Question 5

Flag question Marked out of 6.00 Not yet answered

a) Use the Viterbi algorithm to calculate the most probable part-of-speech sequence for the sentence given in the top row of Table 1. The observation probabilities of the part-of-speech tags computed from an annotated text corpus are given in the next rows of Table 1. The transition probabilities of between the tags are given in Table 2. Show your calculations. (3p)

Table 1. Observation probabilities. Each row indicates the probability of observing one part-of-speech tag (VB = verb, JJ = adjective, DT = determiner or NN = noun) for the given word (column)

	open	a	book
VB	0.6	0	0.8
JJ	0.4	0	0
DT	0	0.9	0
NN	0	0.1	0.2

Table 2. Transition probabilities. Each row indicates the probability of moving from one (preceding) tag (first) to each of the following tags (columns).

	VB	JJ	DT	NN
VB	0	0.2	0.8	0
JJ	0	0.1	0	0.9
DT	0	0.5	0	0.5
NN	0.6	0	0	0.4

b) What are the weaknesses of the method you used in a)? (1p)

c) Describe in detail one method to alleviate these weaknesses. (2p)

Maximum file size: 400 MB, maximum number of files: 1

Files

You can drag and drop files here to add them.

Previous page Finish attempt ...