

EI320A(3) 深度學習使用 Python

Instructors

Tipajin Thaipisutikul (t.greentip@gmail.com)

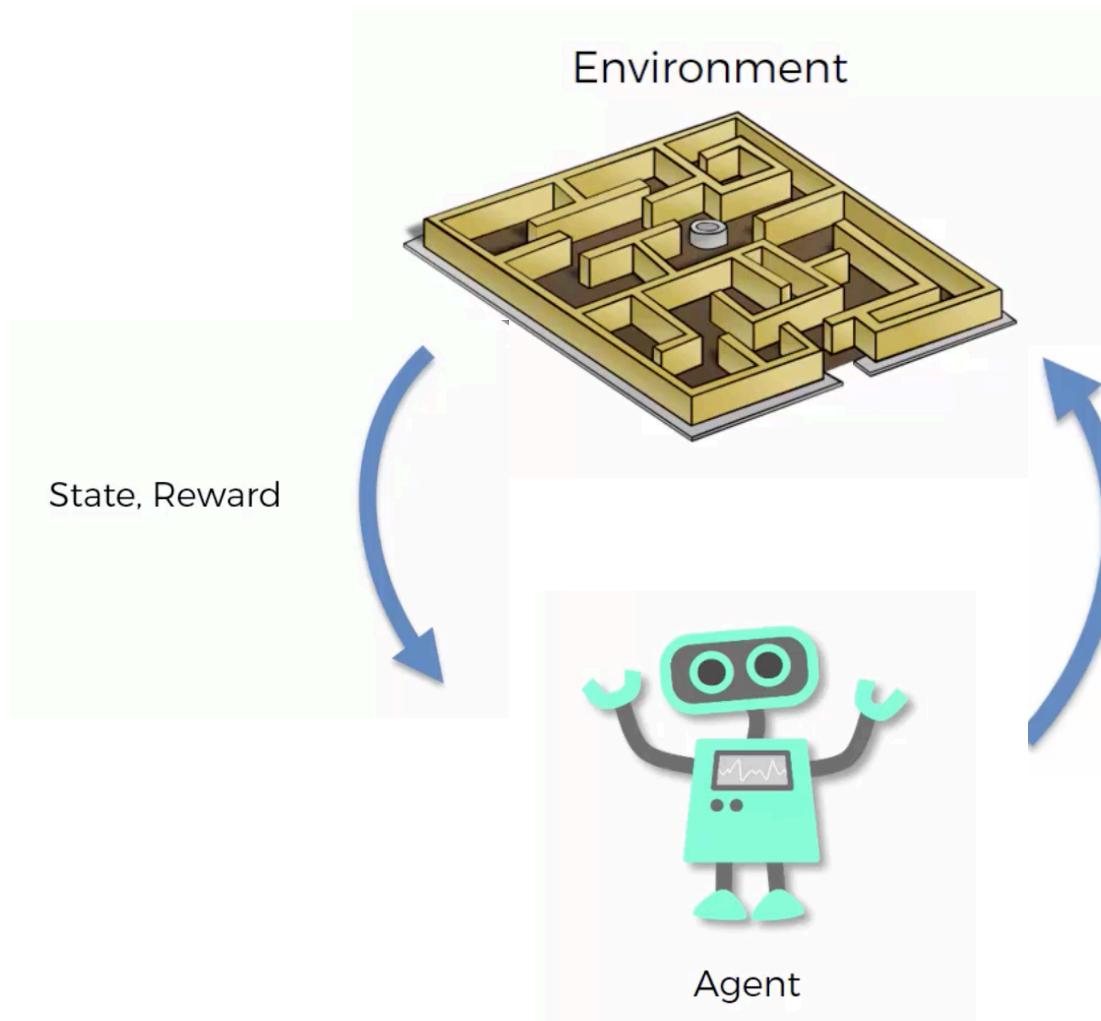
Prof. Huang-Chia Shih (hcshih@Saturn.yzu.edu.tw)

Week	Date	Content	Note	Total
1	2/26	Welcome to the course	Homework (1)	1
2	3/5	Crash Course of Python, NumPy, Pandas, and Matplotlib	In class hands-on (4)	5
3	3/12	Get to know about Data, ML: Classification Models	In class hands-on (5)	10
4	3/19	ML: Regression Models	In class hands-on (5)	15
5	3/26	ML: Clustering/Apriori Models	In class hands-on (5)	20
6	4/2	Holiday		
7	4/9	Introduction to Deep Learning (ANN)		
8	4/16	ANN Labs, Introduction to Convolutional Neural Network (CNN)	In class hands-on (10)	30
9	4/23	Convolutional Neural Network (CNN) & CNN Labs	In class hands-on (5)	35
10	4/30	Introduction to Recurrent Neural Network (RNN)	In class hands-on (5)	40
11	5/7	Recurrent Neural Network (RNN) & RNN Labs	In class hands-on (5)	45
12	5/14	Wrap Up all ANN, CNN, RNN Project Proposal Presentation	Proposal Presentation (10)	55
13	5/21	Generative Adversarial Network (GAN)	In class hands-on (5)	60
14	5/28	Reinforcement Learning (RL)	In class hands-on (5)	65
15	6/4	NLP & S2S & Attention Neural Network	In class hands-on (5)	70
16	6/11	N/A	In class hands-on (5)	75
17	6/18	Final Project Presentation	Final Presentation (25)	100

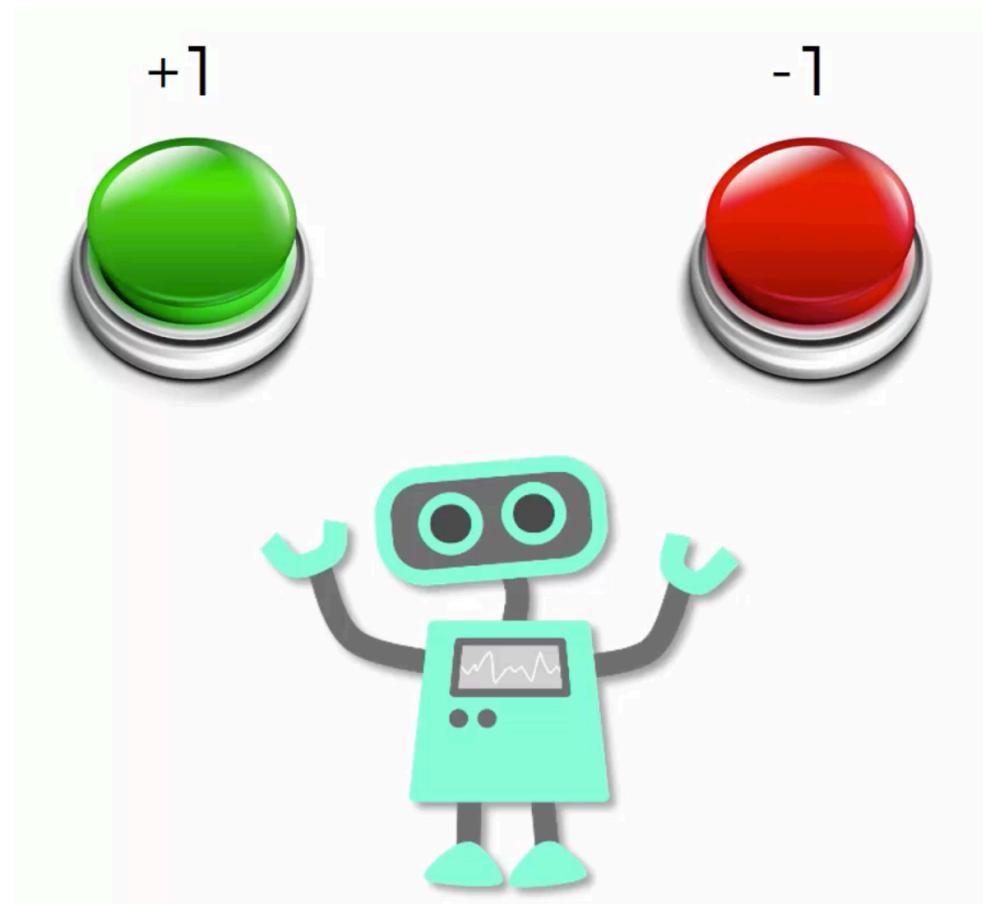
RL Plan of Attack

1. What is Reinforcement Learning? (RL)
2. The Bellman Equation
3. The "Plan"
4. Markov Decision Process (MDP)
5. "Policy" vs. "Plan"
6. Living Penalty
7. Q-Learning Intuition
8. Temporal Difference
9. Deep Q-Learning (Learning & Acting)

1. What is Reinforcement Learning? (RL)



1.What is Reinforcement Learning? (RL)



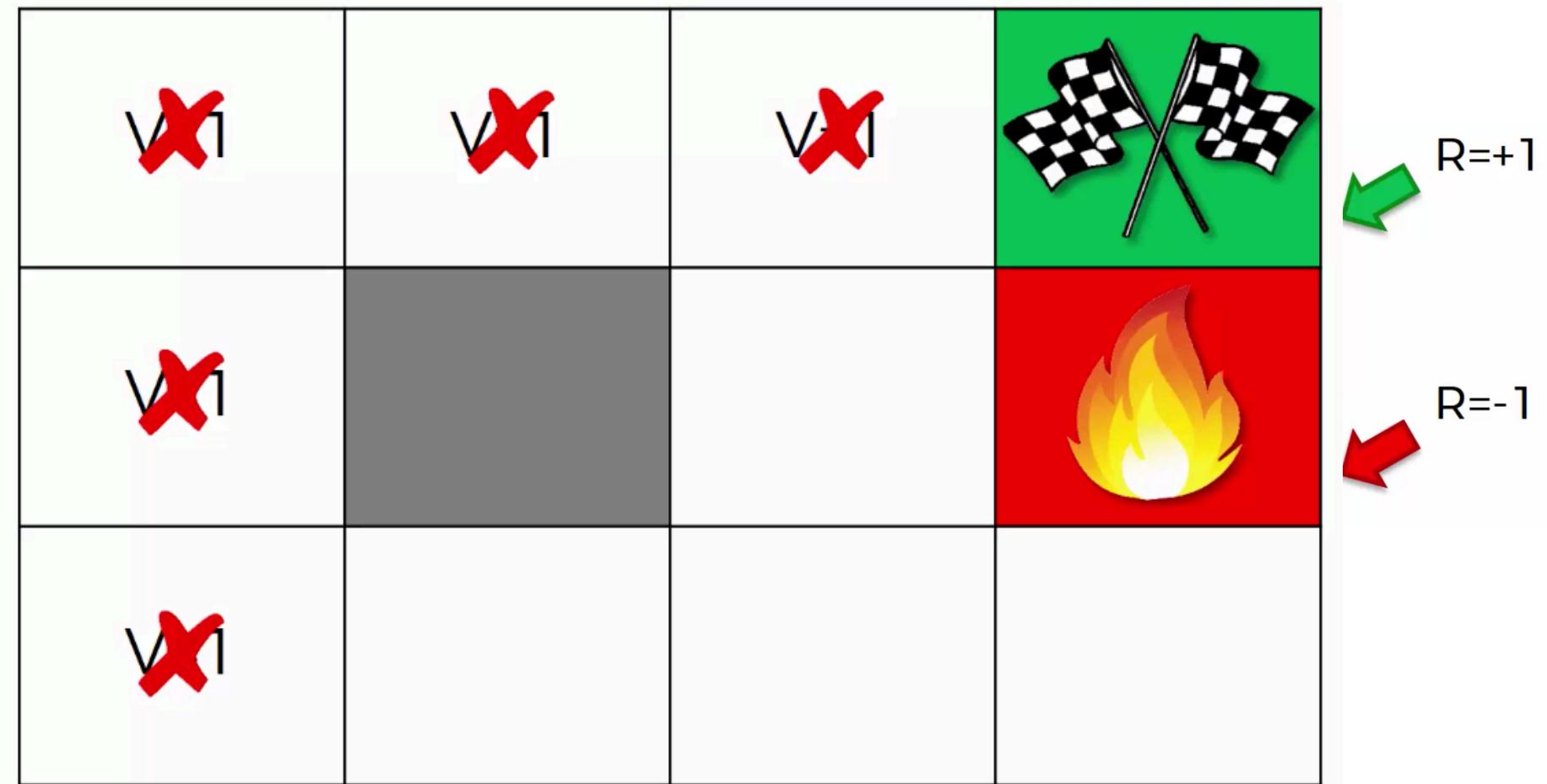
RL Plan of Attack

1. What is Reinforcement Learning? (RL)
2. The Bellman Equation
3. The "Plan"
4. Markov Decision Process (MDP)
5. "Policy" vs. "Plan"
6. Living Penalty
7. Q-Learning Intuition
8. Temporal Difference
9. Deep Q-Learning (Learning & Acting)

2. The Bellman Equation (Dynamic Programming)

Concepts:

- s - State
- a - Action
- R - Reward
- γ - Discount

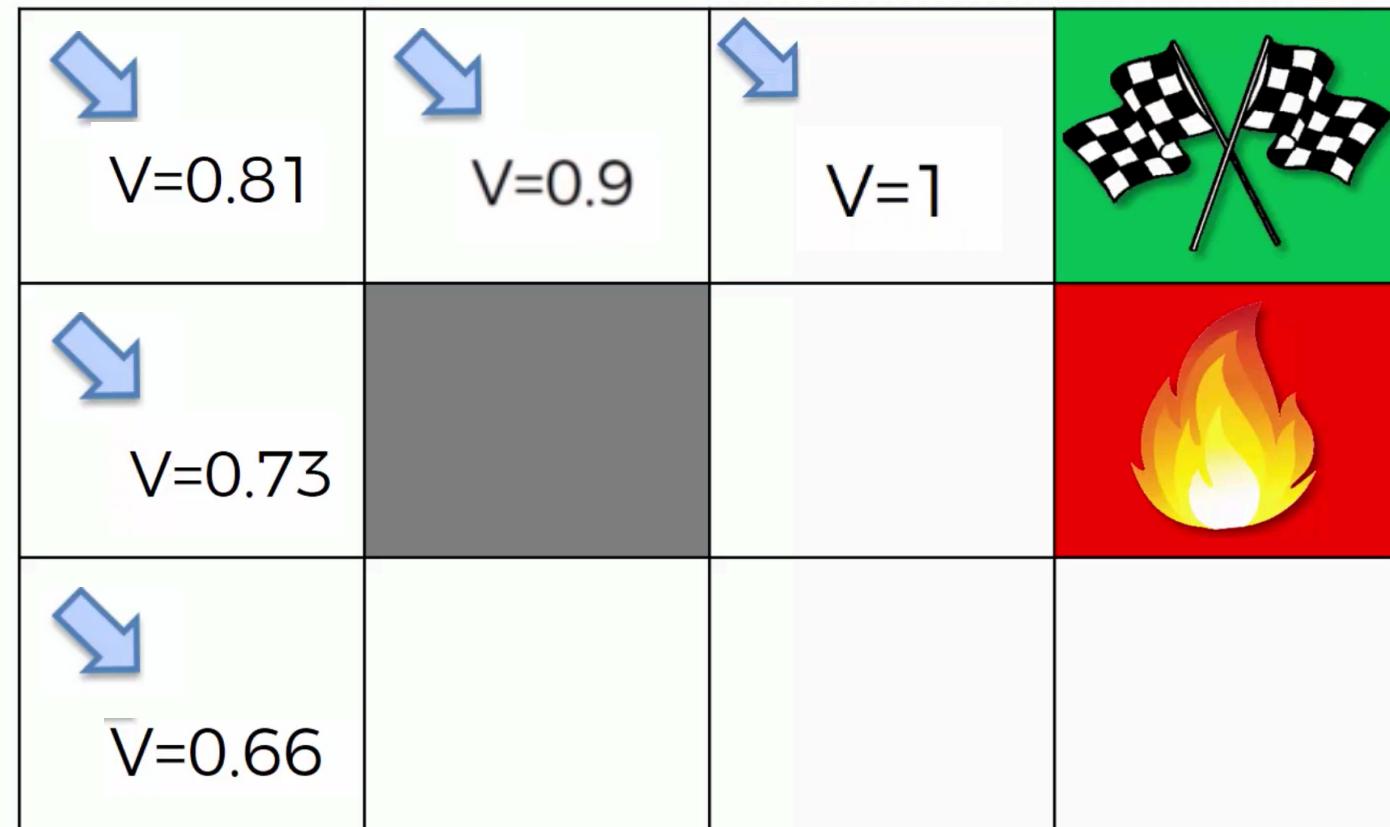


2. The Bellman Equation (Dynamic Programming)

$$V(s) = \max_a (R(s, a) + \gamma V(s'))$$

Concepts:

- s - State
- a - Action
- R - Reward
- γ - Discount

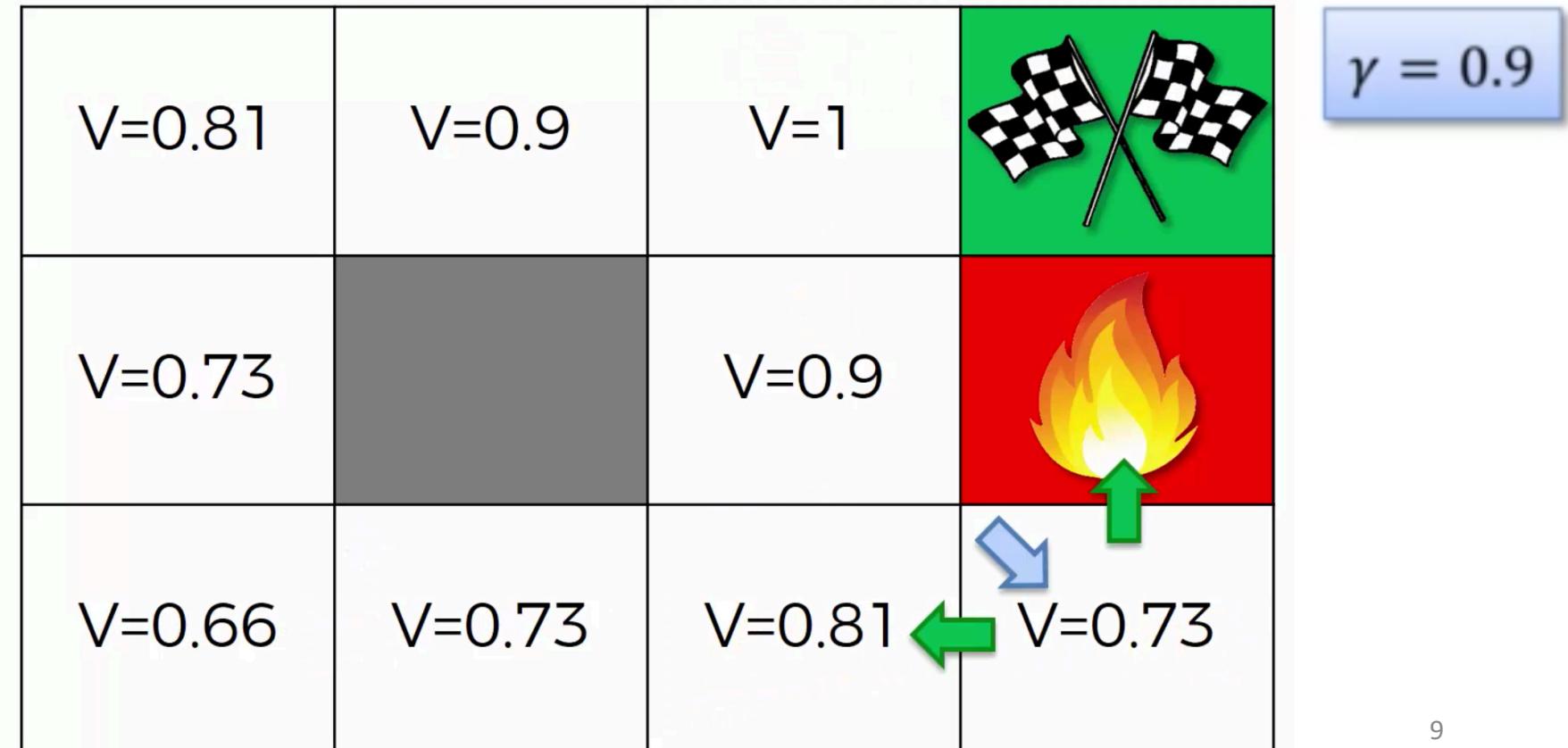


2. The Bellman Equation (Dynamic Programming)

$$V(s) = \max_a (R(s, a) + \gamma V(s'))$$

Concepts:

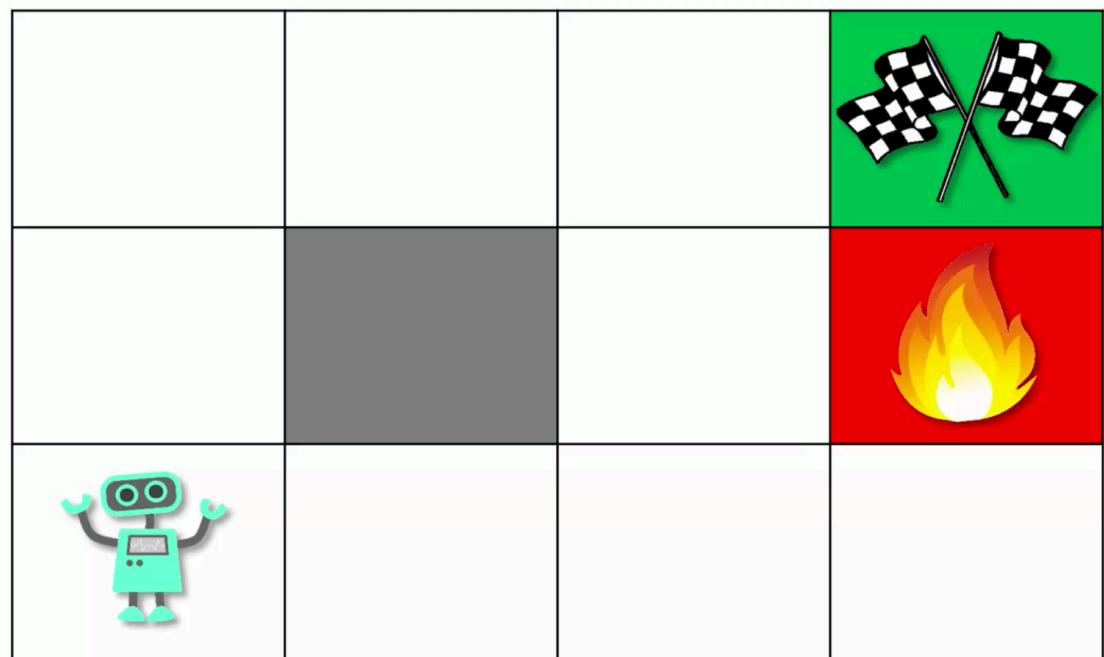
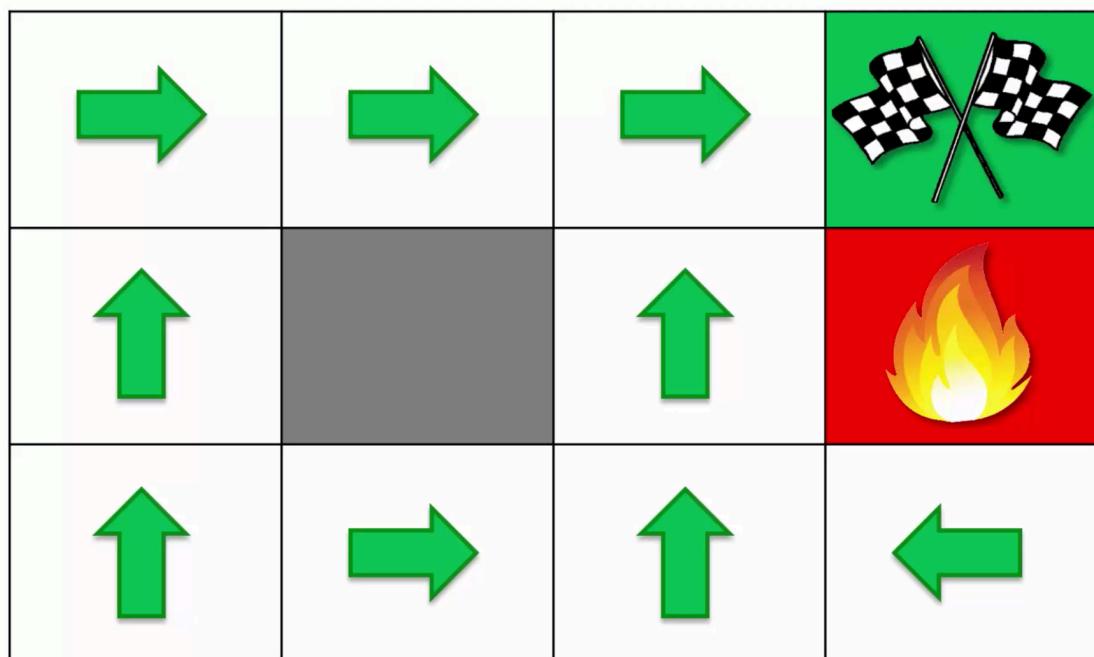
- s - State
- a - Action
- R - Reward
- γ - Discount



RL Plan of Attack

1. What is Reinforcement Learning? (RL)
2. The Bellman Equation
3. **The "Plan"**
4. Markov Decision Process (MDP)
5. “Policy” vs. “Plan”
6. Living Penalty
7. Q-Learning Intuition
8. Temporal Difference
9. Deep Q-Learning (Learning & Acting)

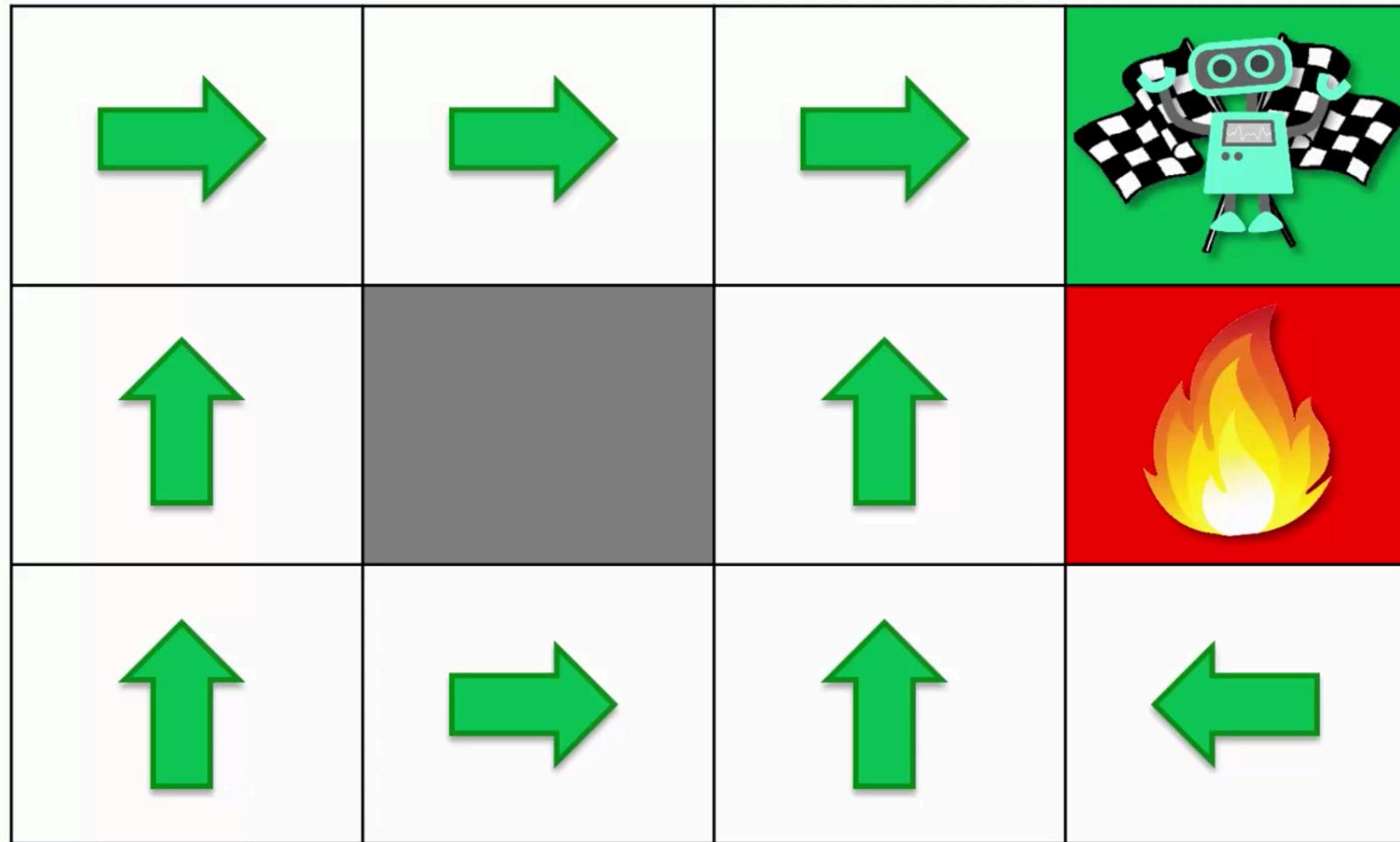
3. The "Plan"



RL Plan of Attack

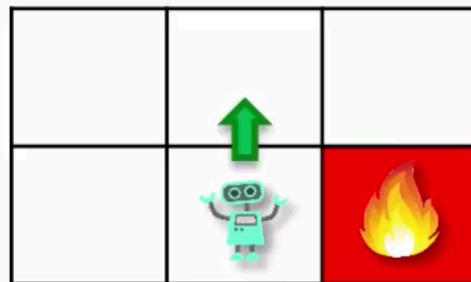
1. What is Reinforcement Learning? (RL)
2. The Bellman Equation
3. The "Plan"
4. **Markov Decision Process (MDP)**
5. "Policy" vs. "Plan"
6. Living Penalty
7. Q-Learning Intuition
8. Temporal Difference
9. Deep Q-Learning (Learning & Acting)

4. Markov Decision Process (MDP)

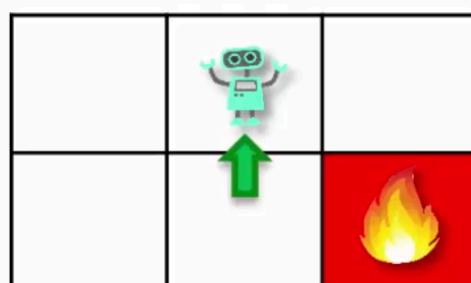


4. Markov Decision Process (MDP)

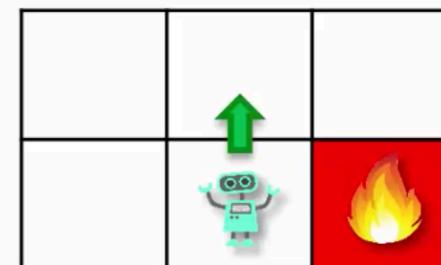
Deterministic Search



100%



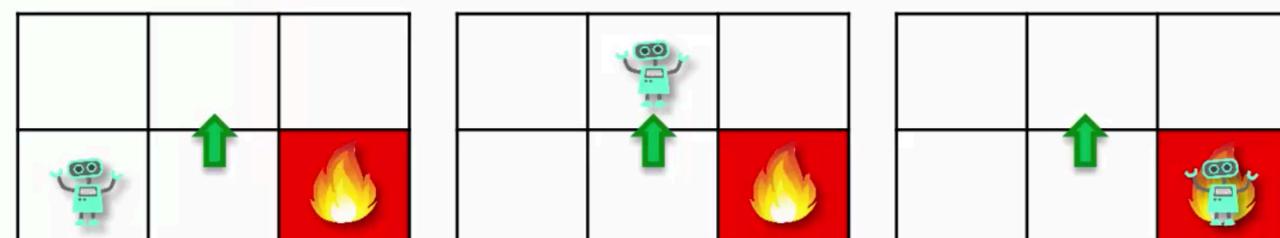
Non-Deterministic Search



10%

80%

10%



4. Markov Decision Process (MDP)

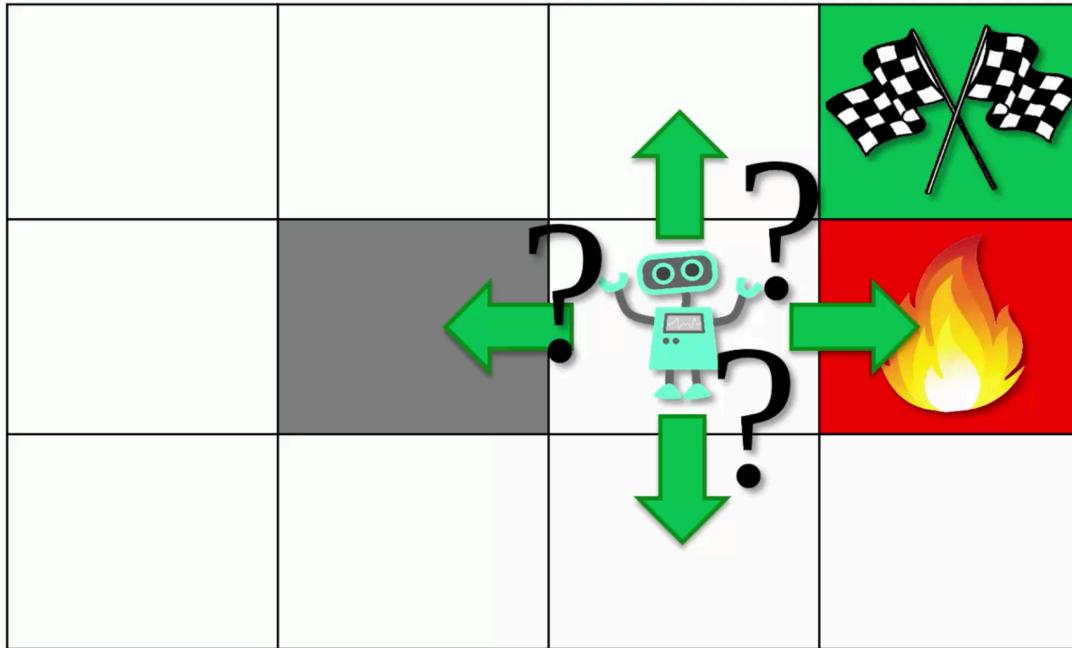
Markov Process

Markov Decision Process (MDP)

A stochastic process has the **Markov property** if the conditional probability distribution of future states of the process (conditional on both past and present states) depends only upon the present state, not on the sequence of events that preceded it. A process with this property is called a Markov process.

Wikipedia

4. Markov Decision Process (MDP)



Markov Decision Processes (MDPs) provide a mathematical framework for modeling decision making in situations where outcomes are partly random and partly under the control of a decision maker.

Wikipedia

4. Markov Decision Process (MDP)

$$0.8 * V(s'_1) + 0.1 * V(s'_2) + 0.1 * V(s'_3)$$

$$V(s) = \max_a (R(s, a) + \gamma \overbrace{V(s')}^{\text{Bellman Equation}})$$

$$V(s) = \max_a \left(R(s, a) + \gamma \sum_{s'} P(s, a, s') V(s') \right)$$

RL Plan of Attack

1. What is Reinforcement Learning? (RL)
2. The Bellman Equation
3. The "Plan"
4. Markov Decision Process (MDP)
5. "Policy" vs. "Plan"
6. Living Penalty
7. Q-Learning Intuition
8. Temporal Difference
9. Deep Q-Learning (Learning & Acting)

5. “Policy” vs. “Plan”

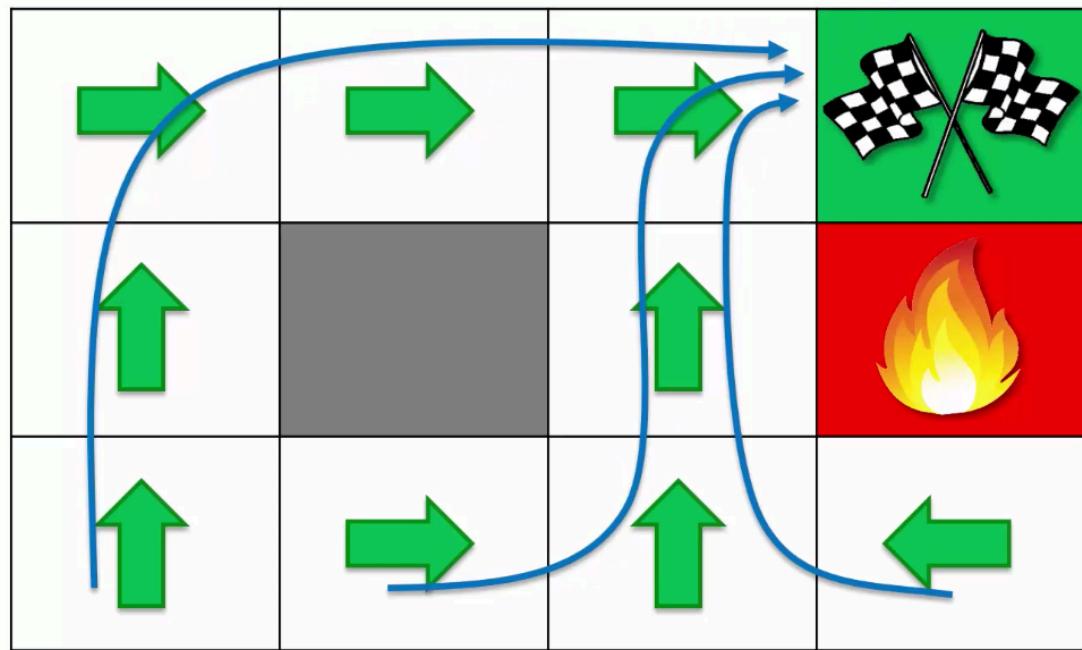
$$V(s) = \max_a \left(R(s, a) + \gamma \sum_{s'} P(s, a, s') V(s') \right)$$

5. “Policy” vs. “Plan”

Deterministic Search

$V=0.81$	$V=0.9$	$V=1$	
$V=0.73$		$V=0.9$	
$V=0.66$	$V=0.73$	$V=0.81$	$V=0.73$

Plan



5. “Policy” vs. “Plan”

Deterministic Search

V=0.81	V=0.9	V=1	
V=0.73		V=0.9	
V=0.66	V=0.73	V=0.81	V=0.73

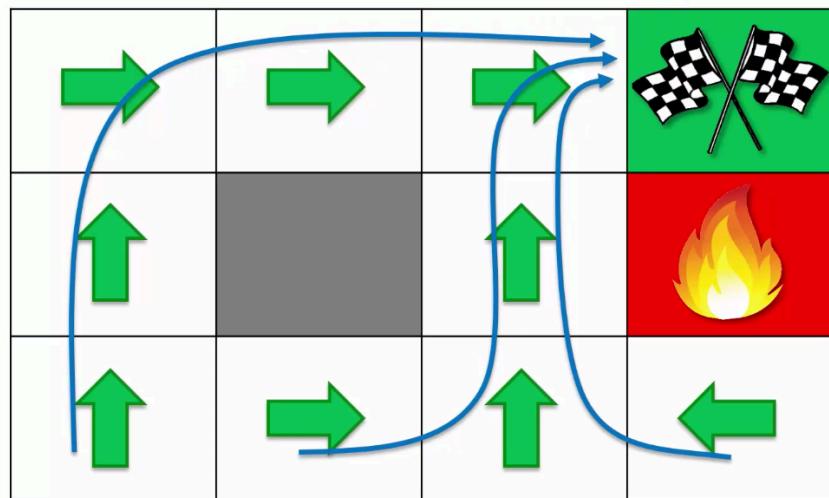
Policy

V=0.71	V=0.74	V=0.86	
V=0.63		V=0.39	
V=0.55	V=0.46	V=0.36	V=0.22

5. “Policy” vs. “Plan”

Plan

$V=0.81$	$V=0.9$	$V=1$	
$V=0.73$		$V=0.9$	
$V=0.66$	$V=0.73$	$V=0.81$	$V=0.73$



Policy

$V=0.71$	$V=0.74$	$V=0.86$	
$V=0.63$		$V=0.39$	
$V=0.55$	$V=0.46$	$V=0.36$	$V=0.22$

