Anales PANEL'81/12 JAIIO
Sociedad Argentina de informática
e Investigación Operativa. Buenos Aires, 1981

SELEÇÃO DE OBJETOS NA DESCRIÇÃO DE SISTEMAS DE INFORMAÇÃO

Maximilian Emil Hehl

João Sahão Junior

Margareth Lauand

Instituto de Ciências Matemáticas de Sâo Carlos - USP
Departamento de Ciências de Computaçâo e Estatística
Av. Dr. Carlos Botelho 1465, Sao Carlos, Brasil

RESUMO

O problema da seleção de objetos, quando se faz a descrição de sistemas de informação com o apoio de computador, é analisado e é apresentado um conjunto de rotinas que constitue um subsistema de programação destinado a permitir que essa seleção seja feita com o auxílio de computador. A seleção é executada através da manipulação de conjuntos com as operações usuais da teoria dos conjuntos. Uma linguagem para a especificação de seleções é apresentada através de descrição formal. A estrutura de dados usada no processo de seleção é apresentada juntamente com a adotada para armazenamento da descrição do sistema de informação. O subsistema de programação desenvolvido para executar a tarefa de seleção é descrito sucintamente.

I. INTRODUÇÃO

A descrição de um sistema consiste basicamente de objetos que representam as entidades que o compõem e relações que indicam a forma pela qual essas entidades interagem entre si.

Quando um sistema atinge um porte muito grande, envolvendo para o seu trabalho de projeto e desenvolvimento uma equipe de pessoas, é conveniente a utilização de um computador para auxiliar a interação de todas essas pessoas. Isso é feito mantendo no computador uma descrição atualizada do sistema, de for

ma a possibilitar que cada um possa ter com facilidade uma visão dos aspectos de seu interesse.

A vantagem da utilização de um computador reside em sua capacidade de poder selecionar na descrição os objetos envolvidos no aspecto requisitado pelo usuário, de forma a lhe possibilitar a obtenção de relatórios significativos.

Essa vantagem deve ser aproveitada de modo a permitir selecionar os objetos necessários para qualquer informação que o usuário necessitar da descrição, de uma forma eficiente e de fácil especificação, reunindo ao mesmo tempo flexibilidade de uso e capacidade de realizar seleções de grande complexidade.

II. ESTRUTURAÇÃO DOS DADOS RELATIVOS À DESCRIÇÃO DE UM SISTEMA

Após o estudo e análise comparativa de várias metodologias para projeto, desenvolvimento, descrição e análise de sistemas de informação, vem sendo elaborado e desenvolvido, no Laboratório de Computação do Departamento de Ciências de Computação e Estatística do Instituto de Ciências Matemáticas de São Carlos - USP, um sistema apoiado por computador que abrange as melhores características das metodologias analisadas.

Este sistema consiste de uma linguagem relacional e de um pacote de programas integrados. À medida que um sistema vai sendo descrito através desta linguagem, é criada na memória do computador uma estrutura de dados que representa a descrição realizada. Tal estrutura pode ser vista como uma representação do modelo matemático de hiperdígrafos categorizados com rótulos, como foi demonstrado em [Ge76] e [Ni77].

A literatura traz numerosas formas de representação de dígrafos categorizados com rótulos. Foi realizado um estudo dessas representações [Tr78], com o objetivo de adequá-las às características específicas dos hiperdígrafos que caracterizam estruturas de dados representativas de descrições de sistemas. A forma de representação que se concluiu ser a mais adequada é a mostrada na figura 1.

A cada objeto da descrição é associado um nome, através do qual a ele se tem acesso, e vários atributos, como por exemplo sinônimos, proteções, datas, etc... Além disso cada objeto está envolvido em relações com outros objetos. Essas relações são armazenadas em estrutura de lista. A cada objeto está associada uma lista que contém todas as relações que partem desse objeto.

As relações entre objetos podem ser classificadas em dois tipos: simples ou múltiplas. Simples são as relações que envolvem apenas dois objetos. Múltiplas são relações que envolvem mais de dois objetos. Cada uma delas tem seu próprio esquema de armazenamento. De uma forma geral, cada célula de uma lista tem as seguintes informações: objeto-destino dessa relação; código da relação; ponteiro para a próxima célula da lista; e ponteiro de multiplicidade da relação (caso a relação seja simples, esse ponteiro é nulo).

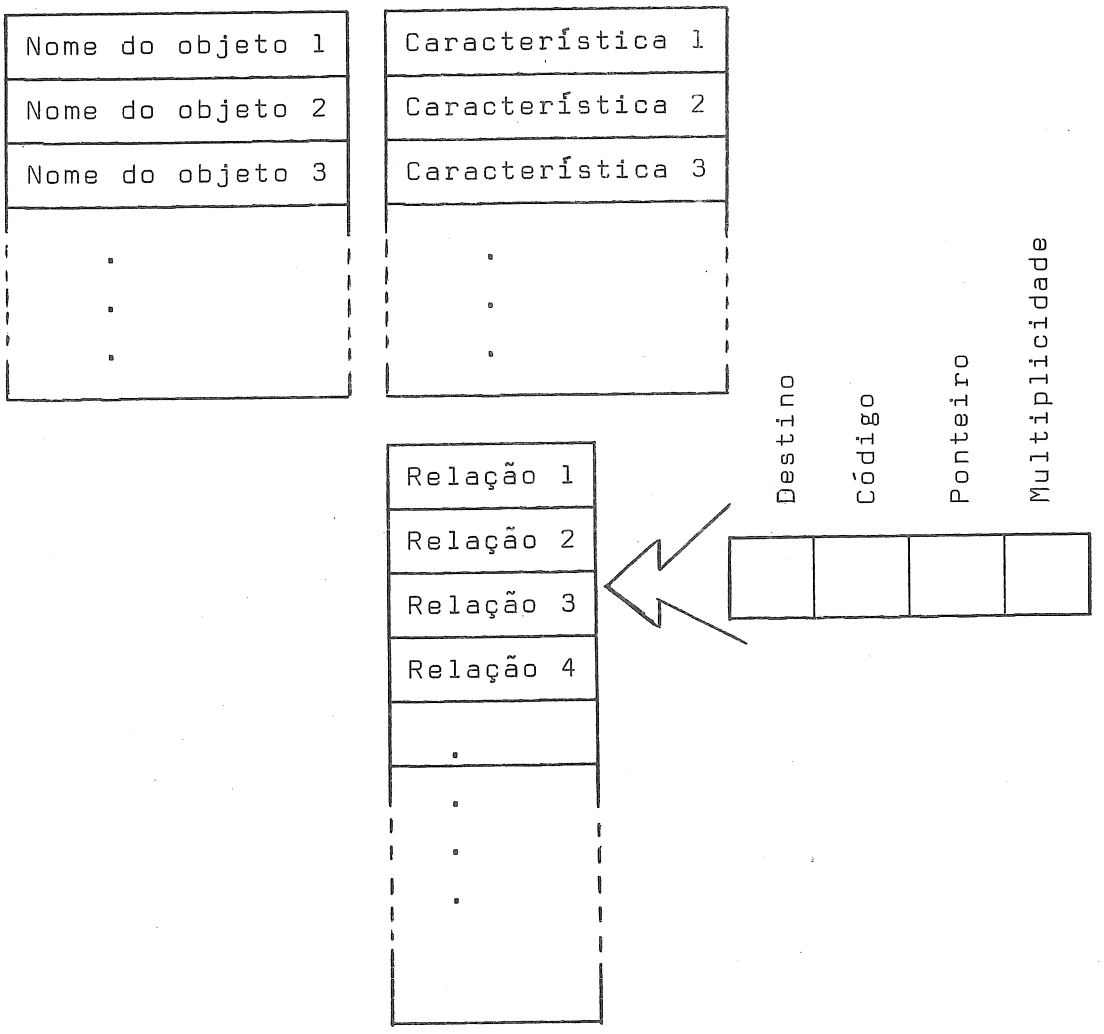

Figura 1

III. O PROBLEMA DA SELEÇÃO

Quando um sistema está em projeto ou em desenvolvimento, e um computador está sendo utilizado para auxiliar essa tarefa, cada membro que dela está participando vai fornecendo informações sobre os aspectos com os quais está trabalhando. Vai sendo elaborada assim a descrição do sistema, de uma forma bastante flexível.

A qualquer instante, durante a elaboração da descrição, um usuário pode querer obter informações do que já foi descrito sobre determinado aspecto. Essas informações lhe são fornecidas através de relatórios que contêm os objetos envolvidos nesse aspecto. Os programas que geram esses relatórios devem ser capazes de organizar a apresentação desses objetos em formato significativo para o usuário. A seleção adequada desses objetos é essencial para a confecção desses relatórios.

A seleção de objetos é feita através da escolha de de

terminadas características envolvidas no aspecto desejado. Cada característica dá origem a um conjunto de objetos. Pode-se, então, pensar em uma seleção como sendo o resultado de operações entre conjuntos de objetos previamente selecionados de acordo com características específicas. Isso faz com que se possa produzir qualquer seleção desejada, não importando a sua complexidade.

Exemplos de características que podem ser usadas para selecionar conjuntos de objetos são: tipo de objeto; data de criação e/ou atualização de objetos; existencia ou não de sinônimos; etc...

As operações entre os conjuntos de objetos são as usuais da Teoria dos Conjuntos, tais como união (+), intersecção (*), diferença (-) e complementação (').

Pode-se então partir de seleções simples e de propósito geral, como aquelas exemplificadas, e através dessas operações chegar a uma seleção de objetos de interêsse específico, associá-la a outras seleções já atualizadas e assim ir construindo qualquer seleção pretendida. Uma vantagem disso é que, ao mesmo tempo em que se utiliza uma teoria matemática bastante difundida, que permite grande especificidade nos critérios de seleção e facilidade de se criar os programas que analisam tais critérios, mantém-se contato muito estreito com a forma que normalmente uma pessoa raciocina quando procura selecionar alguma coisa mentalmente.

Uma característica que frequentemente é usada para selecionar um conjunto de objetos é a forma pela qual esses objetos se relacionam com outros na descrição. Durante a fase de elaboração de uma descrição, usam-se relações de vários tipos para indicar a forma pela qual os objetos interagem. Pode-se então usar esses tipos de relação para selecionar os objetos que se relacionam de determinada forma com os objetos de um certo conjunto.

A linguagem formalmente descrita na figura 2 aplica esses conceitos e sua utilização possibilita o uso do computador para executar o processo de seleção.

Um exemplo de seleção especificada nessa linguagem é:

SEL = ARQUIVO * USADO POR (GERA(SAÍDA)).

Neste exemplo, está especificado que devem ser selecionados todos os objetos do tipo ARQUIVO que sejam usados por qualquer objeto que gere um objeto do tipo SAÍDA.

Para chegar à seleção assim especificada, o sistema em desenvolvimento mencionado na seção II parte da criação de dois conjuntos, os quais chamaremos de C1 e C2 contendo, respectivamente, todos os objetos de tipo ARQUIVO e todos os objetos de tipo SAÍDA da descrição. A seguir é criado um novo conjunto C3 que consiste de todos os objetos da descrição que estão relacionados com algum objeto de C2 através da relação GERA. Então é criado o conjunto C4 formado por todos os objetos que estão relacionados com algum objeto de C3 através da relação USADO POR. Fi

nalmente, chega-se ao conjunto formado por todos os objetos que pertencem tanto a C1 quanto a C4, o qual constitue a seleção especificada. A esse conjunto é atribuído o nome SEL, o qual serve para quaisquer referências posteriores.

```
<COMANDO> ::= <NOME DO CONJUNTO>=<EXPRESSÃO PARA SELEÇÃO>
<EXPRESSÃO PARA SELEÇÃO> ::= <CRITÉRIO DE SELEÇÃO>
                           | <CRITÉRIO DE SELEÇÃO><OPERADOR>
                                <EXPRESSÃO PARA SELEÇÃO>
                           | <EXPRESSÃO PARA SELEÇÃO><OPERADOR>
                                <CRITÉRIO DE SELEÇÃO>
                           | <RELAÇÃO>(<EXPRESSÃO PARA SELEÇÃO>)
                           | <EXPRESSÃO PARA SELEÇÃO>'

<CRITÉRIO DE SELEÇÃO> ::= <TIPO DE OBJETO>
                        | COMENTÁRIOS
                        | SINÔNIMOS
                        | ISOLADO
                        | TODOS
                        | OBJETO=<NOME DO OBJETO>

<OPERADOR> ::= + | - | *
```

Figura 2

IV. ESTRUTURA DE DADOS USADA NO PROCESSO DE SELEÇÃO

A estrutura de dados usada no processo de seleção é esquematizada de forma a permitir uma rápida e eficiente execução das operações necessárias ao processo de seleção. É a estrutura interna principal do programa de seleção de objetos que integra o sistema acima mencionado. Essa estrutura armazena os conjuntos selecionados pelo usuário. Tais conjuntos são armazenados num formato adequado às operações necessárias neste programa. Em se tratando de descrições de sistemas de informação, chegou-se à conclusão de que, para esse arquivo, vinte é um limitante suficiente para o número de conjuntos de seleção que esse arquivo pode armazenar. A figura 3 mostra essa estrutura para facilitar o seu entendimento.

A estrutura geral, que armazena os objetos da descrição e suas características, é mostrada resumidamente à esquerda da figura. Ao lado, estão posicionadas vinte colunas, sendo que cada coluna representa um conjunto selecionado, indicando para cada objeto se ele pertence ou não àquele conjunto. Esses conjuntos selecionados são armazenados num arquivo chamado ARQUIVO DE TRABALHO, onde o primeiro registro desse arquivo é

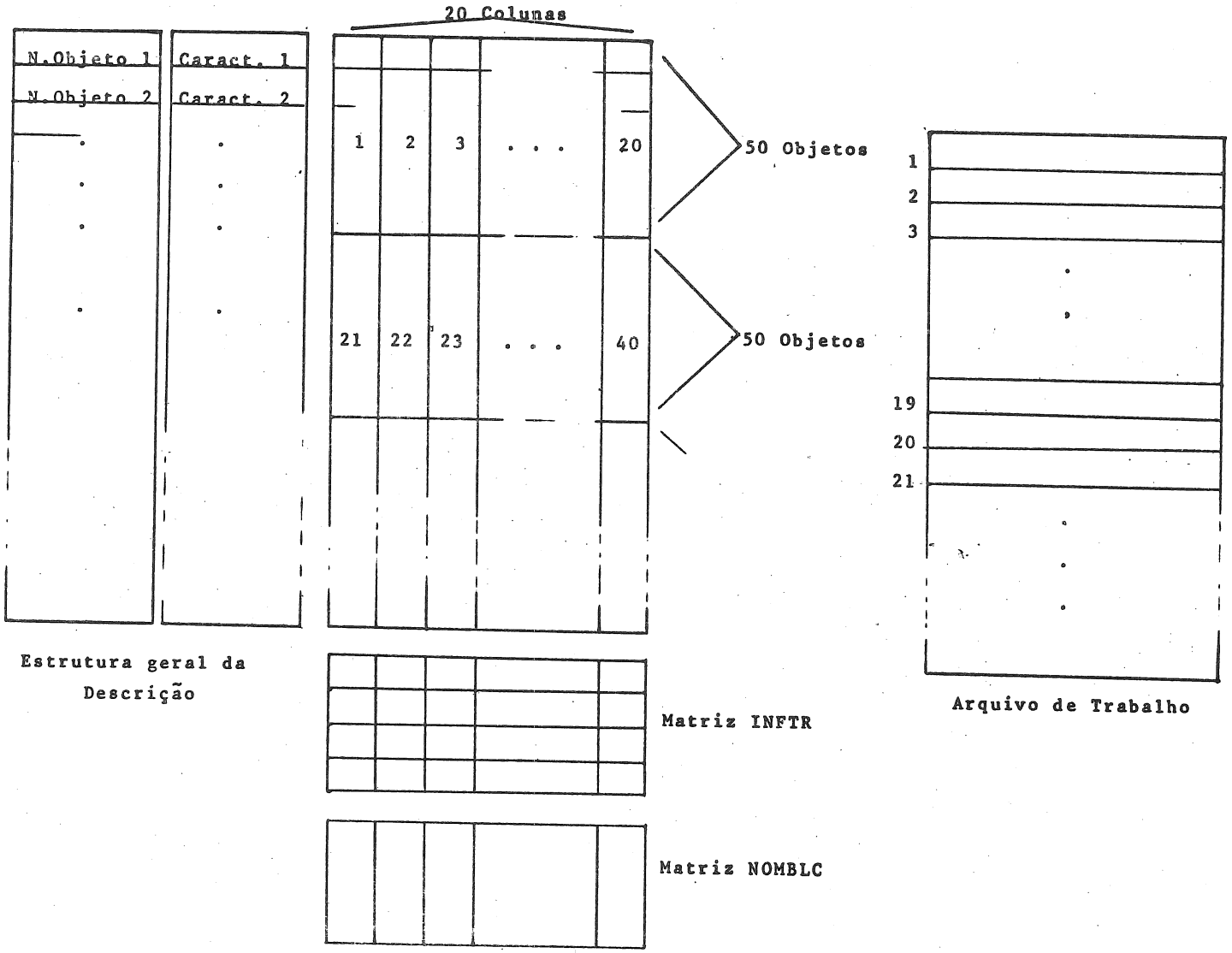
20 Colunas
N.Objeto 1
N.Objeto 2
Caract. 1
Caract. 2
1 2 3 . . . 20
21 22 23 . . . 40
50 Objetos
50 Objetos
1
2
3
19
20
21
Estrutura geral da Descrição
Matriz INFTR
Matriz NOMBLC
Arquivo de Trabalho

Figura 3

constituído pelas cinquenta primeiras posições da primeira coluna; o segundo registro, pelas cinquenta primeiras posições da segunda coluna; e assim por diante, até o vigésimo registro, que armazena as cinquenta primeiras posições da vigésima coluna. As cinquenta posições seguintes da primeira coluna são armazenadas no vigésimo-primeiro registro, e assim por diante, todas as colunas são armazenadas em blocos de cinquenta posições.

Quando se deseja manipular o arquivo de trabalho, as informações são trazidas para a memória, um bloco de cada coluna de cada vez. O bloco de uma dada coluna que é trazido não é necessariamente o mesmo de qualquer outra coluna. Esses blocos são trazidos para uma coluna de uma matriz de 50x20 chamada ARQTR.

Pode-se observar, então, que os blocos de cinquenta posições dos conjuntos selecionados constituem a matriz ARQTR e ainda, que cada bloco é tratado de cada vez, independentemente.

Para trabalhar com essa estrutura, é necessário ainda uma matriz de 20x4 chamada INFTR que contém as seguintes informações para cada conjunto (coluna):

- qual o bloco que está na memória;
- até que palavra do bloco está sendo usada (para o caso em que o último bloco não possua cinquenta posições);
- a que objeto se refere a primeira posição desse bloco;
- estado da coluna:
  - zero = a coluna não está sendo usada,
  - um = a coluna está sendo usada e o arquivo de trabalho está atualizado,
  - dois = a coluna está sendo usada e o arquivo de trabalho não está atualizado.

Existe ainda uma matriz de 10x20 chamada NOMBLC, a qual armazena os nomes dos conjuntos selecionados, atribuídos pelo usuário.

Para cada objeto, indica-se que ele pertence ao conjunto de seleção colocando o valor um na posição correspondente a esse objeto. Para indicar que ele não pertence ao conjunto de seleção, coloca-se zero nessa posição.

V. CONCLUSÕES

O subsistema de seleção de objetos, que faz parte do sistema apoiado por computador mencionado em II, foi desenvolvido em linguagem FORTRAN para um computador PDP 11/45 com preocupação de ser portátil [Br77], de modo a poder ser instalado com poucas adaptações em outros computadores que possuam compilador para a linguagem FORTRAN IV ANSI. Esse subsistema é constituí

do por 42 rotinas específicas, cada uma, em média, com uma página e meia de listagem, além de outras rotinas de uso comum com os outros subsistemas.

No desenvolvimento desse subsistema foram usados conceitos de seleção usuais à Recuperação da Informação como os encontrados em [Sa68] bem como Técnicas de Alocação Dinâmica de Memória conhecidas na literatura técnica, semelhantes às mencionadas em [Ts74]. Por outro lado, o problema de seleção de objetos em sistemas de informação já foi resolvido por outros, como por exemplo Teichroew [Ka79] e Solvberg [Aa74]. O enfoque adotado neste trabalho é de maior generalização, de vez que, em nenhum ponto dos programas, são fixados os objetos a serem selecionados. Assim, apenas através de dados são feitas as particularizações necessárias para que o subsistema funcione para objetos e relações específicos. Essa idéia de generalidade já foi adotada por Teichroew [Te75] mas sua aplicação com as considerações feitas neste trabalho permite que o subsistema de seleção, juntamente com os demais subsistemas, possa ser processado em um mini-computador, como é o caso do PDP 11/45, com limitados recursos de memória principal, em que a máxima partição para usuário é de 64 K-bytes.

Este trabalho foi em parte desenvolvido dentro de um plano de trabalho de bolsa de iniciação científica concedida à estudante de engenharia Margareth Lauand, sob a orientação do professor João Sahão Junior. O sistema mencionado em II, do qual o subsistema a que se refere este trabalho faz parte, está sendo desenvolvido com apoio financeiro da FAPESP e FINEP sob a coordenação do engenheiro Caetano Traina Junior.

BIBLIOGRAFIA

[Aa 74] AANSTAD, P. - Introduction to the Program System CASCADE/2. University of Trondheim, Working Paper nº 38. 1974.

[Br 77] BROWN, P.J. and others - Software Portability, an Advanced Course. Cambridge University Press, 1977.

[Ge 76] GERMANO, F. - Modelos Matemáticos da Teoria dos Sistemas de Informação. Tese apresentada à Universidade de São Paulo como parte dos requisitos para a obtenção do título de Doutor em Ciências. São Carlos, Abril de 1976.

[Ka 79] KANG, K.C. & MAZZILLI, E. - A Query facility for PSA version A5.1. - User's Manual. ISDOS Project.April 1979.

[Ni 77] NICOLETTI, M. - Sobre Modelos Matemáticos do Projeto Lógico de Sistemas de Informação. Dissertação apresentada à Universidade de São Paulo como parte dos requisitos para a obtenção do título de Mestre em Ciências de Computação e Estatística. São Carlos, janeiro de 1977.

[Sa 68] SALTON, G. - Automatic Information Organization and Retrieval. McGraw-Hill, 1968.

[Te 75] TEICHROEW, D., GERMANO, F. & MORAES, P. - Computer-aided Structured Documentation of Information Processing System Requirements. 8º Congresso Nacional de Processamento de Dados. SUCESU - São Paulo - 27 a 31 de outubro - 1975.

[Tr 78] TRAINA JR, C., GERMANO, F. & SAHÃO JR, J. - Uso de Dígrafos para Esquematização de uma Base de Dados. 2º Congresso Brasileiro de Automática. Florianópolis, 4 a 6 de setembro de 1978.

[Ts 74] TSICHRITZIS, D.C. & BERNSTEIN, P.A. - Operating Systems. Academic Press, 1974.