

Miniprojecto 1 de Língua Natural
MEIC-Tagus
2015/2016

1 Objectivos gerais

Criar a baseline para a tarefa de ajudar um agente a escolher a melhor resposta a uma pergunta, tendo em conta a sua base de conhecimento.

2 Dados de entrada

O aluno tem à sua disposição 2 ficheiros:

a) O primeiro ficheiro, PerguntasPosSistema.txt, tem os resultados de um sistema, que dada uma pergunta de um utilizador (User Input), devolve 20 pares do tipo (Trigger, Answer), em que Trigger (T) é uma frase que o sistema considerou semelhante à do utilizador e Answer (A) uma possível resposta ao Trigger. Exemplo de parte deste ficheiro (ignorem as medidas – scores – apresentadas e as mensagens relativas às normalizações):

```
User Input: A tua familia é numerosa?
Normalized question: a tua familia e numerosa
Retrieving Lucene results...
Retrieving QA's from database...
Scoring the QA's...
1
T - A tua família?
A - Não é isso...
M1 score - 0.9021934127035144
M2 score - 0.0
M3 score - 0.6299999999999999
M4 score - 0.873
...
20
T - É a tua família?
A - Olha, eu estou apenas a começar a minha vida, de volta juntos.
M1 score - 0.4498151179067693
M2 score - 0.0
M3 score - 0.72000000000000001
M4 score - 0.0

Final Score - 0.5309445493173137

Best answer score: 0.7044580433770944
Answer: Não é isso...
```

b) O segundo ficheiro, AnotadoAll.txt, tem as perguntas do utilizador (User Input) directamente associadas às 20 respostas (A), anotadas manualmente como sendo plausíveis (y), não plausíveis (n) ou talvez plausíveis num dado contexto (m). Exemplo de uma parte do ficheiro:

```

User Input - A tua familia é numerosa?
A - Não... : y
A - E o teu nome de família. : n
A - Olha, eu estou apenas a começar a minha vida, de volta juntos. :
n
A - É. : y
...
A - Não passo a ser responsável por ela. : n

```

De notar que estes dados são reais, pelo que têm erros, anotações que poderão não fazer sentido, frases em falta, etc. É assim a vida e há que sobreviver da melhor maneira com o que se tem.

3 Tarefas detalhadas

As tarefas a realizar são as seguintes:

- a) Implementar, em Python, um programa que, dada uma frase que deverá ser um dos User Inputs do primeiro ficheiro (caso não seja, o sistema devolve “Frase incorrecta”), tenta encontrar no conjunto dos triggers associados a essa frase, os que sejam **iguais** a esse User Input, a menos de um lowercasing, pontuação e acentuação. Seguidamente, analisa todas as respostas associadas a esses triggers. Devolve a mais frequente. Em caso de empate escolhe uma qualquer. Se nenhum trigger for igual ao User Input, o sistema deverá devolver a string “Não sei responder” (esta será a parte a desenvolver no segundo miniprojecto).

Nome da função: sss

Parâmetro 1: ficheiro PerguntasPosSistema.txt

Parâmetro 2: pergunta

Resultado: resposta escolhida pelo sistema ou “Frase incorrecta” ou “Não sei responder”.

- b) Para avaliar os resultados obtidos, implementar uma função, igualmente em Python, que veja se a resposta escolhida (este módulo deverá chamar a função anterior) foi avaliada como plausível, não plausível ou talvez, de acordo com o segundo ficheiro dado. Cada resposta plausível é considerada correcta; cada resposta não plausível ou talvez plausível deve ser considerada incorrecta. Com base nestes valores devem calcular a accuracy da vossa baseline.

Nome da função: myAvalia

Parâmetro 1: ficheiro AnotaAll.txt

Parâmetro 2: ficheiro com perguntas

Resultado: accuracy do sistema.

4 Entrega

4.1 Quando?

Este mini-projecto deverá ser entregue até às 23h59 horas do dia 16 de Outubro de 2015 (não serão aceites trabalhos a partir dessa data).

4.2 Onde?

As entregas deverão ocorrer via Fénix, num zip (e não rar) constando o número do grupo no nome do ficheiro entregue (será descontado 1 valor, caso este requisito não seja satisfeito).

4.3 O quê?

Material a entregar:

- Código (por favor, nada de paths locais)
- Readme a explicar EXACTAMENTE como correr as duas funções
- Relatório (com um máximo de 3 páginas + potencial bibliografia) que consistirá nas seguintes secções:
 - Capa: identificação clara do grupo e dos autores (não façam uma capa à parte);
 - Introdução: breve introdução ao problema que vão tratar;
 - Arquitectura proposta: pode ser ilustrada com uma figura que explicita que funções/módulos (ex: normalizador) são chamados e onde, qual o fluxo de dados, etc.
 - Experimental Setup: breve descrição dos corpora usados, explicação do que usaram para desenvolvimento, teste, etc., definição da medida de avaliação;
 - Resultados: apresentação da baseline e comentários aos resultados obtidos;
 - Conclusões e eventual bibliografia.

Sugestão: não serão avaliados por isto, mas recomendo que façam o relatório em Latex, em especial se nunca o usaram. É maravilhoso.

5 Avaliação do Projecto

- Qualidade do relatório (organização de acordo com o pedido, clareza do texto, correcção do português, etc.): 5 valores;
- Qualidade da proposta em si, em especial respeitando a metodologia estudada: 10 valores.
- Qualidade do readme (se realmente conseguir pôr tudo a correr sem problemas no meu computador) e do código (em relação a este último só vou ver se está bem comentado e com nomes compreensíveis): 5 valores