Exercise 3.1

	▷	N	N	◁
transition probability:		1/2	1/2	1/3
word probability:		1/2	1/2	

Sequence Probability $= (1/2)^4 * 1/3 = 1/48$

	▷	N	V	◁
transition probability:		1/2	1/6	1/3
word probability:		1/2	1/4	

Sequence Probability $= (1/2)^2 * 1/3 * 1/4 * 1/6 = 1/288$

	▷	V	N	◁
transition probability:		1/2	1/2	1/3
word probability:		3/4	1/2	

Sequence Probability $= (1/2)^3 * 3/4 * 1/3 = 1/32$

	▷	V	V	◁
transition probability:		1/2	1/6	1/3
word probability:		3/4	1/4	

Sequence Probability $= 1/2 * 1/6 * 1/3 * 3/4 * 1/4 = 1/192$

a. The total probability is the sum over these probabilities, 35/576.

b. The most likely tag sequence is $< V, N >$, with probability 1/32.

Exercise 3.2

No. The probability of a word conditioned on a state is independent of the probability of that state. The fact that all possible state sequences have zero probability could just mean that the word sequence is egregiously ungrammatical. For instance, "the the the dog" with a linguistically informed HMM will have no possible state sequences because "the" can only come from state D (determiner), and D cannot follow D. However, the likelihood of state D generating "the" will still be non-zero.

Exercise 3.3

It's not a coincidence! If the Viterbi label for a given word is different from the individually most likely label, then that means that the likelihood of the Viterbi sequence must be less than the sum of the likelihoods of some subset of other sequences; if the Viterbi sequence has more than half of the total probability, then this is impossible. Therefore, if the Viterbi sequence has more than half of the total probability, then it must be the same as the sequence of individually most likely tags.

Exercise 3.4

It suggests that "Daffynition" typically occurs at the end of sentences in the WSJ (along with actual punctuation, closing parentheses, etc.). However, it is capitalized, indicating that it comes at the beginning of a sentence. So "Daffynition" likely occurs as a single-word sentence in the WSJ.