Symbolic Machine Learning - Model Exam Questions

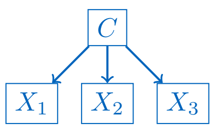
PUBLISHED SPECIMEN

Question 1. (10 points)

We have a bag of three biased coins a, b. and c with probabilities of coming up heads of $20\%, 60\%$, and 80%, respectively. One coin is drawn randomly from the bag (a is twice more likely to be drawn than the other two coins), and then the coin is flipped three times to generate the outcomes X_1, X_2, and X_3.

1. (4 points) Draw the Bayesian network corresponding to this setup and define the necessary CPTs.

   Answer:

   $\mu(C = a) = 1/2, \mu(C = b) = \mu(C = c) = 1/4$

   for $i = 1, 2, 3:$ $\begin{array}{l} \mu(X_i = h|C = a) = 1/5 \\ \mu(X_i = h|C = b) = 3/5 \\ \mu(X_i = h|C = c) = 4/5 \end{array}$.

2. (6 points) Calculate which coin was most likely to have been drawn from the bag if the observed flips come out heads twice and tails once.

   Answer:

   Probability of 2 heads and 1 tail (denote Pr) is the probability of the disjunction of three events (hht, hth, thh), which are mutually exclusive so we can sum their probabilites:

$$\mathrm{Pr} = \mu(C) \cdot \begin{pmatrix} \mu(X_1 = h|C)\mu(X_2 = h|C)\mu(X_3 = t|C) \\ +\mu(X_1 = h|C)\mu(X_2 = t|C)\mu(X_3 = t|C) \\ +\mu(X_1 = t|C)\mu(X_2 = h|C)\mu(X_3 = t|C) \end{pmatrix}$$

   all X_i's have the same CPT so

$$\mathrm{Pr} = 3\mu(C)\mu(X_i = h|C)^2\mu(X_i = t|C)$$

   and

$$\max \mathrm{Pr} = 3 \cdot \max_{C \in \{\, a,b,c \,\}} \left\{ \begin{array}{r} 1/2 \cdot 1/5^2 \cdot 4/5 \text{ (for } C = a) \\ 1/4 \cdot 3^2/5^2 \cdot 2/5 \text{ (for } C = b) \\ 1/4 \cdot 4^2/5^2 \cdot 2/5 \text{ (for } C = c) \end{array} \right\}$$

$$= 3 \cdot \max \left\{ \begin{array}{r} 2/5^3 \text{ (for } C = a) \\ 18/(4 \cdot 5^3) \text{ (for } C = b) \\ 6/(4 \cdot 5^3) \text{ (for } C = c) \end{array} \right\}$$

   so $\arg\max_{C \in \{\, a,b,c \,\}} \mathrm{Pr} = b$

PUBLISHED SPECIMEN

Question 2. (15 points)

Design an on-line learning agent that learns an s-DNF (i.e., a disjunction of conjunctive terms, where each term has at most s literals) on n propositional variables from only *negative observations* (i.e, truth-assignments for which the DNF is not satisfied).

1. (5 points) Describe the initial hypothesized DNF (what terms will be in it?) and describe how the DNF is updated with each negative example.

   Answer:

   The initial DNF will consist of all $\sum_{i=1}^{s} \binom{2n}{i}$ conjunctive terms of at most s literals made out of the $2n$ literals corresponding to the n variables. On each negative observation, all terms satisfied by the example will be deleted from the DNF.

2. (5 points) Provide a bound, as tight as you can, on the maximum number of errors the agents makes, without any assumption on the set of observations received. Is the bound polynomial or exponential in (i) s, (ii) n ?

   Answer:

   On each mistake, at least one term is deleted. So no more than $\sum_{i=1}^{s} \binom{2n}{i}$ mistakes. Because $\binom{2n}{s} \leq (2n)^s$, this is polynomial in n and exponential in s.

3. (5 points) Assume $n = s = 2$. Demonstrate the sequence of hypothesized DNF's, starting with the initial one and updating on two successive observations $o_1 = (1, 0)$, $o_2 = (0, 1)$.

   Answer:

$$\mathrm{p}_1 \vee \neg \mathrm{p}_1 \vee \mathrm{p}_2 \vee \neg \mathrm{p}_2 \vee (\mathrm{p}_1 \wedge \neg \mathrm{p}_1) \vee (\mathrm{p}_1 \wedge \mathrm{p}_2) \vee (\mathrm{p}_1 \wedge \neg \mathrm{p}_2) \vee (\neg \mathrm{p}_1 \wedge \mathrm{p}_2) \vee (\neg \mathrm{p}_1 \wedge \neg \mathrm{p}_2) \vee (\mathrm{p}_2 \wedge \neg \mathrm{p}_2)$$

$$\neg \mathrm{p}_1 \vee \mathrm{p}_2 \vee (\mathrm{p}_1 \wedge \neg \mathrm{p}_1) \vee (\mathrm{p}_1 \wedge \mathrm{p}_2) \vee (\neg \mathrm{p}_1 \wedge \mathrm{p}_2) \vee (\neg \mathrm{p}_1 \wedge \neg \mathrm{p}_2) \vee (\mathrm{p}_2 \wedge \neg \mathrm{p}_2)$$

$$(\mathrm{p}_1 \wedge \neg \mathrm{p}_1) \vee (\mathrm{p}_1 \wedge \mathrm{p}_2) \vee (\neg \mathrm{p}_1 \wedge \neg \mathrm{p}_2) \vee (\mathrm{p}_2 \wedge \neg \mathrm{p}_2)$$

PUBLISHED SPECIMEN

Question 3. (5 points)

Consider a version-space agent whose version space contains all non-contradictory conjunctions on 3 propositional variables.

1. (3 points) How large will the initial (i.e., before seeing any observation) version space be? Explain your reasoning.

   Answer:

   3^n (base: three cases, each variable may be absent, positive, or negative in the conjunction, exponent = number of vars), i.e. $3^3 = 27$.

2. (2 points) What is the maximal possible size of it after the first update, assuming the first class decision was wrong? Explain your reasoning.

   Answer:

   A version-space agent decides by majority of versions so at least 14 versions were wrong on (inconsistent with) the observation; these get deleted and at most 13 remain.

PUBLISHED SPECIMEN

Question 4. (10 points)

Determine the least general generalization of the following two assertions

1. *Superman is mortal or he is not a human.*
2. *Every human who smokes is mortal.*

by representing them as first-order logic clauses and computing their least general generalization with respect to the θ-subsumption order.

Answer:

$\texttt{mortal}(\text{Superman}) \vee \neg\texttt{human}(\text{Superman})$

$\texttt{human}(x) \wedge \texttt{smokes}(x) \rightarrow \texttt{mortal}(x) \equiv \neg\texttt{human}(x) \vee \neg\texttt{smokes}(x) \vee \texttt{mortal}(x)$

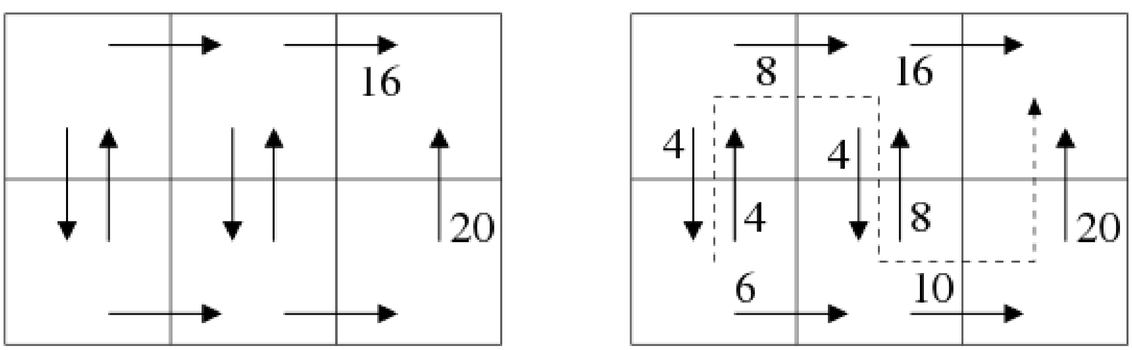
$LGG = \neg\texttt{human}(v_1) \vee \texttt{mortal}(v_1) \equiv \texttt{human}(v_1) \rightarrow \texttt{mortal}(v_1)$

(i.e., Every human is mortal.)

PUBLISHED SPECIMEN

Question 5. (10 points)

Consider the deterministic world below (part (a)). Allowable moves are shown by arrows, and the numbers indicate the reward for performing each action. If there is no number, the reward is zero. Given the Q values in (b), show the changes in the Q estimates when the agent takes the path shown by the dotted line (the agent starts in the lower left cell) when $\gamma = 0.5$ and $\alpha = 0.5$. Show all of your work including intermediate steps.



(a) (b)