



Universidade do Porto
Faculdade de Engenharia

FEUP

Informações sobre Obras Literárias em Linguagem Natural

Relatório Intercalar

Inteligência Artificial

3º ano do Mestrado Integrado em Engenharia Informática e Computação

Elementos do Grupo:

José Ferreira | ei08047@fe.up.pt

Objetivo

Este trabalho está a ser realizado no âmbito da disciplina de inteligência artificial leccionada no 2º semestre do 3º ano do mestrado integrado em engenharia informática e computação da faculdade de engenharia da universidade do porto.

O objectivo deste trabalho é a implementação de um programa que permita a obtenção de dados relativos a obras literárias usando linguagem natural.

Descrição

Especificação

O programa a desenvolver irá conter uma base de conhecimento relativa a obras literárias contendo informações relativas a escritores, livros escritos por estes e prémios atribuídos.

O acesso a estas informações será feito por meio de perguntas e afirmações escritas na língua portuguesa pelo utilizador.

O programa tem de ter capacidade de fazer uma validação sintáctica e semântica dessas mesmas frases e perceber o que lhe é pedido; seja a averiguar a validade de uma afirmação, seja responder correctamente a uma pergunta com um facto contido na sua base de conhecimento.

Tem também de ter a capacidade de se enquadrar num contexto isto é, se posteriormente a dar uma resposta, o utilizador voltar a perguntar algo relacionado com a pergunta anterior usando o prefixo “e..” o programa deve perceber o enquadramento e dar a resposta correcta.

Deve estar preparado para o uso de adjectivos.

Metodologia

- **Análise sintáctica**

A primeira parte do trabalho refere-se à análise sintáctica que o programa deve ser capaz de fazer o que leva à criação de um conjunto de regras que definam uma gramática.

A gramática será definida à custa de uma DCG, Definite clause grammar, que permite estabelecer um conjunto de regras recursivas (estrutura de sintagmas) que permitem gerar ou validar todas as frases possíveis de escrever com um léxico.

- **Levantamento lexical**

O léxico contém o vocabulário, uma lista de palavras da linguagem e um levantamento lexical no contexto do nosso problema é fundamental para testar/gerar frases de acordo com a gramática.

Associadas a cada palavra estão características morfológicas, sintácticas, semânticas e até para outro tipo de aplicações características fonéticas.

A estas características dá-se o nome de features e este conceito é muito importante pois permite-nos usando uma gramática relativamente simples conferir um grau de precisão muito maior à nossa linguagem sem ter de sobre carregar a gramática definida com muitas regras.

- **Análise semântica**

A análise semântica irá acompanhar de perto as regras fundamentais da concordância da gramática de língua portuguesa. São elas: concordância verbo/sujeito, predicativo de sujeito/sujeito, predicativo de sujeito/verbo, atributo/substantivo, aposto/substantivo e predicativo de complemento directo /complemento directo. Sempre que possível usando as features do léxico e evitando adicionar regras à gramática.

• Base de conhecimento

Como foi referido anteriormente a base de conhecimento refere-se a informações sobre escritores, livros e prémios literários.Exemplo :

```
% livro(Nome, autor,publisher,year)
```

```
livro(['Dune'],['Frank' , 'Herbert'], ['Chilton', 'Books'], '1965').
```

```
% escritor(Nome, pais de origem)
```

```
escritor(['Frank' , 'Herbert'] , 'America').
```

```
% premio(Nome do premio,escritor,ano)
```

```
premio('hugo' , ['Frank' , 'Herbert'] , '1966').
```

• Procura de solução

A procura de solução é o culminar deste trabalho, nesta fase terá de se garantir que a query gerada sobre a Base de conhecimento da aplicação coincide na totalidade com a frase em linguagem natural dada como input. Talvez gerar uma query a partir da frase “Frank Herbert escreveu Dune.” seja relativamente simples pois o que temos equivale a Sujeito Acção Objecto. Muito provavelmente terá de ser feito um esforço maior caso nos estejamos a referir a varios sujeitos, acções ou objectos.

Trabalho Efectuado

O processo de escrita deste programa começou pela definição de uma gramática declarativa usando a notação de listas de diferença fornecida pela linguagem prolog.

De modo a testar/gerar a gramática passei para a elaboração de um pequeno léxico adequado ao tema e reformulei a gramática de modo a mostrar a parse tree gerada e conseguindo validar sintacticamente um numero aceitavel de frases comecei a explorar a analise semantica.

Até ao momento apenas de me foquei na concordância de genero e cardinalidade(concordância atributo/substantivo). Seguidamente fiz um esboço da base de conhecimento , preparei uma bateria de testes a efectuar e preparei uma interface com o utilizador.

Resumidamente este é o meu trabalho até á data.

Resultados esperados e forma de avaliação

s(T,[o, escritor, escreveu, o,livro],[]).
algo de errado se passa se não passar neste teste.

s(T,[o,livro, escreveu,o, escritor],[]). teste de concordância sujeito/verbo e verbo/c.d.
n.i.apenas pessoas escrevem livros

s(T,[o,livro, escreveu,o, escritora],[]). teste de concordância de genero
teste já ultrapassado

s(T,[o,escritor, escreveu],[]). teste de verbos transitivos
n.i.

s(T,[o, escritor, escreveu, os,livros],[]).

s(T,[o, escritor, escreveu, o,livros],[]).

n.i. teste de concordância de numero

s(T,[o, livro,foi,escrito,por,o,escritor],[]). teste de agente da passiva
n.i.

s(T,[o,escritor, frank,herbert,escreveu,o,livro,dune],[]). o aposto
implementado com adição de regras na gramatica, em alteração

s(T,[o,escritor, frank,herbert,escreveu,o,livro,dune,em,1966],[]).teste de complemento circunstancial
n.i.

Uma vez ultrapassados estes testes será altura de incorporar na gramatica as frases interrogativas.

Conclusões

Da minha experiência a programar em Prolog poucas linhas de código fazem muito se estiverem devidamente pensadas e afinadas e, também em interfaces de processamento de linguagem natural essa máxima se parece aplicar.

Recursos

Software: SWI-Prolog

Bibliografia:

Prolog and Natural-Language Analysis

<http://www.mtome.com/Publications/PNLA/prolog-digital.pdf>

Natural Language Processing in Prolog

http://www.csupomona.edu/~jrfisher/www/prolog_tutorial/

Natural Language Processing Techniques in Prolog

<http://cs.union.edu/~striegnk/courses/nlp-with-prolog/html/index.html>

Learn Prolog Now!

<http://www.learnprolognow.org/lpnpag.php?pageid=top>