

- Exclusões de Campos

Esta é a ação mais perigosa que o usuário exerce sobre as Bases de Dados, sendo por isso mesmo necessário que se restrinja ao mínimo necessário, o número de pessoas autorizadas a

executá-la. A restrição pode ser parcial, quando aplicada apenas sobre alguns campos, ou total quando aplicada sobre todos os campos. A exclusão pode ainda ter as seguintes interpretações.

– Exclusão de várias ocorrências do campo

É a exclusão qualificada das ocorrências que atendam à condição fornecida.

– Exclusão de todas as ocorrências do campo

É a exclusão completa que altera profundamente a estrutura da Base de Dados, reduzindo o seu tamanho físico.

## II.5 INFORMAÇÕES PRELIMINARES SOBRE O SISTEMA

Existem algumas informações sobre o sistema que podem ser fornecidas ao início da sessão.

- Identificação da Instalação

Denominação da instalação da Base de Dados, bem como a data e hora do processamento que se inicia.

- Identificação do Sistema

Versão do sistema, data de sua instalação e anúncio de novas facilidades.

- Serviço de Comunicação entre Usuários

O sistema informa ao usuário quantos usuários estão utilizando o sistema neste momento e permite que sejam transmitidas mensagens entre eles sem prejuízo das ações normais sobre as Bases de Dados. Quando algum usuário envia mensagem para outro usuário que não esteja operando o sistema no momento, esta mensagem é armazenada até que o usuário destinatário entre em contato com o sistema, quando então todas as mensagens que lhe foram enviadas são apresentadas diretamente no terminal ou em impressora.

- Característica das Bases de Dados

Conjunto de informações sobre a Base de Dados selecionada, que permita ao usuário uma visão global de Base e um eficiente uso de seu potencial.

- Tamanho da Base de Dados em Caracteres

- Data da Última Recuperação

- Período Coberto pelas Informações Armazenadas

- Conteúdo da Base de Dados (campo, itens em cada registro)

- Campos Pesquisáveis (se não forem todos)

- Estratégias para uso ou para fugir de armadilhas próprias da Base de Dados

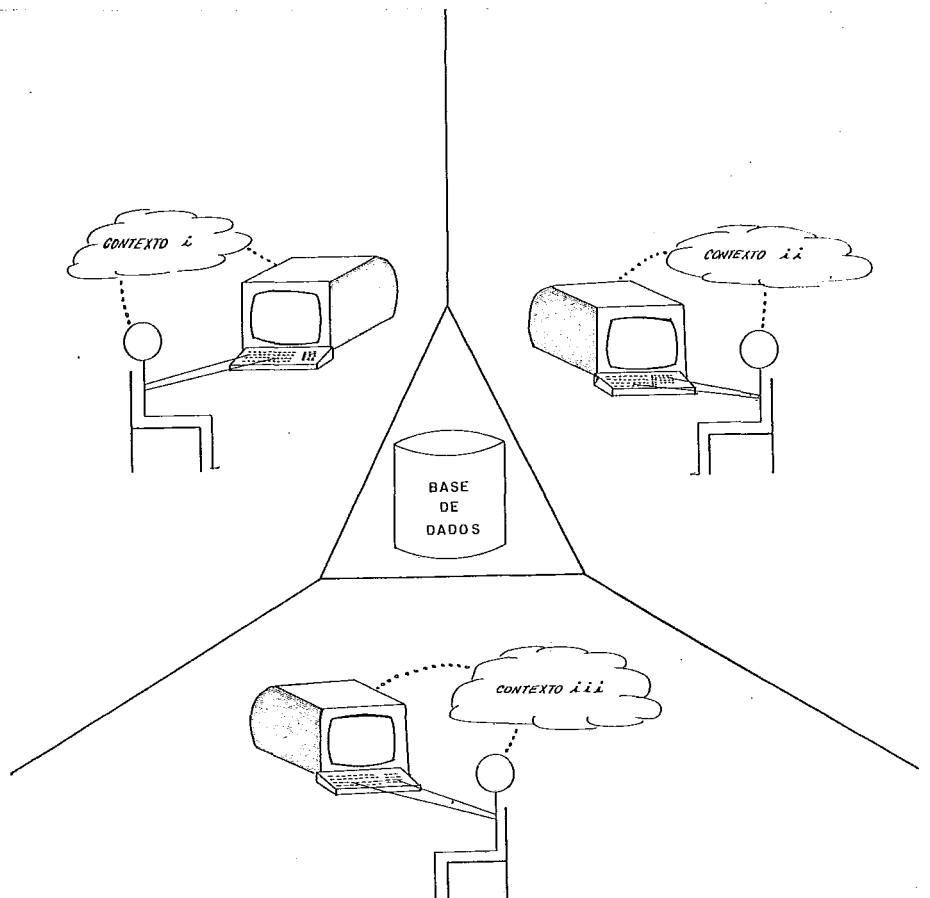
Estas informações poderão ser fornecidas automaticamente, ou mediante solicitação do usuário. Podem ainda ser divididas em dois grupos, segundo sua importância e extensão de apresentação, permitindo-se sempre que o usuário comande a omissão de sua apresentação.

## II.6 ESCOLHA DO CONTEXTO ADEQUADO AO USUÁRIO

O diálogo entre o sistema e o usuário, pressupõe a existência de um contexto previamente estabelecido, ou seja, de acordo com a classificação do usuário define-se o grau de "entendimento" que deverá nortear a interpretação das perguntas feitas.

O estabelecimento deste contexto de diálogo será necessário na medida em que o sistema possua rotinas de interpretação de perguntas capazes de "deduzir" significados de termos parcialmente fornecidos (WAL78) ou ainda se são permitidas segmentações de perguntas, em que o argumento da pergunta anterior vale para a pergunta seguinte.

O contexto pode ainda ser estendido para a interpretação de termos técnicos especiais, comuns dentro de determinada área profissional. Isto facilita a interpretação de perguntas, pois diminui o número de termos a pesquisar quando do exame das perguntas.



### COMPOSIÇÃO DO CONTEXTO

- Estruturas de Informações suportadas
- Terminologia técnica adequada
- Informações adicionais desejáveis
- Forma de entrega dos resultados

Fig. II.6

## II.7 NORMAS DE SEGURANÇA E RECUPERAÇÃO

A segurança de sistemas se faz necessária para a prevenção de situações que coloquem em risco as informações contidas na base de dados. Os riscos a que estão expostas as informações de uma base de dados, são:

- Danificação por manipulação indevida
- Danificação por falhas de equipamentos
- Sinistros
- Furto ou destruição intencional
- Uso das informações por pessoas não autorizadas

### II.7.1 PREVENÇÃO CONTRA DANIFICAÇÃO OU DESTRUIÇÃO DA BASE DE DADOS

As precauções para evitar a danificação ou destruição da Base de Dados dependem do agente danificador.

- O Sistema que Manipula a Base de Dados  
O sistema deve ser desenvolvido segundo uma metodologia que permita analisar todas as situações anormais a que o sistema será submetido ao longo de sua vida útil. Além disso, antes de sua efetiva implantação o sistema deverá ser submetido a testes simulados que demonstrem sua resistência nas situações de exceção. Nos casos de utilização de sistemas prontos, é necessário que ao menos os testes sejam acompanhados por técnicos de PD que auferirão a capacidade do sistema de sobreviver às situações anormais.
- Os Equipamentos que são Utilizados para Operação do Sistema  
A instalação utilizada deve atender as condições de segurança típicas tais como — alarme contra inundação e incêndio, dispositivos de controle de incêndio, equipamentos e prevenção contra sobrecarga de tensão, cofres blindados para segurança de arquivos. Além disso um esquema de manutenção preventiva deverá ser adotado para prevenir quaisquer falhas nos equipamentos que possam afetar a Base de Dados.
- O Usuário do Sistema  
Todos os usuários deverão ser treinados no uso do sistema, dentro de sua faixa de utilização. Poderão ser aceitos usuários não treinados, desde que a linguagem de consulta seja auto-explicativa e de fácil entendimento para o usuário leigo em PD.

Recomenda-se ainda que as ações de alteração da Base de Dados sejam restringidas a um número mínimo de usuários dentro da organização. A restrição pode ainda ser orientada para os campos que o usuário poderá consultar (veja item II.4.3).

O usuário mal intencionado, que venha a ter conhecimento das senhas de outros usuários, pode ser impedido de acessar a Base de Dados através da alteração periódica das senhas (veja item II.3.1).

## **II.7.2 PREVENÇÃO CONTRA SINISTROS, FURTOS OU DESTRUIÇÃO INTENCIONAL**

Para prevenir contra estes danos é necessário a manutenção de cópias da Base de Dados mantidas em lugar seguro, diverso daquele em que o sistema é operado normalmente. A utilização de cofres blindados a prova de fogo e de inundação, na própria instalação ou nos locais alternativos assegura um melhor nível de proteção a Base de Dados.

## **II.7.3 USO DAS INFORMAÇÕES POR PESSOAS NÃO AUTORIZADAS**

Um dos riscos de mais difícil detecção é o uso das informações por pessoas inescrupulosas em proveito próprio ou de grupos externos ou mesmo da própria organização. Nestes casos aplicam-se as proposições feitas nos itens II.2, II.3 e II.4, existindo ainda algumas técnicas de codificação das informações, aplicáveis principalmente nos casos de transferência de dados de um equipamento para outro, ou de uma instalação para outra através de redes de transmissão de dados (HOF77).

## **II.8 CONCLUSÃO**

A segurança das informações é uma preocupação importante quando se trata de sistemas interativos. Entretanto ela não deve ser implementada de forma que o usuário seja inibido de utilizar o sistema pois isto vai de encontro com o principal objetivo das próprias linguagens de consulta, que é de facilitar o acesso e estimular o uso das informações como base dos processos decisórios. Sugere-se pois que se adotem esquemas de proteção por níveis de importância, considerando-se a própria Base de Dados e principalmente que esses níveis de segurança sejam adotados de comum acordo com o usuário principal do sistema.



## CAP. III DIÁLOGO DE ALTO NÍVEL

### III.0 INTRODUÇÃO

Os diálogos de alto nível podem ser descritos como processos de obtenção de informações a partir da interação homem-máquina.

As etapas que compõe este processo serão apresentadas nos itens seguintes, onde procuramos detalhar a lógica funcional de cada etapa, quais sejam:

- Preparação da Pergunta
- Interpretação da Pergunta
- Validação da Pergunta
- Preparação da Resposta
- Entrega de Resultados

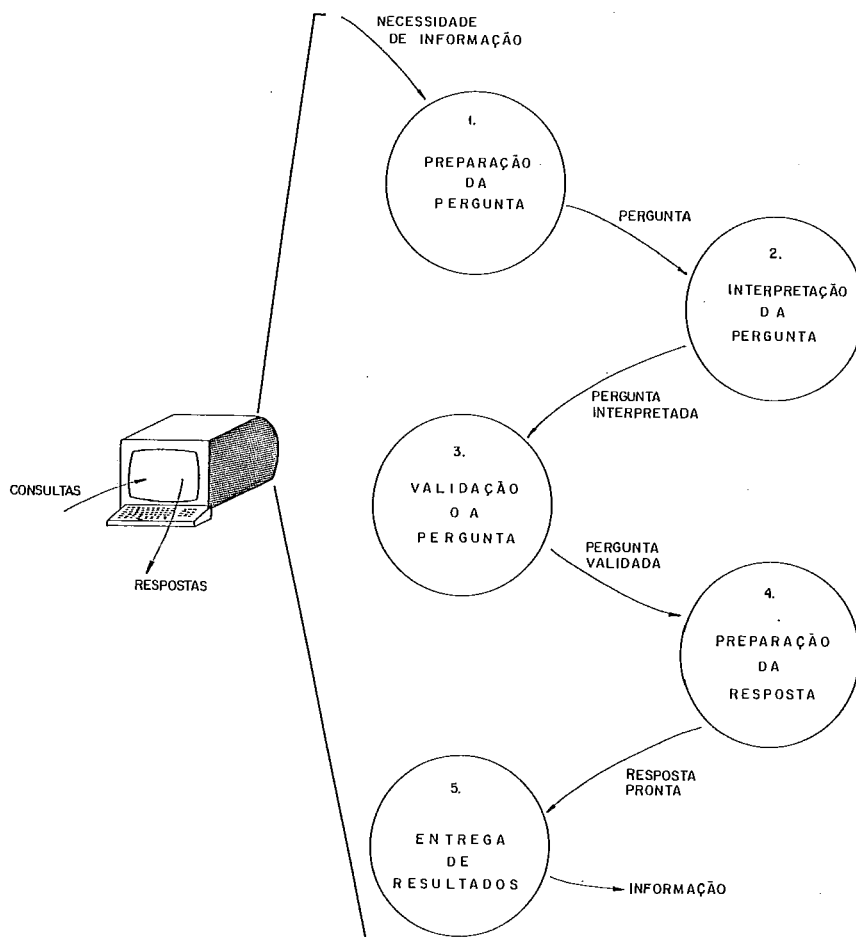


Fig. III.0

### III.1 PREPARAÇÃO DAS PERGUNTAS

A forma de preparar a pergunta depende da forma de diálogo permitida pela linguagem de consulta. As duas formas básicas são:

- Diálogo Orientado ou Menu de Opções

É oferecido ao usuário uma lista de funções disponíveis para que este selecione a que mais lhe convier, por indicação do número ou letra da opção escolhida.

A seguir é aberto novo conjunto de opções sobre a inicial, levando o usuário a percorrer um longo (e às vezes desnecessário) caminho de uma sintaxe pré-fixada. As respostas do usuário poderão sofrer variações quer pelo fornecimento de informações, quer pela indicação de respostas simples do tipo SIM(S) ou NÃO(N).

- Diálogo Livre ou de Alto Nível

É oferecido ao usuário um conjunto de comandos dotados de uma sintaxe própria, composta de termos da linguagem natural e de termos de alguma linguagem técnica orientada para o usuário.

Neste caso a preparação da pergunta compreende o fornecimento de comandos e dados, combinados adequadamente.

A existência de uma sintaxe próxima da linguagem natural facilita a preparação da pergunta. Entretanto mesmo esta sintaxe demanda um mínimo de estudo para sua plena compreensão. Uma solução para evitar essa necessidade é a implantação de rotinas de socorro (veja III.1.3), que auxiliam o usuário na preparação das perguntas.

- DIALOGO LIVRE OU DE ALTO NIVEL

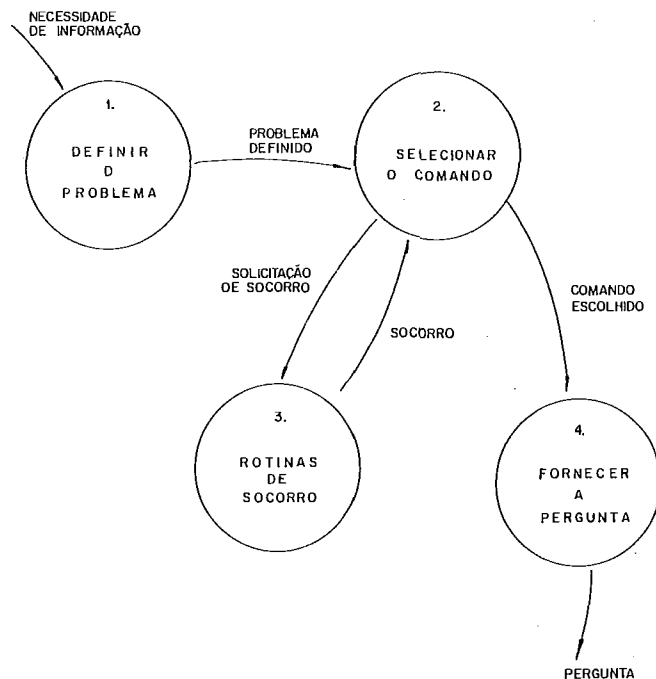


Fig. III.1

### III.1.1 DEFINIR O PROBLEMA

Consiste na análise pelo usuário dos contornos reais do seu problema caracterizando quais os resultados que deverão ser obtidos através da linguagem de consulta.

### III.1.2 SELEÇÃO DO COMANDO

É o estudo da documentação da linguagem para formar a pergunta, sintaticamente correta.

### III.1.3 ROTINAS DE SOCORRO

O usuário poderá necessitar de ajuda na preparação da pergunta no sentido de esclarecer detalhes sobre determinado comando ou função. Nestes casos é útil a existência de comandos especiais de ajuda, reunidos por exemplo num comando de SOCORRO. O comando SOCORRO poderá fornecer uma descrição de todos os recursos da linguagem de forma seletiva (através de menu por exemplo) ou ainda indicar referências bibliográficas adequadas e os telefones das equipes de suporte operacional.

### III.1.4 FORNECER A PERGUNTA

É o processo de digitar a pergunta em algum equipamento de interação segundo as regras de sintaxe da linguagem de consulta.

### III.1.5 OUTROS RECURSOS NECESSÁRIOS NA PREPARAÇÃO DA PERGUNTA

Dado que a funcionalidade de uma linguagem de consulta depende do grau de dificuldade que terá o usuário para preparar a pergunta, damos a seguir algumas características das linguagens que visam facilitar este processo.

- Forma Abreviada dos Comandos
  - A linguagem deve prover 2 formas de comandos
    - Completa
    - Abreviada

A forma completa será utilizada pelos usuários inexperientes, enquanto que a forma abreviada será utilizada por aqueles mais experientes como forma de agilizar consultas e permitir-lhes que formem perguntas mais longas.

- **Variações de Grafia**

A linguagem deve permitir na formação da pergunta o uso de variações de grafia comuns na linguagem natural ou técnica.

Exemplo: JAN – EQUIVALE A – JANEIRO  
 MTS – EQUIVALE A – METROS

- **Consultas Simples ou Grupadas**

A linguagem deve oferecer a possibilidade de segmentação da pergunta. Assim o resultado de uma pergunta poderá ser utilizado por perguntas subseqüentes, que agirão sobre o conjunto de informações obtido pela primeira, ou adicionarão a estes novos dados.

Este recurso é de particular utilidade na redução de erros de sintaxe e rapidez de processamento das perguntas, além de permitir ao usuário uma segmentação de idéias e controle sobre os resultados obtidos.

- **Operadores Aritméticos e Lógicos**

A linguagem deve ser provida de operadores que permitam a composição de perguntas com base nos valores dos campos, tais como:

– Operadores Aritméticos – = > < ≥ ≤  
 – Operadores Lógicos – AND OR NOT

Os operadores aritméticos são utilizados em campos numéricos e datas, sendo essenciais em operações de seleção de informações através de campos destes tipos.

Os operadores lógicos são utilizados para expandir a pergunta ou para interceptar conceitos como método para expandir a pergunta, ou, ainda, para excluir da resposta termos indesejáveis.

- **Condição de Pertinência e Existência**

Operadores especiais que permitem a seleção de informações sobre elementos que possuem um determinado campo com valores dentro de uma faixa ou que simplesmente possuam o campo em questão preenchido com informações.

Exemplo: ... Alunos com Notas entre 5 e 10  
 ... Alunos que possuam C.P.F.

- Função de Descrição dos Campos de Informação

É a função que descreve os campos de informação, fornecendo:

- Códigos dos Campos
- Descrição Detalhada dos Campos
- Tamanho em Caracteres
- Estrutura dos Campos
- Tipo do Campo (se o campo é chave ou não)

Estes dados permitem ao usuário uma melhor composição de suas perguntas.

- Utilização de Sufixos

Nas aplicações de biblioteconomia é comum a necessidade de recuperação de títulos através de palavras-chaves do título. Isto é obtido pela recuperação de todos os títulos que possuam um determinado sufixo, passando o usuário a formar a sua pergunta após a obtenção da palavra chave. Este conceito pode ser estendido a qualquer sufixo dos títulos, tendo o usuário a possibilidade de indicar quantas letras comporão o sufixo ou raiz, sobre a palavra oferecida como argumento da pesquisa.

- Mapeamento de Arquivos por “Strings” de Caracteres

Em algumas aplicações de tratamento de textos é útil que o acesso possa ser feito por qualquer conjunto de caracteres (strings) contidos em campos chaves ou não.

- Função que Produz um Dicionário de Acesso

É a função que fornece informações para análise dos campos, através das seguintes características:

- Número de ocorrências do campo
- Estatística de utilização
- Faixa de valores
- Tempo de acesso médio

### III.2 INTERPRETAÇÃO DAS PERGUNTAS

É o processo de análise, classificação e transformação das perguntas da forma fornecida pelo usuário para uma representação interna que traduza o significado conceitual da pergunta. O processo se compõe de várias etapas, e a interpretação em cada etapa deve ser concluída com sucesso para que se passe a etapa seguinte. Quando numa etapa são detetados um ou mais erros, o processo é interrompido, acionando-se as rotinas de tratamento de erros da linguagem de consulta.

Este processo é dependente da linguagem natural utilizada pois os significados dos termos utilizados podem variar de uma linguagem natural para outra.

É também sensível ao contexto em que a pergunta se insere (veja II.6) já que o significado conceitual da pergunta só fica completo após a interferência dos dados de contexto.

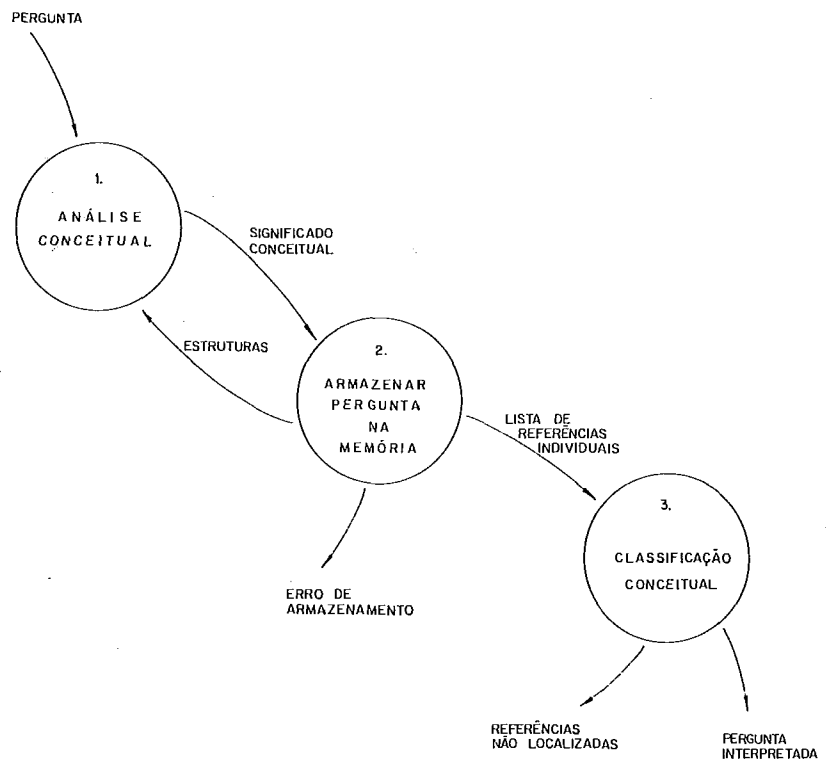


Fig. III.2

### III.2.1 ANÁLISE CONCEITUAL

Consiste na obtenção do significado conceitual da pergunta através da análise da estrutura da pergunta e da sua sintaxe de composição. Ao significado conceitual obtido dos comandos e parâmetros fornecidos na pergunta são acrescentadas as inferências dos dados de contexto. A pergunta sofre uma transformação para uma representação interna que é armazenada na memória, consistindo de listas de referências individualizadas a rotinas da linguagem, estruturas lógicas da base de dados ou informações da base de dados para recuperação simples, armazenamento de novos dados ou alteração dos dados existentes.

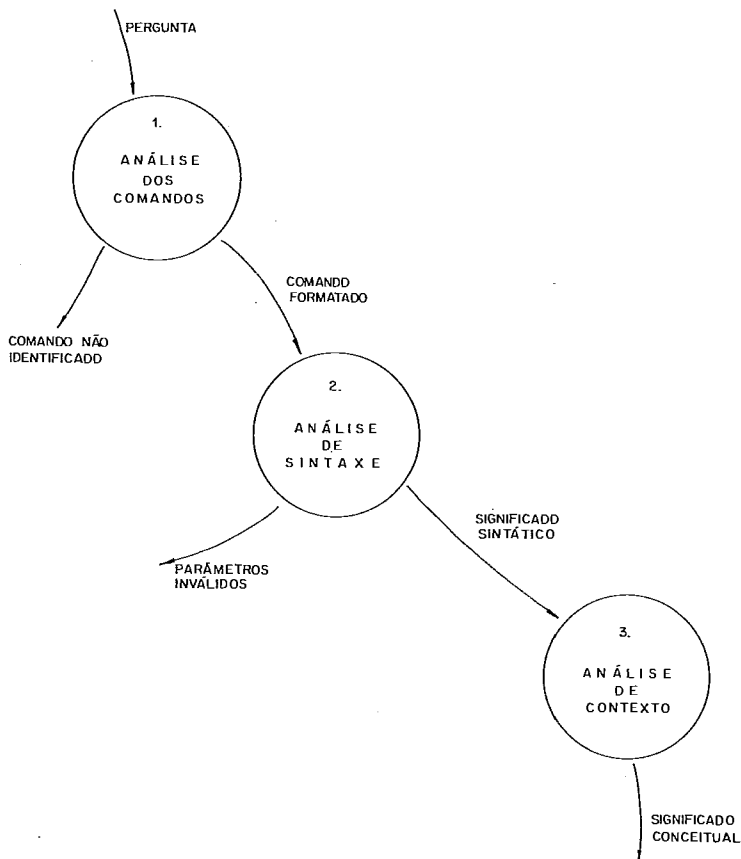


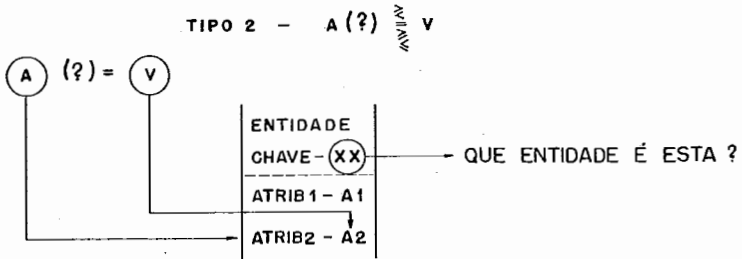
Fig. III.2.1.





- Tipo 2

Dado o valor de um atributo (A), quais as entidades (E) que o possuem ?



- Variação do Tipo 2

Definido um atributo (A), quais os valores (V) que ele assume, e quais as entidades (E) que o possuem?

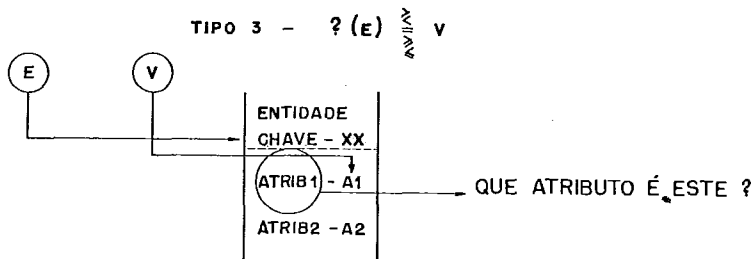
- Tipo 3

Dada uma entidade (E) e um valor (V) para atributo, quais os atributos que possuem este valor ?

Observação:

Este é um tipo raro de pergunta, usado quando diversos atributos descrevem a mesma propriedade da entidade em diferentes períodos de tempo.

Exemplo: Dado o empregado de matrícula número 2622 e salários no valor de Cr\$ 15.000,00 — Em que anos este valor foi excedido ?



— Variação do Tipo 3

Dado um valor (V) para atributos, quais as atributos que os possuem, e quais as suas entidades ?

Exemplo: Forneça os nomes de todos os empregados que tenham recebido mais do que Cr\$ 25.000,00 em qualquer ano.

— Outros Tipos

Existem outros tipos de perguntas, que podem ser reduzidos a estes três, pois são obtidos pela combinação de quaisquer destes tipos com variações na qualificação dos seus atributos, tais como:

— Uso de Múltiplos Valores para os Atributos

Caso em que os atributos são combinados em função de faixas de valores.

Exemplo: Forneça os nomes dos empregados com salários entre Cr\$ 20.000,00 e Cr\$ 30.000,00.

— Ordenação do Resultado

Caso em que o resultado a ser produzido deve ser fornecido ordenado por algum atributo (ou atributos).

Exemplo: Forneça os nomes de todos os empregados classificados pelo primeiro nome.

— Condições Lógicas

Caso em que a entidade é selecionada por condições impostas aos seus atributos.

Exemplo: Forneça os nomes dos empregados com salário maior do que Cr\$ 20.000,00 e idade menor do que 25 anos.

Após esta classificação num dos tipos de pergunta, obtém-se a formatação interna da pergunta e a remoção dos termos que são colocados na pergunta apenas para facilitar a sua legibilidade.

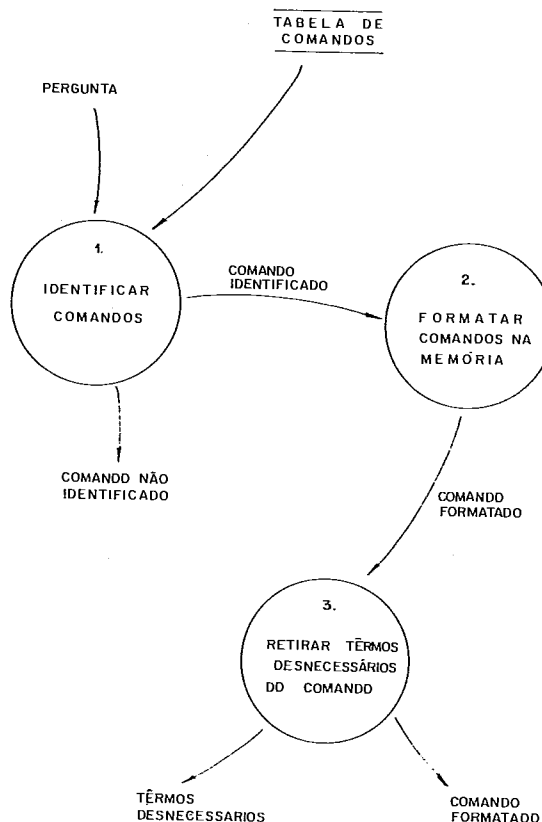


Fig. III.2.1.1

### III.2.1.2 ANÁLISE DE SINTAXE

É o processo de análise das partes do comando para identificação do significado sintático e detecção de erros.

Inicialmente é analisado o comando isolado, ou seja o comando separado de seus parâmetros. Para isso utiliza-se uma tabela de comandos válidos a partir da qual são obtidas as características do comando, que são informações que permitirão interpretar os parâmetros quanto a sua adequação ao comando.

Uma vez obtida a formatação interna do comando (ou comando transformado) faz-se um composição sintática deste com a formatação interna obtida para os parâmetros (ou parâmetros transformados). Esta composição é denominada "Significado Sintático" da pergunta. Os parâmetros transformados são os parâmetros fornecidos na pergunta traduzidos em valores, referência a informações da base de dados ou composições de outros parâmetros da própria pergunta ou de perguntas precedentes.

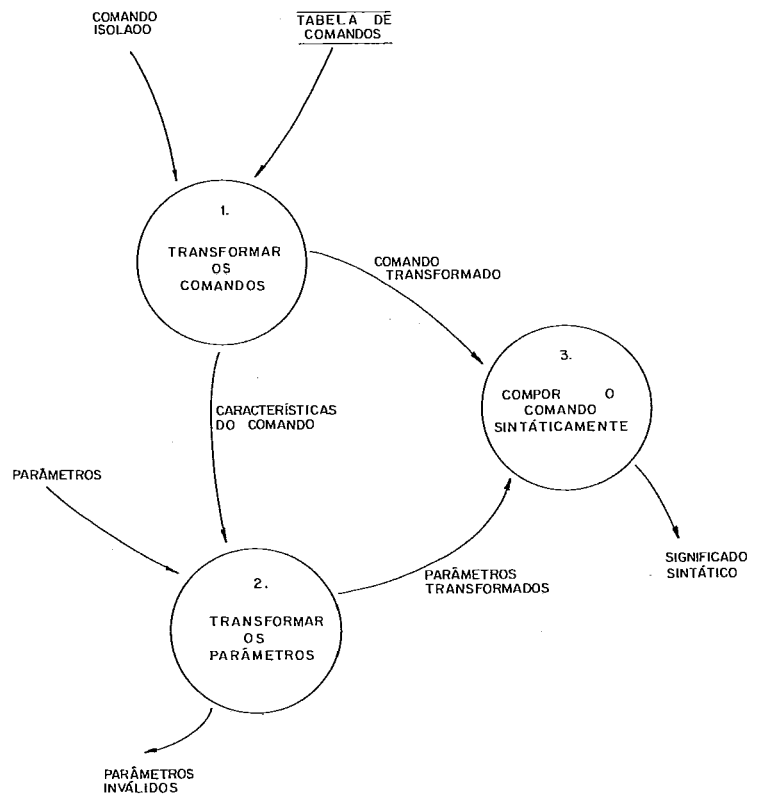


Fig. III.2.1.2

### III.2.1.3 ANÁLISE DE CONTEXTO

Consiste na atribuição à pergunta de dados oriundos do contexto em que ela se processa. Embora estes dados não tenham sido solicitados explicitamente pelo usuário eles tornam as respostas mais claras e objetivas.

Os dados de contexto são informações obtidas no arquivo de contexto relativos a pergunta analisada e se destinam a melhorar o nível de clareza das respostas (veja item II.6).

Ao término desta etapa obtém-se o "Significado Conceitual" da pergunta que é uma tradução da pergunta de seu estado original para uma representação interna contendo valores e referências externas obtidas na própria pergunta e/ou no arquivo de contexto.

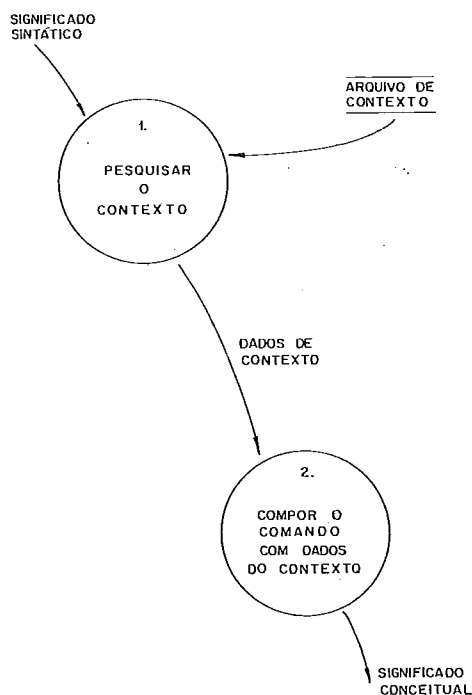


Fig. III.2.1.3

### III.2.2 ARMAZENAMENTO DA PERGUNTA NA MEMÓRIA

As perguntas são tratadas a partir de seu armazenamento na memória. Uma vez detetado o seu significado conceitual, a pergunta é armazenada na memória sob a forma de uma lista de referências individuais, contendo os pontos de real significado. Os pontos mais importantes são devidamente destacados, sendo ainda adicionados a esta lista as informações obtidas do contexto.

No caso de segmentação da pergunta, esta lista é expandida para armazenar os novos dados. Quando no processo de armazenamento são detetadas estruturas de dados fornecidas na pergunta, o processo de Análise Conceitual é reativado para interpretar a estrutura. Antes de acionar o processo de análise conceitual, a memória é pesquisada para verificar se a estrutura em questão já foi anteriormente interpretada.

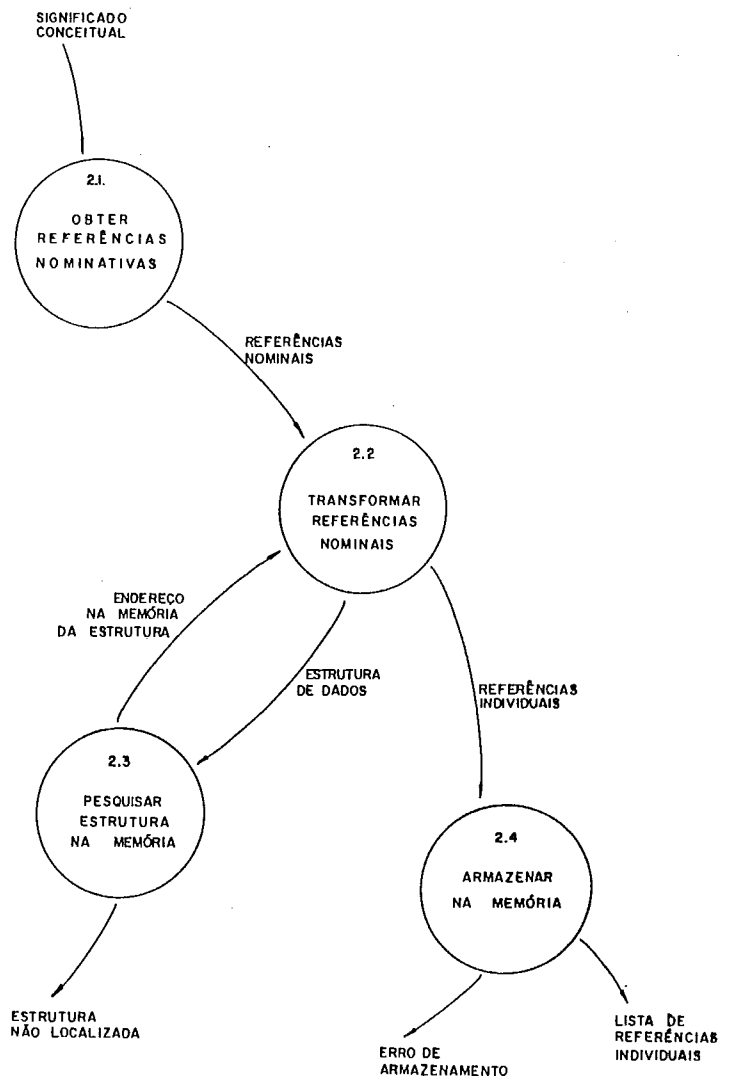


Fig. III.2.2

### III.2.3 CLASSIFICAÇÃO CONCEITUAL

Consiste na deteção do tipo de resultado desejado, quais sejam:

- Respostas às Perguntas  
Informações obtidas da base de dados ou do contexto.
- Ações sobre a Base de Dados  
Manipulações da base de dados.

Esta classificação é possível pois no processo anterior foram obtidas as características do comando que determinam o tipo de ação a ser executada.

Antes da execução da ação do comando, processa-se uma validação das referências explicitadas no comando, detetando-se aquelas não localizadas na memória e deflagrando-se as rotinas de busca ou de composição de resultados.

Os comandos que necessitam de respostas, são aqueles que através de referências a partes específicas da base de dados solicitam informações quantificáveis sob a forma de valores de textos.

Os comandos que executam ações sobre a base de dados são aqueles que alteram a base de dados a partir de informações fornecidas na própria pergunta ou por referências a meios físicos de entrada de dados.

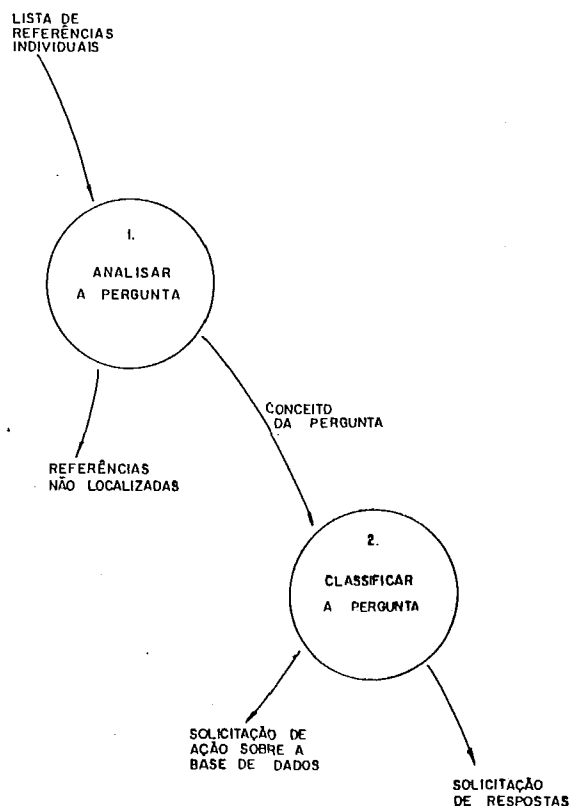


Fig. III.2.3

### III.3 CONFIRMAÇÃO DAS PERGUNTAS

A confirmação da pergunta é obtida pela aprovação do usuário ao diagnóstico da pergunta formada. Estes diagnósticos devem ser bastante claros não só na indicação dos erros como também quanto ao entendimento obtido da pergunta fornecida pelo usuário. Os diagnósticos podem ser produzidos com as seguintes características:

- **Diagnóstico de Erros**  
Quando não seja possível reconhecer algum termo fornecido, a linguagem deve ser a mais explícita possível, indicando qual o termo ininteligível ou mal formado. Uma outra facilidade é a possibilidade de alterar apenas parte da pergunta, sem obrigar o usuário a fornecer por completo a pergunta (WAL78). Isto pode ser facilmente obtido nos terminais de vídeo, separando-se na tela uma região para perguntas e outra para diagnósticos.
- **Significado da Pergunta**  
Quando a pergunta estiver sintaticamente correta ela deve ser mostrada ao usuário para sua confirmação. Isto é necessário pois não raro, por problemas de transmissão de dados, alguns dados são fornecidos mal compostos ou deformados por interferências da transmissão.
- **Resultado Prévio**  
Quando for possível é útil fornecer ao usuário uma estimativa de consumo de recursos necessários para a obtenção dos resultados, tais como o número de linhas a serem impressas no resultado e/ou o tempo/custo provável da pergunta.  
Nestes casos o usuário deve ter a opção de:
  - Deflagar o Processamento
  - Recompilar a Pergunta para Reduzir o Consumo de Recursos.



### III.4 RECOMPOSIÇÃO DAS PERGUNTAS

Nas situações em que a pergunta seja considerada errada ou mal formada, é necessário que o usuário possa recompô-la total ou parcialmente.

A recomposição total só é recomendada em casos extremos (muitos erros numa só pergunta) ou quando o usuário cancelar a pergunta.

#### III.4.1 RECOMPOSIÇÃO PARCIAL

A recomposição parcial pode ser aceita pela referência a partes da pergunta seguidas do seu novo conteúdo.

- Exemplo:
- \* Qual o NUMER de empregados da empresa
  - Pergunta recusada TERMO/NUMER/IRRECONHECÍVEL
  - \* NUMER – NUMERO
  - Pergunta correta
  - Qual o número de empregados da empresa
  - 2.520 empregados.

Outra forma de recomposição parcial é aquela permitida pelos próprios terminais de vídeo que possibilitam a correção dos termos diretamente na tela e sua resubmissão ao sistema.

#### III.4.2 IDENTIFICAÇÃO DE TERMOS COM ERRO DE GRAFIA NUMA PERGUNTA

Ocorre com freqüência que os interpretadores de pergunta, recusem toda uma longa frase, porque dentro desta uma ou mais palavras foram mal escritas por omissão ou troca de caracteres, ou mesmo por adição de caracteres estranhos a palavra. Nestes casos, a simples recusa de toda a frase embora seja de implementação mais simples é óbvia leva o usuário a um duplo esforço no sentido de concentrar-se na digitação da pergunta a fim de evitar erros, pois estes lhe acarretarão a penosa redigitação de todo o texto.

Uma alternativa implementável é a interpretação de palavras por semelhança de grafia, oferecendo-se ao usuário a opção de corrigir apenas as palavras ininteligíveis ou, a seu critério, todo o texto.

Exemplo: Seja a pergunta

\* QUAL O PREÇO MÉDIO DO BATATO INGLESA NO ESTADO DO PARANO?

Como os termos BATATO INGLESA e PARANO são necessários para a consulta, a linguagem de consulta poderá solicitar ao usuário apenas a correção destes termos:

— Termos não reconhecidos: BATATO INGLESA  
PARANO

- Opções para Correção
- BATATO INGLESA
1. BATATA INGLESA
  2. BANANA INGLESA
  3. OUTRO TERMO
- PARANO
1. PARA
  2. PARANA
  3. OUTRO TERMO

### III.4.3 INCLUSÃO DE TERMOS QUE FALTAM NUMA PERGUNTA

Em algumas perguntas a sintaxe fica incompleta pela ausência de algum termo que identifique o objeto da pergunta e que lhe daria o sentido lógico.

Poderá ocorrer inclusive que o usuário por inexperiência com o sistema não forneça todos os elementos necessários, ou que se esqueça de alguns destes. Assim sendo ao permitir que à pergunta sejam adicionados termos faltantes, oferece-se um caminho para a análise parcial da pergunta e uma forma de treinamento ao usuário iniciante.

Exemplo: Seja a pergunta

\* QUAL O SALÁRIO-HORA PERCEBIDO PELO EMPREGADO DE MATRÍCULA NÚMERO 37887-42 ?

A resposta desta pergunta depende do período de trabalho pelo funcionário, logo a linguagem de consulta poderá solicitar ao usuário que adicione o termo que falta.

- \* FALTA NA PERGUNTA A INDICAÇÃO DO PERÍODO QUE PODERÁ SER DA FORMA
1. ANO DE ...
  2. MESES DE ... A ...
  3. MÊS DE ...
  4. PERÍODO DE .../.../... A .../.../...
  5. NO DIA .../.../...

### III.5 SEGMENTAÇÃO DAS PERGUNTAS

É a situação em que após a obtenção da resposta de uma pergunta, utiliza-se a mesma pergunta como base de uma nova pergunta.

Exemplo: \* QUAL A MÉDIA DE SALÁRIOS DA EMPRESA ?  
 — Cr\$ 48.332,10  
 \* E O MAIOR ?  
 — Cr\$ 172.800,00  
 \* E O MENOR ?  
 — Cr\$ 12.406,40

Nos casos de segmentação o usuário fornece o sinal \* , para indicar que se trata de pergunta segmentada. Pode-se ainda utilizar referências a perguntas identificadas, da forma:

\* 1. QUAL A MÉDIA DE SALÁRIOS DA EMPRESA ?  
 — Cr\$ 48.332,10  
 \* 2. QUAL A MÉDIA DE VENDAS DA EMPRESA ?  
 — Cr\$ 73.484.716.839,22  
 \*\* QUAL A RELAÇÃO ENTRE OS RESULTADOS \*1 e \*2 ?  
 — 0,05

### III.6 INFORMAÇÕES PRELIMINARES SOBRE AS RESPOSTAS

Quando o volume de dados produzidos como resposta for muito grande, ou se o tempo para produzir os resultados for proibitivo, é necessário alertar o usuário, dando-lhe opções de evitar o processamento ou diminuir o alcance de suas perguntas.

— SERÃO PRODUZIDAS 10.000 LINHAS, INFORME ONDE IMPRIMIR.

No caso de longos tempos de processamento o usuário deve ter a opção de executar o comando à noite quando a operação possa ser feita sem prejuízo para o sistema e a custos menores.

### III.7 INFORMAÇÕES DE CONTROLE SOBRE O PROCESSAMENTO

Durante o processamento da pergunta em processos interativos é útil permitir ao usuário alguma intervenção no processamento em algumas situações típicas.

- Resultados Demorados

Quando o usuário comanda a execução de um processamento que demande um longo tempo é necessário informá-lo periodicamente sobre o andamento dos trabalhos permitindo-lhe cancelar o processamento a qualquer tempo.

Exemplo:   PROCESSADOS 1.000 ITENS,  
                  QUER INTERVIR NO PROCESSAMENTO ?

- Diagnóstico de Erros Lógicos

Há casos em que uma pergunta mal-sucedida devido a erros de dados produz condições falhas na pergunta do usuário. Nestes casos, antes de processar a pergunta parcialmente, o que poderá ser danoso ou mesmo inútil, deve-se dar ao usuário informações de controle e permitir que ele intervenha no processo, cancelando ou alterando a pergunta.

Exemplo:   NÃO ENCONTREI NENHUM EMPREGADO COM SALÁRIO ENTRE  
                  CR\$ 10.000 E CR\$ 12.000, POSSO PROSEGUIR ?

- Alternativas de Saídas

No decurso do fornecimento dos resultados é às vezes necessário interromper a saída em um meio e continuá-lo posteriormente, ou recomeçá-lo em outro meio melhor situado em termos dos objetivos, ou ainda de maior velocidade que o primeiro. Isto geralmente é obtido por meio de caracteres especiais que quando digitados no terminal, retornam o controle ao usuário. Evidentemente deve-se permitir que o usuário inclusive cancele a saída dos resultados produzidos.

### III.8 PREPARAÇÃO DAS RESPOSTAS E ENTREGA DOS RESULTADOS

O conhecimento que o homem adquire do mundo real e que constituem os elementos de respostas possíveis são obtidos a partir de informações fornecidas pelo próprio homem. Estas informações são armazenadas em algum dispositivo de armazenamento de dados para posterior devolução.

#### III.8.1 FORMAS DE OBTENÇÃO DE RESULTADOS

Existem quatro processos de obtenção de resultados.

- **Informações em seu Estado Natural**

É o mecanismo de respostas mais simples que há, e consiste em devolver ao usuário as informações previamente armazenadas, sem nenhum tratamento computacional.

- **Resultados de Cálculos Imediatos**

Este método consiste em devolver ao usuário resultados de funções da linguagem de consulta aplicadas sobre:

- Dados armazenados individualmente
- Séries de dados armazenados
- Dados fornecidos na pergunta (ou séries de dados)

- **Resultados Previamente Calculados**

É o método em que a linguagem de consulta armazena (temporária ou definitivamente) os resultados de cálculos executados sobre uma informação ou grupo destas, passando a assumir os resultados armazenados como verdadeiros para utilizá-los posteriormente como respostas a perguntas idênticas (na mesma sessão ou em sessões subsequentes). Neste caso surge o clássico problema de inconsistência da informação, vez que os dados utilizados para o cálculo podem ser alterados ao longo do tempo, e a linguagem de consulta deve prover a atualização também dos resultados de cálculos efetuados utilizando estes dados.

- **Composição de Valores**

É o processo de junção de dois ou mais grupos de informações, tomando como base um atributo comum.

### III.8.2 FORMA DE ENTREGA DE RESULTADOS

Os resultados obtidos podem ser entregues diretamente no veículo de entrada da pergunta ou, a critério do usuário em outro veículo mais adequado.

Os fatores decisivos neste caso são:

- Volume de Resultados
- Características Especiais dos Resultados
  - Gráficos
  - Sinais Elétricos/Luminosos/Eletrônicos

### III.8.3 TRATAMENTO DOS RESULTADOS APÓS A ENTREGA

Ao final da entrega dos resultados, estes não devem ser imediatamente destruídos, oferecendo-se antes ao usuário a possibilidade de armazená-los noutra meio, ou ainda de tratá-los convenientemente para uso em outras aplicações, por exemplo como entrada de dados de outras consultas ou mesmo outros Sistemas.

### III.9 TREINAMENTO DO USUÁRIO

A linguagem deve possuir comandos de orientação básica ao usuário que lhe permita conhecer a linguagem não só em termos de seus comandos como também quanto a:

- Características Básicas da Linguagem
- Potencialidades Especiais
- Repertório de Exemplos Comuns de Aplicação de Bases de Dados

É de particular interesse o fornecimento de completa bibliografia sobre os aspectos mais complexos da linguagem.

### III.10 CONCLUSÃO

Apresentamos neste capítulo, o modelo lógico e as características desejáveis de uma linguagem de consulta. Estas proposições, muitas vezes não são implementadas nos sistemas interativos devido ao receio generalizado de que estas características tornem os sistemas lentos e assim pouco produtivos. Entretanto, isto não é sempre verdade. Muitos são os casos em que a "performance" é realmente fundamental para o sistema e que portanto deve-se sacrificar a flexibilidade do diálogo em benefício desta; mas também ocorre com frequência que os técnicos de PD por desconhecimento das formas de implementação, ou descaso quanto a finalidade do Sistema, não implementam as facilidades de diálogo aqui sugeridas. No primeiro caso sugere-se que o técnico em PD reavalie o problema a ser resolvido pelo Sistema, procurando identificar no âmbito da organização os usuários finais e classificá-los segundo sua necessidade de "diálogo", antes de implementar o sistema apenas sob o enfoque de "performance" operacional. No segundo caso, sugere-se que estes técnicos de PD reconsiderem suas opiniões sobre a correta maneira de abordar o problema do usuário, pois acima de qualquer objetivo pessoal deverá estar o da Empresa a que servem. Nos casos de desconhecimento técnico do assunto, a própria Tese e a bibliografia selecionada poderão servir como base para domínio destas técnicas.

## CAP. IV CONCLUSÕES E RECOMENDAÇÕES

### IV.1 IMPLEMENTAÇÕES EM SISTEMAS ATUAIS

As implementações atuais se originam de 4 fontes:

- Sistemas Operacionais de Tempo Compartilhado
- Sistemas de Gerência de Bases de Dados
- Sistemas de Recuperação de Informações
- Sistemas de Consulta a Bases de Dados

#### IV.1.1 SISTEMAS OPERACIONAIS DE TEMPO COMPARTILHADO

Os sistemas operacionais de tempo compartilhado oferecem rotinas especiais para execução de suas tarefas, cuja manipulação é feita via de regra por uma linguagem de consulta implementada sob a forma de menu ou em alguns casos, de comandos de sintaxe simples e direta. Observa-se ainda um grande esforço no sentido de tornar interativas as tarefas de elaboração, depuração e testes de programas. Os sistemas operacionais analisados foram os seguintes:

- RSTS/E da DEC

É um sistema operacional de tempo compartilhado que possibilita a interação com o usuário através de uma linguagem de programação interativa — BASIC-PLUS, e de utilitários específicos para cada atividade, tais como — edição de texto, formatação de arquivos, mensagens entre terminais . . .

- CMS da IBM

É um monitor de interação com o usuário que é processado sob o sistema operacional VM, e dispõe de rotinas especializadas para as atividades normais de processamento de dados tais como — edição de texto, formatação de arquivos, comunicação entre terminais — que são assim executadas interativamente através de terminais.

- MUMPS

É um sistema operacional de ação interativa que permite a execução de todas as operações normais de processamento de dados, dispondo ainda de estruturas de arquivos que permitem a recuperação de informações e o seu manuseio de forma interativa.



#### IV.1.2 SISTEMAS DE GERENCIA DE BASES DE DADOS

Os sistemas de gerência de bases de dados em geral, dispõem de linguagens procedurais executáveis através de linguagens hospedeiras e apenas em alguns casos possuem linguagens de consulta que podem ser utilizadas independentemente (as chamadas linguagens auto-contidas na literatura desses sistemas). Estas linguagens embora dotadas de alguma flexibilidade quanto a forma de composição das perguntas, em virtude da generalização a que se propõem, não oferecem os recursos da ANALISE DE CONTEXTO. Os sistemas de gerência de bancos de dados considerados foram:

- SYSTEM 2000

O Sistema 2000 é um sistema de gerenciamento de bases de dados de grande porte que adota para suas bases de dados a organização hierárquica, mas permite através de ligações lógicas entre os registros a formação de estruturas de rede. Oferece 3 alternativas de consulta às bases de dados — Procedure Language Interface, Report Writer e a Immediate Acces. A primeira se destina ao acesso a bases de dados através de linguagens hospedeiras, a segunda para a produção de relatórios e a terceira, para a consulta através de perguntas formadas pelo próprio usuário através de terminal.

- ADABAS

O ADABAS é o sistema de gerenciamento de bases de dados mais moderno que atualmente está disponível no mercado. Adota para suas bases de dados estruturas dinâmicas que tanto podem ser hierárquicas como em rede, oferecendo inúmeras opções de tratamento, armazenamento e segurança de dados. Oferece, da mesma forma que o SYSTEM 2000, três alternativas de consulta, sendo o ADASCRIP a sua linguagem de consulta interativa que permite ao usuário a formação de perguntas diretamente no terminal.

- TOTAL

O TOTAL é o sistema de gerenciamento de bases de dados mais difundido atualmente, e se caracteriza pela simplicidade de suas soluções. Adota para suas bases de dados a organização em rede e oferece basicamente o acesso a bases de dados através de linguagens hospedeiras. Recentemente o fornecedor tem efetuado esforços no sentido de dotar o TOTAL de recursos mais dinâmicos para consulta as bases de dados, num esforço de atualização tecnológica. O fruto desse esforço e o lançamento de uma linguagem de consulta denominada T-ALK, cujas características se assemelham às das linguagens de consulta dos outros sistemas de gerencia de bases de dados.

### IV.1.3 SISTEMAS DE RECUPERAÇÃO DE INFORMAÇÕES

Os sistemas de recuperação de informações enfatizam o processamento de busca de dados armazenados, sem consideração sobre a semântica das perguntas ou do significado conceitual destas. As linguagens oferecidas, embora dotadas de alguma flexibilidade só permitem obter funções executivas simples. Os sistemas de recuperação de informação considerados foram:

- CICS

É um programa-produto da IBM que se destina a facilitar a intercomunicação entre as bases de dados produzidas no ambiente IBM e os terminais remotos de uma rede de TP. Possui uma linguagem de programação própria de elevada complexidade, formada por macros-comandos utilizados através de uma linguagem hospedeira (COBOL, PL/I). Os sistemas de informação produzidos utilizando o CICS é que poderiam ser considerados como uma linguagem de consulta a bases de dados, desde que observadas as técnicas mostradas anteriormente, e não o próprio CICS.

- STAIRS

É um programa-produto da IBM, produzido com a finalidade de gerenciar bases de dados de grande porte formadas sob o CICS, principalmente em funções de recuperação de textos através de sufixos, ou seja por varredura de toda a base de dados para formação de acervos menores para uma busca seletiva da informação desejada. Da mesma forma que o CICS a consulta é feita através de linguagem hospedeira, o que leva a conclusão de que somente os sistemas produzidos através do STAIRS é que podem ser considerados como uma linguagem de consulta.

- RIQS

É um sistema de recuperação de informações produzido para gerar sistemas de aplicação voltados para a consulta de bases de dados. É contemporâneo do STAIRS da IBM, mas diversamente daquele, dispõe de recursos para consulta direta do usuário as bases de dados através de terminal, não necessariamente através de linguagem hospedeira.

#### IV.1.4 SISTEMAS DE CONSULTA A BASE DE DADOS

São sistemas de apoio a aplicações específicas, que suprem todas as funções de consulta as bases de dados para as quais foram projetadas sem o concurso de nenhum outro sistema, exceto funções do sistema operacional sob os quais são executados.

O principal mérito destas linguagens é a autonomia com que o usuário final manuseia as bases de dados de seu interesse com pouca ou mesmo nenhuma intervenção dos técnicos de PD.

- SHRLDU (WIN75)

A literatura sobre o SHRDLU foi publicada em 1972 como sendo o primeiro programa de computador para entender Inglês. O universo de conhecimento do programa está limitado a blocos de palavras. O SHRLDU simula a manipulação de blocos sobre uma mesa e pode atender a solicitações do tipo:

– POR FAVOR COLOQUE A PIRÂMIDE VERMELHA DENTRO DO BLOCO AZUL

Ou perguntas do tipo:

– ONDE ESTÁ A PIRÂMIDE VERMELHA ?

Segundo seu autor, o SHRDLU é um programa baseado na “Representação Procedural da Linguagem”. Ele considera que o conhecimento necessário para o processamento de uma linguagem natural pode ser representado por “Procedures” dentro do sistema. No SHRLDU, o entendimento de – POR FAVOR PEGUE O BLOCO – consiste na execução do procedimento que simula a tomada do bloco. Esta abordagem de representar (simular) conhecimento, tem forte dependência do domínio das tarefas simuladas. No contexto de aplicações não executivas, esta representação falha completamente.

- LSNLIS (WOO72)

A LSNLIS – Lunar Sciences Natural Language Information Systems, é um protótipo de linguagem natural, projetado para acessar uma base de dados de informações técnicas sobre amostras de rochas lunares coletadas na missão Apolo 11. A LSNLIS utiliza como analisador uma rede de transição aumentada que transforma perguntas feitas em Inglês numa árvore de pesquisa. Um processador de semântica então transforma a árvore de pesquisa sintática numa representação de significado semântico. A representação semântica para a pergunta é atualmente uma expressão de programa executada diretamente para recuperação na memória. A representação na memória e o processo de semântica usado no LSNLIS é bastante adaptado aos dados técnicos para os quais o sistema foi projetado. Entretanto existem imperfeições na heurística de recuperação e no interpretador de semântica, pois não foi adotada a técnica de dedução por análise dos tipos de dados. Assim sendo a cada pergunta faz-se necessária a varredura de todos os itens da BD.

- PLANES (WAL78)

É um sistema de responder perguntas preparado para acessar uma base de dados com informações técnicas sobre manutenção de aviões da marinha americana e sobre vôos. O PLANES parece similar ao LSNLIS, exceto que o analisador inicial substitui a representação da árvore sintática por uma expressão representativa de frases canônicas. Esta representação é então mostrada ao usuário para confirmação, antes que uma fase de interpretação transforme a representação numa expressão formal da linguagem de consulta, que é diretamente executada na memória.

O PLANES é dito ser um sistema independente dos dados. Isto significa que a linguagem de consulta não será substancialmente alterada para acomodar uma extensão da base de dados ou mesmo a manipulação de uma outra base de dados. A alteração será basicamente a adição de um novo vocabulário. Por outro lado os programas de acesso dos dados deverão ser substancialmente modificados para uma nova base de dados com funções de recuperação específicas. A generalidade do PLANES é pois limitada pelo tipo de dados que serão acessados.

- LUIGI (SCR75)

Este sistema simula a rotina de preparação de comida, como se fosse uma cozinheira capaz de responder perguntas como:

- QUE UTENSÍLIOS EU NECESSITO PARA TOSTAR O PÃO ?
- O QUE DEVE SER FEITO PARA COZINHAR UM OVO ?

Quando LUIGI recebe as questões, ele acessa a rotina adequada (TOSTAR O PÃO, COZINHAR UM OVO) e simula o processo especificado. Quando a simulação prove a informação apropriada, a pergunta é respondida. A teoria de representação de conhecimento utilizada pelo LUIGI é denominada pelo seu autor de "Conhecimento Orientado por Rotas".

- GUS (BOB77)

É um programa de diálogo interativo produzido para assumir as ações de um agente de viagens numa conversação objetiva com o cliente. GUS é composto de 4 módulos interativos:

- Analisador Morfológico
- Analisador Sintático
- Estruturador de Raciocínio
- Gerador de Linguagem

O Estruturador de Raciocínio do GUS é ponto central da pesquisa. A noção de estrutura no GUS é coerente com a formulação geral de estruturas propostas por MINKYS (MIN75). GUS utiliza o sistema de estruturas relacionadas com viagens para dirigir o diálogo e produzir as respostas ao tempo em que as perguntas são formuladas. As informações solicitadas possuem alguns valores que são assumidos automaticamente, como por exemplo o Nome do Aeroporto. O projeto para estruturas direcionadoras de diálogo proposta pelo GUS parece ser um promissor começo para um programa de conversação de múltiplas iniciativas.

- SEQUEL (AST75)

A SEQUEL – **Structured English Query Language (AST75)**, é a linguagem de consulta relacional mais desenvolvida atualmente. Originária da linguagem SQUARE – **Specifying Queries as Relational Expression (CHA75)**, executa as mesmas funções que está alterando no entanto a forma de diálogo já que permite a entrada de comandos formados por termos da língua inglesa, enquanto que a SQUARE utiliza basicamente uma notação matemática.

A SEQUEL tem uma versão já implementada, como interface do “Sistema R” da IBM, que é um sistema relacional de grande porte. Nessa versão a SEQUEL pode ser usada como uma linguagem de consulta autônoma, através de terminais, ou como uma linguagem de manipulação de dados executada a partir de uma linguagem hospedeira como por exemplo o PL/I. A SEQUEL permite além das funções de consulta uma completa faixa de ações sobre a base de dados e inúmeras funções de manipulação dos dados. No contexto das operações relacionais executa todas as operações da álgebra relacional. Possui ainda controles de segurança e integridade das informações através da fixação de condições para a alteração da base de dados ou de faixas de valores que determinados campos podem receber (CHA76).

- QUALM (LEH78)

A QUALM é um Sistema de consulta que apresenta uma teoria de entender e responder perguntas que é essencialmente a de um processador de linguagem natural. Esta perspectiva de linguagem natural distingue a QUALM dos Sistemas de Consulta que foram motivados por recuperação de informação ou solução automática de teoremas. Muitos sistemas que tentam responder perguntas escritas em linguagem natural, foram projetados em duas partes:

- Sistema de Recuperação na Memória
- Interface de Linguagem Natural

Freqüentemente o problema de interface é considerado secundário ao sistema de recuperação e os 2 subsistemas são projetados como se eles fossem teoricamente independentes um do outro.

O sistema de responder perguntas baseado no QUALM pode da mesma forma ser decomposto em recuperação na memória (QUALM propriamente dito) e o interface de linguagem natural (Analisador e Gerador). Mas no QUALM esta divisão distingue o processo dependente da linguagem daqueles que são puramente conceituais e independentes da linguagem.

A teoria do QUALM estende as teorias do processo de conhecimento que se originaram com os estudos de memória conceitual, análise (RIE75) e geração (GOL75). Por outro lado o aspecto fundamental do QUALM é o processamento conceitual das informações. Esta orientação caracteriza o abandono das propostas dos sistemas de recuperação de informação e de outros pontos de vista em que o processadores de linguagem são tratados como um simples “Front-end” dos sistemas de consulta.

- QUERY BY EXAMPLE (ZL075)

A QUERY BY EXAMPLE é uma linguagem de consulta relacional projetada para utilização através de Terminais interativos. O diálogo se processa através de tabelas desenhadas no terminal, contendo estruturas de consulta à Base de Dados, em cujas linhas e colunas o usuário fornece os comandos, parâmetros e operadores. Contemporânea da SEQUEL, a QUERY BY EXEMPLE oferece todos os recursos desta linguagem inclusive as funções de manipulação de dados e operadores relacionais. Possui ainda controles de segurança e integridade de informações. Além desses recursos oferece ainda operadores especiais de relações em árvores estruturadas. O que distingue a QUERY BY EXAMPLE das demais linguagens de consulta é que a sua estrutura "tabular" de consulta dá origem a uma "Sintaxe de duas dimensões", enquanto que as linguagens tradicionais possuem uma sintaxe linear. Esta "sintaxe de duas dimensões" oferece ao usuário uma extrema flexibilidade de composição de consultas (ou quaisquer operações), pois possibilita variadas formas de obtenção dos resultados através do adequado preenchimento das colunas com os comandos e operadores da linguagem. A QUERY BY EXAMPLE possui uma versão já implementada pela IBM, cujas características técnicas são semelhantes às apresentadas acima.

## IV.2 PREMISSAS SOBRE IMPLEMENTAÇÕES FUTURAS

Ao enumerarmos as características desejáveis de uma linguagem de consulta, estabelece-se os parâmetros de comparação com um modelo básico que é a própria linguagem natural. Assim sendo pelo grau de proximidade a esta e que se mede a capacidade conversacional oferecida pela linguagem.

Uma linguagem de consulta pode aproximar-se das linguagens naturais por 3 caminhos:

- **FLUIDEZ VERBAL**

Propriedade de uma linguagem que permite a coexistência de palavras de múltiplos significados em diferentes situações de sintaxe.

Esta característica das linguagens naturais é de difícil implementação em qualquer tipo de linguagem artificial devido a sua natureza tipicamente conceitual e temporal em que aspectos altamente complexos como entonação da voz, gestos, expressão facial, são considerados como parte do contexto de entendimento das perguntas e das respostas.

- **PROCESSADOR DE RESPOSTAS**

O Processador nas linguagens artificiais tem como paralelo nas linguagens naturais alguns mecanismos do cérebro humano, cujo processo de funcionamento não são ainda suficientemente conhecidos para que se possa implementá-los artificialmente.

- **INTERPRETAÇÃO DE PERGUNTAS**

Este mecanismo do cérebro humano é razoavelmente conhecido, e pelo menos teoricamente, consistiria na análise de modelos previamente armazenados na memória. Por essa razão este é o mecanismo mais aplicado em todas as linguagens de consulta existentes.

A implementação deste mecanismo é relativamente simples, e depende apenas do acervo de perguntas-respostas disponíveis na linguagem consulta.

As limitações existentes na pesquisa de modelos de Linguagem de Consulta levam que atualmente nem mesmo os interpretadores de perguntas sejam suficientemente próximos das linguagens naturais para que o sistema de informação que dispõe linguagens de consulta ofereça ao usuário atendimentos nos níveis desejados.

### IV.3 OBJETIVOS DE UMA LINGUAGEM DE CONSULTA

A Linguagem de Consulta deve refletir toda a potencialidade de um Sistema de informações, já que ela é o meio de ligação entre o Sistema e os usuários. Desta forma é de se esperar que todos os benefícios das técnicas empregadas no projeto da estrutura dos arquivos do Sistema venham a ser oferecidos aos usuários através da Linguagem.

É importante observar que muitos usuários de um Sistema de Informações podem conhecer a estrutura da informação em si, isto é se são hierárquicas, associativas, etc., mas por certo não terão interesse em detalhes da organização utilizada para a montagem de arquivos, isto é se são utilizadas listas investidas, multilistas ou quaisquer outros recursos técnicos. Isto se deve sobretudo ao fato de que os Sistemas de Informação tendem a se tornar um instrumento de trabalho comum na gestão das grandes Empresas, e os usuários estão interessados antes com os fins e objetivos da Empresa do que com a complexidade interna dos arquivos da Base de Dados utilizada.

É de interesse ressaltar ainda que os comandos de grande capacidade, baseados na sofisticação de equipamentos de armazenamento de dados, tais como discos, tambores ou de entrada e saída de dados tais como os modernos terminais, equipamentos de telecomunicação ou de processadores de alta velocidade, só podem ser oferecidos ao usuário não programador mediante as Linguagens de Consulta. É este pois o desafio maior oferecido pela Linguagem de Consulta ao projetista de um Sistema de Informação, dado que este deve ocultar toda a complexidade dos recursos de "software" e "hardware" disponíveis de modo que apenas a informação seja tratada pelo usuário.

Assim sendo pode-se definir como objetivo da Linguagem de Consulta, o de servir como "**interpretador de conceitos**" entre o homem e a máquina, dado que entre o intelecto humano e o Sistema automatizado (LEF69) existem apenas duas coisas em comum — A Estrutura da Informação (ou conhecimento desta estrutura) e a Linguagem de Consulta. No entanto é possível que cada usuário possua uma visão totalmente diversa de como armazenar os dados, daquela que realmente é empregada, sem que isto ofereça nenhum inconveniente de uso, e até que ao tempo da utilização este sinta que os dados estão armazenados da forma como ele imagina.

Além deste conceito de atendimento pleno ao usuário de sua demanda original e digamos costumeira, deverá a Linguagem prover meios de expandir conceitos, despertar no usuário o interesse pela pesquisa de outras minúcias e objetivos desconhecidos da informação, mantendo contudo uma coerência entre seus resultados. É este deslumbramento pelos aspectos antes inacessíveis da informação, agora possível, que faz da Linguagem de Consulta, o instrumento de trabalho do futuro, pois é possível agora um aprimoramento intenso da comunicação homem-máquina num nível de diálogo simples e rápido o que permite à mente humana um seqüenciamento de idéias mais efetivo e decerto mais produtivo.



#### IV.4 CONCLUSÃO

Em sua finalidade básica as LINGUAGENS DE CONSULTA permitem ao homem um melhor conhecimento do mundo real, dado que as pesquisas sobre as informações armazenadas são facilitadas e até incentivadas.

Nas seções precedentes procuramos analisar algumas características das Linguagens de Consulta que possibilitam esse incentivo e simplificação na interação Usuário-Sistema.

Entretanto, nos Sistemas de Aplicação, muitas vezes apenas algumas destas características são possíveis de implementação, juntamente com outras que não foram abordadas neste trabalho. No entanto este corte de abrangência é decerto útil ao usuário, que se veria em dificuldades para dominar um instrumento muito poderoso e não adequado ao seu contexto. Os tempos de resposta da Linguagem dependem fundamentalmente do número de ações que esta toma após o recebimento de uma pergunta, logo esta abrangência é um fator de dupla consequência, considerando-se os recursos de implementação atuais.

Sugere-se pois que um compromisso de **funcionalidade X rapidez** seja assumido por qualquer Linguagem de Consulta que seja desenvolvida. Isto pode ser obtido pela implementação parcimoniosa das características discutidas neste trabalho, o que deve ser feito evidentemente após um apurado levantamento das reais necessidades do usuário.

Neste sentido, concluímos pela proposição de que as Linguagens de Consulta sejam inseridas nos Sistemas de Aplicação como um requisito fundamental, principalmente naqueles em que o usuário final esteja colocado hierarquicamente em posições decisivas de organização. Justifica-se esta proposição porque estes homens através de suas decisões fazem com que as empresas por eles comandadas executem suas funções econômico-sociais. Desta forma eles necessitam para suas decisões de dados já analisados e estratificados que possam ser combinados intantaneamente, quer por suas características próprias ou por funções de simulação de informação. A forma de oferecer estes recursos é a implementação de linguagens de consulta ágeis, eficientes e moldadas para suas necessidades de decisão. Esperamos que as proposições deste trabalho contribuam de alguma forma para isso.

**BIBLIOGRAFIA**

**ABR74** Data Semantics  
 ABRIAL, J.R.  
 Laboratoire D'Informatique et Medicale de Grenoble  
 France – 1974  
 Data Base Management  
 North-Holland Publishing Company

**AST75 &  
 CHA75** Implementation of a Structured English Query Language  
 ASTRAHAN, M.M. and CHAMBERLIM, D.D.  
 CACM – Volume 18 – N.10 – Outubro/75

Apresenta a versão inicial da Linguagem SEQUEL e sua implementação piloto feita nos laboratórios de pesquisa da IBM.

**BAC78** Can be programming be liberated from de Von Neumann Style.  
 A Functional Style and its Algebra of Programs.  
 BACKUS, John  
 CACM – Volume 21 – N.8 – Agosto/78

Discute as limitações das Linguagens de Programação baseadas no computador de Von Neumann e propõe um modelo de linguagem que dá ênfase a formulação algébrica do problema a ser solucionado.

Apresenta uma profunda crítica às linguagens de programação, observando que as linguagens de programação são dotadas de uma parte fixa e uma parte variável, sendo a parte fixa muito forte (potente) e a parte variável muito fraca. Desta forma, a programação nessas linguagens é fortemente influenciada pela parte fixa destas, criando-se segundo o autor dois universos para o usuário "o lado direito e o lado esquerdo do sinal de atribuição". Demonstra ainda a inadequação das linguagens altamente estruturadas (como o ALGOL) para a execução de consultas a base de dados, pois as variações nas consultas são bastante penosas.

**BOB77** GUS – A Frame-Driven Dialog System  
 BOBROW, D. Gus et all  
 Artificial Intelligence – Abril/1977 – Vol.8-2

Apresenta em detalhes o sistema GUS, exemplificando sua utilização e discutindo sua implementação.

**CAR79** Data Base Management Systems  
 CARDENAS, Alfonso F.  
 Allyn and Bacon Company  
 Boston, Massachusetts — EEUU — 1979

Define os DBMS, apresenta alguns exemplos de implementação importantes enumerando e classificando as funções principais. Caracteriza os tipos de perguntas com base nas combinações de operadores lógicos.

**CHA75** Specifying Queries as Relational Expression — SQUARE  
 CHAMBERLIM, D.D. et All  
 CACM — Volume 18 — N.11 — Novembro/75

Apresenta a Linguagem SQUARE que serviu de base para o projeto da Linguagem SEQUEL.

**CHA76** Sequel-2 — A Unified Approach to Data Definition, Manipulation and Control  
 CHAMBERLIM, D.D. et All  
 IBM Journal R&D — Volume 20 — N.6 — Novembro/76

Apresenta uma nova versão da Linguagem SEQUEL, enfatizando as funções de controle, segurança e integridade das informações.

**DAT77** An Introduction to Data Base Systems — 2ª Edição  
 DATE, C. J.  
 Addison-Wesley Publishing Company, Inc.  
 Reading, Massachusetts — EEUU — 1977

Apresenta com detalhes o significado e a arquitetura de um sistema de gerência de bases de dados. Discute as estruturas destes sistemas sob o ponto de vista das abordagens atuais, dando destaque para os sistemas relacionais. Aborda com profundidade o estudo de um DBMS típico — o IMS — dando a conhecer sua filosofia e organização. Apresenta ainda com razoável grau de detalhes os principais aspectos de uma linguagem de consulta, através da análise de linguagens implementadas experimentalmente (a época do livro). Aborda também os temas de segurança e privacidade de informações, sem no entanto descer a minúcias destes assuntos, limitando-se a fornecer extensa lista bibliográfica devidamente comentada.

**ENG70** A Tutorial Data Base Organization  
 ENGLER, Robert W.  
 Technical Report Number-00-2004-MARCH-1970  
 IBM Corporation

**GH077** Data Base Organization for Data Management  
 GLOSH, Sakti P.  
 Academic Press, Inc.  
 New York – NY – EEUU – 1977

Analisa as linguagens de consulta sob o ponto de vista de sua implementação baseada nas estruturas conhecidas de bases de dados atuais, quais sejam –

Estruturas Hierárquicas  
 Estruturas em Redes  
 Estruturas Relacionais

**GOL75** Conceptual Generation  
 GOLDMAN, N.M.  
 Conceptual Information Processing  
 SCHANK, R.C. (Editor)  
 Amsterdam – North Holland – 1975

**HOF77** Modern Methods for Computer Security and Privacy  
 HOFFMAN, Lance J.  
 Prentice Hall, Inc  
 Englewood Cliffs, New Jersey – EEUU – 1977

Define as informações que necessitam de proteção e os princípios de segurança de dados e privacidade das informações. Apresenta em detalhes os diversos estágios do processo de validação de usuários e os métodos adequados para sua implementação discutindo pormenores e casos de exceção. Apresenta também o processo de transformação de dados como forma de manter a privacidade, dando a conhecer alguns dos principais métodos utilizados.

Discute com propriedade o problema de segurança aplicado aos programas e delinea o processo de proteção. Apresenta ainda considerações sobre proteção através de Hardware e os métodos de proteção de informações estatísticas em bancos de dados providos de linguagem de consulta.

**KNU75** The Art of Computer Programing – 2A. Edição  
 Vol-1 – Fundamental Algoritms  
 KNUTH, Donald E.  
 Addison-Wesley Publishing Company, INC  
 Reading, Massashustes – EEUU – 1975

Define com precisão os conceitos básicos de todas as teorias matemáticas que utiliza nos algoritmos computacionais apresentados ao longo de sua obra, bem como uma linguagem especial denominada MIX, que utiliza para implementação dos algoritmos.

A parte fundamental para a tese, e no entanto o estudo das diversas estruturas de dados utilizadas nos sistemas de informações, que são apresentadas em seus mínimos detalhes teóricos e práticos. Fornece ainda exercícios de diversas gradações de dificuldade e suas respectivas soluções.

**KRI77** The Future MIS  
 KRIEBEL, Charles H.  
 Interactive Decision Oriented Data Bases Systems  
 William C. House, Editor (PP.25-31)  
 Mason/Charters Publishers, Inc  
 New York – NY – EEUU – 1977

Apresenta os aspectos mais relevantes dos sistemas de informações gerenciais, discutindo os vários estágios de maturação atingidos pelos sistemas e preconiza a automação das fases de uma tomada de decisão –

Observação, Inferência, Avaliação e a própria Decisão. A partir das considerações sobre o futuro dos sistemas de informações gerenciais, arrisca algumas previsões, tais como a comercialização de decisões, tomadas com base em sistemas padronizados.

**LEF69** File Structures for On-line Systems  
 LEFKOVITZ, David.  
 Haiden Book Company, Inc  
 Rochelle Park, New Jersey – EEUU – 1969

Analisa os sistemas de informações sob o ponto de vista de suas estruturas de arquivos. Apresenta uma detalhada classificação destes sistemas e discute seus aspectos de implementação. Ressalta no entanto a implementação dos sistemas "On-Line", enumerando as estruturas de arquivos adequadas em cada módulo destes sistemas. Faz ainda uma descrição superficial dos aspectos de "Hardware"

adequados ou desejáveis em cada caso. Os aspectos de organização de arquivos são descritos em detalhes, discutindo-se em cada método, seus méritos e seus problemas, apresentando as soluções possíveis. Aborda superficialmente o problema das linguagens de consulta, enfatizando aquelas utilizadas a partir de programas hospedeiros.

LEH78

The Process of Question Answering  
LEHNERT, Wendy G.  
Lawrence Erlbaum Associates Publishers  
Hisdale, New Jersey — 1978

Estuda o processo do diálogo interativo de perguntas e respostas com um sistema em computador. A abordagem conduz ao realce da importância na modelação do processo humano de perguntas e respostas no computador, ou seja a adoção das linguagens naturais (no caso o ingles).

Estuda com razoável grau de detalhes alguns pontos do complexo mecanismo de diálogo desenvolvido pelo homem. Discute com detalhes todo o processo do diálogo, buscando sua sistematização e apresenta conclusões sobre os problemas, com base em exemplos práticos e referências a implementações existentes.

MAR73

Design of Man-Computer Dialogues  
MARTIN, James  
Prentice-Hall, Inc  
Englewood Clifs, New Jersey — EEUU — 1973

Apresenta uma visão ampla sobre o processo de diálogo entre o homem e o computador, discutindo o funcionamento das diversas técnicas utilizadas, baseando o estudo na análise de exemplos reais ricamente ilustrados, dos quais procura absorver as técnicas utilizadas. Introduz uma sistematização dos processo de diálogos existentes, classificando-os de acordo com a sua aplicabilidade. Analisa ainda o papel do "Operador" ou seja do interlocutor humano, classificando-os em grupos distintos, em função de suas aptidões e nível de interesse pelo trabalho.

MIN75

A Framework for Representation Knowledge  
MINSKY, M.  
The Psycology of Computer Vision  
P. H. Winston (Editor)  
McGraw-Hill — New York — 1975

**MIT75 &  
BOR75**

Personalized Data Bases Systems  
MITTMAN, Benjamin and BORMAN, Lorraine.  
John Wiley & Sons, Inc  
Los Angeles, California — EEUU — 1975

Discute as facilidades de uma linguagem de consulta, ressaltando a implementação feita no Software denominado RIQS. Apresenta as aplicações do RIQS em 3 áreas:

- Ciências Sociais
- Ciências Administrativas e Gerenciais
- Ciências Médicas e Cardiológicas

Sobre cada aplicação, apresenta informações sobre a base de dados usada e da forma de acesso implementada, assim como estatísticas de utilização e normas de segurança adotadas.

Faz ainda em alguns casos, a análise dos benefícios proporcionados pela adoção dos sistemas construídos com base no RIQS.

**MRI75**

System 2000  
MRI Corporation  
User Information Manual — Number 76074000  
Control Data Corporation — Minneapolis — EEUU — 1975

**OLL78**

The Codasyl Approach to Data Base Management  
OLLE, T. WILLIAM  
John Wiley & Sons  
New York — NY — EEUU — 1978

Apresenta com detalhes a abordagem CODASYL, dando ainda uma visão crítica dos seus conceitos. Apresenta também alguns sistemas que atualmente aplicam estes conceitos e que estão disponíveis no mercado, quais sejam:

ADABAS — IMS — TOTAL

Na análise de cada sistema são fornecidas detalhadas explicações sobre sua filosofia e funcionamento.

**PAL75**

Data Base Systems — A Practical Reference  
PALMER, Ian R.  
QED Information Sciences — 1975

- RIE75** Conceptual Analysis  
RIESBECK, C.  
Conceptual Information Processing  
SCHANK, R.C. (Editor)  
Conceptual Information Processing  
Amsterdam – North Holland – 1975
- SCR75** Answering Question About Processes  
SCRAGG, G.W.  
Exploration in Cognition  
Freeman  
San Francisco, California – EEUU – 1975
- SPR77 &  
WAT77** Management Information Systems – Basic Concepts  
SPRAGUE, Ralph H. Jr. and WATSON, Hugh J.  
William C. House, Editor  
Mason/Charters Publishers, Inc  
New York – NY – EEUU – 1977

Define os conceitos básicos de um sistema de informações gerenciais (MIS), utilizado como base para decisões e apresenta três modelos de sistemas, denominados – Estratégicos, Táticos e Operacionais, os quais, deverão atender usuários com responsabilidades de planejamento estratégico, controle gerencial e controle operacional, respectivamente. Enumera as principais características dos modelos e discute o seu inter-relacionamento.

- WAL78** An English Language Question Answering System for a Large Relational Data Base  
WALTZ, David L. University of Illinois at Urbana-Champaign  
CACM – Vol. 21 – Number 7 – July, 1978 – PP.526-539

Apresenta o Sistema PLANES que foi projetado para fornecer informações de uma Base de Dados relacional através de uma linguagem de consulta próxima da linguagem natural (no caso o inglês). Define os objetivos básicos de uma linguagem de consulta, discute suas razões em função da utilização e apresenta aspectos de implementação do PLANES. Enumera os pontos comuns entre o Sistema PLANES e outras pesquisas, fornecendo também extensa bibliografia sobre o assunto.



W0072 The Lunar Sciences Natural Language Interface System – Final Report  
WOODS, W.A., Kaplan, R.M. & Nash-Weber, B.  
BBN Report N.2378  
Bolt Neranek and Newman, Inc  
Cambridge, Mass – June/1972

WIN75 Frame Representation and The Declarative-Procedural  
Controversy  
WINOGRAD, T.  
Representation and Understanding  
D.G. Bobrow & A. Collins (Editor)  
Academic Press – 1975  
New York – New York – EEUU – 1975

YOU72 Desing of On-line Computer Systems  
YOURDON, Edward  
Prentice Hall, Inc  
Englewood Cliffs, New Jersey – EEUU – 1972

Apresenta e conceitua os sistemas "On Line", diferenciando-os dos demais sistemas. Discute aspectos de "Hardware" desejáveis e as características dos sistemas propostos. Discute com profundidade todas as condições básicas dos sistemas "On-Line", seus objetivos aplicabilidade e dificuldades de Implementação. Apresenta também métodos de análise de performance e avaliação de resultados destes sistemas, bem como exemplos de aplicação das técnicas discutidas. Finalmente são apresentados e discutidos os sistemas operacionais e o "Ambiente Operacional" desejável para o funcionamento dos sistemas "On-Line".

YOU79 &  
CON79 Structured Design  
YOURDON, Edward and CONSTANTINE, Larry L.  
Prentice-Hall, Inc  
Englewood Clifs, New Jersey – EEUU – 1979

Apresenta e conceitua a técnica de projeto estruturado, analisando as estruturas de procedimentos e programas. Justifica a técnica com base na funcionalidade de um sistema ou programas. Apresenta com detalhes a técnica e discute a sua aplicação em sistemas reais.

ZL075

QUERY BY EXAMPLE: The Invocation and Definition of Tables and Forms.

M.M.ZLOOF

Proc. Internacional Conference on Very Large Data Bases

ACM — September — 1975