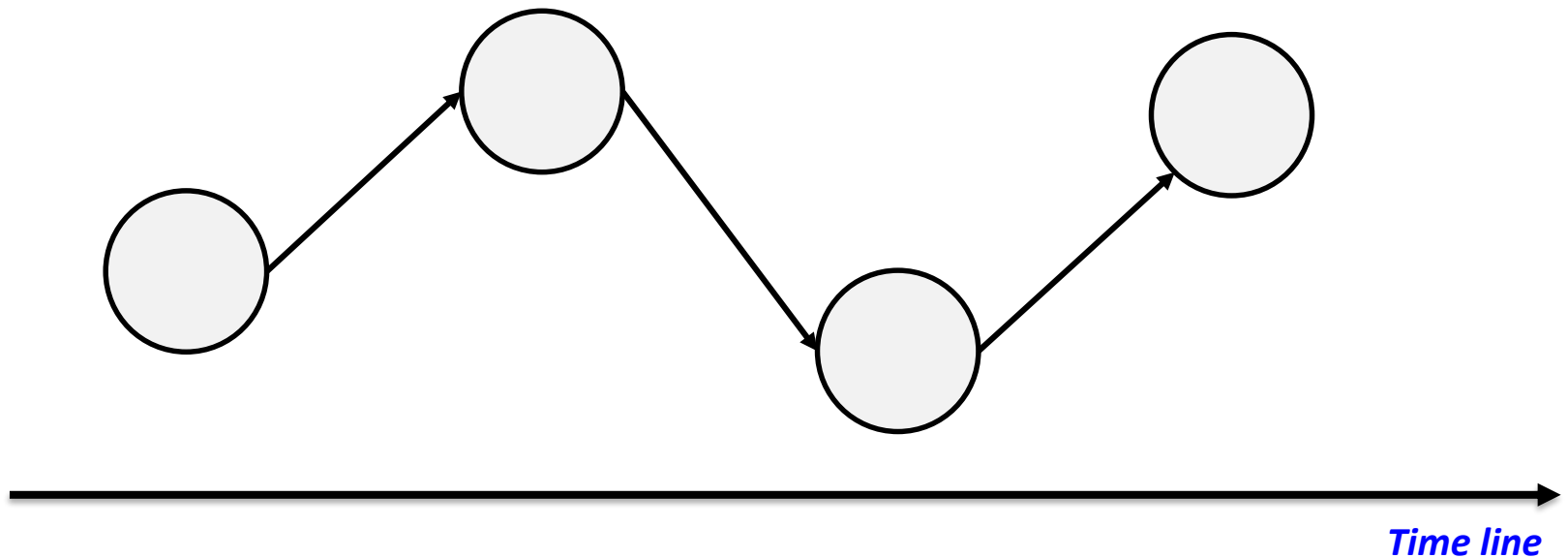


Sequential Decision Making

Sangkeun Jung

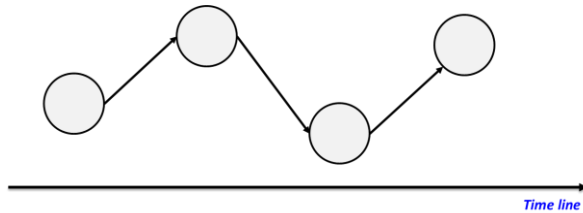
Big Picture



Sequential Decision Making

In artificial intelligence, sequential decision making refers to algorithms that take the dynamics of the world into consideration.

Wide Coverage



[Natural Language Processing]

- Word sequences
- Dialog : Utterance – utterance sequences

[Mechanics]

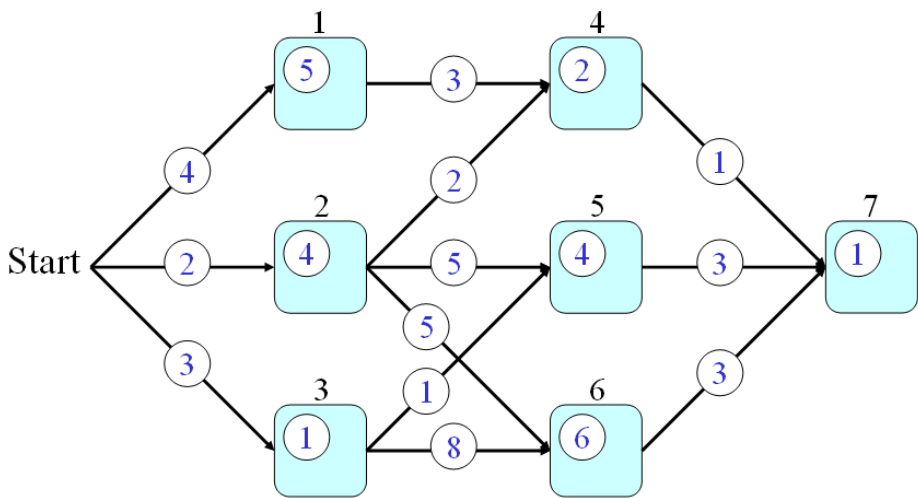
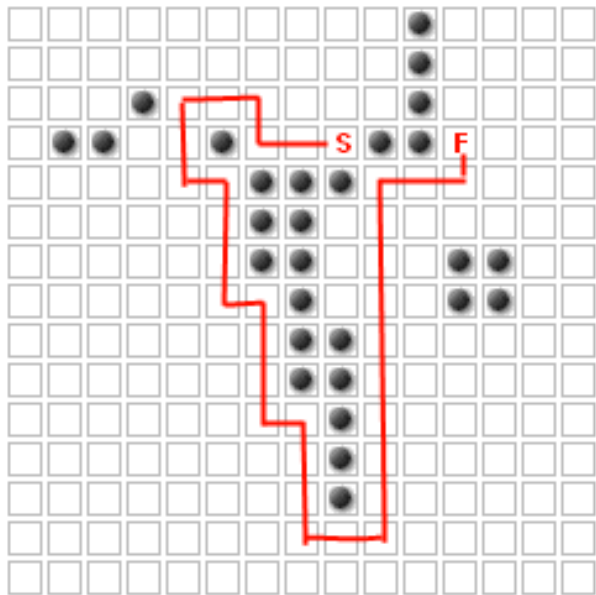
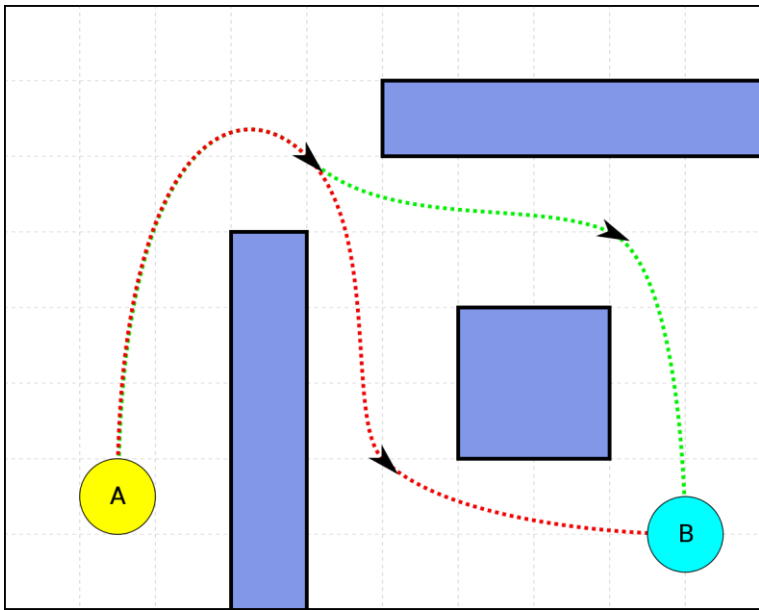
- Robot movements
- Task completion

[Industry Engineering]

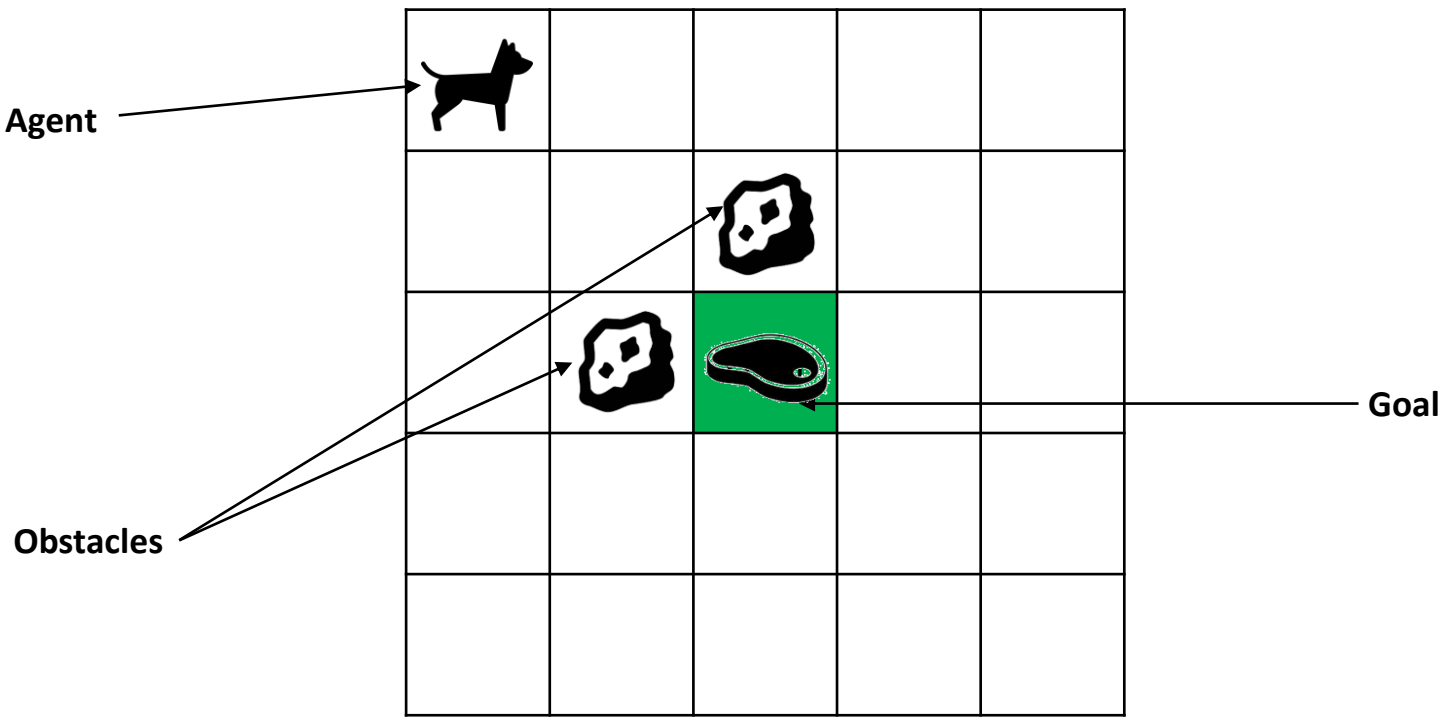
- Production management

....

Typical sequential decision making | Path Finding



Let's start simplest problem | Grid



How to Solve?

How about this?



가장 밝게 빛나는 길만 밟고 가자!

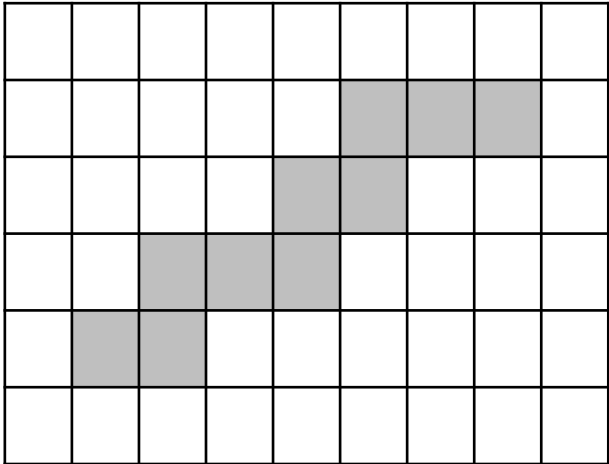


어떻게 길을 빛나게 만들 수 있을까?

Value

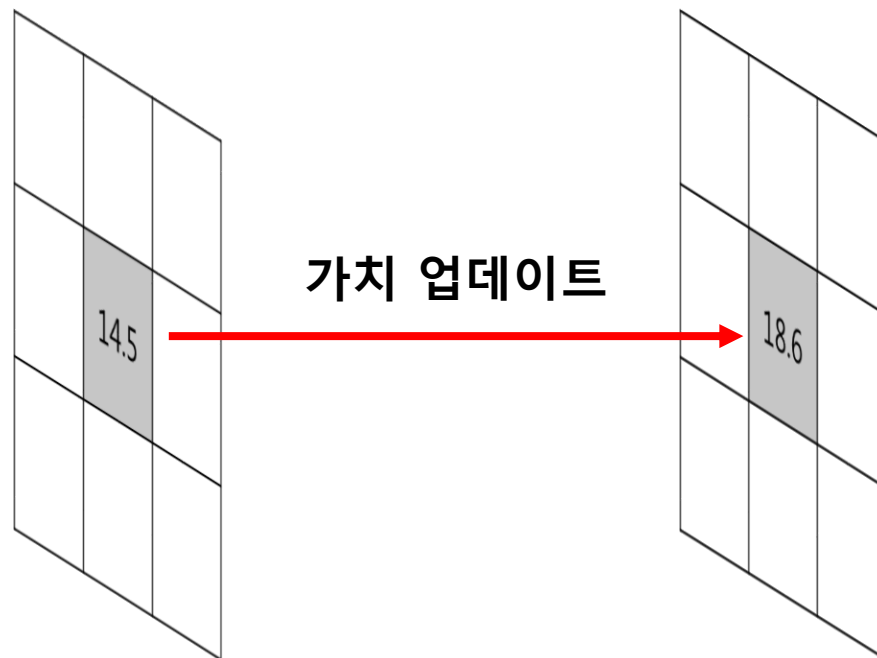
17.3

블록 하나 마다
값(밝기) 를 부여 해 준다.



[Question]
How to find good values?

Value Update

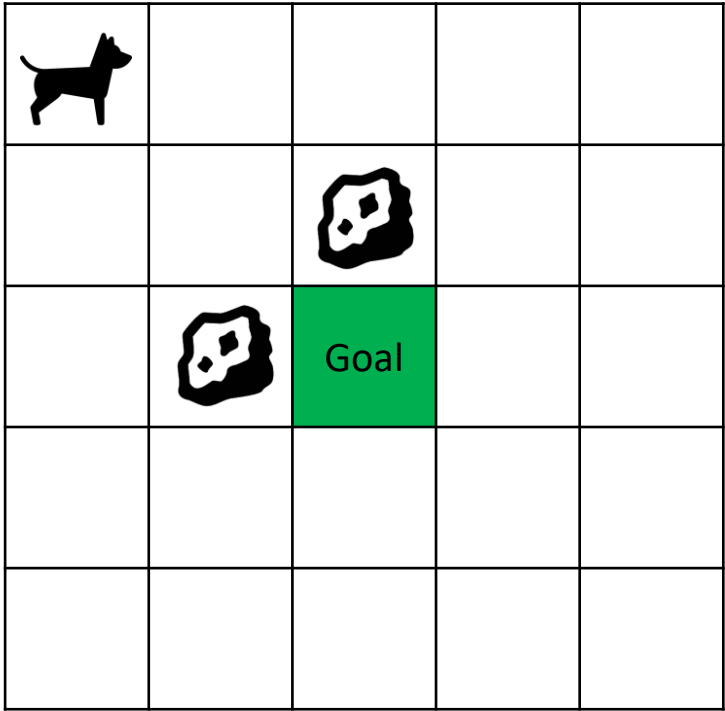


현재 (땅의) 가치

보다 그럴듯한 (땅의) 가치

- 최초에는 땅의 가치를 모르기 때문에(동등, uniform)
- 목표점까지 갈 수 있는 땅들이 높은 값을 가질 수 있게
- 가치를 업데이트 할 수 있는 방법이 필요하다.

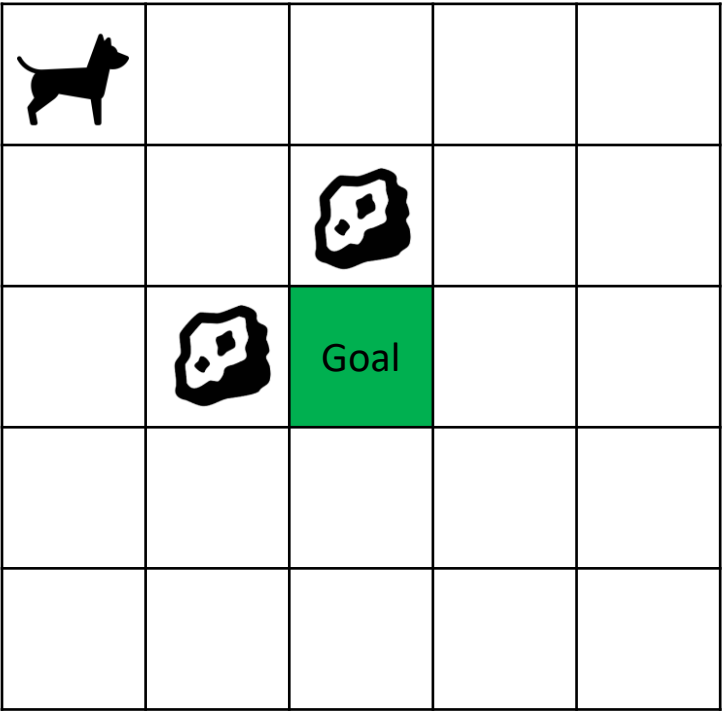
How to update value?



Initial Value

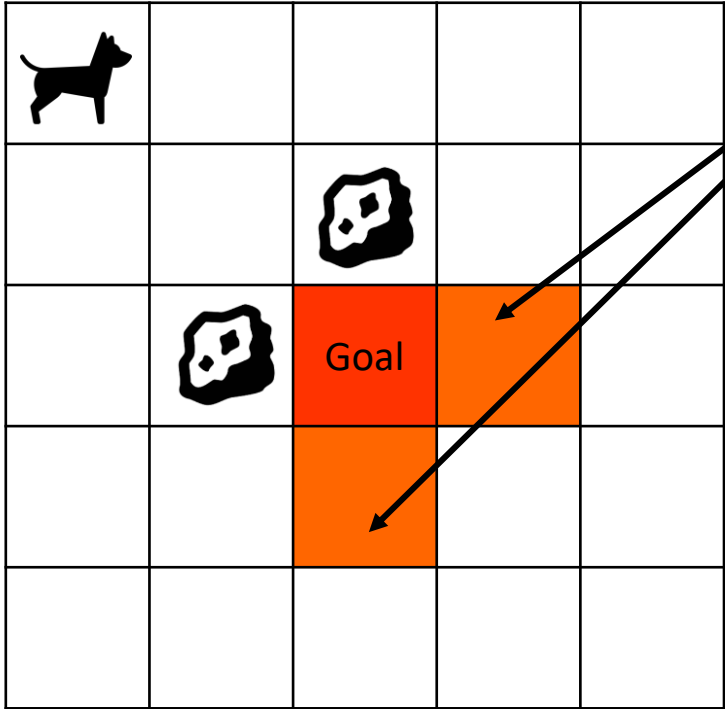
How update value?

Final Value 를 먼저 상상해보고
어떻게 얻을 수 있을지 생각



Final Value

How update value?



이 두 땅의 가치는 매우 높아야 함
(그러나 Goal 보다는 낮아야 함)






이 두 땅의 가치는 Goal 에 도착할 수 있기 때문



그렇다면, 저 두 땅에만 도착할 수 있다면
Goal 에 도착할 가능성이 올라감

How update value?

				
			*	
		Goal	Goal'	*
	*		*	
		*		

* 표시를 한 땅은 가치가 다른 땅보다 높아야 함
(그러나 Goal' 보다는 낮아야 함)






* 땅을 거쳐야만, Goal' 에 도착할 수 있기 때문



(반복) * 땅에 도착할 수 있다면
Goal 에 도착할 가능성이 올라감

How update value?

			*	
			G''	*
		Goal		G''
*	G''		G''	*
	*	G''	*	

$G'' = \text{Goal''}$

* 표시를 한 땅은 가치가 다른 땅보다 높아야 함
(그러나 Goal'' 보다는 낮아야 함)






* 땅을 거쳐야만, Goal'' 에 도착할 수 있기 때문



(반복) * 땅에 도착할 수 있다면
Goal 에 도착할 가능성이 올라감

How update value?

		*	G'''	*
				G'''
*		Goal		
G'''				G'''
*	G'''		G'''	*

$G''' = \text{Goal}'''$

* 표시를 한 땅은 가치가 다른 땅보다 높아야 함
(그러나 Goal''' 보다는 낮아야 함)






* 땅을 거쳐야만, Goal''' 에 도착할 수 있기 때문



(반복) * 땅에 도착할 수 있다면
Goal 에 도착할 가능성이 올라감

How update value? (final)

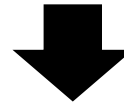
	*	G''''		G''''
*				
G''''		Goal		
G''''				G''''

$G'''' = \text{Goal}''''$

* 표시를 한 땅은 가치가 다른 땅보다 높아야 함
(그러나 Goal'''' 보다는 낮아야 함)

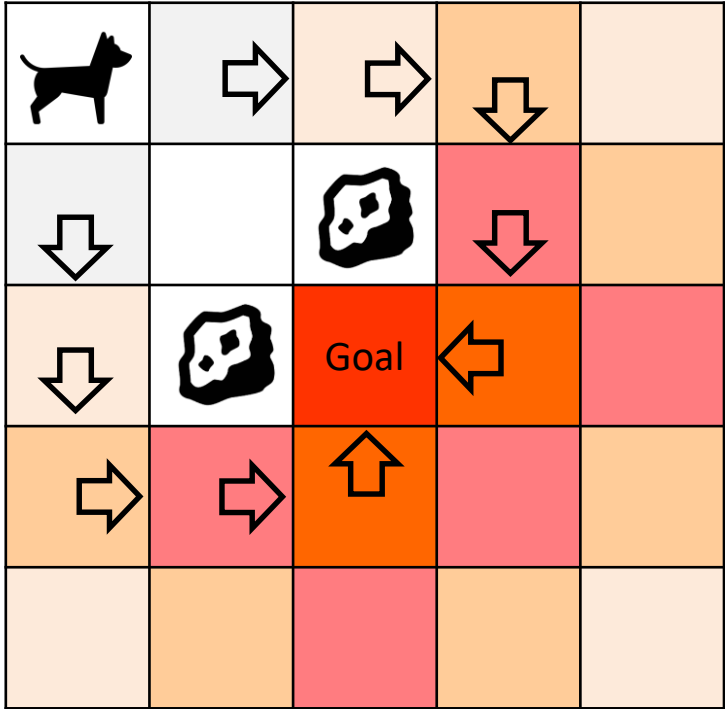


* 땅을 거쳐야만, Goal'''' 에 도착할 수 있기 때문



(반복) * 땅에 도착할 수 있다면
Goal 에 도착할 가능성이 올라감

Path Finding



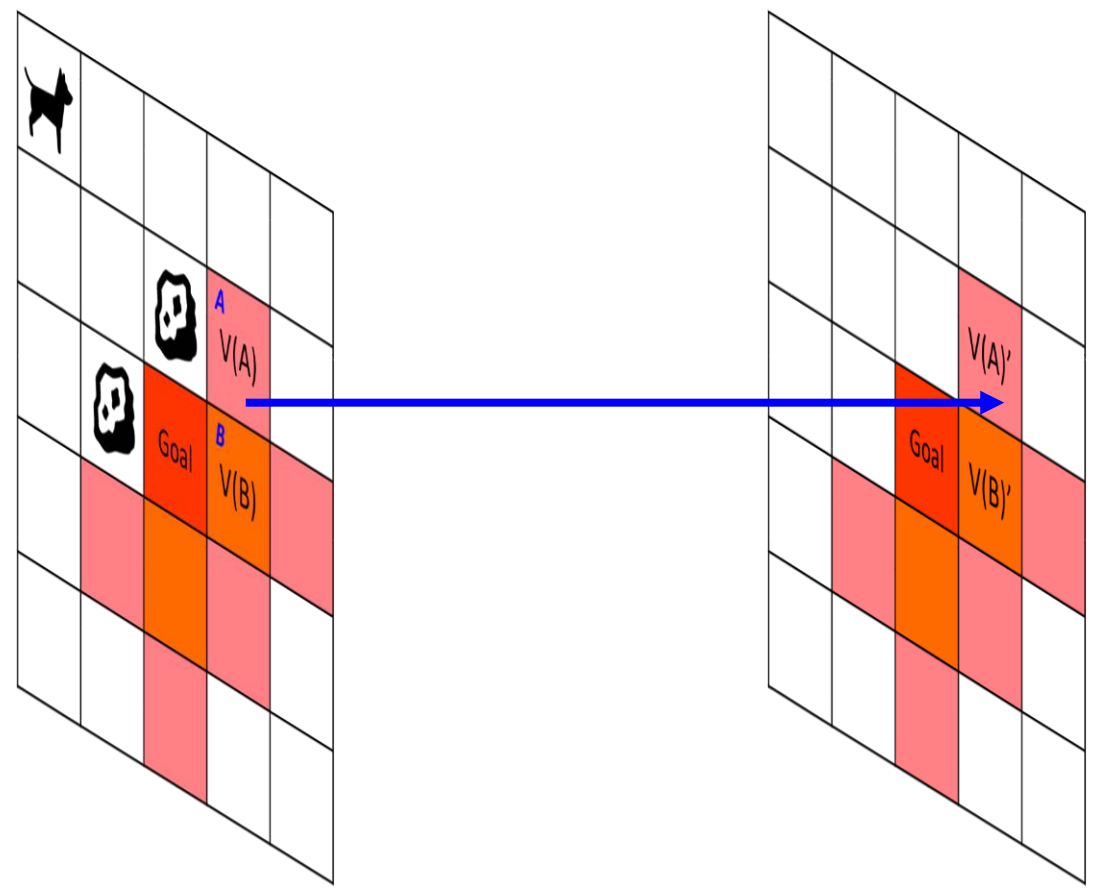
다음 걸음을 내딛을 때
가장 높은 땅의 가치를 가진 곳으로만 간다.

2 개의 최적 Path 가 나온다.

[Question – 10 min]
Make above idea an algorithm

Value Update | $V(A) \rightarrow V(A)'$

A,B는 각각
땅의 이름



땅 A에서 갈 수 있는 모든 땅 중,
가장 가치가 높은 땅의 가치를,
땅 A의 가치로 Update



$$V(A)' = \gamma V(B)$$



그러나, 땅 B 보다는 조금 가치가 낮게

Discount factor: 0.0 ~ 1.0

Coding

SIMPLE VALUE UPDATE METHOD

Implementation Environment

- Anaconda
 - Python 3.6
- OS
 - Windows or Ubuntu
 - (Mac)
- IDE
 - It doesn't matter
 - Recommend
 - Visual Studio
 - Visual Studio code
 - PyCharm

Implementation

- Implement
 - `def update_value(self)`
- Useful Methods

Method or Variables	Description
<code>self.env.possible_actions()</code>	return all possible actions in current state
<code>self.env.state_after_action(state, action)</code>	retrieve next state
<code>self.get_value(state)</code>	return value of the state
<code>self.discount_factor</code>	discount factor

Download code

<https://bit.ly/2m7jdEM>