Computer Science Tripos Part IA and IB

2019-2020
Exam Question
Cover sheet

Student BGN.....2023A.....

Paper.....3.....

Question number.....8.....

How did you answer this question?

Timed	☐	Open Book	☑
Untimed	☑	Closed Book	☐

Questions

List all the questions you have answered for this paper here. 1,4,6,8,9

Computer Science Tripos Honour Code

1. We take it as a principle that maintaining the integrity and fairness of examinations should be regarded as a collaboration between students and the Department.
2. The students undertake that they will not help others in examinations and will not receive any help from others (students or non-students).
3. Students will actively contribute to ensuring that all students adhere to the code.
4. Students will keep to the conditions of the assessment and will accurately report those conditions when asked.
5. The Department will not make any attempt at remote invigilation of online examinations.

I **undertake to respect the Computer Science Tripos honour code**

Tick the box to confirm ☑

8a) A first order HMM consists of the following components:

- A set of N emitting states $S = \{s_1, \ldots, s_N\}$, in this case `F`, `P`, `N`.
- A set of M possible observations $K = \{k_1, \ldots, k_M\}$, in this case, `G`, `A`, `B`.
- A sequence of T hidden states $X = X_1, \ldots, X_T$, each from S.
- A sequence of T observations $O = O_1, \ldots, O_T$ each from K:
- A state transition probability matrix A, with A_{st} indicating the probability of transition from state s to state t. These can be estimated by counting the occurrences of transition from state s to state t, and dividing by the total number of transitions from state s.
- An emission probability matrix B, with B_{os} indicating the probability of emitting observation o from state s. This can be estimated by counting the emissions of observation o from state s, and dividing by the total occurrences of state s.

b)

- The assumption that the next state depends only on the previous state, and not any others:

$$P(X_t \mid X_{t-1}, X_{t-2}, \ldots, X_1) \approx P(X_t \mid X_{t-1})$$

In the context of farming, this indicates that the cultivation method used is only dependent on the previous cultivation method used. i.e. the farmer has a very short memory.

- The assumption that the current observation depends only on the current state, and not any other states, or any observations:

$$P(O_t \mid X_T, X_{T-1}, \ldots, X_1, O_{T-1}, \ldots, O_{t+1}, O_{t-1}, \ldots, O_1) \approx P(O_t \mid X_t)$$

In the context of farming, this indicates that the crop performance for a given year is only dependent on the cultivation method for that year, and not on methods used in other years, or on other crop performances.

c)

(i) Fertilizer is most likely to yield good crop.

$P(G \mid F) = 0.5$

$P(G \mid P) = 0.25$

$P(G \mid N) = 0$

(ii) Since the HMM assumes that the crop performance only depends on the current cultivation method, the succession of methods which will yield the best results for two years is simply the best for one year, doubled. So the best in theory is two rounds of fertilizer. Ignoring the Markov assumptions, this is also evidenced to an extent by the fact that `FF` in the data above resulted in good, then average crops, which is not far off two good crops.

(iii) Ignoring Markov assumptions, the above data gives a probability of $2/3$ that a bad crop is followed by a good crop, and $1/3$ that it is followed by and average crop.

d)

(i) Unlabelled learning

(ii) Evaluation

(iii) Decoding