**Sentences**

The most common production rule for sentences is:

\[ S \rightarrow NP \ VP \]  

which accounts for simple sentences like Abigail ate the kimchi — as we will see, the direct object the kimchi is part of the verb phrase. But there are more complex forms of sentences as well:

\[ S \rightarrow ADVP \ NP \ VP \]
\[ S \rightarrow S \ CC \ S \]
\[ S \rightarrow VP \]

where ADVP is an adverbial phrase (e.g., unfortunately, very unfortunately) and CC is a coordinating conjunction (e.g., and, but).

---

**Noun phrases**

Noun phrases refer to entities, real or imaginary, physical or abstract: Asha, the steamed dumpling, parts and labor, nobody, and the rise of revolutionary syndicalism in the early twentieth century. Noun phrase productions include “bare” nouns, which may optionally follow determiners, as well as pronouns:

\[ NP \rightarrow \{ NN | NNs | NP | PP \} \]

The part-of-speech tags NN, NNs, and NP refer to singular, plural, and proper nouns; PP refers to personal pronouns, and DET refers to determiners. The grammar also contains terminal productions from each of these tags, e.g., PP \rightarrow \{ I \} \{ you \} \{ we \} ... 

Noun phrases may be modified by adjectival phrases (ADJP, e.g., the small Russian dog) and numbers (C (e.g., the five pastries), each of which may optionally follow a determiner:

\[ NP \rightarrow ADJP \ NN | ADJP \ NNs | DET \ ADJP \ NN | DET \ ADJP \ NNs \]

\[ NP \rightarrow \{ CD \} \{ NN | \}

Notice that the grammar does not include the recursive production S \rightarrow ADVP S. It may be helpful to think about why this production would cause the grammar to overgenerate.

Some noun phrases include multiple nouns, such as the liberation movement and an antelope horn, necessitating additional productions:

\[ NP \rightarrow \{ NN \} \{ NNs \} | DET \{ NN \} \{ NNs \} \]  

These multiple noun constructions can be combined with adjectival phrases and cardinal numbers, leading to a large number of additional productions.

Recursively produced noun phrases include coordination, prepositional phrase attachment, subordinate clauses, and verb phrase adjuncts:

\[ NP \rightarrow NP \ CC \ NP | NP \ PP | NP \ SBAR | NP \ VP \]

\[ e.g., \{ the \} \{ red \} \{ the \} \{ black \} \]
\[ e.g., the \{ President \} \{ of \} \{ the \} \{ Georgia \} \{ Institute \} \{ of \} \{ Technology \} \]
\[ e.g., \{ the \} \{ bicycle \} \{ that \} \{ I \} \{ found \} \{ outside \} \]
\[ e.g., \{ a \} \{ bicycle \} \{ made \} \{ of \} \{ titanium \} \]

These recursive productions are a major source of ambiguity, because the VP and PP non-terminals can also generate NP children. Thus, the President of the Georgia Institute of Technology can be derived in two ways, as can a bicycle made of titanium found outside.

---

**Verb phrases**

Verb phrases describe actions, events, and states of being. The PTB tagset distinguishes several classes of verb inflections: base form (VB; she likes to snack), present-tense third-person singular (VBP; she snacks), present tense but not third-person singular (VBP; they snack), past tense (VB; they snacked), present participle (VBG; they are snacking), and past participle (VBN; they had snacked). Each of these forms can constitute a verb phrase on its own:

\[ VP \rightarrow VB | VBG | VBD | VBN | VBP \]

More complex verb phrases can be formed by a number of recursive productions, including the use of coordination, modal verbs (MD; she should snack), and the infinitival to (To):

\[ VP \rightarrow MD \ VP \]
\[ VP \rightarrow VBD \ VP \]
\[ VP \rightarrow VBN \ VP \]
\[ VP \rightarrow To \ VP \]
\[ VP \rightarrow VP \ VP \]

Each of these productions uses recursion, with VP appearing on the right-hand side. This enables the creation of very complex verb phrases, such as she will have wanted to have been snacking.

**Transitive verbs**

Take noun phrases as direct objects, and ditransitive verbs take two direct objects:

\[ VP \rightarrow VBG \ NP \]
\[ VP \rightarrow VBD \ NP \]
\[ VP \rightarrow VBN \ NP \]

These productions are not recursive, so a unique production is required for each verb part-of-speech. They also do not distinguish transitive from intransitive verbs, so the resulting grammar overgenerates examples like she sleeps sushi and she learns Boyang algebra. Sentences can also be direct objects:

\[ VP \rightarrow VBG \ S \]
\[ VP \rightarrow VBD \ SBAR \]

Asha wants to eat the kimchi.

Asha knows that Boyang eats the kimchi.

The first production overgenerates, licensing sentences like *Asha sees Boyang eats the kimchi*. This problem could be addressed by designing a more specific set of sentence non-terminals, indicating whether the main verb can be conjugated.

Verbs can also be modified by prepositional phrases and adverbial phrases:

\[ VP \rightarrow VBG \ PP \]
\[ VP \rightarrow VBD \ ADVP \]
\[ VP \rightarrow VBN \ PP \]

Again, because these productions are not recursive, the grammar must include productions for every verb part-of-speech.

A special set of verbs, known as copula, can take predicator adjectives as direct objects:

\[ VP \rightarrow VBG \ ADJP \]
\[ VP \rightarrow VBD \ ADJP \]

The PTB does not have a special non-terminal for copular verbs, so this production generates non-grammatical examples such as *She wins tall*. The PTB also overgenerates examples such as *She wins tall*.

**Particles**

Particles (FR as a phrase, RP as a part-of-speech) work to create phrasal verbs:

\[ VP \rightarrow VB \ FR \]
\[ VP \rightarrow VB \ RP \]

They gave up their ill-gotten gains.

As the second production shows, particle productions are required for all configurations of verb parts-of-speech and direct objects.

---

**Other constituents**

The remaining constituents require far fewer productions. Prepositional phrases almost always consist of a preposition and a noun phrase:

\[ PP \rightarrow In \ NP \]
\[ PP \rightarrow To \ NP \]

United States of America

He gave his kimchi to Abigail

Similarly, complement clauses consist of a complementizer (usually a preposition, possibly null) and a sentence:

\[ SBAR \rightarrow In S \]
\[ SBAR \rightarrow S \]

She said that it was spicy

She said it was spicy

Adverbial phrases are usually bare adverbs (ADVP \rightarrow RB), with a few exceptions:

\[ ADVP \rightarrow RB \ RB \]
\[ ADVP \rightarrow RB \ PP \]

They went considerably further

They went considerably further than before

The tag RB is a comparative adverb.

Adjectival phrases extend beyond bare adjectives (ADJP \rightarrow JJ) in a number of ways:

\[ ADJP \rightarrow RB \ JJ \]
\[ ADJP \rightarrow RB \ JJ \]
\[ ADJP \rightarrow RB \ JJ \]
\[ ADJP \rightarrow JJ \ CC \ JJ \]
\[ ADJP \rightarrow JJ \ JJ \]
\[ ADJP \rightarrow RB \ VBN \]

very hungry

more hungry

best possible

even bigger

high and mighty

West German

previously reported

The tags JJ and RB refer to comparative and superlative adjectives respectively.

All of these phrase types can be coordinated:

\[ PP \rightarrow PP \ CC \ PP \]
\[ ADVP \rightarrow ADVP \ CC \ ADVP \]
\[ ADJP \rightarrow ADJP \ CC \ ADJP \]
\[ SBAR \rightarrow SBAR \ CC \ SBAR \]

on time and under budget

now and two years ago

quiet and rather deceptive

whether they want control

or whether they want exports

---

**Using this grammar**, and reasonable POS tags, please:

1. Draw parse trees for these phrases:
   - from Denver
   - redeye flights
   - I need to fly between Philadelphia and Atlanta.

2. Invent a new sentence that has a syntactic ambiguity. Draw it and two legitimate parse trees for it.