

Playing Blackjack with Reinforcement Learning

Final Project: Training the artificial player

Mar 29, 2018

Youngho Jung



Game Procedure



Betting



Dealing

- 2 cards for the dealer
- 2 cards for each player



Action

- Stand, Hit, Doubledown
- Split

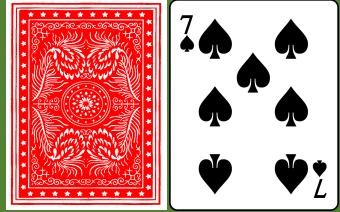


Reward

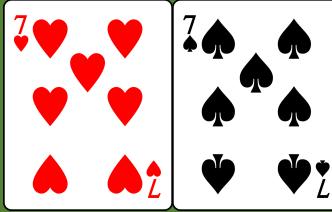
- Sum of cards ≈ 21

Game Procedure: example

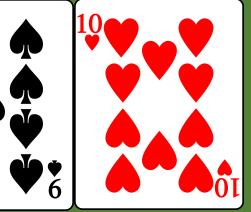
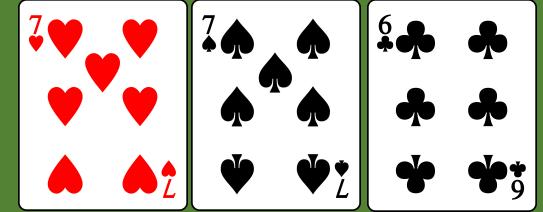
Dealer: 7



Dealer: 14



Dealer: 20



Player: 19



Player: 19

Stand



Dealer draws

Player: 19

I. Infinite decks

Case 1: Each card is drawn with equal probability

Basic Strategy

Basic Strategy		dealer's face-up card				
player's hand	2	3	4	5	6	
	20	S	S	S	S	S
	19	S	S	S	S	S
	18	S	S	S	S	S
	17	S	S	S	S	S
	16	S	S	S	S	S
	15	S	S	S	S	S
	14	S	S	S	S	S
	13	S	S	S	S	S
	12	H	H	S	S	S
	11	D	D	D	D	D
	10	D	D	D	D	D
	9	H	D	D	D	D
	8	H	H	H	H	H
	7	H	H	H	H	H
	6	H	H	H	H	H
	5	H	H	H	H	H

- Re-splitting allowed
- Doubledown allowed after split
- Dealer hits on a soft 17
- 4 – 8 decks



Stand



Hit



Doubledown



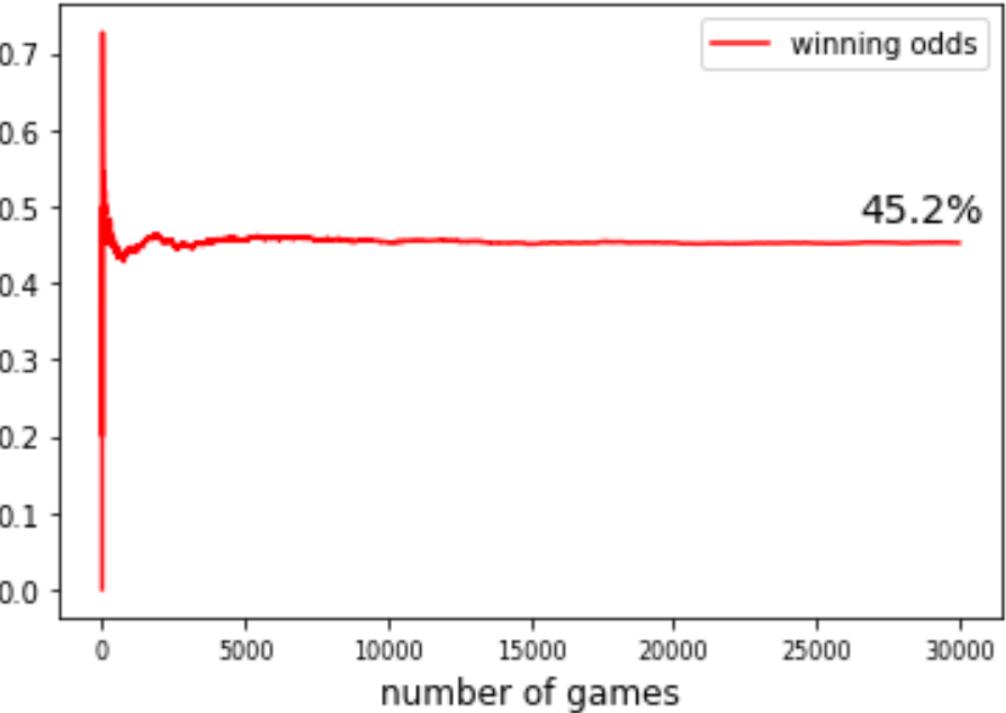
Split

Basic Strategy

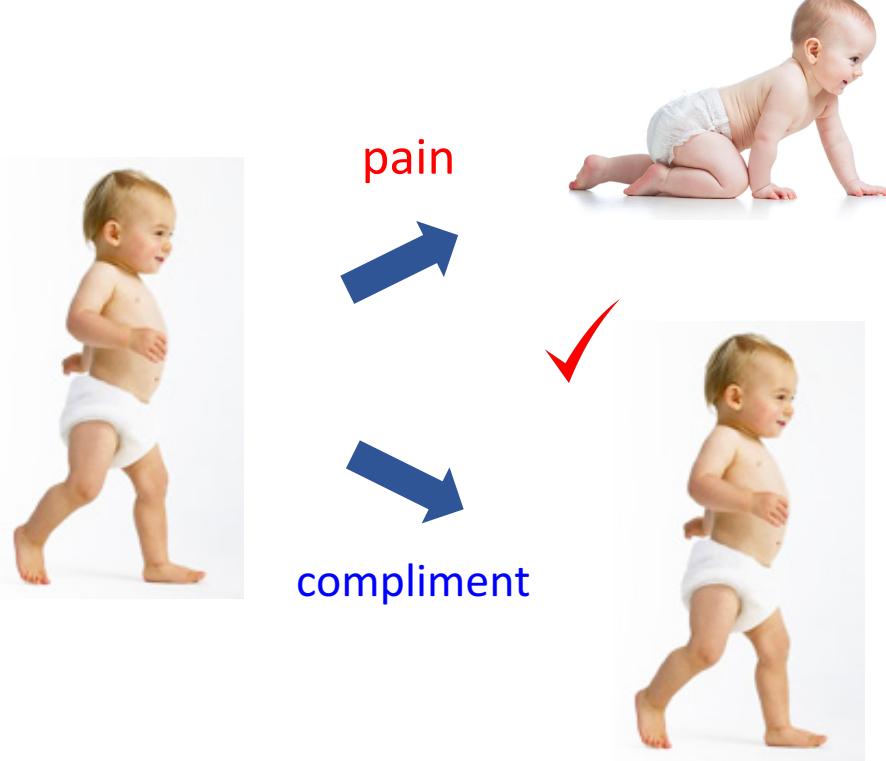
Basic Strategy		dealer's face-up card				
player's hand	2	3	4	5	6	
	20	S	S	S	S	S
	19	S	S	S	S	S
	18	S	S	S	S	S
	17	S	S	S	S	S
	16	S	S	S	S	S
	15	S	S	S	S	S
	14	S	S	S	S	S
	13	S	S	S	S	S
	12	H	H	S	S	S
	11	D	D	D	D	D
	10	D	D	D	D	D
	9	H	D	D	D	D
	8	H	H	H	H	H
	7	H	H	H	H	H
	6	H	H	H	H	H
	5	H	H	H	H	H

- Doubledown allowed after split
- Dealer hits on a soft 17
- 4 – 8 decks

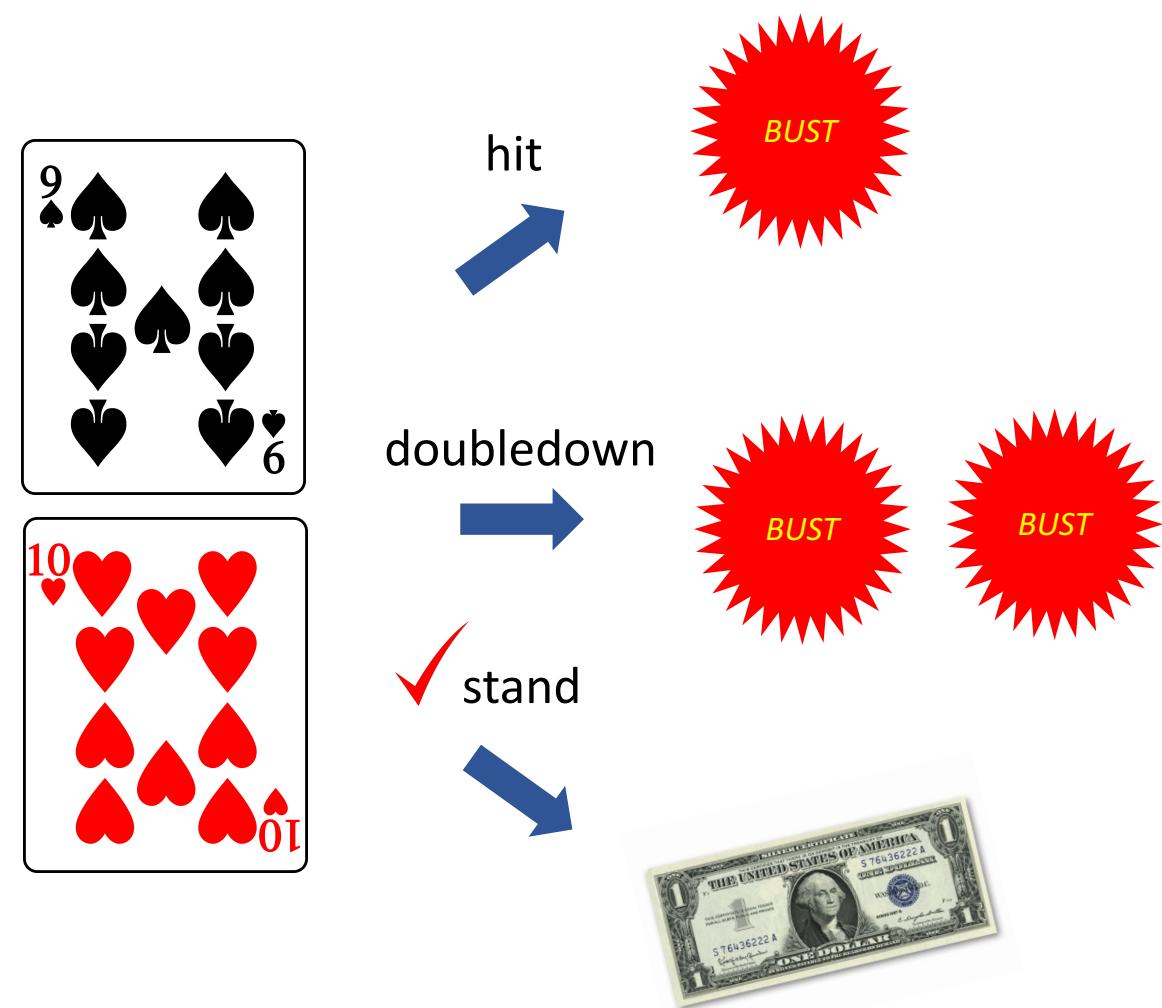
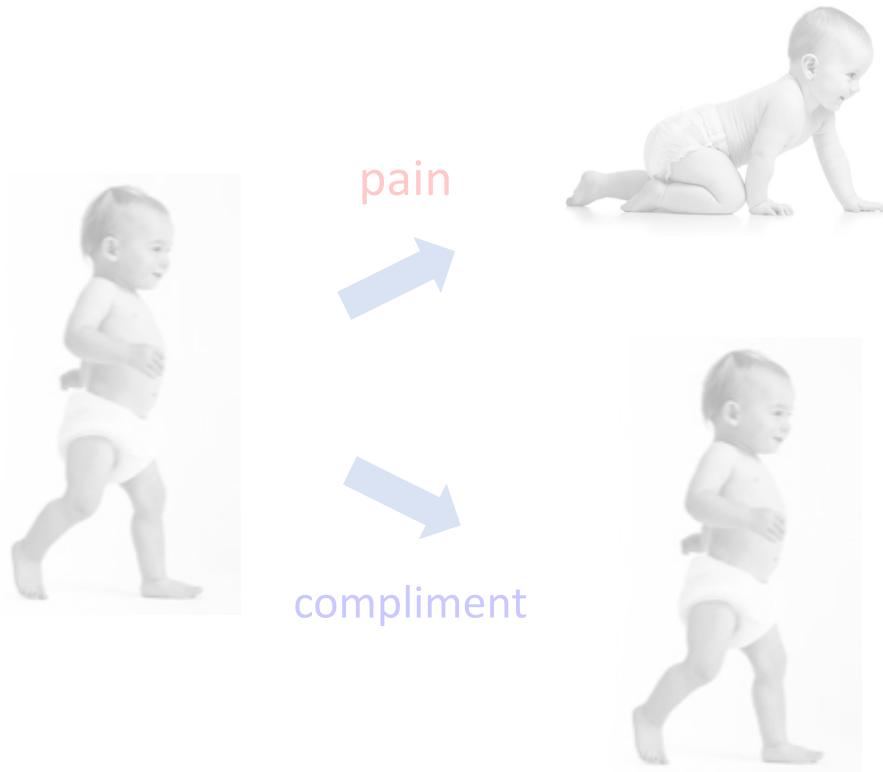
Winning Odds / Basic Strategy



Q-Learning: Reinforcement Learning



Q-Learning: Reinforcement Learning



Q-Learning Strategy

n = 100

player's hand	dealer's face-up card					
	2	3	4	5	6	
20	S	S	S	S	S	
19	S	S	S	S	S	
18	S	S	S	S	S	
17	D	S	S	S	S	
16	S	S	S	S	S	
15	S	S	S	S	D	
14	S	S	S	S	S	
13	S	S	S	S	S	
12	D	D	D	S	S	
11	S	S	S	S	D	
10	S	S	S	S	S	
9	S	S	S	S	S	
8	S	S	S	S	S	
7	S	S	S	D	S	
6	S	S	S	S	S	
5	S	S	S	S	S	
A,9	S	S	D	S	S	
A,8	S	S	S	S	S	
A,7	S	S	S	S	S	
A,6	S	S	S	S	S	
A,5	S	S	S	S	S	
A,4	S	S	S	S	S	
A,3	S	S	S	S	S	
A,2	S	S	S	S	S	
A,A	S	S	S	S	S	
10,10	S	SP	S	S	S	
9,9	S	S	S	S	S	
8,8	S	S	S	S	S	
7,7	S	S	S	S	S	
6,6	S	S	S	S	S	
5,5	S	S	S	S	S	
4,4	S	S	S	S	S	
3,3	S	S	S	S	S	
2,2	S	S	S	S	S	

n = 10000

player's hand	dealer's face-up card					
	2	3	4	5	6	
20	S	S	S	S	S	
19	S	S	S	S	S	
18	S	S	S	S	S	
17	S	S	S	S	D	
16	S	S	S	S	S	
15	S	D	S	S	H	
14	D	S	S	D	S	
13	D	H	S	H	D	
12	D	D	H	D	D	
11	D	H	D	D	H	
10	D	D	H	D	D	
9	D	D	D	D	D	
8	D	H	D	H	D	
7	D	H	D	D	D	
6	D	H	H	H	H	
5	H	D	H	H	H	
A,9	D	D	D	D	D	
A,8	D	H	D	D	S	
A,7	S	S	S	D	D	
A,6	S	H	S	S	S	
A,5	S	D	S	D	D	
A,4	D	D	S	D	D	
A,3	S	D	S	S	H	
A,2	S	S	D	S	S	
A,A	D	D	H	SP	D	
10,10	S	S	S	SP	SP	
9,9	S	SP	H	SP	S	
8,8	D	S	SP	S	D	
7,7	S	SP	H	S	D	
6,6	SP	S	H	H	SP	
5,5	H	D	H	D	D	
4,4	H	SP	D	H	H	
3,3	S	SP	H	SP	SP	
2,2	S	SP	H	H	H	

n = 100000

player's hand	dealer's face-up card					
	2	3	4	5	6	
20	S	S	S	S	S	
19	S	S	S	S	S	
18	S	S	S	S	S	
17	S	S	S	S	S	
16	S	S	S	S	S	
15	S	S	S	S	S	
14	H	H	S	H	H	
13	H	H	H	D	D	
12	H	H	H	D	H	
11	D	D	D	D	D	
10	D	D	D	D	D	
9	D	D	D	D	H	
8	D	H	H	H	H	
7	H	H	H	H	H	
6	H	H	H	H	H	
5	H	H	H	H	H	
A,9	S	S	S	S	S	
A,8	S	S	S	S	D	
A,7	D	D	D	S	S	
A,6	D	D	D	S	S	
A,5	S	D	S	D	D	
A,4	D	D	S	D	H	
A,3	H	H	D	D	D	
A,2	H	H	H	D	H	
A,A	SP	SP	SP	SP	SP	
10,10	S	S	S	S	S	
9,9	S	S	S	S	SP	
8,8	SP	SP	S	SP	D	
7,7	D	H	H	SP	D	
6,6	SP	D	SP	SP	SP	
5,5	D	SP	D	D	D	
4,4	H	SP	SP	H	SP	
3,3	H	SP	H	SP	SP	
2,2	H	SP	H	H	SP	



Stand



Hit



Doubledown



Split

Q-Learning VS Basic Strategy

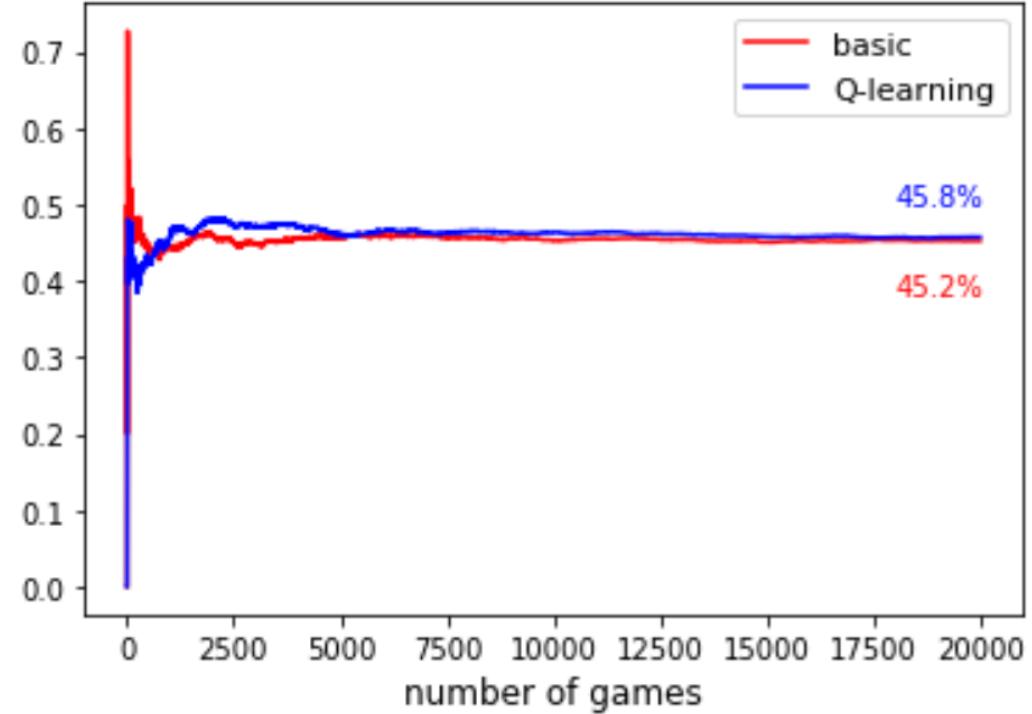
Q-Learning

Basic Strategy		dealer's face-up card			
player's hand		2	3	4	5
	20	S	S	S	S
	19	S	S	S	S
	18	S	S	S	S
	17	S	S	S	S
	16	S	S	S	S
	15	S	S	S	S
	14	H	H	S	H
	13	H	H	H	D
	12	H	H	H	D
	11	D	D	D	D
	10	D	D	D	D
	9	D	D	D	H
	8	D	H	H	H
	7	H	H	H	H
	6	H	H	H	H
	5	H	H	H	H
	A,9	S	S	S	S
	A,8	S	S	S	S
	A,7	D	S	S	S
	A,6	D	D	S	S
	A,5	S	D	S	D
	A,4	D	D	S	D
	A,3	H	H	D	D
	A,2	H	H	H	D
	A,A	SP	SP	SP	SP
	10,10	S	S	S	S
	9,9	S	S	S	SP
	8,8	SP	SP	S	SP
	7,7	D	H	H	SP
	6,6	SP	D	SP	SP
	5,5	D	SP	D	D
	4,4	H	SP	SP	H
	3,3	H	SP	H	SP
	2,2	H	SP	H	H

Basic Strategy

Basic Strategy		dealer's face-up card			
player's hand		2	3	4	5
	20	S	S	S	S
	19	S	S	S	S
	18	S	S	S	S
	17	S	S	S	S
	16	S	S	S	S
	15	S	S	S	S
	14	S	S	S	S
	13	S	S	S	S
	12	H	H	S	S
	11	D	D	D	D
	10	D	D	D	D
	9	H	D	D	D
	8	H	H	H	H
	7	H	H	H	H
	6	H	H	H	H
	5	H	H	H	H
	A,9	S	S	S	S
	A,8	S	S	S	S
	A,7	D	D	D	D
	A,6	H	D	D	D
	A,5	H	H	D	D
	A,4	H	H	D	D
	A,3	H	H	H	D
	A,2	H	H	H	D
	A,A	SP	SP	SP	SP
	10,10	S	S	S	S
	9,9	SP	SP	SP	SP
	8,8	SP	SP	SP	SP
	7,7	SP	SP	SP	SP
	6,6	SP	SP	SP	SP
	5,5	D	D	D	D
	4,4	H	H	H	SP
	3,3	SP	SP	SP	SP
	2,2	SP	SP	SP	SP

Winning Odds



II. 6 decks

Case 2: Cards drawn in the past affect the future outcomes

Q-Learning for a 6-deck game

- Remaining cards matter

Q-Learning for a 6-deck game

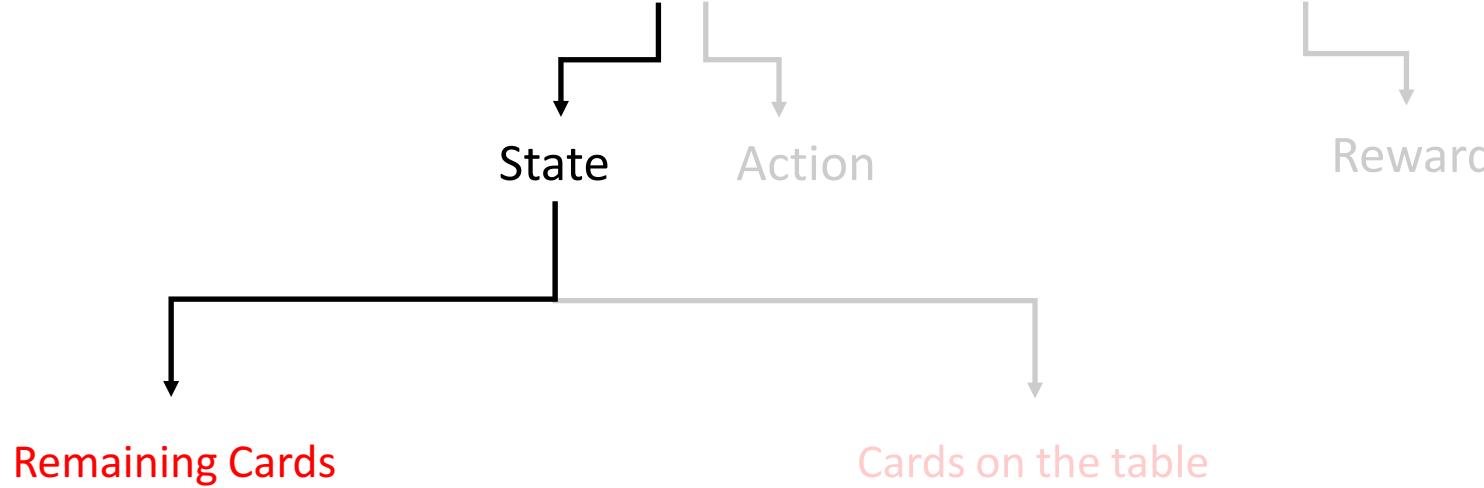
- Remaining cards matter

$$Q(\textcolor{red}{S}, \textcolor{red}{A}) \leftarrow (1 - \alpha)Q(S, A) + \alpha(\textcolor{red}{R} + \gamma \max_A Q(S', A))$$

Q-Learning for a 6-deck game

- Remaining cards matter

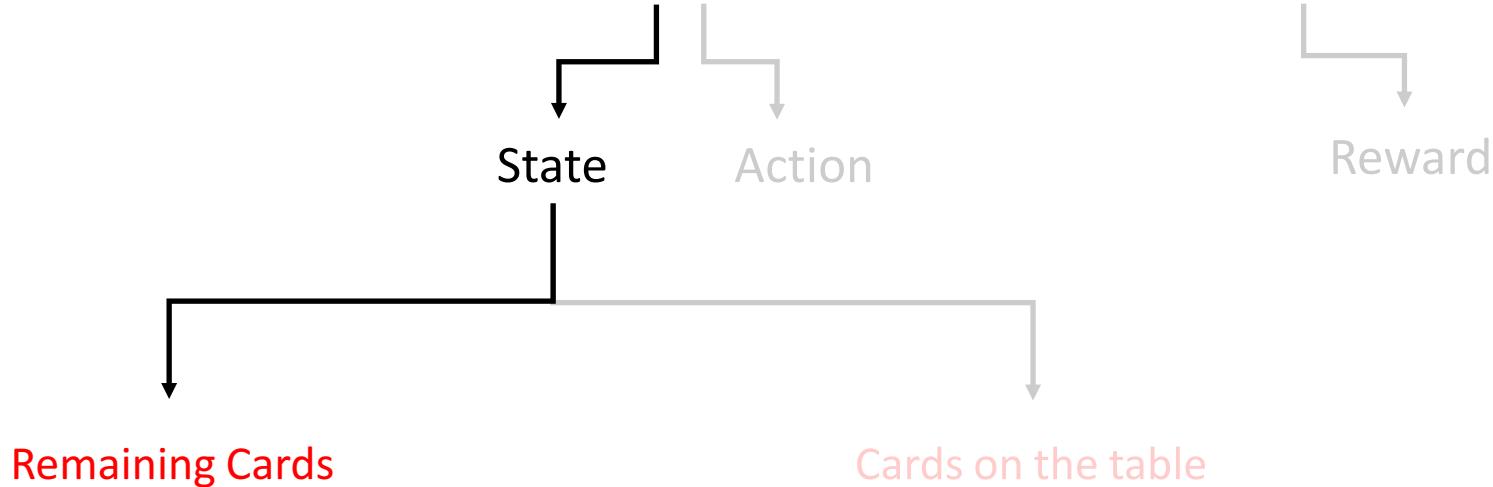
$$Q(S, A) \leftarrow (1 - \alpha)Q(S, A) + \alpha(R + \gamma \max_A Q(S', A))$$



Q-Learning for a 6-deck game

- Remaining cards matter

$$Q(S, A) \leftarrow (1 - \alpha)Q(S, A) + \alpha(R + \gamma \max_A Q(S', A))$$



I. Hi-Low Card Counting

2, 3, 4, 5, 6 → +1

7, 8, 9 → 0

10, J, Q, K, A → -1

II. Omega-II Card Counting

2, 3, 7 → +1

4, 5, 6 → +2

8, A → 0

9 → -1

10, J, Q, K → -2

Hi-Low Card Counting

Low Cards Remaining

Strategy	dealer's face-up card				
	2	3	4	5	6
20	S	S	S	S	S
19	S	S	S	S	S
18	S	S	S	S	S
17	S	S	S	S	S
16	S	S	S	S	S
15	H	H	S	H	H
14	H	H	H	H	H
13	H	H	H	H	H
12	H	H	H	H	H
11	H	H	D	H	H
10	H	D	H	D	H
9	H	H	H	H	H
8	H	H	H	H	H
7	H	H	H	H	H
6	H	H	H	H	H
5	H	H	H	H	H

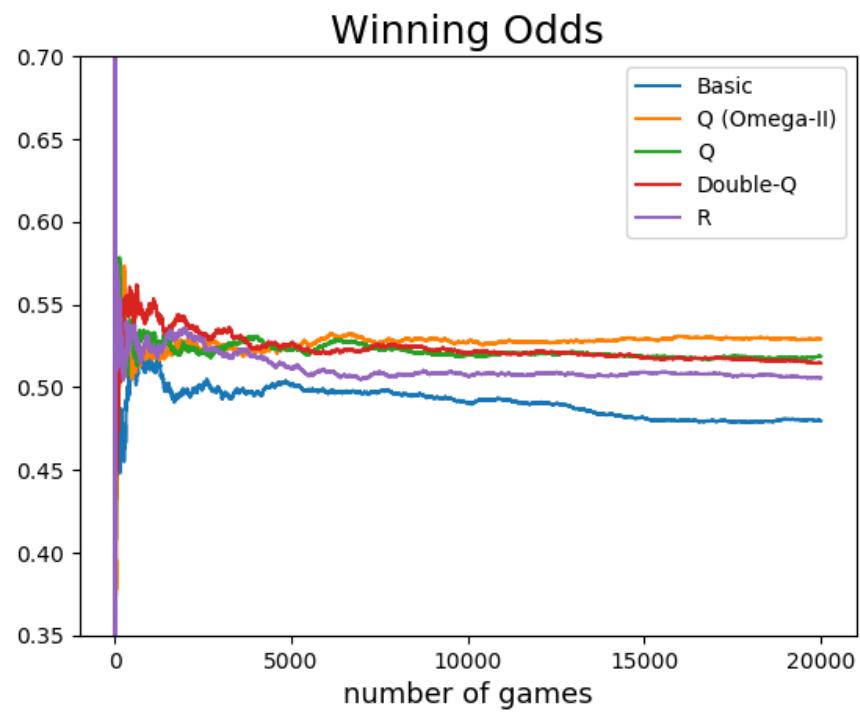
High Cards Remaining

Strategy	dealer's face-up card				
	2	3	4	5	6
20	S	S	S	S	S
19	S	S	S	S	S
18	S	S	S	S	S
17	S	S	S	S	S
16	S	S	S	S	S
15	S	S	S	S	H
14	H	H	H	H	H
13	H	H	H	H	H
12	H	H	H	H	H
11	D	D	D	D	D
10	D	D	H	D	D
9	H	H	H	H	H
8	H	H	H	H	H
7	H	H	H	H	H
6	H	H	H	H	H
5	H	H	H	H	H

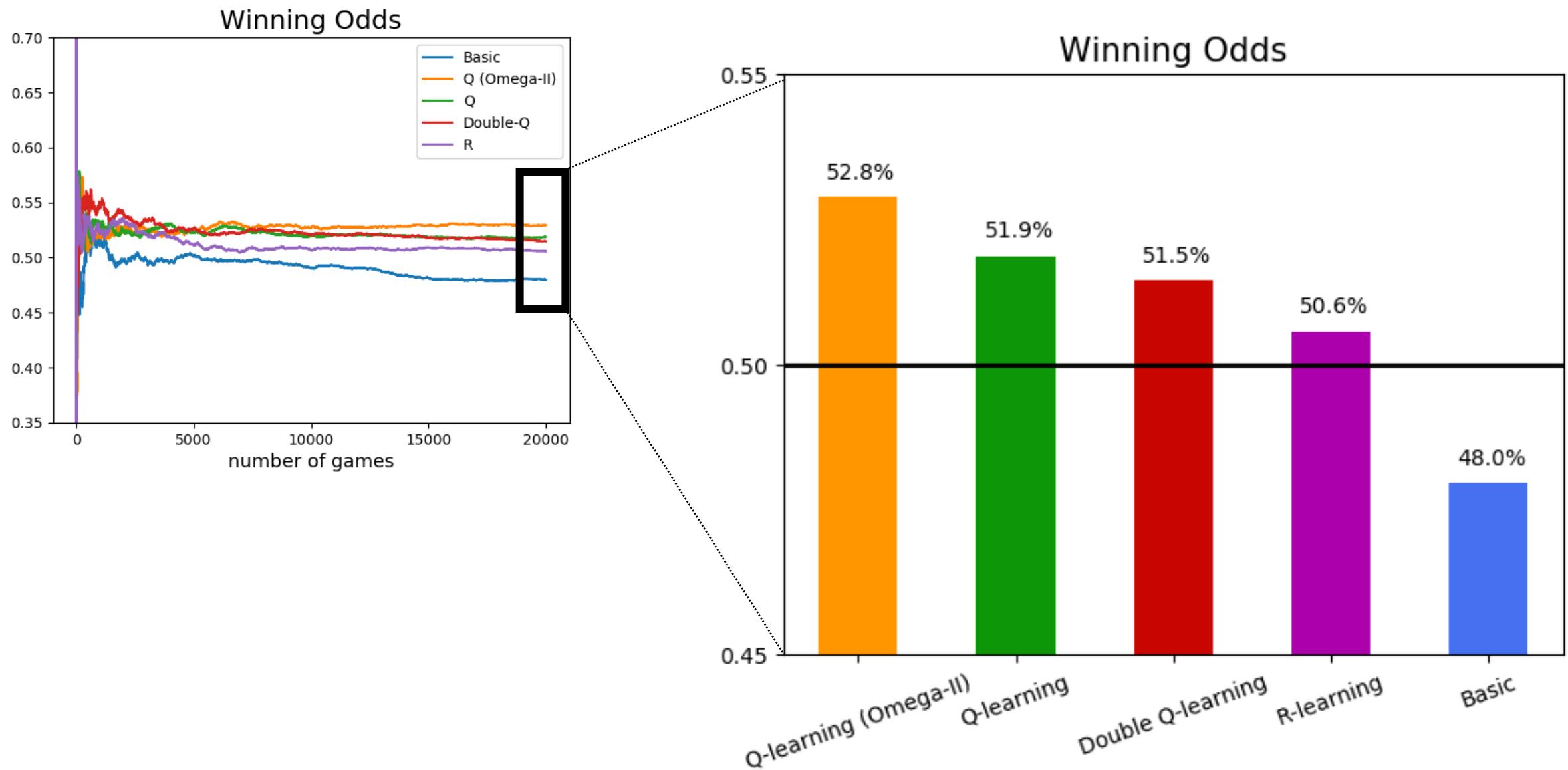
More high cards → It is easier to make dealer go bust



Strategy Comparisons



Strategy Comparisons



Conclusion and Further Studies

- ✓ Omega-II based Q-learning achieved winning odds of 53% for a 6-deck Blackjack.
- ✓ Changing the betting amount dynamically could bring better results.

Thank you

Youngho Jung

youngho.jung@columbia.edu
347.819.0163

