Evaluation of Cross-Language Question Answering Systems

Bernardo Magnini

ITC-irst, Centro per la Ricerca Scientifica e Tecnologica, Trento, ITALY

magnini@itc.it
Question Answering

- QA: find the answer to an open domain question in a large collection of documents
  - IN: questions (instead of keyword-based queries)
  - OUT: answers (instead of documents)
- QA track at TREC
  - Main task, context task
- Scientific Community
  - this year workshops at LREC, COLING, QA session at ACL
  - AQUAINST program in USA
- QA as an applicative scenario
QA Architecture

Question Processing

- Tokenization & POS Tagging
- Question Parsing
- Word Sense Disambiguation
- Answer Type
- Keywords Expansion

Search

- Search Engine
- Query Composition

Answer Extraction

- Answer Ranking
- Answer Identification
- Named Entities Recognition
- Passage Filtering
IN: Quanto è alto il Mont Ventoux?

*(How tall is Mont Ventoux?)*

"Le Mont Ventoux, impérial avec ses 1909 mètres et sa tour blanche telle un étendard, règne de toutes …"

OUT: 1909 metri
Motivations

- Answers may be found in languages different from the language of the question
- Interest in QA systems for languages other than English
- Force the QA community to design real multilingual systems
- Check/improve the portability of the technologies implemented in current English QA systems
Task proposal

• Evaluation for:
  – Monolingual QA systems for languages other than English
  – Cross-Language Question Answering systems
• Task Organization:
  – effort required (no funding)
  – Coordination (one group for language)
• Issues: Question, Corpora, Translations, Evaluation
• Task schedule
CL - QA at CLEF-2003

- Adopt the same rules used at TREC-2002 QA
  - Factoid questions (i.e. no definition questions)
  - Exact answers + document id
- Use the CLEF corpora
- Return the answer in the language of the text collection in which it has been found (i.e. no translation of the answer)
- Consider CLEF-2003 as an initial step toward a more complex task to be organized at CLEF-2004
Tasks at CLEF - 2003

Monolingual Italian
Monolingual Spanish

question

question

Italian

Spanish

answer

answer

ITC-irst
UNED
Tasks at CLEF - 2003

- Monolingual Italian
- Monolingual Spanish
- Bilingual Italian-English
- Bilingual Spanish-English

- ITC-irst
- UNED
- U Texas at Dallas

question
question
question

Italian
Spanish
English

answer
answer
answer
Tasks at CLEF - 2003

Monolingual Italian
Monolingual Spanish
Bilingual Italian-English
Bilingual Spanish-English

question
question
question
question

Italian
Spanish
English

answer
answer
answer

ITC-irst
UNED
U Texas at Dallas

ITC-irst
Questions

- Italian L1
- Spanish L2
- English L3

100 questions
?? Answer patterns

Require: CLEF support for corpora, NIST support for methodology
native speakers for each language
Evaluation

- Manual: exact answers are relatively easy to be judged
- Semi-automatic, using answer patterns
  IF a system-answer satisfies an answer pattern
  THEN the answer is judged correct
  ELSE manual check is required

Require: NIST support for methodology and software
native speakers for each language
Schedule (preliminary)

- November 15: final task proposal
- February 1: corpora distributed to participants
- May 5: Questions to participants
- May 12: Results back to organizer
- July 1: Assessments to participants

- August 21/22: CLEF workshop