



The University of Vermont

Assignment №1: Ch 1 (RL textbook) and Ch 13 (AI textbook)

Ayat Ospanov

September 4, 2018

Contents

| | | |
|----------|----------------------------|----------|
| 1 | Ch 1 (RL textbook) | 1 |
| 1.1 | Exercise 1.1 | 1 |
| 1.2 | Exercise 1.2 | 1 |
| 1.3 | Exercise 1.3 | 2 |
| 1.4 | Exercise 1.4 | 2 |
| 2 | Ch 13 (AI textbook) | 2 |
| 2.1 | Exercise 13.8 | 2 |
| 2.2 | Exercise 13.10 | 3 |
| 2.3 | Exercise 13.13 | 4 |
| 2.4 | Exercise 13.15 | 4 |
| 2.5 | Exercise 13.17 | 5 |
| 2.6 | Exercise 13.21 | 5 |

1 Ch 1 (RL textbook)

1.1 Exercise 1.1

If the algorithms plays against an opponent and as the opponents strategy is fixed, it will only learn to beat it. But if it plays against itself, each step it learns and its strategy changes over time. So it learns better and most probably reaches the optimal strategy. Given tic-tac-toe, to reach the strongest strategy, it should make the first step randomly over games. Otherwise, it only learns how to play for the one side.

1.2 Exercise 1.2

If we take advantage of symmetries, it would increase algorithms speed of learning and decreases memory it uses. But if the opponent did not take advantage of symmetries, it would destroy our game. For example, if the opponent makes a mistake in some of the symmetric situations, but not in others, we might take a wrong action assuming the opponent plays equally in all symmetric cases.

1.3 Exercise 1.3

Only playing a greedy strategy will result in a worse game overall with a high probability. If the algorithm is greedy, it ignores the overall value, and considers only an instant reward. In long run, it will not result in a high value, because there could be a higher reward in the future if it took another action in the past.

1.4 Exercise 1.4

When we learn from exploratory moves, we have a set of probabilities which are expected values of each state with exploration steps. When we don't learn from exploratory moves, we only have values of states given the optimal action to be taken. Learning from exploratory moves reduces a variance from future optimal states, increasing the overall gain. That means winning now, doesn't result in winning in the most of games. So exploratory moves result in more wins.

2 Ch 13 (AI textbook)

2.1 Exercise 13.8

a: $P(\textit{toothache})$

$$\begin{aligned} P(\textit{toothache}) &= \sum_{x \in \textit{Cavity}, y \in \textit{Catch}} P(\textit{toothache}, x, y) = \\ &= 0.108 + 0.012 + 0.016 + 0.064 = 0.2 \end{aligned}$$

b: $P(\textit{Cavity})$

$$\begin{aligned} P(\textit{cavity}) &= 0.108 + 0.012 + 0.072 + 0.008 = 0.2 \\ P(\textit{Cavity}) &= [P(\textit{cavity}), P(\neg \textit{cavity})] = [P(\textit{cavity}), 1 - P(\textit{cavity})] = [0.2, 0.8] \end{aligned}$$

c: $P(\textit{Toothache}|\textit{cavity})$

$$\begin{aligned} P(\textit{toothache}|\textit{cavity}) &= \frac{P(\textit{toothache}, \textit{cavity})}{P(\textit{cavity})} = \frac{0.108 + 0.012}{0.2} = 0.6 \\ P(\textit{Toothache}|\textit{cavity}) &= [P(\textit{toothache}|\textit{cavity}), P(\neg \textit{toothache}|\textit{cavity})] = [0.6, 0.4] \end{aligned}$$

d: $P(\text{Cavity}|\text{toothache} \vee \text{catch})$

$$P(\text{Cavity}|\text{toothache} \vee \text{catch}) = \frac{P(\text{Cavity}, \text{toothache} \vee \text{catch})}{P(\text{toothache} \vee \text{catch})}$$

$$P(\text{Cavity}, \text{toothache} \vee \text{catch}) = [P(\text{cavity}, \text{toothache} \vee \text{catch}), P(\neg \text{cavity}, \text{toothache} \vee \text{catch})]$$

$$P(\text{cavity}, \text{toothache} \vee \text{catch}) = 0.108 + 0.012 + 0.072 = 0.192$$

$$P(\neg \text{cavity}, \text{toothache} \vee \text{catch}) = 0.016 + 0.064 + 0.144 = 0.224$$

$$\begin{aligned} P(\text{toothache} \vee \text{catch}) &= P(\text{cavity}, \text{toothache} \vee \text{catch}) + P(\neg \text{cavity}, \text{toothache} \vee \text{catch}) = \\ &= 0.192 + 0.224 = 0.416 \end{aligned}$$

$$P(\text{Cavity}|\text{toothache} \vee \text{catch}) = [0.192/0.416, 0.224/0.416] = [0.462, 0.538]$$

2.2 Exercise 13.10

a:

$$\begin{aligned} P(\text{BAR}/\text{BAR}/\text{BAR}) &= P(\text{BELL}/\text{BELL}/\text{BELL}) = \\ &= P(\text{LEMON}/\text{LEMON}/\text{LEMON}) = P(\text{CHERRY}/\text{CHERRY}/\text{CHERRY}) = \\ &= \frac{1}{4^3} = \frac{1}{64} \end{aligned}$$

$$\begin{aligned} P(\text{CHERRY}/\text{CHERRY}/?) &= P(\text{CHERRY}/\text{CHERRY}) - \\ &\quad - P(\text{CHERRY}/\text{CHERRY}/\text{CHERRY}) = \frac{1}{4^2} - \frac{1}{4^3} = \frac{3}{64} \\ P(\text{CHERRY}/?) &= P(\text{CHERRY}) - P(\text{CHERRY}/\text{CHERRY}) - \\ &\quad - P(\text{CHERRY}/\text{CHERRY}/\text{CHERRY}) = \frac{1}{4} - \frac{3}{64} - \frac{1}{64} = \frac{12}{64} \end{aligned} \tag{1}$$

$$\begin{aligned} E &= (20 + 15 + 5 + 3) \cdot \frac{1}{64} + 2 \cdot \frac{3}{64} + 1 \cdot \frac{12}{64} = \\ &= \frac{43 + 6 + 12}{64} = \frac{61}{64} \end{aligned}$$

b:

$$P(\text{win}) = \sum P(\text{all above in (1)}) = 4 \cdot \frac{1}{64} + \frac{3}{64} + \frac{12}{64} = \frac{19}{64}$$

2.3 Exercise 13.13

$$Test = \{true, false\}$$

$$Virus = \{present, \neg present\}$$

$$P_A(true|present) = 0.95$$

$$P_A(true|\neg present) = 0.1$$

$$P_B(true|present) = 0.9$$

$$P_B(true|\neg present) = 0.05$$

$$P(present) = 0.01$$

A test is more indicative, if its posterior probability (probability of a person to have a virus given a positive result) is higher. Find it for each of the tests and compare:

$$P_A(present|true) \propto P_A(true|present)P(present) = 0.95 * 0.01 = 0.0095$$

$$P_A(\neg present|true) \propto P_A(true|\neg present)P(\neg present) = 0.1 * 0.99 = 0.099$$

thus:

$$P_A(present|true) = \frac{0.0095}{0.0095 + 0.099} = 0.08756$$

similarly for B:

$$P_B(present|true) = \frac{0.9 * 0.01}{0.9 * 0.01 + 0.05 * 0.99} = 0.15385$$

Here we can see the B test is more reliable.

2.4 Exercise 13.15

$$P(true|present) = 0.99$$

$$P(false|\neg present) = 0.99$$

$$P(present) = 0.0001$$

As $P(present|true)$ is proportional to $P(present)$, it is the good news that $P(present)$ is

low. Let's calculate it:

$$P(\text{present}|\text{true}) \propto P(\text{true}|\text{present})P(\text{present}) = 0.99 * 0.0001 = 0.000099$$

$$P(\neg\text{present}|\text{true}) \propto P(\text{true}|\neg\text{present})P(\neg\text{present}) = 0.01 * 0.9999 = 0.009999$$

thus:

$$P(\text{present}|\text{true}) = \frac{0.000099}{0.000099 + 0.009999} = 0.0098$$

It is a really low chance of having the disease. The number means that approx. 1 of 100 people given the positive result of the test is actually have the disease.

2.5 Exercise 13.17

$$P(X, Y|Z) = P(X|Z)P(Y|Z) \tag{2}$$

$$\begin{aligned} P(X|Y, Z) &= \frac{P(X, Y, Z)}{P(Y, Z)} = \frac{P(X, Y|Z)P(Z)}{P(Y, Z)} = \\ &= \text{using (2)} = \frac{P(X|Z)P(Y|Z)P(Z)}{P(Y, Z)} = \\ &= \frac{P(X|Z)P(Y, Z)}{P(Y, Z)} = P(X|Z) \end{aligned}$$

as X and Y are symmetrical, it is also true for $P(Y|X, Z) = P(Y|Z)$

2.6 Exercise 13.21

a:

$$P(\text{appears blue}|\text{blue}) = 0.75$$

$$P(\text{appears green}|\text{green}) = 0.75$$

We need to calculate $P(\text{blue}|\text{appears blue})$

$$\begin{aligned} P(\text{blue}|\text{appears blue}) &\propto P(\text{appears blue}|\text{blue})P(\text{blue}) \\ P(\text{green}|\text{appears blue}) &\propto P(\text{appears blue}|\text{green})P(\text{green}) = \\ &= (1 - P(\text{appears blue}|\text{blue}))(1 - P(\text{blue})) \end{aligned}$$

Here we need to know $P(\text{blue})$ to find the answer. As we don't have a prior knowledge, let's

consider $P(\text{blue}) = P(\text{green}) = 0.5$. Thus:

$$P(\text{blue}|\text{appears blue}) \propto 0.75 * 0.5 = 0.375$$

$$P(\text{green}|\text{appears blue}) \propto 0.25 * 0.5 = 0.125$$

\implies

$$P(\text{blue}|\text{appears blue}) = \frac{0.375}{0.375 + 0.125} = 0.75$$

b: Now, knowing the prior distribution $P(\text{green}) = 0.9$:

$$P(\text{blue}|\text{appears blue}) = \frac{0.75 * 0.1}{0.75 * 0.1 + 0.25 * 0.9} = 0.25$$

$$P(\text{green}|\text{appears blue}) = 0.75$$

Here the most likely color is green.