Modeling questions and responses

Lecture 1: an introduction to questions and responses

Kyle Rawlins
NASSLLI 2016, 11th July, 2016

Johns Hopkins University, Department of Cognitive Science
Course overview

Basics of questioning and question meaning
  Alternative semantics for questions

Question and response types
  Types of interrogatives
  Types of questions
  Responses and followups
Course overview
Course goals

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

- Introduce, develop, and compare core theories in both fields, from the perspective of linguistics.
- QA running in parallel to linguistic work on questions since *Baseball*, 1963 (see Simmons 1965).
- Focus on questions and responses, not semantics of questions.
• Basics of a range of approaches to questions & responses. Develop a theory of responses.
• Linguistic work on questions, QA research are in fundamentally the same game. (Though cross-citations typically don’t go beyond the 70s.)
  • Even QA systems’ search for supporting information for answers can’t be easily disentangled.
• Can there be productive cross-communication?
1. Basic knowledge of linguistic semantics/pragmatics.
2. Set theory / possible worlds semantics. Some occasional lambda calculus (fairly basic – need to know what a property is).
3. Basic discrete structures. (pairs, tuples, relations)
Course structure

⇒ Lecture 1: Introducing questions and responses.
Course structure

⇒ *Lecture 1*: Introducing questions and responses.
  - *Lecture 2*: Representing question meanings.
  - *Lecture 3*: The architecture of a QA system.
Course structure

⇒ Lecture 1: Introducing questions and responses.
  • Lecture 2: Representing question meanings.
  • Lecture 3: The architecture of a QA system.
  • Lecture 4: Modeling discourses with inquiry.
Course structure

⇒ Lecture 1: Introducing questions and responses.
  • Lecture 2: Representing question meanings.
  • Lecture 3: The architecture of a QA system.
  • Lecture 4: Modeling discourses with inquiry.
  • Lecture 5: Meta-conversation (focus on bias, cg-management, etc).

N.b. course content will be mostly complementary to the Beaver/Roberts/Simons/Tonhauser NASSLLI course Questions Under Discussion.
Goals for today

2. Basic idea of a QA system.
3. Overview of what kind of data a comprehensive system / analysis could cover.
Motivating examples

How would you go about responding to the following?

(1) When was Justin Trudeau born?
How would you go about responding to the following?

(1) When was Justin Trudeau born?

1. Interpret the question (+contextual inferences – do they want just the year?).
How would you go about responding to the following?

(1) When was Justin Trudeau born?

1. Interpret the question (+contextual inferences – do they want just the year?).

2. Options:
   2.1 Consult your memory.
   2.2 Ask someone else who can consult their memory.
   2.3 Look it up on the internet. (Semi-QA systems for facts are pervasive...)
Motivating examples

How would you go about responding to the following?

(1) When was Justin Trudeau born?

1. Interpret the question (+contextual inferences – do they want just the year?).

2. Options:
   2.1 Consult your memory.
   2.2 Ask someone else who can consult their memory.
   2.3 Look it up on the internet. (Semi-QA systems for facts are pervasive...)

3. Turn the information you gained into an answer/response.
More motivating examples

How would you go about answering the following?

(2) Why did Justin Trudeau win the election?
How would you go about answering the following?

(2) Why did Justin Trudeau win the election?

(3) Where should we go for lunch?
   a. Where should I go for lunch?
   b. Where did you go for lunch?

(4) Is Mamoun’s any good?
How would you go about answering the following?

(2) Why did Justin Trudeau win the election?

(3) Where should we go for lunch?
   a. Where should I go for lunch? (A domain-specific QA system: http://wtfsigte.com/)
   b. Where did you go for lunch?

(4) Is Mamoun’s any good?
Basics of questioning and question meaning
1. A question is a speech act.
1. A question is a speech act.

2. In (information-seeking) questioning, a speaker requests information from the hearer(s), as described by the question’s form.
A Searlean gloss on (information-seeking) questions

1. A question is a speech act.
2. In (information-seeking) questioning, a speaker requests information from the hearer(s), as described by the question’s form.
3. A hearer should respond to a question.
4. A hearer’s response should be cooperative:
   4.1 If possible, it should truthfully answer the question to the best of the answerer’s ability.
   4.2 It should indicate the limits of the answerer’s ability.
1. A question is a speech act.
2. In (information-seeking) questioning, a speaker requests information from the hearer(s), as described by the question’s form.
3. A hearer should respond to a question.
4. A hearer’s response should be cooperative:
   4.1 If possible, it should truthfully answer the question to the best of the answerer’s ability.
   4.2 It should indicate the limits of the answerer’s ability.
1. What does it mean for a question to be information-seeking, and what else is there?
1. What does it mean for a question to be information-seeking, and what else is there?
2. What are responses and answers?
Major questions for lectures 1-2

1. What does it mean for a question to be information-seeking, and what else is there?
2. What are responses and answers?
3. How does a human deal with incomplete ability to answer?
Major questions for lectures 1-2

1. What does it mean for a question to be **information-seeking**, and what else is there?
2. What are **responses** and **answers**?
3. How does a human deal with incomplete **ability to answer**?
4. How does the **form** of a question map to a particular request?
Groenendijk’s 1999 Interrogation game:

**Definition 1 (The Game of Interrogation)** Interrogation is a game for two players: the **interrogator** and the **witness**. The rules of the game are as follows:

A. The interrogator may only raise issues by asking the witness non-superfluous questions.

B. The witness may only make credible non-redundant statements which exclusively address the issues raised by the interrogator.

(We will come to the empirical limitations of this sort of game soon enough.)
An analysis that models a Q-A game typically does the following:

- Provide a semantics for arbitrary questions.
- Provide a Question-Answer constraint that relates question-meanings to answers/responses in a predictive way: what answers are licensed by what question-meanings in context?
A computational starting point: Factoid QA


- Two players, user and system. User asks questions, system answers.
- Each question is a fact-based question with (modally) one correct answer (*factoid question*). Systems are supplied with text that contains answers to test questions.
A computational starting point: Factoid QA


- Two players, user and system. User asks questions, system answers.
- Each question is a fact-based question with (modally) one correct answer (factoid question). Systems are supplied with text that contains answers to test questions.
- Answers are short relatively simple phrases (‘1971’).
- System returns one or more ranked answer candidates. (Sometimes with supporting documents.)
- Answer candidates are evaluated (by humans) for correctness according to some criteria.
Number: 10007
What hotel was used for a setting of the Agatha Christie novel, "And Then There Were None"?
FT924-2367
Burgh Island Hotel

Number: 10008
What year was the Magna Carta signed?
FBIS4-66792
1215
Most systems that do a TREC-type QA task break down into answers to three questions:

1. Give some question Q, what information could in principle answer Q?
   - Documents in corpus, passages in documents.

2. Given some possibilities for information that addresses Q, what information correctly answers Q?
   - What documents/passages are the most relevant?

3. Given information that correctly answers Q, how can this information be formulated as an answer?
   - How to extract a real answer from the most relevant passage?
Decomposing Question Answering

Most systems that do a TREC-type QA task break down into answers to three questions:

1. Give some question Q, what information could in principle answer Q?
   • Documents in corpus, passages in documents.

2. Given some possibilities for information that addresses Q, what information correctly answers Q?
   • What documents/passage are the most relevant?
Decomposing Question Answering

Most systems that do a TREC-type QA task break down into answers to three questions:

1. Give some question Q, what information could in principle answer Q?
   - Documents in corpus, passages in documents.

2. Given some possibilities for information that addresses Q, what information correctly answers Q?
   - What documents/passages are the most relevant?

3. Given information that correctly answers Q, how can this information be formulated as an answer?
   - How to extract a real answer from the most relevant passage?
Next: develop a basic semantics/pragmatics for Q-A games.
Hamblin’s first two postulates (Hamblin 1958)

1. An answer to a question is a statement.
2. Knowing what counts as an answer is equivalent to knowing the question.
Hamblin’s first two postulates (Hamblin 1958)

1. An answer to a question is a statement. [read: proposition.]
2. Knowing what counts as an answer is equivalent to knowing the question.

Reduction of meaning of question ⇒ meaning of answer(s).
Hamblin semantics (Hamblin 1973): the meaning of a question is a set of propositions, corresponding to answers.
Alternative semantics for questions

Hamblin semantics (Hamblin 1973): the meaning of a question is a set of propositions, corresponding to answers.

(5) \([\text{What year was Justin Trudeau born in?}] = \)

\[
\begin{align*}
\ldots, \\
\lambda w_S. JT \text{ was born in 1969 in } w, \\
\lambda w_S. JT \text{ was born in 1970 in } w, \\
\lambda w_S. JT \text{ was born in 1971 in } w, \\
\lambda w_S. JT \text{ was born in 1972 in } w, \\
\ldots
\end{align*}
\]
Hamblin semantics has been incredibly influential and there are a variety of alternative versions of alternative semantics. Return to some in lecture 2. Sampling:

- Sets vs. characteristic functions of sets (Karttunen 1977 etc.).
- Partition semantics (questions denote equivalence relations; Groenendijk & Stokhof 1984 etc.).
- Inquisitive semantics (questions denote downward-closed alternative sets; Ciardelli et al. 2013 etc.)
- Two-dimensional accounts (von Stechow 1991), integration with focus semantics, etc.
- Compositional Hamblin semantics (Kratzer & Shimoyama 2002).

- Generally, take question-meaning as (constructed from) a property abstracting over the interrogative DP.
- Direct correspondence between fragment answers and question-meanings (Ginzburg & Sag 2000).
You should be wondering: what is an answer?
You should be wondering: what is an answer? Idea from Groenendijk & Stokhof (1984): distinguish

• *Semantic answers*: roughly, alternative propositions useful in understanding embedded question meanings. (Cf. also Ginzburg & Sag (2000)’s ‘simple answers’.)
You should be wondering: what is an answer? Idea from Groenendijk & Stokhof (1984): distinguish

- **Semantic answers**: roughly, alternative propositions useful in understanding embedded question meanings. (Cf. also Ginzburg & Sag (2000)’s ‘simple answers’.)

from

- **Pragmatic answers**: what (intuitively) counts as answering a question in discourse?
  - Side note: frequent conflation of these notions in the literature. E.g. Hamblin alternatives (from Karttunen onwards) are not typically intended as pragmatic answers.
An influential idea (Roberts 1996) / (Ginzburg 1994, 1996)

Discourse is structured around (potentially implicit) Questions Under Discussion (QUDs).
An influential idea (Roberts 1996) / (Ginzburg 1994, 1996)

Discourse is structured around (potentially implicit) Questions Under Discussion (QUDs).

- Assertions are always addressing some QUD. (There is always a QUD.)
- Overt questions introduce a QUD.
- Hearers can infer covert QUD(s).
- There can be multiple QUDs, hierarchically structured.
Hierarchical QUDs

From Büring (2003):
A puzzle. Why is the following response acceptable (if not necessarily true)?

(6)  A:  Is Mamoun’s any good?
     B:  No, you should go to Evelyn’s.

Standard assumption for polar questions (Hamblin):

(7)  \[ \text{Is Mamoun’s any good?} = \begin{cases} 
\lambda w_s. \text{Mamoun’s is good in } w, \\
\lambda w_s. \text{Mamoun’s is not (any) good in } w
\end{cases} \]
A puzzle. Why is the following response acceptable (if not necessarily true)?

(6)  
A: Is Mamoun’s any good?  
B: No, you should go to Evelyn’s.

Standard assumption for polar questions (Hamblin):

(7) \[
\text{[Is Mamoun’s any good?]} = \\
\left\{ \lambda w_S. \text{Mamoun’s is good in } w, \\
\lambda w_S. \text{Mamoun’s is not (any) good in } w \right\}
\]
A puzzle. Why is the following response acceptable (if not necessarily true)?

(6)  
A: Is Mamoun’s any good?
B: No, you should go to Evelyn’s.

Standard assumption for polar questions (Hamblin):

(7) \[
[\text{Is Mamoun’s any good?}] = \begin{cases} 
\lambda w. \text{Mamoun’s is good in } w, \\
\lambda w. \text{Mamoun’s is not (any) good in } w 
\end{cases}
\]

- Response is addressing some salient QUD: polar question is the tip of an iceberg (Biezma & Rawlins 2012).
A standard pragmatic notion of answerhood from Roberts (1996):

(8) a. A partial answer to a question $q$ is a proposition which contextually entails the evaluation – either true or false – of at least one element of $q$’s alternative set.

b. A complete answer to a question $q$ is a proposition which contextually entails the evaluation for all of $q$’s alternative set.

(Where $p$ contextually entails $p'$ in context $c \subseteq W$ just in case $p \cap c$ entails $p'$.)
Roberts’ relevance constraint

Extending pragmatic answers to strategies:

(9) A strategy of inquiry (intuitively) for a question $q$ is a set of subquestions that contribute to resolving $q$ (their answers at least partially answer $q$).

(10) A move $m$ is Relevant to the question under discussion $q$, i.e., to last(QUD($m$)), iff $m$ either introduces a partial answer to $q$ ($m$ is an assertion) or is part of a strategy to answer $q$ ($m$ is a question).
(11) Is Mamoun’s any good?

In the right context, this question indicates a strategy for a bigger question, ‘where should I go for lunch?’

• (Of course, the real challenge: inferring this bigger question from contextual cues.)
QA from the linguistic perspective

QA has **the same basic ingredients** as the linguistic problem.

- Infer the immediate QUD. (But this is just boringly overt in factoid QA.)
- Find information that supports some answer. This is where differences emerge.
- Extract an answer in the right linguistic form from this information.
Question and response types
I will use **question** to refer to a type of *move in discourse*...

I will use **interrogative** to refer to a linguistic object with interrogative form.
I will use **question** to refer to a type of *move in discourse*...

I will use **interrogative** to refer to a linguistic object with interrogative form.

Coarse diagnostics for interrogative form:

- Subject-aux inversion in root clauses.
- Presence of ‘**whether**’ in embedded clauses.
- Presence of interrogative pronoun/DP.
Constituent questions

(Sometimes called ‘wh-questions’ etc.)

- Constituent of sentence replaced by interrogative pronoun/DP, typically fronted.

Interrogative pronouns in English

- who, what, when, where, why, how, which(+NP), whose(+NP), how many/much(+NP), how+AP
Polar questions (at first glance) ask for a single yes/no choice:

(12) Is Mamoun’s any good?

Alternative questions present a choice:

(13) Should we go to Mamoun’s or should we go to Evelyn’s?
Polar/alternative questions

Polar questions (at first glance) ask for a single yes/no choice:

(12)  Is Mamoun’s any good? [pitch: final rise]

Alternative questions present a choice:

(13)  Should we go to Mamoun’s [pitch: peak] or should we go to Evelyn’s? [pitch: final fall]
Polar/alternative questions

Polar questions (at first glance) ask for a single yes/no choice:

(12) Is Mamoun’s any good? [pitch: final rise]

Alternative questions present a choice:

(13) Should we go to Mamoun’s [pitch: peak] or should we go to Evelyn’s? [pitch: final fall]

Examples with ‘or’ may be string ambiguous:

(14) Should we go to Mamoun’s or Evelyn’s?
Final rise is independent of interrogative form.

(15)  
  a. Is that a persimmon? (Gunlogson 2001)  
  b. That’s a persimmon?

Intuition: (b) sentence is appropriate in a context where someone other than the speaker is already committed to the content proposition.
Extras: falling declaratives can ask questions?

(16) (Context, coffee shop the other day.)
A: You should use this code for anything over 4oz.
B: Which includes a cappucino.
A: (to C) Hey C, I have a question for you...

• Minimally, B’s response seeks confirmation that a cappucino uses more than 4oz. (More examples: Gunlogson 2008)
Extras: bias and tag questions

Questions are very often not ‘neutral’: can convey something about speaker’s expectations/beliefs in context. Range of linguistic tools.

(17) Isn’t that a persimmon?

(18) Is that really a persimmon?

(19) a. That’s a persimmon, isn’t it?
   b. That’s a persimmon, is it?
   c. That’s not a persimmon, is it?
   d. That’s a persimmon, right?

(20) a. That’s a persimmon, don’t you think? (‘interrogative slift’)
   b. Where did John go, do you think? (Haddican et al. 2014 ex. 2)
Worth keeping in mind, from a practical perspective, that there’s all sorts of stuff out there:

(21)  a. (pointing at a book) any good?
     b. A: Max is leaving. B: leaving? (Ginzburg 2012 ex. 30-a)

(22)  What if the dorm doesn’t have AC?

(23)  How about we go get some lunch?
(24) Where did you go for lunch?
Information-seeking questions

(24) Where did you go for lunch?

The unmarked case (Searlean defaults)

- Q does not have the appropriate information to decide between possible answers.
- Q wants that information.
- Q thinks A might be able and willing to decide at least some answers.
Both in linguistics and QA, there has been a tendency to focus on questions with ‘simple’ answers. This is not the only type of question.
Both in linguistics and QA, there has been a tendency to focus on questions with ‘simple’ answers. This is not the only type of question.

(25) What financial relationships exist between drug companies and universities? (ex. from TREC CiQA track, 2006)

(26) Why is the sky blue?

Hirschman & Gaizauskas (2001): “We have evidence that some kinds of questions are harder than others. For example, why and how questions tend to be more difficult, because they require understanding causality or instrumental relations, and these are typically expressed as clauses or separate sentences.”
A puzzle: apparent questions can be used to convey information.

(27)  A: It’s going to rain.
     B: Who cares? (⇝ noone cares)

(28)  Context: room is evidently sweltering hot.
     a. Is it hot or what? (⇝ it’s hot!)
     b. Is it hot or is it hot!
QA systems typically distinguish *factoid* from *list* questions.

Typical linguistics example:

(29) Which student read which novel?

Example from TREC 2007:

(30) What women have worn Chanel clothing to award ceremonies?

⇒ some questions have multiple true answers (in some sense).
Questions with multiple answers differ in context as to how many true answers they need to be ‘finished’. Reading dependent on speaker’s goals (van Rooy 2003)

(31) Where can you get coffee around here? (mention-some)
Questions with multiple answers differ in context as to how many true answers they need to be ‘finished’. Reading dependent on speaker’s goals (van Rooy 2003)

(31) Where can you get coffee around here? (mention-some)
(32) I’m writing a travel guide. Where can you get coffee around here? (mention-all)
Hearers often have to do substantial inference over domain restrictions in order to establish the intended space of answers.

(33) Who wasn’t at the Math lecture today? (Hirschman & Gaizauskas 2001) ↺ list everyone in the world who wasn’t there.
Hearers often have to do substantial inference over domain restrictions in order to establish the intended space of answers.

(33) Who wasn’t at the Math lecture today? (Hirschman & Gaizauskas 2001) \( \not \leftrightarrow \) list everyone in the world who wasn’t there.

Domain restriction is a major practical problem. Light et al. (2001): huge effect of what cues interrogative DP provides to domain, with ‘what’ being the worst case.
What is a pragmatic answer? For typical information-seeking Qs things may seem straightforward:

(34) A: Is Mamoun’s any good?
    B: Yes, it is.

(35) A: Who’s teaching this class?
    B: Kyle is.
Intuitions about answerhood

It isn’t entirely clear that humans have any reliable intuitions about what a pragmatic ‘answer’ in general.
Intuitions about answerhood

It isn’t entirely clear that humans have any reliable intuitions about what a pragmatic ‘answer’ in general.

For example, do you (intuitively) think that B is answering A’s question?

(36)  
A:  What time is the party?  
B:  I don’t know.
Answers and other responses

Intuitions about answerhood

It isn’t entirely clear that humans have any reliable intuitions about what a pragmatic ‘answer’ in general.

For example, do you (intuitively) think that B is answering A’s question?

(36)  A:  What time is the party?
       B:  I don’t know.

For this reason, even for pragmatic answerhood it is helpful to look at ‘answer’ through some theoretical lens. (And lenses vary w.r.t. B’s response!)
Expressions of ignorance or the limits of knowledge do seem to discharge the answerer’s responsibility in discourse.

(37)  A: What time is the party?
     B: I don’t know.
     B’: It might be in the evening?
     B’’: Carmen might know.
Expressions of ignorance or the limits of knowledge do seem to discharge the answerer’s responsibility in discourse.

(37) A: What time is the party?
    B: I don’t know.
    B’: It might be in the evening?
    B”': Carmen might know.

There are other ways to discharge (Asher 2014):

(5) a. N: Excuse me. Could you tell me the time please?
    b. B: Fuck you!
Addressing presuppositions

Judgment: answer or not?

(38)   A: If you go to the party, will you talk to Joanna?
       B: I’m not going to the party.
Addressing presuppositions

Judgment: answer or not?

(38)  A: If you go to the party, will you talk to Joanna?
       B: I’m not going to the party.

Isaacs & Rawlins (2008): B’s response is addressing a presupposition of the conditional (that the antecedent is possible), and in so doing, discharges the question by eliminating a suppositional context.
(39) Where should we go for lunch? Is Mamoun’s any good?
(40) Where’d you go for lunch? Did you end up going to Mamoun’s?
(41) A: Where should we go for lunch?
   B: Do you like middle eastern food?
Clarification questions

(42)  A:  Who’s teaching the NASSLLI course on questions?
    B:  Which course?
    B’: the what course?
    B’: Why?
Answers to English polar questions typically have ‘yes’ or ‘no’ attached.

(43)  Is Mamoun’s any good?
  a.  Yes(, it is).
  b.  No(, it’s not).

Default hypothesis: yes/no mark positive/negative answers respectively.
Answer/polarity particles

Answers to English polar questions typically have ‘yes’ or ‘no’ attached.

(43)  Is Mamoun’s any good?
    a.  Yes(, it is).
    b.  No(, it’s not).

Default hypothesis: yes/no mark positive/negative answers respectively.

- Complication 1: some languages have > 2 answer particles, e.g. reverse particles (Farkas & Bruce 2010).
- Complication 2: complex interaction with negative vs. positive questions (Kramer & Rawlins 2009, Brasoveanu et al. 2012, Krifka 2013)
### Moral 1

The space of responses is much bigger than the space of straightforward answers.
Empirical morals about responses

Moral 1
The space of responses is much bigger than the space of straightforward answers.

Moral 2
In naturalistic discourse, any given question is just the tip of the iceberg.
Empirical morals about responses

Moral 1
The space of responses is much bigger than the space of straightforward answers.

Moral 2
In naturalistic discourse, any given question is just the tip of the iceberg.

Moral 3
A responder may need to do inference about the intent of the questioner.
Four notions of ‘answering’ (lecture 4)

1. Semantic ‘answers’.
2. Does a response fulfill a cooperative responder’s obligation (or otherwise exit the discourse)?
3. Does a response render the context uninquisitive? (With or without resolution.)
4. Does a response contribute to resolving/‘answering’ a question? Or indicate a strategy for doing so?
Next up: question semantics

- Alternative semantics accounts of interrogative clause meanings.
- Structured meanings.
- Questions as knowledge-base queries.
- Questions as vector-space topics.


Farkas, Donka & Kim Bruce. 2010. On reacting to assertions and polar questions. *Journal of Semantics* 27. 81–118.


