Cooperating Intelligent Systems Exam

Stefan Byttner and Sławomir Nowaczyk

26 October 2011

# General Rules

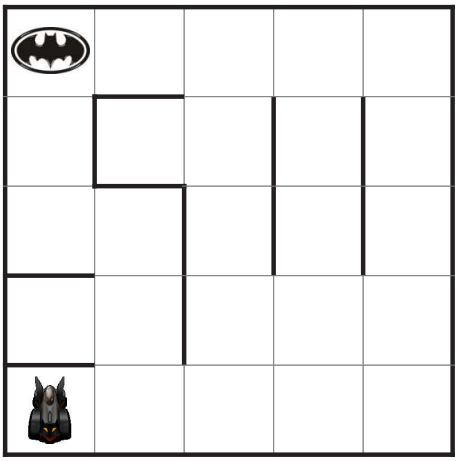
The total number of points for this exam is 40. You must achieve at least 50% of that in order to continue to the oral exam. No books, mobile phones or calculators are permitted during the exam.

## Part A
# Open Questions

You can get a maximum of 20 points in this part.
All answers in this part of the exam must be well-motivated!

## 1. Agents

Imagine that Batman needs to use his Batmobile to reach police commissioner Jim Gordon who is calling him using Bat-Signal:



Batmobile, like most cars, is directional and, at any time, faces one of the directions $d \in \{N, S, W, E\}$. With a single action, Batman can either move forward, at an adjustable velocity $v$, or turn. The turning actions are *left* and *right*, which change the Batmobile's direction by 90 degrees. Turning is only permitted when the velocity is zero, and it leaves it at zero (this is a very early prototype of Batmobile). The movement actions are *accelerate* and *slow down*. *Accelerate* increments the velocity by 1, while *slow down* decrements the velocity by 1. In both cases, Batmobile then moves a number of squares equal to its **new** velocity, in the direction it is facing. Any action which would reduce $v$ below 0 or above a maximum speed $V_{max}$ is illegal. The Batman needs to find a plan which parks it (stationary) on the bat-signal square, using as few actions (time steps) as possible.
As an example: if the Batmobile shown was initially stationary, it might first turn to the east using *left* action, then use *accelerate* action, thus moving one square east, and then use *accelerate* again, thus moving two more squares east. Afterwards, Batman would of course need to slow in order to turn.

a) If the grid is M by N, what is the size of the state space? Give either an exact value, or — if you think that is infeasible — an upper bound. Motivate your answer. (1p)
Upper bound: $M \times N \times 4 \times (V_{max} + 1)$ The state is specified by the position, direction and velocity. Not all configurations may actually be reachable (obvious case is $V_{max} > M > N$).

b) What is the maximum branching factor of this problem? You may assume that illegal actions are not returned by the successor function. Motivate your answer. (1p)
The maximum branching factor is 3. When stopped, Batman can turn left, turn right, or accelerate.

c) Is the Manhattan distance from the Batman's current location to the bat-signal location an admissible heuristic? Why or why not? (1p)
Manhattan distance is not admissible because Batman can travel faster than 1 square per move.

d) Design another non-trivial heuristic for this problem, and justify that it is admissible (2p)
"Manhattan distance divided by Vmax" or "Current velocity"

## 2. Logic

In answering this question please use the following notation:
Propositional symbols: $F$ - Frodo goes to Mordor, $S$ - Sam goes to Mordor, $K$ - The Fellowship enters Khazad-dum, $D$ - Gandalf dies
Object constants: frodo, sam, gandalf, gollum, witch king of Angmar
Predicates: fellowship member$^1$, traitor$^1$, evil$^1$, man$^1$, hobbit$^1$, kills$^2$, saved$^2$, hates$^2$
Variables: x, y and z

### a) Translate the following sentences (2p)

From English into propositional logic:

- Both Frodo and Sam go to Mordor.
  $F \wedge S$
- Gandalf dies if The Fellowship ever enters Khazad-dum.
  $K \Rightarrow D \equiv D \vee \neg K$
- Neither Frodo nor Sam will go to Mordor if The Fellowship does not enter Khazad-dum.
  $\neg K \Rightarrow \neg F \wedge \neg S \equiv \neg(F \vee S) \vee K \equiv (\neg F \wedge \neg S) \vee K$

From propositional logic into English (or Swedish):

- $D \vee \neg F$
  Frodo will not go to Mordor unless Gandalf dies.
- $F \Leftrightarrow S$
  Frodo and Sam can only go to Mordor together.

From English into FOL:

- There is a traitor among fellowship members.
  $\exists_x \; fellowship\;\; member(x) \wedge traitor(x)$
- One of the hobbits was saved by every fellowship member
  $\exists_x \forall_y \; hobbit(x) \wedge fellowship\;\; member(y) \Rightarrow saved(y, x)$
- Witch king of Angmar will not be killed "by the hand of man"
  $\forall_x \; kills(x, witch\;\; king\;\; of\;\; Angmar) \Rightarrow \neg man(x)$

From FOL into English (or Swedish):

- $\exists_x \; hobbit(x) \wedge hates(gollum, x)$
  Gollum hates some hobbit

- $\forall_x \forall_y \; hates(x, y) \wedge hobbit(y) \Rightarrow evil(x)$
  Only evil creatures hate hobbits
- $\forall_x \; evil(x) \Rightarrow \neg fellowship\ member(x)$
  One cannot be evil and be a member of The Fellowship

**b) Use model enumeration to deduce whether Frodo will go to Mordor or not (1p)**

Assume the following knowledge base:
$F \Rightarrow S$
$\neg(D \wedge S)$
$K \Leftrightarrow D$
$K$

| $F$ | $S$ | $K$ | $D$ | $K \Leftrightarrow D$ | $\neg(D \wedge S)$ | $F \Rightarrow S$ | |
|---|---|---|---|---|---|---|---|
| T | T | T | T | T | F | T | - |
| T | T | T | F | F | T | T | - |
| T | T | F | T | F | F | T | - |
| T | T | F | F | T | T | T | - |
| T | F | T | T | T | T | F | - |
| T | F | T | F | F | T | F | - |
| T | F | F | T | F | T | F | - |
| T | F | F | F | T | T | F | - |
| F | T | T | T | T | F | T | - |
| F | T | T | F | F | T | T | - |
| F | T | F | T | F | F | T | - |
| F | T | F | F | T | T | T | - |
| F | F | T | T | T | T | T | + |
| F | F | T | F | F | T | T | - |
| F | F | F | T | F | T | T | - |
| F | F | F | F | T | T | T | - |

Answer: Frodo will not go to Mordor

**c) Use resolution to check whether Gollum is a member of The Fellowship (2p)**

Assume the following knowledge base:
$\exists_x \; hobbit(x) \wedge hates(gollum, x)$
$\forall_x \forall_y \; hates(x, y) \wedge hobbit(y) \Rightarrow evil(x)$
$\forall_x \; evil(x) \Rightarrow \neg fellowship\ member(x)$
Rewrite KB in CNF:
$hobbit(h)$
$hates(gollum, h)$
$\neg hates(x, y) \vee \neg hobbit(y) \vee evil(x)$
$\neg evil(x) \vee \neg fellowship\ member(x)$
We can try to prove that Gollum **is** a member of The Fellowship, by adding
$\neg fellowship\ member(gollum)$
to the KB, but this does not lead anywhere.
Thus, we try to prove that Gollum **is not** a member of The Fellowship, by adding
$fellowship\ member(gollum)$
to the KB.
Using resolution repeatedly, we obtain:
$\neg evil(gollum)$
$\neg hates(gollum, y) \vee \neg hobbit(y)$
$\neg hates(gollum, h)$
$\emptyset$
Thus, Gollum is not a member of The Fellowship.

# 3. Uncertainty

You are given a task of building a new decision support system that can be used by the Enterprise for detecting Romulan starships.
For simplicity, assume that Enterprise is going to be approached by at most one other vessel at the time, and that in this region of the galaxy, 75% of encountered ships will belong to United Federation of Planets, and all others will be Romulan.
Starship Enterprise spends 75% of its travels in deep space, where probability of encountering another vessel is 40%. Of the remaining time, 80% is spent in planetary systems (where probability of encountering another vessel is 60%) and 20% is spent in black hole boundary (where probability of encountering another vessel is 20%).
Assuming normal operation, Enterprise sensors can detect and recognise other vessels without any problems. However, Romulan ships can be in "cloaking mode", which makes them invisible to the Enterprise's sensors. No other race possesses this technology. Any Romulan vessel has a 20% chance of being cloaked. Even for the most current sensor technology, the only hint of a nearby starship in cloak mode is a slight magnetic disturbance caused by the enormous amount of energy required for cloaking. The Enterprise has a magnetic disturbance sensor, but it is very hard to distinguish background magnetic disturbance from that generated by a nearby starship in cloak mode, as shown in the following probability distribution (numbers correspond to probabilities of that the sensor will report magnetic disturbance, depending on location and presence of cloaked ships):

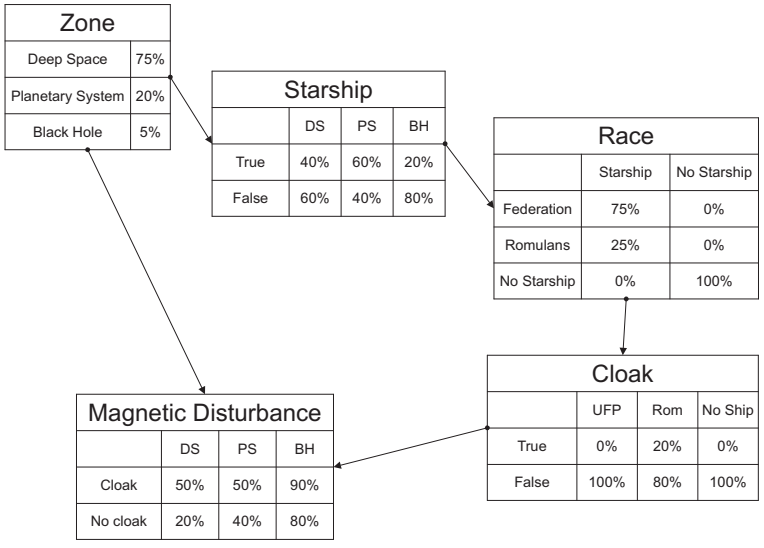
| | Deep Space | Planetary System | Black Hole |
|---|---|---|---|
| Cloaked ship nearby | 50% | 50% | 90% |
| No cloaked ships | 20% | 40% | 80% |

**a) Specify a full joint distribution table for this problem (2p)**

Attention: the way you come up with those probabilities is more important than actual numbers, so make sure your line of reasoning is easy to follow.

| | | DS | | PS | | BH | |
|---|---|---|---|---|---|---|---|
| | | MD | No MD | MD | No MD | MD | No MD |
| UFP | Cloak | 0.000% | 0.000% | 0.000% | 0.000% | 0.000% | 0.000% |
| | No cloak | 4.500% | 18.000% | 3.600% | 5.400% | 0.600% | 0.150% |
| Romulan | Cloak | 0.750% | 0.750% | 0.300% | 0.300% | 0.045% | 0.005% |
| | No cloak | 1.200% | 4.800% | 0.960% | 1.440% | 0.160% | 0.040% |
| No ship | Cloak | 0.000% | 0.000% | 0.000% | 0.000% | 0.000% | 0.000% |
| | No cloak | 9.000% | 36.000% | 3.200% | 4.800% | 3.200% | 0.800% |

**b) Draw a Bayesian Network modelling this problem (2p)**



**c) Explain what is a *causal* Bayesian Network (1p)**

# 4. Learning

**a) Explain the concept of inductive bias (1p)**

**b) Calculate Version Spaces (2p)**

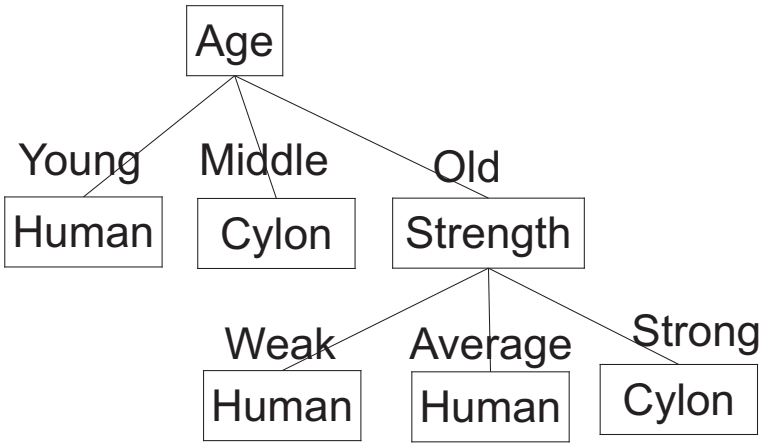
for predicting whether a particular character is a human (as opposed to being a Cylon), based on the data from the table below. Assume that the hypothesis can be represented as either a conjunction or a disjunction of exactly two "attribute operator value" expressions, where each "attribute" is one of "Sex, Age, Strength, Eye Colour, Known Clones", each "operator" is $=$ or $\neq$ and each "value" is an actual value from the data (there are no wildcards).
Is this VS unique?

| | Name | Sex | Age | Strength | Eye Colour | Known Clones |
|---|---|---|---|---|---|---|
| Human | William Adama | Male | Old | Average | Blue | No |
| Human | Laura Roslin | Female | Old | Weak | Brown | No |
| Human | Billy Keikeya | Male | Young | Strong | Green | No |
| Human | Tom Zarek | Male | Young | Strong | Brown | No |
| Cylon | Sharon Agathon | Female | Middle | Strong | Black | No |
| Cylon | Aaron Doral | Male | Old | Strong | Brown | Yes |
| Cylon | D'Anna Biers | Female | Middle | Strong | Green | Yes |
| Cylon | Cavil | Male | Old | Strong | Blue | No |
| Cylon | Leoben Conoy | Male | Old | Strong | Black | No |

$S = G = \text{Strength} \neq \text{Strong} \vee \text{Age} = \text{Young}$

**c) Learn decision tree (2p)**

Use the same data as in point 4.b) However, instead of information gain, use "number of examples misclassified by a majority classifier trained on this node" as node quality measure. Explain some of the problems with such a measure?
A majority classifier is a simple classifier which assigns every example to whichever class is more common in the training set.



## Part B

# Multiple Choice Questions

You can get a maximum of 20 points in this part: 1 point for each correct answer. However, each incorrect answer will result in deduction of 1 point — so if you do not know, do not guess.
There is exactly one correct answer for each question in this part of the exam.

### 1. A complete search algorithm will

- ☐ Make rational expansions of nodes given an admissible heuristic
- ☐ Guarantee to find the lowest cost path among all solutions
- ☐ Guarantee to find a solution if there is one accessible X
- ☐ Find the lowest cost path solution if a solution exist

### 2. For DFS and BFS search algorithms, imagine the situation that some paths are extremely long (or even infinite)

- ☐ Depth-first search will perform well
- ☐ Depth-first search may perform badly X
- ☐ Breadth-first search will perform badly
- ☐ None of the two algorithms will perform efficiently in this setting

### 3. For DFS and BFS search algorithms, imagine the situation where all paths are of similar length and all paths lead to a goal state

- ☐ Depth-first search will perform well X
- ☐ Depth-first search will perform badly
- ☐ Breadth-first search will perform well
- ☐ Both algorithms will perform efficiently in this setting

### 4. For DFS and BFS search algorithms, imagine the situation of a very high branching factor and long paths to the goal

- ☐ Both algorithms will perform efficiently in this setting
- ☐ Depth-first search will perform badly
- ☐ Breadth-first search will perform well
- ☐ Breadth-first search will perform badly X

### 5. An optimal solution to a relaxed (simplified) problem

- ☐ Is an admissible heuristic to the original problem X
- ☐ Is an admissible heuristic to the original problem if the heuristic is bounded
- ☐ Is an admissible heuristic to the original problem if the triangle inequality is fulfilled
- ☐ Is a consistent heuristic to the original problem if the algorithm is complete

**6. Which of the following is true for a multilayer perceptron using continuous activation functions?**

- ☐ It can exactly model linear functions only if it has a hidden layer with logistic sigmoidal activation functions
- ☐ It can exactly model any mathematical function
- ☐ It can exactly model any mathematical function except discrete functions
- ☐ It can exactly model linear functions using one hidden layer with hyperbolic tangent activation functions X

**7. For kernels in support vector machines, what is a correct statement?**

- ☐ A valid kernel must be positive-definite
- ☐ A valid kernel must be negative-definite
- ☐ A valid kernel must be positive semi-definite X
- ☐ A valid kernel must be negative semi-definite

**8. What statement is most true regarding a game such as a crossword puzzle?**

- ☐ Is an environment which is static and episodic
- ☐ Is an environment which is dynamic and episodic
- ☐ Is an environment which is static and sequential X
- ☐ Is an environment which is dynamic and sequential

**9. A reflex agent is**

- ☐ An example of a stochastic agent
- ☐ An agent that may have problems with a partially observable environment X
- ☐ An agent that always has a deterministic agent function
- ☐ An agent that is better in a discrete environment than a model-based agent

**10. An utility function is not used in game playing to**

- ☐ Estimate the expected utility of the game from a given position
- ☐ Compute the winner of the game based on number of pieces left X
- ☐ Compute the utility of a terminal leaf in a game tree
- ☐ Compute the best move for the successor function

**11. In order for the minimax algorithm to be computing optimal strategy, the game must be**

- ☐ Turn-based, episodic, partial-information, deterministic
- ☐ Turn-based, discrete, full-information, deterministic
- ☐ Episodic, discrete, full-information, stochastic
- ☐ Turn-based, zero-sum, full-information, deterministic X

## 12. A entails B means that:

- ☐ if A is true, then B has to be true as well X
- ☐ if B is true, then A has to be true as well
- ☐ A it true if and only if B is true
- ☐ if A is true, then B may also be true

## 13. A sentence is valid if

- ☐ it is in CNF form
- ☐ it does not contain any quantifiers
- ☐ it is true in some model
- ☐ it is true in all models X

## 14. The following is the truth table for:

- ☐ conjunction
- ☐ disjunction X
- ☐ implication
- ☐ exclusive-or

| | True | False |
|---|---|---|
| True | True | True |
| False | True | False |

## 15. Which of the following is a tautology:

- ☐ $(a \wedge b) \vee (b \wedge a)$
- ☐ $(a \vee b) \equiv (a \wedge b)$
- ☐ $\neg\neg a$
- ☐ $a \Rightarrow b \equiv \neg a \vee b$ X

## 16. Which of the following is not a definition of BN

- ☐ X is a Bayesian network with respect to G if it satisfies the local Markov property, i.e. each variable is conditionally independent of its descendants given its parent variables X
- ☐ X is a Bayesian network with respect to G if every node is conditionally independent of all other nodes in the network, given its Markov blanket
- ☐ X is a Bayesian network with respect to G if its joint probability density function can be written as a product of the individual density functions, conditional on their parent variables
- ☐ X is a Bayesian network with respect to G if, for any two nodes i and j, i and j are conditionally independent given the set of nodes that d-separate i and j

**17. Which of the following is Bayes theorem**

- ☐ $P(A|B) = \frac{P(B|A)P(B)}{P(A)}$
- ☐ $P(A|B) = \frac{P(A|B)P(A)}{P(B)}$
- ☐ $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$ X
- ☐ $P(B|A) = \frac{P(B|A)P(A)}{P(B)}$

**18. Each node in a Bayesian Network is a**

- ☐ joint posterior probability
- ☐ random variable
- ☐ a priori probability
- ☐ conditional probability distribution function X

**19. What is true accuracy of the following hypothesis "If you roll a 6-sided die twice, the sum will be larger than 3"**

- ☐ $\frac{1}{6}$
- ☐ $\frac{11}{12}$ X
- ☐ $\frac{31}{36}$
- ☐ $\frac{8}{9}$

**20. Which of the following types of reasoning is not an example of machine learning**

- ☐ supervised
- ☐ reinforced
- ☐ unsupervised
- ☐ deductive X