Computational Linguistics: Human Computer Interaction (II)

Raffaella Bernardi

KRDB, Free University of Bozen-Bolzano

P.zza Domenicani, Room: 2.28, e-mail: bernardi@inf.unibz.it

Contents

1 .. via natural language 4
2 NLDB: Architecture 5
3 Question Answering 6
3.1 Sample of QA architecture 7
3.1.1 Question Classification: Why and How? 8
3.1.2 Question Classification: Problems 9
3.1.3 Document Retrieval in QA 10
3.1.4 Document Retrieval in QA: Approaches 11
3.1.5 Candidate Answer Extraction 12
3.1.6 Answer Re-ranking and Selection 13
3.2 Enriched QA Architecture 14
3.2.1 Other QA issues 15
3.3 Understanding in QA? 16
3.4 Challenges 17
4 Interactive Question Answering 18
4.1 Context in IQA 19
4.2 Context in Open Domain IQA 20

4.3 System-initiated Interactions 21
4.3.1 Type of system initiatives 22
4.3.2 System-initiated contributions 23
4.4 System initiative in DB IQA 24
4.5 Ontology in IQA 25
4.6 Example 26
5 Conclusions: LCT 27
5.1 You and LCT 28

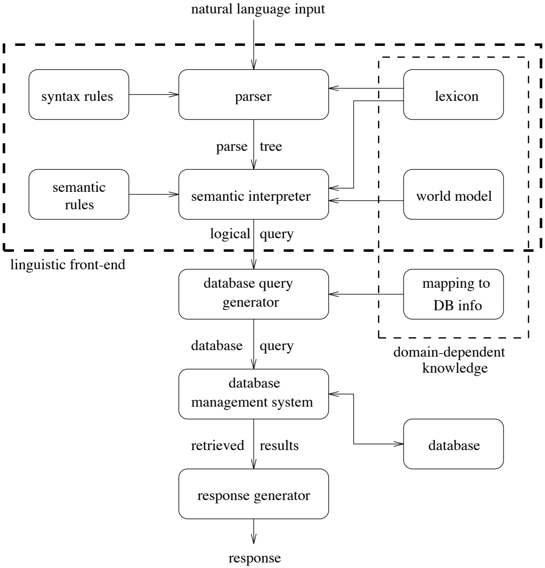
1. .. via natural language

Now, researchers are back to deal with unrestricted extended text and dialogues.

1. NLDB
2. Dialogue Systems,
3. QA
4. IQA

Slides taken from Bonnie Webber's presentation at the BIT Seminars at FUB (December 2006).

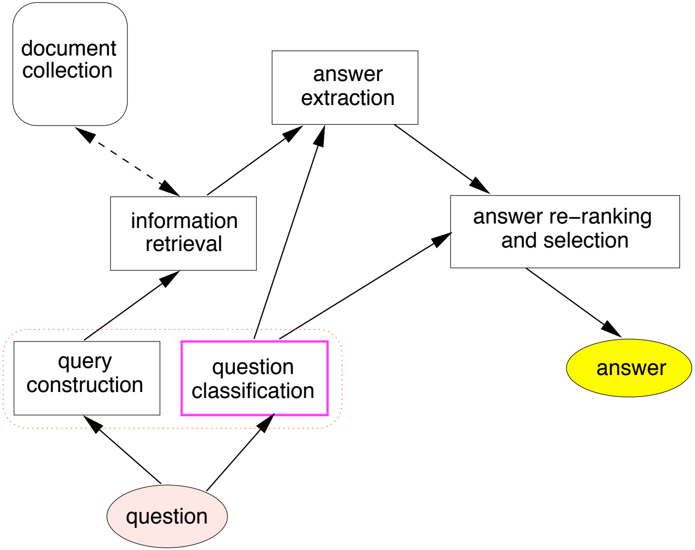
2. NLDB: Architecture



3. Question Answering

- Take as input an NL question, give as answer part of sentences extracted from documents.
- Open domain
- Might give more than one possible answer –the given answers might be ranked.

3.1. Sample of QA architecture



3.1.1. Question Classification: Why and How?

Strategies for extracting answer candidates depends in part of the type of question: factoid, list, definition/other, event etc.

Within factoid QA, classifying questions (e.g., Person, Location, Date, Quantity) can provide semantics constraints on answer candidates.

Classification can be obtained by means of manually constructed rules (if a question starts with Who or Whom: type Person).

Manually constructed rule sets (can) have large coverage and reasonable accuracy, but they are not sufficient to support fine-grained classification.

Current approaches either just consider coarse classes, or they organize QA types in a hierarchy, picking up answer requirements from the most specific class that the classifier is able to reliably identify.

3.1.2. Question Classification: Problems

Questions can be ambiguous: different senses may fit into different classes. E.g., “What is an SME”?

- A Small-to-Medium Enterprise (abbreviation)
- A company with fewer than 100 employees (definition)

Classes are not really disjoint, so even unambiguous question can fit into more than one class. Eg., “What is a bipolar disorder?”

Variability of expressing the same question, e.g.,

- What tourist attraction are there in Reims
- Where do tourist go in Rome
- What can I see in Reims
- What is worth seeing in Reims

3.1.3. Document Retrieval in QA

- QA over free text arose as an off-shoot of document retrieval
- In document retrieval, queries are interpreted as requests for documents about a topic
- QA requires answers not the documents containing them
- The answers to factoid questions are very locally in a document. Documents containing such local information are easily not tagged as relevant by standard document retrieval models.
- So document ranking by relevance may not correlate with a document's likelihood of containing an answer to a particular question
- If document retrieval fails to return documents containing answers, all subsequent processing is useless.

3.1.4. Document Retrieval in QA: Approaches

Use standard IR engine (eg., Lucene) in a way more appropriate for QA:

- Break documents into passages and perform passage retrieval
- Use more complex strategies to index the documents/passage collection (stem, PoS, dependencies)

Apply additional filters.

3.1.5. Candidate Answer Extraction

Goal: locate candidate answers within the retrieved texts (documents or passages).

Evidence used for candidate answers:

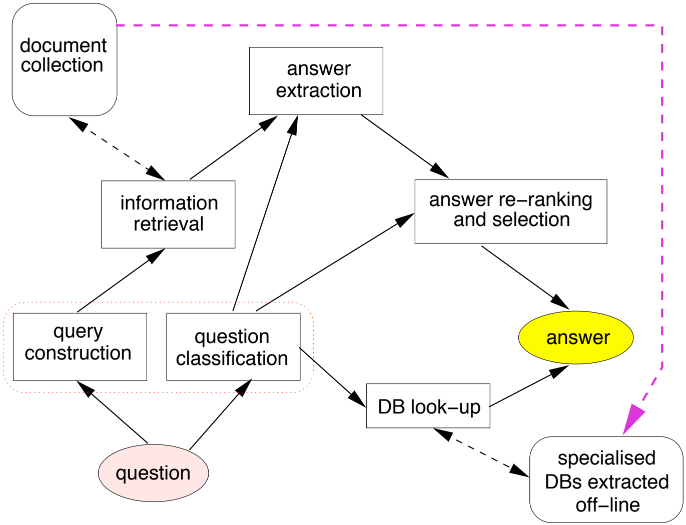
- string patterns
  - specifying string context in which answer string may be found.
  - specifying properties of the answer string
- presence of answer type linked to the question type
  - from prior indexing of the different types of named entities
  - on-the-fly named entity recognition
- patterns in light-weight logical form (e.g., arguments to particular predicates

3.1.6. Answer Re-ranking and Selection

Answer selection could reflect:

- confidence in query's ability to return a correct answer
- document ranking and query/document match
- conformance with the answer type
- weight contributed by overlapping N-grams.

3.2. Enriched QA Architecture



3.2.1. Other QA issues

- Improving extraction of answer candidates through better linguistic analysis
- Improved techniques for answer re-ranking and selection
- Answering list Qs, definition/other Qs, event Qs, Why Qs, How Qs, What's the difference Qs, Qs of opinion, etc.
- Fusing information when more than one piece of information is relevant to an answer
- Providing extended answers (summarising opinions, laying out what's known about an issue, providing justification for an answer.)
- Interactive Question Answering (different types of user and system-initiated interactions)

3.3. Understanding in QA?

In 2000 (TREC-9) semantically equivalent questions received some attention. For 54 of the questions additional variants were generated by paraphrasing the original question. All systems were sensible to the changes of the surface structure. This showed the need of further research aimed at deeper understanding of questions.

Since 2001, TREC also includes questions with no answers in the documents and questions with "list" answers. Both aspects requires some more level of understanding: systems to measure their certainty about an answer, on the one hand, and on the other they may need to synthesize the answer from multiple documents.

3.4. Challenges

This trend will be pushed further by the answer evaluation criteria the required presentation modes:

- answer must be correct, succinct, coherent and justified.
- Systems will need to handle cases when:
  - multiple answers are found,
  - when no answer is found, or
  - when contradictory answers are found.
- System will have to go beyond extraction of snippets of text to provide answer synthesis across sentences and across documents.

Providing appropriate coherent answers will be a major research area and will depend heavily on progress in text summarization.

4. Interactive Question Answering

- Extended interaction –hence the need to model the context of the interaction and established by the interaction.

  No IQA without context modelling.

- System-initiated contributions and interactions:
  - System-initiated contributions to its answer turn
  - System-initiated interactions outside its answer turn: Context modelling does not itself constitute interaction.
- User-initiated interactions, including but not limited to:
  - asking domain-related questions
  - answering clarification questions posed by the system

    Users need to do more to achieve knowledge-related goals.

4.1. Context in IQA

Context is needed to resolve context-dependent linguistic phenomena that occur in any extended text or dialogue (pronoun, anaphora, ellipsis, fragments or e.g., "other countries")

- Q1: What museum in Florence was damaged by a major bomb explosion?
- A: Uffizi Gallery
- Q2: What kind of explosives were used?

- Q1: In what country did the game of croquet originate?
- A1: Ireland
- Q2: How about bowls?

4.2. Context in Open Domain IQA

At every turn, a system should check whether the topic of the interaction has:

- stayed the same?
- switched to an item from the previous context?
- switched to something new?

4.3. System-initiated Interactions

In Open Domain QA issues related to this problem have not been addressed, in DBQA/NLDB, most of the work dates back to the 80's and early 90's. E.g,

- What information is needed to take a particular type of initiative?
- When does the system have such information?
- What leads the system to take the initiative?

4.3.1. Type of system initiatives

- Question clarification:
  - Q1: What is the largest country in Southeast Asia?
  - S: By “largest” do you mean in land area of population?

  But, how much clarification is it wise to ask for? E.g., By “country in South Asia” do you mean countries south of China and east of India?

- Question Verification:
  - User: Restaurants on Page Mill Road?
  - Systems: I found 258 restaurants on Page Mill Road. Would you like to try searching by type of cuisine?

  N.B. CO-OP (Kaplan, 1983) used a similar strategies.

4.3.2. System-initiated contributions

- Justification
  - User: What's the largest country in Southeast Asia?
  - User: Largest by land mass:
    1. Indonesia: 1,919,440 Km^2 (CIA World Factbook)
    2. ...
- Facilitate users exploration of additional hypothesis.
- Contribute to user understanding of domain by correcting misconceptions and by answering questions related to one that has failed. E.g
  - User: Does China have the largest population of Muslims in Southeast Asia?
  - System: China is not considered part of Southeast Asia. Estimates of the Muslim population of China range from 20 million to 100 million. Indonesia has the largest population of Muslims in Southeast Asia. As of 2006, approximately 88 % of its 222 million people are Muslims.

4.4. System initiative in DB IQA

System initiative facilitated through update model: Justification

- User: Can I borrow Jurafsky & Martin?
- System: No. Until it is removed from the Reference Section, it is not available for borrowing.

4.5. Ontology in IQA

Annika Flycht-Eriksson investigates the use of ontologies in IQA.

The role of the interaction is to correctly interpret user requests, hence an important role of the ontology is to:

- organize the semantics of natural language expressions and
- organize the knowledge of the world as perceived by the user.
- provide an interface to external information sources from which the system collects answers to user requests.

4.6. Example

Example of a information exchange where ontological knowledge is used to guide clarification requests and database access:

- U: What do waterfowl look like?
- S: Choose one of the following Waterfowl: Barncale Goose, Tufted Duck, Bean Goose, Whooper Swan, Gadwall
- U: Gadwall
- S: You can receive information about the size and plumage of a Gadwall. Please, specify what you are interested in.
- U: size
- S: A Gadwall has a length of 56 cm from beak tip to tail. Information about wingspan is missing for Gadwall.

Similarly, the interaction could be used to try to map the vague proprieties to ones suitable for database access.

5. Conclusions: LCT

QA lesson:

- Never forget to look at what has been done in the past!
- Work on modules
- Flow of ideas across fields
- what was not successful in the past could be successful now

Current keywords in NLP:

- systems
- hybrid approaches
- evaluation

5.1. You and LCT

Each of you could contribute to its development by bringing your own expertise.

But do not forget you would be dealing with Natural Language and you should know about it, a bit.

5.1. You and LCT

Each of you could contribute to its development by bringing your own expertise.

But do not forget you would be dealing with Natural Language and you should know about it, a bit.

This afternoon: LCT Student Day and IBM prize!

```
* Oral presentations and KRDB & IBM Best Thesis Awards Ceremony:
  Room D003 (15:30-16:30)
* Coffee Break: UniBar
* Poster presentations: In front of Room D102 (17:00-18:00)
```

Further info can be found at: http://www.inf.unibz.it/mcs/lct/phd08.php