

NLP Homework2

Shuchen Liu (shl174@pitt.edu)

2.1

End	0	0	0	0
H	0	.32	.0448	0.0125
C	0	.02	.048	0.00288
Start	1.0	0	0	0
T=	0	1	2	3

Probabilities with T=3:

$$V2(2)*P(H|H)*P(3|H) = 0.0448*0.7*0.4 = 0.0125$$

$$V2(1)*P(H|C)*P(3|H) = 0.048*0.4*0.4 = 0.00768$$

$$\text{So } V3(2) = \max(0.0125, 0.0054) = 0.0125$$

$$V2(2)*P(C|H)*P(3|C) = 0.0448*0.3*0.1 = 0.00134$$

$$V2(1)*P(C|C)*P(3|C) = 0.048*0.6*0.1 = 0.00288$$

$$\text{So } V3(1) = \max(0.00288, 0.00134) = 0.00288$$

Do calculation in log:

$$\ln(V2(2)) + \ln(P(H|H)) + \ln(P(3|H)) = -3.10 - 0.35 - 0.91 = -4.36$$

$$\ln(V2(1)) + \ln(P(H|C)) + \ln(P(3|H)) = -3.03 - 0.91 - 0.91 = -4.85$$

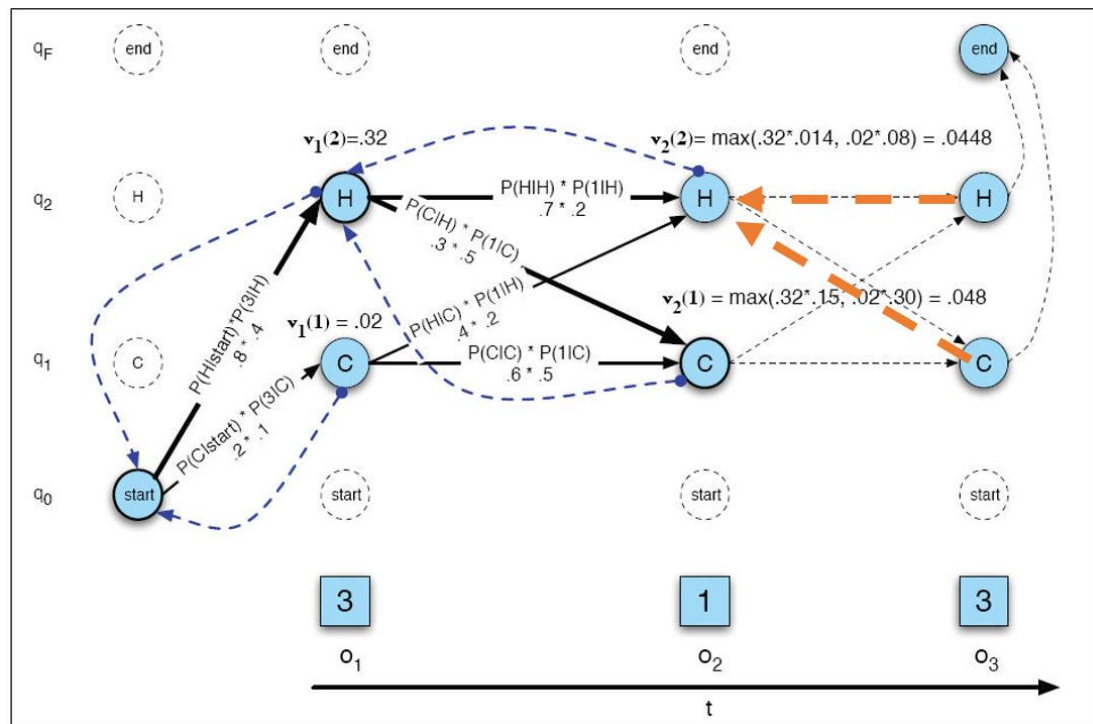
$$\text{So } \ln(V3(2)) = \max(-4.36, -4.85) = -4.36$$

$$\ln(V2(2)) + \ln(P(C|H)) + \ln(P(3|C)) = -3.10 - 1.20 - 2.30 = -6.60$$

$$\ln(V2(1)) + \ln(P(C|C)) + \ln(P(3|C)) = -3.03 - 0.51 - 2.30 = -5.84$$

$$\text{So } \ln(V3(1)) = \max(-6.60, -5.84) = -5.84$$

Best path:



Add a backtrace link to the backtrace figure going from STATE Hot at time $t = 3$ to STATE Hot at time $t = 2$.

Add a backtrace link to the backtrace figure going from STATE Cold at time $t = 3$ to STATE Hot at time $t = 2$.

2.2

The correctness:

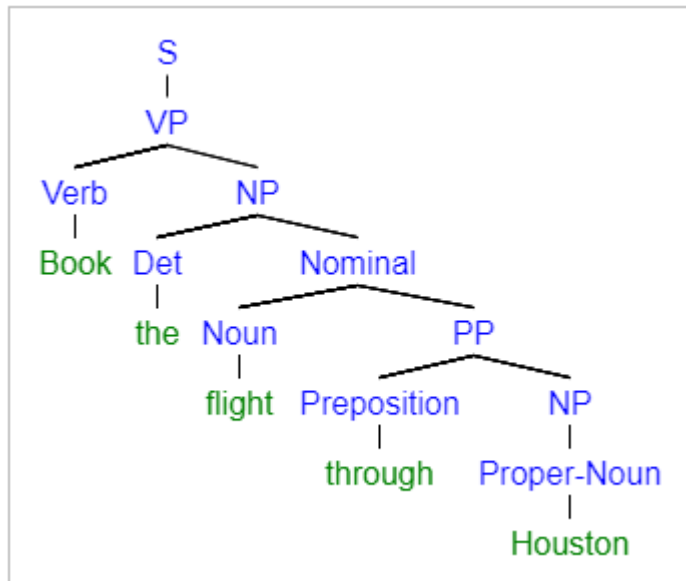
According to my CKY parser, I got several parses of those tow sentences.

First sentence:

"Book the flight through Houston"

The gold standard of this sentence is

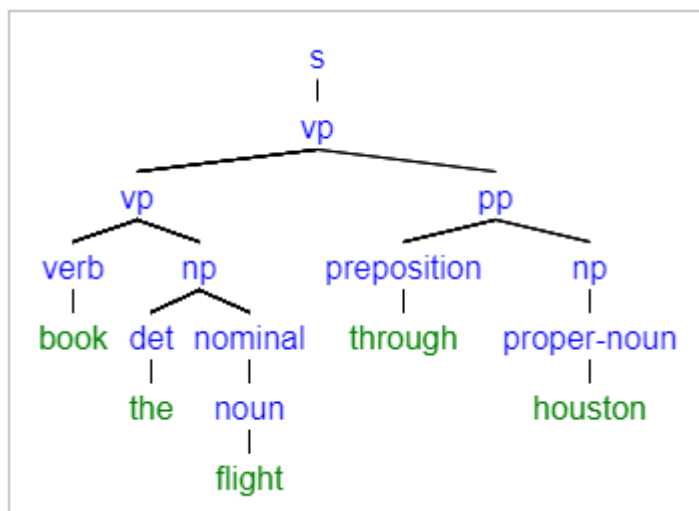
"[S [VP [Verb Book] [NP[Det the][Nominal [Noun flight][PP[Preposition through][NP[Proper-Noun Houston]]]]]]]". And the tree looks like:



My parser generated three possible parses. They are:

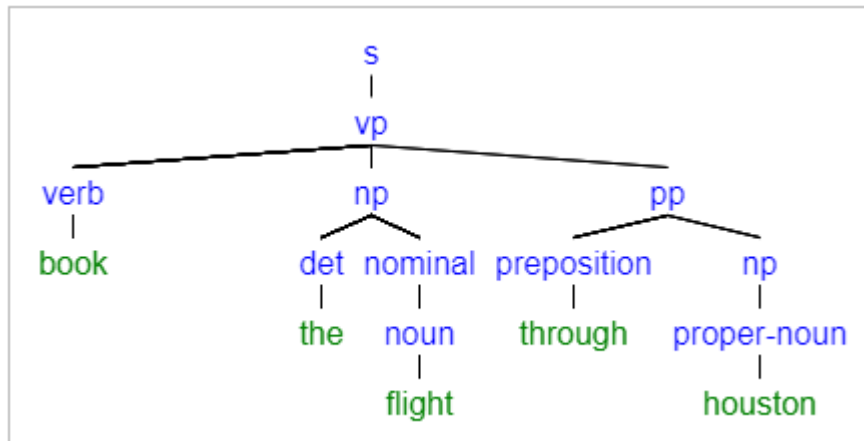
First:

One parse: [s[vp[vp[verb book][np[det the][nominal[noun flight]]]][pp[preposition through][np[proper-noun houston]]]]]
 Precision: 0.75
 Recall: 0.8181818
 Probability: 1.09350026E-7



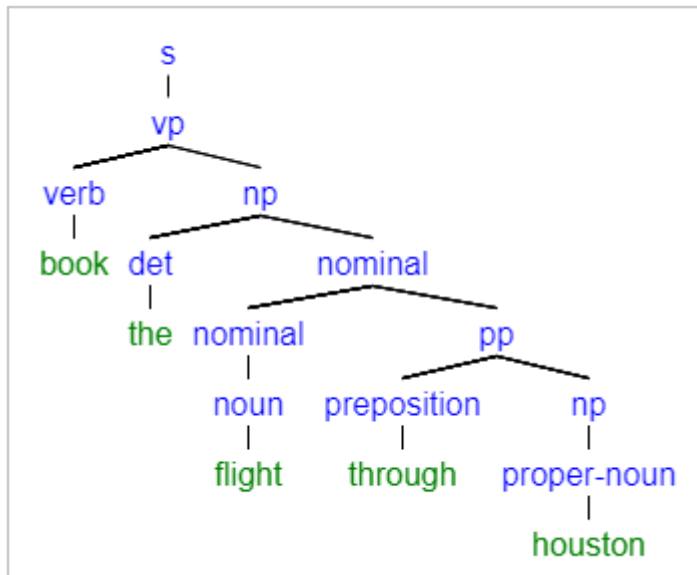
Second:

One parse: [s[vp[verb book][np[det the][nominal[noun flight]]]][pp[preposition through][np[proper-noun houston]]]]]
 Precision: 0.8181818
 Recall: 0.8181818
 Probability: 3.6450007E-7



Third:

One parse: [s[vp[verb book][np[det the][nominal[nominal[noun flight]][pp[preposition through][np[proper-noun houston]]]]]]]
 Precision: 0.9166667
 Recall: 1.0
 Probability: 3.645001E-8



The third is almost the same as the gold stander. And I think there is an error in the gold stander since "flight" can be a noun and noun can be nominal. However in the gold stander, the noun is eliminated.

Based on my CKY parser the probability of this sentence is 5.103001E-7.

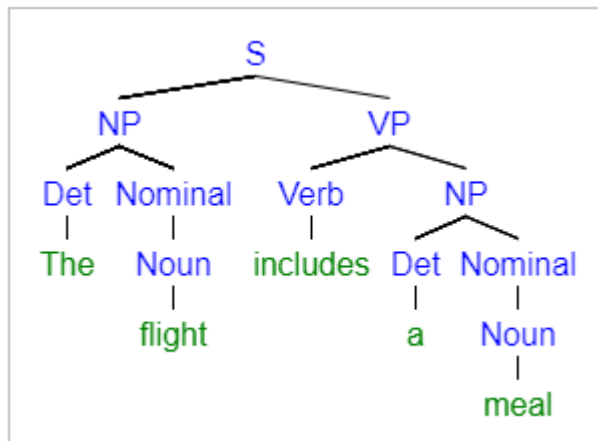
Second sentence:

"The flight includes a meal"

The gold stander is

"[S[NP[Det The][Nominal[Noun flight]]][VP[Verb includes][NP[Det a][Nominal[Noun meal]]]]]"

And the tree looks like:



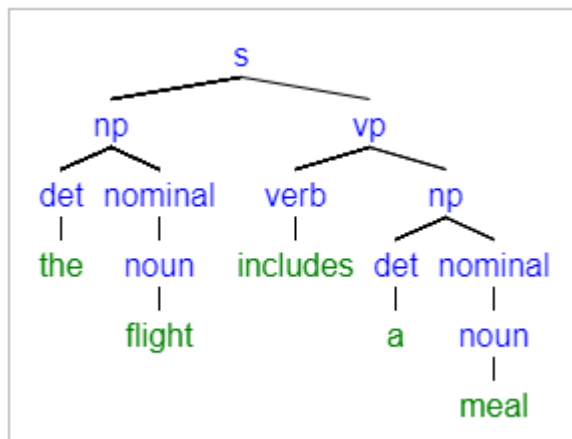
My parser only generated one parse and it is exactly the same as the gold stander.

One parse: [s[np[det the][nominal[noun flight]]][vp[verb includes][np[det a][nominal[noun meal]]]]]

Precision: 1.0

Recall: 1.0

Probability: 2.9160008E-6



Based on my CKY parser the probability of this sentence is 2.9160008E-6.