1. Context-Free Grammars (CFGs)

Also known as phrase structure grammars

• And Backus-Naur form is a standardised approach to notating CFGs

Making explicit the restrictions on rewriting we started with earlier, a CFG consists of

Terminals
or terminal symbols: words (for now)

Non-terminals
or non-terminal symbols: Names for constituents in a language
• E.g., NP (noun phrase), VP (verb phrase), V (verb), S (sentence)

Rules
or productions, each of which is a pair of
a left-hand side
a single non-terminal
a right-hand side
a sequence of any number of terminals and non-terminals

For example: \[ VP \rightarrow V \ NP \]

Distinguished symbol
One of the non-terminals
• The starting point for all analyses
• Usually S
2. Some preliminary NP Rules

Some overly-simple rules for noun phrases:

\[
\begin{align*}
\text{NP} & \rightarrow \text{Det Nominal} \\
\text{NP} & \rightarrow \text{ProperNoun} \\
\text{Nominal} & \rightarrow \text{Noun} \mid \text{Nominal Noun}
\end{align*}
\]

These rules describe two kinds of NPs:

- One that consists of a determiner followed by a nominal
- And another that says that proper names are NPs.

The third rule illustrates:

- A disjunction
  - Two kinds of nominals
  - Not strictly speaking a rule
  - Rather a shorthand notation for two rules
- A recursive definition
  - Same non-terminal on the right and left-side of the rule
  - We can see how this works if we consider a noun phrase such as a delivery truck repair manual

\[\text{S} \rightarrow \text{NP VP}\]

3. A bit more detail on English Grammar

Huddleston and Pullum's *The Cambridge Grammar of the English Language* is 1860 pages long

- !

So we won't cover all of English by a very long way

- Just enough to uncover some key shortcomings of CFGs

We'll look briefly at

- Sentences
- Noun phrases
  - Agreement
- Verb phrases
  - Subcategorisation

4. Sentence Types

Declaratives

\[\text{A plane left}\]
Imperatives

Leaves!

Yes-no questions

Did the plane leave?

WH questions

When did the plane leave?

5. Noun Phrases, more carefully

We can identify three quite distinct types of noun phrases:

Pronouns

she, he, we, ...

Proper Nouns

Edinburgh, Star Wars, the Eiffel Tower, ...

Complex noun phrases

the next prime minister after Thatcher

Consider the following moderately complicated noun phrase:

the first three morning flights from Denver to Tampa leaving before 10

We’ll need something along the lines of the following tree
6. NP Structure

That big NP is really about flights

- That’s its central critical noun
- Let’s call that the head of the NP

We can dissect this kind of NP into:

- The constituents that can come before the head
- The constituents that can come after it

7. Before the nominal: Determiners

Complex noun phrases can start with determiners

\[ \text{CNP} \rightarrow \text{Det} \ 	ext{CNP} \]

Determiners can be

Simple lexical items

\[ \text{the, this, a, her} \]

(Recursive) possessives

\begin{align*}
\text{simple} & \quad \text{Robin's car} \\
\text{complex} & \quad \text{Robin’s youngest child’s toy}
\end{align*}

8. Before the nominal: Other premodifiers

Other premodifiers include

- Quantifiers, cardinals, ordinals:
every flight

three flights

first flight

- Adjectives and Adjective phrases:
  
  large cars
  extremely sleepy baby

There are constraints we haven't captured on the order of pre-modifiers:

- Between adjectives and quantifiers:
  
  every eligible candidate
  *eligible every candidate

- Between one adjective and another:
  
  big red bus
  ?red big bus

- Following a common linguistic convention, I'm using an initial asterisk to indicate a word sequence which is not in a (natural) language or cannot (should not) be accepted by a formal grammar.
- Likewise an initial question mark for a borderline in/out word sequence

9. The nominal: the head and its postmodifiers

Eventually (or even right away), we get to the Nominal

- Including the head, with or without compounding

The postmodifiers which stack up behind the head may include

- Prepositional phrases:
  
  flight from Seattle

- Non-finite clauses (gerundive, infinitive):
  
  flights arriving before noon
  first flight to depart

- Relative clauses:
  
  flights that serve breakfast
  people whom the pilot knows

Similar general (recursive) rules to handle these
10. Agreement

**Agreement**: when constraints hold among constituents that take part in a rule or set of rules

For example, in English, as in many other languages, determiners and the head nouns in NPs have to agree in number

- *this flight*  
- *this flights*

- *those flight*  
- those flights

11. The agreement problem for CFGs

Our earlier NP rules are clearly deficient since they don’t capture this constraint

That rule accepts, and assigns correct structures, to grammatical examples (*this flight*)
- But also accepts incorrect examples (*these flight*)

Such a rule is said to **overgenerate**
- *We’ll come back to this...*

12. Verb Phrases

English verb phrases consist of

- some optional modifiers
- a main verb
  - which we will once again call the **head**
- and zero or more **arguments**
  - also called **complements**

13. Verb Phrases: modifiers

We have to account for a range of structures ahead of the main verb

- Including adverbs, modals and auxiliary verbs
We get a familiar-looking right-branching structure when these combine

14. Subcategorisation

We need some rules for different patterns of arguments:

**intransitive**

- disappear

**transitive**

- prefer a morning flight

**transitive with PP**

- leave Boston in the morning

**ditransitive**

- buy Robin a ticket

Not all verbs are allowed to participate in all the VP rules

We **subcategorise** verbs in a language according to the sets of VP rules they participate in

This is a modern take on the traditional notion of transitive/intransitive.

Modern grammars may have 100s of subcategorisation classes
Sometimes called subcategorisation frames

15. Subcat examples and counterexamples

Some examples of the diversity of complement patterning

<table>
<thead>
<tr>
<th>John sneezed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please find a flight to Edinburgh</td>
</tr>
<tr>
<td>Can you help me with a flight</td>
</tr>
<tr>
<td>Give me a cheaper fare</td>
</tr>
<tr>
<td>Give a cheaper fare to my children</td>
</tr>
<tr>
<td>I prefer to leave earlier</td>
</tr>
<tr>
<td>I was told (that) KLM has a flight</td>
</tr>
</tbody>
</table>

And some counterexamples

| *John sneezed the book |
| *I prefer KLM has a flight |
| *Give with a flight |

As with agreement phenomena, we need a way to formally express the constraints