DGfS-CL Fall School 2011:
Introduction to Tree Adjoining Grammar

Exercises for the first week

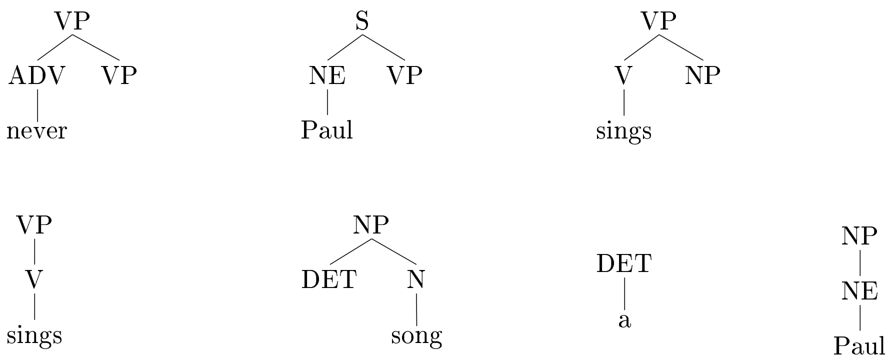
1. (a) Let G be a context-free grammar with the following rewriting rules:

S → NE VP	NP → DET N
VP → V	NP → NE
VP → V NP	VP → ADV VP
V → sings	
ADV → never	
NE → Paul	
DET → a	
N → song	

Give a TSG which strongly lexicalizes this CFG. Why is the resulting TSG not linguistically satisfying?

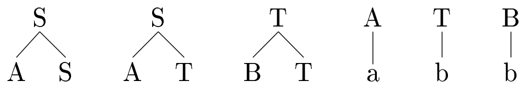
Solution:

The following TSG lexicalizes G. Note that the subcategorization frames are no longer localized:

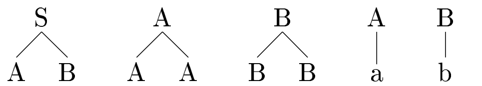


(b) Consider the following TSGs:

G_1:



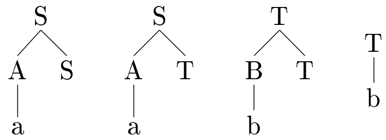
G_2:



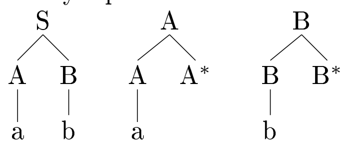
We assume that these TSGs have both a start symbol, namely S. I.e, only trees with root symbol S are in the tree language.

i. What are the string languages generated by the two TSGs?
Solution:
Both generate the same language, namely $\{a^n b^k \mid n, k \geq 1\}$.

ii. Decide for each of the TSGs whether it can be strongly lexicalized, i.e., whether a lexicalized TSG exists that generates the same set of trees. If so, give such a TSG. If not, explain why not and give a weakly equivalent LTAG.
Solution:
G_1 can be strongly lexicalized:

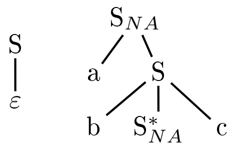


G_2 cannot even be weakly lexicalized, since lexicalizing the S tree would inhibit the adding of either more bs or as.
Weakly equivalent LTAG:

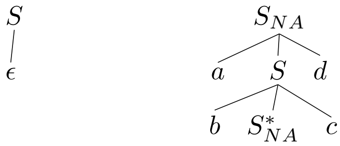


2. Let $L_1 := \{a^n b^n c^n \mid n \geq 0\}$, $L_2 := \{a^n b^n c^n d^n \mid n \geq 0\}$.

   (a) Give a TAG (with adjunction constraints) that generates L_1.
   Solution:
   TAG for L_1:



   (b) Give a TAG (with adjunction constraints) that generates L_2.
   Solution:
   TAG generating L_2:



3. (a) Generalize the CYK algorithm in order to make it work for a TAG which has elementary trees with more than two daugthers per node. Hint: You have to exchange the deduction rules *MoveUnary* and *MoveBinary* for a single new rule, which you might want to call *MoveUp*.
   Solution:
   Replace the deduction rules **move-unary** and **move-binary** with a single new rule **move-up**:

$$\textbf{Move-up:}\quad \frac{[\gamma,(p\cdot 1)_\top,i_0,f_{11},f_{12},i_1],\ldots,[\gamma,(p\cdot m)_\top,i_{m-1},f_{m1},f_{m2},i_m]}{[\gamma,p_\perp,i_0,f_{11}\oplus\cdots\oplus f_{m1},f_{12}\oplus\cdots\oplus f_{m2},i_m]}$$

   As a side condition, we require that the node address $p \cdot (m+1)$ does not exist in γ.

4. Give the tree sets of **two non-local MCTAGs** that derive the following sentences. Try to give a linguistically sound analysis, and say why it is sound, and why it is non-local (if that is the case). You do **not** have to use features! Also provide the derivation tree and the derived tree.

   (a) Extraction out of a complex NP:

   (1) [Which painting]$_i$ did you see a picture of $\_\__i$?

   (b) Extraposition of a relative clause:

   (2) Somebody$_i$ lives nearby [who$_i$ has a CD burner].