Where is future QA?: Where is future QA?: One pilot and two main proposals
Diana Santos
Luís Costa
Linguateca, www.linguateca.pt
and SINTEF ICT
Alicante, September 22th

QoIA: collaboration among Web services

Proposal for a pilot in QA
Pilot: a different setup that should not replace the main track, but whose (positive or negative) results hopefully influence future main tracks

Rationale (generic)

Semantic Web technology is a buzzword (funding, visibility, available technolog(ies))
An evaluation contest should also promote collaboration

QA rationale

Most interesting questions have multiple answers
The path to get the answer is often more informative than the answer alone
Most (naive?) people consult the Web (i.e., they don’t care whether it was system X or system Y or person A or B)
Integration of different methodologies, different answer snippets, different answer rankings is the key to future QA technology (not monolithic systems with a proprietary database)
The size/form of the perfect answer depends on the questioner (and on the answer set)

QA@CLEF rationale

We have 30-36-40 separate systems which provide answers together with snippets that justify them
The systems do not know which kind of question they are answering
don’t need to know, either, where did the answer came from (people, or other systems)
Some systems even provide their own confidence judgement
Wide variety of approaches (and languages)

But
There does not seem to have been significant individual progress
There is very little crosslinguality
There is no cooperation

Practical setup

Some systems would have to be “refactored” as Web services
Eventually only invocable by CLEF registered participants
A basic infrastructure of communication among QA agents has to be deployed
Communication protocol
Service identification
Enough participants (both systems and groups, for more than one language) interested

Is there any interest out there?

Two additions for the main track: the Web

Add the Web to our collections!
If you can justify your answer with a Web document (cached), do it!
There is absolutely nothing that changes in the main track
Assessors still have only to look at snippets
Who cares which Web you have indexed or which major search engine API you invoke behind the scenes?
Two additions for the main track: reasoning

- Ask non-trivial questions!
- Devise relational questions, that is, that require the system (the answerer) to understand and use a particular relationship between 2 (or more) facts
- Examples:
  - Whose politician’s wife runs for the White House?
  - Who was the first African Nobel prize winner?
  - How long was X’s murderer in prison?
- Systems can always find more (factoid) answers when we thought there was only one in our collection... so a reanalysis of question types is always necessary.

Summing up

Main track

30 difficult questions next year
- with the Web (in whatever format) as an additional/alternative justification source

The QoA pilot

- a pilot exploring cooperation between different systems
- comparing answers
- validating answers
- merging different information
- in a way, the logical conclusion/extension of WiQA and AVE