INF5390 – Kunstig intelligens

Natural Language Communication

Roar Fjellheim

Outline

- Communication and action
- Language structures
- Parsing and semantics
- Steps of communication
- Machine translation
- Summary

AIMA Chapter 23: Natural Language for Communication

Communication and language

- One definition of *communication*
  - √ Communication is the intentional exchange of information brought about by the production and perception of signs drawn from a shared system of a limited number of conventional signs
- Humans use *language* to communicate
  - √ Language is a “shared system of a limited number of conventional signs”
  - √ Its structure is sufficiently rich to allow an unbounded number of qualitatively different messages

Communication as action

- To produce messages in a language is one of the *actions* available to an agent
- This action is called a *speech act* (can be spoken, written, etc.)
- In a speech act, an *utterance* consisting of *words* is delivered from a *speaker* to a *hearer*
- Different types of speech acts serve different purposes

Some types of speech acts

Inform	Provide information to hearer
Query	Ask for information
Answer	Inform in response to query
Request	Ask hearer to perform action
Deny	Refuse to perform action
Command	Request with no option to deny
Promise	Commit to future action
Offer	Propose to do future action
Acknowledge	Confirm e.g. request or offer
….	

Planning and understanding speech acts

- Deciding *when* a speech act is called for, and decide *which* one to use, is equivalent to *planning*
- *Understanding* a speech act is similar to *diagnosis* or *plan recognition*
- I.e., one can use methods from other parts of AI in implementing *perception* and *action* in communicating agents

Natural and formal languages

- *Natural languages* are a rich field of empirical and logical study, including in AI
- *Formal languages* are invented ones, in contrast to natural languages, and include logic, etc.
- Formal language *concepts* are being used in analysis of natural languages

Formal language concepts

- A formal language is a set of *strings* (*sentences*)
  - "The wumpus is dead"
- A string is a sequence of symbols taken from a finite set called the *terminal symbols* (*words*)
  - "dead", "is", "wumpus", "the"
- A *phrase* is a substring of a sentence. There are different categories (symbolized by *nonterminal symbols*) of phrases
  - *NP* (noun phrase): "the wumpus"
  - *VP* (verb phrase): "is dead"

Formal language concepts (cont.)

- The structure (grammar) of a language can be defined using a *phrase structure*, i.e. combinations of terminal and nonterminal symbols
  - √ *NP VP*
- *Rewrite rules* define how a single nonterminal symbol (phrase) may be replaced by a structure
  - √ *S* → *NP VP*

A grammar for a fragment of English

- Lexicon
  - List of valid words
  - Categories: Noun, verb, adjective, ..
- Grammar
  - Rules for valid sentences
  - Nonterminals: Sentence (S), noun phrase (NP) ..
- Parsing
  - Analyze a given sequence of lexicon words as a tree-structure allowed by grammar rules

Lexicon of the fragment

$$
\begin{aligned}
Noun &\rightarrow \mathbf{\mathit{stench}} \mid \mathbf{\mathit{breeze}} \mid \mathbf{\mathit{glitter}} \mid \mathbf{\mathit{nothing}} \\
&\quad \mid \mathbf{\mathit{wumpus}} \mid \mathbf{\mathit{pit}} \mid \mathbf{\mathit{pits}} \mid \mathbf{\mathit{gold}} \mid \mathbf{\mathit{east}} \mid \ldots \\
Verb &\rightarrow \mathbf{\mathit{is}} \mid \mathbf{\mathit{see}} \mid \mathbf{\mathit{smell}} \mid \mathbf{\mathit{shoot}} \mid \mathbf{\mathit{feel}} \mid \mathbf{\mathit{stinks}} \\
&\quad \mid \mathbf{\mathit{go}} \mid \mathbf{\mathit{grab}} \mid \mathbf{\mathit{carry}} \mid \mathbf{\mathit{kill}} \mid \mathbf{\mathit{turn}} \mid \ldots \\
Adjective &\rightarrow \mathbf{\mathit{right}} \mid \mathbf{\mathit{left}} \mid \mathbf{\mathit{east}} \mid \mathbf{\mathit{south}} \mid \mathbf{\mathit{back}} \mid \mathbf{\mathit{smelly}} \mid \ldots \\
Adverb &\rightarrow \mathbf{\mathit{here}} \mid \mathbf{\mathit{there}} \mid \mathbf{\mathit{nearby}} \mid \mathbf{\mathit{ahead}} \\
&\quad \mid \mathbf{\mathit{right}} \mid \mathbf{\mathit{left}} \mid \mathbf{\mathit{east}} \mid \mathbf{\mathit{south}} \mid \mathbf{\mathit{back}} \mid \ldots \\
Pronoun &\rightarrow \mathbf{\mathit{me}} \mid \mathbf{\mathit{you}} \mid \mathbf{\mathit{I}} \mid \mathbf{\mathit{it}} \mid \ldots \\
Name &\rightarrow \mathbf{\mathit{John}} \mid \mathbf{\mathit{Mary}} \mid \mathbf{\mathit{Boston}} \mid \mathbf{\mathit{UCB}} \mid \mathbf{\mathit{PAJC}} \mid \ldots \\
Article &\rightarrow \mathbf{\mathit{the}} \mid \mathbf{\mathit{a}} \mid \mathbf{\mathit{an}} \mid \ldots \\
Preposition &\rightarrow \mathbf{\mathit{to}} \mid \mathbf{\mathit{in}} \mid \mathbf{\mathit{on}} \mid \mathbf{\mathit{near}} \mid \ldots \\
Conjunction &\rightarrow \mathbf{\mathit{and}} \mid \mathbf{\mathit{or}} \mid \mathbf{\mathit{but}} \mid \ldots \\
Digit &\rightarrow \mathbf{0} \mid \mathbf{1} \mid \mathbf{2} \mid \mathbf{3} \mid \mathbf{4} \mid \mathbf{5} \mid \mathbf{6} \mid \mathbf{7} \mid \mathbf{8} \mid \mathbf{9}
\end{aligned}
$$

Divided into closed and open classes

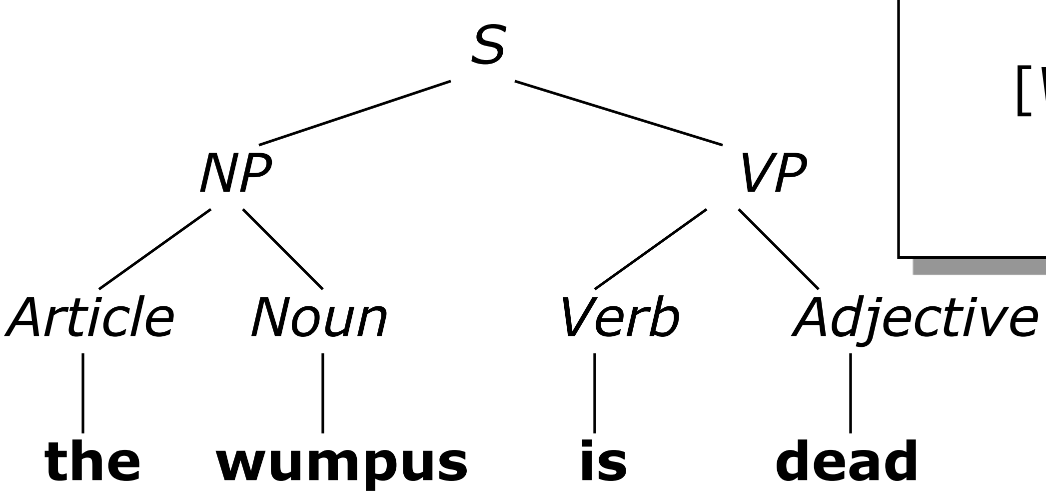
Grammar of the fragment

S	$\rightarrow$	$NP\ VP$	I + feel a breeze
	$\mid$	$S\ Conjunction\ S$	I feel a breeze + and + I smell a wumpus
NP	$\rightarrow$	$Pronoun$	I
	$\mid$	$Noun$	pits
	$\mid$	$Article\ Noun$	the + wumpus
	$\mid$	$Digit\ Digit$	3 4
	$\mid$	$NP\ PP$	the wumpus + to the east
	$\mid$	$NP\ RelClause$	the wumpus + that is smelly
VP	$\rightarrow$	$Verb$	stinks
	$\mid$	$VP\ NP$	feel + a breeze
	$\mid$	$VP\ Adjective$	is + smelly
	$\mid$	$VP\ PP$	turn + to the east
	$\mid$	$VP\ Adverb$	go + ahead
PP	$\rightarrow$	$Preposition\ NP$	to + the east
$RelClause$	$\rightarrow$	$\textbf{\textit{that}}\ VP$	that + is smelly

Parsing

- Search for a *parse tree* for a given sentence, e.g.

  PARSE(”**the wumpus is dead**”, grammar, *S*)



[*S*: [*NP*: [*Article*: **the**]
[*Noun*: **wumpus**]]
[*VP*: [*Verb*: **is**]
[*Adjective*: **dead**]]]

Top-down vs. bottom-up parsing

- Top-down parsing
  - Initial parse tree is the root with unknown children [*S*: ?]
  - At each step, select leftmost node in the tree with unknown children and look for grammar rules with LHS that matches the node. Replace ? with RHS and repeat
  - Stop when leaves of the tree exactly matches the string
- Bottom-up parsing
  - Initial list of words, seen as list of singleton parse trees
  - At each step, replace each sequence of parse trees that matches an RHS of a grammar rule, with the corresponding LHS, and repeat
  - Stop when the tree is the single node *S*

Semantic interpretation

- Having analyzed the sentence, we need to interpret its *meaning*; i.e. decide its semantic content
- We adopt first-order logic (FOL) as the representation language
  - √ E.g., ”the wumpus is dead and John loves Mary” has the meaning: $Dead(Wumpus) \wedge Loves(John, Mary)$
- Compositional semantics
  - √ The meaning of the entire sentence is composed of the meanings of its constituents

Augmenting grammar for semantics

- Each category of the grammar is *augmented* with a single argument that represents the semantics
  - √ *NP* becomes *NP*(*obj*) - where *obj* is the FOL term that represents the noun phrase
  - √ VP becomes VP(*rel*) - where *rel* is the FOL relation (predicate) that represents the verb
  - √ Also needs λ-expressions for verbs:
    - λ*x Loves*(*x* , *Mary*) - the predicate of variable *x* such that *x* loves *Mary*
    - (λ*x Loves*(*x* , *Mary*))(*John*) - the predicate applied to the argument *John*, yielding *Loves*(*John*, *Mary*)

Semantically augmented grammar fragment

$S(rel(obj)) \rightarrow NP(obj)\ VP(rel)$

$VP(rel(obj)) \rightarrow Verb(rel)\ NP(obj)$

$NP(obj) \rightarrow Name(obj)$

$Name(John) \rightarrow$ **John**

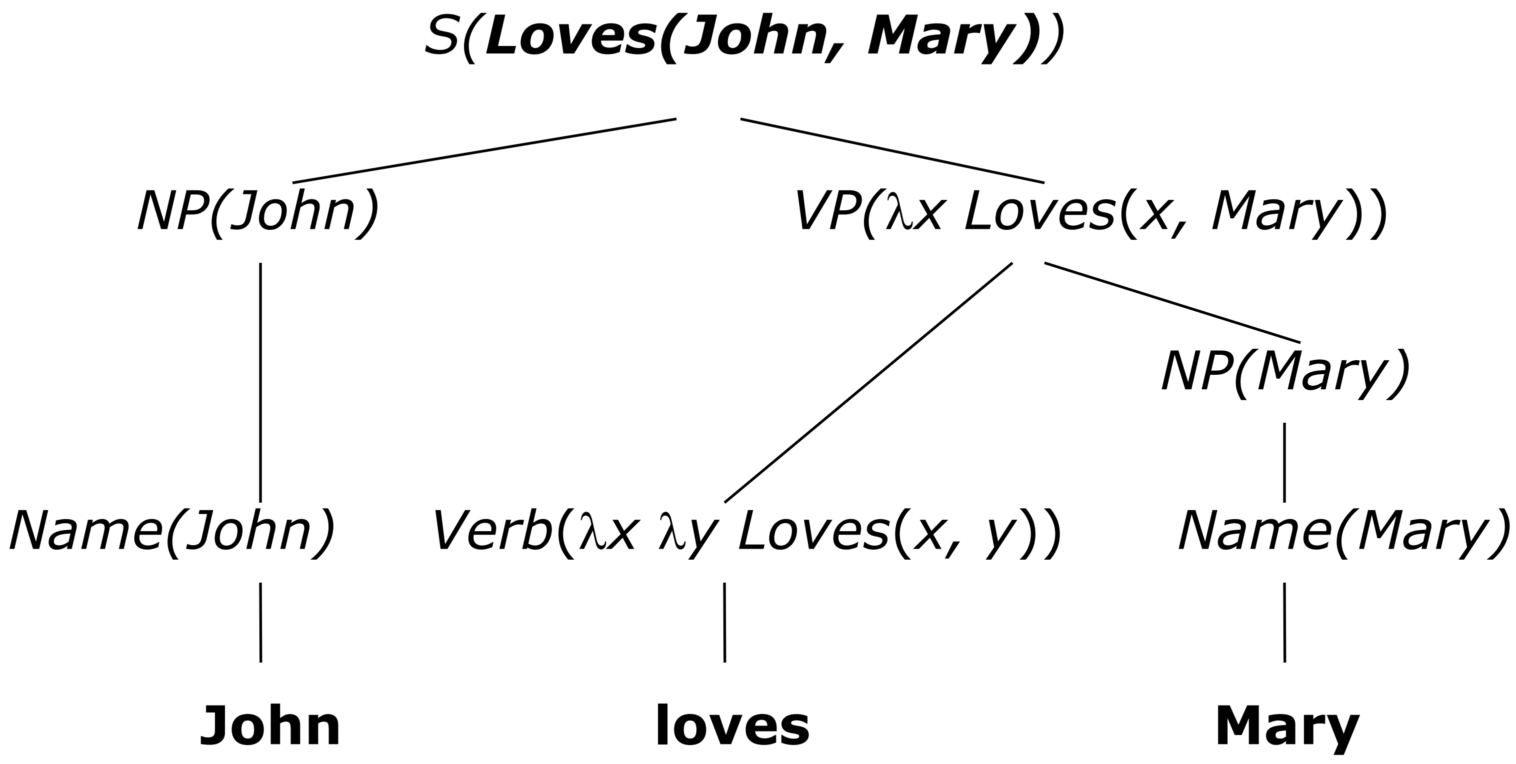
$Name(Mary) \rightarrow$ **Mary**

$Verb(\lambda x\ \lambda y\ Loves(x, y)) \rightarrow$ **loves**

Can be extended:

- Time
- Tense
- Quantification
- Pragmatics
- Etc.

Deriving semantics during parsing



Steps of communication

Speaker *S* wants to convey proposition *P*
to hearer *H* using words *W*

Speaker S

- *Intention*
  S wants *H* to believe *P*
- *Generation*
  S chooses the words *W*
- *Synthesis*
  S utters the words *W*

Hearer H

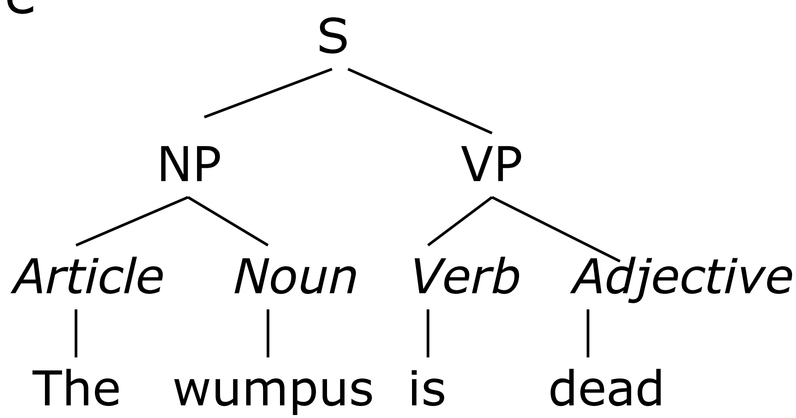
- *Perception*
  H perceives *W'* (ideally=*W*)
- *Analysis*
  H infers that *W'* may mean P_1, .., P_n
- *Disambiguation*
  H infers that *S* intended P_i (ideally=*P*)
- *Incorporation*
  H decides to (dis)believe P_i

Speaker steps in more detail

- *Intention*
  - Speaker decides *that* there is something to say, e.g. by reasoning about beliefs and goals of hearer
    - *Know*(*H*, ¬*Alive*(*Wumpus*, *S3*))
- *Generation*
  - Speaker uses knowledge about language in deciding *what* to say
    - "The wumpus is dead"
- *Synthesis*
  - Finally, the sentence is *uttered* via the "speech act organ" (printer, screen, speech synthesizer, ..)

Hearer steps in more detail

- *Perception*
  - The utterance is received, e.g. by speech recognition, scene analysis, ..
- *Analysis*
  - *Parsing*: Recognizing constituent phrases (*parse tree*)
  - *Interpretation*: Extract meaning as expression in e.g. logic



$\neg$*Alive*(*Wumpus*, *S3*)

Tired(*Wumpus*, *S3*)

Hearer steps in more detail (cont.)

- *Disambiguation*
  - Analysis may yield different interpretations, and the agent must choose the most *probable* one, e.g. using probabilistic reasoning
    - $\neg$*Alive*(*Wumpus*, *S3*)
- *Incorporation*
  - Finally, the agent updates its knowledge base with the new information
    - TELL(*KB*, $\neg$*Alive*(*Wumpus*, *S3*))

Machine translation (MT)

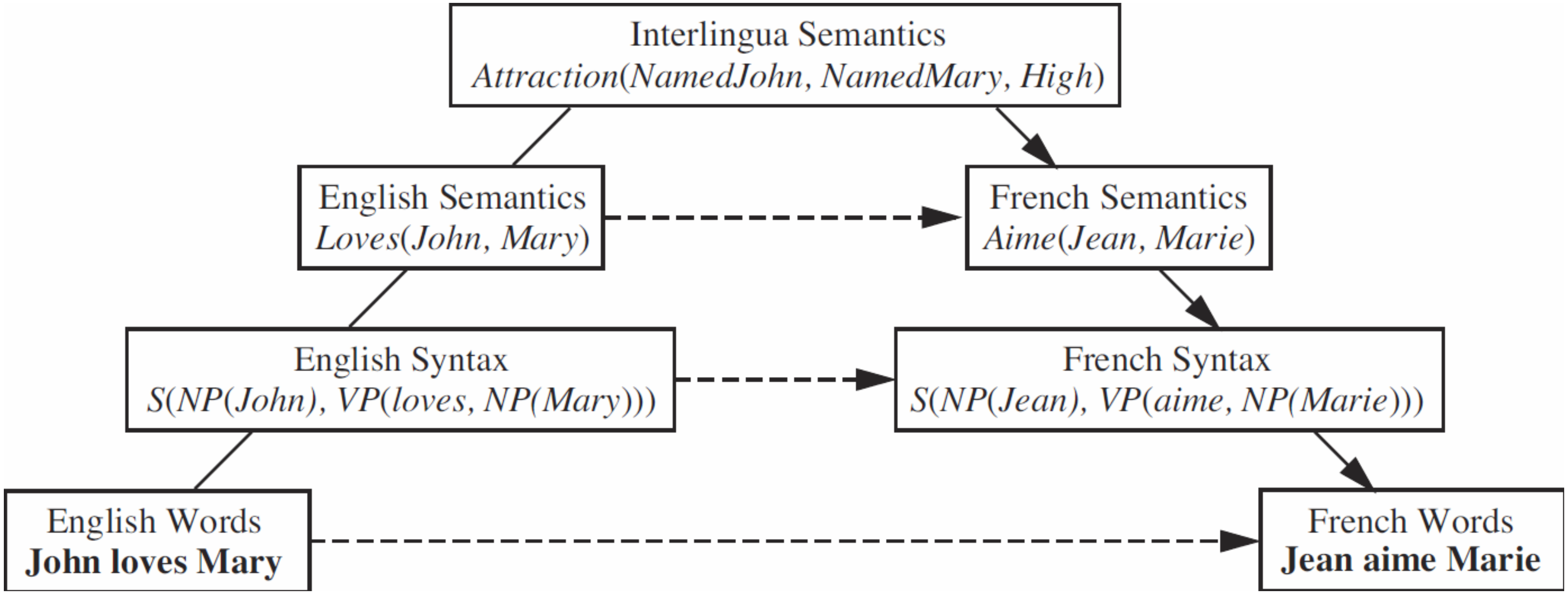
- Machine translation is the automatic translation of one natural language (the *source*) to another language (the *target*)

MT by deep linguistic analysis

- MT by three-step process
  1. Analyze source text syntactically and semantically
  2. Create deep knowledge representation of meaning of source text
  3. Generate target text representing the same meaning in target syntax
- Can use methods described earlier for natural language communication, but problematic
  - √ Requires rich semantic model (FOL not sufficient?)
  - √ Requires strong parsing and generation capabilities

MT by using transfer model

- Large database of translation rules and examples on lexical, syntactic and semantic levels
- Can match rules/examples on any level



Statistical machine translation

- Successful MT systems (e.g. Google Translate) are built by training probabilistic models using statistics from large text collections
- Does not need complex ontologies or grammars of source and target languages
- Relies on large amounts of sample translations from which a transfer model can be learned

Summary

- Agents need to *communicate* in order to achieve certain goals, such as getting the other agent to believe something or to do something
- Sending a signal is called a *speech act*, of which many types may be identified: inform, request, deny, promise, etc.
- *Formal languages* (incl. subsets of natural language) used for communication may be defined by a *lexicon* and a *grammar*
- Efficient techniques have been developed for *parsing* the structure of sentences and interpreting the intended *semantics*

Summary (cont.)

- Communication involves *speaker* and *hearer* steps, methods exist to handle each of the steps for a range of formal languages
- *Machine translation* (MT) systems automatically translate from a *source* to a *target* language
- Most successful MT systems are based on probabilistic models built from large collections of translation samples
- In addition to language communication, (some) agents need to interact with their environment through vision, tactile sensing, robotic locomotion and manipulation, etc.