

Lab Sheet 3 – HMM

Set: Thur 2nd Apr 2020

Due: **8pm, Friday 24th April 2020**

Marks: 5%

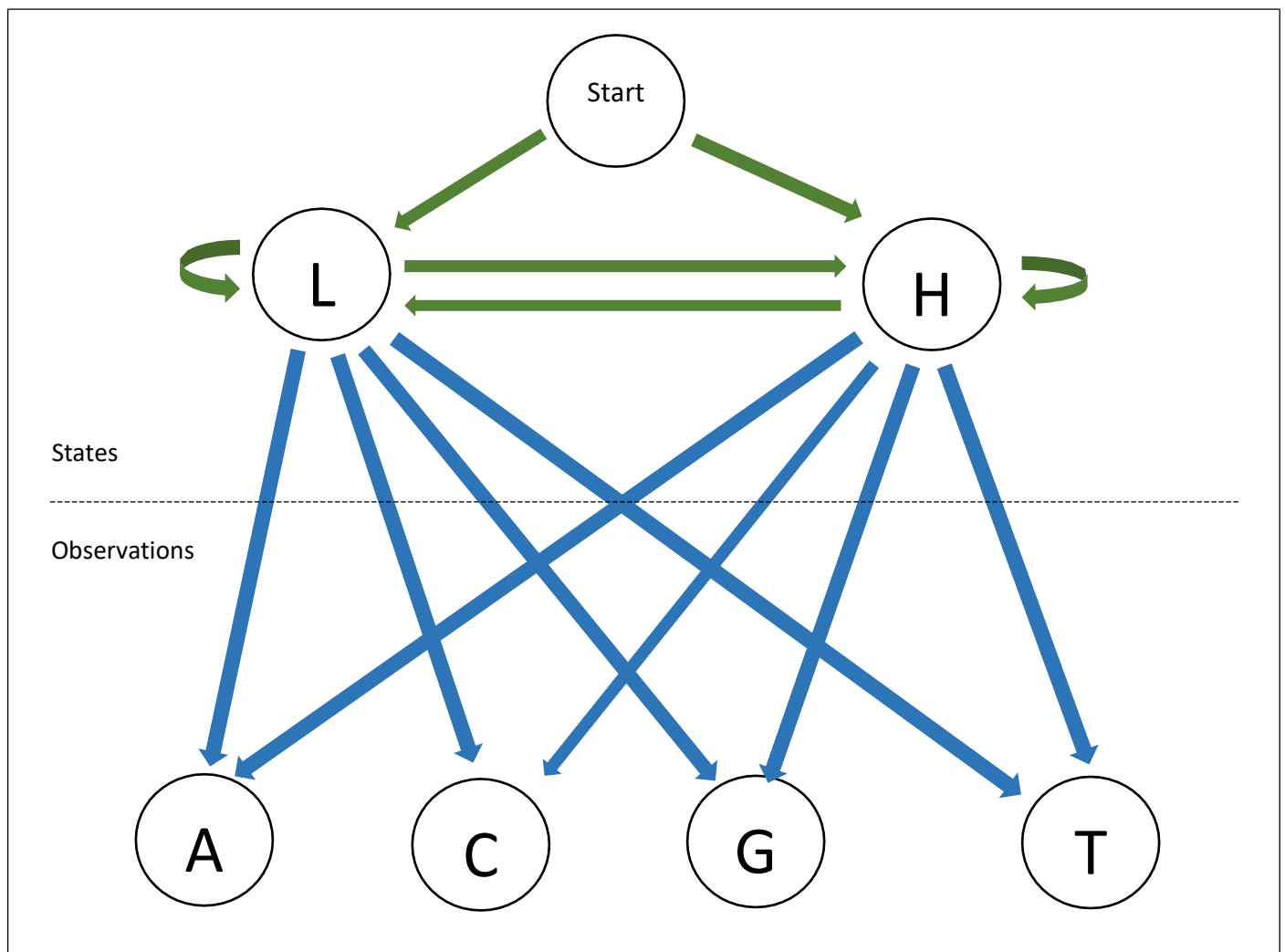
Submission: Moodle

Upload: Jupyter Notebook, including output.

In this lab you will implement the Viterbi algorithm with backtracking to determine the most likely set of states that would produce a given set of observations.

$$p_l(i, x) = e_x(i) \max_k (p_k(j, x - 1) \cdot p_{kl})$$

Here is the HMM you will use:



States: L, H

Observations: A, C, T, G

All probabilities given as $\log_2()$ values.

Start Probability:

	L	H
	-1	-1

Transitions:

		To L	To H
From	L	-0.737	-1.322
From	H	-1	-1

Observation Probabilities:

	A	C	G	T
L	-1.737	-2.322	-2.322	-1.737
H	-2.322	-1.737	-1.737	-2.322

Table 1

Task:

Given the sequence of observations **GGCACTGAA**

1. Implement the dynamic Viterbi algorithm to compute the missing entries in the table below. It is necessary to also record the max decisions you made to perform the second, backtracking step.
2. Implement backtracking to identify the most probable path and therefore the most likely sequence of states that would produce the observations.

	G	G	C	A	C	T	G	A	A
H	-2.73								-25.65
L	-3.32								-24.49

Table 2

Marking Guidance

You should upload a Jupyter Notebook to Moodle that contains both the code you have created and the results of running it. The deadline for this lab sheet is **Friday 24th April 8pm**.

Given that demonstration in the lab sessions is no longer expected, to increase your confidence in the software you are developing, the calculated values at the first and last step are shown in Table 2.

The marking scheme will be as follows:

If you provide the complete set of values expected in the table above and provide the correct backtracked set of states you will receive the full 5/5 marks allocated to this lab.

If you provide the complete set of values expected in the table above but do not provide the correct backtracked set of states you will receive 4/5 marks.

Any other output can receive between 0/5 to 3/5 based on my evaluation of how correct your code.

Ken.