Unit 6 - Week 5-Learning enabled by Prior Theories

Assignment 5

The due date for submitting this assignment has passed.
As per our records you have not submitted this assignment.

Due on 2020-04-01, 23:59 IST.

1) Explanation Based Learning typically takes a predefined Domain Theory and modifies it in such a way that new more complex rules are created as specific combinations of several existing rules. The goal is to make problem-solving more efficient and the rule formation process is guided by the problem examples considered for training. Sometimes these more complex rules are called Macro operators but there is more general term used for this phenomenon. Which? *1 point*

- Aggregation
- Unification
- Compilation
- Abstraction

No, the answer is incorrect.
Score: 0
Accepted Answers:
Compilation

2) **Domain Theory:** *3 points*

SafeToDrive(P,V) :- Sober(P), Errorfree(V), Familiar with vehicle (P,V)
Errorfree(V):- Passed an inspection (V). Errorfree(V):- Worked when driven last (V).
Familiar with vehicle (P,V):- Have driven it before (P, V). Familiar with vehicle (P,V):- Have taken a course on it (P, V).
Sober(P) :- Have a not tasted alcohol since two days (P). Sober(P) :- Passed an alcohol test (P)
Have taken a course on it (P,V) :- Theoretical course (P,V). Have taken a course on it (P,V) :- Practical course (P,V)
Passed an alcohol test (X):- Self test (X). Passed an alcohol test (X):- Police test (X).
Passed an inspection (X):- Private inspection(X). Passed an inspection(X):- Authority inspection(X).

Operational predicates:
Worked when driven last, Have driven it before, Practical course, Theoretical course, Have a not tasted alcohol since two days, Self test, Police test, Private inspection, Authority inspection

Training Example:
SafeToDrive (John,Buggy)
Authority inspection (Buggy), Practical course (John, Buggy), Police test(John).

Goal concept:
SafeToDrive(P, V)

Explain and generalize from SafeToDrive(John, Buggy). How does the resulting rule look like after applying an EBL algorithm?

- SafeToDrive(John,Buggy):- Authority inspection(Buggy), Practical course (John,Buggy), Police test (John).
- SafeToDrive(P,V):- Authority inspection (V), Practical course (P,V), Police test (P).
- SafeToDrive(John,Buggy):-Passed an inspection (Buggy), Have taken a course on it (John,Buggy), Passed an alcohol test (John).
- SafeToDrive(P,V):-Passed an inspection (V), Have taken a course on it (P,V), Passed an alcohol test (P).

No, the answer is incorrect.
Score: 0
Accepted Answers:
SafeToDrive(P,V):-Authority inspection (V), Practical course (P,V), Police test (P).

3) Which of the following inference schemes correspond to Abduction? *1 point*

- All men are mortal;
  Socrates is a man;
  ∴ Socrates is mortal.
- Socrates is a man;
  Socrates is mortal;
  ∴ All men are mortal.
- All men are mortal;
  Socrates is mortal;
  ∴ Socrates is a man.

No, the answer is incorrect.
Score: 0
Accepted Answers:

All men are mortal;
Socrates is mortal;
∴ Socrates is a man.

4) **A general ILP algorithm:** *1 point*

Initialize (Hypotheses, Examples)
Repeat
1. H= select (Hypotheses, Examples)
2. Hs = refine (H, Examples)
3. Hypotheses = reduce (Hypotheses+Hs, Examples)

Until stopping-criterion (Hypotheses, Examples)
Return Hypotheses.

Example

Initial hypothesis: mammal(X),
Background knowledge:
givemilk(gorilla), givemilk(elephant), givemilk(wolf),
havefourlegs(elephant), havefourlegs(wolf), predator(wolf).
Positive Examples: mammal(gorilla), mammal(elephant), mammal(wolf)

How does the Hypotheses set (theta subsumptions/specializations) look like after a first iteration of a topdown ILP search based on the first positive example mammal(gorilla) ?

- (mammal(gorilla):- givemilk(gorilla).)
- (mammal(X):- givemilk(X), mammal(X):-havefourlegs(X), mammal(X):-predator(X).)
- (mammal(X):- givemilk(X) .)

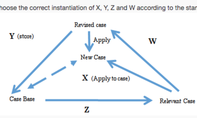
No, the answer is incorrect.
Score: 0
Accepted Answers:
(mammal(X):- givemilk(X), mammal(X):-havefourlegs(X), mammal(X):-predator(X).)

5) Tuning of indexing mechanisms may be used in one of the main phases of Case Based Reasoning. Which? *5 points*

- RETRIEVE
- REVISE
- REUSE
- RETAIN

No, the answer is incorrect.
Score: 0
Accepted Answers:
RETAIN

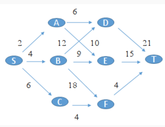
6) Please choose the correct instantiation of X, Y, Z and W according to the standard Case-based Reasoning Model *1 point*



- X Y Z W = REUSE RETAIN RETRIEVE REVISE
- X Y Z W = RETRIEVE REVISE REUSE RETAIN
- X Y Z W = RETAIN REVISE REUSE RETRIEVE
- X Y Z W = RETAIN RETRIEVE REVISE REUSE

No, the answer is incorrect.
Score: 0
Accepted Answers:
X Y Z W = REUSE RETAIN RETRIEVE REVISE

7) A dynamic programming approach to calculating the shortest distance from S to T gives the result= 14. *2 points*
Which result would a greedy forward search give?



- 26
- 27
- 29

No, the answer is incorrect.
Score: 0
Accepted Answers:
29

8) Monte Carlo Reinforcement Learning MC methods learn directly from samples of complete episodes of experience. *3 points*

- First-visit MC: average returns only for first time s is visited in an episode.
- Every-Visit MC: average returns for every time s is visited in an episode

Simplified Algorithm for first-visit Monte Carlo
Initialize state-value functions. Return list(s) ← empty list
2. For all s in all episodes E
3. return ← sum of rewards r in episode from state to termination.
4. if this is the first occurrence of this state s add the calculated return to the returns list (s).
5. Calculate values of all s as average over all return lists (s).

Example: An undiscounted Markov Reward Process with two states A and B. The Transition matrix and reward function are unknown.
Two sample episodes: E1 A->A, r=1 A->B, r=2 B->A, r= -2 A->B, r=2 B->terminate, r= -1.
E2 B->A, r=-2 A->B, r=2 B->terminate, r = -1

Which are the estimated values of A and B after considering the two episodes using the simplified first-visit algorithm?

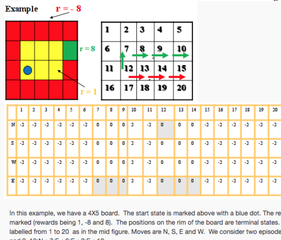
- V(A)= 1,V(B)= -1/2
- V(A)= 3,V(B)= -1
- V(A)= 3/2,V(B)= -1
- V(A)= 1/2,V(B)= -2

No, the answer is incorrect.
Score: 0
Accepted Answers:
V(A)= 3/2,V(B)= -1

9) *5 points*

Initialize $Q(s, a)$ arbitrarily
Repeat (for each episode)
Initialize s
Repeat (for each step of the episode)
Take action a, observe r, s' $Q(s, a) \leftarrow Q(s, a) + \alpha[r + \gamma \max_{a} Q(s', a') - Q(s, a)]$; $s \leftarrow s'$

Example



In this example, we have a 4X5 board. The start state is marked above with a blue dot. The rewards for all positions are also marked (rewards being 1, -5 and 5). The positions on the rim of the board are terminal states. The board elements are labelled from 1 to 20 as in the mid figure. Moves are N, S, E and W. We consider two episodes: 1. 12:E->13:E->14:E-> 15 and 2. 12:N->7:E->8:E->9:E->10.
What are the updated versions of the Q elements for state 12, 13, 14, 7,8 and 9 after considering these two episodes? α = γ = 1.

- 12,E=1 13,E=1 14,E= -10 12,N=1 7,E=1 8,E=1 9,E=10
- 12,E=0 13,E=0, 14,E= -10 12,N=0, 7,E=0 8,E=0 9,E=10
- 12,E=1 13,E=2 14,E= -8 12,N=1 7,E=2 8,E=3 9,E=-8
- 12,E=1 13,E=2 14,E= -8 12,N=1, 7,E=2 8,E=4 9,E=10

No, the answer is incorrect.
Score: 0
Accepted Answers:
12,E=1 13,E=1 14,E= -10 12,N=1 7,E=1 8,E=1 9,E=10