Lecture 1: Overview

CS 181
Spring 2016
Kim Bruce

Some slide content taken from Unger and Michaelis

Questions for Today

- What is computational semantics?
  - What is formal semantics?
- Why should you care?
- What do we need to know to do it?
- What are you going to do in this course?

Computational Semantics is ...

- ... the study of how to *automate* the process of *constructing* and *reasoning* with meaning representations of natural language expressions.
  Wikipedia
- ... the art and science of *computing* meanings for the expressions of a natural language.
  Van Eijck & Unger

Computational Semantics is

- ... not the same as natural language processing
  - Depends on machine learning and other probability and statistics-based techniques
  - More AI, with less dependence on linguistics theory
- Computational semantics depends on:
  - logic
  - lambda calculus (to describe functions), finite automata
  - declarative-style programming languages
    - traditionally PROLOG, but we use functional languages.
- Two approaches are complementary

Demand for Language Comprehension/Production

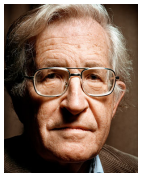
- Audio:
  - Recognition: Siri, dictation, voice-mail systems
  - Speech synthesis for warnings, directions, etc.
- Textual: Google search
  - Tell me all the science classes that meet after 11 a.m. that have a professor that has won the Wig teaching award and that are designed for non-majors.

Natural Language Hard!!

- Many things make natural language hard.
- Ambiguity in structure, word meaning:
  - I saw the girl in the park *with a telescope*
  - I made her *duck*
- Ambiguity in reference:
  - The police refused the students a permit because *they* feared violence.
  - The police refused the students a permit because *they* advocated violence.

Formal Study of Language

- Noam Chomsky, 1957, proposed to identify a language with the set of all correct sentences of that language.
- The ability of language users to recognize members of this set is called *competence*.
- Goal: Build a model of our linguistic knowledge, abstracting from language *performance* (speech disabilities, memory limitations, errors, etc).
- Such a model is called grammar.



Linguistics subfields

- **Phonology** explores what the smallest meaning-distinguishing units (sounds) are and how they are combined into the smallest meaning-carrying units (morphemes).
- **Morphology** is concerned with how morphemes are combined into words.
- **Syntax** studies how words are combined into phrases and sentences.

Linguistics subfields

- **Semantics** investigates meanings of basic expressions and how meaning is assigned to complex expressions based on the meaning of simpler expressions and syntactic structure.
- **Pragmatics** studies the ways in which context contributes to meaning. It encompasses speech act theory, conversational implicature, talk in interaction and other approaches to language behavior.

Focus on Meaning

- ... based on the structure imposed by a grammar.
- Must develop grammar that will enable us to:
  - determine the structure of well-formed phrases & sentences
  - determine the constituents of complex expressions as well as their internal structure
  - assign appropriate meanings to syntactically well-formed expressions, on the basis of their structure.

Trichotomy of Aspects

- **Syntax**: study of strings and structure imposed on them by grammars generating them.
- **Semantics**: study of relationship between strings and their meanings (i.e., the extra-linguistic structure they are about)
- **Pragmatics**: study of the use of meaningful strings to communicate about extra-linguistic structure in an interaction process between users of the language.

Requires

- Description of structure (formal grammars)
- Account of meaning of atomic components (words — or sometimes parts of words)
- Mechanism to build meaning of more complex phrases from those of the atomic components and the structure of the phrases.

What is meaning?

- Lexical Semantics
  - What are the meanings of words?
- Compositional Semantics
  - What are the meanings of phrases and sentences?
  - How are the meanings of phrases and sentences derived from the meanings of words?

Lexical semantics

- Relation between words and the world
- Main concern will be with formal representation
- Leave subtleties to philosophers:
  - http://plato.stanford.edu/entries/word-meaning/

Compositional Semantics

- What are the meanings of phrases and sentences?
  - *Wittgenstein*: To know the meaning of a sentence is to know how the world would have to be for the sentence to be true.
  - E.g.
    - It is Wednesday morning.
    - Pomona College is a great school.
    - The king of France is bald. (????)

Compositional Semantics

- How are the meanings of phrases and sentences derived from the meanings of words?
  - *The meaning of a complex expression is a function of the meanings of its parts and of the syntactic rules by which they are combined.*
  - Reflects our approach rather than statement about how humans understand language.

Our Focus

- Compositional Semantics.
  - Assume lexical semantics and focus on how to build up meanings.
  - However lexical is construed, can use compositional approach to build meanings of more complex phrases and sentences.

Formal Semantics

- Study of meaning based on logic
- Use tools to
  - formally describe language
  - compute meanings in a systematic way

Birth of Formal Semantics

- Roots in logic with Aristotle, Frege, Tarski, ...
- Richard Montague:
  - *There is in my opinion no important theoretical difference between natural languages and the artificial languages of logicians; indeed, I consider it possible to comprehend the syntax and semantics of both kinds of languages within a single natural and mathematically precise theory.*
  - 1970a: 'English as a formal language'
  - 1970b: 'Universal grammar'
  - 1973: 'The proper treatment of quantification in ordinary English' (PTQ)

Impact

"To us, the revolutionary idea in Montague's PTQ paper (and earlier papers) is the claim that natural language is not impossibly incoherent, as his teacher Tarski had led us to believe, but that large portions of its semantics can be treated by combining known tools from logic, tools like functions of finite type, the λ-calculus, generalized quantifiers, tense and modal logic, and all the rest."

"Montague had a certain job that he wanted to do and used whatever tools he had at hand to do it. If the product looks a bit like a Rube Goldberg machine, well, at least it works pretty well." (Barwise and Cooper 1981)

Our Approach

- Introduce ideas of formal semantics
  - with focus on how to implement them
- Combine ideas from:
  - Math (sets and functions)
  - Mathematical logic
    - Meaning of sentence w.r.t. formal model
  - Lambda calculus
- Implement them in Haskell

Next Few Weeks

- Sets, Relations, & Functions
- Lambda Calculus
- Haskell
- Logic as a formal language
- Semantics of natural language, implemented
- Then selection of topics in computational semantics.
  - Goal to provide enough foundation that can learn more on own.

Structure

- Build up foundations
  - You do weekly homework, text and programs
- Midterm — after spring break
- Final Project
  - Based on material in course
  - Written & oral components
  - Work individually or in pairs

Syllabus is a guess!

Why Programming is a Good Medium for Expressing Poorly Understood and Sloppily Formulated Ideas

We have been programming universal computers for about 50 years. Programming provides us with new tools to express ourselves. We now have intellectual tools to describe "how to" as well as "what is". This is a profound transformation: it is a revolution in the way we think and in the way we express what we think. For example, one often hears a student or teacher complain that the student knows the "theory" of some subject but cannot effectively solve problems.We should not be surprised: the student has no formal way to learn technique. We expect the student to learn to solve problems by an inefficient process: the student watches the teacher solve a few problems, hoping to abstract the general procedures from the teacher's behavior on particular examples. The student is never given any instructions on how to abstract from examples, nor is the student given any language for expressing what has been learned. It is hard to learn what one cannot express. But now we can express it!
Expressing methodology in a computer language forces it to be unambiguous and computationally effective. The task of formulating a method as a computer-executable program and debugging that program is a powerful exercise in the learning process. The programmer expresses his/her poorly understood or sloppily formulated idea in a precise way, so that it becomes clear what is poorly understood or sloppily formulated. Also, once formalized procedurally, a mathematical idea becomes a tool that can be used directly to compute results.
Gerry Sussman, 2005

Questions?