Tutorial 4: POS tagging and HMM

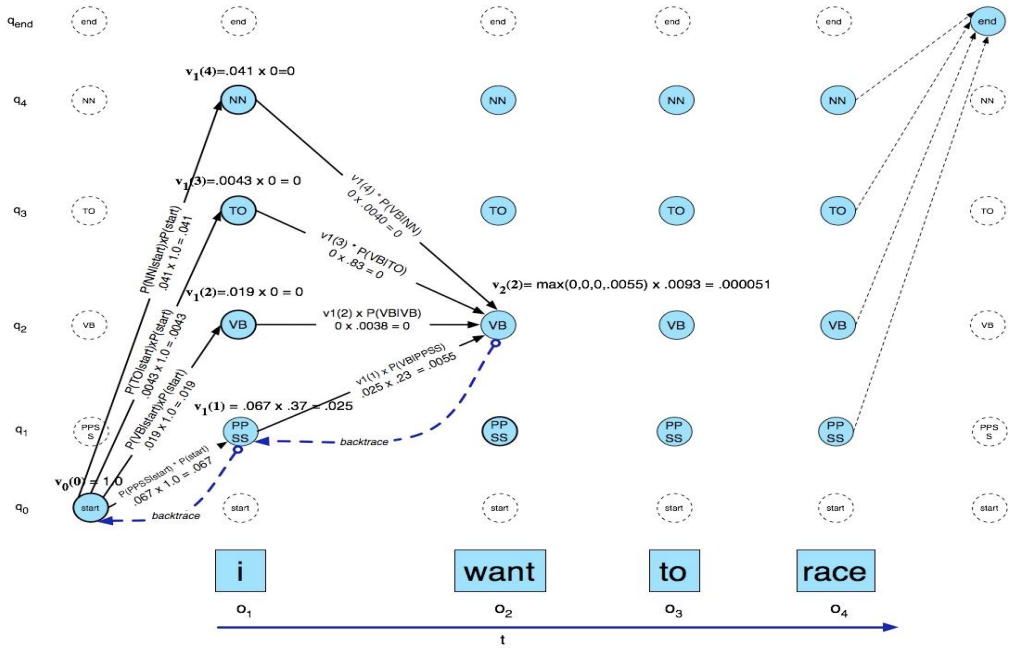
Q1. Find one tagging error in each of the following sentences that are tagged with the Penn Treebank tagset:

1. How/WRB do/MD I/PRP get/VB to/TO Singapore/NN
2. Do/VBP you/PRP have/VB any/DT vacancies/NN
3. This/DT room/NN is/VBZ too/JJ noisy/JJ
4. Can/VB you/PRP give/VB me/PRP another/DT room/NN

Q2. Finish the computation of the Viterbi algorithm in the Viterbi example used in the lecture for HMM. The transition probability and word likelihood probabilities are in the following tables.

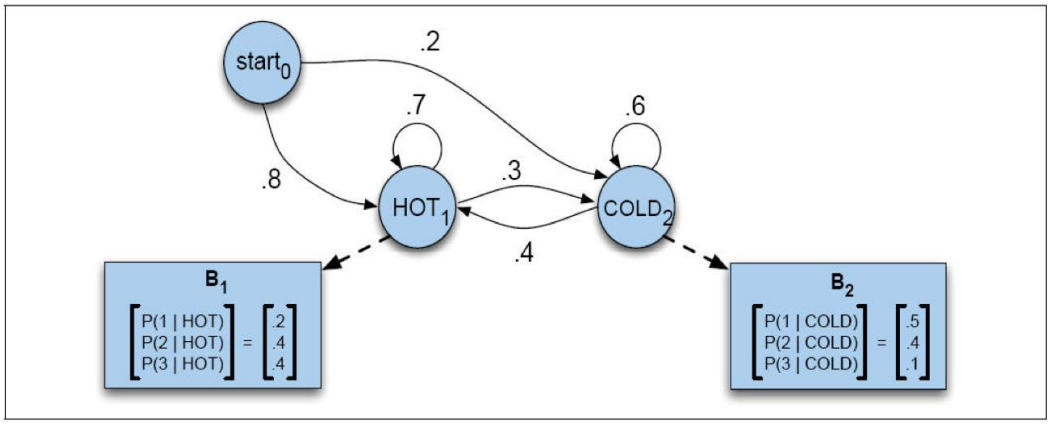
	VB	TO	NN	PPSS
<s>	.019	.0043	.041	.067
VB	.0038	.035	.047	.0070
TO	.83	0	.00047	0
NN	.0040	.016	.087	.0045
PPSS	.23	.00079	.0012	.00014

	I	want	to	race
VB	0	.0093	0	.00012
TO	0	0	.99	0
NN	0	.000054	0	.00057
PPSS	.37	0	0	0



Q3. Run the Viterbi algorithm with the HMM in the figure below to compute the most likely weather sequences for each of the two following observation sequences:

- Sequence: 312312312
- Sequence: 311233112



Q4. The Church tagger (1988) is different from the HMM tagger since it incorporates the probability of the tag given the word.

- HMM: p(word|tag) * p(tag|previous n tags)
- Church: p(tag|word) * p(tag|previous n tags)

Interestingly, this use of a kind of "reverse likelihood" has proven to be useful in the modern log-linear approach to machine translation (see textbook page 939). As a gedanken-experiment, construct a set of tag transition probabilities, and a set of lexical tag probabilities that demonstrate the two taggers (Church tagger and HMM) produce different tags for a word.

Hint: The Church and HMM taggers will perform differently when, given two tags, tag_1 and tag_2:

$p(tag_1|word) > p(tag_2|word)$ vs. $p(word|tag_1) < p(word|tag_2)$