History and goals of NLU; course plan and goals

Bill MacCartney and Christopher Potts

CS 244U: Natural language understanding
Jan 10

Goals of NLU

A handful of broad goals, and ignoring internal tensions:

- Insights into language and society
- Insights into computation
- Insights into human cognition
- Solve a major sub-problem of Artificial Intelligence
- Computers that can do our most tedious, dangerous, and/or high-precision language-oriented tasks

Technological and cognitive goals

Allen (1987:2):

> *[T]here can be two underlying motivations for building a computational theory. The technological goal is simply to build better computers, and any solution that works would be acceptable. The cognitive goal is to build a computational analog of the human-language-processing mechanism; such a theory would be acceptable only after it had been verified by experiment.*
>
> *[. . .]*
>
> *Thus, the technological goal cannot be realized without using sophisticated underlying theories that are on the level being developed by theoretical linguists. On the other hand, the present state of knowledge about natural language processing is so preliminary that attempting to build a cognitively correct model is not feasible.*

What is understanding? Some possible answers

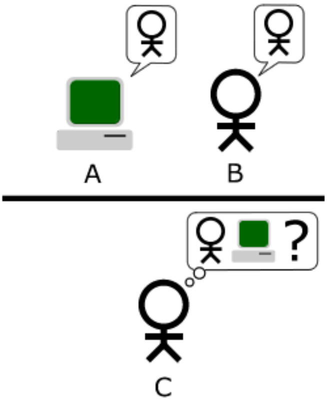
To understand a statement is to

- determine its truth (perhaps with justification); and/or
- calculate its entailments; and/or
- take appropriate action in light of it; and/or
- translate it accurately into another language; and/or
- ground it in a cognitively realistic conceptual space; and/or
- …

Turing's (1950) 'imitation game' (the Turing test)

Turing replaced "Can machines think?", which he regarded as "too meaningless to deserve discussion" (p. 442), with the question whether an interrogator could be tricked into thinking that a machine was a human using only conversation (no visuals, no demands for physical performance, etc.).

"*May not machines carry out something which ought to be described as thinking but which is very different from what a man does? This objection is a very strong one, but at least we can say that if, nevertheless, a machine can be constructed to play the imitation game satisfactorily, we need not be troubled by this objection.*" (p. 435)



http://en.wikipedia.org/wiki/Turing_test

Objections to the imitation game

Turing considers nine objections. The following have played significant roles in the scientific debate since then:

1. The theological objection (p. 443)

"*Thinking is a function of man's immortal soul. God has given an immortal soul to every man and woman, but not to any-other animal or to machines. Hence no animal or machine can think.*"

Objections to the imitation game

Turing considers nine objections. The following have played significant roles in the scientific debate since then:

1. The theological objection (p. 443)

"*Thinking is a function of man's immortal soul. God has given an immortal soul to every man and woman, but not to any-other animal or to machines. Hence no animal or machine can think.*"

Turing's reply: we want to allow that God could imbue an elephant with the power to think, so why not a machine?

Proponent: this has no prominent modern proponents, as far as I know, but Chomsky takes the position that, as a matter of usage, we use *think* only for humans and human-like entities, but he adds that this question is a uninteresting as holding birds up as the definitive case of flying and then asking whether jets really fly.
`http://www.framingbusiness.net/archives/1366`

Objections to the imitation game

Turing considers nine objections. The following have played significant roles in the scientific debate since then:

3. The Mathematical Objection (p. 444)

"*There are a number of results of mathematical logic which can be used to show that there are limitations to the powers of discrete-state machines.*"

Objections to the imitation game

Turing considers nine objections. The following have played significant roles in the scientific debate since then:

3. The Mathematical Objection (p. 444)

"*There are a number of results of mathematical logic which can be used to show that there are limitations to the powers of discrete-state machines.*"

Turing's reply: it hasn't been shown that the brain doesn't suffer from the same limitations.

Proponent: Penrose (1990), who calls on inferences from Gödel's incompleteness results (anticipated by Turing).

Objections to the imitation game

Turing considers nine objections. The following have played significant roles in the scientific debate since then:

4. The Argument from Consciousness

The machine must have a rich, human-like cognitive life, and we must be able to verify that.

Objections to the imitation game

Turing considers nine objections. The following have played significant roles in the scientific debate since then:

4. The Argument from Consciousness

The machine must have a rich, human-like cognitive life, and we must be able to verify that.

Turing's reply: in weak (behavioral) form, this is fine. In strong form, it leads us to question whether other *humans* can think.

Proponents: Penrose (1990) and Chalmers (1997).

Searle's Chinese Room Argument

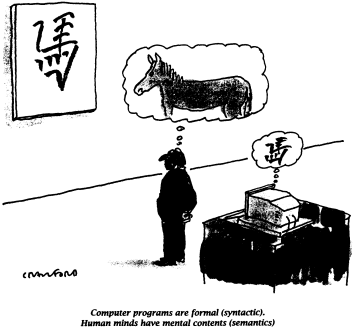
The thought experiment (see Searle 1980, 1990; Cole 2009)

Imagine yourself in a room containing a basketful of symbols from a language L that you don't understand, along with a rule book (written in English) for matching symbols in L with other symbols in L. People outside the room pass you strings of symbols in L, you follow your rules, and pass them back symbols in L. The rule book is so good that the symbols you pass back are indistinguishable from the replies of a native speaker of L. You would pass the Turing test, but (Searle says) no one would say you understand.

Searle's Chinese Room Argument

The thought experiment (see Searle 1980, 1990; Cole 2009)

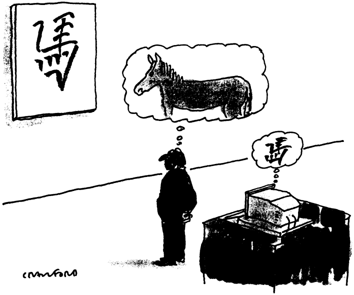
Imagine yourself in a room containing a basketful of symbols from a language L that you don't understand, along with a rule book (written in English) for matching symbols in L with other symbols in L. People outside the room pass you strings of symbols in L, you follow your rules, and pass them back symbols in L. The rule book is so good that the symbols you pass back are indistinguishable from the replies of a native speaker of L. You would pass the Turing test, but (Searle says) no one would say you understand.



Searle's Chinese Room Argument

The thought experiment (see Searle 1980, 1990; Cole 2009)

Imagine yourself in a room containing a basketful of symbols from a language *L* that you don't understand, along with a rule book (written in English) for matching symbols in *L* with other symbols in *L*. People outside the room pass you strings of symbols in *L*, you follow your rules, and pass them back symbols in *L*. The rule book is so good that the symbols you pass back are indistinguishable from the replies of a native speaker of *L*. You would pass the Turing test, but (Searle says) no one would say you understand.



Computer programs are formal (syntactic).
Human minds have mental contents (semantics)

Responses:

- The *room* understands.
- There is no reason to believe Searle's axiom "syntax by itself is neither constitutive of nor sufficient for semantics" (Churchland and Churchland 1990).

A worry based on the first Turing Test (Loebner prize)

Shieber (1994) offers an entertaining and insightful overview.

> *Cynthia Clay, the Shakespeare aficionado, was thrice misclassified as a computer. At least one of the judges made her classifications on the premise that "[no] human would have that amount of knowledge about Shakespeare" [26]. Lisette Gozo was honored as the most human of the agents for her discussion of women's clothing, although one judge rated two computer programs above her.*

> *One reporter noted that Weizenbaum himself was "disturbed" by how easily people were fooled by these programs [26], and more than one of the judges reported that they were disappointed in the programs' capabilities after their expectations had been raised by interacting with ELIZA in the interviewing process.*

[Egads; see slide 11 —CP]

Human-level or complementary to humans?

From ‘The Challenge and Promise of Artificial Intelligence’, with Peter Norvig and Eric Horvitz (Nov 30, 2011; 51 min. in)

`http://www.kqed.org/a/radiospecials/R201111302000`

Moderator: How far are we away from human intelligence? Just take a gamble.

Peter Norvig: Well, first of all, I object to that, because I think that’s a low target to aim at. [Audience laughs.] Right, ’cause certainly there’s lots of things already that computers are much, much better than people at. [. . .] We want to be able to say, “What is it that humans can’t do that computers can do better?” Now, part of that may be that the computers want to have some basic competency at the human-level in order to interact with us better. But the goal shouldn’t be human-level performance. The goal should be super-human partnership.

The history of NLU

Rough overview of the progression:

1. Pattern-matching using a small set of hand-built rules
2. Pattern-matching with large corpora/databases
3. Linguistically rich, logically-informed systems, usually grounded in highly structured databases
4. Statistical models that use mixtures of the above and strive for wide coverage and error tolerance

Eliza (Weizenbaum 1966)

1. User: You are like my father in some ways.
2. ELIZA: What resemblance do you see
3. User: You are not very aggressive but I think you don't want me to notice that.
4. ELIZA: What makes you think I am not aggressive?
5. User: You don't argue with me.
6. ELIZA: Why do you think I don't argue with you?
7. User: You are afraid of me.
8. ELIZA: Does it please you to believe I am afraid of you?

Rewrite user's reply by (i) swapping 1st and 3rd person, (ii) interpolating stock phrases, and (iii) using scores to rank possible transformations.

STUDENT (Bobrow 1964)

Sample input (p. 54)

"If the number of customer Tom gets is twice the square of 20% of the number of advertisement he runs, and the number of advertisements he runs is 45, what is the number of customers Tom gets?"

Overview of the method

1. Map referential expressions to variables.
2. Use regular expression templates to identify and transform mathematical predications.
3. Let the computer solve the resulting transformed set of statements using a kind of constraint satisfaction.

Corpus-based approaches

Method

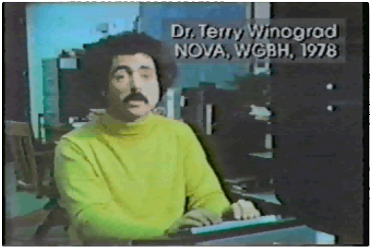
1. Store and index a large collection of texts.
2. (Impose additional structure: tagging, parsing, contextual metadata, etc.)
3. Accept input sentences and match them to relevant sentences using some metric.

For an overview of early examples, see Simmons 1970. This is still how most modern question-answering systems work, including Watson (Ferrucci et al. 2010).

Procedural systems

Winograd (1972): full-grounded system that parses the user's input, maps it to logical form, interprets that logical form in its world, and then tries to take appropriate action.

`http://hci.stanford.edu/winograd/shrdlu/`



One project did succeed. Terry Winograd's program SHRDLU could use English intelligently, but there was a catch: the only subject you could discuss was a micro-world of simulated blocks.

Compare with Watson (Ferrucci et al. 2010):

`http://www-03.ibm.com/innovation/us/watson/`

Logical systems

Focus on mapping text to logical forms and interpreting them in a (usually hand-built) database.

- Chat-80 (Pereira and Warren 13; Warren and Pereira 1982): Prolog program with a database that allowed users to pose queries about world geography. Chat-80 is still distributed, and there is a Python/NLTK module for working with it:
  - `http://www.cis.upenn.edu/~pereira/oldies.html`
  - `http://www.nltk.org/`
- SRI's Core Language Engine (Alshawi et al. 1988): Prolog system for mapping texts to a predicate calculus.

Discourse models

Beyond sentence meaning

In the late 1970s, researchers began focussing on studying discourse-level phenomena like intonational meaning, anaphora and coreference, conversational implicature, presuppositions, connotations, and other phenomena that typically involve going beyond the encoded content to resolve underspecification and extract implicit meaning (Grosz 1977; Hobbs 1979; Sidner 1979; Webber 1979).

Scripts and comprehensive understanding

The Story Understanding Paradigm of Roger Schank and his students and fellow-travelers (Yale, 1969-1988) sought to achieve comprehensive understanding of unconstrained texts. The over-arching idea was to reduce all communication to a set of scripts (Schank 1969, 1977; Schank and Abelson 1977; Lehnert 1977; DeJong 1982).

The statistical revolution

- The statistical revolution of the mid-to-late 1990s profoundly affected all aspects of NLP.
- It was initially detrimental to NLU researchers turned to more constrained problems to explore the new approaches. (But see Ng and Zelle 1997.)
- At present, NLU is enjoying a renaissance. There is a feeling that NLP has sufficient mastery of the statistical techniques and the initial set of problems that it is appropriate to build on them to achieve the goals of NLU.
- As a result, all of the above approaches are being employed, often with modern new mixes: logical approaches (MacCartney 2009), including probabilistic logics (Richardson and Domingos 2006; McCallum et al. 2008), rule-based and script-based interpretation (Lee et al. 2011; Chambers 2011), typically with the backing of very large semi-structured or unstructured databases (like the Web).

Connections with nearby fields

- NLU is a large part of NLP. The boundaries are unclear; even apparently superficial tasks might be influenced by semantic and pragmatic inference.
- (Relationship of our class to CS224N: similar structure and style, significant overlap in content, but substantial shift in emphasis. This one omits many NLP topics, expands on others, and includes some topics that don't get covered at all over there. Also somewhat less math-heavy.)
- NLU has relatively little overlap with formal semantics, though this might change as NLU re-embraces logical approaches
- NLU arguably contains computational semantics; whereas the goal of computational semantics is an understanding of sentence meaning and inference, NLU seeks true utterance understanding.

And now over to Bill . . .

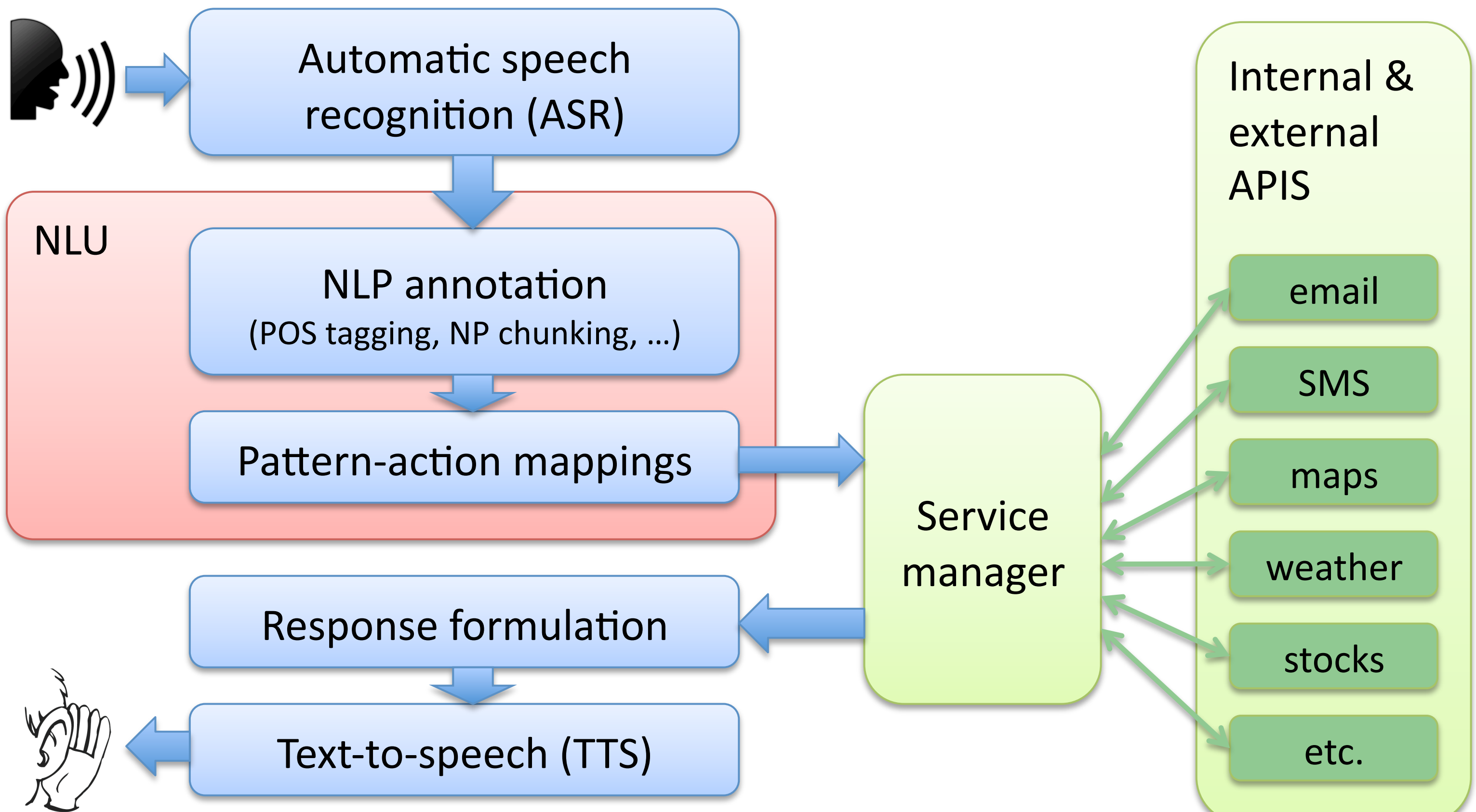
NLU today & tomorrow

- It's an exciting time to be doing NLU!
- In academia, a resurgence of interest in NLU
  (after a long winter)
- Widespread perception that NLU is poised to break through & have huge impact
- Explosion in businesses, products, and services that do NLU
  (or promise to)
- White-hot job market for Stanford grads with mad NLU skillz!

Siri

- The voice-driven personal assistant on your iPhone 4S
- Perhaps the most visible & exciting application of NLU today
- If you believe the marketing hype & the breathless press:
  - Siri represents a major breakthrough in artificial intelligence (AI)
  - Sets the standard for the next generation of UI design

How does Siri work?



What Siri can't do [yet]

Where is A Bug's Life playing in Mountain View?

A Bug's Life is playing at the Century 16 Theater.

When is it playing there?

It's playing at 2pm, 5pm, and 8pm.

OK. I'd like 1 adult and 2 children for the first show.
How much would that cost?

Need domain knowledge, discourse knowledge, world knowledge

QA systems & answer engines

- Question-answering (QA) systems: IBM's Watson
- Answer engines: Wolfram Alpha
- Google and Bing are slipping natural-language QA functionality into search unobtrusively

Sentiment analysis

- Analyze user-generated text content
  - Product reviews: Amazon, Yelp, TripAdvisor, …
  - Social media: Twitter, Facebook, Google+, blogs, …
- Extract coarse- and fine-grained sentiment
  - Coarse-grained: attitudes toward products, brands, personalities, …
  - Fine-grained: salient features of products & attitudes toward them
- Zillions of startups:
  - For marketers: what do people like/dislike about their products?
  - For consumers: which products do other people prefer?
  - For traders: market sentiment from Twitter feeds [Bollen et al. 2011]

Automated trading

- Most financial trading is now done by automated systems ("high-frequency trading", HFT)
- Most HFT strategies rely in part on automated analysis of unstructured data feeds — i.e., natural language text
- You can make vast profits if you can discover and act on market-moving news a few milliseconds faster than your rivals
- Essentially, they're using NLU to predict the markets
- Can go spectacularly wrong: 2008 incident with UAL stock

Business intelligence

- Extracting actionable intelligence from millions of unstructured documents
- Cataphora, H5: legal discovery, compliance, and information management
- Palantir, Quid: intelligence for government & business
- Autonomy: "Meaning Based Computing", acquired by HP in October for $10B

User modeling & personalization

- The better an internet service understands you and what you care about, the more value it can deliver to you
- One of the best ways to figure out what you care about is to understand your natural language
- Example
  - You posted something on your social network about the World Series
  - You recently read a book by David Brooks
  - We recommend an op-ed piece by David Brooks about the business of baseball
- Google, Facebook, Bing, Apple, many others

Why now?

Why is NLU experiencing a resurgence now?

Three key reasons:

- More data
- More compute power
- Better ideas

… and, synergies among these factors.

More data

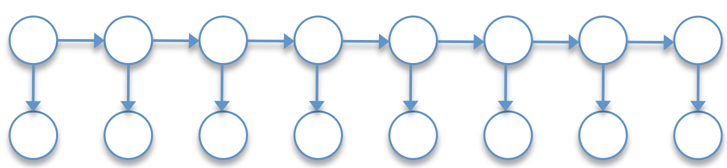
- The explosion of machine-readable natural language text
- Exabytes (10^{18} bytes) of text, doubling every year or two
- 100x growth of web in last ten years
- Web pages, emails, IMs, SMSs, tweets, docs, PDFs, …
- Opportunity — and necessity — to extract meaning

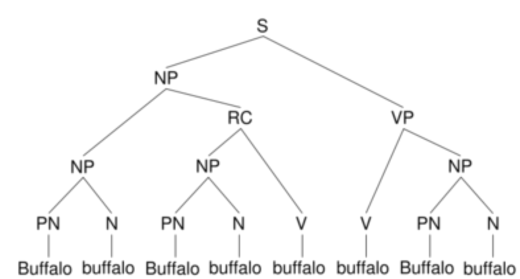
More compute power

- Surprisingly, it's not so much about faster chips
- Rather, it's about massive server farms
- Above all, the software and services to exploit them (e.g., Amazon EC2, Borg, MapReduce, Hadoop, Flume, Pig, Hive, Mahout, …)
- Also: more memory, more disk, & more network bandwidth

Better ideas

- Recent work in NLU builds on advances in statistical NLP
- Often uses proven NLP components as building blocks
- Leverages ideas & techniques developed in statistical NLP
- More sophisticated statistical modeling and machine learning algorithms
- Focus shifting from supervised to unsupervised learning





Three levels of meaning

- One way to carve up NLU is by *levels of meaning*:
  - Meanings of words: lexical semantics
  - Meanings of sentences: compositional / formal semantics
  - Meanings of paragraphs, dialogs, discourses
- That's how we've structured this course, too

Lexical semantics

Why does lexical semantics matter? Because the meanings of sentences are built up from the meaning of words!

There's a large iguana on my sofa.

There's a big lizard on my couch.

In order to determine whether this is a valid inference, you need to know the relation between the meanings of *large* and *big*, *iguana* and *lizard*, *sofa* and *couch*.

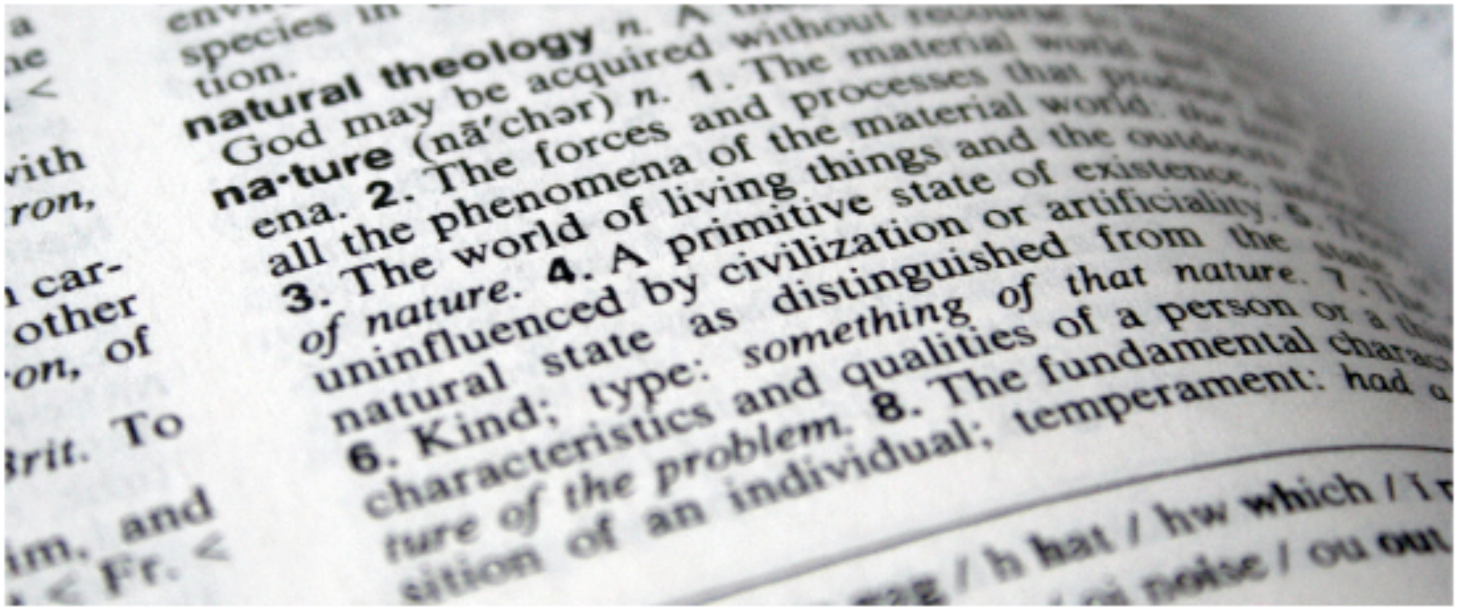
Questions for lexical semantics

What questions do we want to be able to answer?

- How many distinct senses does each word have?
- What's the *meaning* of each word sense?
  (And, how should we represent these meanings?)
- How are word senses related to each other?
  (Synonymy, antonymy, hypernymy, etc.)
- How can we identify the sense of word in context?

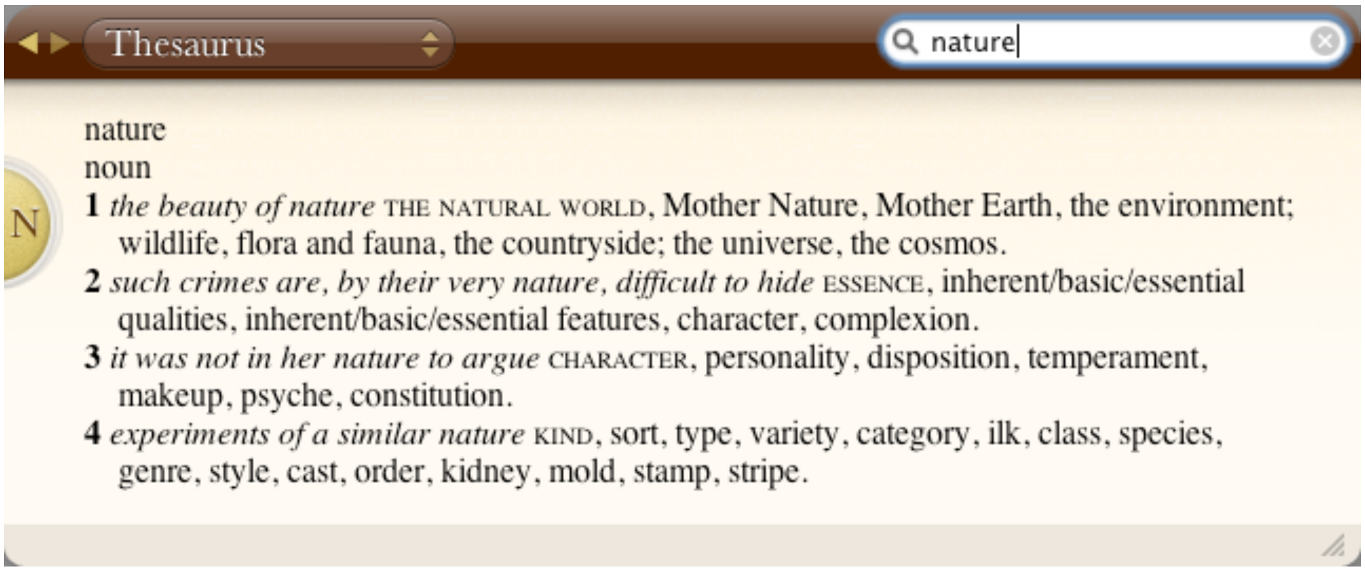
Some of these questions can be answered with the help of *manually constructed resources*.

Dictionaries



Dictionaries enumerate the senses of a word, and indicate what each sense means using glosses and (sometimes) examples.

Thesauri



Thesauri don't directly enumerate senses, but they do provide meanings in terms of synonyms and (sometimes) antonyms.

WordNet

- S: (n) **door** (a swinging or sliding barrier that will close the entrance to a room or building or vehicle) *"he knocked on the door"; "he slammed the door as he left"*
  - *direct hyponym* / *full hyponym*
  - ***part meronym***
    - S: (n) lock (a fastener fitted to a door or drawer to keep it firmly closed)
  - *direct hypernym* / *inherited hypernym* / *sister term*
    - S: (n) movable barrier (a barrier that can be moved to allow passage)
  - *part holonym*
- S: (n) doorway, **door**, room access, threshold (the entrance (the space in a wall) through which you enter or leave a room or building; the space that a door can close) *"he stuck his head in the doorway"*

WordNet enumerate senses and indicates their meanings using glosses and examples, and several types of relations between senses.

Problems

Manually-constructed resources can be very useful.
But, they:

- Require lots of time, money, & expertise to build
  - You're pretty much stuck with what's already available
- Don't cover neologisms: *retweet*, *iPad*, *blog*, ...
- Don't include jargon: *poset*, *LIBOR*, *hypervisor*, …
- Don't cover all languages

Can we *learn* lexical semantics?

Key idea of statistical revolution:

- Don't construct knowledge resources manually — it's too laborious and expensive.
- Instead, automatically *induce* knowledge by discovering statistical regularities in large corpora.

Learning to *distinguish* senses

Lots of work on unsupervised clustering of word senses
[e.g., Schütze 1992, 1998]

Can use features of context (defined in various ways):

The *bass* guitar is a stringed instrument played primarily …
… the string and wind *bass* instruments are usually …

Smallmouth *bass* anglers often use smaller versions …
… coverage of professional tournament *bass* fishing …

Can also rely on parallel translations:

The *bass* guitar is a stringed instrument played primarily …
El *bajo* es un instrumento de cuerdas que se toca principalmente …

… coverage of professional tournament *bass* fishing …
... la cobertura de los profesionales torneo de pesca de la *lubina* ...

Learning what senses *mean*

Well, no one has devised a way to construct conventional dictionary definitions automatically (AFAIK).

But there are other ways to represent meaning …

Inducing word similarity

One way to represent the meaning of a word is by listing words with similar meanings — like a thesaurus!

And word similarity can be induced from context.

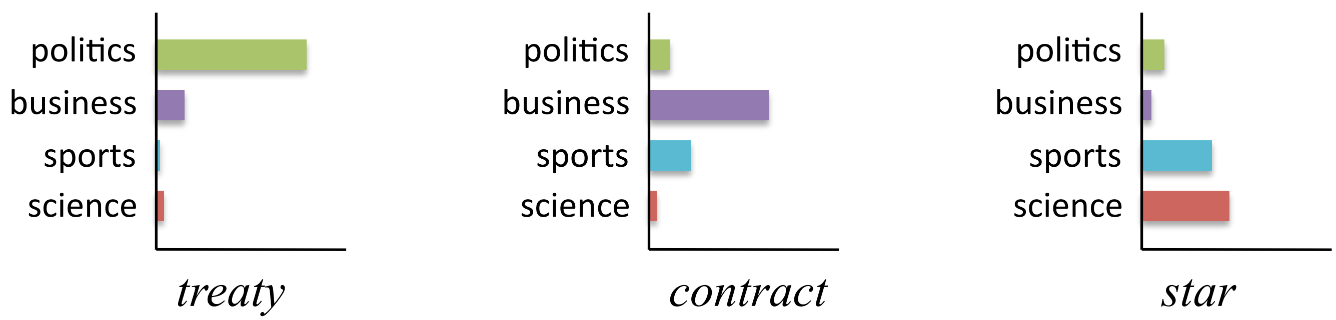
"drank a bottle of *"

Hi I'm Noreen and I once drank a bottle of **wine** in under 4 minutes
SHE DRANK A BOTTLE OF **JACK**?! harleyabshireblondie.
he drank a bottle of **beer** like any man
I topped off some salted peanuts and drank a bottle of **water**
The partygoers drank a bottle of **champagne**.
MR WEST IS DEAD AS A HAMMER HE DRANK A BOTTLE OF **ROGAINE**
aug 29th 2010 i drank a bottle of **Odwalla Pomegranate Juice** and got **...**
The 3 of us drank a bottle of **Naga Viper Sauce ...**
We drank a bottle of **Lemelson pinot noir** from Oregon ($52)
she drank a bottle of **bleach** nearly killing herself, "to clean herself from her wedding"

Vector-space models of meaning

Another way we can represent word meanings is in terms of vector spaces of (relatively) low dimension.

Many possible approaches here. A super-simple one would look at word frequencies in different news topics.



(Real VSMs are considerably more sophisticated.)

Inducing *relations* between senses

We’ve already looked at ways to induce synonymy.

What about other lexical semantic relations?

Hearst [1992, 1998] proposed a clever way to infer hypernymy from large corpora using patterns.

"* such as *"

Legal scholars such as Erwin Chemerinsky, dean of the UC Irvine School of Law, …
… in the form of generalized online communities such as Theglobe.com …
Many animals such as tubeworms, vent mussels, vent crabs, and vent shrimps, …
Remove all MIME encodings, such as content-transfer encoding, …
is abundant in green, leafy vegetables such as collard greens, spinach, and kale.
Should performance enhancing drugs (such as steroids) be accepted in sports?
Some states, such as Colorado, don't require much farming to get a tax-saving …

Word sense disambiguation

- Lexical ambiguity is *pervasive* in natural language text
  - Most of the 1000 most common words are *multiply* ambiguous
  - In WordNet, *set* has 13 senses as noun, 25 as verb, 7 as adjective!
- Creates grievous challenges in MT, IR, QA, text cat, …
- WSD is the task of identifying the sense of a word *in context*
  - Depends on having predefined sense inventory for each word
- Most work in WSD uses the supervised learning paradigm
  - Get annotated training data; define feature representation; apply machine learning algorithms

Between words and sentences

Next, we'll look at two topics which lie somewhere "above" lexical semantics and "below" sentential semantics:

- Relation extraction
- Semantic role labeling

In these tasks, we're extracting meaning from sentence fragments. We're looking at more than individual words, and things like word order and syntax begin to matter, but it's less than full semantic interpretation of complete sentences.

Relation extraction

In relation extraction, we aim to derive structured information (esp. instances of binary relations) from unstructured text:

Born in 1955, Bill Gates is the co-founder, chairman, and former CEO of Microsoft, the world's largest software company, with revenues of $70 billion per year.

⇒

```
yearOfBirth(Bill Gates, 1955)
founder(Bill Gates, Microsoft)
chairman(Bill Gates, Microsoft)
revenues(Microsoft, $70 billion)
```

Can be used to populate databases, to answer questions, …

Semantic role labeling

The same event can be described in many different ways, and *syntactic roles* may not correspond to *semantic roles*.

John broke the window.
John broke the window with a rock.
The rock broke the window.
The window broke.
The window was broken by John.

breaking
AGENT: John
PATIENT: window
INSTRUMENT: rock

Semantic roles can act as a shallow meaning representation that can let us make simple inferences that aren't possible from the pure surface string of words, or even a syntactic parse tree.

Sentential semantics

Further up the value chain, we'd like to do full semantic interpretation of complete sentences.

This goal was a major focus of early work in NLU.

STUDENT [Bobrow 1964]
If the number of customers Tom gets is twice the square of …

SHRDLU [Winograd 1971]:
Move the red block on top of the smaller green one.

CHAT-80 [Warren & Pereira 1982]
What are the capitals of the countries bordering the Baltic?

But these systems were rule-driven and closed-domain.

Compositional semantics

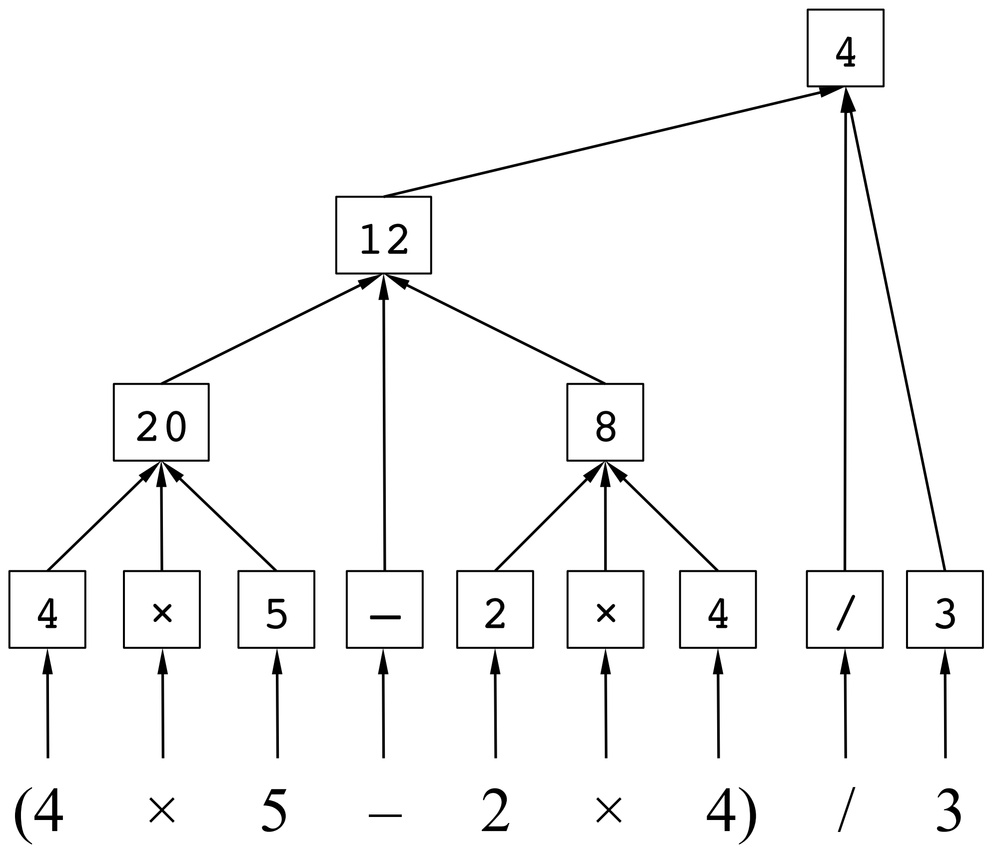
Sentential semantics is often known as *compositional semantics*, because of this big idea:

The Principle of Compositionality

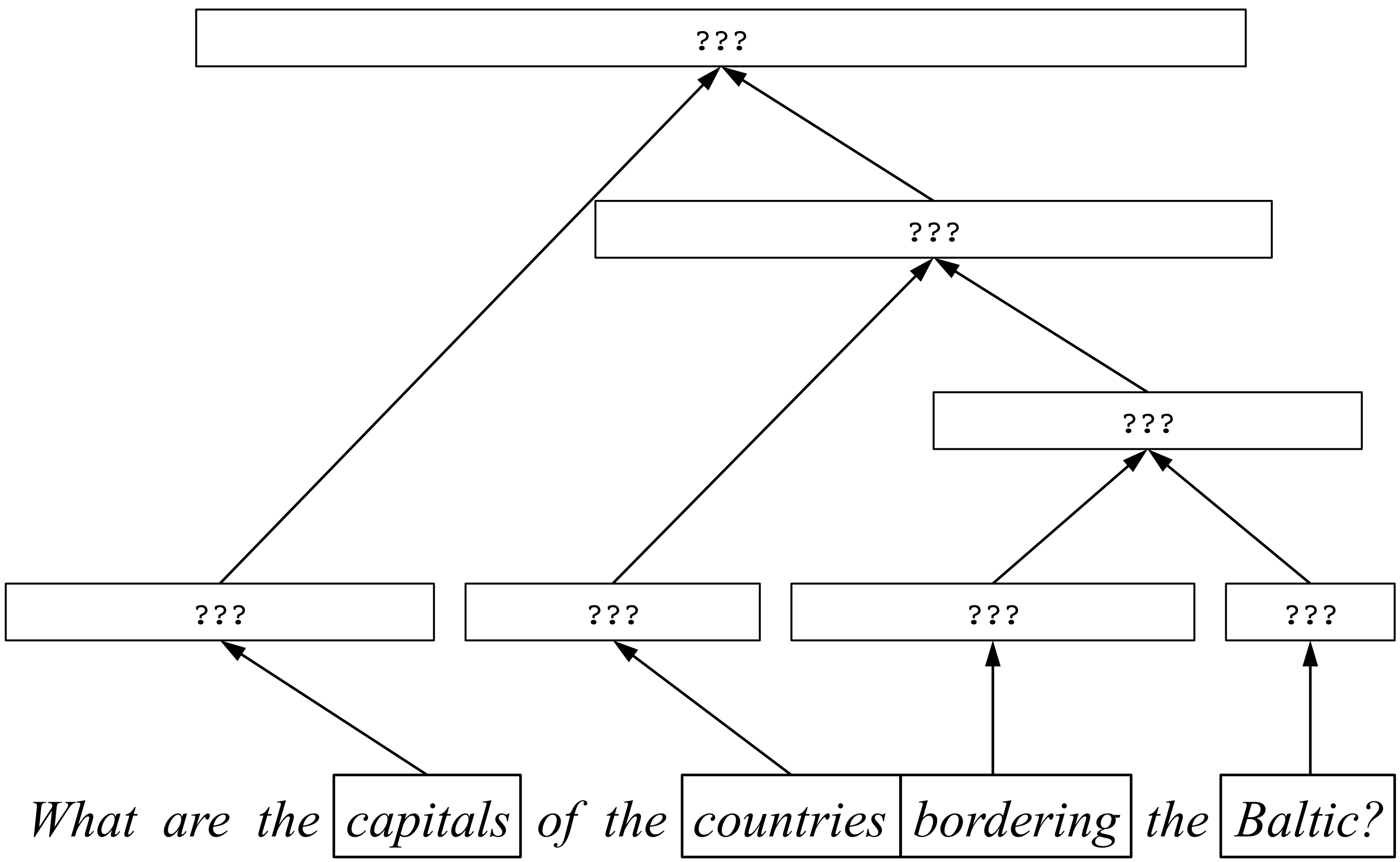
The meaning of the whole
is a function of the meanings of the parts
and the way they are combined.

(Widely attributed to Frege, though anticipated by Yāska and Plato.)

Compositionality in arithmetic



Compositionality in language



Formal semantics

But how do we represent meanings during composition?

Indeed, when we do full semantic interpretation of a sentence, what kind of final output are we aiming at?

One answer: expressions of formal logic.

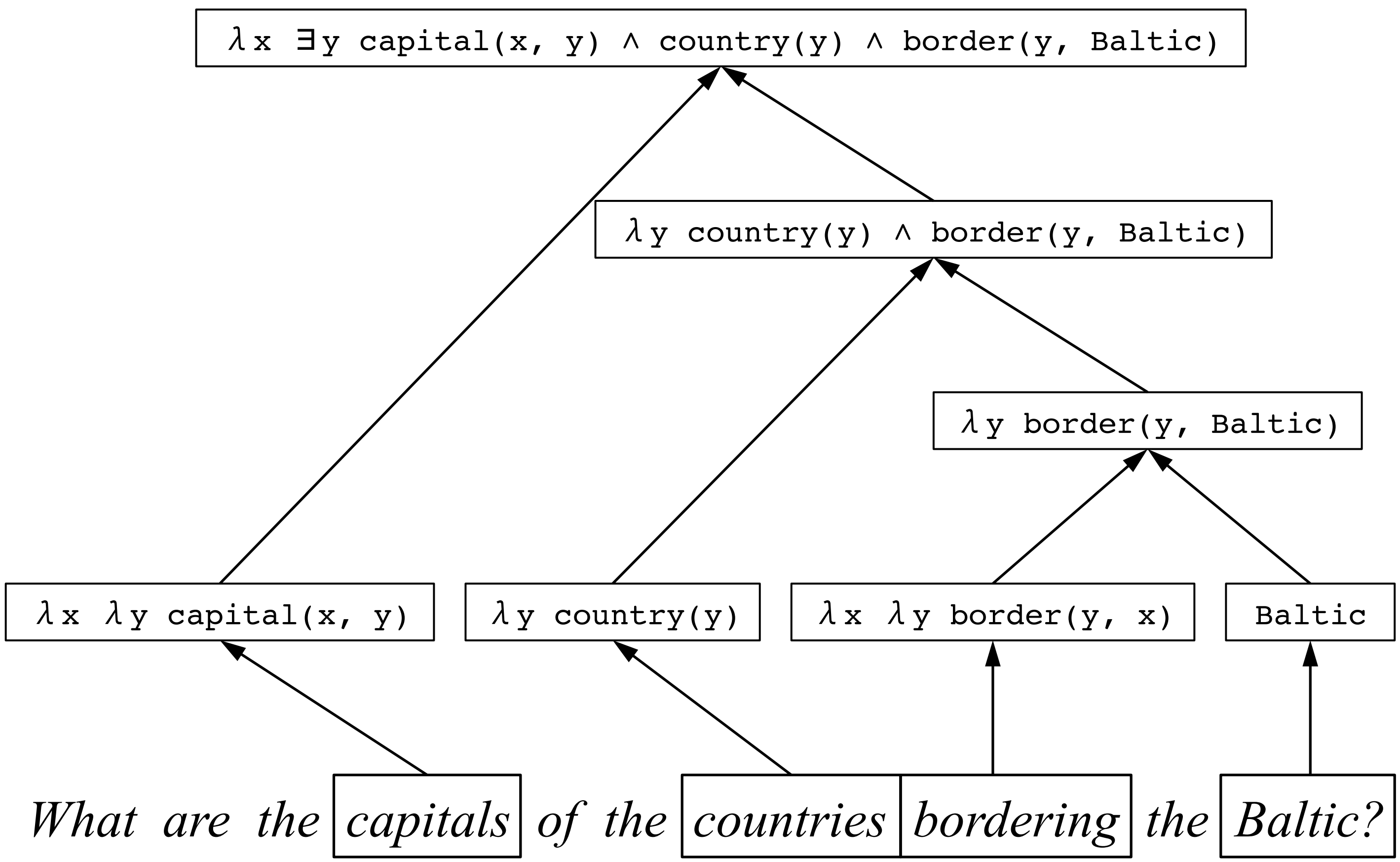
Or more generally: expressions of lambda calculus.

$$\lambda x\ \exists y\ \texttt{capital}(x, y) \wedge \texttt{country}(y) \wedge \texttt{border}(y, \texttt{Baltic})$$

What are the capitals of the countries bordering the Baltic?

Hence, compositional semantics ⇒ formal semantics

Montague semantics



Can we use MT techniques?

Maybe we can *learn* how to do full semantic interpretation by borrowing the highly successful strategies of statistical machine translation?

Texas borders Kansas.	`border(Texas, Kansas)`
What is the capital of Utah?	λx `capital(x, Utah)`
What states border Maine?	λx `state(x)` $\wedge$ `border(x, Maine)`
Maine does not border Utah.	$\neg$`border(Maine, Utah)`
Every state has a capital.	$\forall$x `state(x)` $\rightarrow$ $\exists$y `capital(y, x)`
Austin is the capital of Texas.	`capital(Austin, Texas)`

But where will we get the parallel corpora (training data)?

Natural language inference

- One test of understanding is the ability to draw inferences
- The Recognizing Textual Entailment (RTE) challenges

  P. *Twenty-five of the dead were members of the law enforcement agencies and the rest of the 67 were civilians.*
  H. *25 of the dead were civilians.*

- A broad variety of approaches have been explored
  - Full semantic interpretation ⇒ theorem provers
  - Measuring lexical similarity between bags of words
  - Intermediate approaches, e.g. shallow semantic interpretation
- A closely related task: recognizing paraphrases
  - *X causes Y* ≈ *X can lead to Y* ≈ *Y is triggered by X* ≈ *Y is attributed to X* ≈ ...

Discourse and context

In the discourse and context unit, we will go (way) beyond the encoded semantic content to try to capture something more like utterance meaning. The specific topics we've chosen:

- Hedges: the methods speakers employ for indicating their confidence and commitment
- Sentiment analysis: how information about attitudes, perspectives, and emotions is conveyed.
- Textual coherence: the often invisible glue that binds sentences together
- Discourse and dialogue: interactional aspects of language

These topics do not exhaust the space, but we think they provide a good sample of the problems and approaches, which should make it possible for you to pursue other areas (coreference, presupposition, connotation, speech acts, . . .) if you wish.

References I

Allen, James F. 1987. *Natural Language Understanding*. Menlo Park, CA: The Benjamin/Cummings Publishing Company, Inc.

Alshawi; H.; Carter; D. M.; van Eijck; J.; Moore; R. C.; Moran; D. B.; Pereira; F. C. N.; Pulman; S. G.; Smith; and A. G. 1988. Interim report on the SRI core language engine. Technical Report CCSRC-5, SRI International, Cambridge Research Centre, Cambridge, England.

Bobrow, Daniel G. 1964. *Natural Language Input for a Computer Problem Solving System*. Ph.D. thesis, Rensselaer Polytechnic Institute.

Chalmers, David. 1997. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.

Chambers, Nathanael. 2011. *Inducing Event Schemas and their Participants from Unlabeled Text*. Ph.D. thesis, Stanford.

Churchland, Paul M. and Patricia Smith Churchland. 1990. Could a machine think? *Scientific American* 32–37.

Cole, David. 2009. The Chinese room argument. In Edward N. Zalta, ed., *Stanford Encyclopedia of Philosophy (Winter 2003 Edition)*. CSLI. URL `http://plato.stanford.edu/entries/chinese-room/`.

References II

DeJong, Gerald F. 1982. An overview of the FRUMP system. In Wendy G. Lehnert and Martin H. Ringle, eds., *Strategies for Natural Language Processing*, 149–176. Hillsdale, NJ: Erlbaum.

Ferrucci, David; Eric Brown; Jennifer Chu-Carroll; James Fan; David Gondek; Aditya A. Kalyanpur; Adam Lally; J. William Murdock; Eric Nyberg; John Prager; Nico Schlaefer; and Chris Welty. 2010. Building Watson: An overview of the deepQA project. *AI Magazine* 31(3):59–79.

Grosz, Barbara J. 1977. *The Representation and Use of Focus in Dialogue Understanding*. Ph.D. thesis, UC Berkeley.

Hobbs, Jerry R. 1979. Coherence and coreference. *Cognitive Science* 3(1):67–90.

Lee, Heeyoung; Yves Peirsman; Angel Chang; Nathanael Chambers; Mihai Surdeanu; and Daniel Jurafsky. 2011. Stanford's multi-pass sieve coreference resolution system at the CoNLL-2011 shared task. In *Proceedings of the CoNLL-2011 Shared Task*.

Lehnert, Wendy. 1977. Human and computational question answering. *Cognitive Science* 1(1):47–73.

MacCartney, Bill. 2009. *Natural Language Inference*. Ph.D. thesis, Stanford University.

References III

McCallum, Andrew; Khashayar Rohanimanesh; Michael Wick; Karl Schultz; and Sameer Singh. 2008. FACTORIE: Efficient probabilistic programming for relational factor graphs via imperative declarations of structure, inference and learning. In *NIPS Workshop on Probabilistic Programming*.

Ng, Hwee Tou and John Zelle. 1997. Corpus-based approaches to semantic interpretation in natural language processing. *AI Magazine* 18(4):45–64.

Penrose, Roger. 1990. *The Emperor's New Mind: Concerning Computers, Minds, and The Laws of Physics*. Oxford University Press.

Pereira, Fernando C. N. and David H. D. Warren. 13. Definite clause grammars for language analysis — a survey of the formalism and a comparison with augmented transition networks. *Artificial Intelligence* 3(231–278).

Richardson, Matthew and Pedro Domingos. 2006. Markov logic networks. *Machine Learning* 62(1–2):107–136.

Schank, Roger. 1969. A conceptual dependency representation for a computer-oriented semantics. Technical Report 83, Stanford Computer Science.

Schank, Roger. 1977. Sam–a story understander. Technical Report 43, Yale University Computer Science, New Haven, CT.

References IV

Schank, Roger and Robert P. Abelson. 1977. *Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures*. Hillsdale, NJ: Erlbaum.

Searle, John R. 1980. Minds, brains and programs. *Behavioral and Brain Sciences* 3(3):417–457.

Searle, John R. 1990. Is the brain's mind a computer program? *Scientific American* 26–31.

Shieber, Stuart. 1994. Lessons from a restricted Turing test. *Communications of the ACM* 37(6):70–78.

Sidner, Candace L. 1979. *Towards a Computational Theory of Definite Anaphora*. Ph.D. thesis, City University of New York.

Simmons, Robert F. 1970. Natural language question-answering systems: 1969. *Computational Linguistics* 13(1):15–30.

Turing, Alan M. 1950. Computing machinery and intelligence. *Mind* 59(236):433–460.

Warren, David H. D. and Fernando C. N. Pereira. 1982. An efficient easily adaptable system for interpreting natural language queries. *American Journal of Computational Linguistics* 8(3–4):110–122.

References V

Webber, Bonnie Lynn. 1979. *A Formal Approach to Discourse Anaphora*. New York: Garland.

Weizenbaum, Joseph. 1966. ELIZA — a computer program for the study of natural language communication between man and machine. *Communications of the ACM* 9(1):36–45.

Winograd, Terry. 1972. Understanding natural language. *Cognitive Psychology* 3(1):1–191.