

Planning and Management of Uncertainty (103211/01)

Section A Planning (Dr M Fox and Dr A Coddington)**Question 1**

- (a) Briefly describe the planning strategy used by Graphplan-based planners. **[2 Marks]**
- (b) Construct the graph for the following problem, stopping when the goals are pairwise non-mutex in a layer. Make sure you mark all of the binary mutex relations present in the graph.

Initial state: {}

Goal state: q & r

Actions:

A

pre:

add: p,q

del:

B

pre: q

add: r

del: q

Show how a plan to achieve the goal state can be extracted from the graph you have constructed. **[4 Marks]**

- (c) Why is Graphplan able to terminate efficiently on unsolvable problems? Explain the termination criteria used. **[4 Marks]**
- (d) Suppose that a researcher is implementing a modified version of Graphplan in the hopes of optimising its search performance. It is proposed that memoized bad goal sets could be managed using a data structure that does not always identify a set as bad even when it is, but with a small chance of making this mistake. When it identifies a set as bad it does so correctly. The consequences of this will be that the identification of bad goal sets is speeded up.
- What will be the impact of this design decision on the search process? **[2 Marks]**
 - What is the impact on the soundness of the planner? **[2 Marks]**
 - What is the impact on completeness of the search? **[2 Marks]**
 - Will the planner still be able to offer the same termination properties as the original Graphplan algorithm? Explain your answer. **[4 Marks]**

continued

Question 2

- (a) Explain what is meant by a relaxed plan. **[2 Marks]**
- (b) Describe the method by which relaxed distance estimates are constructed in:
- i. HSP **[2 Marks]**
 - ii. FF **[2 Marks]**
 - iii. Sapa **[3 Marks]**
- (c) An AI researcher is constructing a partial order planner and has decided to use a heuristic distance estimate based on the use of a relaxed plan. The nodes in the partial order search space are partial plans, represented as 4-tuples (Os, Ts, Cs, E, Ps) , where Os is the collection of actions selected so far, Ts is a collection of temporal constraints, which might be both relative and metric, and which specify how the time points associated with actions are organised, Cs is a collection of causal links, E is an environment and Ps is the set of outstanding goals.
- Describe the transition function that generates states in the search process. **[2 Marks]**
- (d) With reference to the standard strategy (as used by FF and HSP) for relaxing action descriptions, propose one way in which the transition function could be relaxed to enable an estimate of the distance between one node and a solution node to be computed. Explain how the relaxed plan construction would work using your modification. **[5 Marks]**
- (e) What would be the difficulty in computing an informative relaxed distance estimate for partial order planning? Propose two ways in which this difficulty might be alleviated. **[4 Marks]**

Question 3

- (a) Show how the PDDL2.1 level 3 durative action DISEMBARK-TRUCK would be encoded for TGP and describe the differences between the two encodings. **[3 Marks]**

```
(:durative-action DISEMBARK-TRUCK
  :parameters (?driver - driver ?truck - truck
              ?loc - location)
  :duration (= ?duration 1)
  :condition (and (over all (at ?truck ?loc))
                  (at start (driving ?driver ?truck)))
  :effect (and (at start (not (driving ?driver ?truck)))
               (at end (at ?driver ?loc))
               (at end (empty ?truck))))

(:durative-action DRIVE-TRUCK
  :parameters (?truck - truck ?loc-from - location
              ?loc-to - location ?driver - driver)
  :duration (= ?duration 10)
  :condition (and (at start (at ?truck ?loc-from))
                  (over all (driving ?driver ?truck))
                  (at start (link ?loc-from ?loc-to)))
  :effect (and (at start (not (at ?truck ?loc-from)))
               (at end (at ?truck ?loc-to))))
```

- (b) Using the PDDL2.1 level 3 durative actions shown above, explain how VHPOP generates a plan to solve following goal. The initial state is also given below.

```
(:goal (and (at driver1 s1) (at truck1 s1))

(:initial ((at truck1 s0) (link s0 s1) (driving driver1 truck1)))
```

In your discussion you should:

- i. show how the underlying algorithm has been adapted to reason with PDDL2.1 level 3 durative actions; **[4 Marks]**
 - ii. explain how any temporal constraints are generated and solved. **[4 Marks]**
- (c) Describe what techniques may be used by VHPOP to control search. **[4 Marks]**
- (d) Compare and contrast the advantages and disadvantages associated with VHPOP and another temporal planner. **[5 Marks]**

Question 4

- (a) Describe briefly how conformant planning and contingency planning deal with uncertainty in the world. **[2 Marks]**
- (b) The ADL operator MEDICATE represents the process of medicating a patient where I means that the patient is infected, H means they are hydrated and D means that they are dead.

```
(:action MEDICATE
  :condition ()
  :effect (and (when I (not I))
               (when (not H) D)))
```

```
(:action DRINK
  :precondition: ()
  :effect H)
```

What are the *aspects* of the MEDICATE operator? **[2 Marks]**

- (c) Suppose there are two possible initial worlds, $w1$ and $w2$, shown below and the goal is to get the patient uninfected and not dead.

$w1 = ((\text{not } I) (\text{not } H) (\text{not } D))$

$w2 = (I H (\text{not } D))$

```
(:goal (and (not I) (not D)))
```

How does the conformant planner CGP generate plans? In your answer you should explain the changes made to the Graphplan algorithm to enable CGP to reason about

- i. uncertainty in the initial conditions; **[5 Marks]**
- ii. actions with uncertain outcomes. **[3 Marks]**

You may use the ADL operator above together with the ADL operator DRINK and the initial conditions and goal to illustrate your answer.

- (d) Suppose the sensing action EXAMINE senses whether the patient is ill or not (i.e. senses the truth of I). Describe how CGP might be adapted to incorporate sensing actions and generate contingent plans (for example the planner SGP)? **[4 Marks]**
- (e) Compare and contrast two other approaches used in planning which deal with uncertainty in the world. **[4 Marks]**

Section B Management of Uncertainty (Dr A Hunter)

Question 5

- (a) Discuss the meaning of the word “uncertainty” in the context of this module, and explain why it is important to model it. **[3 Marks]**
- (b) For predictive two-class classification, one approach is to define decision functions for each class k which equal the posterior probability of class membership (formula below), and assign to the class with the highest value. Assuming that the class conditional probability distributions, $p(x|\mathcal{C}_k)$, are normal (formula below) with equal covariance matrices, state and justify the form of the decision function. **[5 Marks]**

$$y_k(\mathbf{x}) = p(x|\mathcal{C}_k)p(\mathcal{C}_k)$$

$$p(x) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left(\frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^{-1}(\mathbf{x} - \mu)\right)$$

- (c) Define *parametric*, *non-parametric* and *semi-parametric* pattern recognition, briefly state examples of each type in the context of classification, and discuss the relative advantages and disadvantages of these three broad approaches. **[12 Marks]**

Question 6

- (a) Describe the Schema Theorem, and the implications it has, with the limitations imposed by the crossover disruption issue, for the representation of problems for solution by the Genetic Algorithm. **[8 Marks]**
- (b) You are presented with a two-class classification problem; there are 10 continuous input features, and a single output feature. A supervised data set of 10,000 examples is available. One analyst suggests using Genetic Programming (GP) with the operator set $\{*, +, \phi\}$ (where ϕ is the logistic sigmoid function) to solve the problem; a second analyst objects that this is simply an inferior method of specifying a standard neural network. Discuss this issue. You should discuss any similarities and differences between the resulting models and training processes, and the consequences of these differences. State and justify your own preferred approach (Genetic Programming or Neural Network) for this problem domain. Outline the attributes that might lead you to select the other approach. **[12 Marks]**

Question 7

- (a) Explain how the gradient vector is obtained in feedforward neural network training by the back propagation algorithm, and discuss the efficiency of the technique. **[6 Marks]**
- (b) One problem with basic gradient descent is slow convergence. Briefly explain how second-order training algorithms attempt to address this problem. **[6 Marks]**
- (c) Kohonen's algorithm is used to train Self-Organizing Feature Maps. Discuss this algorithm in detail, and explain how it achieves "topological ordering" of the SOFM's neurons. **[8 Marks]**

Question 8

- (a) Describe the stages in training a radial basis function neural network. Briefly describe (in a single sentence each) several alternative algorithms that can be used for each stage, and describe in detail one particular algorithm for each stage, stating the particular advantages of the chosen algorithms over alternatives. **[8 Marks]**
- (b) Overfitting is a major problem with all types of neural network. Briefly describe the approaches that can be taken to mitigate the problem for the following network types: Probabilistic Neural Network (PNN), Radial Basis Function (RBF) and Multilayer Perceptron (MLP). **[6 Marks]**
- (c) Compare and contrast *feature selection* and *feature extraction* as approaches to dimensionality reduction, giving advantages and disadvantages of the two approaches. **[6 Marks]**