AI exam, 2014. 12. 17. NAME:
TOTAL POINTS:60 (minimum to pass: 24) NEPTUN ID:

1. Knowledge is power:
   a. What is the purpose of a heuristic function?
   b. Define the admissibility and dominance of a heuristic function.
   c. Assume that there are three heuristic functions, none of them is admissible, but we know that at each node only one of them estimates the future cost larger than the real future cost. How could we build an admissible heuristics?
   d. Along a path to goal G the following costs and an admissible heuristic function *h* is available. Could we improve it?

Step i	1	2	3	4	5	6	7	goal G
Cost of step i	1	2	3	2	1	4	1	
cost to the goal G	14	13	11	8	6	5	1	0
h	7	11	8	6	5	4	1	
improved h								

(10 points)

2. What would you prefer: "a sound, but not complete" or "a complete, but not sound" inference engine? Explain why and give an example for "a sound, but not complete" inference method. (5 points)

3. The DARII **syllogism** is as follows:

- $\forall x.\ B(x) \rightarrow A(x)$
- $\exists x.\ C(x) \wedge B(x)$
- $\exists x.\ C(x) \wedge A(x)$

Prove it with resolution! (10 points)

4. Consider the naive Bayesian network in Fig.2 with discrete variables (number of values are indicated in parentheses). Explain the semantics of the structure (the implied independencies). Indicate the necessary parameters and construct the corresponding joint distribution in a factorized format according to the structure of this Bayesian network. Demonstrate the application of the Bayes rule in computing P(D|E,F,G) and explain its advantage. (10 p)

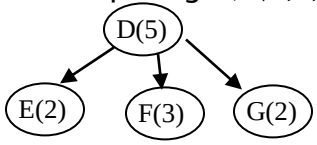

Fig.1. Naive Bayesian network with discrete variables (number of values are indicated in parentheses).

5. Define a decision tree for the following expression: $(A \rightarrow \neg B) \rightarrow (C \rightarrow B)$. Prove that any Boolean expression can be represented by a decision tree (for example, give a constructive proof by defining the main steps of such an algorithm). (10 points)

6. Derive the full Bayesian learning formula (in predicting the next observation from the earlier observations). How could we accommodate the „Ockham's razor" principle within this framework? (5 p)