

utensílio documentário que o põe atualmente ao serviço da comunidade científica. Que nossos colegas, responsáveis por instituições, escavações, sítios e museus que têm desejado nos acolher para cumprir nossas missões encontrem na realização de nossa empresa o testemunho de nosso reconhecimento e da importância que nós damos ao tratamento e à difusão documentária nas ciências da Antiguidade. Nosso desejo é que ao nos consultar e ao prosseguir em sua colaboração, eles estejam entre os primeiros a ter acesso aos sistemas de informação que a tecnologia contemporânea pôs à nossa disposição, e que eles possam no futuro nos ajudar a superar novas etapas em nossa pesquisa sobre as técnicas da documentação.

O LOGICIAL SIGMI/SIGMINI

CHARLINE BRISBOIS
Union Minière
Bruxelas

PATRICK MORDINI
Centre d'Automatique et
d'Informatique
École Nationale Supérieure des
Mines - Paris

Introdução

O logicial SIGMI/SIGMINI foi concebido para gerar informações muito heterogêneas tendo entre elas relações complexas. Ele se inspira por sua vez nos sistemas de gestão de bases de dados (SGBD) e sistemas documentários (SD). Uma das idéias originais foi a de não declarar previamente nem os dados nem as suas relações.

A gestão de dados sob o SIGMINI consiste em uma manipulação de pares (nome do dado ou elemento, valor do dado) estruturados hierarquicamente por um jogo de parênteses. Esta relação permite ao usuário estocar todos

os dados e as relações desejadas no momento de sua preparação, sem precisar se referir a um sistema prévio.

A interrogação de um banco cujos dados são estocados seguindo um tal modelo deve compreender uma interrogação explícita da estrutura dos dados mais do que operadores de comparação habituais.

Segue-se a uma interrogação do tipo SD sem estrutura, através de um fichário invertido, uma interrogação de dados pré-selecionados do tipo SGBD com estruturas, e isto devido a uma linguagem muito simples.

A primeira versão do sistema, SIGMI, funciona pelo tratamento de lotes sobre a série IBM 370. A segunda versão do sistema, SIGMINI, é interativa e destinada mais particularmente aos minicomputadores que disponham da rede do sistema UNIX. Ele é fruto da colaboração entre o Centre de Recherche de l'École des Mines de Paris e a Union Minière.

A concepção do SIGMINI foi particularmente estudada sob o ângulo de sua portabilidade. Os programas são escritos em FORTRAN IV. Com exceção das rotinas de Entrada/Saída, o logicial se adapta facilmente a outros sistemas.

Os exemplos utilizados a seguir são tomados de um dos bancos constituídos pelo Centre de Recherche sur les Traitements Automatisés en Archéologie Classique (CNRS - Université de Paris X) e foram fornecidos por A.-M. Guimier-Sorbets.

I - Entrada de Dados

1. O dado elementar

Cada dado elementar é introduzido graças a um par (elemento = valor). O elemento dá uma informação genérica: local de descoberta, tipo de cena, número de partes, datação, bibliografia.

O valor oferece uma informação específica para um elemento dado.

Existem muitos tipos de elementos:

tipo semântico – por ex. Local de descoberta = Kourion / Negrópolis /

tipo padrão – por ex. Tipo de cena = Cena marinha /

tipo numérico – por ex. Número de partes = 1 /

tipo numérico com dois termos – por ex. Datação = -250 A -185 /

tipo comentário – por ex. Bibliografia = RUPP D.W., 1978 /

Os tipos padrão e semântico são apresentados no capítulo dicionário.

O valor numérico pode ser introduzido seja sob a forma de número inteiro, seja sob a forma de número decimal. O valor numérico com dois termos é um par de valores numéricos que limita a região na qual se encontra o valor. Se a extensão desta região é nula, o tipo numérico com dois termos é interpretado como um tipo numérico simples.

O valor do tipo comentário é uma cadeia de caracteres quaisquer, de extensão inferior a 32.000 caracteres.

2. As relações entre os dados

Cada objeto homogêneo de informações (ou ficha) pode ser estruturado hierarquicamente graças à introdução de parênteses.

Exemplo: a análise abaixo descreve um mosaico encontrado em Delos (para o desenho deste mosaico e seu esquema de análise, ver acima o artigo de A.-M. Guimier-Sorbets). O exemplo permite ver que os dados e a estrutura são fornecidos simultaneamente.

Cada elemento pode se apresentar não importa onde e com as repetições necessárias.

A apresentação abaixo da análise de um mosaico em formato SIGMI/SIGMINI não corresponde exatamente ao formato de entrada, pois o programa

de edição dos documentos simplifica a apresentação ao suprimir os parênteses de fecho; no mais, as diferentes linhas são automaticamente deslocadas de um espaço, para a direita ou para a esquerda, segundo sua correspondência a um nível superior ou inferior do esquema arborescente, o que as torna mais legíveis e fáceis de controlar.

MOSAICO = DELOS

Número = 16

Comentário = mosaico de Orbius

Local de descoberta = Delos/Ágora/dos Italianos/nicho/10

Local de conservação = *in situ*/Delos

Comentário = restaurações modernas/incerto

Datação = -110 A -90

Dimensão 1 = 420

Dimensão 2 = 226

Orientação = para a porta

Número de cores = 99

Número de tapetes = 2

Agenciamento = soleira/independente

(**Parte = registro/meio**

Técnica = opus tessellatum/meio

Decoração = monocroma

Cor = branco

(**Parte = tapete**

Forma = retângulo

Dimensão 1 = 322

Dimensão 2 = 158

Técnica = opus tessellatum/meio/fim/lâminas de chumbo

(**Parte = borda**

Número de partes = 7

Número de bandas decoradas = 2

(**Parte = banda**

Número = 4

Decoração = monocroma

Cor = branco

(**Parte = filete/triplo**

Decoração = monocroma

Cor = branco

(**Parte = banda**

Posição = exterior

Decoração = composição/linear/manchas/esquerda

Cor = preto
 Cor fundo = branco
 (Parte = banda
 Posição = interior
 Técnica = pasta de vidro
 Decoração = composição/
 linear/guilochê/L/direita/
 perspectiva
 Cor = branco/bege/vermelho/
 azul/verde

Tratamento = degradê

Cor fundo = preto

(Parte = fundo
 Técnica = pasta de vidro
 Decoração = composição/de
 superfície/meandro/de suásticas/
 2/quadrado/direito/perspectiva
 Cor = branco/vermelho/verde/
 azul/nuance

Cor fundo = preto

(Decoração = quadrado/
 bipartido/em diagonal

(Parte = tapete/soleira/sem borda

Forma = retângulo

Dimensão 1 = 224

Dimensão 2 = 7

Técnica = opus tessellatum/fim/
 meio

(Parte = fundo

Decoração = inscrição/
 dedicatória

Comentário = I. ORBIUS

M.F.HOR*ATIA*

Cor = preto

Cor fundo = vermelho

Bibliografia = EAD, XXIX, p. 24-26,
 40-41, 43, 48, 53-55, 69, 71, 96, 104, 115,
 133-135, fig. 19 planta, fig. 21-25/EAD,
 XIX, P. 57, 59, 97-98, FIG. 2, 6 plantas,
 fig. 49, pr. XVII.2/FOUGERES G.,
 1887/BULARD M., 1908, P. 191-192, fig.
 67/BRUNEAU PH., J. DUCAT, 1983,
 p. 168, fig. 42, planta II

Endereço = 721/1710/932/940/913/718/803

3. Possibilidade de adaptar e de completar a estrutura de dados

Se a descoberta posterior de um

outro fragmento de um mosaico já analisado permite completar sua análise, podemos ser obrigados a modificar sua estrutura: na verdade, a parte nova pode não apresentar o mesmo nível técnico que as outras, e será então necessário relevar esta indicação do nível 1 e relacioná-la a cada uma das partes concernentes, nos níveis 2, 3, 4 ou 5.

Este exemplo permite constatar que a estrutura é dinâmica e pode seguir a evolução dos dados do ponto de vista tanto de suas relações como de seus valores.

II - Dicionário

O dicionário tem dois objetivos:

- o controle dos nomes dos elementos e dos valores padrões,
- a elaboração do código destes nomes de modo a estocá-los no banco.

a) Conteúdo do dicionário

Os tipos numérico e comentário são caracterizados pelo fato de que não importa qual o valor numérico ou o conjunto de caracteres que podem ser atribuídos a um elemento. Estes valores não são estocados no dicionário, apenas os elementos.

O tipo padrão restringe os elementos a uma lista de valores possíveis. Esta lista pode ser aberta e aumentar automaticamente à medida em que novos termos correspondentes apareçam durante a preparação das fichas. Ela pode também ser fechada durante a preparação de maneira a excluir os termos incorretamente escritos ou descartados pelo responsável pelo dicionário e, neste caso, a incorporação de novos termos se faz através de uma operação específica.

No dicionário se encontram duas categorias de informação, de um lado os elementos e de outro os diferentes conjuntos os valores atribuídos aos elementos de tipo padrão e semântico. Em

cada um destes conjuntos é possível vincular os termos através de relações de sinonímia. Assim um mesmo valor pode ser introduzido sob a forma completa e resumida e/ou em línguas estrangeiras. Um usuário pode empregar indiferentemente não importa qual sinônimo.

Por exemplo:

local de conservação = Museu do Louvre /

ou

conservação = Louvre /

Um dispositivo permite identificar a qual língua pertence um termo e, desde que esta língua tenha sido introduzida, uma ficha poderá então aparecer, quando da edição, em uma língua diferente daquela em que foi preparada.

Em princípio, cada valor pertence a um único elemento. Entretanto, um mecanismo dito de compatibilidade permite indicar que o conjunto de valores de um elemento é compatível com o conjunto de valores de outro elemento.

Por exemplo: local de descoberta e local de conservação são elementos compatíveis, eles têm o mesmo valor.

b) Dicionário semântico

O dicionário semântico explicita a significação e as relações recíprocas dos valores de alguns elementos. Ele é mais poderoso do que a maior parte dos tesouros hierarquizados dos sistemas documentários.

Por exemplo: local de descoberta é um elemento do tipo semântico.

Se uma ficha é introduzida com o valor "Local de descoberta = Pella", esta ficha será selecionada por sua vez para uma questão referente a Pella e para uma questão referente à Grécia. Isto é possível porque o dicionário semântico contém a informação de que Pella pertence à Grécia. Se se introduz uma ficha com o valor "Local de descoberta = Macedônia", a ficha aparecerá quando de uma questão referente à Grécia, Bulgária, Iugoslávia, Albânia ou

Turquia. Isto mostra que o dicionário semântico permite regular os problemas de relações entre valores não hierarquizados.

Para mais informações sobre a codificação semântica, ver bibliografia, 12.

III - Interrogação

A interrogação de um banco de dados gerado pelo sistema SIGMINI se faz em dois tempos. O primeiro é denominado pré-seleção. Esta se realiza através do fichário invertido. O segundo tempo é denominado seleção, quando cada uma das fichas pré-selecionadas é examinada em seu conteúdo e sua estrutura. É neste nível que intervêm os operadores que descrevem as relações que deveriam existir entre os dados.

1. Pré-Seleção

O fichário invertido permite dar, para o elemento ou o valor, os números das fichas que os contém.

Todos os elementos ou valores não são invertidos, e a escolha daqueles que o são é determinada pela sua introdução no dicionário, em função do caráter seletivo de cada um deles.

Exemplos:

- O elemento Decoração é invertido sobre cada um de seus valores. Quando da interrogação, o fichário invertido restitui o conjunto do número de fichas concernentes por cada um dos valores do elemento Decoração.
- O elemento Assinatura é invertido tal qual, sem seus valores. O conjunto das fichas que comporta o elemento Assinatura é fornecido na interrogação.

Tomemos o exemplo da questão: "Mosaico com fundo ornado de meandros em suásticas com borda de manchas, e uma inscrição".

Em linguagem de interrogação SIGMINI, ela se torna:

1. decoração = meandro
2. : JU :
3. decoração = em suásticas
4. : PE :
5. parte = borda
6. : PE :
7. decoração = manchas
8. : ET :
9. decoração = inscrição

As linhas 1, 3, 5, 7 e 9 formam o que se chama de critérios de seleção. As linhas 2, 4, 6 e 8 contêm os operadores de estrutura.

A pré-seleção consiste em interrogar o banco sem a estrutura, os operadores de estrutura das linhas 2, 4 e 6 sendo substituídos por ET. As questões podem ser postas critério por critério e se pode obter assim a cada etapa o número de fichas pré-selecionadas.

Para a questão citada acima, serão selecionadas todas as fichas que contenham todos os valores pedidos.

Esta pré-seleção reduz o banco a um conjunto de fichas que poderão convir se a estrutura for respeitada, entre as quais as fichas de Delos 16 e Delos 195.

O conjunto de fichas obtidas na pré-seleção pode parecer muito grande ou muito pequena para o usuário. Elas

podem colocar a sua questão de maneira mais ampla ou mais restrita segundo o caso. Mas se o resultado da pré-seleção lhe for conveniente, ele prossegue para a seleção.

2. Seleção

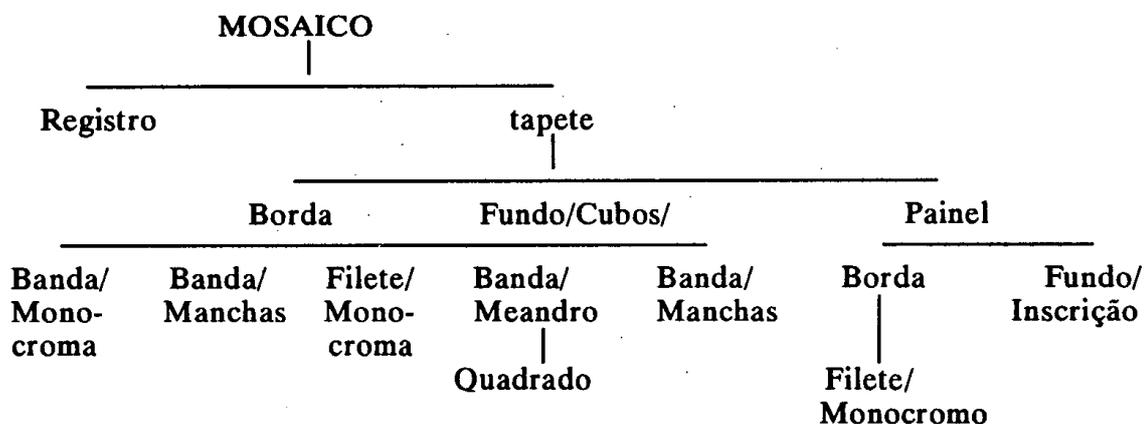
Quando da pré-seleção, a ficha de Delos 195, cujo esquema hierárquico é dado abaixo, é mantida para a questão dada em exemplo. Se examinarmos esta ficha perceberemos que ela não responde realmente à questão colocada já que o meandro e as manchas ornaram as duas bandas da mesma borda, enquanto que se desejava que o meandro ornasse um fundo (do tapete ou do painel).

Os operadores de estrutura permitem suprimir o "bruto" na seleção das fichas.

No exemplo dessa questão, apenas três operadores de estrutura são utilizados, embora existam muitos outros.

Exemplos:

JU C1:JU:C2 significa que os critérios C1 e C2 estão no mesmo nó da estrutura.



Esquema arborescente do mosaico 195 de Delos

- AS C1:AS:C3 significa que C1 e C2 têm elos hierárquicos na árvore e que C1 está antes de C2.
- FR C1:FR:C2 significa que C1 e C2 têm o mesmo pai.
- PE C1:PE:C2 significa que C1 é o pai de C2.

Todos os operadores têm o seu equivalente negativo.

Por exemplo:

- NJU C1:NJU:C2 significa que deve existir um nó que contém C1 sem a presença de C2.

Também é possível impedir que dois critérios estejam presentes no mesmo nó de uma ficha. Isto se escreve:

NÃO : (C1 : JU : C2).

A interrogação de um banco SIGMINI se faz inteiramente de forma interativa. O operador pode adaptar sua questão de acordo com o ruído ou o silêncio que obtém no nível da pré-seleção inicialmente, e depois na seleção.

Todas as questões são mantidas mesmo após a sessão do usuário, as questões permanecendo disponíveis seja para interrogar uma outra sub-base, seja para serem inseridas em outra questão.

Uma vez terminada a seleção, os números das fichas selecionadas são mantidas. Isto permite a gestão da difusão seletiva da informação, ou perfil.

ANEXO

Descrição Sucinta das Principais Funções do SIGMINI

Introdução e atualização do dicionário.

A introdução pode ser feita:

- seja em interativo, em um monitor,
- seja em tratamento por lotes, a partir de uma fita ou de um fichário disco.

Os comandos disponíveis são:

- criação de elemento ou valor

- atualização das relações de sinonímia
- listagem do dicionário
- estabelecimento da compatibilidade e dos tipos de inversão.

Introdução das fichas no banco.

A introdução se faz através seja de uma fita, seja de um fichário disco. Ao nível do sistema, as fichas erradas são rejeitadas e devem ser reintroduzidas. O usuário tem no entanto o meio de corrigir as fichas erradas por um programa de atualização sem ter que regravá-los inteiramente.

Um editor especializado permite efetuar uma preparação controlada interativa. Fichas-tipo que consistem em esqueletos de ficha são denominados em função dos domínios tratados para orientar a introdução.

Interrogação de um banco.

O usuário deve previamente designar o banco ou a parte do banco que deseja interrogar. No caso em que a parte do banco sobre a qual o usuário deseja colocar questões não esteja instalada nas unidades de disco, as questões são registradas e não são executadas senão quando as fichas a que se referem estiverem presentes.

Cada usuário dispõe de um ou mais questionários nos quais ele pode introduzir até 42 questões diferentes. No interior do questionário cada questão tem um número de referência, o que permite reutilizá-la para a redação de uma outra questão ou para um perfil.

Ao fim da introdução de uma questão três possibilidades se apresentam:

- análise sintática
- análise sintática e pré-seleção
- análise sintática, pré-seleção e seleção.

Edição de dados.

As fichas pré-selecionadas e se-

leccionadas podem ser editadas no monitor ou impressas. É possível editar apenas uma parte da ficha.

DESCRIÇÃO DA ESTAÇÃO DE CONSULTA

MICHEL LENCI
Diretor do Centre d'Automatique
et d'Informatique
École Normale Supérieure des
Mines de Paris

O aparelho de interrogação do banco de dados de imagens foi concebido no Centre d'Automatique et d'Informatique da École Normale Supérieure des Mines de Paris. Ele se compõe de um microcomputador, de um terminal e de seu monitor.

O microcomputador é ligado por meio de um modem ao computador central no qual estão estocados os bancos de dados. Seu papel é duplo: ele conserva localmente as respostas enviadas pelo computador central e controla o videoleitor, oferecendo ao usuário uma série de comandos para percorrer as respostas e obter no monitor as vistas

correspondentes (Figura nº 1).

Num primeiro momento o microcomputador funciona no modo "transparente", isto é, tudo que é digitado no terminal é enviado ao computador central. Inversamente, tudo que vem do computador central é enviado à tela do terminal. O usuário pode assim conduzir uma sessão normal. Ele pode, particularmente, submeter uma questão aos bancos de dados. O computador responde então enviando o texto dos documentos pertinentes. O microcomputador que vigia o fluxo de dados detecta primariamente o fluxo de respostas. Ele passa então a um segundo modo de funcionamento, dito modo "local".

No modo local, o fluxo proveniente do computador central – isto é, as respostas à questão colocada – não é mais dirigida para a tela do terminal, mas é estocada na memória do microcomputador. O usuário está então informado que pode começar a consultar os documentos recebidos, enquanto outros continuam a chegar.

Ele dispõe para isto de um certo número de comandos que lhe permitem obter na tela os documentos recebidos

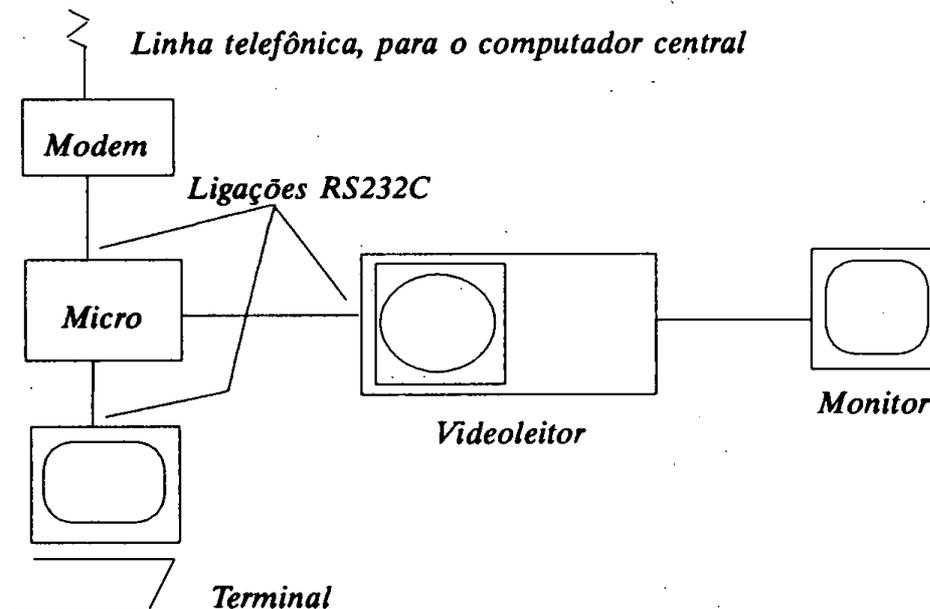


Figura nº 1: Esquema das conexões da estação de consulta