

Visão do futuro do HAREM

Diana Santos
Linguatca
Encontro HAREM, 15 de Julho de 2006

Estrutura

- Visão de conjunto
 - gregos e troianos
 - contexto1 e contexto2
- Questões por determinar
 - uso dos recursos: a questão do género, das abordagens, dos recursos
 - a questão de gregos e troianos: conseguimos agradar a todos?
- Futuro
 - Mais do mesmo, com algumas melhorias
 - Novos desafios (um harém de HAREMs?)

REM, reconhecimento de entidades mencionadas

- Identificação e classificação de nomes próprios (e expressões numéricas) em texto -- em português

Eça de Queirós nasceu na Póvoa de Varzim em 1845, e faleceu 1900, em Paris. Estudou na Universidade de Coimbra.



Eça de Queirós nasceu na Póvoa de Varzim em 1845, e faleceu 1900, em Paris. Estudou na **Universidade de Coimbra**.

Simples, né?

- A Mona Lisa foi roubada do Louvre e encontrada ao pé da Torre Eiffel, junto com Picasso e Braques, no cais de Orsay. O Libération já anunciou a detenção do "ruivo de Lyon", que vai voltar ao célebre D76 da prisão de Vincennes. Belém emitiu um comunicado de solidariedade oferecendo o SMI, depois de Tony Blair ter colocado a Scotland Yard à disposição da França.
- Metonímia, metáfora, comparação, falta de precisão...
- Este "artigo" é sobre quê?
 - roubo de obras de arte em França
 - colaboração entre polícias europeias

O que é?

- É uma espécie de primeira passagem num texto para ter ideia do seu conteúdo...
- Semântica "light"
- Um pré-processamento dos textos com informação que os "agarra" ao mundo
- Uma ajuda a toda e qualquer tarefa de PLN...

Para que serve? Aplicações em que dá jeito:

- IR: indexar e buscar, visualizar
- TA: traduzir como deve ser
 - Rio de Janeiro
 - Prestes
- Análise sintáctica
 - ela foi a Lisboa de TGV
 - lá foi a Maria de Adidas para a festa
- Síntese e reconhecimento de fala
 - PUCi, TAP, IPO, Universidade de Aveiro
- Sumarização

O nosso problema: pouca massa crítica

- Conjunto de grupos com ambições e problemas diferentes
 - detecção de localizações na Web
 - análise sintáctica de cobertura larga
 - anotação de entrevistas no Museu da Pessoa
- Generalizámos a tarefa de forma a abranger todas as tarefas, mas reduzimos a sua dificuldade (sem co-referência, sem interpretação final)
- Aumentámos a quantidade de mini-problemas (vários tipos de texto)
 - identificação
 - vários sentidos (objectivos) possíveis

As categorias contempladas pelo HAREM

- PESSOA
- ORGANIZACAO
- LOCAL
- TEMPO
- OBRA
- ABSTRACCAO
- ACONTECIMENTO
- COISA
- QUANTIDADE
- VARIADO

Os tipos: subcategorias com motivação linguística

- PESSOA
 - INDIVIDUAL
 - CARGO
 - MEMBRO
- OBRA
 - (PRODUTO)
 - ARTE
- ABSTRACCAO
 - DISCIPLINA
 - ESTADO
 - ESCOLA
 - OBRA
 - ESTADO
- GRUPOIND
- GRUPOCARGO
- GRUPOMEMBRO
- PUBLICACAO
- REPRODUZIDA
- MARCA
- PLANO
- IDEIA
- NOME

Os tipos: subcategorias com motivação linguística

- LOCAL
 - GEOGRAFICO
 - ADMINISTRATIVO
 - ALARGADO
- ORGANIZACAO
 - INSTITUICAO
 - EMPRESA
- COISA
 - OBJECTO
 - CLASSE
- ACONTECIMENTO
 - EFEMERIDE
 - ORGANIZADO
- VIRTUAL
- CORREIO
- ADMINISTRACAO
- SUB
- SUBSTANCIA
- EVENTO

Os tipos: subcategorias com motivação linguística

- TEMPO
 - DATA
 - HORA
- VALOR
 - MOEDA
 - CLASSIFICACAO
- VARIADO
 - OUTRO
- PERIODO
- CICLICO
- QUANTIDADE

A noção de contexto

- Ficar-se pelos tipos: um país é sempre um país
- Ir ao significado no contexto: conforme o significado da própria frase, a categoria varia... de duas maneiras
 - previsível e sistemática (qualia): faz parte do significado de um país poder ser visto de várias maneiras
 - imprevisível (tb afecta o de cima)
- O segundo caso é mais difícil... mas mais útil

Avaliação conjunta: o ideal

- Concordar numa tarefa e discutir os pormenores em conjunto
- Criar um cenário de avaliação
 - medidas
 - recursos
 - procedimento
- Comparar o desempenho dos vários sistemas, obtendo o estado da arte
- Tornar públicos os recursos, programas e as saídas dos sistemas para
 - validação externa
 - pesquisa tanto sobre a tarefa como sobre a metodologia de avaliação
 - organização de avaliações conjuntas futuras
 - treino de novos actores/participantes

Avaliação conjunta: na prática...

- As pessoas confiam na organização
- Não testam os programas
- Não analisam as directivas
- Não seguem as normas
- Quando instadas a sugerir medidas, até recusam...

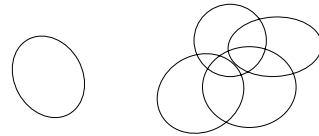
- Na prática, a organização tem muito mais responsabilidade do que na teoria

Vantagens de uma avaliação conjunta

- Acordo sobre detalhes que geralmente tornam as medidas individuais de avaliação incomensuráveis
- Aumentar a visibilidade de uma tarefa particular, dos seus problemas e soluções: construção de comunidade
 - vários novos sistemas nasceram com o HAREM
- Produzir muita documentação que de outra forma nunca seria produzida
 - directivas do HAREM: discussão de problemas morfológicos; discussão sobre questões de RAP no CLEF
- Pode dar origem a "baselines" (mínimos garantidos) e recursos (sistemas, almanaques) para trabalho futuro

Desvantagens de uma avaliação conjunta

- Torcer demasiado o que se quer avaliar de forma a agradar a todos
- Preocupar-se com um conjunto de detalhes irrelevantes aos outros
- Depender do conjunto de sistemas



Problemas do I HAREM

- Inconsistência em relação ao tratamento das categorias e dos tipos
 - Entropia dos tipos considerada, não das categorias
 - Categorias abertas (VARIADO), tipos fechados (sem OUTRO)
- Inconsistência em relação à relação entre a identificação e a classificação semântica
 - para algumas categorias, a identificação depende da classificação (VALOR, COISA), donde a identificação conceptualmente podia ser medida depois
- A questão do contexto não foi levada aos seus últimos limites
 - A dona da barraquinha das faturas da Feira Popular é uma PESSOA! independentemente da Feira Popular ser um LOCAL
 - A prisão de Vincennes pode ser um LOCAL em Vincennes (outro LOCAL)

Coisas que podiam ser melhoradas

- Premiar subpartes que fazem sentido; penalizar o inverso
 - Reunião da Comissão Política do PC do B de Minas Gerais Outubro de 1999
- Premiar a identificação de grandes EMs ou de EMs teoricamente ambíguas Santos e Silva
- Pesar os sucessos pelo seu peso relativo
- Tratar da entropia a nível das várias categorias
- Usar uma matriz de proximidade ou de risco

- Resultado: programas de avaliação ainda mais complicados, mas mais justos

Futuro: duas opções

- prosseguir afinando o armamento
 - possibilidade de usar aprendizagem automática
 - possibilidade de crítica e refinamento das directivas e CDs
 - possibilidade de ir mais a fundo em alguns casos
 - possibilidade de fazer validação estatística
- mudar de rumo e experimentar outras coisas
 - uma tarefa muito mais concreta: um tipo de textos, um tipo de problema
 - uma tarefa que use REM: RAP, indexação, ...
- fazer ambas...
 - neste caso, precisamos decididamente de mais apoio da comunidade...