Natural Language Processing (NLP) Chapter 22: Communication

- "Classical" view (pre-1953):

language consists of sentences that are true/false (logic)

- "Modern" view (post-1953):

language is a form of action

- We have speaker, hearer, and utterance terms referring to any mode of communication. We will also use the term word to refer to any kind of conventional communicative sign.

Speech acts

Speech acts achieve the speaker's goals:

- Inform: " There's a pet in front of you"
- Query: " Can you see the gold?"
- Command: " Pick it up"
- Promise: " I'll share the gold with you"
- Acknowledge: " OK"

Communication

- Communication Typical communication
  S (speaker) wants to convey P (proposition) to H (hearer) using W (words in a formal or natural language)

1. Speaker
   - **Intention:** S wants H to believe P
   - **Generation:** S chooses words W
   - **Synthesis:** S utters words W

2. Hearer
   - **Perception:** H perceives words W" (ideally W" = W)
   - **Analysis:** H infers possible meanings P1,P2,…,Pn for W"
   - **Disambiguation:** H infers that S intended to convey Pi (ideally Pi=P)
   - **Incorporation:** H decides to believe or disbelieve Pi

Communication

How could this go wrong?

- Insincerity: (S doesn't believe P)
- Speech wreck ignition failure
- Ambiguous utterance
- Differing understanding of current context

NLP Processing

- Natural Language Understanding
- Taking some spoken/typed sentence and working out what it means
- Natural Language Generation
- Taking some formal representation of what you want to say and working out a way to express it in a natural (human) language (e.g., English)

Applications of NLP

- Machine Translation
- Database Access
- Information Retrieval

Selecting from a set of documents the ones that are relevant to a query

- Text Categorization

Sorting text into fixed topic categories

- Extracting data from text

Converting unstructured text into structure data

- Spoken language control systems
- Spelling and grammar checkers

Natural language understanding

Raw speech signal

↓ **Speech recognition**

Sequence of words spoken

↓ **Syntactic analysis** using knowledge of the grammar

Structure of the sentence

↓ **Semantic analysis** using info. about meaning of words

Partial representation of meaning of sentence

↓ **Pragmatic analysis** using info. about context

Final representation of meaning of sentence

NLP Topics

- **Phonology** – concerns how words are related to the sounds that realize them.
- **Morphology** – concerns how words are constructed from more basic meaning
- **Syntax** – concerns how can be put together to form correct sentences – Grammar
- **Semantics** – concerns what words mean and how these meaning combine in sentences to form sentence meaning.
- **Pragmatics** – concerns how sentences are used in different situations
- **Discourse** – concerns how the immediately preceding sentences affect the interpretation of the next sentence. "context"
- **World Knowledge** – includes general knowledge about the world.

Lexical Processing

- The purpose of lexical processing is to determine meanings of individual words.
- Basic methods is to lookup in a database of meanings – **lexicon**
- We should also identify non-words such as punctuation marks.
- Word-level ambiguity -- words may have several meanings, and the correct one cannot be chosen based solely on the word itself. – bank in English
- Solution -- resolve the ambiguity on the spot (if possible) or pass-on the ambiguity to the other levels.

Syntactic Processing

- **Parsing --** converting a flat input sentence into a hierarchical structure that corresponds to the units of meaning in the sentence.

  • There are different parsing formalisms and algorithms.
  • Most formalisms have two main components:
  → **grammar --** a declarative representation describing the syntactic structure of sentences in the language.
  → **parser --** an algorithm that analyzes the input and outputs its structural representation (its parse) consistent with the grammar specification.
  CFGs are in the center of many of the parsing mechanisms. But they are complemented by some additional features that make the formalism more suitable to handle natural languages.

Semantic Analysis

- Assigning meanings to the structures created by syntactic analysis.

  • Mapping words and structures to particular domain objects in way consistent with our knowledge of the world.

  • Semantic can play an important role in selecting among competing syntactic analyses and discarding illogical analyses.
  – I robbed the bank -- bank is a river bank or a financial institution
  • We have to decide the formalisms which will be used in the meaning representation.

Pragmatics

- Uses context of utterance
  - Where, by who, to whom, why, when it was said
  - Intentions: *inform, request, promise*, *criticize,* …
- Handling Pronouns
  - "Mary eats apples. She likes them."
    - She="Mary", them="apples".
- Handling ambiguity
  - Pragmatic ambiguity: "you're late": What's the speaker's intention: informing or criticizing?

Discourse

- Discourses are collection of coherent sentences (not arbitrary set of sentences)
- Discourses have also hierarchical structures
  - **anaphora resolution -- to resolve referring expression**
  – Mary bought a book for Kelly. **She didn't like it.**
  - **She refers to Mary or Kelly. -- possibly Kelly**
  - **It refers to what -- book.**
  – Mary had to lie for Kelly. **She didn't like it.**

  Discourse structure may depend on application.
  – Monologue
  – Dialogue
  – Human-Computer Interaction

World Knowledge

Includes knowledge of the physical world

- the world of human social interaction,
- The role of goals and intentions in communication.
- This general background knowledge is essential to understand the full meaning of a text or conversation