Modeling questions and responses

Lecture 2: Representing question meanings

Kyle Rawlins

NASSLLI 2016, 12th July, 2016

Johns Hopkins University, Department of Cognitive Science

Course structure

- *Lecture 1*: Introducing questions and responses.

⇒ *Lecture 2*: Types of questioning, and representing question meanings.

- *Lecture 3*: The architecture of a QA system.
- *Lecture 4*: Modeling questions in discourse.
- *Lecture 5*: Spillover, meta-conversation.

Outline

Within the bounds of a Q-A game:

1. Give some question Q, what information could in principle answer Q?
2. Given some possibilities for information that addresses Q, what information *correctly* answers Q?
3. Given information that correctly answers Q, how can this information be formulated as *an answer*?

Types of questions and responses

Types of interrogatives (summary)

Type	main cues	pitch
polar	aux inversion, ‘whether’	rising (root only)
alternative	inversion, ‘whether’, disjunction	final fall
constituent	wh-item, inversion	flat (see Bartels 1999)

Types of interrogatives (summary)

Type	main cues	pitch
polar	aux inversion, ‘whether’	rising (root only)
alternative	inversion, ‘whether’, disjunction	final fall
constituent	wh-item, inversion	flat (see Bartels 1999)

Further cues for interrogative-like properties:

- Tags and slifts.
- Rising pitch independent of interrogative form.
- Fixed constructions (‘what about’ etc.).
- Focus structure (See QUD course).

Information-seeking questions

(1) Where did you go for lunch?

Information-seeking questions

(1) 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.

Complex information-seeking questions

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.

Complex information-seeking questions

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.

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

(3) 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."

Rhetorical/information-conveying questions

A puzzle: apparent questions can be used to convey information.

Cf. Biased questions from lecture 1.

(4) A: It's going to rain.
B: Who cares? ($\rightsquigarrow$ noone cares)

(5) Context: room is evidently sweltering hot.
a. Is it hot or what? ($\rightsquigarrow$ it's hot!)
b. Is it hot or is it hot!

List questions

QA systems typically distinguish *factoid* from *list* questions.

Typical linguistics example:

(6) Which student read which novel?

Example from TREC 2007:

(7) 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)

(8) Where can you get coffee around here? (mention-some)

Mention-some vs. all

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)

(8) Where can you get coffee around here? (mention-some)

(9) I'm writing a travel guide. Where can you get coffee around here? (mention-all)

The role of domain restriction and inference

Hearers often have to do substantial inference over domain restrictions in order to establish the intended space of answers.

(10) Who wasn't at the Math lecture today? (Hirschman & Gaizauskas 2001) $\not\rightsquigarrow$ list everyone in the world who wasn't there.

The role of domain restriction and inference

Hearers often have to do substantial inference over domain restrictions in order to establish the intended space of answers.

(10) Who wasn't at the Math lecture today? (Hirschman & Gaizauskas 2001) $\not\rightsquigarrow$ 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.

Questions and requests

Polite requests in English are typically constructed with interrogative form, not imperative form:

(11) Can you open the window?

(12) #Open the window.

Questions and requests

Polite requests in English are typically constructed with interrogative form, not imperative form:

(11) Can you open the window?

(12) #Open the window.

An appropriate response is an action. Moreover, some questions can be 'answered' with actions:

(13) A: What are you reading?
B: (shows book cover to A)

(This is already exploited by many QA systems that can't accurately form a linguistic answer!)

Responses and followups

Next, a brief survey of some ways one can respond to a question.

The obvious kind of response: answers

What is a pragmatic answer? For typical information-seeking Qs things may seem straightforward:

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

(15) A: Who's teaching this class?
B: Kyle is.

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.

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?

(16) 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?

(16) 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!)

Inability / refusal to answer

Expressions of ignorance or the limits of knowledge do seem to discharge the answerer's responsibility in discourse.

(17) A: What time is the party?
B: I don't know.
B′: It might be in the evening?
B″: Carmen might know.

Inability / refusal to answer

Expressions of ignorance or the limits of knowledge do seem to discharge the answerer's responsibility in discourse.

(17) 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?

(18) 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?

(18) 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.

Question-question sequences

(19) Where should we go for lunch? Is Mamoun's any good?

(20) Where'd you go for lunch? Did you end up going to Mamoun's?

(21) A: Where should we go for lunch?
B: Do you like middle eastern food?

Clarification questions (etc.)

Ginzburg (1998, 2012), Schlöder & Fernández (2015) a.o.

(22) A: Who's teaching the NASSLLI course on questions?
B: Which course?
B′: the what course?
B′: Why?

(23) (Fernández (2006) ch 2 ex. 12b, from BNC)
A: [...] You lift your crane out, so this part would come up.
B: The end?
A: The end would come up and keep your load level on the ground you see.

Answer/polarity particles

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

(24) 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.

(24) 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)

Responses beyond questions

Is it really possible to partition off questions when understanding responses?

- Farkas & Bruce (2010) a.o.: polarity particles get reused across different speech act types. (Eng: ‘ok’, ‘no’, ‘yeah’)

Responses beyond questions

Is it really possible to partition off questions when understanding responses?

- Farkas & Bruce (2010) a.o.: polarity particles get reused across different speech act types. (Eng: 'ok', 'no', 'yeah')

(25) A: It's raining.
B: Yeah / ok / no it isn't.

(26) A: Is it raining?
B: Yeah / #ok / no (it isn't).

(27) A: Open the window.
B: ??yeah / ok / no.

Generalization of the problem

Assertions, imperatives, and questions enter a proposal to coordinate on what should be common ground (in some sense); this coordination always involves inquiry at some level. (Continue in lecture 4)

Empirical morals about responses

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.

Embedded questions and embedding predicates

Root and embedded questions

The analysis of interrogative clauses must handle the parallels between root and embedded questions (Karttunen 1977, Groenendijk & Stokhof 1984):

(28) When was Justin Trudeau born?

(29) Kyle knows when Justin Trudeau was born.
$\rightsquigarrow$ K would be able to correctly answer the question: when was JT born?

Root and embedded questions

The analysis of interrogative clauses must handle the parallels between root and embedded questions (Karttunen 1977, Groenendijk & Stokhof 1984):

(28) When was Justin Trudeau born?

(29) Kyle knows when Justin Trudeau was born.
$\rightsquigarrow$ K would be able to correctly answer the question: when was JT born?

Another paraphrase: K's epistemic state decides every proposition in $[\![$**When was JT born?**$]\!]$.

Root and embedded questions

While there has been a focus on 'know', there are hundreds of Q-embedding predicates in English, of great variety.

- Hacquard & Wellwood (2012) appendix C: address, advertise, announce, answer, clarify, communicate, confirm, defend, demonstrate, demystify, depict, describe, detail, dictate, disclose, discuss, emphasize, explain, highlight, illuminate, illustrate, indicate, mention, negotiate, overstate, pinpoint, prove, publicize, recommend, relay, remind, report, reveal, say, show, signal, state, suggest, summarize, tell, underscore, anticipate, estimate, forecast, foresee, guess, overestimate, predict, project, speculate control, decide, determine, plan, prejudge, settle, specify, analyze, ask, assess, calculate, check, compute, consider, contemplate, debate, evaluate, examine, explore, figure, gauge, inquire, investigate, judge, misjudge, ponder, probe, ques- tion, re-evaluate, re-evaluating, reassess, reconsider, reevaluate, reexamine, research, re- view, study, think, wonder, ascertain, detect, discern, discover, find, found, gather, hear, identify, infer, learn, observe, realize, record, sense, baffle, disregard, doubt, envision, fathom, forget, grasp, ignore, imagine, know, mind, overlook, recall, remember, see, surprise, underestimate, understand, visualize, appreciate, dig, hate, like, care, influence, matter

Related constructions

A generalization: interrogative morphology is often rather multi-purpose.

- Relative structures (especially free relatives).
- Cleft structures.
- Exclamatives.
- Indefinite items.

(30) Alfonso ate what he was given.

Two standard diagnostics (Baker 1968, 1970):

- Free relatives don't accept 'else'.
  (31) *Alfonso ate what else he was given.
- Free relatives don't accept 'the hell'.
  (32) *Alfonso ate what the hell he was given.
- (The lines get blurry when '-ever' is in the picture.)

Alternative semantics

Alternative semantics for questions

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

(33) ⟦**What year was Justin Trudeau born in?**⟧ =

$$\left\{\begin{array}{l} \ldots, \\ \lambda w_s \,.\, \text{JT was born in 1969 in } w, \\ \lambda w_s \,.\, \text{JT was born in 1970 in } w, \\ \lambda w_s \,.\, \text{JT was born in 1971 in } w, \\ \lambda w_s \,.\, \text{JT was born in 1972 in } w, \\ \ldots \end{array}\right\}$$

Hamblin's third postulate (Hamblin 1958)

The possible answers to a question are an exhaustive set of mutually exclusive possibilities.

- Exhaustive: the alternatives cover the set of possibilities. (E.g. $\mathscr{W}$, or the context set, or the epistemic state, or...)
- Mutually exclusive: the alternatives don't overlap.

There is a large literature on getting exclusivity right that I will not go into, starting with (Groenendijk & Stokhof 1984, Heim 1994).

An argument against exclusivity based on Yablo:

(34) How many stars are there in the sky, plus or minus 10?

Constructing alternative sets compositionally

Various approaches to this. One standard approach:

- Compositional Hamblin semantics: denotations are sets of ordinary denotations. Enrich composition with Pointwise Function Application (PFA). (Kratzer & Shimoyama 2002, Kratzer 2005)
- Does well with languages where interrogative DPs appear in situ.
- Does well with languages where interrogative DPs have indefinite uses. (Japanese)

See lambda notebook Hamblin semantics demo.

Partition semantics for questions

Groenendijk & Stokhof (1984)

Question-meanings are equivalence relations in sets of possible worlds.

- Or, an interrogative clause at *w* denotes the true exhaustive answer to the question at *w*.
- Equivalence relation: reflexive, symmetric, transitive.

Suppose that in w_1, w_2 J goes to the party, and in w_3, w_4 she doesn't.

(35) $[\![$Is J going to the party?$]\!] =$

$$\left\{ \begin{array}{llll} \langle w_1, w_1 \rangle, & \langle w_1, w_2 \rangle, & & \\ \langle w_2, w_1 \rangle, & \langle w_2, w_2 \rangle, & & \\ & & \langle w_3, w_3 \rangle, & \langle w_3, w_4 \rangle, \\ & & \langle w_4, w_3 \rangle, & \langle w_4, w_4 \rangle \end{array} \right\}$$

Converting partitions to alternative sets

- Two worlds belong in the same alternative iff they are connected.
- Conversion can go both ways as long as the alternative set is a partition (obeys the third postulate).

(36) $\text{Alts}(Q_{\langle s,\langle s,t\rangle\rangle}) = \{p_{\langle st\rangle} \mid p \neq \emptyset \wedge \exists u_s : \forall v_s : Q(u)(v) \leftrightarrow p(v)\}$

Inquisitive semantics adjust the alternative semantics picture in two ways that I'll discuss here (using Ciardelli et al. (2013) as a reference):

- Hamblin's third postulate is dropped.
- Alternative sets are downward closed.

Non-exhaustivity and hybrids

Challenges to exhaustivity:

- At best, we want exhaustivity relative to a local context.
- A *hybrid*: a question-meaning that consists of alternatives in some form but does not exhaust the relevant space of possibilities. E.g. it is both informative and inquisitive (Groenendijk 2009, Ciardelli et al. 2013).
- Many uses for dropping exhaustivity: presuppositions, clause-embedding (Rawlins 2013), etc.

Illustration of non-exhaustivity in a partition semantics

Suppose that in w_1, w_2 J goes to the party, and in w_3, w_4 she doesn't. In w_5 there is no party.

(37) $[\![$Is J going to the party?$]\!] =$

$$\left\{ \begin{array}{llll} \langle w_1, w_1 \rangle, & \langle w_1, w_2 \rangle, & & \\ \langle w_2, w_1 \rangle, & \langle w_2, w_2 \rangle, & & \\ & & \langle w_3, w_3 \rangle, & \langle w_3, w_4 \rangle, \\ & & \langle w_4, w_3 \rangle, & \langle w_4, w_4 \rangle \end{array} \right\}$$

Now, this set is an equivalence relation (and therefore exhaustive) relative to $\{w_1, w_2, w_3, w_4\}$, worlds where the presupposition is satisfied, but not relative to $\mathscr{W}$.

Another argument against third postulate

Suppose you want the *denotation* of a conditional interrogative clause (Isaacs & Rawlins 2008; inquisitive semantics work e.g. Ciardelli et al. 2013, talk at most recent SALT).

(38) If it rains, will you go to the party?

(39) John hasn't decided whether he will go the party if it rains.

Another argument against third postulate

Suppose you want the *denotation* of a conditional interrogative clause (Isaacs & Rawlins 2008; inquisitive semantics work e.g. Ciardelli et al. 2013, talk at most recent SALT).

(38) If it rains, will you go to the party?

(39) John hasn't decided whether he will go the party if it rains.

If you want this denotation in static terms (e.g. as an alternative set), it is *extremely hard* to avoid overlapping alternatives on worlds where the antecedent is false, or to avoid having a non-exhaustive alternative set excluding those worlds (your choice).

Downward-closed alternative sets

Ciardelli et al. (2013):

Definition 2 (Issues).
Let s be an information state, and $\mathcal{I}$ a non-empty set of enhancements of s. Then we say that $\mathcal{I}$ is an issue over s if and only if

1. $\mathcal{I}$ is *downward closed*: if $t \in \mathcal{I}$ and $t' \subset t$ then also $t' \in \mathcal{I}$.
2. $\mathcal{I}$ forms a *cover* of s: $\bigcup \mathcal{I} = s$.

- Intuition: an issue contains any information state that resolves that issue. A Hamblin issue contains only maximal information states that resolve it.
- Search in QA is over something closer to downward-closed issues.

Side note: downward closure vs. pragmatic answerhood

- One reason for downward closure is to directly license any resolving response.
- Hamblin semantic answers don't do this.

Side note: downward closure vs. pragmatic answerhood

- One reason for downward closure is to directly license any resolving response.
- Hamblin semantic answers don't do this.
- Should this be done in the semantics or pragmatics? Roberts (1996) definition from yesterday accomplishes the same thing via pragmatic answerhood: any completely resolving response is a complete answer by that def.

Summary: alternative semantics

Interrogative clause denotation is an alternative set whose members correspond to answers.

- Potential constraint on alternative sets (with caveats): Hamblin's third postulate.
- Alternatives are internally undifferentiated. (E.g. propositional.)
- Use in QA would require deep natural language understanding.

Structured meanings

Fragment answers

(40) A: Who is teaching this class?
B: Kyle.

- Hamblin: reject fragment answers as the primary form of answers.
- Structured approaches start from fragment answers as primary. (Some key cites: Ginzburg 1992, Ginzburg & Sag 2000, Krifka 2001)

Fragment answers

(40) A: Who is teaching this class?
B: Kyle.

- Hamblin: reject fragment answers as the primary form of answers.
- Structured approaches start from fragment answers as primary. (Some key cites: Ginzburg 1992, Ginzburg & Sag 2000, Krifka 2001)

(41) $[\![$**Who is teaching this class?**$]\!] =$
$\lambda x_e . \lambda w_s . \text{teaching-this-class}'(w,x)$

Need to incorporate the meaning of the interrogative DP. The denotation of an interrogative clause as a pair (Krifka 2001):

(42) $[\![$**Who is teaching this class?**$]\!] =$
$\langle \lambda x_e \,.\, \lambda w_s \,.\, \text{teaching-this-class}'(w, x), \textsc{Person} \rangle$

Note: if Person is treated as a Hamblin set and these compose via PFA, you get the Hamblin meaning.

Basic notion of aboutness/relevance

(43) A (fragment) response X is relevant to a structured question-meaning Q just in case $[\![X]\!] \in Q[1]$.

(This is pretty dumb as-is, see Ginzburg & Sag (2000), Ginzburg (2012) for real versions.)

Even richer structured meanings

See Ginzburg (2012) for an extremely sophisticated development of this in the TTR framework:

$$(\mathrm{r} : \begin{bmatrix} \mathrm{x} & : & \mathrm{Ind} \\ \mathrm{rest} & : & \mathrm{person(x)} \end{bmatrix}) \begin{bmatrix} \mathrm{sit} & = & r_1 \\ \mathrm{sit\text{-}type} & = & \begin{bmatrix} \mathrm{c} : \mathrm{run(r.x)} \end{bmatrix} \end{bmatrix}$$

Questions and (unstructured) topic models

Information Retrieval (IR) / vector spaces 101

If you're in ~1999 and want to build a search engine, you'd probably start with some variant of the following:

1. Collect all documents you want to search.
2. Preprocess them (remove stop words (mostly function words), clean up formatting).
3. Build a vector of unigram word counts for all remaining words in the document.
    - Each vector now becomes a point in a high-dimensional space.
4. Turn query into vector, and find documents via some distance metric in this space. Typically, cosine distance.

$$\frac{\vec{a} \cdot \vec{b}}{||\vec{a}||||\vec{b}}$$

Vector spaces 102

- Previous slide is extremely simplistic, but it is still the nutshell of vector space models.
- Nowadays, one would probably start with something a bit less blunt, such as LDA (Latent Dirichlet Allocation) – a document is a mixture of topics. (Overview: Blei 2012). But LDA is only the beginning of this rabbit hole…

Vector spaces 102

- Previous slide is extremely simplistic, but it is still the nutshell of vector space models.
- Nowadays, one would probably start with something a bit less blunt, such as LDA (Latent Dirichlet Allocation) – a document is a mixture of topics. (Overview: Blei 2012). But LDA is only the beginning of this rabbit hole...
- Some other starts on making it more sophisticated: better preprocessing (e.g. tf-idf, normalize by how common a term is across many documents), more interesting vectors than unigram frequency (n-gram, integrated tagging/parses, etc), better distance metrics.

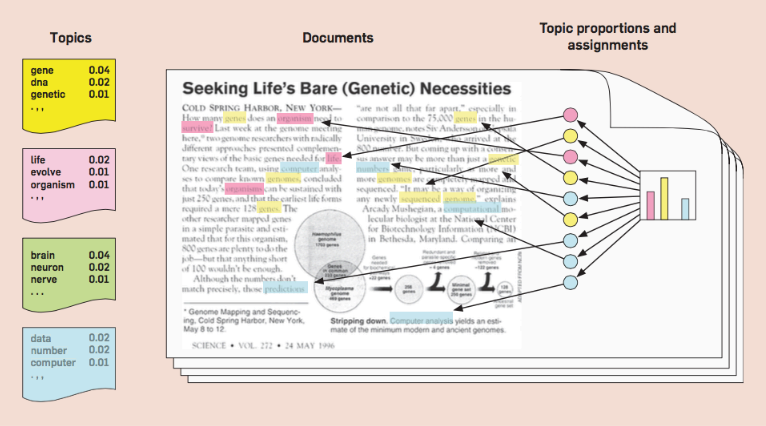
Vector spaces 102

- Previous slide is extremely simplistic, but it is still the nutshell of vector space models.
- Nowadays, one would probably start with something a bit less blunt, such as LDA (Latent Dirichlet Allocation) – a document is a mixture of topics. (Overview: Blei 2012). But LDA is only the beginning of this rabbit hole...
- Some other starts on making it more sophisticated: better preprocessing (e.g. tf-idf, normalize by how common a term is across many documents), more interesting vectors than unigram frequency (n-gram, integrated tagging/parses, etc), better distance metrics.
- A vector is proxy for a topic (in some sense of the word topic).

  ⇒ Question can be represented as, effectively, a topic.

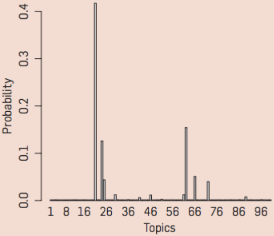
Blei 2012 fig 1

Figure 1. The intuitions behind latent Dirichlet allocation. We assume that some number of "topics," which are distributions over words, exist for the whole collection (far left). Each document is assumed to be generated as follows. First choose a distribution over the topics (the histogram at right); then, for each word, choose a topic assignment (the colored coins) and choose the word from the corresponding topic. The topics and topic assignments in this figure are illustrative—they are not fit from real data. See Figure 2 for topics fit from data.



Blei 2012 fig 2

Figure 2. Real inference with LDA. We fit a 100-topic LDA model to 17,000 articles from the journal *Science*. At left are the inferred topic proportions for the example article in Figure 1. At right are the top 15 most frequent words from the most frequent topics found in this article.



"Genetics"	"Evolution"	"Disease"	"Computers"
human	evolution	disease	computer
genome	evolutionary	host	models
dna	species	bacteria	information
genetic	organisms	diseases	data
genes	life	resistance	computers
sequence	origin	bacterial	system
gene	biology	new	network
molecular	groups	strains	systems
sequencing	phylogenetic	control	model
map	living	infectious	parallel
information	diversity	malaria	methods
genetics	group	parasite	networks
mapping	new	parasites	software
project	two	united	new
sequences	common	tuberculosis	simulations

Questions as (vector-space) topics

Summary

to find supporting information for a question from a large set of unstructured text, can use vector-space techniques to identify likely documents / likely passages.

- These typically won't get you all the way to the answer to the question, though they may get far enough for certain tasks.
- Need more fine-grained natural language understanding techniques (e.g. semantic parsing).

Questions as queries / structured topics

A very old idea: questions as db queries

Suppose you have a SQL table that looks like this:

(44) Table `People`:

firstname	lastname	birthdate	birthplace	height	...
Justin	Trudeau	12/25/71	Ottawa	6’ 2”	...
Kyle	Rawlins	12/30/79	Boston	5’ 10”	...
...					

Suppose you have a SQL table that looks like this:

(44) Table `People`:

firstname	lastname	birthdate	birthplace	height	...
Justin	Trudeau	12/25/71	Ottawa	6' 2"	...
Kyle	Rawlins	12/30/79	Boston	5' 10"	...
...					

Then, if you can convert 'When was JT born?' into the following, you're all set:

```
from People select birthdate where first-
name="Justin" and lastname="Trudeau";
```

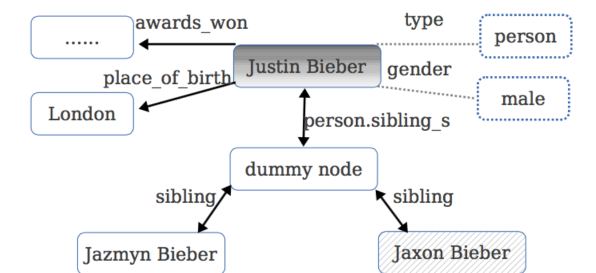
Knowledge-base queries

In fact, many organizations are collecting just this sort of structured information. (e.g. Freebase ⇒ Google Knowledge Graph)

Converting NL to queries

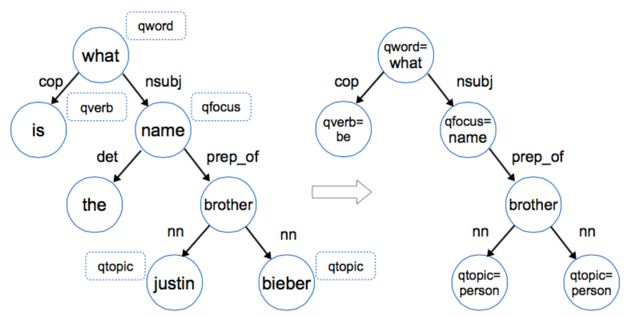
Getting the question ⇒ query mapping in general is extremely hard.

- Traditional databases are much too rigid to do this in a general way.
- Need probabilistic models, flexible data / query language.
- Modern viewpoint: query/knowledgebase as semi-structured topic graphs (Yao & Durme 2014). Freebase:



Convert NL to queries

Yao & Durme (2014) query format: convert a Stanford dependency parse to a more general graph:



Learn a model for aligning question graphs to topic graphs. (A lot of hard machine learning elided here.)

Summarizing

- All four types of question-meaning discussed today involve a notion of what the question is about.
- They differ in how structured the notion of aboutness is.
- They differ in how much the structure and form of the question determine what the question is about.

Bibliography I

Asher, Nicholas. 2014. The non-cooperative basis of implicatures. In *Formal Approaches to Semantics and Pragmatics*, Springer.

Baker, Carl Lee. 1968. *Indirect Questions in English*: University of Illinois Ph.D. dissertation.

Baker, Carl Lee. 1970. Notes on the description of English questions: the role of an abstract question morpheme. *Foundations of Language* 6. 197–219.

Bartels, Christine. 1999. *The intonation of English statements and questions*. Garland Publishing.

Blei, David. 2012. Probabilistic topic models. *Communications of the ACM* 55(4). 77–84.

Bibliography II

Brasoveanu, Adrian, Donka Farkas & Floris Roelofsen. 2012. Polarity particles and the anatomy of n-words. In *Proceedings of SuB 2011*, .

Ciardelli, Ivano, Jeroen Groenendijk & Floris Roelofsen. 2013. Inquisitive semantics: a new notion of meaning. *Language and linguistic compass* 7(9). 459–476.

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

Fernández, Raquel. 2006. *Non-sentential utterences in dialogue: classification, resolution, and use*: University of London dissertation.

Ginzburg, Jonathan. 1992. *Questions, queries and facts: A semantics and pragmatics for interrogatives*: Stanford University Ph.D. dissertation.

Bibliography III

Ginzburg, Jonathan. 1998. Clarifying utterances. In *Proceedings of the 2nd workshop on the formal semantics and pragmatics of dialogue*, Twente.

Ginzburg, Jonathan. 2012. *The Interactive Stance: meaning for conversation.* Oxford University Press.

Ginzburg, Jonathan & Ivan Sag. 2000. *Interrogative investigations: the form, meaning, and use of English interrogatives.* CSLI Publications.

Groenendijk, Jeroen. 2009. Inquisitive semantics: Two possibilities for disjunction. In P. Bosch, D. Gabelaia & J. Lang (eds.), *Seventh International Tbilisi Symposium on Language, Logic, and Computation*, Springer.

Groenendijk, Jeroen & Martin Stokhof. 1984. *Studies in the semantics of questions and the pragmatics of answers*: University of Amsterdam PhD dissertation.

Hacquard, Valentine & Alexis Wellwood. 2012. Embedding epistemic modals in English: a corpus-based study. *Semantics and Pragmatics* 5(4). 1–29. doi:10.3765/sp.5.4.

Hamblin, C. L. 1958. Questions. *Australasian Journal of Philosophy* 36. 159–168.

Hamblin, C. L. 1973. Questions in Montague English. *Foundations of Language* 10. 41–53.

Heim, Irene. 1994. Interrogative semantics and Karttunen's semantics for *know*. In Rhonna Buchalla & Anita Mitwoch (eds.), *IATL 1: Proceedings of the Ninth Annual Conference and of the Workshop on Discourse*, The Israel Association for Theoretical Linguistics.

Hirschman, L. & R. Gaizauskas. 2001. Natural language question answering: the view from here. *Natural Language Engineering* 7(4). 275–300.

Isaacs, James & Kyle Rawlins. 2008. Conditional questions. *Journal of Semantics* 25. 269–319.

Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1. 3–44.

Kramer, Ruth & Kyle Rawlins. 2009. Polarity particles: an ellipsis account. In *Proceedings of NELS 39*, .

Kratzer, Angelika. 2005. Constraining premise sets for counterfactuals. *Journal of Semantics* 22(2). 153–158.

Kratzer, Angelika & Junko Shimoyama. 2002. Indeterminate pronouns: the view from Japanese. In Yukio Otsu (ed.), *Proceedings of the 3rd Tokyo Conference on Psycholinguistics*, 1–25. Hituzi Syobo.

Krifka, Manfred. 2001. Quantifying into question acts. *Natural Language Semantics* 9. 1–40.

Krifka, Manfred. 2013. Response particles as propositional anaphors. In *Proceedings of SALT 23*, 1–18.

Light, Marc, Gideon S. Mann, Ellen Riloff & Eric Breck. 2001. Analyses for elucidating current question answering technology. *Natural Language Engineering* 7(4). 325–342.

Rawlins, Kyle. 2013. About 'about'. In *Proceedings of SALT 2013*, .

Roberts, Craige. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics, 1998 revision. In Jae Hak Yoon & Andreas Kathol (eds.), *OSUWPL vol. 49: Papers in Semantics*, The Ohio State University, Department of Linguistics.

van Rooy, Robert. 2003. Questioning to resolve decision problems. *Linguistics and Philosophy* 26. 727–763.

Schlöder, Julian J. & Raquel Fernández. 2015. Clarifying intentions in dialogue: A corpus study. In *Proceedings of the 11th International Conference on Computational Semantics*, 46–51.

Yao, Xuchen & Benjamin Van Durme. 2014. Information extraction over structured data: question answering with Freebase. In *Proceedings of ACL 52*, 956–966.