

Scripting no Processamento de Linguagens Naturais
(4º ano de Curso de MiEI)
Trabalho Prático 2
Relatório de Desenvolvimento

Catarina Cardoso
(a75037)

Paulo Guedes
(a74411)

Pedro Cunha
(a73958)

02/02/2018

Conteúdo

1	Introdução	2
2	Decisões e Implementação	3

Capítulo 1

Introdução

Descrição

No âmbito da unidade curricular de Scripting no Processamento de Linguagens Naturais, foi proposta a resolução de um trabalho prático usando a linguagem Perl capaz criar um bot. Foram tidas em conta as funcionalidades e metodologias de implementação propostas pelo docente.

Estrutura do Relatório

Após o capítulo introdutório, segue-se o capítulo 2 onde se expõe detalhadamente as decisões de implementação tomadas aquando a realização do projeto.

Capítulo 2

Decisões e Implementação

A primeira fase do projeto passou pela utilização de uma ferramenta já anteriormente utilizada nas aulas da unidade curricular, o **freeling**. Através do uso do *system command* **analyze**, foi possível executar o processo de **tokenizing** e **lemmatization** do *input*, técnicas essenciais para o bom funcionamento do bot (de modo a facilitar a comparação entre o *input* feito pelo utilizador e as regras).

A segunda etapa do trabalho passou pela criação de uma lista de **regras/respostas** capaz de servir como primeira base de apoio para a **interação** com o utilizador. Assim, a frase enviada ao bot será comparada com todas as regras. Para esta comparação, foi implementado um método em que todas as frases teriam um **"valor" de comparação**. Este valor é obtido através da comparação de todas as **palavras** da frase enviada ao bot com todas as **palavras** da regra. Caso encontre uma palavra igual, é adicionado um valor. Caso as palavras estejam na mesma **ordem** que na regra (com margem de duas palavras), é adicionado mais um valor. É apresentada a resposta da regra com **maior** valor de comparação.

Posteriormente, procedeu-se à implementação de várias **funcionalidades** capazes fornecer informação relativa à **temperatura** e **locais a visitar** de uma determinada **cidade/país**. Para além disso, ainda foi possível implementar funcionalidades capaz de fornecer **informações genéricas** e **notícias** relativas a vários tópicos (figuras públicas, países, cidades, etc). O processo de implementação das funcionalidades descritas anteriormente é bastante simples, com a obtenção da página **html** das várias fontes de informação(**google** e **wikipédia**) e proceder ao seu **parsing**, capaz de extrair o conteúdo pretendido. No que diz respeito ao tratamento do **standart input**, são procuradas determinadas palavras-chave (visitar, conhecer, notícias, etc), salientando-se que o tópico sobre o qual se pretende obter a informação deve sempre ter capitalização da primeira letra de cada um dos seus componentes. Estas funcionalidades têm prioridade relativamente à lista de regras descrita acima. Finalmente, caso não exista qualquer palavra-chave e um grau de coincidência menor que **30%** com a lista de regras, a estratégia passa por selecionar aleatoriamente um provérbio (de uma lista de mais de 1000) e apresentá-lo ao utilizador.