

Final Examination

CS 540: Introduction to Artificial Intelligence

May 14, 2007

LAST NAME: _____

FIRST NAME: _____

<u>Problem</u>	<u>Score</u>	<u>Max Score</u>
1	_____	15
2	_____	20
3	_____	10
4	_____	15
5	_____	20
6	_____	20
Total	_____	100

1. [15] **Probabilistic Reasoning**

In her guest lecture, Prof. Burnside presented data related to whether a woman has breast cancer or not, and whether the results of a mammography test are positive (i.e., indicating breast cancer is likely) or not. Let Boolean random variables B stand for "has breast cancer" and M stand for "mammography test is positive." The prior probability of having breast cancer is 1%. The probability of testing positive when you *have* breast cancer is 90%. The probability of testing negative when you do *not have* breast cancer is 89.9%.

(a) [5] What is the prior probability of having a positive mammography test, i.e., what is $P(M)$?

(b) [5] If a patient has a positive mammography test, what is the probability that she has breast cancer? That is, compute $P(B \mid M)$.

(c) [5] If a patient gets a negative mammography test, what is the probability that she has breast cancer? That is, compute $P(B \mid \neg M)$.

2. [20] **Naive Bayes**

Consider the problem of detecting if an email message contains a virus. Use 4 Boolean random variables to model this problem: V indicates if the message contains a virus (=1) or not (=0), A indicates if the message has an attachment (=1) or not, L indicates if the message is long (=1) or not, and K indicates if the sender is known (=1) to the receiver or not.

(a) [3] If we use a Naive Bayes Classifier to solve this problem, we make the assumption $P(A, L, K | V) = P(A|V)P(L|V)P(K|V)$. What does this assumption mean in terms of independence and/or conditional independence of any of the variables A , L , K , and V ?

(b) [3] Draw the Bayesian Network (just the nodes and arcs) that corresponds to this problem with the Naive Bayes assumption in (a).

(c) [10] Given an email message with $A=1$, $L=0$, and $K=0$, compute the probability that it contains a virus. That is, compute $P(V = 1 | A = 1, L = 0, K = 0)$. Use the following probabilities as needed.

V	$P(A=1 V)$	$P(L=1 V)$	$P(K=1 V)$	$P(V)$
0	0.1	0.9	0.9	0.9
1	0.99	0.2	0.05	0.1

(d) [4] What constraints are there, in general, on the values filled in the 8 entries in the table above? That is, can we fill in the table with 8 arbitrary real numbers? If not, what properties must hold to make it valid?

3. [10] **Bayesian Networks**

Consider the following cell phone domain. Having a charged battery (CB) enables a cell phone to be operational (PO). Having an operational cell phone, being in the coverage area (CA), and knowing the number to call (KN) enables the placing of a call (PC). Assume there are no other dependencies. All random variables are Boolean.

(a) [7] Draw the Bayesian Network for this domain, including nodes, arcs and the form of each conditional probability table (CPT) at each node.

(b) [3] How many independent values are in the full joint probability distribution for this problem?

4. [15] **Neural Networks**

(a) [3] Can a Perceptron learn the function defined by the following 5 data points? If so, construct a Perceptron that classifies all these values correctly; if not, argue why there is none.

Input1	Input2	Desired Output
-1	0	0
0	0	1
2	1	0
1	0	1
1	-2	0

(b) [2] Which of the following describes the process of *classifying* a test example using a feed-forward neural network that uses back-propagation learning?

- (i) Activation values are propagated from the input nodes through the hidden layers to the output nodes.
- (ii) Activation values are propagated from the output nodes through the hidden layers to the input nodes.
- (iii) Weights on the arcs are modified based on values propagated from input nodes to output nodes.
- (iv) Weights on the arcs are modified based on values propagated from output nodes to input nodes.
- (v) Arcs in the network are modified, gradually shortening the path from input nodes to output nodes.
- (vi) Weights on arcs from the input nodes are compared to the weights on the arcs coming into the output nodes, and then these weights are modified to reduced the difference.

(c) [2] Which of the options in (b) describes the process of *learning* in a feed-forward neural network that uses back-propagation?

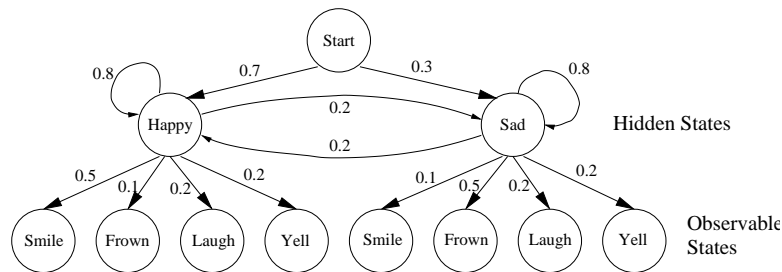
4. (cont.)

(d) [4] It is often stated that local minima can cause difficulties for a feed-forward neural network that uses back-propagation. Local minima of what function of what arguments? Also explain why this is a problem in terms of the type of search method that is performed by back-propagation.

(e) [4] How can over-fitting be avoided when training a feed-forward neural network?

5. [20] **Hidden Markov Models**

The following HMM describes Chuck's emotional state. Each day he is either *Happy* or *Sad*. But he hides his emotional state so all you can observe is his facial expression each day, which is either a *Smile*, *Frown*, *Laugh*, or *Yell*. Let q_t define his emotional state on day t . Let O_t indicate his observed facial expression on day t . The arcs from the Start state give the probabilities of his initial emotional state on day 1 (i.e., they specify the π vector in the HMM).



(a) [6] What is $P(q_2 = \text{Happy})$?

(b) [7] What is $P(O_2 = \text{Frown})$?

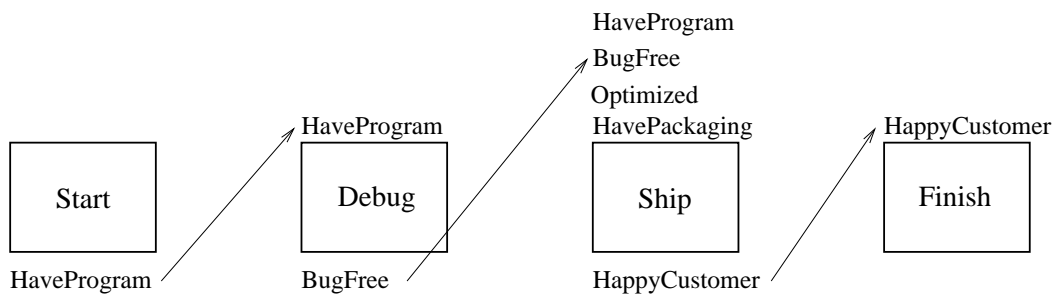
(c) [7] What is $P(q_1 = \text{Sad} \mid O_1 = \text{Frown})$?

6. [20] **Planning**

(a) [12] Consider using a partial-order planner for the development and release of a software product. Assume there exist the following four STRIPS operators:

Operator	Precondition	Effects
Optimize	HaveProgram	Optimized \wedge \neg BugFree
Debug	HaveProgram	BugFree
Ship	HaveProgram \wedge BugFree \wedge Optimized \wedge HavePackaging	HappyCustomer
DesignPackaging	HaveProgram	HavePackaging

The initial state specifies that *HaveProgram* is true, and the goal state is one in which *HappyCustomer* is true. From the above information, say we have constructed the following partial-order plan (which shows only causal links, and has preconditions on top and effects on the bottom of each step):



(i) [6] Add steps and causal links to the above figure to create a new partial-order plan in which there are *no open preconditions*.

(ii) [4] What *threat(s)* exist, if any, in your partial-order plan produced in (i)? If there are threats, add temporal links (as dotted arcs) to the above figure to resolve them all.

6. (cont.)

(iii) [2] Give one possible solution plan (i.e., total ordering of steps) that is implied by your partial-order plan produced in (i) and (ii).

(b) [4] Define a STRIPS operator, $move(x,y)$, that gives the preconditions and effects for a move of the blank square, B , from position x to position y in the 8-puzzle. Use the following predicates:

- (i) $At(t, x)$ means tile t is at position x
- (ii) $At(B, x)$ means the blank square is at position x
- (iii) $Movable(t)$ means tile t is adjacent (north, south, east or west) to the blank square

(c) [4] Define one or more effect axioms in the Situation Calculus that capture the same meaning as the operator $move(x,y)$ in (b). (You may modify the arguments of the predicates to include a situation variable, if desired.)