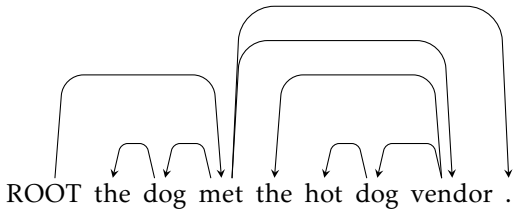
Context free dependency parsing

April 28, 2021

1. Consider the following unlabeled dependency tree:



   (a) Give the (smallest) bilexical context-free grammar allowing to parse the sentence and obtain the desired dependencies

   Solution: First the easy part:

   - $\Sigma = \{the, dog, met, hot, vendor, .\}$, $N = \{Xt : t \in \Sigma\} \cup \{ROOT\}$, $S = ROOT$

   Next we determine the production rules. We need a rule for the root of the tree:

   - $ROOT \rightarrow Xmet$

   The dependents of *met* are *dog* (on the left), *vendor* and . (on the right). We need 3 production rules to account for that:

   - $Xmet \rightarrow Xdog\ Xmet \mid Xmet\ Xvendor \mid Xmet\ X.$

   We do the same for all the other words with dependents:

   - $Xdog \rightarrow Xthe\ Xdog \mid Xhot\ Xdog$
   - $Xvendor \rightarrow Xthe\ Xvendor \mid Xdog\ Xvendor$

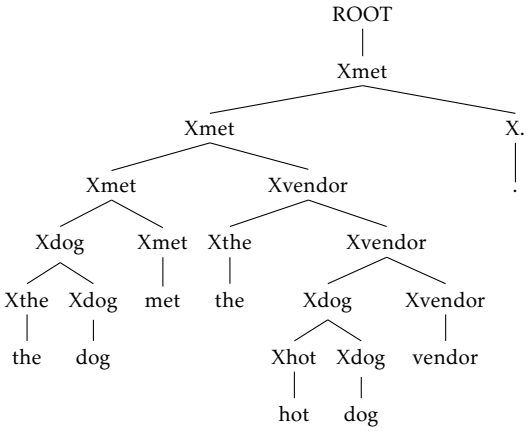
   Finally we have to add the terminal dependencies:

   - $Xthe \rightarrow the$
   - $Xdog \rightarrow dog$
   - $Xmet \rightarrow met$
   - $Xhot \rightarrow hot$
   - $Xvendor \rightarrow vendor$
   - $X. \rightarrow .$

   (b) Show one CYK parsing derivation for the sentence *the dog met the hot dog vendor .* (including the full stop) using the grammar determined in the previous step and draw the corresponding constituency tree

   Solution:

1.	[*Xthe*, 1, 1]	axiom
2.	[*Xdog*, 2, 2]	axiom
3.	[*Xmet*, 3, 3]	axiom
4.	[*Xthe*, 4, 4]	axiom
5.	[*Xhot*, 5, 5]	axiom
6.	[*Xdog*, 6, 6]	axiom
7.	[*Xvendor*, 7, 7]	axiom
8.	[*X.*, 8, 8]	axiom
9.	[*Xdog*, 1, 2]	combine(1, 2)
10.	[*Xmet*, 1, 3]	combine(9, 3)
11.	[*Xdog*, 5, 6]	combine(5, 6)
12.	[*Xvendor*, 5, 7]	combine(11, 7)
13.	[*Xvendor*, 4, 7]	combine(4, 12)
14.	[*Xmet*, 1, 7]	combine(10, 13)
15.	[*Xmet*, 1, 8]	combine(14, 8)
16.	[*ROOT*, 1, 8]	root(15)



(c) How many different constituency trees can be obtained in (b) ? Do they all encode the same dependencies?

Solution: *met* has one dependent on the left, two on the right, and there are 3 possible orders in which these dependents can be attached (lrr, rlr, rrl). All the other words have all their dependents either on the left or on the right. However, the grammar also allows to attach the second *the* to the second *dog* (which gives a different dependency tree), which brings the total number of different constituency trees to 6.

2. Show the split-head representation (constituency tree) of the dependency tree in Ex. 1

Solution: