

Date: \_\_\_\_\_

Name: \_\_\_\_\_

**Note: In the “check correct answer” questions, more than one answer may be correct!****1. (1 pt) HTN planning uses the principles of:**

- Planning using model checking.
- Planning by solving a SAT problem.
- Planning using decomposition of abstract operators.
- Planning using MDP.

**2. (1 pt) Agent function maps**

- current perceptions of agent to agent’s action.
- agent’s percept sequence to agent’s action.
- current perceptions of agent to its internal state.
- agent’s percept sequence to its internal state.

**3. (2 pts) Describe the components of conceptual planning model****4. (3 pts) Sort the following lotteries using the preference relation ‘<’:**

- a) [0.3, A; 0.3, B; 0.4, C]
- b) [0.1, A; 0.2, B; 0.7, C]
- c) [0.2, A; 0.2, B; 0.6, C]

Assume that agent’s preferences of individual results are sorted as  $C < A < B$ , and the utility function  $U(C) > 0$ .

**5. (7 pts) Rational tickets purchase:**

I want to buy tickets for 2 theater performances in National Theater. I have two options: (a) I can buy both tickets in advance for a decreased price (so called “combined ticket”), or (b) I can buy each ticket right before the performance (so called “single tickets”). The probability that I will eventually have the time to go to the theater is 0.3 for both performances. Single ticket costs 200 CZK for each performance, combined ticket costs 350 CZK for both performances together. Seeing each performance is worth 200 CZK for me. Should I buy combined ticket or single tickets?

6. (2 pts) Define STRIPS planning domain.

7. (7 pts) Labyrinth

Agent lives in an environment with 5 states

	s12	s13
s21	s22	s23

with the following rewards in individual states

	-2	+10
2	-2	-10

Agent can move in all 4 main directions, but its movement is not deterministic: it moves in the required state with the probability 0.8, with probability 0.2 it will move in one of the 2 perpendicular directions (0.1 and 0.1). If it hits the wall, it will stay in place. States s13 and s23 are absorbing, after reaching them, agent does not get any more rewards.

Assume that after several steps of value iteration, the value function looks like this:

	9	+10
-4	-5	-10

Questions:

- for non-absorbing states s21, s12 and s22 determine the current optimal actions from the set {UP, DOWN, LEFT, RIGHT}.
- for state s12, compute the new value after a single step of value iteration. Use a discount factor 0.9.

8. (2 pts) By choosing from the available options, describe the properties of MDP and RL, respectively (fill a character to all 8 cells of table).

	MDP	RL
reward function is known/unknown (K/U)		
environment transition function is known/unknown (K/U)		
environment transition function is deterministic/stochastic (D/S)		
optimal strategy (policy) is deterministic/stochastic (D/S)		

9. (3 pts) Define the main components of Belief-Desire-Intention (BDI) architecture

10. (2 pts) What is a reactive planner? Describe also its control loop.