Entrance test - B4M36SMU (sample solution)

1 Do not play with the matches

There are 4 matches lying on a table. The goal of an automated robot is to gradually remove them such that there is no match remaining on the table. The robot can remove 1 or 2 matches in one step. The problem is that the robot's arm is unreliable, it can remove more matches than the robot planned. To be precise, in half the attempts the arm removes one more match than planned. If the robot tries to remove more matches than actually available, the task becomes cyclic (-1 turns into 4, -2 turns into 3). Propose the optimal robot control strategy, the goal is to minimize the number of steps.

(a) Propose a task formalization based on Markov Decision Process (MDP).

(b) Formally derive the optimal strategy. If the derivation turns out difficult, show a few steps only and define the termination conditions.

(c) Use the derivation ad b and for each state select one out of two available actions.

(d) How many steps the robot with the optimal control strategy needs to reach zero matches?

Solution

(a) Define set of states as the number of matches, i.e. $\{0, 1, 2, 3, 4\}$. Let -1 be the penalty for any action and set discount factor $\gamma = 1$. Actions will be $\{1, 2\}$ and they represent the number of matches that the robot intends to remove.

State 0 is terminal and there are no actions available. For any other state s and action a we get to state $s - a \mod 5$ with probability $\frac{1}{2}$ and to state $s - a - 1 \mod 5$ with probability $\frac{1}{2}$.

There are other valid possibilities how to choose penalties and discount factor, however this choice allows us to find solution to the point d easily.

(b) For each state we set initial value $V_0(s) = 0$. Now we modify the initial estimate to obtain a better solution. We use the Bellman equation

$$V_{k+1}(s) = \max_a \sum_{s'} P(s' \mid s, a) \cdot (R(s, a, s') + \gamma V_k(s')),$$

which is in our case (for $s \neq 0$)

$$V_{k+1}(s) = -1 + \max_{a \in \{1,2\}} \left\{ \frac{V_k(s - a \mod 5) + V_k(s - a - 1 \mod 5)}{2} \right\}. \tag{1}$$

By applying this equation we get $V_k(s)$

s	$V_0(s)$	$V_1(s)$	$V_2(s)$	$V_3(s)$	$V_4(s)$	$\cdots$	$V_\infty(s)$
4	0	-1	-2	$-\frac{5}{2}$	$-\frac{11}{4}$	$\cdots$	$-\frac{10}{3}$
3	0	-1	$-\frac{3}{2}$	$-\frac{7}{4}$	-2	$\cdots$	$-\frac{7}{3}$
2	0	-1	$-\frac{3}{2}$	$-\frac{7}{4}$	-2	$\cdots$	$-\frac{7}{3}$
1	0	-1	$-\frac{3}{2}$	-2	$-\frac{9}{4}$	$\cdots$	$-\frac{8}{3}$
0	0	0	0	0	0	0	0

(c) At $V_3(s)$ we encounter the optimal policy for the first time. From equation (1) after evaluation of $\max_{a \in \{1,2\}}$ we see, which action leads to higher reward.

The optimal strategy consists of removing 1 match for all states but state 3, where the robots tries to remove 2 matches.

(d) From $V_\infty(s)$ we see that the robot needs $\frac{10}{3}$ actions on average to remove all matches. Remember that cost for each action was -1.

Recommended literature MDPs were part of A4B33ZUI course. Visit lecture `https://cw.fel.cvut.cz/wiki/_media/courses/a4b33zui/mdps_show.pdf`. For more detailed review you can visit Chapter 17 of AIMA book.

2 CNF

Convert to CNF:

$$\exists x \forall y \forall z \left(\text{person}(x) \wedge ((\text{likes}(x, y) \wedge \neg\text{equal}(y, z)) \Rightarrow \neg\text{likes}(x, z))\right).$$

Solution First eliminate the implication:

$$\exists x \forall y \forall z \left(\text{person}(x) \wedge (\neg\left(\text{likes}(x, y) \wedge \neg\text{equal}(y, z)\right) \vee \neg\text{likes}(x, z))\right).$$

Now use de Morgan's law to move the negations inwards:

$$\exists x \forall y \forall z \left(\text{person}(x) \wedge (\neg\text{likes}(x, y) \vee \text{equal}(y, z) \vee \neg\text{likes}(x, z))\right).$$

Instead of existential variable x we use a *Skolem constant* X to obtain:

$$\forall y \forall z \left(\text{person}(X) \wedge (\neg\text{likes}(X, y) \vee \text{equal}(y, z) \vee \neg\text{likes}(X, z))\right).$$

As a last step we can drop universal quantifiers:

$$\text{person}(X) \wedge (\neg\text{likes}(X, y) \vee \text{equal}(y, z) \vee \neg\text{likes}(X, z)).$$

Recommended literature A guide how to convert FOL sentence to CNF is available on `https://april.eecs.umich.edu/courses/eecs492_w10/wiki/images/6/6b/CNF_conversion.pdf`. For more detailed guide consult with Section 2.5 of `http://math.feld.cvut.cz/demlova/teaching/lgr/text_lgr_2015.pdf`.

3 Bayesian Decision Making

Imagine you are a shepherd, whose sheep are constantly attacked by wild beasts. You care about your beloved flock and recall the machine learning course you took at the farming university during childhood.

Three species of predators live in the surrounding forests: bear, wolf and lynx. You know from experience that – from eight cases there are four predator attacks by wolves, three attacks by bears and one by lynx. The wolves are known to attack in pack and usually kill more sheep at once (two or more sheep are killed in 8 out of 10 attacks). Bears are solitary, but may also kill more sheep (two or more sheep are killed in 5 of 10 attacks). Lynx always kills one sheep. Wolves never return to the same place for a new attack next night, a lynx returns only in one out of ten cases, bear returns in half of cases. Assume that the event of predator returning to the same place and the number of sheep eaten are independent.

Answer the following questions:

(a) What is the probability of an attack, in which a lynx killed more than one sheep?

For the following questions, suppose exactly one of your sheep was killed by a wild animal during the night.

(b) What is the probability that sheep was attacked by bear?

(c) Which predator killed the sheep? Decide merely on the basis of maximum likelihood (with the assumption of uniform a priori probability). Explain the decision.

(d) Which predator killed the sheep? Use the method of maximum a posterior probability. Explain the decision.

(e) What is the probability that a predator will return the next night? Assume the MAP hypothesis about the killer from the previous question and provide a simplified approximation.

(f) What is the probability that a predator will return the next night? Consider full Bayesian learning.

Solution Define the following events:

B	bear attacked
W	wolf attacked
L	lynx attacked
O	one sheep was killed
R	the attacker returns

(a) Lynx always kills one sheep, therefore the probability is 0.

(b) We use Bayes formula to evaluate the probability $P(B \mid O)$

$$P(B \mid O) = \frac{P(O, B)}{P(O)} = \frac{P(O \mid B)P(B)}{P(O)} = \frac{P(O \mid B)P(B)}{P(O \mid B)P(B) + P(O \mid W)P(W) + P(O \mid L)P(L)}$$

$$= \frac{\frac{3}{8} \times \frac{1}{2}}{\frac{3}{8} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{5} + \frac{1}{8} \times 1} = \frac{5}{11} = 45\,\%.$$

(c) To get MLE we need to find an attacker $X \in \{B, W, L\}$ that maximizes the likelihood $P(O \mid X)$.

$$
\begin{aligned}
P(O \mid B) &= \frac{1}{2}, \\
P(O \mid W) &= \frac{1}{5}, \\
P(O \mid L) &= 1.
\end{aligned}
$$

We see that *lynx* attacked the sheep.

(d) To get MAP we need to find an attacker $X \in \{B, W, L\}$ that maximizes $P(X \mid O)$. Similarly to b we calculate probability of wolf and lynx being the attacker

$$
\begin{aligned}
P(W \mid O) &= \frac{8}{33}, \\
P(L \mid O) &= \frac{10}{33}.
\end{aligned}
$$

We compare the probabilities to see that *bear* killed the sheep.

(e) If bear attacked last night, then it will return with probability $P(R \mid B) = 0.5$.

(f) If we consider full Bayesian learning, we need to consider all three animals being the attacker.

$$
\begin{aligned}
P(R \mid O) &= \sum_{X \in \{B,W,L\}} P(R \mid X) \cdot P(X \mid O) = P(R \mid B)P(B \mid O) + P(R \mid W)P(W \mid O) + P(R \mid L)P(L \mid O) \\
&= \frac{1}{2}\frac{5}{11} + \frac{1}{10}\frac{8}{33} + 0\frac{10}{33} = \frac{17}{66} = 26\,\%.
\end{aligned}
$$

Recommended literature Review statistics course (A1B01PSI), namely the conditional probability in Section 3 of `http://cmp.felk.cvut.cz/~navara/psi/PMS_print.pdf`.

4 Resolution principle

Using resolution, prove from

$$\forall x \exists y : human(x) \rightarrow mother(x, y) \wedge human(y)$$

that

$$\forall x : human(x) \rightarrow mother(\texttt{mother\_of}(x), \texttt{mother\_of}(\texttt{mother\_of}(x)))$$

i.e. that every human has a grandmother.

Solution First convert to CNF. Skolemization leads to

$$\forall x : human(x) \rightarrow mother(x, \texttt{mother\_of}(x)) \wedge human(\texttt{mother\_of}(x))$$

Now we eliminate the implication to obtain

$$\forall x : \neg human(x) \vee (mother(x, \texttt{mother\_of}(x)) \wedge human(\texttt{mother\_of}(x)))\,.$$

In clausal form we obtain 2 clauses. Rename variables in one to prevent variable clash

$$\neg human(x_1) \vee mother(x_1, \texttt{mother\_of}(x_1)), \tag{2}$$
$$\neg human(x_2) \vee human(\texttt{mother\_of}(x_2)). \tag{3}$$

Unify

$$x_1 = \texttt{mother\_of}(x_2).$$

(2) and (3) resolve towards

$$\neg human(x_2) \vee mother(\texttt{mother\_of}(x_2), \texttt{mother\_of}(\texttt{mother\_of}(x_2)))$$

i.e

$$human(x) \rightarrow mother(\texttt{mother\_of}(x), \texttt{mother\_of}(\texttt{mother\_of}(x))).$$

Recommended literature A guide how to convert FOL sentence to CNF is available on `https://april.eecs.umich.edu/courses/eecs492_w10/wiki/images/6/6b/CNF_conversion.pdf`. For more detailed guide consult with Section 2.5 of `http://math.feld.cvut.cz/demlova/teaching/lgr/text_lgr_2015.pdf`.