Försättsblad till skriftlig tentamen vid Linköpings universitet

Datum för tentamen	2018-01-05
Sal (2)	G33(1) TER1(70)
Tid	14-18
Kurskod	TDDC17
Provkod	TEN1
Kursnamn/benämning **Provnamn/benämning**	Artificiell intelligens En skriftlig tentamen
Institution	IDA
Antal uppgifter som ingår i tentamen	8
Jour/Kursansvarig Ange vem som besöker salen	Olov Andersson
Telefon under skrivtiden	070-5473343
Besöker salen ca klockan	ca kl. 15
Kursadministratör/kontaktperson (namn + tfnr + mailaddress)	Anna Grabska Eklund, anna.grabska.eklund@liu.se, ankn. 2362
Tillåtna hjälpmedel	Miniräknare/Hand calculators
Övrigt	
Antal exemplar i påsen	

Linköpings Universitet
Institutionen för Datavetenskap
Patrick Doherty

Tentamen
TDDC17 Artificial Intelligence
5 January 2018 kl. 14-18

Points:

The exam consists of exercises worth 37 points.
To pass the exam you need 22 points.

Auxiliary help items:

Hand calculators.

Directions:

You can answer the questions in English or Swedish.
Use notations and methods that have been discussed in the course.
In particular, use the definitions, notations and methods in appendices 1-3.
Make reasonable assumptions when an exercise has been under-specified.
State these assumptions explicitly in your answer.
Begin each exercise on a new page.
Write only on one side of the paper.
Write clearly and concisely.

Jourhavande: Olov Andersson, 070 547 33 43. Olov will arrive for questions around 15.00.

1. The following questions pertain to automated planning:

   (a) What characterizes *satisficing* planning in comparison to other common types of planning? Also, do satisficing planners typically require *admissible* heuristics? Motivate clearly why admissibility is, or is not, required in this context. **[2p]**

   (b) Informally, a *relaxation* to a planning problem "removes constraints" from the problem. The lectures (and lecture notes) provided a more formal definition of relaxation which is not based on how you change the planning problem but on what happens with the *solutions*. This allowed us to prove more directly that optimal solutions to a relaxed problem can be used to define admissible heuristics. **[3p (for i and ii below)]**

      i. Given two classical planning problems P and P', we know that P' is a relaxation of P if and only if ... (what is the correct criterion?)

      ii. Also provide a general example of how a relaxation could modify the state transition system of a problem in a way that satisfies the criterion above.

2. The following questions pertain to machine learning. Give *short and informative* answers.

   (a) Two major classes of learning algorithms are supervised learning and reinforcement learning. Contrast these in terms of what inputs are needed for *training*. **[2p]**

   (b) The Q-learning algorithm is often paired with an exploration strategy. Explain why. **[2p]**

   (c) This question pertains to deep learning. Assume you want to train a classifier from image inputs. Explain why a convolutional neural network would, or would not, be suitable for such a task? **[2p]**

3. The following questions pertain to Answer Set Programming. Appendix 3 may be useful to use:

   (a) Given the program Π_1, consisting of the following rules:

   r1: $a \leftarrow not\ b.$

   r2: $b \leftarrow not\ a.$

   r2: $q \leftarrow a.$

      1 what is the reduct Π_1^S for Π_1 given that $S = \{q\}$? **[1p]**

      2 what is the reduct Π_1^S for Π_1 given that $S = \{a\}$? **[1p]**

   (b) What are the *actual* answer sets for Π_1? Explain using the associated reducts for Π_1. **[1p]**

   (c) Let the program $\Pi_2 = \Pi_1 \cup \{\neg q \leftarrow a\}$.Given the *actual* answer sets for Π_1 above provided in your answer to (c), are any of the answer sets for program Π_1 also answer sets for program Π_2? Explain why or why not using reducts of Π_2 as part of your answer. **[2p]**

   (d) Why is Answer Set Programming considered to be a nonmonotonic reasoning formalism? **[1p]**

4. The following questions pertain to Bayesian Networks. Figure 1 below provides a Bayesian network for a satellite monitoring problem. Conditional tables for the problem are also provided.

   (a) Provide an equivalent equation for the joint distribution $\mathbf{P}(B, S, E, D, C)$, as a product of conditional distributions based on the independence assumptions associated with the Bayesian network in Figure 1. **[2p]**

   (b) Suppose the battery system does not fail, the solar panel does and there is trajectory deviation, but no communication loss: $P(\neg b, s, d, \neg c)$. What is the probability that this can happen? **[1p]**

   (c) Suppose the battery system fails and the solar panel system does not. What is the probability of communication loss: $P(c \mid b, \neg s)$? **[2p]**

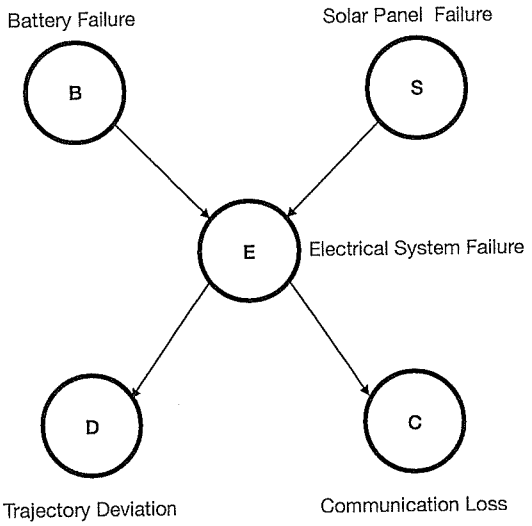

Figure 1: Bayesian Network Example

S	P(S)
T	0.02

B	P(B)
T	0.05

B	S	$P(E)$
T	T	0.99
T	F	0.87
F	T	0.85
F	F	0.10

E	$P(D)$
T	0.99
F	0.10

E	$P(C)$
T	0.99
F	0.15

5. A* search is the most widely-known form of best-first search. The following questions pertain to A* search:

   (a) Explain what an *admissible* heuristic function is using the notation and descriptions in (c). **[1p]**

   (b) Can a *consistent* heuristic function be inadmissible? Explain why or why not. **[1p]**

   (c) Let $h(n)$ be the estimated cost of the cheapest path from a node n to the goal. Let $g(n)$ be the path cost from the start node n_0 to n. Let $f(n) = g(n) + h(n)$ be the estimated cost of the cheapest solution through n.

   Provide a sufficiently rigid proof that A* is optimal if $h(n)$ is admissible. You need only provide a proof for either tree-search (seminar slides) or graph-search (in course book). If possible, use a diagram to structure the contents of the proof to make it more readable. **[2p]**

6. The following questions pertain to the course article by Newell and Simon entitled *Computer Science as an Empirical Enquiry: Symbols and Search.*

   (a) What is a *physical symbol system* (PSS) and what are its structural and conceptual components? **[2p]**

   (b) What is the Physical Symbol System Hypothesis? **[1p]**

   (c) Do you think the Physical Symbol System Hypothesis provides an adequate description of the structural and conceptual components required for a system exhibiting intelligence? Provide reasonable justifications for your opinion. **[1p]**

   (d) What is the heuristic search hypothesis? **[1p]**

7. Constraint satisfaction (CS) problems consist of a set of variables, a value domain for each variable and a set of constraints. A solution to a CS problem is a consistent set of bindings to the variables that satisfy the constraints.

   (a) Suppose there are 5 territories T1, T2, T3, T4, and T5, each with a sensor that monitors the area associated with that territory. Each sensor has three possible radio frequencies, F1,F2, F3. Sensors overlap if they are in adjacent areas. The adjacency relation between two territories is symmetric. Let $Adj(x, y)$ represent the adjacency relation where $Adj(T1, T2)$, $Adj(T1, T3)$,$Adj(T2, T3)$, $Adj(T3, T4)$. If two sensors overlap, they can not use the same frequency.

      1. Define a constraint satisfaction problem for this scenario. **[1p]**
      2. Provide a constraint graph for the CS problem. **[1p]**
      3. Provide one solution for the CS problem. **[1p]**

8. Consider the following logical theory (where a, b and 27, 28 are constants) and we view grounded atomic formulas as propositional atoms. (In this case unification of two grounded atomic formulas is successful when they are identical):

$$(Package(\mathsf{b}) \wedge Package(\mathsf{a}) \wedge Inroom(\mathsf{b}, 27) \wedge Inroom(\mathsf{a}, 28)) \Rightarrow Smaller(\mathsf{b}, \mathsf{a}) \quad (1)$$

$$Package(\mathsf{a}) \quad (2)$$

$$Package(\mathsf{b}) \quad (3)$$

$$Inroom(\mathsf{a}, 27) \vee Inroom(\mathsf{a}, 28) \quad (4)$$

$$Inroom(\mathsf{b}, 27) \quad (5)$$

$$\neg Smaller(\mathsf{b}, \mathsf{a}) \quad (6)$$

   We would like to show using resolution that package $Inroom(\mathsf{a}, 27)$. To do this, answer the following questions:

   (a) Convert formulas (1) - (6) into conjunctive normal form (CNF) with the help of appendix 1. **[1p]**

   (b) Prove that $Inroom(\mathsf{a}, 27)$ is a logical consequence of (1) - (6) using the resolution proof procedure. **[2p]**

   - Your answer should be structured using resolution refutation trees (as used in the book or course slides).

Appendix 1

Converting arbitrary wffs to CNF form: (Propositional/grounded 1st-order formula case)

1. Eliminate implication signs using the equivalence: $\alpha \rightarrow \beta \equiv \neg\alpha \vee \beta$.
2. Reduce scopes of negation signs using De Morgan's Laws:
   - $\neg(\omega_1 \vee \omega_2) \equiv \neg\omega_1 \wedge \neg\omega_2$
   - $\neg(\omega_1 \wedge \omega_2) \equiv \neg\omega_1 \vee \neg\omega_2$
3. Remove double negations using the equivalence: $\neg\neg\alpha \equiv \alpha$.
4. Put the remaining formula into conjunctive normal form. Two useful rules are:
   - $\omega_1 \vee (\omega_2 \wedge \omega_3) \equiv (\omega_1 \vee \omega_2) \wedge (\omega_1 \vee \omega_3)$
   - $\omega_1 \wedge (\omega_2 \vee \omega_3) \equiv (\omega_1 \wedge \omega_2) \vee (\omega_1 \wedge \omega_3)$
5. Eliminate $\wedge$ symbols so only clauses remain.

Appendix 2

A generic entry in a joint probability distribution is the probability of a conjunction of particular assignments to each variable, such as $P(X_1 = x_1 \wedge \ldots \wedge X_n = x_n)$. The notation $P(x_1, \ldots, x_n)$ can be used as an abbreviation for this.

The chain rule states that any entry in the full joint distribution can be represented as a product of conditional probabilities:

$$P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid x_{i-1}, \ldots, x_1) \tag{7}$$

Given the independence assumptions implicit in a Bayesian network a more efficient representation of entries in the full joint distribution may be defined as

$$P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid \mathit{parents}(X_i)), \tag{8}$$

where $\mathit{parents}(X_i)$ denotes the specific values of the variables in $\mathit{Parents}(X_i)$.

Recall the following definition of a conditional probability:

$$\mathbf{P}(X \mid Y) = \frac{\mathbf{P}(X \wedge Y)}{\mathbf{P}(Y)} \tag{9}$$

The following is a useful general inference procedure:

Let X be the query variable, let $\mathbf{E}$ be the set of evidence variables, let $\mathbf{e}$ be the observed values for them, let $\mathbf{Y}$ be the remaining unobserved variables and let α be the normalization constant:

$$\mathbf{P}(X \mid \mathbf{e}) = \alpha \mathbf{P}(X, \mathbf{e}) = \alpha \sum_{\mathbf{y}} \mathbf{P}(X, \mathbf{e}, \mathbf{y}) \tag{10}$$

where the summation is over all possible $\mathbf{y}$'s (i.e. all possible combinations of values of the unobserved variables $\mathbf{Y}$).

Equivalently, without the normalization constant:

$$\mathbf{P}(X \mid \mathbf{e}) = \frac{\mathbf{P}(X, \mathbf{e})}{\mathbf{P}(\mathbf{e})} = \frac{\sum_{\mathbf{y}} \mathbf{P}(X, \mathbf{e}, \mathbf{y})}{\sum_{\mathbf{x}} \sum_{\mathbf{y}} \mathbf{P}(\mathbf{x}, \mathbf{e}, \mathbf{y})} \tag{11}$$

Appendix 3: Answer Set Programming

Computing Answer Sets for a program Π:
Given a program Π:

1. Compute the possible answer sets for Π:
   (a) Powerset 2^{Π} of all atoms in the heads of rules in Π.
2. For each $S \in 2^{\Pi}$:
   (a) Compute the reduct Π^S of Π.
   (b) If $Cn(\Pi^S) = S$ then S is an answer set for Π.
   (c) If $Cn(\Pi^S) \neq S$ then S is not an answer set for Π.

The following definitions may be useful:

Definition 1 A program Π consists of a signature Σ and a collection of rules of the form:

$$l_0 \vee, \ldots, \vee l_i \leftarrow l_{i+1}, \ldots, l_m, not\ l_{m+1}, \ldots, not\ l_n$$

where the l's are literals in Σ. □

Definition 2 [Satisfiability]
A set of (ground) literals satisfies:

1. l **if** $l \in S$;
2. *not* l **if** $l \notin S$;
3. $l_1 \vee \ldots l_n$ **if** *for some* $1 \leq i \leq n, l_i \in S$;
4. a set of (ground) extended literals **if** S satisfies every element of this set;
5. rule r **if** , whenever S satisfies r's body, it satisfies r's head. □

Definition 3 [Answer Sets, Part I]
Let Π be a program not containing default negation (i.e., consisting of rules of the form):

$$l_0 \vee, \ldots, \vee l_i \leftarrow l_{i+1}, \ldots, l_m.$$

An *answer set* of Π is a consistent set S of (ground) literals such that

1. S satisfies the rules of Π and
2. S is minimal (i.e., there is no proper subset of S that satisfies the rules of Π. □

Appendix 3 is continued on the next page.

Definition 4 [Answer Sets, Part II]
Let Π be an arbitrary program and S be a set of ground literals. By Π^S we denote the program obtained from Π by

1. removing all rules containing *not* l such that $l \in S$;
2. removing all other premises of the remaining rules containing *not*.

S is an answer set of Π if S is an answer set of Π^S.
We refer to Π^S as the *reduct* of Π with respect to S. □

Definition 5 [Consequence operator T_Π]
The smallest model, $Cn(\Pi)$, of a positive program Π can be computed via its associated *consequence operator* T_Π. For a set of atoms X we define,

$$T_\Pi X = \{head(r) \mid r \in \Pi \text{ and } body(r) \subseteq X\}.$$

Iterated applications of T_Π are written as T_Π^j for $j \geq 0$, where

$$T_\Pi^0 X = X$$
$$T_\Pi^i X = T_\Pi T_\Pi^{i-1} X \text{ for } i \geq 1.$$

For any positive program Π, we have $Cn(\Pi) = \bigcup_{i \geq 0} T_\Pi^i \emptyset$.
Since T_Π is monotonic, $Cn(\Pi)$ is the smallest fixpoint of T_Π. □