Foundations of AI

16. Natural Language Processing

Wolfram Burgard, Bernhard Nebel, and Luc De Raedt
Contents

- Communication
- Phrase Structure Grammar
- Syntactic Analysis
- Perspective
Speech Acts

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 in context C
- **Synthesis:** S utters words W
- **Perception:** H perceives W in context C’
- **Analysis:** H infers possible meanings P1, ..., Pn
- **Disambiguation:** H infers intended meaning Pi
- **Incorporation:** H incorporates Pi into KB

How could this go wrong?

- **Insincerity:** S doesn’t believe P
- Speech wreck ignition failure
- Ambiguous utterance
- Differing understanding of current context C<>C’
Stages in communication (informing)

Intention:

\[ \text{Know}(H, \neg \text{Alive}(\text{Wumpus}, S_3)) \]

Generation:

"The wumpus is dead"

Synthesis:

[thaxwahmpaxsihnzdhehd]

Perception:

"The wumpus is dead"

Analysis: (Parsing):

\[ S \]

\[ NP \]

\[ \text{Article} \]

\[ \text{The} \]

\[ \text{Noun} \]

\[ \text{wumpus} \]

\[ \text{Verb} \]

\[ \text{is} \]

\[ \text{Adjective} \]

\[ \text{dead} \]

(Semantic Interpretation):

\[ \neg \text{Alive}(\text{Wumpus}, \text{Now}) \]

\[ \text{Tired}(\text{Wumpus}, \text{Now}) \]

(Pragmatic Interpretation):

\[ \neg \text{Alive}(\text{Wumpus}, S_3) \]

\[ \text{Tired}(\text{Wumpus}, S_3) \]

Disambiguation:

\[ \neg \text{Alive}(\text{Wumpus}, S_3) \]

Incorporation:

TELL( KB,

\[ \neg \text{Alive}(\text{Wumpus}, S_3) \]
Speech Acts

- Animals use isolated symbols for sentences
  - restricted set of communicable propositions, no \textit{generative capacity}
- \textbf{Grammar} specifies the compositional structure of complex message (speech, language, music)
- A \textbf{formal language} is a set of \textit{strings} of \textit{terminal symbols}
- Each string in the language can be analyzed/generated by the grammar
- The grammar is a set of rewrite rules,
  - e.g.,
    \[
    S \rightarrow NP \ VP \\
    Article \rightarrow \textit{the} \mid \textit{a} \mid \textit{an} \mid \ldots
    \]
- Here \(S\) is the \textit{sentence} symbol, \(NP\) and \(VP\) are \textit{non-terminals}, and \textit{the}, \textit{a} ... \textit{terminals}
Grammar Types

Regular: \( \text{nonterminal} \rightarrow \text{terminal}[\text{nonterminal}] \)

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

Context-free: \( \text{nonterminal} \rightarrow \text{anything} \)

\[
S \rightarrow aSb
\]

Context-sensitive: more nonterminals on right-hand side

\[
ASB \rightarrow AAbBB
\]

Recursively enumerable: no constraints

Natural language probably context-free
The 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
    | right | left | east | south | back | ...
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 open and closed classes
The Wumpus Grammar

\[ S \rightarrow NP \ VP \quad \text{I + feel a breeze} \]
\[ | \quad S \ \text{Conjunction} \ S \quad \text{I feel a breeze + and + I smell a v} \]

\[ NP \rightarrow Pronoun \quad \text{I} \]
\[ | \quad Noun \quad \text{pits} \]
\[ | \quad Article \ Noun \quad \text{the + wumpus} \]
\[ | \quad Digit \ Digit \quad 3 \ 4 \]
\[ | \quad NP \ PP \quad \text{the wumpus + to the east} \]
\[ | \quad NP \ \text{RelClause} \quad \text{the wumpus + that is smelly} \]

\[ VP \rightarrow Verb \quad \text{stinks} \]
\[ | \quad VP \ NP \quad \text{feel + a breeze} \]
\[ | \quad VP \ Adjective \quad \text{is + smelly} \]
\[ | \quad VP \ PP \quad \text{turn + to the east} \]
\[ | \quad VP \ Adverb \quad \text{go + ahead} \]

\[ PP \rightarrow Preposition \ NP \quad \text{to + the east} \]
\[ \text{RelClause} \rightarrow \text{that} \ VP \quad \text{that + is smelly} \]
Grammatical Judgments

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

Adjusting is a learning problem

Real grammars typically 10-500 pages, insufficient for proper English
Parse Trees

Exhibit the grammatical structure of a sentence

```
S
  / \   \
NP  VP
  /  \  /
Pronoun Verb NP
  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.”
Parsing CFGs

Many different parsing algorithms
- Top-down
- Bottom-up
- Chart-parsing (aka CYK algorithm)
  - Dynamic programming
  - \(O(n^3)\)

\[
[S: [NP :[Pronoun:I]]
  [VP :[VP:[Verb:shoot]]
    [NP:[Article:the][Noun:wumpus]]]]
\]
Top-down parsing

- Initial state : \([S: \, ?]\)
- Successor function:
  - Select rules for leftmost node in tree with unknown children and apply
  - \([S:[S:?]\text{[Conjunction:?]\,}[S:?]\])
  - \([S:[NP:?]\text{[VP:?]\,}]\)
- Goal Test:
  - Check whether leaves of the parse tree correspond to (complete) input

Problem for Top-down: left-recursive rules \(S \rightarrow S \text{ Conjunction } S\)
infinite loops
Top-down parsing

Successful top-down parse

[S: ?]
[S:[NP:?][VP:?]]
[S:[NP:[Art:?][Noun:?]][VP:?]]
[S:[NP:[Art:The][Noun:wumpus]][VP:?]]
[S:[NP:[Art:The][Noun:wumpus]][VP:[Verb:?]]]
[S:[NP:[Art:The][Noun:wumpus]][VP:[Verb:stinks]]]

Alternative notation

S  [The, wumpus, stinks]
NP VP [The, wumpus, stinks]
Art Noun VP [The, wumpus, stinks]
Noun VP [wumpus, stinks]
VP [stinks]
Verb [stinks]
[] []
Bottom-up parsing

- Initial state
  - [the, wumpus, is, dead]

- Successor function
  - If subsequence at pos i matches right-hand of rule then replace by left hand
  - [[Art:the],wumpus, is, dead]

- Goal state :
  - A single state with root S

Problem for Bottom up:

Art -> []
Bottom-up parsing

- [the, wumpus, is, dead]
- [[Art:the],wumpus,is,dead]
- [[Art:the][Noun:wumpus],is,dead]
- [[NP:[Art:the][Noun:wumpus]],is,dead]
- [[NP:[Art:the][Noun:wumpus]],[Verb:is],dead]
- [[NP:[Art:the][Noun:wumpus]],[VP:[Verb:is]],dead]
- [[NP:[Art:the][Noun:wumpus]],[VP:[Verb:is]],[Adj:dead]]
- [[NP:[Art:the][Noun:wumpus]],[VP:[VP:[Verb:is]],[Adj:dead]]]
- [S:{{NP:[Art:the][Noun:wumpus]],[VP:[VP:[Verb:is]],[Adj:dead]]}]}
## Bottom-up parsing

<table>
<thead>
<tr>
<th>step</th>
<th>list of nodes</th>
<th>subsequence</th>
<th>rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>Init</td>
<td>the wumpus is dead</td>
<td>the</td>
<td>Article → the</td>
</tr>
<tr>
<td>2</td>
<td>Article wumpus is dead</td>
<td>wumpus</td>
<td>Noun → wumpus</td>
</tr>
<tr>
<td>3</td>
<td>Article Noun is dead</td>
<td>Article Noun</td>
<td>NP → Article Noun</td>
</tr>
<tr>
<td>4</td>
<td>NP is dead</td>
<td>is</td>
<td>Verb → is</td>
</tr>
<tr>
<td>5</td>
<td>NP Verb dead</td>
<td>dead</td>
<td>Adjective → dead</td>
</tr>
<tr>
<td>6</td>
<td>NP Verb Adjective</td>
<td>Verb</td>
<td>VP → Verb</td>
</tr>
<tr>
<td>7</td>
<td>NP VP Adjective</td>
<td>VP Adjective</td>
<td>VP → VP Adjective</td>
</tr>
<tr>
<td>8</td>
<td>NP VP</td>
<td>NP VP</td>
<td>S → NP VP</td>
</tr>
<tr>
<td>Goal</td>
<td>$S$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Chart parsing

- Have the students in B.Sc. Informatik take the exam of AI.
- Have the students in B.Sc. Informatik taken the exam of AI?
- Double work
  - Dynamic programming - combine bottom-up and top-down
- Chart:
  - $N+1$ vertices
  - Labeled edges, e.g., $[0,2 \ S \rightarrow \ NP \ *VP]$
    - Denotes that from 0 to 2 we have a NP, and if we find a VP, from 2 to k then we have an S from 0 to k
Chart parsing

- **Initialization Convention**: 
  - Add \([0,0 \text{ S'} -> * \text{ S}]\) non-terminal - upper case
  - Add edge \([i,j \text{ A} -> \text{ B} * \text{ c}]\) sequence - lower case
    - if c is empty then call extender
    - else call predictor
- **Predictor** \([i,j \text{ A} -> \text{ b} * \text{ C e}]\) - top-down
  - With (all) rules \(\text{ C} -> \text{ d}\)
  - Add edge \([i,j \text{ A} -> \text{ b} * \text{ d e}]\)
- **Extender** \([j,k \text{ B} -> \text{ c} *]\) - bottom-up
  - With (all) edges \([i,j \text{ A} -> \text{ e} * \text{ B f}]\)
  - Add edge \([i,k \text{ A} -> \text{ e c} * \text{ f}]\)
- **Scanner** \([j,k \text{ A} -> \text{ c} * \text{ D e}]\) - bottom-up
  - Word of type \(\text{ D}\) at position \(k\)
  - Add edge \([j,k+1 \text{ A} -> \text{ c D} * \text{ e}]\)
function CHART-PARSE(words, grammar) returns chart

chart ← array[0... LENGTH(words)] of empty lists
ADD-EDGE([0, 0, S' → • S])
for i ← from 0 to LENGTH(words) do
  SCANNER(i, words[i])
return chart

procedure ADD-EDGE(edge)
  /* Add edge to chart, and see if it extends or predicts another edge. */
  if edge not in chart[END(edge)] then
    append edge to chart[END(edge)]
    if edge has nothing after the dot then EXTENDER(edge)
  else PREDICTOR(edge)

procedure SCANNER(j, word)
  /* For each edge expecting a word of this category here, extend the edge. */
  for each [i, j, A → α • B β] in chart[j] do
    if word is of category B then
      ADD-EDGE([i, j+1, A → α B • β])

procedure PREDICTOR([i, j, A → α • B β])
  /* Add to chart any rules for B that could help extend this edge */
  for each (B → γ) in REWRITES-FOR(B, grammar) do
    ADD-EDGE([j, j, B → • γ])

procedure EXTENDER([j, k, B → γ •])
  /* See what edges can be extended by this edge */
  e_B ← the edge that is the input to this procedure
  for each [i, j, A → α • B' β] in chart[j] do
    if B = B' then
      ADD-EDGE([i, k, A → α e_B • β])

Figure 22.9
A completed chart
## Chart Parse Trace

<table>
<thead>
<tr>
<th>Edge</th>
<th>Procedure</th>
<th>Derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>INITIALIZER</td>
<td>([0, 0, S' \rightarrow S] )</td>
</tr>
<tr>
<td>b</td>
<td>PREDICTOR(a)</td>
<td>([0, 0, S \rightarrow NP \ VP] )</td>
</tr>
<tr>
<td>c</td>
<td>PREDICTOR(b)</td>
<td>([0, 0, NP \rightarrow Pronoun )</td>
</tr>
<tr>
<td>d</td>
<td>SCANNER(c)</td>
<td>([0, 1, NP \rightarrow Pronoun \bullet] )</td>
</tr>
<tr>
<td>e</td>
<td>EXTENDER(b,d)</td>
<td>([0, 1, S \rightarrow NP \bullet VP] )</td>
</tr>
<tr>
<td>f</td>
<td>PREDICTOR(e)</td>
<td>([1, 1, VP \rightarrow Verb] )</td>
</tr>
<tr>
<td>g</td>
<td>PREDICTOR(e)</td>
<td>([1, 1, VP \rightarrow VP \ NP] )</td>
</tr>
<tr>
<td>h</td>
<td>SCANNER(f)</td>
<td>([1, 2, VP \rightarrow Verb \bullet] )</td>
</tr>
<tr>
<td>i</td>
<td>EXTENDER(g,h)</td>
<td>([1, 2, VP \rightarrow VP \bullet NP] )</td>
</tr>
<tr>
<td>j</td>
<td>PREDICTOR(g)</td>
<td>([2, 2, NP \rightarrow Pronoun] )</td>
</tr>
<tr>
<td>k</td>
<td>SCANNER(j)</td>
<td>([2, 3, NP \rightarrow Pronoun \bullet] )</td>
</tr>
<tr>
<td>l</td>
<td>EXTENDER(i,k)</td>
<td>([1, 3, VP \rightarrow VP NP \bullet] )</td>
</tr>
<tr>
<td>m</td>
<td>EXTENDER(e,l)</td>
<td>([0, 3, S \rightarrow NP \ VP \bullet] )</td>
</tr>
</tbody>
</table>

**Figure 22.9** Trace of a parse of “I feel it.” For each edge a-m, we show the procedure used to derive the edge from other edges already in the chart. Some edges were omitted for brevity.
Definite Clause Grammars

- A form of **unification** based grammar
- Non-terminals become atomic expressions
  - $s(\text{Num}) \rightarrow np(\text{Num}), vp(\text{Num})$
  - $\text{Numb}$ is a variable
  - Binds to **singular** and **plural**
- Employ unification during parsing.
- Directly executable in Prolog
- Rules directly translate to (definite) clause logic
- Next slides employ Prolog notation
  - Running examples with, e.g., YAP Prolog or SWI Prolog
Non-terminals with arguments

sentence --> noun_phrase(N), verb_phrase(N).
noun_phrase --> article(N), noun(N).
verb_phrase(N) --> intransitive_verb(N).
article(singular) --> [a].
article(singular) --> [the].
article(plural) --> [the].
noun(singular) --> [turtle].
noun(plural) --> [turtles].
intransitive_verb(singular) --> [sleeps].
intransitive_verb(plural) --> [sleep].

Agreement
Case Marking

pronoun(singular,nominative) --> [he]; [she]
pronoun(singular,accusative) --> [him]; [her]
pronoun(plural,nominative) --> [they]
pronoun(plural,accusative) --> [them]
sentence --> np(Number,nominative), vp(Number)
vp(Number) --> v(Number), np(X,accusative)
np(Number,Case) --> pronoun(Number,Case)
np(Number,X) --> det, n(Number)

He sees her. She sees him. They see her.
But not Them see he.
Top-down parsing with DCG

sentence: [He, sees, her]
np(Num, nom), vp(Num): [He, sees, her]
pronoun(Num, nom), vp(Num): [He, sees, her]
    Num = sing !!!  pronoun(sing, nom) --> [he]
vp(sing): [sees, her]
v(sing), np(sing, accusative): [sees, her]
np(sing, accusative): [her]
pronoun(sing, accusative]: [her]
[]
Sub-categorization

\[
\begin{align*}
vp & \quad \rightarrow \quad v(1). & \quad v(1) & \rightarrow \quad [\text{sleep}] \\
vp & \quad \rightarrow \quad v(2), \ np. & \quad v(2) & \rightarrow \quad [\text{chase}] \\
vp & \quad \rightarrow \quad v(3), \ np, \ np. & & \quad \ldots \\
vp & \quad \rightarrow \quad v(4), \ s. & & \\
\end{align*}
\]

<table>
<thead>
<tr>
<th>Verb</th>
<th>Complement</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sleep</td>
<td>None</td>
<td>The cat slept</td>
</tr>
<tr>
<td>Chase</td>
<td>One NP</td>
<td>The cat chased the dog</td>
</tr>
<tr>
<td>Give</td>
<td>Two NP</td>
<td>John gave Bill the book</td>
</tr>
<tr>
<td>Say</td>
<td>sentence</td>
<td>John said he loved Mary.</td>
</tr>
</tbody>
</table>
Constructing parse trees

\[
\begin{align*}
\text{sentence}(s(NP,VP)) & \quad \rightarrow \quad \text{noun_phrase}(NP),\text{verb_phrase}(VP). \\
noun_phrase(np(N)) & \quad \rightarrow \quad \text{proper_noun}(N). \\
noun_phrase(np(Art,Adj,N)) & \quad \rightarrow \quad \text{article}(Art),\text{adjective}(Adj),\text{noun}(N). \\
noun_phrase(np(Art,N)) & \quad \rightarrow \quad \text{article}(Art),\text{noun}(N). \\
\text{verb_phrase}(vp(IV)) & \quad \rightarrow \quad \text{intransitive_verb}(IV). \\
\text{verb_phrase}(vp(TV,NP)) & \quad \rightarrow \quad \text{transitive_verb}(TV),\text{noun_phrase}(NP). \\
\text{article}(art(the)) & \quad \rightarrow \quad [the]. \\
\text{adjective}(adj(lazy)) & \quad \rightarrow \quad [lazy]. \\
\text{adjective}(adj(rapid)) & \quad \rightarrow \quad [rapid]. \\
\text{proper_noun}(pn(achilles)) & \quad \rightarrow \quad [achilles]. \\
\text{noun}(n(turtle)) & \quad \rightarrow \quad [turtle]. \\
\text{intransitive_verb}(iv(sleeps)) & \quad \rightarrow \quad [sleeps]. \\
\text{transitive_verb}(tv(beats)) & \quad \rightarrow \quad [beats].
\end{align*}
\]
Semantic Interpretation

np(fido)   -->  [fido].
np(felix)  -->  [felix].

v(X^slept(X))   -->  [slept].
v(Y^(X^chased(X,Y))) -->  [chased].

s(Pred)  -->  np(Subj), vp(Subj^Pred).

vp(Subj^Pred)  -->  v(Subj^Pred).
vp(Subj^Pred)  -->  v(Obj^(Subj^Pred)), np(Obj).

?-prhase(s(X),[fido, chased, felix]).
X = chased(fido,felix)
Lambda Expressions

$X^{\text{slept}(X)}$ stands for $\lambda X. \text{slept}(X)$

$Y^{(X^{\text{chased}(X,Y)})}$ stands for $\lambda Y \lambda X. \text{chased}(X,Y)$

These lambda-expressions denote relationships
They can be instantiated by unification

With unification

$X^{\text{slept}(X)}$ with $X=\text{fido}$ yields $\text{fido}^{\text{slept}(\text{fido})}$

unify $\text{fido}^{\text{slept}(\text{fido})}$ with $X^R$

Yields $R=\text{slept}(\text{fido})$

In lambda-calculus

$(\lambda X. \text{slept}(X)) \text{ fido}$ gives $\text{slept}(\text{fido})$
Natural Language Processing

- Complex process
  - Requiring knowledge, reasoning and AI in general
- Grammars to represent natural languages
- Parsing techniques for syntactic analysis
- DCG add expressive power to CFGs
  - Convenient in practice
  - Unification based grammars