Question 1

The only tag sequence $y_1 \ldots y_{n+1}$ for which $p(y_1 \ldots y_{n+1}) > 0$ is `D N V STOP`. Thus the only sequences that satisfy the conditions are

the dog dog, D N V STOP
the barks dog, D N V STOP
the dog barks, D N V STOP
the barks barks, D N V STOP

Question 2a

Input: a sentence $x_1 \dots x_n$, parameters $q(s|u,v,w)$ and $e(x|s)$.
Definitions: Define $\mathcal{K}$ to be the set of possible tags. Define $\mathcal{K}_{-2} = \mathcal{K}_{-1} = \mathcal{K}_0 = \{*\}$, and $\mathcal{K}_k = \mathcal{K}$ for $k = 1 \dots n$.
Initialization: Set $\pi(0, *, *, *) = 1$.
Algorithm:

- For $k = 1 \dots n$,
  - For $u \in \mathcal{K}_{k-2}$, $v \in \mathcal{K}_{k-1}$, $w \in \mathcal{K}_k$,

$$\pi(k, u, v, w) =$$
$$\max_{s \in \mathcal{K}_{k-3}} \left(\pi(k-1, s, u, v) \times q(w|s, u, v) \times e(x_k|w)\right)$$

- **Return**
  $\max_{u \in \mathcal{K}_{n-2}, v \in \mathcal{K}_{n-1}, w \in \mathcal{K}_n} \left(\pi(n, u, v, w) \times q(\text{STOP}|u, v, w)\right)$

Question 2b

Input: Parameters $q(s|u, v)$ and $e(x|s)$.
Definitions: Define $\mathcal{K}$ to be the set of possible tags. Define $\mathcal{K}_{-1} = \mathcal{K}_0 = \{*\}$, and $\mathcal{K}_k = \mathcal{K}$ for $k = 1 \ldots n$. Define $\mathcal{V}$ to be the set of possible words.
Initialization: Set $\pi(0, *, *) = 1$.
Algorithm:

- For $k = 1 \ldots n$,
  - For $u \in \mathcal{K}_{k-1}$, $v \in \mathcal{K}_k$,

$$\pi(k, u, v) =$$
$$\max_{s \in \mathcal{K}_{k-2}, x \in \mathcal{V}} (\pi(k-1, s, u) \times q(v|s, u) \times e(x|v))$$

- **Return** $\max_{u \in \mathcal{K}_{n-1}, v \in \mathcal{K}_n} (\pi(n, u, v) \times q(\mathsf{STOP}|u, v))$

Question 3

Input: Parameters $q(s|u)$ and $e(x|s)$.
Definitions: Define $\mathcal{K}$ to be the set of possible tags. Define $\mathcal{K}_{-1} = \mathcal{K}_0 = \{*\}$, and $\mathcal{K}_k = \mathcal{K}$ for $k = 1 \ldots n$.
Initialization: Set $\pi(0, *, *) = 1$.
Algorithm:

- For $k = 1 \ldots n$,
  - For $u \in \mathcal{K}_{k-1}$, $v \in \mathcal{K}_k$,

$$\pi(k, u, v) = \max_{s \in \mathcal{K}_{k-2}} \left(\pi(k-1, s, u) \times q(v|s) \times e(x|v)\right)$$

- **Return** $\max_{u \in \mathcal{K}_{n-1}, v \in \mathcal{K}_n} \left(\pi(n, u, v) \times q(\text{STOP}|u)\right)$

Question 4

The only word which is seen with more than one tag is *can*: hence *can* is the only word such that $e(\text{word}|y)$ is greater than 0 for more than one tag y.

Now if we consider the two incorrect taggings,

the/DT can/VB is/VB in/IN the/DT shed/NN

the/DT dog/NN can/NN see/VB the/DT cat/NN

we can see that both have probability 0 under the maximum likelihood model. The first sentence has the tag bigram *DT VB* that is never seen in the training corpus, and hence has an associated parameter $q(\text{VB}|\text{DT}) = 0$, resulting in the entire tagged sentence having probability zero. The second sentence has a tag bigram *NN NN* that is also never seen, and hence has probability 0. It is easy to verify that the two correct tag sequences both have probability greater than 0.

Question 5

Consider a training set consisting of the single tagged sentence

the/DT the/NN the/DT the/DT the/DT the/DT

In this case we have maximum likelihood estimates

$$e(\text{the}|\text{DT}) = e(\text{the}|\text{NN}) = 1$$

$$q(\text{NN}|\text{DT}) = 1/5 \quad q(\text{DT}|\text{DT}) = 3/5 \quad q(\text{STOP}|\text{DT}) = 1/5$$

It can then be verified that the most likely tag sequence for *the the the the the the* is *DT DT DT DT DT DT STOP*.