There are six questions on this exam. You must answer them all. All answers must be in the test booklet. All questions have the same value for grading. Your grade depends on the correctness, completeness and style of your answers.

1. Consider the problem of transporting a number of cars among islands, using a ferry. Each island is accessible from all other islands. Cars can be boarded onto the ferry or debarked from it. The ferry can carry only one car at a time.

There are five islands: A, B, C, D and E. There are four cars: 1, 2, 3 and 4. Car 1 starts on island A, cars 2 and 3 start on island B, and car 4 starts on island C. The ferry starts on island D. The goal is to get cars 1, 2, and 4 to island D and car 3 to island E, in the minimum time.

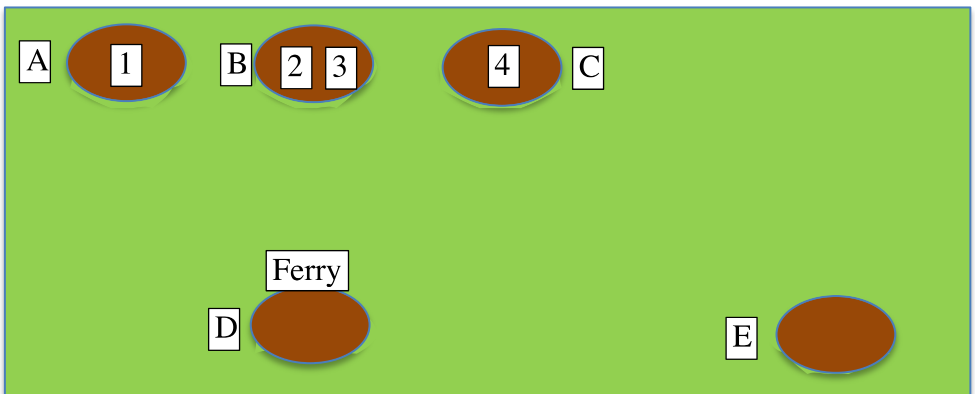


Figure 1. The islands and cars.

The distances between the islands are (in kilometers):

	A	**B**	**C**	**D**	**E**
A	0	1	2	4	18
B	1	0	1	6	15
C	2	1	0	4	10
D	4	6	4	0	12
E	18	15	10	12	0

You will formulate this as a planning problem using STRIPS (PDDL) action schemas.

(a) Describe the start and goal states for this problem using STRIPS notation. Indicate what predicates are necessary for complete descriptions of these states. You don't have to write out *every* necessary literal (all islands and cars and all their relations) but indicate what literals are needed and give a few examples.

(b) Write action schemas for three actions: boarding a car on the ferry, sailing from one island to another, and debarking the car from the ferry.

(c) Using your schemas from part (b), show how resolution is used to make the sequence of three moves: sailing from island D to island B, boarding car 2, and sailing to island D.

(d) Discuss what search method you would choose to guide this planner, and why.

Answer:

(a) At(Car1, IslandD), At(Car2, IslandD), At(Car4, IslandD), At(Car3, IslandE)

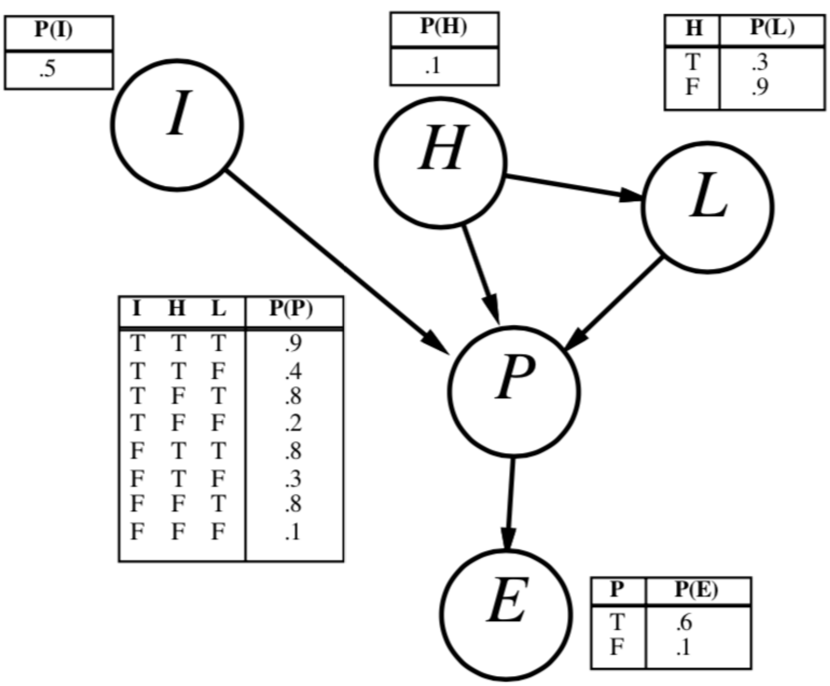
(b)
Action(Board(Cari),
PRECOND: -OccupiedFerry(), At(Cari, Islandj)
EFFECT: OccupiedFerry(), - At(Cari, Islandj), At(Cari, Ferry)

etc. – for Sail use moves of Ferry

(c) Init state is conjunction of literals about cars, ferry, islands. Show binding of preconditions to state, then replacement/addition of effects.

(e) Backchaining is good. Hierarchical planning using moves between islands is good. Examination of the distances between the islands reveals that island E is far away from the others, which are clustered together. Since I did not specify where the ferry must end up, the shortest solutions are those that move the three cars 1, 2, 4 to island D, then move car 3 to island E and ends there. So the move to island E must be last. This means any greedy algorithm will work well here. A* works perfectly if we use the distance table as the heuristic, because the distances are accurate so the heuristic is perfect.

2. Bayes Nets



P(I)
.5

P(H)
.1

H	P(L)
T	.3
F	.9

I	H	L	P(P)
T	T	T	.9
T	T	F	.4
T	F	T	.8
T	F	F	.2
F	T	T	.8
F	T	F	.3
F	F	T	.8
F	F	F	.1

P	P(E)
T	.6
F	.1

Figure 2: A Simple Bayes Net with Boolean variables I=Intelligent, H=Honest, P=Popular, L=LotsOfCampaignFunds, E=Elected.

Consider the Bayes net shown in Figure 2.

a. Which, if any, of the following are asserted by the network structure (ignoring the CPTs for now)?

(i) P(I, L) = P(I) P(L)
(ii) P(E | P, L) = P(E | P, L, H)
(iii) P(P | I, H) = P(P | I, H, L)

b. Calculate the value of P (i, h, ¬l, p, ¬e).

c. Calculate the probability that someone is intelligent given that they are honest, have few campaign funds, and are elected.

Answer:

(a) (3) (i) and (ii).
(i) is true because L is a nondescendant and I has no parents.
(ii) is true—both are equal to $\mathbf{P}$(E—P).
(iii) cannot be asserted by the network structure—the CPTs may imply equality.

(b) (3) $P(i, h, \neg l, p, \neg e) = P(i)P(h)P(\neg l|h)P(p|i, h, \neg l)P(\neg e|p)$
$= .5 \times .1 \times .7 \times .4 \times .4 = .056$

(c) (4) $\mathbf{P}(I|h, \neg l, e) = \alpha \mathbf{P}(I, h, \neg l, e)$
$= \alpha(\mathbf{P}(I, h, \neg l, p, e) + \mathbf{P}(I, h, \neg l, \neg p, e))$
$= \alpha(\langle .084, .063\rangle + \langle .021, .0245\rangle) = \alpha\langle .105, .0875\rangle \approx \langle .545, .455\rangle$

Every variable is conditionally independent of all its non-descendants in the graph given the value of all its parents.

3. A simplified model for the stock market is that it can be in one of two states: Bull Market or Bear Market. Each day, the market can be observed to be Up or Down. In a Bull Market, the market is Up 80% of the days and Down 20% of the days. A Bull Market is 70% likely to continue on the next day, and 30% likely to switch to a Bear Market. In a Bear Market, the market is down 70% of the days and Up on the other days. A Bear market is 80% likely to continue on the next day, and 20% likely to switch to a Bull Market. On the first day, the market is equally likely to be either Bull or Bear.

a. Show how the market can be formulated as an HMM. What is the probability of transitioning from every state to every other state? What is the probability of observing each output (Up or Down) in each state?

b. Use the filtering algorithm to determine the probability of being in each state at time *t* after observing Up-Down-Down, for *t*=1,2,3.

Answer:

(a) State variable is Market, with values Bull, Bear. Sensor variable is Change, with values Up, Down. State CPT is State(t-1) and P(State(t)).

P(M0) = [.5, .5] from given conditions.

We will write the distributions as <Bull, Bear>.

On day 1, observe Change = Up.

The state prediction step is P(M1) = SUM P(M1 | m0) P(m0)   sum over all m0

= <.7, .3> * .5 + <.2, .8> * .5 = <.35, .15> + <.1, .4> = <.45, .55>

Updating for the sensor observation on day 1:

P(M1 | Up) = P(Up | M1) P(M1) = <.8, .3> <.45, .55> = <.360, .165>

normalized is <.686, .314>

The state prediction step to day 2 is:

P(M2 | Up) = P(M2 | m1=Bull)P(m1=Bull | c1=Up) + P(M2 | m1=Bear)P(m1=Bear | c1=Up) =

<.7, .3> * .686 + <.2, .8> * .314 = <.480, .206> + <.063, .251> =

<.543, .457>

Updating for the sensor observation on day 2 is:

P(M2 | c1=Up, c2=Down) = P(c2=Down | M2) P(M2 | c1=Up) =

<.2, .7> <.543, .457> = <.109, .320> normalized is <.254, .746>

The state prediction step to day 3 is:

P(M3 | c1=Up, c2=Down) = P(M3 | m2=Bull) P(m2=Bull | c1=Up, c2=Down) +

P(M3 | m2=Bear) P(m2=Bear | c1=Up, c2=Down) =

<.7, .3> * .254 + <.2, .8> * .746 = <.178, .076> + <.149, .597> =

<.327, .673>

Updating for the sensor observation on day 3 is:

P(M3 | c1=Up, c2=Down, c3=Down) = P(c3=Down | M2) P(M2 | c1=Up, c2=Down) =

<.2, .7> <.254, .746> = <.051, .522> = <.089, .911> normalized

4. Markov Decision Processes

a. The Bellman update equation for value iteration is:

$$U_{i+1}(s) \leftarrow R(s) + \gamma \max_{a \in A(s)} \sum_{s'} P(s' \mid s, a) U_i(s')$$

Discuss the advantages and disadvantages of having the discount γ close to zero. Then discuss the advantages and disadvantages of having γ close to one.

b. The equation in part (a) is the Bellman update for value iteration. Write the corresponding update equation for policy iteration. This equation is simpler than the value iteration update equation. How does this help us solve this equation?

Answer:

a. Discount close to 0 speeds convergence, but can lead to missing long-term effects of actions. Discount close to 1 gives slow convergence, but leads to good results.

b. Same equation but without max, and a replaced by policy. This makes the equation solvable by linear algebra, since it is n equations in n unknowns.

5. Consider a concept learning problem in which each instance is a real number, and in which each hypothesis is an interval over the reals. More precisely, each hypothesis in the hypothesis space H is of the form $a < x < b$, where a and b are any real constants, and x refers to the instance. For example, the hypothesis $4.5 < x < 6.1$ classifies instances between 4.5 and 6.1 as positive, and all other numbers as negative. Explain why the Candidate Elimination algorithm cannot create a maximally specific consistent hypothesis for some possible sets of training examples (so no version space can be created.) Suggest a slight modification to the hypothesis representation so that there will be maximally specific hypotheses for every possible set of training examples.

Answer:

If the set of training examples is not a single interval, e.g. it looks like this:

- - - - - - - - - + + + - - - + + + - - - - - + + - - - -

then H will not be able to model it.

The modification of the hypothesis space is to use disjunctions, which solves this problem.

6. Q-learning

a. How long a sequence of training examples is needed to guarantee that Q-learning will learn the optimal policy?

b. One effective TD learning approach is to use a very optimistic (high) estimate for the initial utilities of actions. Why does this help in TD learning (what problem does it help avoid)?

c. Another approach is for a Q-learning agent to act randomly on some fraction of actions, while slowly decreasing this fraction. Why does this help in Q-learning (what problem does it help avoid)?

Answer:

a. Infinite

b. Increases exploration – helps avoid early convergence to suboptimal policy.

c. Increases exploration – helps avoid early convergence to suboptimal policy.