

Introduction to Deep Processing Techniques for NLP

Deep Processing Techniques for NLP

Ling 571

January 4, 2016

Gina-Anne Levow

Roadmap

- Motivation:
 - Applications
- Language and Thought
- Knowledge of Language
 - Cross-cutting themes
 - Ambiguity, Evaluation, & Multi-linguality
- Course Overview

Motivation: Applications

- Applications of Speech and Language Processing
 - Call routing
 - Information retrieval
 - Question-answering
 - Machine translation
 - Dialog systems
 - Spell- , Grammar- checking
 - Sentiment Analysis
 - Information extraction....

Building on Many Fields

- Linguistics: Morphology, phonology, syntax, semantics,...
- Psychology: Reasoning, mental representations
- Formal logic
- Philosophy (of language)
- Theory of Computation: Automata,...
- Artificial Intelligence: Search, Reasoning, Knowledge representation, Machine learning, Pattern matching
- Probability..

Language & Intelligence

- Turing Test: (1950) – Operationalize intelligence
 - Two contestants: human, computer
 - Judge: human
 - Test: Interact via text questions
 - Question: Can you tell which contestant is human?
- Crucially requires language use and understanding

Limitations of Turing Test

- ELIZA (Weizenbaum 1966)
 - Simulates Rogerian therapist
 - User: You are like my father in some ways
 - ELIZA: WHAT RESEMBLANCE DO YOU SEE
 - User: You are not very aggressive
 - ELIZA: WHAT MAKES YOU THINK I AM NOT AGGRESSIVE...
 - Passes the Turing Test!! (sort of)
 - “You can fool some of the people....”
- Simple pattern matching technique
- True understanding requires deeper analysis & processing

Turing Test Revived

- “On the web, no one knows you’re a....”
 - Problem: ‘bots’
 - Automated agents swamp services
 - Challenge: Prove you’re human
- Test: Something human can do, ‘bot can’t
- Solution: CAPTCHAs
 - “Completely Automated Public Turing Test To Tell Computers and Humans Apart”
 - Initially: distorted images: easy for human; hard for ‘bot’
 - Driven by perception
 - Drives improvements in AI – vision, audio, OCR
 - “Arms race”: better systems, harder CAPTCHAs
 - Images, word problems, etc

Knowledge of Language

- What does HAL (of 2001, A Space Odyssey) need to know to converse?
- *Dave: Open the pod bay doors, HAL.*
- *HAL: I'm sorry, Dave. I'm afraid I can't do that.*

Knowledge of Language

- What does HAL (of 2001, A Space Odyssey) need to know to converse?
 - *Dave: Open the pod bay doors, HAL.*
 - *HAL: I'm sorry, Dave. I'm afraid I can't do that.*
- Phonetics & Phonology (Ling 450/550)
 - Sounds of a language, acoustics
 - Legal sound sequences in words

Knowledge of Language

- What does HAL (of 2001, A Space Odyssey) need to know to converse?
 - *Dave: Open the pod bay doors, HAL.*
 - *HAL: I'm sorry, Dave. I'm afraid I can't do that.*
- Morphology (Ling 570)
 - Recognize, produce variation in word forms
 - Singular vs. plural: Door + sg: → door; Door + plural → doors
 - Verb inflection: Be + 1st person, sg, present → am

Knowledge of Language

- What does HAL (of 2001, A Space Odyssey) need to know to converse?
 - *Dave: Open the pod bay doors, HAL.*
 - *HAL: I'm sorry, Dave. I'm afraid I can't do that.*
- Part-of-speech tagging (Ling 570)
 - Identify word use in sentence
 - Bay (Noun) --- Not verb, adjective

Knowledge of Language

- What does HAL (of 2001, A Space Odyssey) need to know to converse?
 - *Dave: Open the pod bay doors, HAL.*
 - *HAL: I'm sorry, Dave. I'm afraid I can't do that.*
- Syntax
 - (Ling 566: analysis;
 - Ling 570 – chunking; Ling 571 – parsing)
 - Order and group words in sentence
 - I'm I do , sorry that afraid Dave I can't.

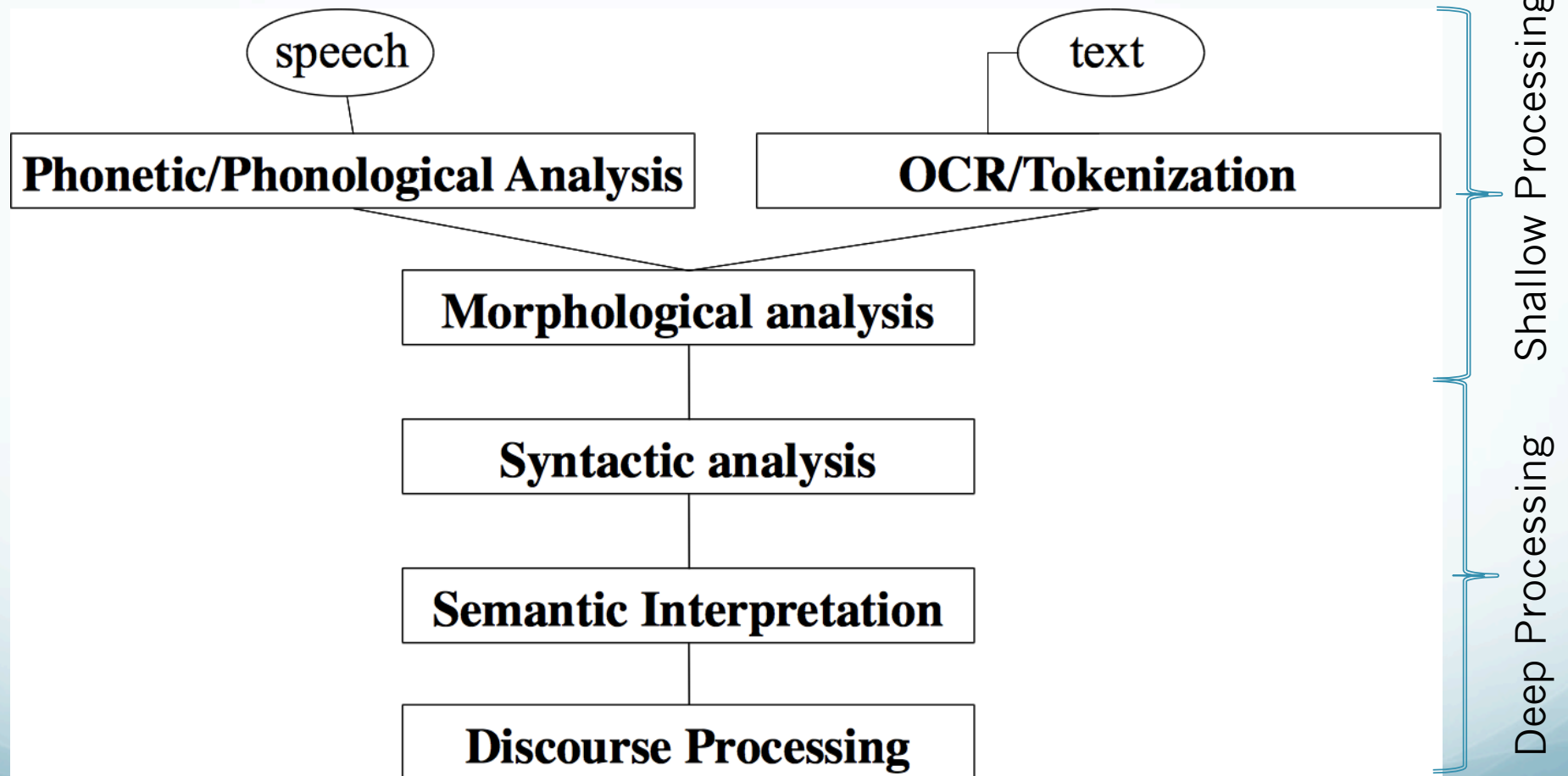
Knowledge of Language

- What does HAL (of 2001, A Space Odyssey) need to know to converse?
 - *Dave: Open the pod bay doors, HAL.*
 - *HAL: I'm sorry, Dave. I'm afraid I can't do that.*
- Semantics (Ling 571)
 - Word meaning:
 - individual (lexical), combined (compositional)
 - ‘Open’ : AGENT **cause** THEME to become *open*;
 - ‘pod bay doors’ : (pod bay) doors

Knowledge of Language

- What does HAL (of 2001, A Space Odyssey) need to know to converse?
 - *Dave: Open the pod bay doors, HAL.* (request)
 - *HAL: I'm sorry, Dave. I'm afraid I can't do that.* (statement)
- Pragmatics/Discourse/Dialogue (Ling 571)
 - Interpret utterances in context
 - Speech act (request, statement)
 - Reference resolution: I = HAL; that = 'open doors'
 - Politeness: I'm sorry, I'm afraid I can't

Language Processing Pipeline



Shallow vs Deep Processing

- Shallow processing (Ling 570)
 - Usually relies on surface forms (e.g., words)
 - Less elaborate linguistics representations
 - E.g. HMM POS-tagging; FST morphology
- Deep processing (Ling 571)
 - Relies on more elaborate linguistic representations
 - Deep syntactic analysis (Parsing)
 - Rich spoken language understanding (NLU)

Cross-cutting Themes

- Ambiguity
 - How can we select among alternative analyses?
- Evaluation
 - How well does this approach perform:
 - On a standard data set?
 - When incorporated into a full system?
- Multi-linguality
 - Can we apply this approach to other languages?
 - How much do we have to modify it to do so?

Ambiguity

- “I made her duck”
- Means....
 - I caused her to duck down
 - I made the (carved) duck she has
 - I cooked duck for her
 - I cooked the duck she owned
 - I magically turned her into a duck

Ambiguity: POS

- “I made her duck”

- Means....

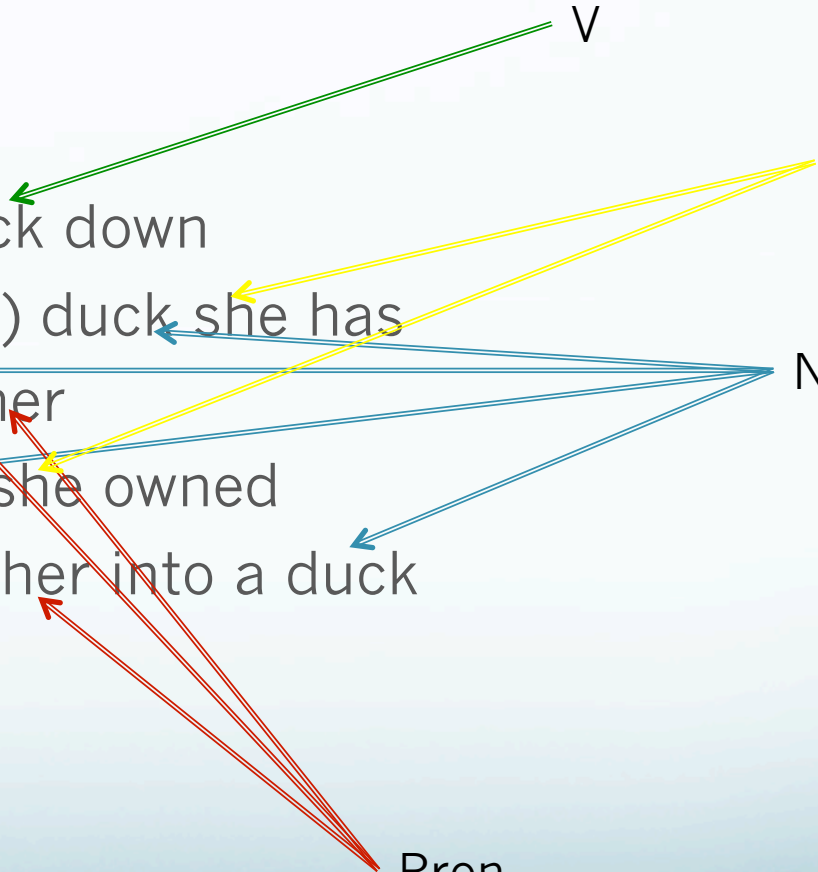
- I caused her to duck down
- I made the (carved) duck she has
- I cooked duck for her
- I cooked the duck she owned
- I magically turned her into a duck

V

Poss

N

Pron



Ambiguity: Syntax

- “I made her duck”
- Means....
 - I made the (carved) duck she has
 - ((VP (V made) (NP (POSS her) (N duck))))
 - I cooked duck for her
 - ((VP (V made) (NP (PRON her)) (NP (N (duck))))

Ambiguity: Semantics

- “I made her duck”
- Means....
 - I caused her to duck down
 - Make: AG **cause** TH to do sth
 - I cooked duck for her
 - Make: AG cook TH for REC
 - I cooked the duck she owned
 - Make: AG cook TH
 - I magically turned her into a duck
 - Duck: animal
 - I made the (carved) duck she has
 - Duck: duck-shaped figurine

Ambiguity

- Pervasive
- Pernicious
- Particularly challenging for computational systems
- Problem we will return to again and again in class

Course Information

- <http://courses.washington.edu/ling571>

Syntax

Ling 571

Deep Processing Techniques for Natural Language Processing

January 4, 2016

Roadmap

- Sentence Structure
 - Motivation: More than a bag of words
 - Constituency
- Representation:
 - Context-free grammars
 - Formal definition of context free grammars
 - Chomsky hierarchy
 - Why not finite state?
 - Aside: Context-sensitivity

More than a Bag of Words

- Sentences are structured:
 - Impacts meaning:
 - Dog bites man vs man bites dog
 - Impacts acceptability:
 - Dog man bites

Constituency

- Constituents: basic units of sentences
 - word or group of words that acts as a single unit
- Phrases:
 - Noun phrase (NP), verb phrase (VP), prepositional phrase (PP), etc
 - Single unit: type determined by head (e.g., $N \rightarrow NP$)

Constituency

- How can we tell what units are constituents?
- On September seventeenth, I'd like to fly from Sea-Tac Airport to Denver.
 - September seventeenth
 - On September seventeenth
 - Sea-Tac Airport
 - from Sea-Tac Airport

Constituency Testing

- Appear in similar contexts
 - PPs, NPs, PPs
- Preposed or Postposed constructions
 - On September seventeenth, I'd like to fly from Sea-Tac Airport to Denver.
 - I'd like to fly from Sea-Tac Airport to Denver on September seventeenth.
 - Must move as unit
 - *On I'd like to fly September seventeenth from Sea-Tac Airport to Denver.
 - *I'd like to fly on September from Sea-Tac airport to Denver seventeenth.

Representing Sentence Structure

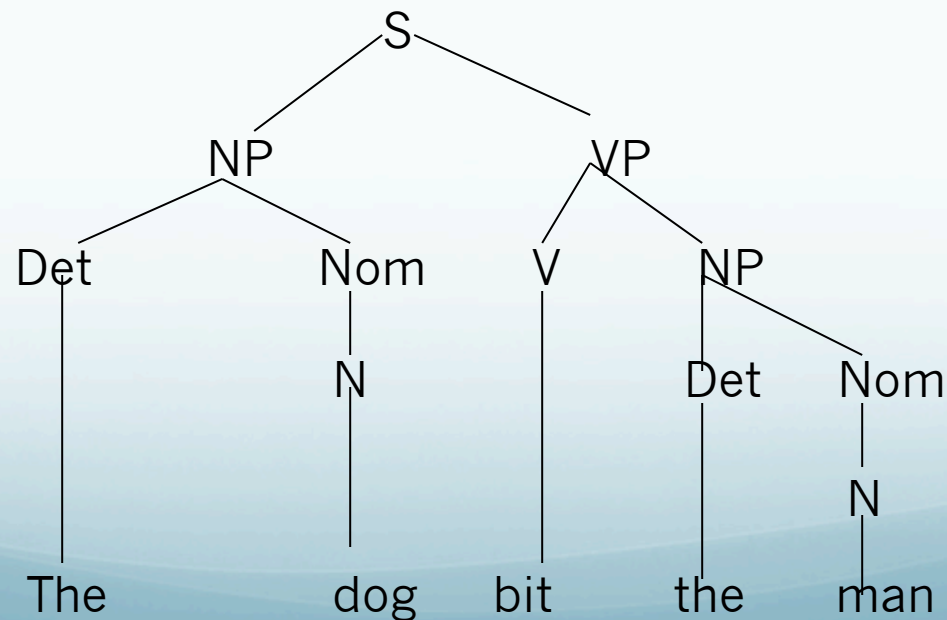
- Captures constituent structure
 - Basic units
 - Phrases
 - Subcategorization
 - Argument structure
 - Components expected by verbs
- Hierarchical

Representation: Context-free Grammars

- CFGs: 4-tuple
 - A set of terminal symbols: Σ
 - A set of non-terminal symbols: N
 - A set of productions P : of the form $A \rightarrow \alpha$
 - Where A is a non-terminal and α in $(\Sigma \cup N)^*$
 - A designated start symbol S
- $L = \{w \mid w \text{ in } \Sigma^* \text{ and } S \Rightarrow^* w\}$
 - Where $S \Rightarrow^* w$ means S derives w by some seq

Representation: Context-free Grammars

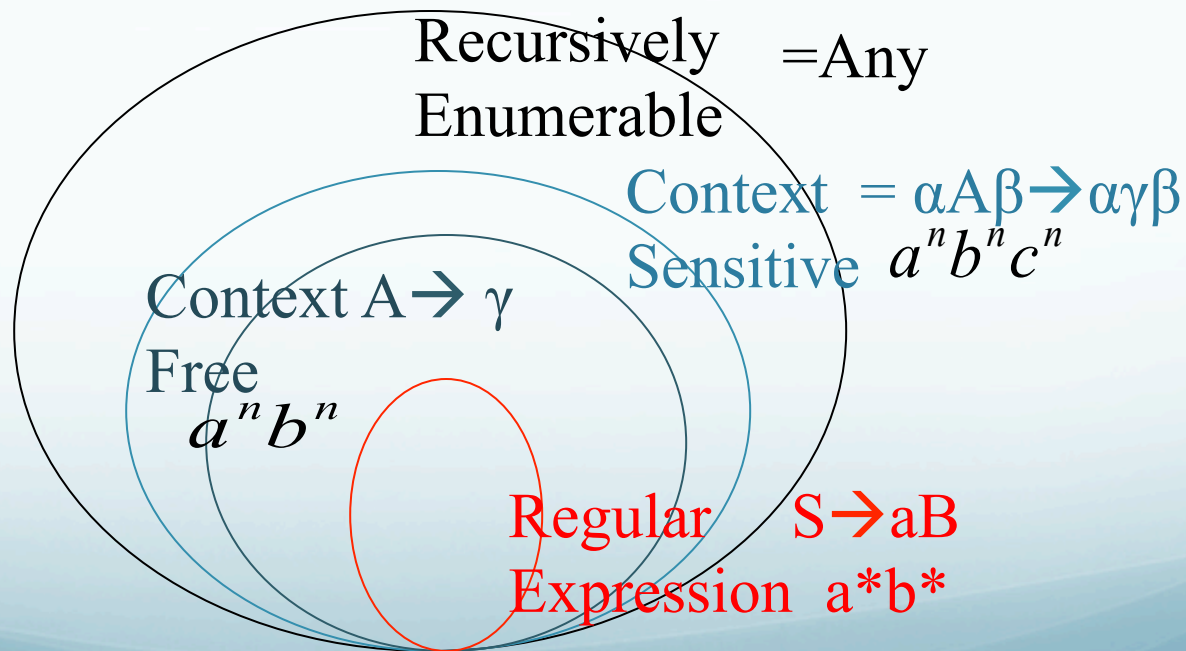
- Partial example
 - Σ : the, cat, dog, bit, bites, man
 - N: NP, VP, Nom, Det, V, N, Adj
 - P: $S \rightarrow NP VP$; $NP \rightarrow Det Nom$; $Nom \rightarrow N Nom | N$;
 $VP \rightarrow V NP$, $N \rightarrow cat$, $N \rightarrow dog$, $N \rightarrow man$, $Det \rightarrow the$, $V \rightarrow bit$,
 $V \rightarrow bites$
 - S



Sentence-level Knowledge: Syntax

- Different models of language
 - Specify the expressive power of a formal language

Chomsky
Hierarchy



Representing Sentence Structure

- Why not just Finite State Models?
 - Cannot describe some grammatical phenomena
 - Inadequate expressiveness to capture generalization
- Center embedding
 - Finite State: $A \rightarrow w^*; A \rightarrow w^*B$
 - Context-Free: $A \Rightarrow \alpha A \beta$
 - Allows recursion
 - The luggage arrived.
 - The luggage that the passengers checked arrived.
 - The luggage that the passengers that the storm delayed checked arrived.

Parsing Goals

- Accepting:
 - Legal string in language?
 - Formally: rigid
 - Practically: degrees of acceptability
- Analysis
 - What structure produced the string?
 - Produce one (or all) parse trees for the string
- Will develop techniques to produce analyses of sentences
 - Rigidly accept (with analysis) or reject
 - Produce varying degrees of acceptability