Review Final Exam COSC 6368 Fall 2016

1) D-Separability

Assume that the following belief network is given, consisting of nodes A, B, C, D, and E that can take values of true and false.

a)

a) Are C and E independent; is C|∅ and E|∅ d-separable? Give a reason for your answer! ∅ denotes “no evidence given”[2]

There are two paths from C to E

C-B-E

C-D-E

Neither path is blocked —it would only be blocked if both arrows would be pointing to B, and both arrow would be pointing to D respectively; consequently, C and E are not d-separable.

b) Is E|CD d-separable from A|CD? Give a reason for your answer! [3]

There are 2 paths between A and E:

A-C-B-E neither node C(in evidence) satisfies patterns 1a,1b or 2 nor node B (not in evidence) satisfies pattern 3; therefore this path is not blocked

A-C-D-E node D(in evidence) satisfies pattern 1a; consequently, this path is blocked.

However, as not all paths are blocked between E and A, A|CD is not d-separable from E|CD/

1. Learning

a) What is the main differences between reinforcement learning and inductive learning, also called learning from examples? [

b) We would like to predict the gender of a person based on two binary attributes: leg-cover (pants or skirts) and beard (beard or bare-faced). We assume we have a data set of 20000 individuals, 10000 of which are male and 10000 of which are female. 75% of the 10000 males are barefaced. Skirts are present on 50% of the females. All females are bare-faced and no male wears a skirt.

Compute the information gain of using the attribute leg-cover for predicting gender! [4]

c) What is the key contribution of the backpropagation algorithm? What problems does it solve? Why is there no backpropagation algorithm for perceptrons?

3) Evolutionary Computing

a) What roll do crossover operators play in EC systems?

b) What role do mutation operators play in EC systems?

c) What advantages do you see in using probabilistic search algorithms—in contrast to deterministic search algorithms, such as simulated annealing or evolutionary computing?

4) Reinforcement Learning

a) Apply temporal difference learning[[1]](#footnote-1) to the DEF World, depicted below, relying on the following assumptions: [4]



* If state 1 or 4 are visited a reward of 0 is obtained
* Utilities of the 4 states are initialized with 0
* The agent starts in state 2

The agent applies s(ending in state 3)-w-e. What are the utilities of states 2, 3, 4, and after those 3 operators have been applied? Do not only give the final result but also how you derived the final result including formulas used!

c) Give the Bellman equation for state 2 of the DEF world! [2]

d) Assume you have a policy that always selects the action that leads to the state with the highest expected utility. Present arguments that this is usually not a good policy by describing scenarios in which this policy leads to suboptimal behavior of the agent [4]!

e) Assume you use temporal difference learning in conjunction with a random policy which choses actions randomly assuming a uniform distribution. Do you believe that the estimations obtained are a good measurement of the “goodness” of states, that tell an intelligent agent (*assume the agent is smart!!*) what states he/she should/should not visit? Give reasons for your answer! [3]

f) What role does the learning rate play in temporal difference learning; how does running temporal difference learning with low values of  differ from running it with high values of ? [2]

g) Assume you run temporal difference/Q- learning with high values of γ—what are the implications of doing that? [2]

5) Naïve Bayes

Naïve Bayesian systems make the conditional independence assumption when for example computing P(D|S1,S2,S4). What assumptions are exactly made? What advantages do you see in the approach? What are the drawbacks of making the conditional independence assumption?

a) What are the main differences between supervised learning and reinforcement learning? [4]

SL: static world[0.5], availability to learn from a teacher/correct answer[1]

RL: dynamic changing world[0.5]; needs to learn from indirect, sometimes delayed feedback/rewards[1]; suitable for exploration of unknown worlds[1]; temporal analysis/worried about the future/interested in an agent’s long term wellbeing[0.5], needs to carry out actions to find out if they are good—which actions/states are good is (usually) not know in advance1[0.5]

e) Assume you use temporal difference learning in conjunction with a random policy which choses actions randomly assuming a uniform distribution. Do you believe that the estimations obtained are a good measurement of the “goodness” of states, that tell an intelligent agent (*assume the agent is smart!!*) what states he/she should/should not visit? Give reasons for your answer! [3]

Not really; as we assume an intelligent agent will take actions that lead to good states and avoids bad states, the agent that uses the random policy might not recognize that a state is a good state if both good and bad states are successors of this state; for example,

 S2: R=+100

S1:R=-1

 S3: R=-100

Due to the agent’s policy the agent will fail to realize the S1 is a good state, as the agent’s average reward for visiting the successor states of S1 is 0; an intelligent agent would almost always go from S1 to S2, making S1 a high utility state with respect to TD-learning.

c) What role does the learning rate play in temporal difference learning; how does running temporal difference learning with low values of  differ from running it with high values of ? [2]

It determines how quickly our current beliefs/estimations are updated based on new evidence.

d) Assume you run temporal difference learning with high values of γ—what are the implications of doing that? [2]

If γ is high the agent will more focus on its long term wellbeing, and will shy away from taking actions—although they lead to immediate rewards—that will lead to the medium and long term suffering of the agent.

γ is 1.0 and  is 0 e) Assume you use temporal difference learning in conjunction with a random policy which choses actions randomly assuming a uniform distribution. Do you believe that the estimations obtained are a good measurement of the “goodness” of states, that tell an intelligent agent (*assume the agent is smart!!*) what states he/she should/should not visit? Give reasons for your answer! [3]

Not really; as we assume an intelligent agent will take actions that lead to good states and avoids bad states, the agent that uses the random policy might not recognize that a state is a good state if both good and bad states are successors of this state; for example,

 S2: R=+100

S1:R=-1

 S3: R=-100

Due to the agent’s policy the agent will fail to realize the S1 is a good state, as the agent’s average reward for visiting the successor states of S1 is 0; an intelligent agent would almost always go from S1 to S2, making S1 a high utility state with respect to TD-learning.

c) What role does the learning rate play in temporal difference learning; how does running temporal difference learning with low values of  differ from running it with high values of ? [2]

It determines how quickly our current beliefs/estimations are updated based on new evidence.

d) Assume you run temporal difference learning with high values of γ—what are the implications of doing that? [2]

If γ is high the agent will more focus on its long term wellbeing, and will shy away from taking actions—although they lead to immediate rewards—that will lead to the medium and long term suffering of the agent.

* .5
1. and **not** Q-learning. [↑](#footnote-ref-1)