Artificial Intelligence

Lecture 17: Introduction to Natural Language Processing

Dr. Benjamin Inden

Department of Computing & Technology, Nottingham Trent University

February 27, 2017

Lecture for this week

- Based on Russel & Norvig, Chapter 23; Lucci & Kopec, Chapter 12; Lecture Notes by Dr. Paul Bowden; Bird, Klein & Loper, Natural Language Processing with Python – Analyzing Text with the Natural Language Toolkit, 2nd ed.; some images and other materials from Wikipedia.
- Please do not post these slides to the internet or use other means of public distribution. They are for your personal use only.

Outline

- What is communication?
- Grammar for natural language

What is communication?

Communication is the intentional exchange of information brought about by the production and perception of signs drawn from a shared system of conventional signs.
This "shared system of conventional signs" is called a *language*.

Speech acts

SITUATION

Speaker → Utterance → Hearer

Speech acts achieve the speaker's

	Inform	"There's a pit in front of you"
	Query	"Can you see the gold?"
goals:	*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

A model of stages in communication (Act: Informing)

Intention	S wants to inform H that P
Generation	S selects words W to express P in context C
Synthesis	S utters words W
Perception	H perceives W' in context C'
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)
- Ambiguous utterance
- Differing understanding of current context ($C \neq C'$)

Language features: Vocabulary

a set of words, usually many hundreds or thousands, to name and describe objects in the world and actions (processes) that those objects - people or animals (or even inanimate things, such as the wind) - may perform.
A language's vocabulary will not usually be the same as, or even similar to, another randomly-selected language's vocabulary. There is an arbitrary relation of form to meaning (dog, cane, Hund, etc)
Therefore all NLP systems will have to contain a lexicon.

Grammar

Vervet monkeys, antelopes etc. use isolated symbols for sentences
$\Longrightarrow$ restricted set of communicable propositions, no generative capacity
(Chomsky (1957): *Syntactic Structures*)
Grammar specifies the compositional structure of complex messages
e.g., speech (linear), text (linear), music (two-dimensional)
A *formal language* is a set of *strings* of *terminal symbols*
Each string in the language can be analyzed/generated by the grammar
The grammar is a set of *rewrite rules*, e.g.,
$S \rightarrow VP\ NP$
$Article \rightarrow the|a|an|...$
Here S is the sentence symbol, NP and VP are *nonterminals*

Grammars for natural languages

A grammar must allow to

- ask questions
- form commands
- make a negative
- talk about past, present and future events
- describe things
- speculate on merely possible events
- make statements about the mental state of some other entity.

English grammar: Nouns

noun	a word used to name things
proper noun	names of individual towns, people, etc.
count noun	things you can count (e.g., book, person)
mass noun	things you cannot count (e.g., water, butter, ...)

singular vs. *plural*

English grammar: Adjectives

adjectives qualify nouns. E.g., “dog” -> “big dog”

comparative: “bigger”, “faster”, ...

superlative: “biggest”, “fastest”, ...

English grammar: Verbs

verbs describe actions / processes ("run", "drink", "happen", "be", ...)

transitive verbs have an object ("read a book")

intransitive verbs don't have an object ("sleep")

modal verbs "can", "should", "must", ...

Typically, verbs have

- tense (when?)
- aspect (how?)

English grammar: Adverbs

Adverbs qualify verbs, adjectives, or other adverbs (“slowly”, “loudly”, “beautifully”)

English grammar: More parts of speech

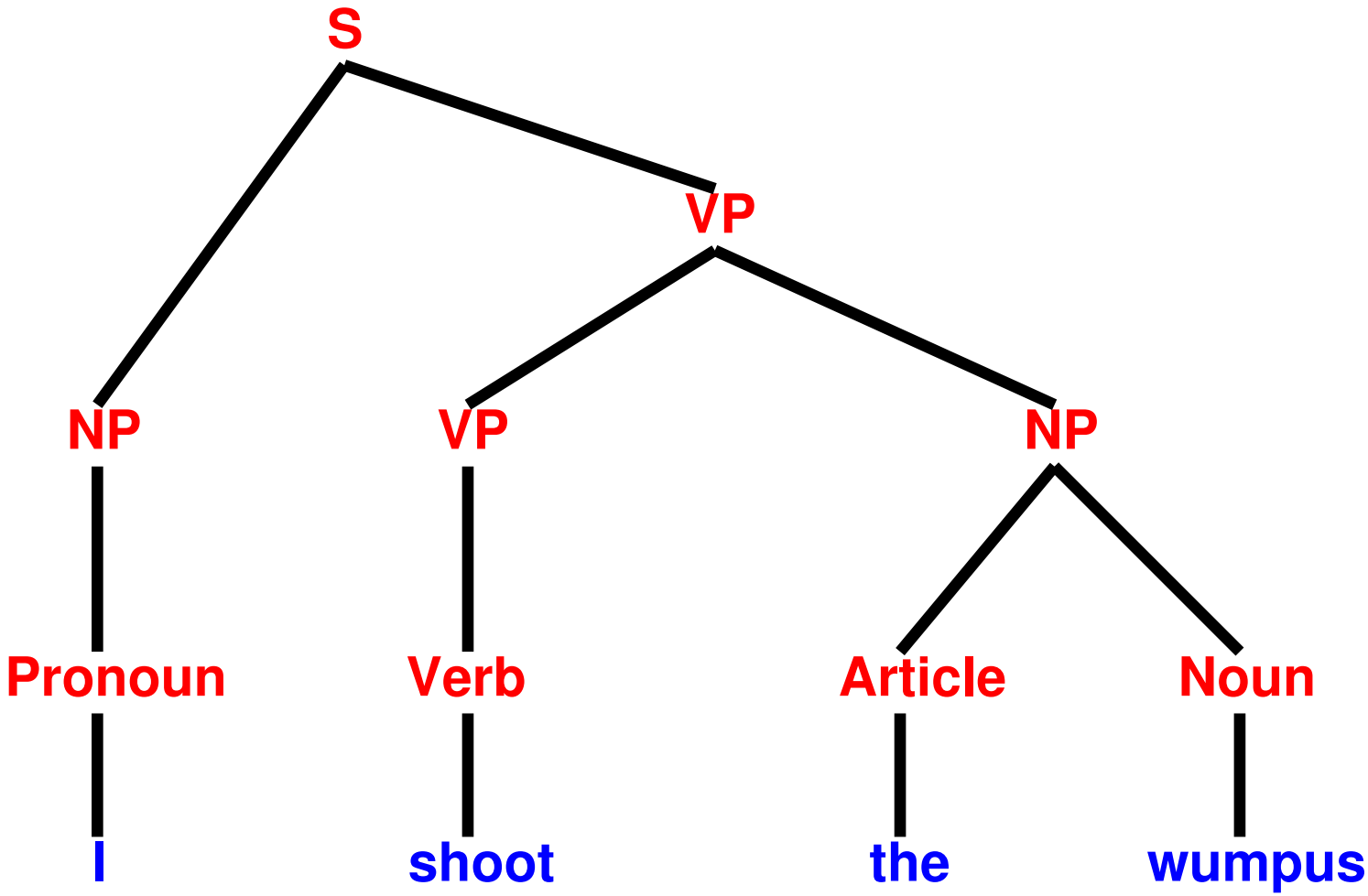
- Pronouns
  - possessive pronouns (“his”, “mine”, ...)
  - personal pronouns (“I”, “them”, “they”, ...)
- Prepositions (“in”, “with”, “onto”, ...)
- Conjunctions
  - Co-ordinating conjunctions (“and”, “but”, ...)
  - Sub-ordinating conjuctions (“because”, “unless”, “if”, ...)
- Articles
  - Definitive article (“the”)
  - Indefinitive articles (“a”, “an”)

Sentences

A sentence is usually described as a unit of text which holds a single thought or idea.
Words within the sentence can have different functions, which can be indicated by inflection or word order.
In English, many sentences have a structure of subject - verb - object. E.g.:

1. The cat caught the mouse.
2. The mouse caught the cat.

Parse trees



Texts

A text is a group of sentences. The text as a whole has a purpose, and so the sentences in it must form a coherent, readable whole. Some challenges for processing of text:

Anaphoric reference "Harry came over to my house last night. I wanted *him* to help me with my maths homework."

Ellipsis "Two buses came round the corner, and then another two."

Dialogues

In a dialogue, both parties make speech acts until the purpose of the communication session has been satisfied, to one or both of the conversants' satisfaction.
Features:

- Turn taking
- Extralinguistic events (gestures, facial expressions,...)
- Anaphoric references / ellipsis
- Topic / topic change

Conversational agents

ELIZA (c. 1966) by Joseph Weizenbaum

Men are all alike.
IN WHAT WAY
They're always bugging us about something or other.
CAN YOU THINK OF A SPECIFIC EXAMPLE
Well, my boyfriend made me come here.
YOUR BOYFRIEND MADE YOU COME HERE
He says I'm depressed much of the time.
I AM SORRY TO HEAR YOU ARE DEPRESSED
It's true. I am unhappy.
DO YOU THINK THAT COMING HERE WILL HELP YOU NOT TO BE UNHAPPY
I need some help, that much seems certain.
WHAT WOULD IT MEAN TO YOU IF YOU GOT SOME HELP
Perhaps I could learn to get along with my mother.
TELL ME MORE ABOUT YOUR FAMILY
My mother takes care of me.
WHO ELSE IN YOUR FAMILY TAKES CARE OF YOU

Pattern matching in Eliza

Input pattern	Response pattern
X me Y	X you Y
I remember X	Why do you remember X right now?
My {family member} is Y	Who else in your family is Y?
X {family member}Y	Tell me more about your family
<no recognised pattern>	Tell me more about that

More recent directions in research on conversational agents

- Gestures
- Backchanneling
- Emotions
- Question answering using structured and unstructured knowledge bases (e.g., the Internet)
- ...

Max: an example of an embodied conversational agent

courtesy of Ipke Wachsmuth, `https://www.techfak.uni-bielefeld.de/~ipke/`

Summary

- Stages of communication
- English grammar
- Parse trees
- Conversational agents