Modeling questions and responses

Lecture 5b: Integration

Kyle Rawlins
NASSLLI 2016, 15th July, 2016

Johns Hopkins University, Department of Cognitive Science
Course structure

- Lecture 1: Introducing questions and responses.
- Lecture 2: Representing question meanings.
- Lecture 3: The architecture of a QA system.
- Lecture 4-5: The dynamics of responses.
- Lecture 5b: Integration.
Course goals

Questions in discourse: two parallel fields
Linguistics/philosophy (semantics/pragmatics)
Question Answering (QA; NLP/Computational Linguistics)

- Develop, introduce and compare core theories in both fields, from the perspective of linguistics.
- QA running in parallel to linguistic work on questions since *Baseball*, 1963. This course focused on open-domain QA.
- What could the linguistics ↔ QA interaction look like (if it existed)?
Big questions:

1. What are the commonalities in ‘semantics’ (representations of question meaning)?
2. What are the commonalities in the pragmatics of responses?
3. How could a QA system in 2050 be impacted (in principle) by linguistic theory?
4. What do QA systems do that linguistic theory doesn’t cover?
Moldovan & Surdeanu (2003) fig 4:

**Fig. 4.** A generic QA System architecture.
Responses come in two kinds:

1. Address an element on the table (relevance).
   (What is the common ground like?)
   - Accept or dispute an assertion.
   - Contribute to resolving a question.
Responses come in two kinds:

1. Address an element on the table (relevance).
   (What is the common ground like?)
   - Accept or dispute an assertion.
   - Contribute to resolving a question.

2. Manipulate the table. (What is the discourse like?)
   - Enter/exit the discourse.
   - Reject a move (question/ assertion) altogether.
   - Contribute to dispelling a QUD (deny presupposition, express ignorance, ...)

Reminder of linguistic response model (lecture 5)
Alternative semantics: directly represent (in various ways) information that would be relevant to resolving a question.
Alternative semantics: directly represent (in various ways) information that would be relevant to resolving a question.

- Closest thing we find: unstructured topic models. Distribution over content words is a proxy for distribution over relevant information.
- Clusters of words ~ what people tend to say when using those words. (Given enough text).
Alternative semantics: directly represent (in various ways) information that would be relevant to resolving a question.

- Closest thing we find: unstructured topic models. Distribution over content words is a proxy for distribution over relevant information.
- Clusters of words $\leadsto$ what people tend to say when using those words. (Given enough text).
- More sophisticated approaches (partly used in e.g. Watson): topic model over pieces of semantics representations.
- Can these match up?
Structured meanings: explicitly represent how a fragment answer should fit into the semantic structure of the question. (Property denotation)
Structured meanings: explicitly represent how a fragment answer should fit into the semantic structure of the question. (Property denotation)

- Classical query approach: questions translated as database queries.
- Closely analogous to modern structured meaning approaches (Ginzburg).
Structured meanings: explicitly represent how a fragment answer should fit into the semantic structure of the question. (Property denotation)

- Classical query approach: questions translated as database queries.
- Closely analogous to modern structured meaning approaches (Ginzburg).
- Modern descendent of query combines Information Retrieval techniques: questions as graph-structured topics. Align topics probabilistically with knowledge-base.
- Meeting point: linguistically informed semantic parsing ↔ current IR techniques?
Pragmatic comparison

We focused on QA systems with extremely limited pragmatics - basically Q-A games.

- A lot of effort to make sure that extracted fragment answers are highly relevant (as complete answers) to question. That’s it.
We focused on QA systems with extremely limited pragmatics - basically Q-A games.

- A lot of effort to make sure that extracted fragment answers are highly relevant (as complete answers) to question. That’s it.
- In contrast: linguistic responses are a rich class that goes well outside core complete answers. Not to mention, hierarchical discourse structures.
Pragmatic comparison

We focused on QA systems with extremely limited pragmatics - basically Q-A games.

- A lot of effort to make sure that extracted fragment answers are highly relevant (as complete answers) to question. That’s it.
- In contrast: linguistic responses are a rich class that goes well outside core complete answers. Not to mention, hierarchical discourse structures.
- There exist systems that try to deal with more of this. Complex Interactive Question Answering (TREC track): allow for clarification questions, objections to proposed responses, etc.
- **Major challenge**: evaluation. (Diekema et al. 2003)
Pragmatic comparison

We focused on QA systems with extremely limited pragmatics - basically Q-A games.

- A lot of effort to make sure that extracted fragment answers are highly relevant (as complete answers) to question. That’s it.
- In contrast: linguistic responses are a rich class that goes well outside core complete answers. Not to mention, hierarchical discourse structures.
- There exist systems that try to deal with more of this. Complex Interactive Question Answering (TREC track): allow for clarification questions, objections to proposed responses, etc.
- **Major challenge**: evaluation. (Diekema et al. 2003)
- (Also: closed-domain dialogue systems. Things like Siri.)
The major thing QA has that linguistics doesn’t is a plethora of methods for searching bodies of information to actually find answers to questions.
The major thing QA has that linguistics doesn’t is a plethora of methods for searching bodies of information to actually find answers to questions.

- Suggestion: when thinking about constituent questions in an alternative semantics way, this search is *inextricably tied* to inferring the alternatives that a questioner has in mind.
The major thing QA has that linguistics doesn’t is a plethora of methods for searching bodies of information to actually find answers to questions.

- Suggestion: when thinking about constituent questions in an alternative semantics way, this search is *inextricably tied* to inferring the alternatives that a questioner has in mind.
- Particular case of interest is domain restriction for constituent questions. Watson (for example) has a substantial amount of engineering for this.
The major thing QA has that linguistics doesn’t is a plethora of methods for searching bodies of information to actually find answers to questions.

- Suggestion: when thinking about constituent questions in an alternative semantics way, this search is *inextricably tied* to inferring the alternatives that a questioner has in mind.
- Particular case of interest is domain restriction for constituent questions. Watson (for example) has a substantial amount of engineering for this.
- Can we learn anything about semantic representations from successful computational semantics approaches? One potential moral is *flexibility* (counter-intuitive from linguistic perspective).
A rich theory of pragmatics and responses.

- A full QA system with human-like behavior will have to infer from the asked question to goals of the asker. Covert QUDs, domain goals, etc.
A rich theory of pragmatics and responses.

- A full QA system with human-like behavior will have to infer from the asked question to goals of the asker. Covert QUDs, domain goals, etc.
- Ideally: tailor response to these goals. For example: inference from single polar question to bigger QUD.
A rich theory of pragmatics and responses.

- A full QA system with human-like behavior will have to infer from the asked question to goals of the asker. Covert QUDs, domain goals, etc.
- Ideally: tailor response to these goals. For example: inference from single polar question to bigger QUD.
- General theory of responses and coordination between agents – can this be adopted by QA (and related) systems? Explicit modeling of the context?

Evaluations – human performance?