1 Probability Model

$$\begin{aligned} P(w_{1,n}) &= \Sigma_{t_{1,n}} P(w_{1,n}, t_{1,n}) \\ &= \Sigma_{t_{1,n}} P(t_{1,n}) P(w_{1,n} \mid t_{1,n}) \\ &= \Sigma_{t_{1,n}} P(t_{1,n}) \end{aligned}$$

A PCFG:

s → np vp	: .8
s → vp	: .2
np → noun	: .4
np → noun pp	: .4
np → noun np	: .4
vp → verb	: .3
vp → verb np	: .3
vp → verb pp	: .2
vp → verb np pp	: .2

The lexical part:

prep →like	: 1.0
verb →swat	: .2
verb →flies	: .4
verb →like	: .4
noun →swat	: .05
noun →flies	: .45
noun →ants	: .5

(1) Swat flies like ants.

1. $[_{np}$ swat flies $][_{vp}$ like ants $]$

2. [$_{vp}$ swat [$_{np}$ flies][$_{pp}$ like ants]]

3. Rules used in parse 1:

s → np vp	:	.8	:	swat flies like ants
np → noun np	:	.2	:	swat flies
np → noun	:	.4	:	flies, ants
vp → verb np	:	.3	:	like ants

The lexical part:

verb →like	:	.4
noun →swat	:	.05
noun →flies	:	.45
noun →ants	:	.5

4. Probability:

$$.8 \times .2 \times .4 \times .3 \times .4 \times .4 \times .05 \times .45 \times .5 = 3.5 \times 10^{-5}$$

5. Prob(Parse 1) = Prob($NP_{0,2}$, $VP_{2,4}$, $Noun_{0,1}$, $swat_{0,1}$, $NP_{1,2}$, $Noun_{1,2}$, $flies_{1,2}$, $Verb_{2,3}$, $likes_{2,3}$, $NP_{3,4}$, $Noun_{3,4}$, $ants_{3,4}$ | $S_{0,4}$)

6. We can apply chain rule. The Prob(parse 1) must equal:

   (a) Prob($NP_{0,2}$,$VP_{2,4}$ | $S_{0,4}$) X
   (b) Prob($Noun_{0,1}$,$NP_{1,2}$ | $S_{0,4}$, $NP_{0,2}$, $VP_{2,4}$) X
   (c) Prob($swat_{0,1}$ | $S_{0,4}$, $NP_{0,2}$,$VP_{2,4}$, $Noun_{0,1}$, $NP_{1,2}$) X
   (d) Prob($Noun_{1,2}$| $S_{0,4}$, $NP_{0,2}$, $VP_{2,4}$, $NP_{1,2}$, $Noun_{0,1}$, $NP_{1,2}$, $swat_{0,1}$) X
   (e) Prob($flies_{1,2}$| $S_{0,4}$, $NP_{0,2}$, $VP_{2,4}$, $NP_{1,2}$, $Noun_{0,1}$, $NP_{1,2}$, $swat_{0,1}$, $Noun_{1,2}$) X

(f) Prob($\text{Verb}_{2,3}$, $\text{NP}_{3,4}$| $\text{S}_{0,4}$, $\text{NP}_{0,2}$, $\text{VP}_{2,4}$, $\text{NP}_{1,2}$, $\text{Noun}_{0,1}$, $\text{NP}_{1,2}$, $\text{swat}_{0,1}$, $\text{Noun}_{1,2}$, $\text{flies}_{1,2}$) X

(g) Prob($\text{like}_{2,3}$| $\text{S}_{0,4}$, $\text{NP}_{0,2}$, $\text{VP}_{2,4}$, $\text{NP}_{1,2}$, $\text{Noun}_{0,1}$, $\text{NP}_{1,2}$, $\text{swat}_{0,1}$, $\text{Noun}_{1,2}$, $\text{flies}_{1,2}$, $\text{Verb}_{2,3}$, $\text{NP}_{0,2}$) X

(h) Prob($\text{Noun}_{3,4}$| $\text{S}_{0,4}$, $\text{NP}_{0,2}$, $\text{VP}_{2,4}$, $\text{NP}_{1,2}$, $\text{Noun}_{0,1}$, $\text{NP}_{1,2}$, $\text{swat}_{0,1}$, $\text{Noun}_{1,2}$, $\text{flies}_{1,2}$, $\text{Verb}_{2,3}$, $\text{NP}_{0,2}$ $\text{like}_{2,3}$) X

(i) Prob($\text{ants}_{3,4}$| $\text{S}_{0,4}$, $\text{NP}_{0,2}$, $\text{VP}_{2,4}$, $\text{NP}_{1,2}$, $\text{Noun}_{0,1}$, $\text{NP}_{1,2}$, $\text{swat}_{0,1}$, $\text{Noun}_{1,2}$, $\text{flies}_{1,2}$, $\text{Verb}_{2,3}$, $\text{NP}_{3,4}$, $\text{like}_{2,3}$. $\text{Noun}_{3,4}$)

7. We assume the probability that a constituent exists is independent of anything above it in the tree or the left or right of it. It is context free. For example:

   (a) Prob($\text{Noun}_{1,2}$| $\text{S}_{0,4}$, $\text{NP}_{0,2}$, $\text{VP}_{2,4}$, $\text{NP}_{1,2}$, $\text{Noun}_{0,1}$, $\text{NP}_{1,2}$, $\text{swat}_{0,1}$) = Prob($\text{Noun}_{1,2}$| $\text{NP}_{1,2}$)

   (b) Prob($\text{Verb}_{2,3}$, $\text{NP}_{3,4}$| $\text{S}_{0,4}$, $\text{NP}_{0,2}$, $\text{VP}_{2,4}$, $\text{NP}_{1,2}$, $\text{Noun}_{0,1}$, $\text{NP}_{1,2}$, $\text{swat}_{0,1}$, $\text{Noun}_{1,2}$, $\text{flies}_{1,2}$) = Prob($\text{Verb}_{2,3}$, $\text{NP}_{3,4}$| $\text{VP}_{2,4}$)

8. WE rewrite our parse probability as:

$$Prob(VPNP \mid S) \times Prob(NounNP \mid NP) \times Prob(swat \mid Noun) \times Prob(Noun \mid NP) \times Prob(flies \mid Noun) \times Prob(likeNP \mid Verb) \times Prob(Noun \mid NP) \times Prob(ants \mid Noun)$$

9. We define Prob($\text{Verb}_{2,3}$, $\text{NP}_{3,4}$| $\text{VP}_{2,4}$) as

$$Prob(\text{VP}_{2,4} \rightarrow \text{Verb}_{2,3}, \text{NP}_{3,4})$$

   Think of this of the probability of THIS VP rule applying, given we'vce got a VP from 2 to 4.