

Natural Language Processing



- Stages in understanding natural language
- Why it's hard

14-11-2011

(Some slides adapted from presentations by Dan Jurafsky and Bonnie Dorr.)

Computer Speech and Language Processing

- What is it?
 - ▶ Getting computers to perform useful tasks involving human languages whether for:
 - Enabling human-machine communication
 - Improving human-human communication
 - Doing things with spoken or textual material
 - ▶ Examples:
 - Spoken Conversational Agents
 - Machine Translation
 - Question Answering
 - ...

Conversational Agents: An Example

Consider the following interaction with HAL the computer from *2001: A Space Odyssey*

Dave: Open the pod bay doors, Hal.

HAL: *I'm sorry Dave, I'm afraid I can't do that.*

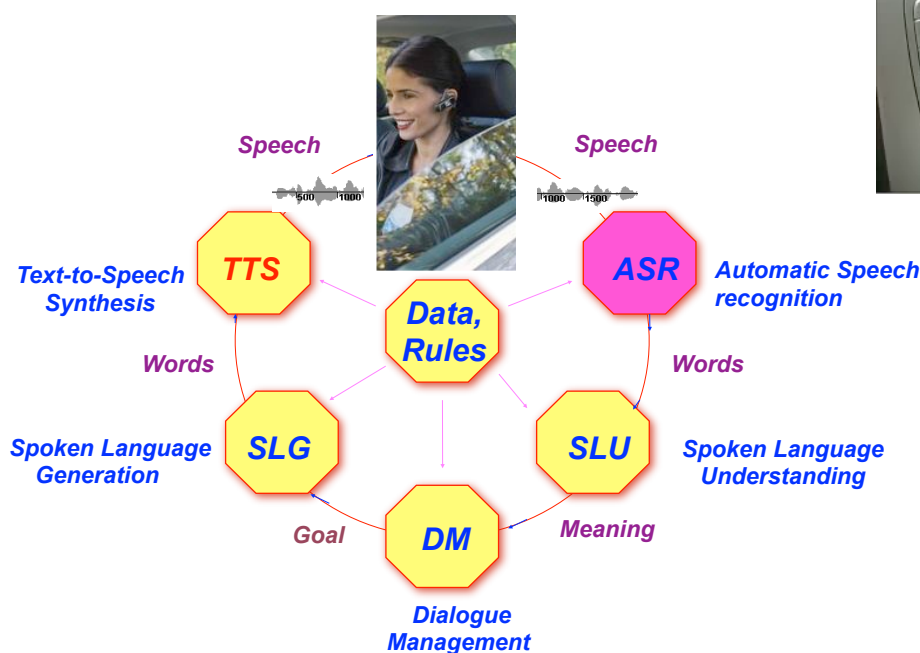
Knowledge needed to build HAL?

- Speech recognition and synthesis
 - ▶ Dictionaries (how words are pronounced)
 - ▶ Phonetics (how to recognize/produce each sound of English)
- Natural language understanding
 - ▶ Knowledge of the English words involved
 - What they mean
 - How they combine (what is a “pod bay door”?)
 - ▶ Knowledge of syntactic structure
 - I'm I do, Sorry that afraid Dave I'm can't

And more ...

- Dialog and pragmatic knowledge
 - ▶ “Open the door” is a REQUEST (as opposed to a STATEMENT or QUESTION)
 - ▶ It is polite to respond, even if you’re planning to kill someone.
 - ▶ It is polite to pretend to want to be cooperative
 - I’m afraid, I can’t...
 - ▶ Discourse structure
 - What is “that” in “I can’t do that.” ?
- Language Generation & Speech Synthesis
 - ▶ what to say
 - ▶ what words and discourse and syntactic structures to use to say it
 - ▶ how to map the words to the sounds of the language

Spoken Dialogue Systems



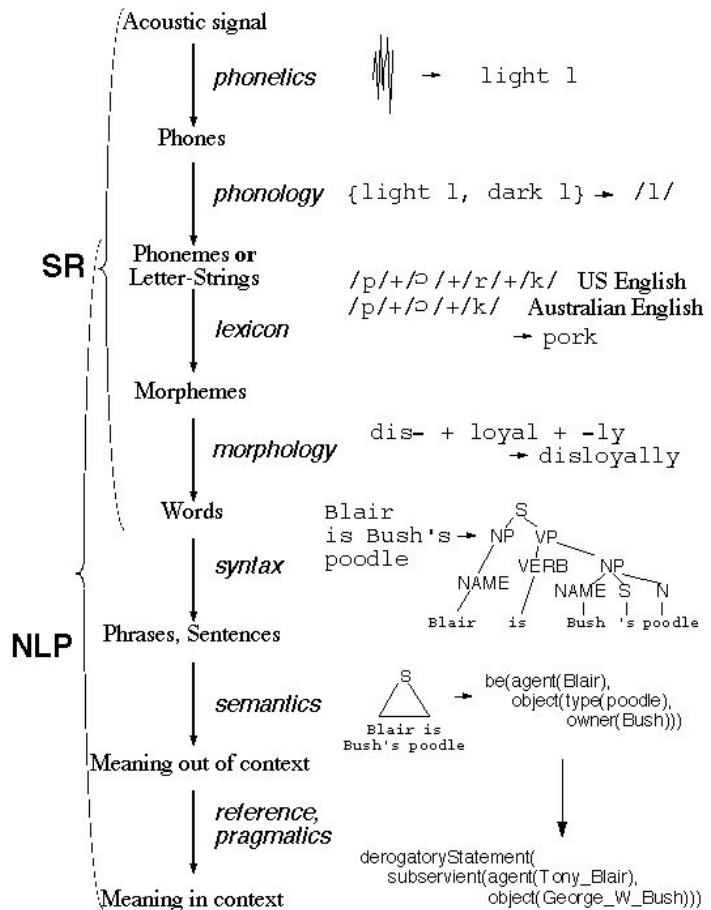
Two related tasks in any application of NLP:

- Represent the knowledge in a computationally tractable form;
- Design and implement algorithms that employ the knowledge to achieve communicative goals

Many NLP tasks can be viewed as transforming one representation (letters, sounds, words, syntactic structures, meanings) into another

Each step needs different knowledge: phonetic, orthographic, lexical, morphological, grammatical, semantic, pragmatic, common-sensical

Speech Recognition and NLP Model:



Speech/Character Recognition

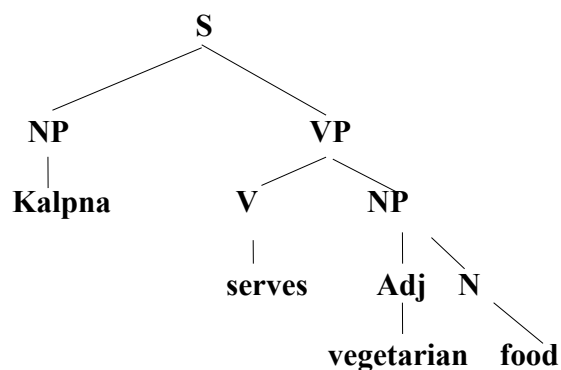
- Listener needs to break a continuous stream of sound/sequence of characters into smaller units
 - ▶ decomposition into words
 - ▶ segmentation of words into appropriate phones or letters
- Requires knowledge of phonological patterns:
 - I'm **en**ormously proud.
 - I **mean** to make you proud.
- ▶ need extra cues, e.g., duration, to determine which it is

Morphological Analysis

- Morphology: word formation
- Inflectional
 - ▶ duck + s = [N duck] + [plural s]
 - ▶ duck + s = [V duck] + [3rd person s]
- Derivational, e.g., nominalization
 - ▶ kind, kindness: [Adj kind] + “ness” = [N kindness]
- Spelling changes
 - ▶ drop, dropping
 - ▶ hide, hiding

Syntactic Analysis

- Associate constituent structure with string
- Prepare for semantic interpretation



Semantics

- A way of representing meaning
- Abstracts away from syntactic structure
- Example:

- ▶ First-Order Logic:

- `cuisine-type(Tanjore, Vegetarian)`

- ▶ Can be:

- Tanjore serves Vegetarian food*

- Tanjore has Vegetarian food*

- Tanjore has Vegetarian cuisine*

- Vegetarian food is served by Tanjore*

Discourse Analysis

- Discourse: How propositions fit together in a conversation or multi-sentence text

- ▶ Pronoun reference:

- The professor told the student to finish the assignment.

- He was pretty aggravated at how long it was taking to hand it in.

- ▶ Multiple reference to same entity:

- David Cameron, Prime Minister of the UK.

- ▶ Relation between sentences:

- John hit the man. He had stolen his bicycle.

- Max fell. John pushed him.

Two Fundamental Problems for NLP

- **Ambiguity:** the transformation from one representation to another is often one-to-many
- **Context:** At all levels, a lot is left out and must be supplied from context

Ambiguity

- Find at least 5 meanings of this sentence:

I made her duck

Ambiguity

Find at least 5 meanings of this sentence:

I made her duck

- I cooked waterfowl for her benefit (to eat)
- I cooked waterfowl belonging to her
- I created the (plaster?) duck she owns
- I caused her to quickly lower her head or body
- I waved my magic wand and turned her into undifferentiated waterfowl

Ambiguity is Everywhere

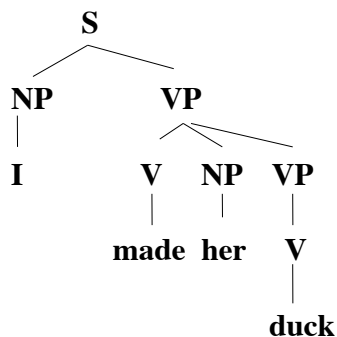
- Lexical category: part of speech
 - ▶ **Duck** can be a **Noun** or **Verb**
 - **V:** *Duck!* I caused her to quickly lower her head or body.
 - **N:** I cooked waterfowl for her benefit
 - ▶ **Her** can be **possessive** (of her) or **dative** (for her)
 - **Possessive:** I cooked waterfowl belonging to her.
 - **Dative:** I cooked waterfowl for her benefit
- Lexical Semantics:
 - ▶ **Make** can mean **create** or **cook**
 - **create:** I made the (plaster) duck statue she owns
 - **cook:** I cooked waterfowl for her benefit

Really Everywhere

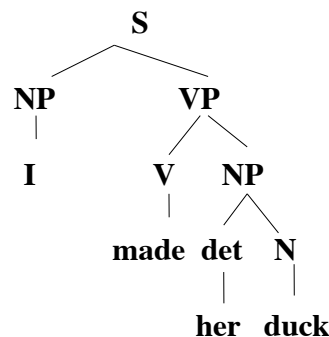
- Grammar: **Make** can be:
 - ▶ Transitive: (**verb** has a **noun direct object**)
 - I cooked [waterfowl belonging to her]
 - ▶ Ditransitive: (**verb** has **2 noun objects**)
 - I (magically) made [her] (into) [undifferentiated waterfowl]
 - ▶ Action-transitive (**verb** has a **direct object and another verb**)
 - I caused [her] [to move her body]

Syntactic Ambiguity

- Structural ambiguity: one sentence can have many syntactic representations



*I caused her to quickly
lower her head or body*



*I created her plaster
waterfowl*

Ambiguity is Everywhere

- Phonetics!
 - I mate or duck
 - I'm eight or duck
 - Eye maid; her duck
 - Aye mate, her duck
 - I maid her duck
 - I'm aid her duck
 - I mate her duck
 - I'm ate her duck
 - I'm ate or duck
 - I mate or duck

*It's hard to wreck
a nice beach!*

Vagueness

- Example:
 - “I want to eat Indian food for lunch.”

Exactly what do I want to eat?

When?

Context to the rescue

Q1: What did you cook for Mary last night?

A1: I made her duck.

Q2: Where did Mary get that great plaster duck?

A2: I made her duck.

Summary

- Participating in a natural language conversation is hard:
 - ▶ ambiguity and vagueness
 - ▶ need to interpret and generation language **in context**
 - ▶ real world knowledge
 - ▶
- Language use is intertwined with general human cognitive abilities
- Ability to process language as people do will signal the arrival of truly intelligent machines

[We're getting there!](#)