# Report Reinforcment Learning

## 1) 3 states



**STATE 1**

$x_i \in \{0;1\}$ ... přítomnost i-té karty
$i \in \{1,...,52\}$
$S \in N_0$ ... hodnota mojí ruky
$X_D \in \{1,...,52\}$ ... dealerova ukázaná karta
$V_D \in N$ ... hodnota dealerovy karty
$a \in \{0,1\}$ ... 1 hit, 0 stick

dict = {state i : [v, num.visits]}

www.alcronhotel.com

**STATE 2**

$S_i$ $i \in \{1-4\}$ number of i-th suit in players hand
$R_j$ $j \in \{1,-,13\}$ number of j-th rank in players hand
$M\$ \in N$ value of my hand
$S_D$ suit of dealr card
$R_D$ rank of dealrs hand

**STATE 3**

$M \in N_0^*$ value of my hand
$a$ number of aces in hand
$C_m$ number of cards in hand
$V_D$ value of dealrs card

-21     21-4-44
0-4     19.4.9
8-11    =675
12 - 101

## 2) rough estimate of overall number of states



**STATE 1**

$$\left(\binom{52}{2}+\binom{52}{3}+\binom{52}{4}+\binom{52}{5}+\binom{52}{6}\right)\cdot \underset{S}{20}\cdot\underset{X_D}{52}\cdot\underset{V_D}{11}=$$

$$=(1326+22100+270725+2598960)\cdot 20\cdot 52\cdot 11$$

$$\doteq 32\ 032\ 000\ 000$$

**STATE 2**

$\binom{n+n-1}{n} \propto$ combinations with repetition

$n=4$ $S_i$

$$\left(\binom{1+n-1}{1}+\binom{2+n-1}{2}+\binom{3+n-1}{3}+\binom{4+n-1}{4}+\binom{5+n-1}{5}\right)\cdot$$

$R_j$

$$\left(\binom{1+13-1}{1}+\binom{2+13-1}{2}+\binom{3+13-1}{3}+\binom{4+13-1}{4}+\binom{5+13-1}{5}\right)\cdot \underset{M}{20}\cdot \underset{S_D}{4}\cdot \underset{R_D}{13}$$

$$= (4+10+20+35+56)\cdot(13+91+455+1820+6188)\cdot 21\cdot 4\cdot 13$$

$$\doteq 1\ 169\ 395\ 500$$

**STATE 3**

$$\underset{M}{20}\cdot \underset{a}{5}\cdot \underset{c}{9}\cdot \underset{V_D}{10} = 9000$$

Calculations for representations 1 and 2 contain many invalid states so the result is only a rough estimate.

For state 1 we calculate combinations of cards that we can have in hand and multiply this with the all possible values in hand, dealers card and its values.

For state 2 we calculate combinations of suits in hand multiplied by combinations of rank, value of hand, suit of dealers card and rank of dealers card.

For state 3 we multiply possible number of values in player's hand, number of aces in hand, cards in hand and value of dealers hand.

3) I have chosen state 3. Simply for the reason that the other two state representations are too large. It contains the value of the player's hand, number of aces in hand because that inforamtion might help us since they can contain two values, number of cards in hand and value of dealers card. This state does not contain all the information. The simplifications is in that we do not know ranks or suits of any of the cards in player's or dealer's hand.
We cannot use exact methods like policy or value iteration since we do not have an MDP. Number of episodes and discount factor were chosen based on trial and error method.

6) Comparison of the agents can be found in the repository as various pdfs from averages to moving averages and so on. Estimates of the U/Q values are also in the pdfs.