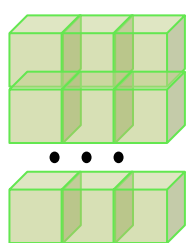


MMORPG NPC 강화 학습 개념과 설계 방법 구상

Joanne



state-mine

state-p1

state-p2

state-last
action

state-p1&p
2

state-last
action

•
•
•

state-map

encoder network

encoder network

encoder network

encoder network

encoder network

encoder network

행동 선택기
(공격 or 방어?)

아
요

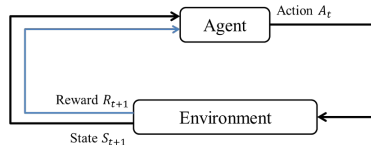
공격 skill 선택기

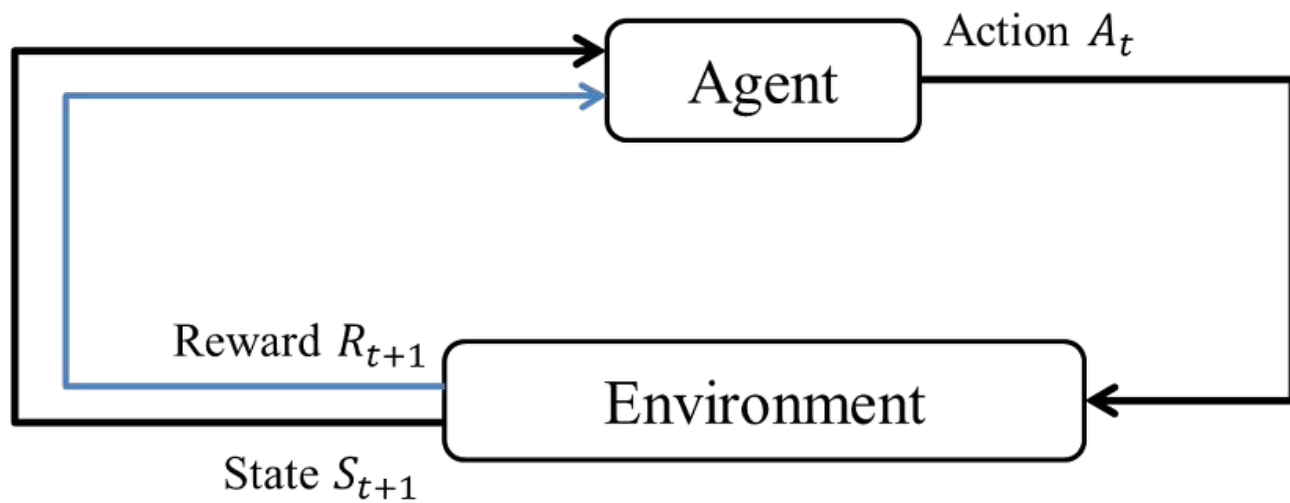
공격 대상 **P** 선택기

이동 위치/거리 선택기

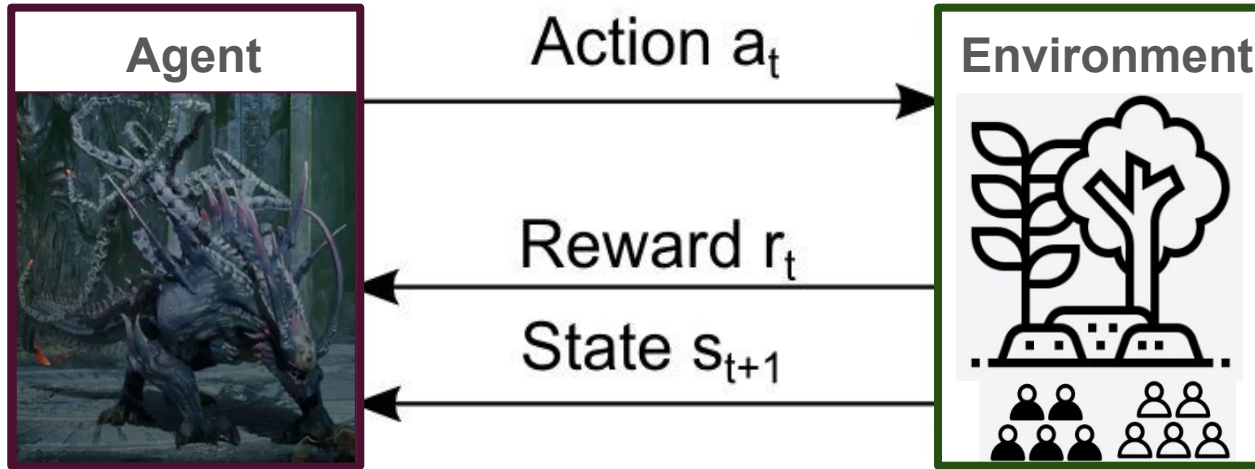
출력 행동 변환

계층적 행동 선택





Boss as an Agent



Agent acts with Goals

Environment rewards Agent

Agent observe state

OpenAI Gym 강화학습 알고리즘 개발 도구

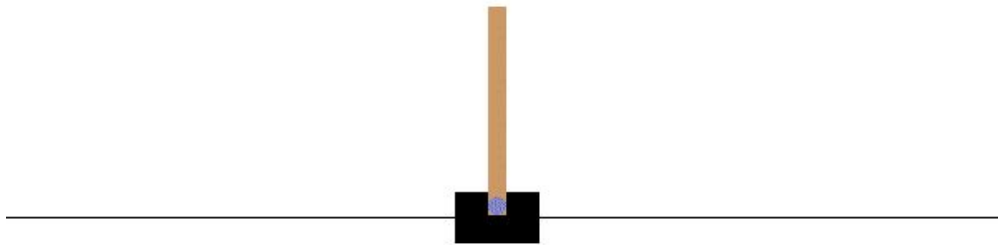
prerequisites: Anaconda3 (python >3.6), pip 라이브러리 (openAI Gym, open cv)

cartpole 예제:

딥러닝을 강화학습에 응용한 알고리즘들을 적용해보기

- Deep Q Network(DQN)
- [Actor Critic \(A2C\)](#)
- [Continuous Actor Critic \(A2C\)](#)

Num	Observation	Min	Max
0	Cart Position	-2.4	2.4
1	Cart Velocity	-Inf	Inf
2	Pole Angle	~ -0.418 rad (-24°)	~ 0.418 rad (24°)
3	Pole Velocity At Tip	-Inf	Inf



OpenAI Gym(2014) 강화학습 알고리즘 개발 도구

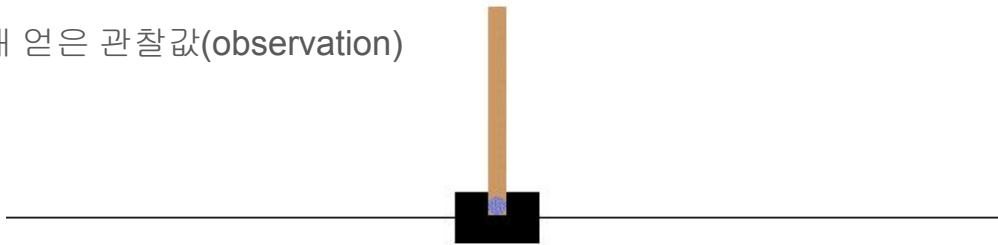
cartpole 예제:

```
import gym
env = gym.make('CartPole-v0') env.reset()
for _ in range(1000):
    env.render() env.step(env.action_space.sample())
env.close()
```

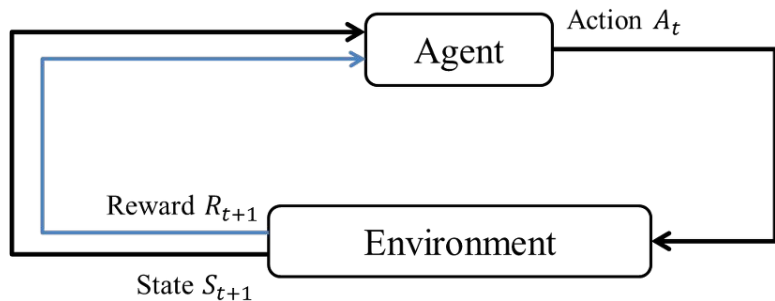
`env.reset()` : 새로운 에피소드(initial environment)를 불러옴

`env.render()` : 행동(action)을 취하기 이전에 환경에 대해 얻은 관찰값(observation)

`env.step()` : 행동(action)을 취한 이후에 환경에 대해 얻은 관찰값(observation)



MIR5 Boss 보상체계



1. 에이전트 (**Agent**): 학습하고 결정을 내리는 주체, **Boss Asterion**
2. 환경 (**Environment**): 에이전트가 상호작용하는 세계, **Mir 5 환경**
3. 상태 (**State**): 현재 상황에 대한 정보, 모든 플레이어 위치, 상태값, 스킬 정보, 히스토리
4. 행동 (**Action**): 에이전트가 취할 수 있는 선택들, 모든 플레이어 위치, 상태값, 스킬 정보, 히스토리
행동: 사용가능한 모든 명령
5. 보상 (**Reward**): 각 행동에 대한 피드백, 보상: 플레이어를 kill했을때 +1, 데미지 받을때 -1

RL combined with deep neural networks

Creating Pro-Level AI for a Real-Time Fighting Game Using Deep Reinforcement Learning(2019)

<https://arxiv.org/pdf/1904.03821>

기존 RL+신경망은 고정된 환경, 1:1 turn-based 게임에서는 인간 수준을 초월
→ 방대한 액션, 액션 간의 상호 의존성, 불완전한 정보가 생기는 격투
게임에서는 한계

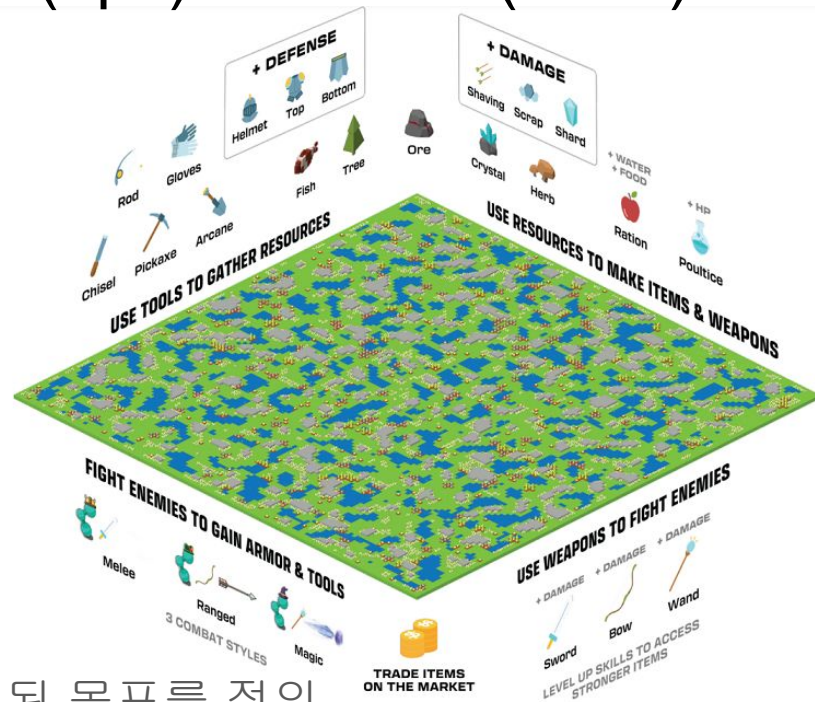


Figure 1. A scene from the B&S Arena Battle

블레이드 & 소울(ncsoft)에서 1대1 전투 AI 에이전트를 개발 사례

강화학습으로 훈련된 에이전트가 다섯 명의 프로 게이머와 대결하여 62%의
승률을 기록

Neural MMO 2.0: 멀티 에이전트(npc) 강화학습 (2023)



기존 강화학습 → 환경은 reward를 통해 고정된 목표를 정의

Neural MMO 2.0 → 사용자들이 에이전트별 또는 팀별 목표와 보상을 정의할 수 있는 유연한 태스크 시스템을 도입

Neural MMO 2.0 Flexible Tasks Systems



가상 세계에서 에이전트 집단을 시뮬레이션하는 오픈소스 연구 플랫폼

- 여러 NPC가 등장하는 SIMS 특성의 게임에 각 성격, 특정 목표를 부여하기 좋음
- NPC가 플레이어를 돕거나 보조적인 역할을 할 때 유용함

Neural MMO 2.0 Flexible Tasks Systems



Game State

Predicates

Tasks

A Game State

```
# True if any agent in subject can see a tile of tile_type
any(tile_type.index in t for t in subject.obs.tile.material_id)

# True if all subjects are alive.
np.count_nonzero(subject.health > 0) == len(subject)

# Computes the maximum l-inf distance between teammates
current_dist = max(subject.row.max()-subject.row.min(),
                    subject.col.max()-subject.col.min())

# Tracks hits scored with a specific combat style
hits = subject.event.SCORE_HIT.combat_style == combat_style.SKILL_ID

# Computes the summed gold of all teammates
subject.gold.sum()

# Evaluates to >= 1 if the current game tick is >= the specified tick
gs.current_tick / num_tick
```

MIR5 Boss' Game States

Unique ID	Description	Skill ID	Main Attack Stat	Next Attack ID	Movement Type	Movement Location	Movement Location Without	Movement Delay	Movement
ID	Note	ID_Skill	IsMainAttack	ID_NextAttack	SkillMoveType	MoveRange	EmptyTargetMoveRange	DelayMove	MoveTim
NoZero:true									
All	Refer	All	All	All	All	All	All	All	All
int32	string	int32	bool	int32	enum<SkillMoveType>	double	double	float	float
4020005	볼트 스킬6 낙뢰 떨어기토템 공격	402005	0	0	SkillMoveType_TargetMov	700	700	0.15	0.0
4090601	볼트 스킬6 낙뢰 떨어기토템 공격	40906	1	0	0	0	0	0	
4901001	볼트 구르기	49010	1	0	SkillMoveType_DirectMov	800	800	0.27	0.0
4901101	볼트 뒤 구르기	49011	1	0	SkillMoveType_DirectMov	-700	-700	0.27	0.0
90003301	알테리온 몬스터 정면 평타1	900033	1	0	0	0	0	0	
90003401	알테리온 몬스터 정면 평타2	900034	1	0	0	0	0	0	
90003501	알테리온 몬스터 후방 평타1	900035	1	0	0	0	0	0	
90003601	알테리온 몬스터 좌측 회전 평타1-1	900036	1	90003602	SkillMoveType_DirectMov	1000	1000	0.7	0.2
90003602	알테리온 몬스터 좌측 회전 평타1-2	900036	0	0	0	0	0	0	
90003701	알테리온 몬스터 우측 회전 평타1-1	900037	1	90003702	SkillMoveType_DirectMov	1000	1000	0.7	0.0
90003702	알테리온 몬스터 우측 회전 평타1-2	900037	0	0	0	0	0	0	
90003801	알테리온 몬스터 소용돌이	900038	1	90003802	0	0	0	0	
90003802	알테리온 몬스터 소용돌이	900038	0	90003803	0	0	0	0	
90003803	알테리온 몬스터 소용돌이	900038	0	90003804	0	0	0	0	
90003804	알테리온 몬스터 소용돌이	900038	0	90003805	0	0	0	0	
90003805	알테리온 몬스터 소용돌이	900038	0	90003806	0	0	0	0	
90003806	알테리온 몬스터 소용돌이	900038	0	90003807	0	0	0	0	
90003807	알테리온 몬스터 소용돌이	900038	0	0	0	0	0	0	
90003901	알테리온 몬스터 광포한 돌격	900039	1	90003902	SkillMoveType_DirectMov	2000	2000	0.85	2.8