COMS W4705: Natural Language Processing
Written Homework 2

Minghao Li (ml4025)

March 22, 2018

Problem 1

Initial state: $([\text{root}]_\alpha,[\text{he,sent,her,a,funny,meme,today}]_\beta,\{\}_A)$

shift: $([\text{root,he}]_\alpha,[\text{sent,her,a,funny,meme,today}]_\beta,\{\}_A)$

left-arc: $([\text{root}]_\alpha,[\text{sent,her,a,funny,meme,today}]_\beta,\{(\text{he,nsubj,sent})\}_A)$

shift: $([\text{root,sent}]_\alpha,[\text{her,a,funny,meme,today}]_\beta,\{(\text{he,nsubj,sent})\}_A)$

right-arc: $([\text{root}]_\alpha,[\text{sent,a,funny,meme,today}]_\beta,\{(\text{he,nsubj,sent}),(\text{sent,jobj,her})\}_A)$

shift: $([\text{root,sent}]_\alpha,[\text{a,funny,meme,today}]_\beta,\{(\text{he,nsubj,sent}),(\text{sent,jobj,her})\}_A)$

shift: $([\text{root,sent,a}]_\alpha,[\text{funny,meme,today}]_\beta,\{(\text{he,nsubj,sent}),(\text{sent,jobj,her})\}_A)$

shift: $([\text{root,sent,a,funny}]_\alpha,[\text{meme,today}]_\beta,\{(\text{he,nsubj,sent}),(\text{sent,jobj,her})\}_A)$

left-arc: $([\text{root,sent,a}]_\alpha,[\text{meme,today}]_\beta,\{(\text{he,nsubj,sent}),(\text{sent,jobj,her}),(\text{funny,amod,meme})\}_A)$

left-arc: $([\text{root,sent}]_\alpha,[\text{meme,today}]_\beta,\{(\text{he,nsubj,sent}),(\text{sent,jobj,her}),(\text{funny,amod,meme}),(\text{a,det,meme})\}_A)$

right-arc: $([\text{root}]_\alpha,[\text{sent,today}]_\beta,$
$\{(\text{he,nsubj,sent}),(\text{sent,jobj,her}),(\text{funny,amod,meme}),(\text{a,det,meme}),(\text{sent,dobj,meme})\}_A)$

shift: $([\text{root,sent}]_\alpha,[\text{today}]_\beta,$
$\{(\text{he,nsubj,sent}),(\text{sent,jobj,her}),(\text{funny,amod,meme}),(\text{a,det,meme}),(\text{sent,dobj,meme})\}_A)$

right-arc: $([\text{root}]_\alpha,[\text{sent}]_\beta,$
$\{(\text{he,nsubj,sent}),(\text{sent,jobj,her}),(\text{funny,amod,meme}),(\text{a,det,meme}),(\text{sent,dobj,meme}),(\text{sent,advmod,today})\}_A)$

right-arc: $([]_\alpha,[\text{root}]_\beta,$
$\{(\text{he,nsubj,sent}),(\text{sent,jobj,her}),(\text{funny,amod,meme}),(\text{a,det,meme}),$
$(\text{sent,dobj,meme}),(\text{sent,advmod,today}),(\text{root,pred,sent})\}_A)$

terminate: $([\text{root}]_\alpha,[]_\beta,$
$\{(\text{he,nsubj,sent}),(\text{sent,jobj,her}),(\text{funny,amod,meme}),(\text{a,det,meme}),$
$(\text{sent,dobj,meme}),(\text{sent,advmod,today}),(\text{root,pred,sent})\}_A)$

Problem 2

(a) (1) (Entry 0)

$$S \to \bullet\ NP\ VP\ [0,0]$$
$$NP \to \bullet\ Adj\ NP\ [0,0]$$
$$NP \to \bullet\ PRP\ [0,0]$$
$$NP \to \bullet\ N\ [0,0]$$
$$Adj \to \bullet\ \text{baking}\ [0,0]$$
$$PRP \to \bullet\ \text{they}\ [0,0]$$
$$N \to \bullet\ \text{potatoes}\ [0,0]$$

(2) (Entry 1)

$$PRP \to \text{they}\ \bullet\ [0,1]$$
$$NP \to PRP\ \bullet\ [0,1]$$
$$S \to NP\ \bullet\ VP\ [0,1]$$
$$VP \to \bullet\ V\ NP\ [1,1]$$
$$VP \to \bullet\ Aux\ V\ NP\ [1,1]$$
$$V \to \bullet\ \text{baking}\ [1,1]$$
$$V \to \bullet\ \text{are}\ [1,1]$$
$$Aux \to \bullet\ \text{are}\ [1,1]$$

(3) (Entry 2)

$$V \to \text{are}\ \bullet\ [1,2]$$
$$Aux \to \text{are}\ \bullet\ [1,2]$$
$$VP \to V\ \bullet\ NP\ [1,2]$$
$$VP \to Aux\ \bullet\ V\ NP\ [1,2]$$
$$NP \to \bullet\ Adj\ NP\ [2,2]$$
$$NP \to \bullet\ PRP\ [2,2]$$
$$NP \to \bullet\ N\ [2,2]$$
$$V \to \bullet\ \text{baking}\ [2,2]$$
$$V \to \bullet\ \text{are}\ [2,2]$$
$$Adj \to \text{baking}\ [2,2]$$
$$PRP \to \bullet\ \text{they}\ [2,2]$$
$$N \to \bullet\ \text{potatoes}\ [2,2]$$

(4) (Entry 3)

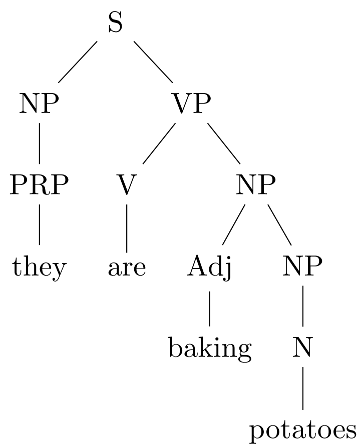
$$V \to \text{baking}\ \bullet\ [2,3]$$
$$Adj \to \text{baking}\ \bullet\ [2,3]$$
$$NP \to Adj\ \bullet\ NP\ [2,3]$$
$$VP \to AUX\ V\ \bullet\ NP\ [1,3]$$
$$NP \to \bullet\ Adj\ NP\ [3,3]$$
$$NP \to \bullet\ PRP\ [3,3]$$
$$NP \to \bullet\ N\ [3,3]$$
$$Adj \to \bullet\ \text{baking}\ [3,3]$$

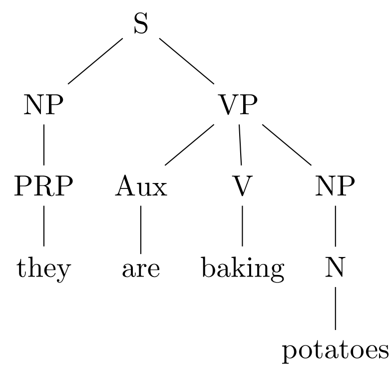
$$\text{PRP} \rightarrow \bullet \text{ they } [3,3]$$
$$\text{N} \rightarrow \bullet \text{ potatoes } [3,3]$$

(5) (Entry 4)

$$\text{N} \rightarrow \text{potatoes} \bullet [3,4]$$
$$\text{NP} \rightarrow \text{N} \bullet [3,4]$$
$$\text{NP} \rightarrow \text{Adj NP} \bullet [2,4]$$
$$\text{VP} \rightarrow \text{Aux V NP} \bullet [1,4]$$
$$\text{VP} \rightarrow \text{V NP} \bullet [1,4]$$
$$\text{S} \rightarrow \text{NP VP} \bullet [0,4]$$
$$\text{S} \rightarrow \text{NP VP} \bullet [0,4]$$

(b) parser trees





Probability for the left tree:

$$P = 1.0 * 0.1 * 1.0 * 0.8 * 0.5 * 0.3 * 1.0 * 0.6 * 1.0 = 0.0072$$

Probability for the right tree:

$$P = 1.0 * 0.1 * 1.0 * 0.2 * 1.0 * 0.5 * 0.6 * 1.0 = 0.006$$

Problem 3

(a) For the left tree, we have a sequence of tags as:

[PRP,V,Adj,N]

corresponding to the word sequence:

[they,are,baking,potatoes]

The original formula for computing HMM should be like:

$$P = P(PRP|START)P(they|PRP)P(V|PRP)P(are|V)P(Adj|V)P(baking|Adj)P(N|Adj)P(potatoes|N)$$

From the structure of the left tree, we can know that the probabilities of required sequence of tags should be tree parser without terminal symbols, so we need to multiply all the probabilities up from top to last possible nonterminals in the branches. That is:

$$P(tags) = P(PRP|START)P(V|PRP)P(Adj|V)P(N|Adj) = 1.0*0.1*0.8*0.3*0.6 = 0.0144$$

On the other hand, the joint probabilities from nonterminals to terminals can be directly calculated:

$$P(words) = P(they|PRP)P(are|V)P(baking|Adj)P(potatoes|N) = 1.0*0.5*1.0*1.0 = 0.5$$

As a result, the whole joint probability for P(tags, words) is

$$P(tags, words) = P(tags)P(words) = 0.0144*0.5 = 0.0072$$

Similarly, for the another tree I draw on the right, using the same logic:

$$P(tags) = 1.0*0.1*0.2*0.6 = 0.012$$

$$P(words) = 1.0*1.0*0.5*1.0 = 0.5$$

$$P(tags, words) = P(tags)P(words) = 0.012*0.5 = 0.006$$

(b) Inspired by the above calculation, I design an HMM which can get the same joint probability as PCFG does.

(1) For the word sequence: [they,are,baking,potatoes], obtain all the possible tag sequences that point to these terminals.
[PRP,V,Adj,N]
[PRP,Aux,Adj,N]
[PRP,V,V,N]
[PRP,Aux,V,N]

(2) Next we are going to calculate each joint sequence probability based on the parsing trees.

$$P(PRP, V, Adj, N, they, are, baking, potatoes) =$$

some probability based on PCFG from the left tree I draw above.

$$P(PRP, Aux, Adj, N, they, are, baking, potatoes) = 0$$

since we can't get a parse tree under this tag sequence.

$$P(PRP, V, V, N, they, are, baking, potatoes) = 0$$

since we can't get a parse tree under this tag sequence.

$$P(PRP, Aux, V, N, they, are, baking, potatoes) =$$

some probability based on PCFG from the right tree I draw above.

Now we can get the joint probability from the above 4 calculations.

(c) It is indeed possible to translate some PCFG to HMM that produce an identical joint probability as PCFG does. The general process is kind of like that in part(b), but a little more than it: we need to consider one more situation under which different trees may share the same tag sequence. So when calculating the HMM using PCFG(step2 above), we need to sum the probabilities of these trees up in order to get their HMM.

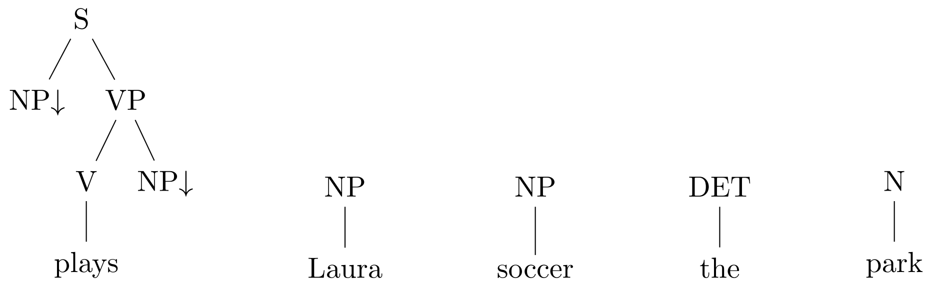
Through thinking, it really makes sense, because HMM model just considers tags' sequence, while PCFG considers more like how to generate the tags through an overall begin S. Therefore one tag sequence may contain different tree structures, which are multiple PCFGs.

However, there still exist some counter cases under which PCFG doesn't have its equal HMM. For example, if PCFG contains a rule like: $A \to \alpha B$, where α is a terminal symbol. But when we are trying to give α a tag, there's no one directly pointed to it. It works well in part(b) because all the given words of [they,are,baking,potatoes] can be directly derived from one tag. In other words, they doesn't share the tag with other non terminals. Therefore, if $A \to \alpha B$ appears, we can't get the tag for α, thus no HMM chain.

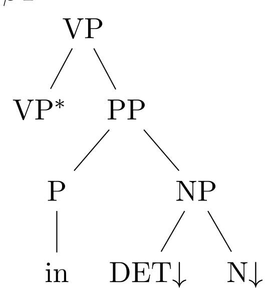
In conclusion, it's impossible to generate an equivalent HMM from any given PCFG.

Problem 4

(a) I design an LTAG grammar as below:
Initial Trees:
$\alpha 1$ $\alpha 2$ $\alpha 3$ $\alpha 4$ $\alpha 5$



Auxiliary Tree:
$\beta 1$

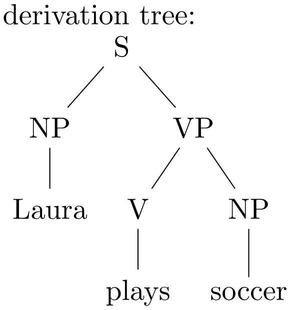


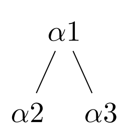
The trees I designed can satisfy the requirement:

(1) Each tree has exactly one anchor.
(2) There is at least one auxiliary tree.

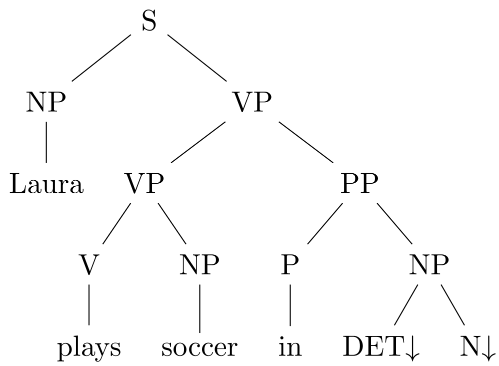
(b) I will explain how we can get to the final state step by step.

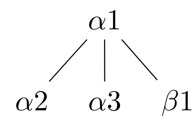
Firstly, substitute two NP in the $\alpha 1$ with $\alpha 2$ and $\alpha 3$, we can get the following derived and derivation tree:





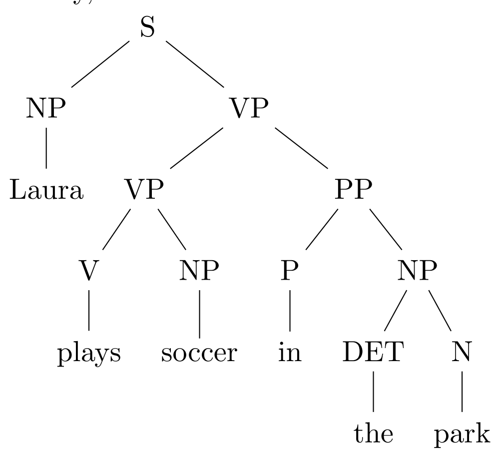
Then we can adjunct tree $\beta 1$ into where NP is in $\alpha 1$, getting the derived and derivation tree:

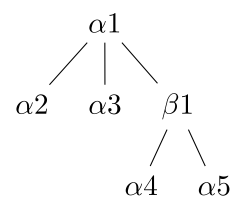




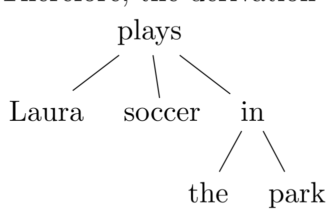
Clarify: the line from $\alpha1$ to $\beta1$ should be imaginary line to represent adjunction relationship.

Finally, I will substitute DET and N with $\alpha4$ and $\alpha5$, getting the following results:





Therefore, the derivation tree using all terminals should be like:



Clarify: the line from *plays* to *in* should be an imaginary one since it stands for adjunction relationship.