



# LUX AI CHALLENGE 2

Presentation by Hyeonsu Nam



# 프로젝트 개요

☆ Lux AI 챌린지는 경쟁자가 다른 경쟁자와 1v1 시나리오에서 다중 변수 최적화, 리소스 수집 및 할당 문제를 해결하기 위해 에이전트를 설계하는 대회입니다. 최적화 외에도 성공적인 에이전트는 상대방을 분석하고 우위를 점하기 위해 적절한 정책을 개발할 수 있어야 합니다.

# 프로젝트 세부 사항

## ☆ 필요 지식

- Pytorch 코드 리뷰 관련 지식
- Gym Library
- SB3(DQN-MlpPolicy)
- SB3-Contrib(Maskable PPO)
- 강화학습 관련 전반적인 지식

# PROBLEM



- 처음 시도해보는 강화학습 관련 프로젝트였기에 전반적으로 관련 지식 부족
- 관련 지식이 부족하여 코드로 구현하는 것과 코드 리뷰하는 것에 많은 문제점 발생  
=> 강화학습 관련 기초 이론과 Basic 한 코드를 리뷰하는 시간을 가짐

# 참고 자료

- Youtube 혁펜하임(이론)
- CartPole based on Deep Q-Learning 코드 리뷰
- 마리오 강화학습 튜토리얼  
([https://github.com/wonseokjung/moduyeon\\_supermario](https://github.com/wonseokjung/moduyeon_supermario))
- Github 에 공유된 LUX AI Season 1 자료
- Kaggle 에 공유된 LUX AI Season 2 자료
  - Single-map-based-observation
  - Rule Based model
  - Lux AI S1 3rd code review

# RULE BASED AGENT IN PYTHON

```
# self.faction_names = 각 플레이어 이름에 해당하는 진영 이름을 나타냄
self.faction_names = {
    'player_0': 'TheBuilders',
    'player_1': 'FirstMars'
}

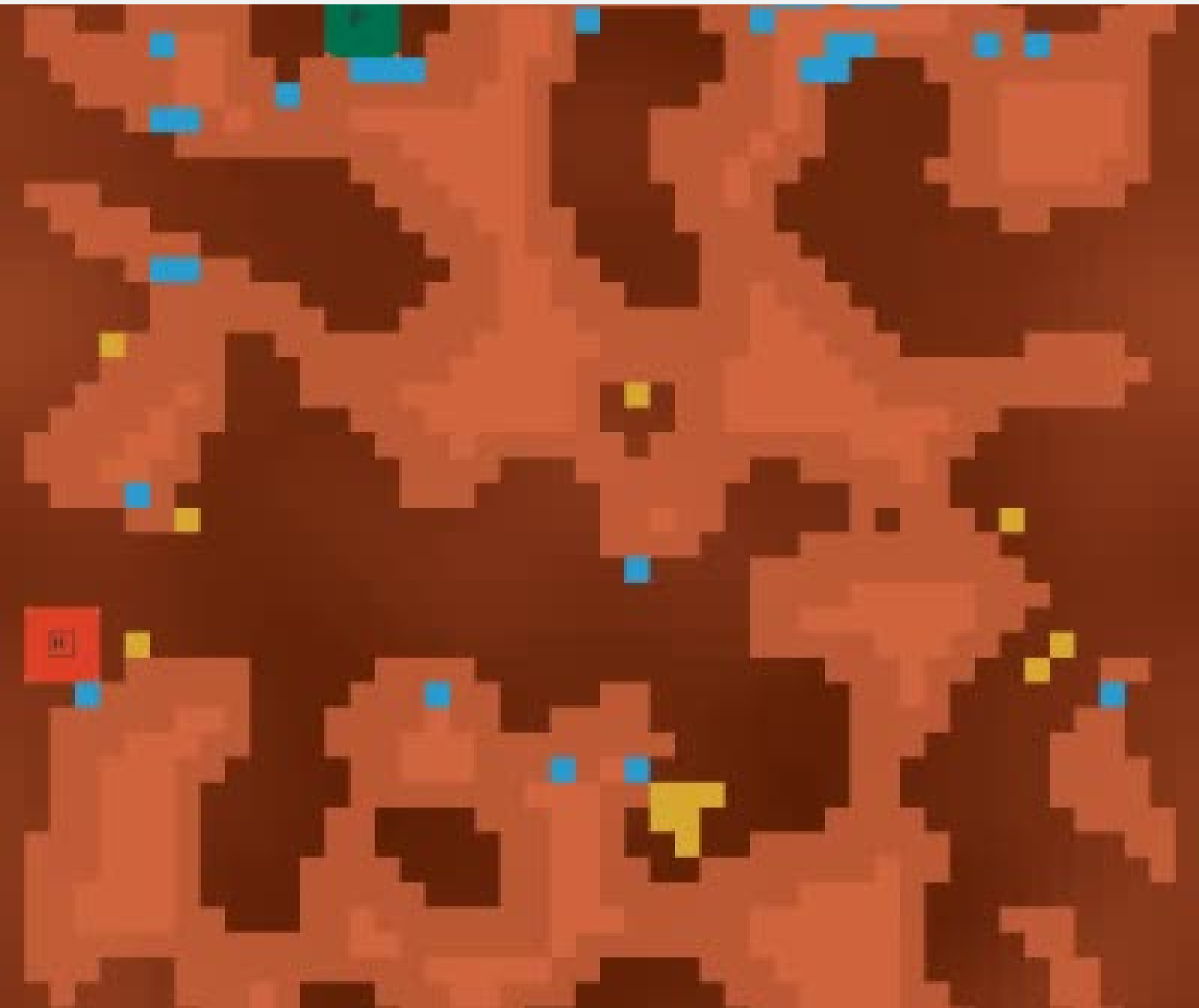
self.bots = {}
self.botpos = []
self.bot_factory = {}
self.factory_bots = {}
self.factory_queue = {}
self.move_deltas = np.array([[0, 0], [0, -1], [1, 0], [0, 1], [-1, 0]]) # Lux 게임에서 유닛들이 이동할 때 사용되는 방향 벡터들을 담고 있는 배열

# 게임의 초기 설정 단계 / 이 메소드는 'obs' = 현재 게임 상태 와 게임 플레이어가 할 수 있는 행동을 결정하는데 필요한 정보들을 사용하여 게임 플레이어의 행동을 반환
def early_setup(self, step: int, obs, remainingOverageTime: int = 60):
    ...

    Early Phase
    ...

    actions = dict() # 플레이어가 할 수 있는 행동을 담은 dic
    if step == 0:
        # Declare faction / 자신의 소속 진영과 입찰가를 설정
        actions['faction'] = self.faction_names[self.player]
        actions['bid'] = 0 # Learnable
    else:
        # Factory placement period / 게임 시작 후 => 공장을 배치하는 단계
        # optionally convert observations to python objects with utility functions
```

# SUBMITTING AN AGENT(TEST)



total_timesteps		20000	
train/			
approx_kl		7.31647e-09	
clip_fraction		0	
clip_range		0.2	
entropy_loss		-1.98	
explained_variance		0.942	
learning_rate		2.41e-05	
loss		-7.78e-07	
n_updates		8	
policy_gradient_loss		-1.28e-06	
value_loss		3.25e-05	
train_metrics/			
action_queue_updates_success		125	
action_queue_updates_total		196	
ice_dug		1.25	
metal_produced		0	
ore_dug		0	
water_produced		0.312	

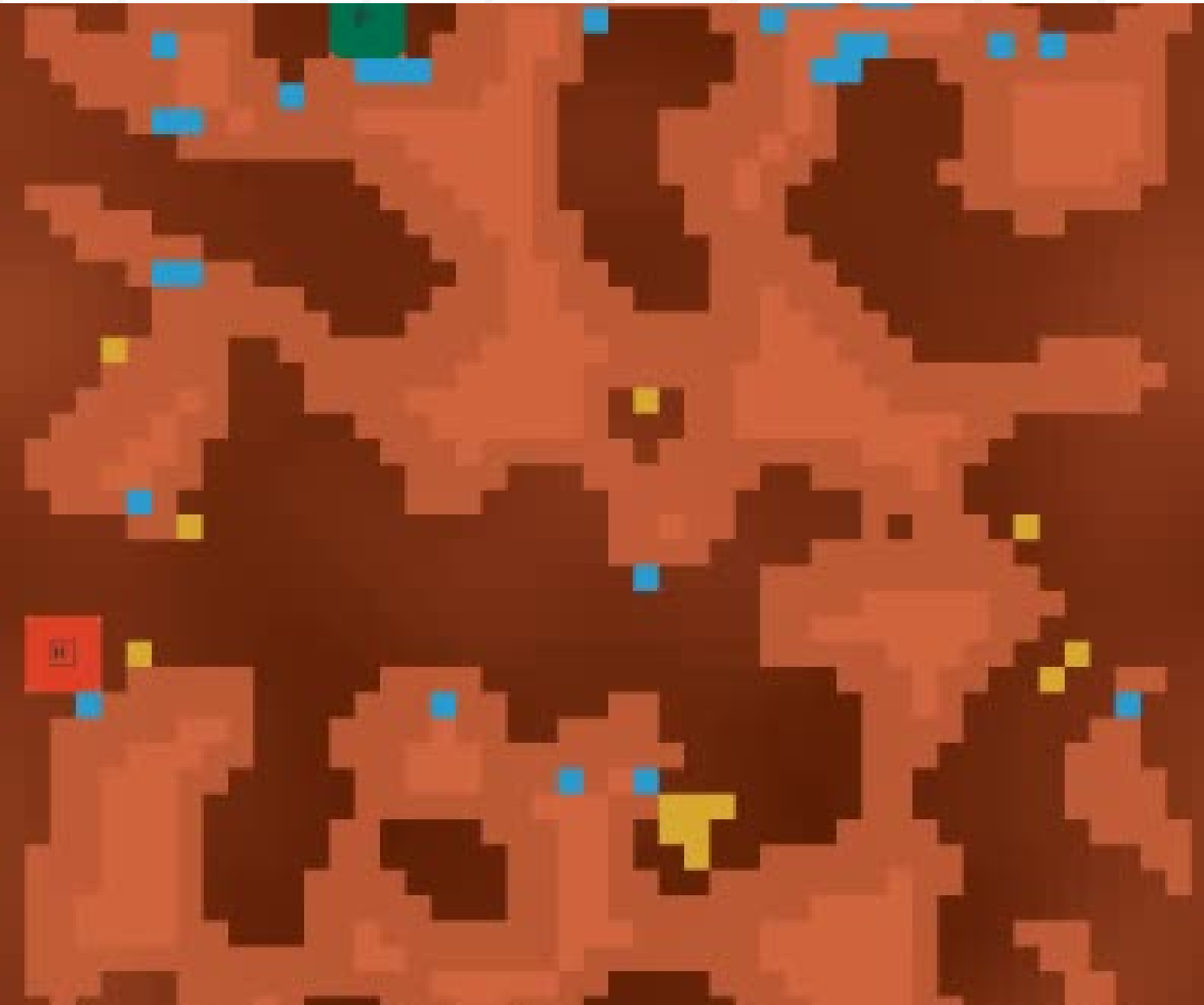
# PROBLEM



- **Heavy Robot 1개 생산**  
=> Light Robot 대량 생산하는 방향으로 코드 수정
- **Factory 1개 생산**  
=> 개선점 Factory 5개 생산하는 방향으로 코드 수정



# SUBMITTING AN AGENT(FINAL)



action_queue_updates_success	268
action_queue_updates_total	419
ice_dug	4
metal_produced	0
ore_dug	0
water_produced	1

rollout/	
ep_len_mean	591
ep_rew_mean	12.6
time/	
fps	249
iterations	2
time_elapsed	32
total_timesteps	8000
train/	
approx_kl	4.066527e-08
clip_fraction	0
clip_range	0.2
entropy_loss	-2.07
explained_variance	0.00157
learning_rate	2.59e-05
loss	1.3
n_updates	2
policy_gradient_loss	-6.06e-06

# FIRST ISSUE

- 두 종류의 로봇(Heavy, Light)을 생산할 수 있으나 한 개의 로봇만 조종이 가능함.
- 이러한 문제점으로 인해 두번째 이후로 생산되는 로봇은 조종이 불가능하여 새로운 Action을 부여해 주지 못하는 상황이 생김.
- ⇒ Gym Library document 에서 MultiDiscrete 개념과 스타크래프트 게임을 Python 코드로 만든 예제코드를 참고하여 코드 수정하려고 했으나 메소드 오버라이딩에서 계속된 에러가 생김으로 인해 최종적으로 코드를 완성하지 못함.

# SECOND ISSUE

- 공장이 한 개만 spawn되는 상황

→ 게임 내 최대 공장 허용치인 5개를 spawn 하기 위해 코드 수정이 필요.

# THIRD ISSUE

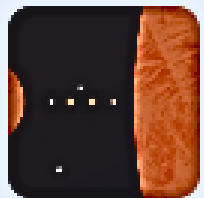
- 얼음과 광물을 같이 채굴하여 각각 물, 철을 생산해야하는데 기존 작성했던 코드로는 얼음만을 채굴하는 상황

→ 광물을 같이 채굴하여 철을 생산할 수 있는 코드 작성 필요

⇒ reward를 얼음을 캐을 때와 광물을 채굴했을 때 같이 줄 수 있도록 코드를 수정하였으나 ore ⇒ metal 정제하는 과정이 log상에서 누락되는 걸 확인

⇒ action\_queue 관련 문제인 것 같은데 아마도 Factory 관련 Action을 Specific 하게 정의하지 못한 게 문제이지 않을까 생각한다.

# 결과



**Lux AI Season 2**

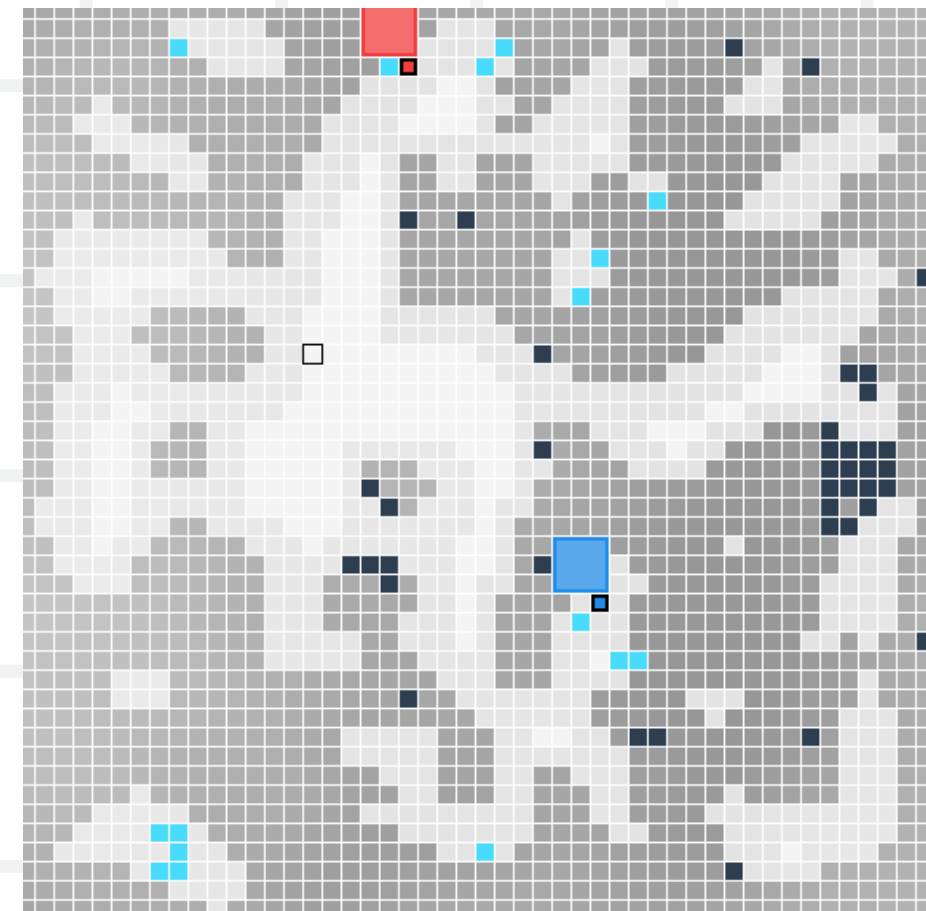
Terraform Mars!

Featured · Simulation Competition · 637 Teams

365/637



**최종 결과: 365등 / 637 팀**



Presentation by Hyeonsu Nam

# 결론

- 비교1 : 강화 학습이라는 것을 처음 접해 (파이썬+Gym Library) 코드로 구현을 하는 것에 많은 어려움을 겪었음.
- 비교2 : 강화 학습 관련 Gym Library, Pytorch 코드 리뷰 등 교육 과정에서 배우지 못한 내용들을 공모전 준비하면서 익히고 적용할 수 있었던 좋은 경험이었다.
- 비교3 : 본인을 제외한 팀원 3명의 추천으로 공모전에 참여하였으나 결국 시간이 지남에 따라 2주가 넘어가는 시점에 다들 포기하고 싶어했지만 전 끝까지 하자고 독하여 365등 결과를 만들어 냈.
- 결론 : 강화 학습에 대한 경험을 한 공모전이어서 좋았음.

# THANK YOU

Presentation by Hyeonsu Nam

