Introduction to Ling571

Scott Farrar
CLMA, University of Washington
farrar@u.washington.edu

January 4, 2010

Today's lecture

1. Ling571 in the CLMA Program
   - Shallow processing
   - Deep processing
   - Cross-cutting themes
2. Linguistic Structure

Shallow processing

Shallow processing means less reliance on linguistic structures, more reliance on surface (textual/signal) patterns in the data. Some tasks for **shallow** processing.

- speech recognition using hidden Markov models
- part-of-speech tagging using n-gram techniques
- information extraction based on text patterns (making minimal use of linguistic knowledge)

Shallow $\neq$ easy or simple (cf. Ling570).

Shallow processing

Shallow processing means less reliance on linguistic structures, more reliance on surface (textual/signal) patterns in the data. Some tasks for **shallow** processing.

- speech recognition using hidden Markov models
- part-of-speech tagging using n-gram techniques
- information extraction based on text patterns (making minimal use of linguistic knowledge)

Shallow $\neq$ easy or simple (cf. Ling570).

Shallow processing

Shallow processing means less reliance on linguistic structures, more reliance on surface (textual/signal) patterns in the data. Some tasks for **shallow** processing.

- speech recognition using hidden Markov models
- part-of-speech tagging using n-gram techniques
- information extraction based on text patterns (making minimal use of linguistic knowledge)

Shallow $\neq$ easy or simple (cf. Ling570).

Shallow processing

Shallow processing means less reliance on linguistic structures, more reliance on surface (textual/signal) patterns in the data. Some tasks for **shallow** processing.

- speech recognition using hidden Markov models
- part-of-speech tagging using n-gram techniques
- information extraction based on text patterns (making minimal use of linguistic knowledge)

Shallow $\neq$ easy or simple (cf. Ling570).

Shallow processing

Shallow processing means less reliance on linguistic structures, more reliance on surface (textual/signal) patterns in the data. Some tasks for **shallow** processing.

- speech recognition using hidden Markov models
- part-of-speech tagging using n-gram techniques
- information extraction based on text patterns (making minimal use of linguistic knowledge)

Shallow $\neq$ easy or simple (cf. Ling570).

Shallow processing task

Morpheme identification

testing, fling, going, bling, go, test

Shallow processing task

Morpheme identification

testing, fling, going, bling, go, test

Morphemes

test, *fl, go, *bl, ing

Shallow processing task

Morpheme identification

testing, fling, going, bling, go, test

Morphemes

test, *fl, go, *bl, ing

In fact, shallow processing is often used to derive structure for further deeper processing.

Deep processing

Deep processing means utilizing elaborated linguistic structures.
Some tasks for deep processing:

- deriving structural descriptions of natural language sentences (NL parsing)
- deriving meaning representations from speech (NL understanding)
- generating accurate NL based on meaning representations (NL generation)
- clustering documents based on extracted meaning

Deep processing requires more linguistic knowledge than shallow processing.

Deep processing

Deep processing means utilizing elaborated linguistic structures.
Some tasks for deep processing:

- deriving structural descriptions of natural language sentences (NL parsing)
- deriving meaning representations from speech (NL understanding)
- generating accurate NL based on meaning representations (NL generation)
- clustering documents based on extracted meaning

Deep processing requires more linguistic knowledge than shallow processing.

Deep processing

Deep processing means utilizing elaborated linguistic structures.
Some tasks for deep processing:

- deriving structural descriptions of natural language sentences (NL parsing)
- deriving meaning representations from speech (NL understanding)
- generating accurate NL based on meaning representations (NL generation)
- clustering documents based on extracted meaning

Deep processing requires more linguistic knowledge than shallow processing.

Deep processing

Deep processing means utilizing elaborated linguistic structures.
Some tasks for deep processing:

- deriving structural descriptions of natural language sentences (NL parsing)
- deriving meaning representations from speech (NL understanding)
- generating accurate NL based on meaning representations (NL generation)
- clustering documents based on extracted meaning

Deep processing requires more linguistic knowledge than shallow processing.

Deep processing

Deep processing means utilizing elaborated linguistic structures.
Some tasks for deep processing:

- deriving structural descriptions of natural language sentences (NL parsing)
- deriving meaning representations from speech (NL understanding)
- generating accurate NL based on meaning representations (NL generation)
- clustering documents based on extracted meaning

Deep processing requires more linguistic knowledge than shallow processing.

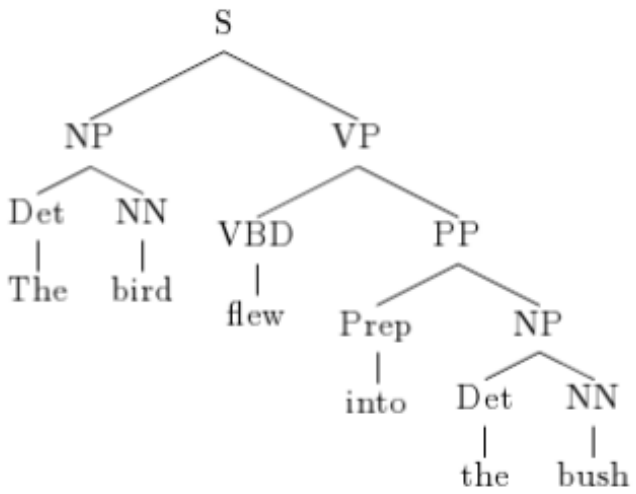
Deep processing

Deep processing means utilizing elaborated linguistic structures. Some tasks for deep processing:

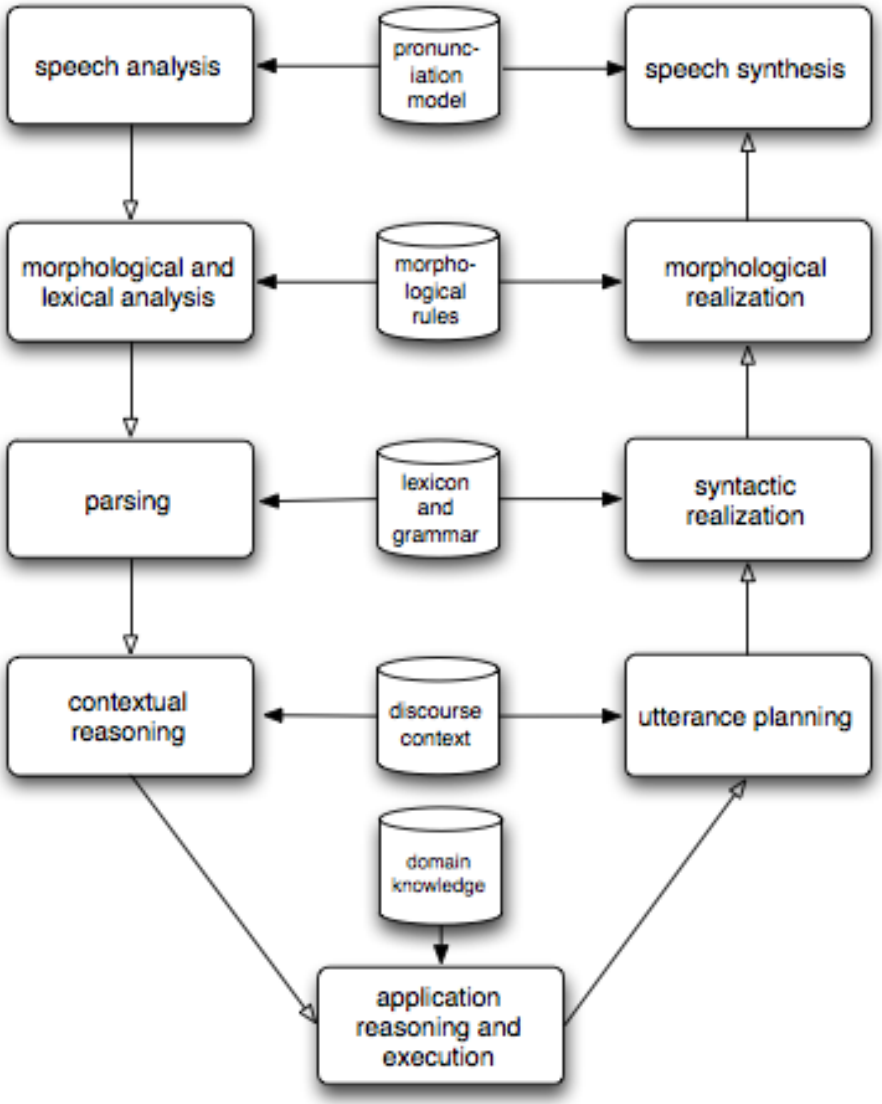
- deriving structural descriptions of natural language sentences (NL parsing)
- deriving meaning representations from speech (NL understanding)
- generating accurate NL based on meaning representations (NL generation)
- clustering documents based on extracted meaning

Deep processing requires more linguistic knowledge than shallow processing.

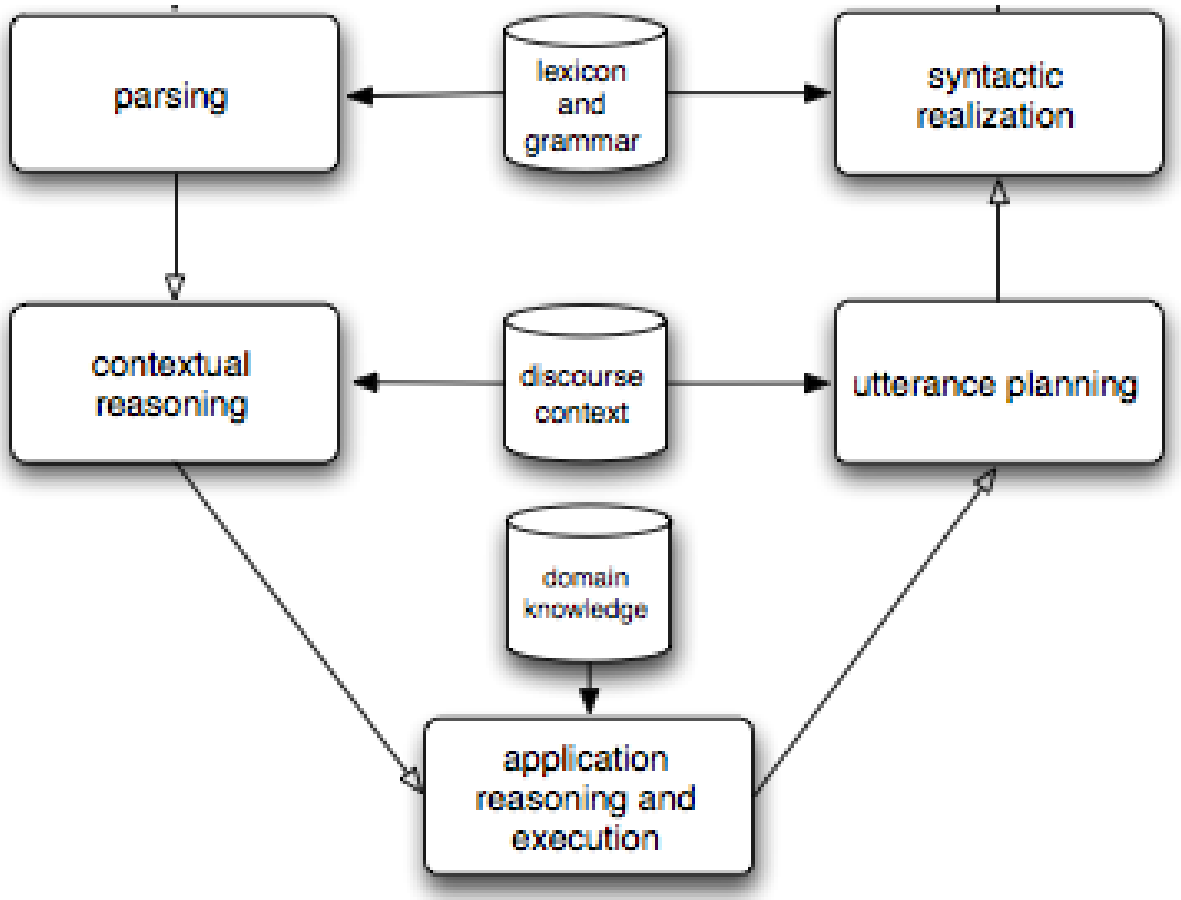
Linguistic structure



End-to-end system



Focus of Ling571



Deep and shallow processing: similarities

- Both require and can benefit from stochastic and rule-based techniques
- Both require lots of data
- Each has its own core set of algorithms
- The end goal is the same: deriving useful information from natural language

What about the CLMA themes?

1. ambiguity resolution
2. evaluation
3. multilingual processing

Ambiguity resolution

Language is inherently ambiguous, at every linguistic level (phonological, morphological, syntactic, etc.):

phon /aiskrim/ *ice-cream* or *I scream*

morph *un-doable* or *undo-able*

synt *Flying planes can be dangerous.*

sem *Every boy kissed a girl.*

Ambiguity resolution

Language is inherently ambiguous, at every linguistic level (phonological, morphological, syntactic, etc.):

phon /aiskrim/ *ice-cream* or *I scream*

morph *un-doable* or *undo-able*

synt *Flying planes can be dangerous.*

sem *Every boy kissed a girl.*

Ambiguity resolution

Language is inherently ambiguous, at every linguistic level (phonological, morphological, syntactic, etc.):

phon /aiskrim/ *ice-cream* or *I scream*

morph *un-doable* or *undo-able*

synt *Flying planes can be dangerous.*

sem *Every boy kissed a girl.*

Ambiguity resolution

Language is inherently ambiguous, at every linguistic level (phonological, morphological, syntactic, etc.):

phon	/aiskrim/ *ice-cream* or *I scream*
morph	*un-doable* or *undo-able*
synt	*Flying planes can be dangerous.*
sem	*Every boy kissed a girl.*

Evaluation

For each NLP task, we require some measure of success.

Consider an information retrieval system: TREC competition, Ask.com

$$\textit{Precision} = \frac{\#\text{of correct answers given by the system}}{\#\text{ of answers given by system}}$$

$$\textit{Recall} = \frac{\#\text{of correct answers given by the system}}{\text{total }\#\text{ of possible correct answers given by system}}$$

What about a parser?

Evaluation

For each NLP task, we require some measure of success.

Consider an information retrieval system: TREC competition, Ask.com

$$\textit{Precision} = \frac{\#\text{of correct answers given by the system}}{\#\text{ of answers given by system}}$$

$$\textit{Recall} = \frac{\#\text{of correct answers given by the system}}{\text{total }\#\text{ of possible correct answers given by system}}$$

What about a parser?

```
Parsing [sent. 2 len. 8]: There is a fly in my soup .
(ROOT
  (S
    (NP (EX There))
    (VP (VBZ is)
      (NP
        (NP (DT a) (VB fly))
        (PP (IN in)
          (NP (PRP$ my) (NN soup)))))
    (. .)))

(ROOT
  (S
    (NP (EX There))
    (VP (VBZ is))
    (NP
      (NP (DT a) (NN fly))
      (PP (IN in)
        (NP (PRP$ my) (NN soup))))))
```

```
Parsing [sent. 2 len. 8]: There is a fly in my soup .
(ROOT
  (S
    (NP (EX There))
    (VP (VBZ is)
      (NP
        (NP (DT a) (VB fly))
        (PP (IN in)
          (NP (PRP$ my) (NN soup)))))
    (. .)))

(ROOT
  (S
    (NP (EX There))
    (VP (VBZ is))
    (NP
       (NP (DT a) (NN fly))
       (PP (IN in)
         (NP (PRP$ my) (NN soup))))))
```

Multilingual processing

- Linguistic structure, to some extent, applies to all languages.
- But each language has its own particular structures:
  - Word order varies.
  - What's a word?
  - How does the language carve up the semantic space?
- We'll look at other languages when necessary, but mostly stick to English.

Multilingual processing

- Linguistic structure, to some extent, applies to all languages.
- But each language has its own particular structures:
  - Word order varies.
  - What's a word?
  - How does the language carve up the semantic space?
- We'll look at other languages when necessary, but mostly stick to English.

Multilingual processing

- Linguistic structure, to some extent, applies to all languages.
- But each language has its own particular structures:
  - Word order varies.
  - What's a word?
  - How does the language carve up the semantic space?
- We'll look at other languages when necessary, but mostly stick to English.

Multilingual processing

- Linguistic structure, to some extent, applies to all languages.
- But each language has its own particular structures:
  - Word order varies.
  - What's a word?
  - How does the language carve up the semantic space?
- We'll look at other languages when necessary, but mostly stick to English.

Multilingual processing

- Linguistic structure, to some extent, applies to all languages.
- But each language has its own particular structures:
  - Word order varies.
  - What's a word?
  - How does the language carve up the semantic space?
- We'll look at other languages when necessary, but mostly stick to English.

Multilingual processing

- Linguistic structure, to some extent, applies to all languages.
- But each language has its own particular structures:
  - Word order varies.
  - What's a word?
  - How does the language carve up the semantic space?
- We'll look at other languages when necessary, but mostly stick to English.

What's not included in Ling571

- word-level processing and below (570)
- machine learning (572)
- dialogue processing (573)
- machine translation (575)
- speech processing (575)
- information extraction/retrieval, Q/A (575)

[see next set of slides]