



UNIVERSIDADE NOVA DE LISBOA
DEPARTAMENTO DE INFORMÁTICA

ORBI-UM SISTEMA PERICIAL PARA O ORDENAMENTO
BIOFISICO DO TERRITÓRIO INTERROGÁVEL EM POR
TU GÜES

Luís Moniz Pereira
Eugénio de Oliveira
Paul Sabatier

FCT/DI-4/82

Março 82

ORBI-UM SISTEMA PERICIAL PARA O ORDENAMENTO
BIOFISICO DO TERRITÓRIO INTERROGÁVEL EM POR
TUGUÊS

Luís Moniz Pereira
Eugénio de Oliveira
Paul Sabatier

FCT/DI-4/82

Março 82

ORBI - UM SISTEMA PERICIAL PARA O ORDENAMENTO
BIOFISICO DO TERRITORIO INTERROGAVEL EM PORTUGUES

Luis Moniz Pereira
Eugenio Oliveira
Paul Sabatier

Departamento de Informatica
Universidade Nova de Lisboa

Marco 1982

RESUMO

Descreve-se aqui um sistema informatico, interrogavel num subconjunto do Portugues, e implementado em Prolog num pequeno computador, o qual incorpora e assimila conhecimento de especialistas no dominio do ordenamento biofisico do territorio. Tal sistema e' capaz de explicar a aplicacao desse conhecimento a uma base de dados territoriais assim como de responder a perguntas sobre as suas possibilidades linguisticas. O sistema ORBI foi desenvolvido no espaco de um ano, sob contracto de prestacao de servicos entre o Nucleo de Inteligencia Artificial do Departamento de Informatica da UNL e a Direccao de Estudos do Ambiente da Secretaria de Estado do Ambiente e Qualidade de Vida.

Palavras chave : Sistemas Periciais, Linguagem Natural, Prolog, Ordenamento.

INDICE

Introducao
Dominio da aplicacao
Protocolo
Configuracao do sistema
Organizacao dos programas
Dominio de conhecimento
 Base de conhecimentos
 Deducao, avaliacao e explicacao
 Acesso 'a base de dados
 Aquisicao do conhecimento e
 Metaconhecimento
Competencia de linguagem natural
 Competencia linguistica
 Analise lexical
 Analise sintactica
 Analise semantica
 Competencia metalinguistica
 Desenvolvimentos linguisticos futuros
Conclusoes
Agradecimentos
Referencias

INTRODUCAO

=====

O nosso sistema pericial compreende uma base de dados com factos e uma base de conhecimentos com regras de inferencia, sendo capaz de responder a questoes usando essas regras para raciocinar sobre os factos. Tambem explica tais raciocinios, com maior ou menor detalhe, de acordo com as necessidades do utilizador. Regras e factos podem ser actualizados pelo utilizador e novos items serem automaticamente incorporados sem necessidade de reprogramacao. As perguntas podem ser feitas quer em linguagem natural (Portugues) quer seguindo as opcoes de um menu. Podem dizer respeito ao dominio de conhecimento, incluindo a maneira como tal conhecimento esta' organizado, ou sobre as possibilidades de linguagem natural que o sistema possui, aliviando em consequencia o utilizador da necessidade de consultar um manual.

-- E' o primeiro sistema pericial escrito em Prolog incluindo linguagem natural, explicacoes e metaconhecimento. Alem disso abarca uma applicacao realista (do mundo real com um tamanho realista).

-- E' o primeiro dos sistemas periciais envolvendo uma tal combinacao de possibilidades a ser implementado num pequeno computador (PDP11/02). Para que tal fosse possivel foi crucial a utilizacao da linguagem Prolog.

-- Inclui a primeira interface em linguagem natural com capacidades metalinguisticas. Alem disso a gramatica compreende o tratamento das elisoes, extrapolicao e a coordenacao de frases e de relativas. Adicionalmente assegura grande numero de testes quer sintacticos quer semanticos, que despoletam mensagens de erro apropriadas. Tambem a formula semantica e' gerada sob forma otimizada para permitir a sua execucao sem necessidade de planeamento extra.

DOMINIO DA APLICACAO

=====

O dominio de conhecimento envolve a avaliacao da oferta biologica e fisica do meio. A dimensao tipica da regioa e' a do concelho, a qual e' descrita por intermedio de 23 descriptors. A cada descriptor corresponde um mapa da regioa coberto por uma grelha uniforme, formando quadrados de 200 metros de lado. Em cada quadrado os valores dos descriptors sao digitalizados em inteiros de 1 a 5. Alem disso a cada valor e' associada uma representatividade, tambem um inteiro de 1 a 5, descrevendo o grau de homogeneidade do valor no quadrado considerado.

Cada concelho compreende em media 6000 quadrados. Exemplos de descriptors:

d1 -- qualidade dos microclimas
d9 -- riscos de erosao
d11-- equilibrio da vegetacao

Dada a base de dados dos factos observados expressos pelos descriptors, informacao derivada mais abstracta pode ser obtida definindo factores e aptidoes. Um factor e' o resultado de uma combinacao de descriptors. Tem um valor inteiro entre 1 e 5 ou um valor s, significando ser esse factor selectivo. Tambem aqui a cada valor e' associada uma representatividade entre 1 e 5. O valor de cada factor e' computado atraves dos valores dos descriptors que o suportam atraves de uma regra fornecida pelos especialistas. Uma regra pode ter varios casos alternativos dependendo dos valores actuais dos descriptors. Por pto, um descriptor pode ter um valor que se sobrepe a todos os outros e como tal definir por si so' completamente o factor, ou ainda o valor do factor

pode ser uma soma pesada dos valores de todos, ou de alguns, descriptors em que se baseia de acordo com os valores presentes. A representatividade dos factores contudo, e' sempre uma soma ponderada das representatividades dos descriptors intervenientes. Ha' 16 factores. Exemplos sao:

- f1 -- conforto climatico
- f3 -- resistencia do solo 'a erosao
- f6b-- riquezas cinegeticas e paisagisticas

Aptidoes sao analogamente definidas em termos de factores, tal como estes em termos de descriptors, atraves de regras de inferencia fornecidas por uma equipa de especialistas. Existem 5 aptidoes :

- a1-- habitacao
- a2-- agricultura nao intensiva
- a3-- agricultura intensiva
- a4-- industria
- a5-- infraestruturas de recreio

Finalmente uma regra sobre a sintese biofisica baseia-se nos valores e representatividade das aptidoes para extrair indicacoes de areas privilegiadas, de acordo com os objectivos de desenvolvimento.

O conhecimento dos especialistas e' tambem tornado explicito atraves de um grande numero de textos explicativos, um para cada valor possivel de cada descriptor.

PROTOCOLO =====

Qual o valor da aptidao 3 no ponto de coordenadas 103,709 ?

No ponto de coordenadas 103-709

a aptidao 3 tem o valor muito apto (com representatividade 2)

Pretende explicacoes sobre as razoes que me levaram a concluir terem os pontos considerados o(s) valor(es) indicado(s) para a(s) aptidao(oes) em causa ?
Responda sim ou nao .

sim.

Quer explicacoes muito detalhadas ? Sim ou nao?

sim.

No ponto de coordenadas 103,709

a aptidao 3 tem o valor muito apto (5 com representatividade 2) porque:

5 foi calculado por

$$(25*(6-1)+10*4+5*4+60*5+50)/100$$

por substituicao em

$$(25*(6-F6d)+10*F3+5*F2d+60*F4+50)/100$$

***o factor 2b tem o valor 4 (com representatividade 5) porque:

o descriptor 9 tem o valor 4 (com representatividade 5)

porque segundo o especialista:

Os aquiferos ocorrentes mais importantes apresentam uma protecção ELEVADA dado pertencerem ao manto mio-pliocenico cativo.

*** o factor 3 tem o valor 4 (com representatividade 4) porque:

4 foi calculado por
 $(30*(6-1)+70*4+50)/100$ por substituição em
 $(30*(6-07)+70*010+50)/100$

o descriptor 7 tem o valor 1 (com representatividade 5)
porque segundo o especialista:

O escoamento pluvial superficial é MUITO BAIXO dado que as características da bacia hidrográfica onde se encontra incluído e a pluviosidade média anual de terminam um número de aspereza corrigido INFERIOR A 150.

o descriptor 10 tem o valor 4 (com representatividade 3)
porque segundo o especialista:

Os solos são considerados como tendo risco de erosão BAIXO dado apresentarem como limitações as associadas à sua reduzida espessura efectiva.

*** o factor 4 tem o valor 5 (com representatividade 1) porque :

o descriptor 10a tem o valor 5 (com representatividade 1)
porque segundo o especialista :

Os solos são susceptíveis de utilização agrícola de grande intensidade. A aptidão pedológica à agricultura pode ser considerada MUITO ELEVADA. Estes solos estão muito provavelmente total ou parcialmente protegidos pelo artigo 1 do D.L.308/79.

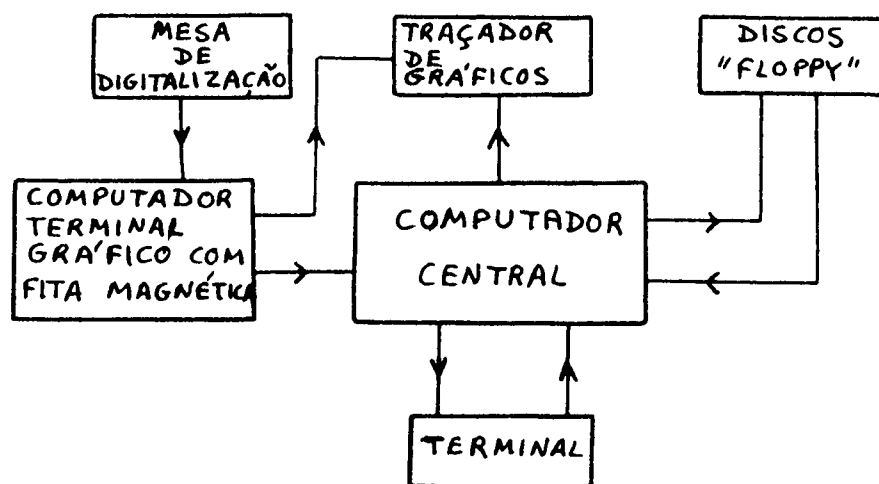
***o factor 6d tem o valor 1 (com representatividade 1) porque :

o descriptor 11 tem o valor 1 (com representatividade 1)
porque segundo o especialista :

O equilíbrio da vegetação é considerado MUITO BAIXO devido ao facto das fitocenoses naturais se encontrarem num estado muito degradado com regresso extremamente difícil ao climax potencial.

CONFIGURACAO DO SISTEMA

=====



A digitalização é feita "off line" usando um digitalizador Tektronix 4956 e o computador terminal gráfico (com fita magnética) Tektronix 4052. Para tal é usada a linguagem Basic. O leitor interessado pode consultar [DIA] para detalhes. Basta dizer que o utilizador pode indicar que todos os pontos de uma área fechada tem o mesmo par valor/representatividade fornecendo apenas o seu contorno. O programa atribui esse mesmo par a todos os pontos da grelha dentro desse contorno. Qualquer conflito entre contornos é figurado graficamente e uma decisão é pedida ao utilizador.

O "plotter", um Tektronix 4662, pode ser usado para monitorar e fornecer uma "hard copy" do que foi digitalizado, assim como periférico de saída onde as respostas às perguntas podem ser dadas em mapas.

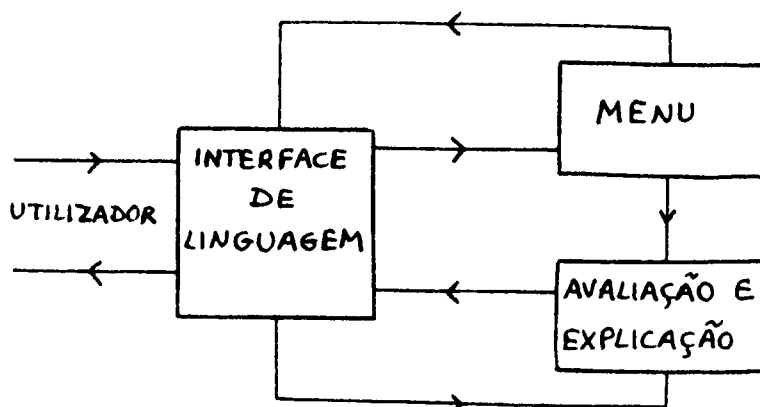
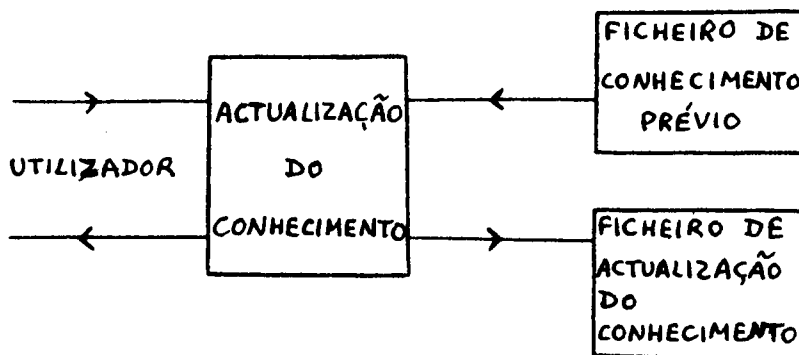
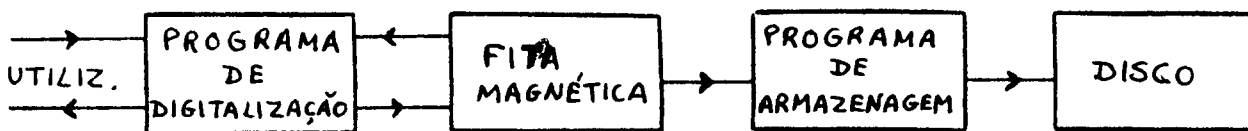
Os pontos digitalizados e os seus descritores são enviados do terminal gráfico para o computador central, o qual os armazena no disco em cláusulas unitárias.

O computador central, um Heathkit 11/02 com um CPU LSI11 endereça 64K bytes de memória central, e corre um interpretador Prolog sob o sistema de exploração RT11. Comunica com a unidade Heathkit H27 de 2 discos flexíveis de densidade simples. O terminal é um VT100 da DEC. Esta é planeada a substituição a curto prazo do computador central por um PDP11/23 com 256 Kbytes, três vezes mais rápido e dos discos flexíveis por um disco duro de tecnologia Winchester de 80 Mbytes. Este último permitirá a utilização do sistema de exploração UNIX e do interpretador Prolog correspondente [CLO].

Estamos bastante satisfeitos com o sistema Prolog, que do ponto de vista externo é semelhante ao Prolog para o sistema DEC10 [CLO]. O interpretador completo ocupa 17 Kbytes mas pode ser comprimido para 12 Kbytes removendo o pacote de "debug" ou até para menos se for necessário. O sistema de exploração RT11 tem 5 Kbytes aproximadamente. Enriquecemos o sistema Prolog com a adição de novos predicados, como "all" para obter conjuntos de soluções [PER-2], e ainda a facilidade de carregar um módulo programado em Prolog a partir de outro módulo. Esta possibilidade é indispensável uma vez que os módulos do nosso programa não cabem conjuntamente em 64 Kbytes de memória. A comunicação entre módulos é feita através de ficheiros. Presentemente o carregamento de um módulo leva 4 segundos. Com o PDP11/23, em vez de ter de transferir o próximo programa do disco, ele poderá ser armazenado noutra banca de memória para uso imediato.

ORGANIZACAO DO SISTEMA

=====
 Ha' quatro programas separados : um em Basic para a digitalizacao atras re-ferida; outro em Pascal para receber os pontos digitalizados e armazena'-los em disco; outro ainda em Prolog para actualizacao do conhecimento; e um quarto compreendendo tres modulos basicos, todos em Prolog- a interface em lingua-gem natural, o menu e o modulo de avaliacao e explicacao- todos ligados como se mostra a seguir.



Organizacao do programa

Cada modulo basico e' ele proprio decomponivel noutros modulos. Eles serao descritos detalhadamente nas seccoes apropriadas. Para os nossos propositos actuais basta realcar que esses modulos nao coexistem na memoria. Inicialmente o utilizador comunica com a interface em linguagem natural, seja para ex-

primir uma pergunta, um comando ou um facto, seja para dar a iniciativa ao programa enviando pelo teclado a palavra 'fala'. Neste caso o modulo menu e' carregado em memoria central e o utilizador guiado atraves de um menu de questoes possiveis ou ajudado na construcao da sua pergunta. No entanto, a qualquer momento pode recuperar a iniciativa batendo a palavra 'ouve' como resposta a qualquer pergunta do programa. Nesse caso a interface em linguagem natural e' entao de novo carregada. Por outro lado, depois de identificada a pergunta o modulo de avaliacao e explicacao e' carregado e a pergunta lida de um ficheiro.

Se uma pergunta, comando ou instrucao e' enviado 'a interface de linguagem natural, esta produz a semantica correspondente, escreve-a num ficheiro e carrega o modulo de avaliacao e explicacao. Durante a avaliacao, modulo que tem acesso 'a base de dados externa em disco com os pontos e seus descriptors, uma estrutura explicativa e' gerada caso o utilizador mais tarde a requeira. Depois da resposta ser produzida no terminal e/ou desenhada no "plotter", e uma explicacao eventualmente pedida e recebida, o modulo de linguagem natural e' recarregado ficando pronto para novas perguntas ou, em caso do utilizador querer, terminar a sessao ao ser-lhe enviada a palavra 'adeus'.

O programa de actualizacao do conhecimento le a base de conhecimentos, interactua com o utilizador para aceitar quaisquer actualizacoes, faz-lhes testes de forma e consistencia e produz um ficheiro com informacao actualizada. O utilizador podera' depois instruir o programa principal para os incorporar.

BASE DE CONHECIMENTOS =====

A notacao propria ao Prolog de Edimburgo sera' usada a partir daqui. A base de conhecimentos compreende um conjunto de regras para computar os valores de factores, aptidoes e

sintese do melhor uso. A base de dados consiste em clausulas unitarias, uma por ponto, da forma :

ponto(X,Y, d(v1-r1,v2-r2,...)).

onde X e Y sao coordenadas e d e' o functor dos descriptors, contendo os pares valor-representatividade para cada descriptor.

Regras tipicas para uma aptidao e para um factor sao apresentadas ao utilizador como termos, como se mostra seguidamente, com operadores infixos apropriados. Tais regras transformam-se em clausulas do programa como se ilustra no caso abaixo da regra para um factor.

a3=Va3 com Ra3=R tal que

Va3=0 e Ra3=R

se Vf4=0 onde f4=Vf4 com R
ou Vf6a=s onde f6a=Vf6a com R
ou Vf6b=s onde f6b=Vf6b com R
ou Vf6c=s onde f6c=vf6c com R

senao Va3= (25*(6-Vf6d)+10*Vf3+5*Vf2b+60*Vf4+50)/100
e Ra3= (25*Rf6d+10*Rf3+5*Rf2b+60*Rf4+50)/100

onde $f6d=Vf6d$ com $Rf6d$
 e $f2b=Vf2b$ com $Rf2b$
 e $f3=Vf3$ com $Rf3$
 e $f4=Vf4$ com $Rf4$.

$f6c=Vf6c$ com $Rf6c$ tal que

$d6a=Vd6a$ com $Rd6a$
 e $d23a=Vd23a$ com $Rd23a$

e ($Vf6c=0$ e $Rf6c=R$

se $Vd6a=0$ e $R=Rd6a$ ou $Vd23a=0$ e $R=Rd23a$

senao $Vf6c=s$ e $R= \text{maximo}(Rd6a,Rd23a)$) .

Esta regra e' transformada na clausula :

```
f('6c',D,Vf6c-Rf6c) :- d('6a', D, Vd6a-Rd6a )
                        e d('23a',D,Vd23a-Rd23a)
                        e ( Vf6c=0 e Rf6c=R
                            se Vd6a=0 e R=Rd6a ou Vd23a=0 e R=Rd23a
                            senao Vf6c=s e maximo(Rd6a,Rd23a,Rf6c) ) .
```

onde D recebe o functor dos descriptors dos pontos para os quais se deve calcular $f6c$.

Como exemplo de clausula de descriptors temos :

```
d(9,D,D9) :- arg(13,D,D9), tabela(9,D9,E9).
tabela(9,V-R,E9) :- t9(V,E9), !.
t9(1,mesg(m1)).
...
```

significando que o par valor-representatividade do descriptor 9 esta' ligado a D9 que e' o argumento 13 do functor dos descriptors ligado a D. O predicado tabela liga a variavel explicacao E9 a um termo da forma mesg(m), de acordo com o valor de V, onde m e' o nome de um ficheiro e 'mesg' um predicado do sistema desenvolvido por nos, usado mais tarde para fazer sair pelo terminal a mensagem explicativa guardada no ficheiro 'm', elucidando o significado de o descriptor 9 ter o valor V. De notar contudo que E9 so' tem uma ocorrencia. Isto e' assim porque as clausulas da base de conhecimentos nao sao directamente chamadas. Elas sao usados sob o controlo de um interpretador especial. O mesmo sucedendo com as regras para os factores e aptidoes, cujo formato as torna mais faceis de serem compreendidas pelo utilizador depois de traduzidas para a notacao original em que deram entrada.

DEDUCAO, AVALIACAO E EXPLICACAO

=====

Feita uma pergunta, tres processos tem lugar em sequencia : deducacao, avaliacao e explicacao. Na fase de deducacao, todas as chamadas e subchamadas da semantica da pergunta sao executadas, e a parte deterministica da 'arvore das

explicacoes e' tambem gerada. O objectivo e' conseguir eficiencia em perguntas envolvendo um conjunto de solucoes diferindo apenas nas suas partes nao-deterministicas. Seguidamente, na fase de avaliacao, as partes nao-deterministicas da pergunta e da estrutura explicativa sao avaliadas. Finalmente, os resultados sao apresentados e explicados. Esquemticamente teremos :

```
execucao( Pergunta ) :-  
    ^deducao( Pergunta, Parte_nao_deterministica, Explicacao ),  
    avaliacao( Parte_nao_deterministica ),  
    explicacao( Resposta, Explicacao ).
```

onde ^ e' um predicado de "garbage collection" que recupera o espaco usado na fase de deducao :

```
^G :- G, assert(G), fail.  
^G :- retract(G), trimcore.
```

Para um entendimento mais preciso, considere-se a clausula :

```
deducao( all(X,P,S), all(explicacao(X,E), NDP, S), E ) :- !, deducao(P,NDP,E).
```

Toma uma pergunta da forma "all" e transforma-a, onde P e' substituido pela sua parte nao-deterministica NDP, e a explicacao E resultante da execucao das partes deterministicas de P e' ligada a cada objecto X.

A proxima clausula vai tomar uma clausula para uma chamada 'a' aptidao A (com A especificada), executa a parte determinada do seu corpo B, e devolve a parte nao-determinada NDB, e a parte determinada da explicacao de B, 'a' qual acrescenta a explicacao da chamada :

```
deducao( a(A,D,V), NDB, a(A,V)&E ) :- nonvar(A), clause( a(A,D,V), B), !,  
    deducao( B, NDB, E).
```

As partes nao-determinadas da explicacao sao obtidas durante a fase de avaliacao. Assim :

```
deducao( A senao B, solve(NDA,E=EA,! ; NDB,E=EB ), E ) :-!,deducao(A,NDA,EA),  
    deducao(B,NDB,EB).
```

onde solve(G):- G.

E' necessario o '!' em 'solve' para nao afectar outras partes da semantica.

Os nos terminais da estrutura explicativa sao nao-deterministicos. Sao gerados pelas clausulas dos descriptors como as chamadas ao predicado 'tabela', que produz as mensagens relevantes para a explicacao, como ja' vimos, durante a fase de avaliacao :

```
deducao( d(N,D,V), Corpo, d(N,V)&E ) :- nonvar(N),  
    clause( d(N,D,V), Corpo), !,  
    Corpo= ( arg(A,D,V),tabela(N,V,E) ).
```

Para as chamadas nao-deterministicas que nao contribuem para a explicacao existe uma clausula final onde elas resolvem : deducao(G,G,nil).

ACESSO À BASE DE DADOS

=====

A base de dados relacional dos pontos não está em memória mas como cláusulas unitárias em disco. Porque assumimos que cada pergunta tem quando muito uma chamada $p(X,Y,D)$, podemos usar o apontador do ficheiro para manter o registo da última cláusula lida :

```
ponto(X,Y,D) :-see(dados), seen,          /* ficheiro seguramente fechado */
                see(dados),              /* abre ficheiro */
                repeat,                  /* cria um ponto de retorno */
                read(P),                  /* le próxima cláusula unitária */
(P=end_of_file, !, seen, fail ;         /* não há mais soluções */
 (var(X);var(Y)), P=ponto(X,Y,D) ;      /* solução não-determinística */
 nonvar(X),nonvar(Y), P=ponto(X,Y,D), !, /* solução determinística, logo */
                seen ).                  /* fecha o ficheiro */
```

AQUISIÇÃO DO CONHECIMENTO E METACONHECIMENTO

=====

O programa de aquisição de conhecimento, ainda em desenvolvimento, pretende ajudar o utilizador autorizado a reformular a base de conhecimentos, em particular através de uma notação apropriada, fazendo testes de consistência e não requerendo qualquer familiaridade com pormenores de implementação. O metaconhecimento é necessário para tal, assim como para o sistema poder responder a perguntas sobre si próprio.

Tentamos não duplicar informação na base de conhecimentos, para facilitar a actualização automática das estruturas relevantes, e ter estruturas de dados facilmente lidas e usadas por outras partes do programa, e facilmente executáveis como programas que também são.

Numa frase, as cláusulas da base de conhecimentos devem ser usadas quer como agentes processadores de conhecimento activo quer como informação passiva para os agentes do metaconhecimento. Idealmente os agentes do metaconhecimento deverão ser auto-descritivos. Achamos o Prolog útil também neste caso, e ainda para escrever interpretadores [PER-3] para conseguir os regimes de controlo requeridos pelo processamento do conhecimento. Sobre estas possibilidades daremos conta em futuras publicações.

COMPETÊNCIA DE LINGUAGEM NATURAL

=====

COMPETÊNCIA LINGÜÍSTICA

A competência linguística de DRBI é conseguida por meio da análise lexical e da análise sintática/semântica, que transformam uma frase em linguagem natural numa lista de chamadas Prolog imediatamente avaliável, e cuja ordem é a ótima. Por exemplo, a pergunta :

>Quais são os pontos tendo o mesmo valor para a aptidão agricultura intensiva?

é transformada na expressão Prolog (cf.[PER-2] para a definição de 'all') :

```
all ( V-S ,
      all ( [ponto:(X,Y)] , (p(X,Y,D),a(3,D,V-R)) same V , S ),
      L )
```

onde 3 e' o codigo de 'agricultura intensiva', e L e' a resposta sob a forma de uma lista de pares V-S, tal que para cada valor V da aptidao 3, S e' a lista dos pontos com coordenadas X,Y que tem valor V nessa aptidao.

ANALISE LEXICAL

Como em Pique e Sabatier [PIQ] a analise lexical substitui cada palavra da frase entrada pela correspondente categoria lexical. Isto e' feito indexando o vocabulario de ORBI pelas palavras. Associada a palavra 'pontos' temos a seguinte informacao lexical, sob a forma de clausula unitaria :

```
pontos( nome_comum( ponto, pontos, masculino, plural) ).
```

de acordo com o esquema :

```
<palavra>( <cat.lex.>( <morf.prof.>, <morf.superf.>, <genero>, <numero> ) ).
```

Quando nao existe entrada para uma palavra ela e' dada como desconhecida, a nao ser que seja detectada como parte de um nome composto :

```
nome_composto( [agricultura,intensiva!L], L, numero(3) ).
```

ANALISE SINTACTICA

A analise sintactica/semantica e' realizada por meio de uma gramatica nuclear contendo regras independentes do e tambem sensiveis ao contexto, expressas no formalismo das gramaticas de clausulas definidas de Pereira e Warren 1980 [PER], e incluindo testes sintacticos e semanticos com mensagens de erro. As regras lidam com as estruturas fundamentais do Portugues, nomeadamente :

- perguntas sim/nao
- perguntas que/qual/quantos/onde/porque (pronomes interrogativos)
- comandos
- clausulas afirmativas, negativas, relativas, preposicionais, coordenadas extrapostas e elididas
- qualificacao, abreviacao e complementacao complexa de substantivos
- quantificacao universal, existencial, cardinal, definida e indefinida
- verbos e advverbios
- adjectivos
- terminologia metalinguistica, permitindo interrogar o programa sobre a sua propria competencia linguistica

Esta gramatica nuclear e' independente das applicacoes, sendo transportavel para outros dominios. Tal foi ja' efectivamente feito (Pereira,Porto 1982 [PER-1,PER-4]).

A gramatica nuclear e' completada por um modulo dependente da applicacao (10 vezes mais pequeno) contendo estruturas, vocabulario (basicamente substantivos e verbos) e controle semantico dependentes do dominio de ORBI.

Dedicamos muita energia a incorporacao de estruturas elididas e extrapostas na competencia linguistica de ORBI porque elas sao essenciais a uma boa e

natural interacção. Uma interface de linguagem natural incapaz de perceber frases envolvendo estas estruturas não é verdadeiramente natural.

O control sintactico e semantico verifica os acordos de numero e genero, qualificacao de nomes complexos, e a compatibilidade entre substantivos e verbos e seus complementos, indicando quaisquer erros. Exemplos :

> Em cada ponto, quais sao a factor de cada aptidoes ?

* Falta de acordo entre 'a' e 'factor', entre 'valem' e 'factor', entre 'cada' e 'apitidoes' !

> Da'-me o valor da aptidao resistencia dos solos 'a erosao dos pontos cujo descriptor potencial do recreio e' superior a 2 !

* 'resistencia dos solos 'a erosao' e' um factor, nao uma aptidao !

As analises sintactica e semantica nao sao separadas como no CHAT-80(Warren ,Pereira 1981[PER]) mas misturadas como no MICROSIAL(Pique,Sabatier1981[PIQ]). Tal solucao e' preferivel para parar a analise sintactica logo que e' detectado um erro semantico. E' o caso, por exemplo, quando e' empregue um complemento ilegitimo para um verbo. Nesse caso, ORBI informa o utilizador. Proporciona ainda uma gramatica mais compacta. Exemplos de erros semanticos :

> Sera' que a regioa de cada ponto tem o valor muito apto para a aptidao 3 ?

* 'ponto' nao poder ser complemento de 'regiao' neste contexto !

> Quantos factores sao superiores a cada ponto ?

* 'factores' nao pode ser sujeito do verbo 'ser' com o complemento 'ponto' !

FRASES COMPLETAS E COM ELISOES. ESTRUTURA DOS SINTAGMAS

Estrutura das FRASES COMPLETAS das perguntas e comandos, com possiveis extraposicoes (os elementos facultativos indicam-se entre parenteses) :

a) As perguntas sim/nao tem a forma : (sera' que) SN_SUJEITO SV ?

> Sera' que cada ponto cujo factor resistencia ao incendio e' igual a 2 nao tem nenhuma aptidao superior a 3 ? (notar a dupla negacao)

> Alguns pontos que tem o mesmo valor para a aptidao industria tem o factor riquezas do subsolo entre 1 e 3 ?

> A media dos valores da aptidao 2 dos pontos da regioa 101,106;120,140 e' inferior 'a media da aptidao 3 dos pontos da regioa 202,215;240,260 ?

b) As perguntas com pronome interrogativo tem a estrutura :

PRONOME_E_SN_COMPLEMENTO SN_SUJEITO (NEGACAO) VERBO ?

PRONOME_E_SN_COMPLEMENTO (NEGACAO) VERBO (DUPLA NEGACAO) SN_SUJEITO ?

>Que valor tem cada aptidao dos pontos cujo descriptor riscos de erosao e' 5 ?

- > A que factores cada aptidao de cada ponto que existe e' igual ?
- > Qual e' a regioao que nao tem nenhum ponto com o valor 1 para a aptidao industria ? (notar a dupla negacao)
- > Que valor e' que tem cada descriptor da aptidao 2 dos pontos tendo 2 factores iguais a 1 ?
- > O que e' a aptidao A2 ?
 - ou a estrutura : PRONOME_SN_SUJEITO SV ?
- > Que pontos tem a aptidao agricultura intensiva com valor muito apto ?
- > Que factores sao superiores ao valor da aptidao 2 do ponto 11,40 ?
- > Quantos pontos tendo o mesmo valor para cada aptidao existem ?
- > Porque tem a aptidao industria para o ponto 121,57 o valor nao apto ?
 - c) Os comandos tem a estrutura: (IMPERATIVA_VERBO_TRANSITIVO) SN_OBJECTO !

- > Da'-me a media dos valores do factor conforto climatico dos pontos que tem o mesmo valor para a aptidao 4 !
- > O valor de cada factor do ponto 205,102 !
- > Os pontos da regioao 105,206;118,219 que tem 2 aptidoes com o valor 1 ?
 - d) As extraposicoes sao permitidas nas perguntas e comandos. ORBI compreende a extraposicao esquerda dos complementos de qualquer SN_SUJEITO. Os complementos extrapostos podem ocorrer no inicio da frase e/ou imediatamente antes do SN_SUJEITO. Exemplos :
- > Em cada ponto, quais sao os factores de cada aptidao que sao superiores a 2 ?
- > Quanto vale, para o ponto 28,105, o descriptor permeabilidade do substracto ?
- > Sera' que, em alguns pontos, a aptidao A3 e' igual ao factor F6 ?
 - Cada extraposicao e' primeiro analisada sintaticamente e depois concatenada ao SN_SUJEITO quando e' detectada a sua terminacao.
- As ELISOES podem ocorrer na mesma frase ou entre frases consecutivas. Nos exemplos abaixo as sequencias de palavras elididas sao indicadas entre [e], e o ponto onde foram elididas por [...] . No primeiro caso ORBI lida com as elisoes do sujeito e/ou verbo das clausulas coordenadas :
 - > [O ponto 69,103 tem] o valor apto para a aptidao industria e [...] o valor do factor resistencia do solo 'a erosao inferior ao F4 ?
 - > [Cada ponto] tendo o F4 inferior a 3 ou [...] nao tendo o F12 inferior a 2 tem o valor muito apto para a aptidao habitat concentrado ?

A interpretação destas elisões não é complexa. Os diferentes tipos de hiatos são esperados pelas regras de gramática para a coordenação de cláusulas. A elisão entre frases é bem mais complexa e difere de acordo com o tipo de frase. Abaixo, fornecemos primeiro uma frase completa e depois uma elidida precedida por *, e introduzida por 'E' (ou 'e') :

a) Elisões entre frases sim/não

- elisão de todo o SN sujeito :

> [Será que a aptidão habitat concentrado dos pontos cujo descriptor D8 é igual a 3], é entre 1 e 4 ?

* E [...] é igual a 3 ?

- elisão de todo o SV :

> Será que a aptidão habitat concentrado dos pontos cujo descriptor piroresistência vegetal é igual a 3 [é entre 1 e 4] ?

* E os factores do ponto cuja aptidão indústria é igual a 5 [...] ?

- elisão de parte do SN sujeito e de todo o SV :

> [O descriptor D10] dos pontos na região 91,707;103,900 que é entre 3 e 5 [é inferior ao descriptor escoamento pluvial superficial do ponto 26,40] ?

* E [...] dos pontos cuja aptidão 2 é superior a 3 [...] ?

> [O descriptor D8] dos pontos na região 91,707;103,900 [que é entre 3 e 5, é inferior ao descriptor piroresistência vegetal do ponto 26,40] ?

* E [...] na região 100,109;140,150 [...] ?

> O descriptor D6 [dos pontos da região 91,707;103,900 que é entre 3 e 5 é inferior ao descriptor D2 do ponto 26,40] ?

* E o factor riquezas do subsolo [...] ?

b) Elisões entre perguntas com pronome interrogativo

- elisão de todo o SN sujeito :

> Que [pontos da região 23,78;84,98] tem o valor apto para a aptidão A1 ?

* E [...] tem o valor 5 para a aptidão indústria ?

- elisão de todo o SV :

> Que pontos da região 54,57;78,80 [tem o valor não apto para a aptidão agricultura intensiva] ?

* E que pontos tendo o descriptor escoamento pluvial superficial com o valor 5 [...] ?

- elisão de parte do SN sujeito e de todo o SV :

- > [Que descriptors] da aptidao 3 no ponto 84,12 que sao superiores a 2 [sao iguais aos factores do ponto 32,10] ?
- * E [...] da aptidao 4 dos pontos da regioao 45,67;12,34 [...] ?
- > [Que descriptors] do factor 5 [no ponto 14,36 que sao entre 2 e 4, sao iguais aos factores do ponto 45,56] ?
- * E [...] da aptidao recreio [...] ?
- > Que factores [da aptidao industria no ponto 67,54 que sao inferiores a 4 sao superiores ao factor riquezas geograficas do ponto 45,78] ?
- * E que descriptors [...] ?

c) Elisoes entre comandos

- elisao de todo o verbo :

- > [Da'-me] o valor do factor F8 do ponto 23,175 ?
- * E [...] os pontos cujo factor F8 e' entre 3 e 5 ?
- elisao de parte do SN sujeito :
- > [Os descriptors] da aptidao 4 do ponto 184,78 que sao superiores a 2 ?
- * E [...] dos pontos da regioao 91,45;76,62 ?
- > [Os descriptors] da aptidao 6 [de cada ponto] que sao maiores que 3 ?
- * E [...] de cada factor [...] ?
- > Os descriptors [da aptidao industria do ponto 64,81 que sao menores que 2] ?
- * E os factores [...] ?

d) Como sao resolvidas as elisoes entre frases

As elisoes que ocorrem entre frases sao resolvidas por ORBI de acordo com a estrutura da frase anterior no dialogo. Para o conseguir, para cada frase gramatical aceite por ORBI sao memorizadas as sequencias morfologicas que poderiam ser elididas na frase seguinte, caso ela comece por 'E' (ou 'e'). Por exemplo, relativamente 'a frase :

- > Que descriptors da aptidao agricultura intensiva no ponto 56,78 que sao superiores a 2, sao iguais aos factores do ponto 12,95 ?

ORBI memorizara' as sequencias morfologicas correspondentes 'as estruturas :

[que descriptors]
 [da aptidao agricultura intensiva]
 [no ponto 56,78]
 [que sao superiores a 2]
 [sao iguais aos factores do ponto 12,95]

que poderao ser elididas na frase seguinte se ela for eliptica ; caso contrario elas sao esquecidas. Se a proxima frase for, por exemplo :

* E no ponto 65,78 ?

ORBI tenta preencher os hiatos da frase com as diferentes estruturas de acordo com certas regras, tentando encontrar uma analise gramaticalmente correcta. No exemplo acima a reconstituicao correcta e' :

* Que descriptors da aptidao agricultura intensiva no ponto 65,78 que sao superiores a 2, sao iguais aos factores do ponto 12,95 ?

A tecnica consiste em transformar a frase eliptica numa frase completa, e analisa-la em seguida. A nova frase ex-eliptica e' formada morfologicamente por partes de uma frase diferente, e por isso o controle dos acordos de genero e numero e' suspenso durante a sua analise.

E' claro que uma frase eliptica pode ser seguida por ainda outra :

> [Que factores] do ponto 21,66 [que sao diferentes de 1 sao iguais ao descriptor riscos de erosao] ?

* E [...] do ponto 45,67 [...] ?

* E [...] dos pontos da regioao 68,23;89,20 [...] ?

* E [...] do ponto com o valor pouco apto para a aptidao recreio [...] ?

Damos seguidamente alguns exemplos de ESTRUTURAS SINTAGMATICAS nominais (SN) e verbais (SV) que ORBI compreende.

a) Os sintagmas nominais tem a estrutura :

(DETERMINANTE) (ADJECTIVO) SUBSTANTIVO (COMPLEMENTOS) (CLAUSULAS RELATIVAS)

[o mesmo valor para o descriptor equilibrio da vegetacao nos pontos da regioao 23,41;67,90 que e' entre 2 e 5]

[os pontos cuja aptidao habitat concentrado e' superior a 4, cuja aptidao A2 e' inferior a 3, ou tendo o factor conforto climatico diferente do factor D8]

[a media dos valores do factor F12 dos pontos que tem o valor pouco apto para cada aptidao]

b) Os sintagmas verbais tem as estruturas :

(NEGACAO) VERBO ou (NEGACAO) VERBO (PREPOSICAO) (DUPLA NEGACAO) SN

[nao tem a aptidao 3 inferior a 3]

[e' igual ao factor f5]

ORBI tambem entende a sua conjuncao ou dijuncao, como em :

[sao iguais a 5 ou tem o factor 2 superior ao descriptor D9]

[tem a aptidao agricultura nao intensiva com o valor apto, a aptidao 5 igual a 1 e nenhum factor da aptidao 4 entre 1 e 2]

(notar neste exemplo as elisoes do verbo, permitidas nas frases coordenadas)

VOCABULARIO

O vocabulario de ORBI e' dividido em duas partes: nuclear e especifico.

O VOCABULARIO NUCLEAR (tal como a gramatica nuclear) e' independente da applicacao, e transportavel para outros dominios. Divide-se em duas partes: a linguistica e a metalinguistica. A parte LINGUISTICA contem :

a) Determinantes

o, os, a, as ; um, uns, uma, umas ; nenhum, nenhuns, nenhuma, nenhuma;
algum, alguns, alguma, algumas ; qualquer, quaisquer ; cada ;
todo, todos, toda, todas ; 0, 1, 2, 3, ...

b) Preposicoes de ; a ; em ; com ; sem ; para

c) Contraccoes

ao, aos, as ; no, nos, na, nas ; num, nuns, numa, numas ;
do, dos, da, da ; dum, duns, duma, dumas

d) Verbos

ser: e' , sao ; ter: tem , tendo ;
existir: existe , existem ; haver: ha'

e) Pronomes relativos e interrogativos

que ; porque ; qual, quais ; onde ;
cujo, cuja ; quanto, quantos, quanta, quantas

f) Conectores de frases e clausulas, e negacao e ; ou ; mas ; nao

g) Expressoes pre'-locutorias

sera' que ; e' que ; da'-me, de-me ; diga-me ; desenhe ; adeus

O VOCABULARIO METALINGUISTICO contem as palavras utilizadas para interrogar ORBI sobre a sua competencia linguistica, em particular :

vocabulario[s] ; palavra[s] ; verbo[s] ; nome[s] ; artigo[s] ; adjectivo[s] ;
preposicao[oes] ; pronome[s] ; conjuncao[oes] ; sujeito[s] ; complemento[s]...

O VOCABULARIO ESPECIFICO, especialmente os substantivos, as abreviaturas, os nomes de entidades, verbos, adjectivos e adverbios, referem-se ao dominio especifico de ORBI. Apresentamos aqui apenas um subconjunto dele :

a) Substantivos, abreviaturas e nomes de entidades

regiao[oes], rg ; ponto[s], p ; aptidao[oes], a ; factor[s], f ;
descriptor[es], d ; representatividade[s], rp ; valor[es] ; media[s] ;
agricultura intensiva ; recreio ; industria ; riscos de erosao ;
habitat concentrado ; riquezas geograficas ...

b) Adjectivos mesmo, mesmos, mesma, mesmas

c) Verbos e adverbios

valer: vale, valem ;
ser/e'/sao: igual[ais]/superior[es]/inferior[es]/maior[es]/menor[es]/
diferente[s]/entre .. e ..

ANALISE SEMANTICA -----

Para cada frase gramaticalmente correcta e' produzida uma expressao logica, pronta a ser avaliada, que descreve a sua semantica, mais um termo que expressa o foco de atencao e e' utilizado para a dar a resposta. A semantica de uma frase e' obtida a partir da do sintagma nominal sujeito, representada por SNS, e da do sintagma verbal, SV, liberta de quantificadores como se mostra abaixo. Se o SNS e' quantificado universalmente, a forma final avaliavel da expressao logica difere consoante a pergunta e' sim/nao ou outra. Exemplo :

sim/nao: cada([X,Y], (SNS,SV)) --> nao (SNS, nao SV)

(ie. dizer que todos os [X,Y] que SNS entao SV, e' o mesmo que dizer nao ha' nenhuns que SNS e nao SV, sendo a resposta apenas 'sim' ou 'nao')

outras: cada([X,Y], (SNS,SV)) --> all([X,Y], (SNS,SV), S)

(aqui a resposta passara' a ser conjunto S de todos os [X,Y] que SNS e SV)

Os sintagmas nominais podem conter quantificadores imbricados, cujo escopo e' o inverso relativamente 'a ordem por que aparecem na frase. Alem disso, sao reduzidos a um unico, onde a ordem das variaveis quantificadas exprime o escopo : "cada X de cada Y Propriedades" --> cada([Y,X], Propriedades) . Isto torna o tratamento subsequente da quantificacao mais facil, evitando a recrusao sobre as estruturas semanticas.

Nao ha' tratamento de pressuposicoes (os artigos definidos e indefinidos sao tratados do mesmo modo) pois elas sao desnecessarias no dominio de ORBI. A quantificacao existencial so' precisa ser explicitada no caso de sintagmas nominais com a construcao 'mesmo' , utilizando-se entao a construcao 'same' do predicado de sistema 'all' [PER-2].

Tem-se o cuidado de produzir a boa ordem de avaliacao das chamadas Prolog em todas as ocaes, em particular quando existem clausulas relativas. Quando uma relativa se reporta a um substantivo embora so' apareca depois dos seus complementos, a ordem (S,R,C) e' utilizada na matriz da expressao logica do sintagma nominal. Se a relativa se refere a um complemento, a sua semantica vem imediatamente a seguir 'a dele: (S,C,R). As quantificacoes universais das relativas sao sempre reduzidas 'a forma de avaliacao nao (_ , nao _), antes da sua imbricacao semantica no sintagma nominal se efectuar.

A semantica dos sintagmas verbais e' obtida a partir da semantica V do verbo (com possivel negacao) mais a do possivel sintagma nominal que se lhe segue, onde qualquer eventual quantificacao universal no SN e' previamente reduzida 'a forma nao (_ , nao _) :

V + cada(L, SN)	-->	nao (SN, nao V)
nao V + cada(L, SN)	-->	nao (SN, V)
V + SN	-->	(SN, V)
nao V + SN	-->	(SN, nao V)

O dicionario nao contem a semantica de cada palavra : ela encontra-se incluida em duas tabelas. Numa, as entradas sao 5-tuplos, cada um correspondendo a um par substantivo/complemento. Dadas as respectivas morfologias, os seus predicados individuais, com as variaveis comuns apropriadamente ligadas, podem ser obtidos, mais uma condicao predicativa adicional, se necessaria 'a ligacao (como quinto elemento do tuplo). Exemplo :

```
tabela_de_complementos( factor: F - FV , aptidao: A - AV ,
                        f( F, D, FV) , a( A, D, AV) ,
                        parte_de( f( F, D, FV) , a( A, D, AV) ).
```

A semantica de cada substantivo individual pode igualmente ser obtida desta tabela quando ocorre sem complementos, bastando ignorar os argumentos irrelevantes. Uma tabela semelhante a esta e' usada para os verbos e seus complementos.

COMPETENCIA METALINGUISTICA

Devido 'a riqueza e variedade de estruturas, vocabulario e usos envolvidos na linguagem natural, uma interface para ela falhara' sempre algumas frases produzidas por um utilizador, seja ele casual ou nao. Portanto a interface devera' ser realmente informativa, mostrando explicitamente porque rejeita uma frase, e quais sao as suas capacidades.

Como no MICROSIAL [PIQ], ORBI aponta as palavras desconhecidas, os desacordos gramaticais, as designacoes erroneas de nomes compostos, e certas falsas pressupoicoes ; mas tambem as complementacoes incorrectas de substantivos e os sujeitos e complementos incompativeis dos verbos.

Na verdade, estes diagnosticos nao chegam para assegurar uma boa interacao em linguagem natural. Nalguns sistemas, o utilizador abalanca-se 'a consulta de um manual de referencia, geralmente sem grande sucesso porque demasiado complexo. Uma solucao melhor e' oferecer ao utilizador a possibilidade de inquirir o sistema sobre a sua competencia linguistica, como fizemos ate' certo ponto no nosso sistema. ORBI pode por exemplo responder 'as questoes :

> Da'-me o vocabulario !

> A palavra 'factor' existe ?

> Da'-me as preposicoes
os nomes que existem !
os verbos

> Quais sao os possiveis complementos do substantivo 'ponto' ?

> Quais sao os possiveis sujeitos e complementos de cada verbo ?

O conhecimento necessario para responder a estas questoes metalinguisticas encontra-se agrupado numa base de dados especifica. Desenvolvimentos futuros considerarao permitir-se ser a propria gramatica a servir de base de dados metalinguistica para fornecer exemplos de perguntas com certas palavras desejadas, e para explicitar as estruturas de frase que admite.

DESENVOLVIMENTOS LINGUISTICOS FUTUROS

Os desenvolvimentos da COMPETENCIA LINGUISTICA do ORBI dizem respeito a :

- tipificacao implicita de entidades, atraves da mencao apenas de uma ocorrencia especifica : 'a permeabilidade do substracto' em vez de 'o descriptor permeabilidade do substracto' .
- conjuncao e dijuncao de designacoes : 'os factores 9 e 12' , 'a aptidao industria ou recreio' .
- conjuncao do sintagma nominal sujeito : 'A aptidao 5 e o factor 12 do ponto 21,66 tem o mesmo valor ?' .
- construcoes com 'o maior' e 'o menor' .
- construcoes com 'so' ' e 'apenas' .
- pronominalizacao : na mesma frase e entre frases.
- alargamento dos casos de extraposicao e de elisao.

No que toca 'a COMPETENCIA METALINGUISTICA esperamos providenciar :

- definicao de sinonimos e abreviaturas :
- > 'erosao' e' sinonimo de 'riscos de erosao' .
- > 'RMFA' e' uma abreviatura de 'resistencia 'a poluicao dos mantos aquiferos'.
- exemplificacao de frases :
- > da'-me 2 perguntas sim/nao com o verbo 'valer' e com o substantivo 'aptidao' como sujeito !
- diagnosticó explicito sobre analises incompletas, exibindo as maiores subestruturas analisaveis, gracias a um sistema de 'debugging' especial 'a disposicao do utilizador.

CONCLUSOES

=====

Encontramos no Prolog uma excelente linguagem para a implementacao de sistemas periciais ("expert systems"). Varias razoes podem ser aduzidas :

--- Integra expeditamente num mesmo simples e poderoso formalismo todos os desejaveis componentes dos sistemas periciais : linguagem natural, base de conhecimentos [OLI], facilidades explicativas, base de dados relacional, meta-conhecimento e interpretadores para control especializado [PER-3].

--- A sua compaticidade de expressao, juntamente com uma boa implementacao, permite a um programa com a complexidade de um sistema pericial ser realizado numa maquina pequena para uso efectivo.

--- A semantica dual, declarativa e operacional, das clausulas da base de conhecimentos facilita o desenvolvimento do metaconhecimento.

AGRADECIMENTOS

Agradecemos 'a Junta Nacional de Investigacao Cientifica pelo seu suporte financeiro adicional ao contracto com a Direccao de Estudos do Ambiente da Secretaria de Estado do Ambiente.

Os nossos agradecimentos vao tambem para A. Porto e M. Filgueiras pela sua ajuda na melhoria do sistema Prolog.

De maneira alguma agradecemos ao Instituto Nacional de Investigacao Cientifica o nao ter suportado este projecto.

REFERENCIAS

-
- [CLO] Clocksin, W. ; Mellish, C.
Programming in Prolog
Springer-Verlag, 1981

 - [DIA] Dias, V. ; Pereira, L.M.
Digitalizacao de dados territoriais
Departamento de Informatica, 1982

 - [OLI] Oliveira, E.
A representacao do conhecimento num sistema pericial visto como uma base de dados logica
2 Congresso Portugues de Informatica, 1982

 - [PER-1] Pereira, L.M. ; Porto, A.
A Prolog implementation of a large system with natural language on a small machine
Submetido 'a First International Conference on Logic Programming
Marselha 1982

 - [PER-2] Pereira, L.M. ; Porto, A.
All solutions
Logic Programming Newsletter n.2, Autumn 1981

 - [PER-3] Pereira, L.M.
Logic control with logic
Submetido 'a First International Conference on Logic Programming
Marselha 1982

- [PER-4] Pereira,L.M. ; Porto,A.
Implementacao em Prolog de um grande sistema com linguagem natural
numa pequena maquina
2 Congresso Portugues de Informatica, 1982
- [PER] Pereira,F. ; Warren,D.H.D.
Definite clause grammars for language analysis - a survey of the
formalism and its comparison with augmented transition networks
Artificial Intelligence 13, 1980, 231-278
- [PIQ] Pique,J.F. ; Sabatier,P.
An informative, adaptable and efficient natural language
consultable database system
Proceedings of ECAI-82, Orsay, 1982
- [SEA] Servicos de Estudos do Ambiente
Ordenamento biofisico do concelho de Sesimbra,
tratamento automatico.
relatorio preliminar,1981
- [WAR] Warren,D. ; Pereira,F.
An efficient easily adaptable system for interpreting
natural language queries
Dept. of Artificial Intelligence, University of Edinburgh, 1981