Introduction to Artificial Intelligence

Language

Bernhard Beckert

UNIVERSITÄT KOBLENZ-LANDAU

Summer Term 2003

Outline

- **Communication**

- **Grammar**

- **Syntactic analysis**

- **Problems**

Communication

"Classical" view (pre-1953)

Language consists of sentences that are true/false

"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)

Intention   S wants to inform H that P

Generation   S selects words W to express P

Synthesis   S utters words W

Perception   H perceives W'

Analysis   H infers possible meanings $P_1, \ldots P_n$

Disambiguation   H infers intended meaning P_i

Incorporation   H incorporates P_i into KB

How could this go wrong?

– **Insincerity (S doesn't believe P)**

– **Speech wreck ignition failure**

– **Ambiguous utterance**

– **Differing understanding of current situation**

Grammar

Purpose

Grammar specifies the compositional structure of complex messages e.g., speech (linear), text (linear), music (two-dimensional)

Formal language

A set of strings of terminal symbols

Each string in the language can be analyzed/generated by the grammar

Grammar

Rewrite rules

The grammar is a set of rewrite rules

Example

$$S \rightarrow NP\ VP$$

$$Article \rightarrow the \mid a \mid an \mid \ldots$$

S: **the sentence symbol**

NP, VP: **nonterminals**

the, a: **terminal symbols**

Wumpus Lexicon

$Noun \rightarrow$ *stench | breeze | glitter | nothing | wumpus | pit | pits | gold | east | ...*

$Verb \rightarrow$ *is | see | smell | shoot | feel | stinks | go | grab | carry | kill | turn | ...*

$bAdjective \rightarrow$ *right | left | east | south | back | smelly | ...*

$Adverb \rightarrow$ *here | there | nearby | ahead | right | left | east | south | back | ...*

$Pronoun \rightarrow$ *me | you | I | it | S/HE | Y'ALL...*

$Name \rightarrow$ *John | Mary | Boston | UCB | PAJC | ...*

$Article \rightarrow$ *the | a | an | ...*

$Preposition \rightarrow$ *to | in | on | near | ...*

$Conjunction \rightarrow$ *and | or | but | ...*

$Digit \rightarrow$ *0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9*

Divided into closed and open classes

Wumpus Grammar

$$S \rightarrow NP\ VP$$ **I + feel a breeze**
$$|\quad S\ Conjunction\ S$$ **I feel a breeze + and + I smell a wumpus**

$$NP \rightarrow Pronoun$$ **I**
$$|\quad Noun$$ **pits**
$$|\quad Article\ Noun$$ **the + wumpus**
$$|\quad Digit\ Digit$$ **3 4**
$$|\quad NP\ PP$$ **the wumpus + to the east**
$$|\quad NP\ RelClause$$ **the wumpus + that is smelly**

$$VP \rightarrow Verb$$ **stinks**
$$|\quad VP\ NP$$ **feel + a breeze**
$$|\quad VP\ Adjective$$ **is + smelly**
$$|\quad VP\ PP$$ **turn + to the east**
$$|\quad VP\ Adverb$$ **go + ahead**

$$PP \rightarrow Preposition\ NP$$ **to + the east**

$$RelClause \rightarrow that\ VP$$ **that + is smelly**

Grammaticality judgements

Formal language L_1 may differ from natural language L_2

L_1 L_2
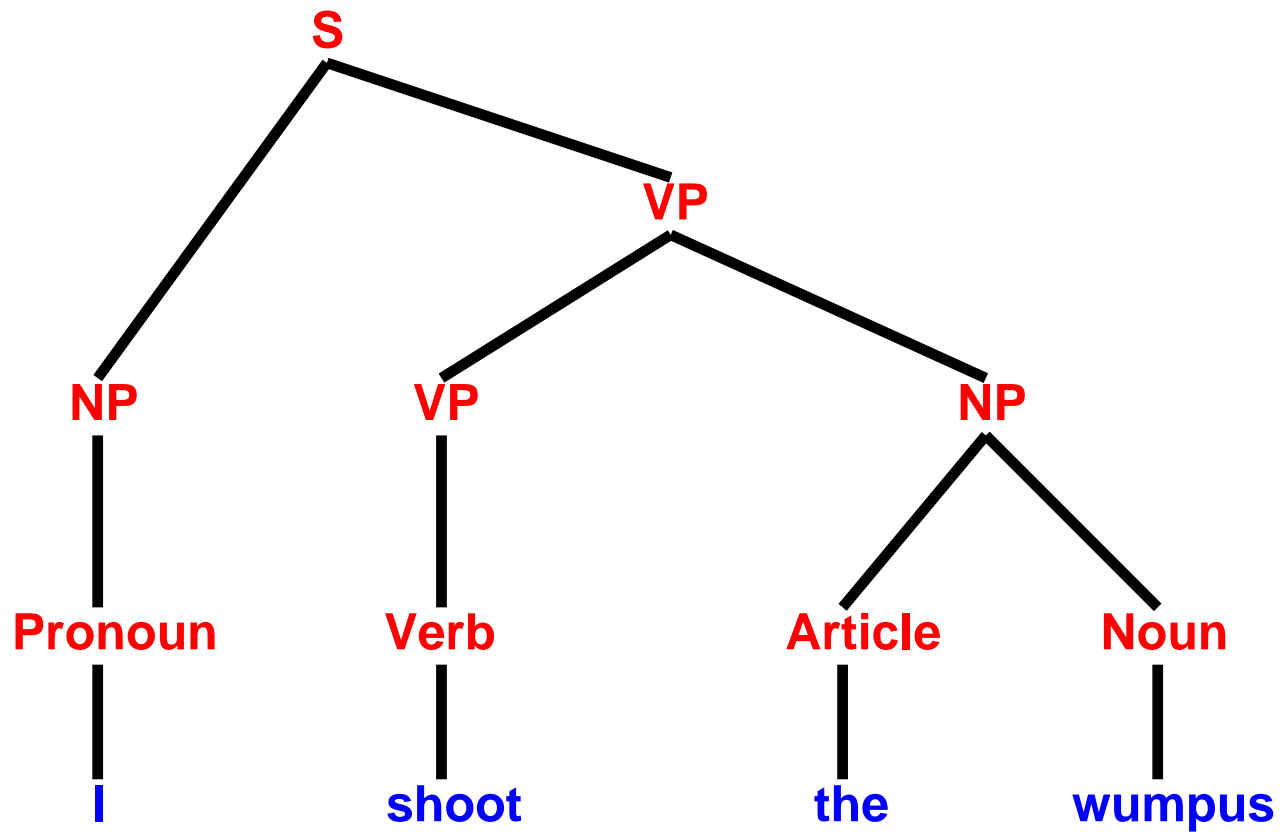
false positives   **false negatives**
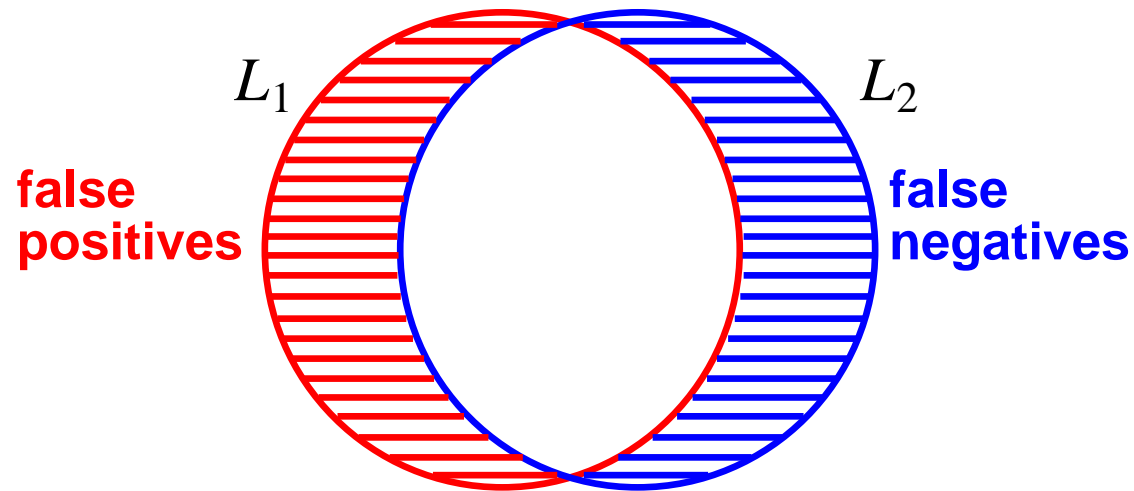
Examples

+   the gold grab the wumpus
+   I smell the wumpus the gold
−   I give the wumpus the gold
+   I donate the wumpus the gold

Parse Trees

Exhibit the grammatical structure of a sentence

Syntax in Natural Language Processing

Syntactic structure is 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 And and and, and and and Chips, as well as after Chips?"**

Real Language

Real human languages provide many problems for natural language processing

- **ambiguity**

- **anaphora**

- **indexicality**

- **vagueness**

- **noncompositionality**

- **discourse structure**

- **metonymy**

- **metaphor**

Ambiguity

Examples

- **Squad helps dog bite victim**

- **Helicopter powered by human flies**

- **I ate spaghetti with meatballs**
  salad
  abandon
  a fork
  a friend

Note

Ambiguity can be

– **lexical**

– **syntactic**

– **semantic**

– **referential**

Indexicality

Indexical sentences

Refer to utterance situation (place, time, etc.)

Examples

- **I am over here**

- **Why did you do that?**

Anaphora

Anaphora

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

Examples

- **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**

Metonymy

Metonomy

Using one noun phrase to stand for another

Example

- I've read **Shakespeare**

- **Chrysler** announded record profits

- The **ham sandwich** on Table 4 wants another beer

Metaphor

Metaphor

"Non-literal" usage of words and phrases, often systematic:

Example

I've tried killing the process but it won't die.

Noncompositionality

Examples

- **basketball shoes**

- **baby shoes**

- **alligator shoes**

- **designer shoes**

- **brake shoes**

Disambiguation

Requires knowledge of different kind

- **World model**

- **Mental model (of the speaker)**

- **Language model**

- **Acoustic model**