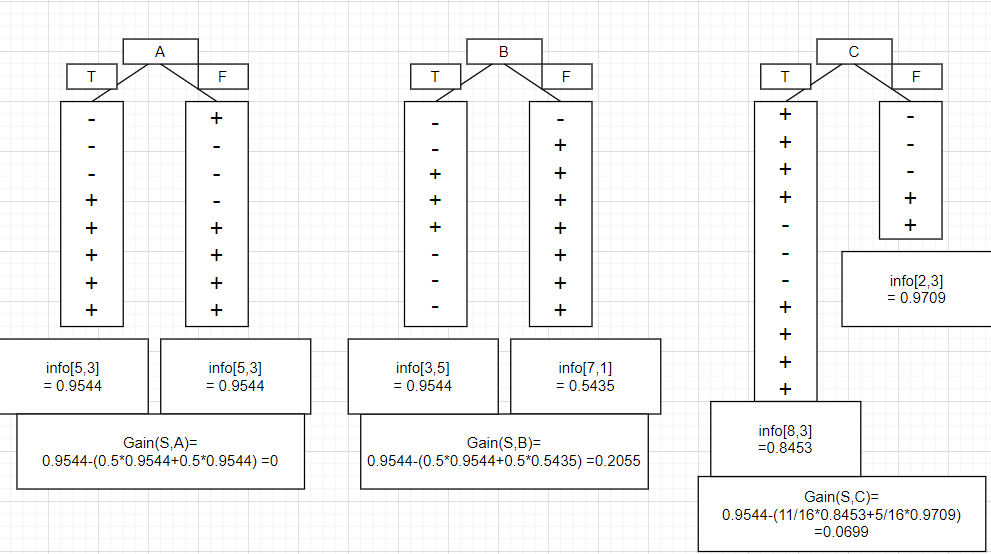
Homework5

JAESANGPARK

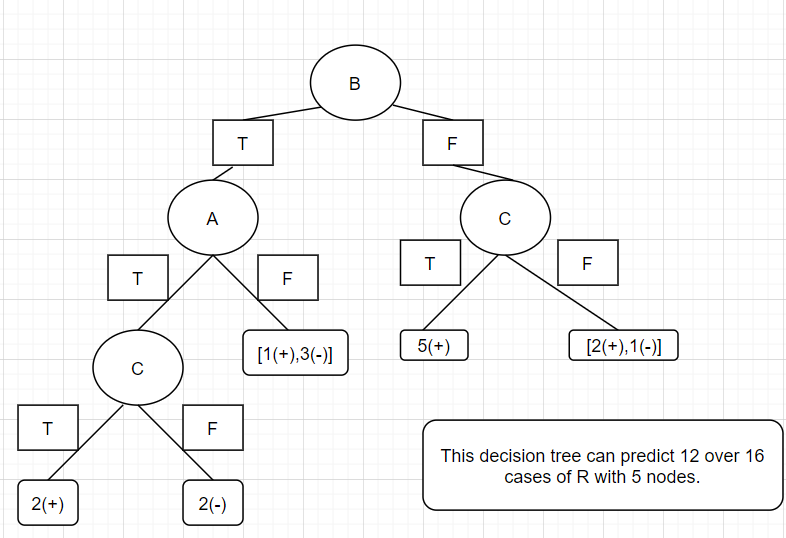
5.1

Entropy[10/16 , 6/16] = 0.9544



Since the information gain of B is the largest, so the decision tree would be built with B.

5.2



The tree with only one node in previous problem can’t predict any case of R.

Since there is no data about (A,B,C) = (T,T,T) , (F,T,F) and (F,F,F), we can’t make adequate decision tree.

I think I need more general training set to make the performance better.

5.3

1. unigram

If the probability of the occurrence of each word is independent, then the probability of the sentence is the product of the probability of each word.

That means,

=0.05 \* 0.000075 \* 0.0035 \* 0.00015 \* 0.00002\* 0.000015

= 0.000000000000000000000590625

1. bigram

If the probability of the occurrence of each word is dependent on previous word, then the probability of the sentence is :

0.05\*0.0004\*0.00000 \* 0.0006\*0.0000\*0.00000

= 0

1. pos-driven model

With this model, I have to consider the hidden feature of each word, which means is that noun? verb?.

And like the bigram, I need to consider the order of each word.

0.6\* 0.65\*0.0002\*0.15\*0.03\*0.15\*0.001\*0.15\*0.000025\*0.3\*0.00004

=0.00000000000000000000236925

5.4

a. unigram

My utterance is “the the the the the the.”

It is definitely not that good English.

And the probability is =0.000000015625 >>>0.000000000000000000000590625

b. bigram

My utterance is “the approved the approved the approved.”

Since the is the largest, I would use that pattern.

And the is the best.

So the probability of my utterance is

which is absolutely larger than 0.

1. pos-driven model

My utterance is “has has has has has has.”

Since the is the largest followed by ,

I choose this combination.

The probability of my utterance is .

which is 0.0000000000000553584375.

It is 23,365,384 times higher than the previous “The council has approved wind farms.”