Communication and Language

Chapter 22

Outline

♦ Communication
♦ Grammar
♦ Syntactic analysis
♦ Problems

Communication

“Classical” view (pre-1953):
  language consists of sentences that are true/false (cf. logic)

“Modern” view (post-1953):
  language is a form of action

Wittgenstein (1953) Philosophical Investigations
Austin (1962) How to Do Things with Words
Searle (1969) Speech Acts

Why?

To change the actions of other agents
Speech acts

**SITUATION**

**Speaker** → **Utterance** → **Hearer**

Speech acts achieve the speaker’s goals:
- Inform “There’s a pit in front of you”
- Query “Can you see the gold?”
- Command “Pick it up”
- Promise “I’ll share the gold with you”
- Acknowledge “OK”

Speech act planning requires knowledge of
- Situation
- Semantic and syntactic conventions
- Hearer’s goals, knowledge base, and rationality

### Stages in communication (informing)

<table>
<thead>
<tr>
<th>Stage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intention</td>
<td>S wants to inform H that ( P )</td>
</tr>
<tr>
<td>Generation</td>
<td>S selects words ( W ) to express ( P ) in context ( C )</td>
</tr>
<tr>
<td>Synthesis</td>
<td>S utters words ( W )</td>
</tr>
<tr>
<td>Perception</td>
<td>H perceives ( W' ) in context ( C' )</td>
</tr>
<tr>
<td>Analysis</td>
<td>H infers possible meanings ( P_1, \ldots, P_n )</td>
</tr>
<tr>
<td>Disambiguation</td>
<td>H infers intended meaning ( P_i )</td>
</tr>
<tr>
<td>Incorporation</td>
<td>H incorporates ( P_i ) into KB</td>
</tr>
</tbody>
</table>

How could this go wrong?
- Insincerity (S doesn’t believe \( P \))
- Speech wreck ignition failure
- Ambiguous utterance
- Differing understanding of current context \( (C \neq C') \)

### Grammar types

**Regular:** nonterminal → terminal | nonterminal

\[ S \rightarrow aS \]
\[ S \rightarrow \Lambda \]

**Context-free:** nonterminal → anything

\[ S \rightarrow aSb \]

**Context-sensitive:** more nonterminals on right-hand side

\[ ASB \rightarrow AAaBB \]

**Recursively enumerable:** no constraints

Related to Post systems and Kleene systems of rewrite rules

Natural languages probably context-free, parsable in real time

### Wumpus lexicon

**Nouns** → stench | breeze | glitter | nothing
            | wumpus | pit | pits | gold | east | ...

**Verbs** → is | see | smell | shoot | feel | stinks
           | go | grab | carry | kill | turn | ...

**Adjectives** → right | left | east | back | smelly | ...

**Adverbs** → here | there | nearby | ahead
            | right | left | east | south | back | ...

**Pronouns** → me | you | I | it | ...

**Names** → John | Mary | Boston | UCB | PAJC | ...

**Articles** → the | a | an | ...

**Prepositions** → to | in | on | near | ...

**Conjunctions** → and | or | but | ...

**Digits** → 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9

Divided into **closed** and **open** classes
Wumpus lexicon

Noun → stench | breeze | glitter | nothing
  | wumpus | pit | pits | gold | east | ...
Verb → is | see | smell | shoot | feel | stinks
  | go | grab | carry | kill | turn | ...
Adjective → right | left | east | south | back | smelly | ...
Adverb → here | there | nearby | ahead
Pronoun → me | you | I | it | S/HE | Y’ALL | ...
Name → John | Mary | Boston | UCB | PAJC | ...
Article → the | a | an | ...
Preposition → to | in | on | near | ...
Conjunction → and | or | but | ...
Digit → 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9

Divided into closed and open classes

Wumpus grammar

S → NP VP
  | S Conjunction S
  | I + feel a breeze
NP → Pronoun
  | Noun
  | Article Noun the + wumpus
  | Digit Digit 3 4
  | NP PP the wumpus + to the east
  | NP RelClause the wumpus + that is smelly
VP → Verb
  | VP NP feels + a breeze
  | VP Adjective is + smelly
  | VP PP turn + to the east
  | VP Adverb go + ahead
PP → Preposition NP
  | to + the east
RelClause → that VP
  | that + is smelly

Grammaticality judgements

Formal language $L_1$ may differ from natural language $L_2$

Adjusting $L_1$ to agree with $L_2$ is a learning problem!

* the gold grab the wumpus
* I smell the wumpus the gold
  * I give the wumpus the gold
* I donate the wumpus the gold

Intersubjective agreement somewhat reliable, independent of semantics!

Real grammars 10–500 pages, insufficient even for “proper” English

Parse trees

Exhibit the grammatical structure of a sentence

I shoot the wumpus

Exhibit the grammatical structure of a sentence

I shoot the wumpus

Exhibit the grammatical structure of a sentence

I shoot the wumpus
### Syntax in NLP

Most view syntactic structure as an essential step towards meaning:

- "Mary hit John" ≠ "John hit Mary"

“And since I was not informed—as a matter of fact, since I did not know that there were excess funds until we, ourselves, in that checkup after the whole thing blew up, and that was, if you’ll remember, that was the incident in which the attorney general came to me and told me that he had seen a memo that indicated that there were no more funds.”

“Wouldn’t the sentence ‘I want to put a hyphen between the words Fish and And and And and Chips in my Fish-And-Chips sign’ have been clearer if quotation marks had been placed before Fish, and between Fish and and, and and and And, and And and and, and and and and, and and and Chips, as well as after Chips?”

### Logical grammars

BNF notation for grammars too restrictive:
- difficult to add “side conditions” (number agreement, etc.)
- difficult to connect syntax to semantics

Idea: express grammar rules as logic

\[
X \rightarrow YZ \quad \text{becomes} \quad Y(s_1) \land Z(s_2) \Rightarrow X(Append(s_1, s_2))
\]

\[
X \rightarrow word \quad \text{becomes} \quad X(["word"])\]

\[
X \rightarrow Y \mid Z \quad \text{becomes} \quad Y(s) \Rightarrow X(s) \quad Z(s) \Rightarrow X(s)
\]

Here, \(X(s)\) means that string \(s\) can be interpreted as an \(X\)
Now it’s easy to augment the rules

\[
NP(s_1) \land EatsBreakfast(Ref(s_1)) \land VP(s_2) \\
\Rightarrow NP(Append(s_1, "who"), s_2))
\]

\[
NP(s_1) \land Number(s_1, n) \land VP(s_2) \land Number(s_2, n) \\
\Rightarrow S(Append(s_1, s_2))
\]

Parsing is reduced to logical inference:

Ask\(\langle KB, S(["I\ " am\ " a\ " wumpus"]\rangle)\)

(Can add extra arguments to return the parse structure, semantics)

Generation simply requires a query with uninstantiated variables:

Ask\(\langle KB, S(x)\rangle\)

If we add arguments to nonterminals to construct sentence semantics, NLP generation can be done from a given logical sentence:

Ask\(\langle KB, S(x, At(Robot, [1, 1]))\rangle\)

Real human languages provide many problems for NLP:

- ambiguity
- anaphora
- indexicality
- vagueness
- discourse structure
- metonymy
- metaphor
- noncompositionality

Squad helps dog bite victim

Helicopter powered by human flies

American pushes bottle up Germans

I ate spaghetti with meatballs
Ambiguity

Squad helps dog bite victim
Helicopter powered by human flies
American pushes bottle up Germans
I ate spaghetti with meatballs
    salad
abandon

Squad helps dog bite victim
Helicopter powered by human flies
American pushes bottle up Germans
I ate spaghetti with meatballs
    salad
abandon
a fork
a friend

Ambiguity can be lexical (polysemy), syntactic, semantic, referential

Anaphora

Using pronouns to refer back to entities already introduced in the text

After Mary proposed to John, they found a preacher and got married.
Anaphora

Using pronouns to refer back to entities already introduced in the text
After Mary proposed to John, they found a preacher and got married.
For the honeymoon, they went to Hawaii

Mary saw a ring through the window and asked John for it
Mary threw a rock at the window and broke it

Indexicality

Indexical sentences refer to utterance situation (place, time, S/H, etc.)
I am over here
Why did you do that?

Metonymy

Using one noun phrase to stand for another
I’ve read Shakespeare
Chrysler announced record profits
The ham sandwich on Table 4 wants another beer

Metaphor

“Non-literal” usage of words and phrases, often systematic:
I’ve tried killing the process but it won’t die. Its parent keeps it alive.
Noncompositionality

basketball shoes
baby shoes
alligator shoes
designer shoes

Noncompositionality

basketball shoes
baby shoes
alligator shoes
designer shoes
brake shoes

Noncompositionality

basketball shoes
baby shoes
alligator shoes
designer shoes
brake shoes
red book
Noncompositionality

basketball shoes
baby shoes
alligator shoes
designer shoes
brake shoes
red book
red pen
red hair

Noncompositionality

basketball shoes
baby shoes
alligator shoes
designer shoes
brake shoes
red book
red pen
red hair
red herring
small moon
large molecule
mere child
Noncompositionality

basketball shoes
baby shoes
alligator shoes
designer shoes
brake shoes
red book
red pen
red hair
red herring
small moon
large molecule
mere child
alleged murderer
real leather
artificial grass