Esfinge at CLEF 2007: First Steps in a Multiple Question and Multiple Answer Approach

Luís Miguel Cabral, Luís Fernando Costa, Diana Santos

{Luis.M.Cabral,Luis.Costa,Diana.Santos}@sintef.no

Linguateca / SINTEF ICT

PB 124, Blindern NO-0314 Oslo, Norway

Esfinge on the Web: http://www.linguateca.pt/Esfinge/

Esfinge Overview

➤ General domain question answering system, participation in QA@CLEF since 2004.

➤ The starting point was the redundancy-based architecture described in [Brill 2003].

➤ Exploiting the redundancy in the Web where Portuguese is one of the most used languages.

➤ For 2007: use of the M,N-O,P model for QA (multiple questions, multiple answers), anaphoric resolution, and Wikipedia.

Answer Selection

➤ Merging different runs by analysing the result (3 to 6 runs)
  ➤Number of times an answer occurs in the results
  ➤The relevance of the support snippet regarding the question text

➤Automatic Selection achieves 55% - 69% of the correct answers in the best possible selection

➤Combination of several runs did not produce better results than some of the individual runs combined and therefore needs further attention

Experiments

➤Comparing runs with two different regular expression pattern generations
  ➤Esfinge regular expression patterns
  ➤PALAVRAS generated patterns

➤Measuring the influence of the Web
  ➤Web + Newspapers + Wiki
  ➤Newspapers + Wiki

➤Measuring the import of using Wikipedia
  ➤Web + Newspapers + Wiki
  ➤Web + Newspapers

➤Using phrase patterns based on the syntactical analysis by PALAVRAS seems to slightly increase the performance

➤Addition of Wikipedia brought 37.5% (9/24) increase in performance

➤Use of the Web brought 32% (8/25) increase in performance

Future Work

➤ Extend the question reformulation module
  ➤ create expressions that include the answer expected position
  ➤ use ontologies
  ➤ use further syntactic analysis
➤ Create high precision filters for closed-class types of answers
➤ Improve choice among several answers
➤ Weigh redundancy by lexical counts in each document collection
➤ Use link structure of Wikipedia
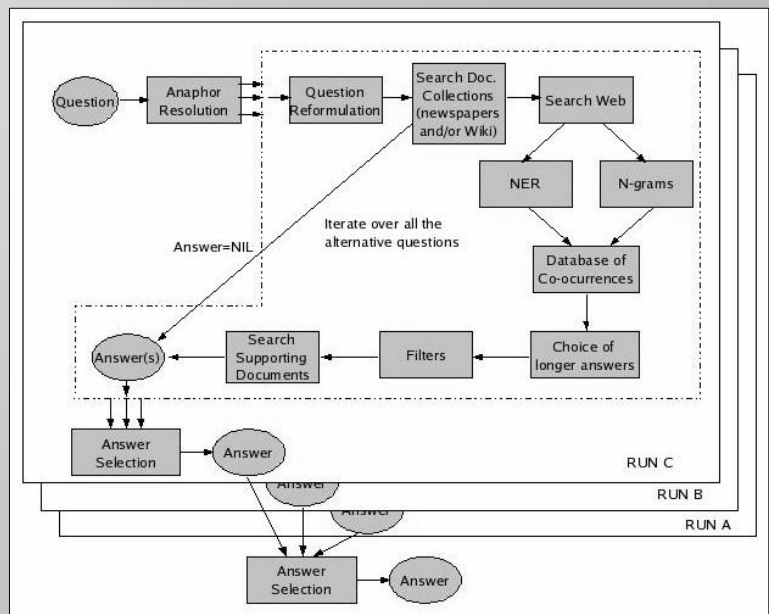➤ Do an empirical study for weighting alternative questions

➤Bick, Eckhard. The Parsing System "Palavras": Automatic Grammatical Analysis of Portuguese in a Contraint Grammar Framework. PhD thesis, Aarhus University, 2000

➤Brill, Eric. 'Processing Natural Language without Natural Language Processing', in A. Gelbukh (ed.), CICLing 2003, LNCS 2588, Springer-Verlag Berlin Heidelberg, 2003, pp. 360-9.

Anaphor Resolution

➤ Coping with five kinds of anaphoric expressions (see paper)
➤ Invoking PALAVRAS, a broad-coverage parser for Portuguese [Bick 2000]
➤ Dealing with sentence arguments (object, subject, etc.) as candidates
➤ Shallow reference resolution algorithm: 71% accuracy in general (over 122 questions)
➤ Producing several questions from the original one
➤ As a by-product, producing "syntactically-based patterns" as an alternative to regular expression reformulation of Esfinge
➤ Example (3 questions out of 1):
  ➤ *Contra quem é que Steffi Graaf não jogou nas semi-finais de Roland Garros?*
  ➤ *Contra quem é que o pai de Steffi Graaf não jogou nas semi-finais de Roland Garros?*
  ➤ *Contra quem é que ela não jogou nas semi-finais de Roland Garros?*

This module had relatively good results (although it had little influence in the final performance given the few cases it had to resolve)

Results

#	Description	Right Answers (all questions)			Unsupported Answers	Inexact Answers -	Inexact Answers +	Right Answers (1st questions in 150 topics)			Total NIL
		i)	ii)	iii)				i)	ii)	iii)	
1	Esfi071PTPT Official	15	0	8	2	4	0	12	0	4	143
2	Esfi072PTPT Official	11	0	10	3	2	0	6	0	5	181
3	Web + News + Wiki	33	2	6	2	7	0	27	2	5	74
4	News + Wiki	25	0	6	1	3	0	21	0	5	74
5	Web + news	24	1	8	4	6	0	19	1	6	107
6	Automatic selection 3-5	31	1	6	3	7	0	27	1	5	74
7	Best Selection 3-5	46	2	--	4	8	0	38	2	--	--
8	Web + News + Wiki	35	3	5	1	6	1	28	3	3	67
9	News + Wiki	25	2	7	3	7	0	19	2	4	98
10	Web + News	28	0	5	1	3	1	21	0	3	67
11	Automatic Selection 8-10	34	2	5	2	6	1	27	2	3	68
12	Best Selection 8-10	49	3	--	2	8	1	38	3	--	--
13	Automatic Selection 3-5, 8-10	34	1	6	2	6	1	30	1	4	73
14	Best Selection 3-5, 8-10	61	3	--	5	10	1	48	3	--	--
15	Best Run in 2006	50	--	--	3	7	2	--	--	--	--
16	CLEF2006A	57	--	--	6	10	2	--	--	--	--
17	CLEF2006B	56	--	--	4	7	1	--	--	--	--

(rows 8–14 grouped under label: PALAVRAS)

i) Right answers (including NIL)   ii)Partial right answers on lists   iii)NIL right answers