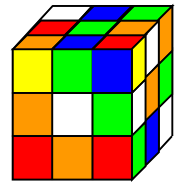
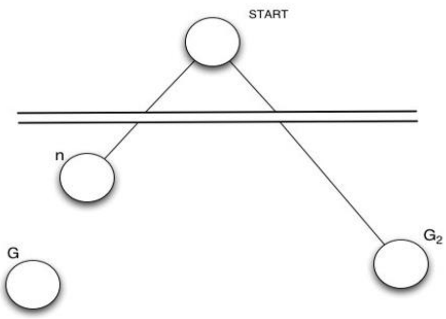
1. Define admissibility and suggest an admissible heuristic functions for the Rubik's cube. (half-twist is counted as a single move). (5p)



2. Prove that A^* never terminates with the suboptimal goal G_2 (see figure). Use the following notation $f(n)=g(n)+h(n)$.



(10p)

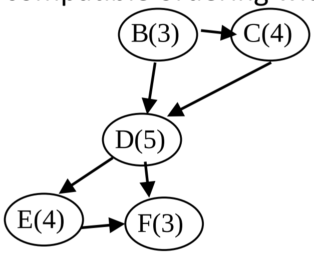
3, What does it mean that propositional logic is compositional (truth-functional)? Demonstrate it evaluating the expression $(A \rightarrow C) \rightarrow (\neg B \rightarrow \neg A)$ for A=TRUE, B=FALSE, C=TRUE. Give a natural language example (for example a story, sentence, operator) where meaning is not compositional.
(5 points)

4. The **FESTIMO syllogism** is as follows:

$$\forall x.\ B(x) \rightarrow \neg A(x)$$
$$\underline{\exists x.\ C(x) \wedge A(x)}$$
$$\exists x.\ C(x) \wedge \neg B(x)$$

Prove it with resolution.
(15 points)

5. What is the number of free parameters in the Bayesian network below assuming that all variables are discrete, the number of values are indicated in parentheses, and we use table models to specify the conditional probabilites. What is the gain compared to a naive table representation? Give an example for the specification of P(F|E,D) using a decision tree. Decompose the joint distribution **P(B,C,D,E,F)** using a compatible ordering with the graph, then simplify it by accepting the assumptions implied by the graph!



(15 points)

6, Define formally the Maximum Expected Utility principle. Select the best action in the following case:

Action	Probabilities of outcomes	Utilities of outcomes
1	0.3, 0.3,0.4	3, 2, 10
2	0.7, 0.2, 0.1	2, 9, 40

(5 points)

6, Define and explain the concept of probably approximately correct learning, the corresponding sample complexity, and its inequality.
(10 points)

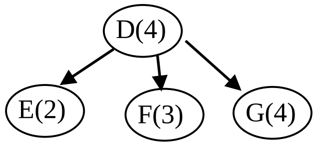
1. In case of an ideal, admissible heuristic function in A* search, what is the number of expanded nodes ? Give a definition for the expanded nodes. (5 p)

2. What is the type of the next statement: valid, satisfiable, not satisfiable, none of these. Prove your answer with truth tables. (10 points)

$$(A \rightarrow \neg B) \rightarrow \neg (C \rightarrow B).$$

3. Define completeness and soundness in theorem proving. Give an example. (5p)

4. Consider the following naive Bayesian network with discrete variables(number of values are indicated in parenthesis). Assuming that the edges are necessary and sufficient, construct a Bayesian network structure using the following ordering of the variables: E, F, G, D (meaning that arrows are allowed only from earlier to later variables). Compare the number of parameters used in the networks. (10 p)



5. Define a decision tree with PopularDog as an outcome variable for the following observations:
Table 2.

CaseID.	Small	White	Barking	PopularDog
X1	Yes	Yes	Yes	Yes
X2	Yes	Yes	No	Yes
X3	No	No	Yes	Yes
X4	No	No	No	No
X5	No	Yes	No	No
X6	Yes	No	No	Yes

in which Small, White, and Barking are the predictor variables. Define a Booleanl expression for PopularDog using these variables both in a conjunctive and disjunctive normal forms. What is the cardinality of decision trees (as logical functions) over *n* binary attributes?
(15 points)

6. In a diagnostic problem three devices can be faulty A, B, C (A,B,C denote respectively that A, B, C is faulty). The devices cannot be investigated separately, but we know the following facts about their system (knowledge base KB).

S1: The system is fault tolerant, which means that it works with one faulty device, but it breaks down with two or more faulty devices.

S2: A and B cannot be faulty at the same time.

S3:The system breaks down.

a) Define the statements S1, S2, S3 using the propositions A, B, C.
b) Convert the KB to conjunctive normal form (CNF).
c) Indicate the models of the knowledge base.
d) Show with truth-table that the knowledge base KB entails that C is faulty.
e) Prove that C is faulty with resolution assuming the knowledge base KB.

(15p)

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)**
- <u>$\exists$**x. C(x) $\wedge$ B(x)**</u>
- $\exists$**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)

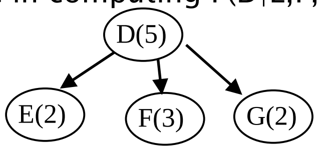


Fig.1. Naive Bayesian network with discrete variables (number of values are indicated in parentheses).

5. Define a decision tree for the following expression: $(A \rightarrow \neg B) \rightarrow (C \rightarrow B)$. (5 points)

2. What are the main differences between propositional and first-order logic? (consider the syntactic, semantic and practical aspects as well) (5 points)

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

$\forall$**x. B(x) → A(x)**
$\forall$**x. C(x) → ¬A(x)**
$\forall$**x. C(x) → ¬B(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)

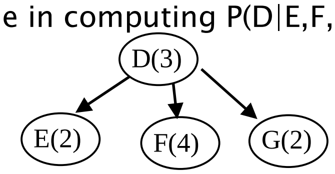
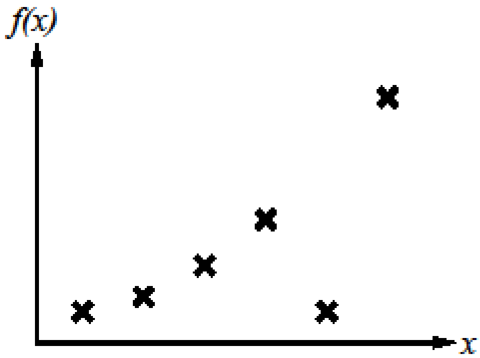


Fig.2. Naive Bayesian network with discrete variables (number of values are indicated in parentheses).

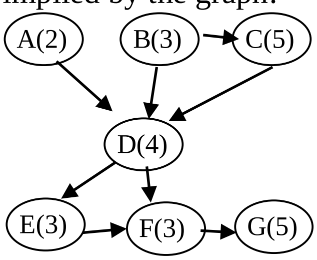
5. Define a decision tree for the OR and for the parity functions over three variables A, B, C, that is for the expression „$A \vee B \vee C$" and for the expression „A XOR B XOR C".Prove that any Boolean expression can be represented by a decision tree (for example, give a constructive proof by defining the main steps of such an algorithm). (10 points)

6. What is the „Ockham's razor" principle? Explain its application in the following task, which seeks to model the relation between *x* and *f(x)=y*.



(10 p)

4. What is the number of free parameters in the Bayesian network below assuming that all variables are discrete, the number of values are indicated in parentheses, and we use table models to specify the conditional probabilites. What is the gain compared to a naive table representation? Decompose the joint distribution **P(A,B,C,D,E,F,G)** using a compatible ordering with the graph, then simplify it by accepting the assumptions implied by the graph!



10p

5. Adam, Betty, and Chris played and a window got broken. Adam says: ‘Betty made, Chris is innocent.' Betty says: ‘If Adam is guilty, then Chris too'. Chris says: ‘I am innocent; someone else did it'.

Is it possible that none of them is lying? If they are telling the truth, who broke the window? Answer these questions with truth table method using the following notation:

- A: Adam is innocent.
- B: Betty is innocent.
- C: Chris is innocent. (10p)

(10 points)