

# Starcraft2 Multi-agent

RL for micro-control of units

# Contents

- Overview
- Approach and Goal
- Architecture
- Basic Spec
- Current Status & Further Plan
- Distribution of project load
- Schedule

# Overview

- **SMAC**(The Starcraft Multi-Agent Challenge)
  - 좀 더 real world에 적용될만한 **multi-agent RL**을 연구하기 위해 제안된 benchmark problem
  - 각각의 unit들이 하나의 agent로써, 강화학습을 통해 게임 승리를 위한 행동이 어떤 것인지를 학습
- **pysc2**(Deepmind), **pymarl**(WhiRL) 등
- 전체 게임 X (너무 큰 규모) → unit들 간의 전투상황에 한정

# Approach and Goal

- **multi-agent 강화학습**에 대한 기본지식 습득
- 실험환경 setting
  - **SMAC** & **pyMARL** (python 환경)
- 기존에 개발되어있는 플랫폼 분석
  - 사용된 알고리즘(QMIX, COMA 등)에 대한 이해
  - 성능 평가 (각 전투상황 별 승률)
- **성능 개선**
  - 기존에 구현된 알고리즘의 한계점 및 개선할 점 파악
  - MARL 최신 연구들을 조사하여 새로운 알고리즘 구현
  - 라운드 별, 상황 별 기존 모델과 개발한 모델과의 scoring 비교

# Demo video

- [https://www.youtube.com/watch?v=VZ7zmQ\\_obZ0](https://www.youtube.com/watch?v=VZ7zmQ_obZ0)

# Basic Spec

- **Scenarios**

- **Circumstances**

- 각 agent는 **본인 시야범위** 내의 agent들만 관찰 가능
- 체력 상황, 공격력, unit type 등

- **Action**

- **move**(direction : N, S, E, W), **attack**(or heal)

- **Rewards**

- 기본적으로 전투 승리 시 **+1**, 패배 시 **-1**
- 가한 공격량 및 받은 피해량, enemy unit kill 횟수에 따라 **bonus**

The StarCraft Multi-Agent Challenge			
Name	Ally Units	Enemy Units	Type
3m	3 Marines	3 Marines	homogeneous & symmetric
8m	8 Marines	8 Marines	homogeneous & symmetric
25m	25 Marines	25 Marines	homogeneous & symmetric
2s3z	2 Stalkers & 3 Zealots	2 Stalkers & 3 Zealots	heterogeneous & symmetric
3s5z	3 Stalkers & 5 Zealots	3 Stalkers & 5 Zealots	heterogeneous & symmetric
MMM	1 Medivac, 2 Marauders & 7 Marines	1 Medivac, 2 Marauders & 7 Marines	heterogeneous & symmetric
5m.vs.6m	5 Marines	6 Marines	homogeneous & asymmetric
8m.vs.9m	8 Marines	9 Marines	homogeneous & asymmetric
10m.vs.11m	10 Marines	11 Marines	homogeneous & asymmetric
27m.vs.30m	27 Marines	30 Marines	homogeneous & asymmetric
3s5z.vs.3s6z	3 Stalkers & 5 Zealots	3 Stalkers & 6 Zealots	heterogeneous & asymmetric
MMM2	1 Medivac, 2 Marauders & 7 Marines	1 Medivac, 3 Marauders & 8 Marines	heterogeneous & asymmetric
2m.vs.1z	2 Marines	1 Zealot	micro-trick: alternating fire
2s.vs.1sc	2 Stalkers	1 Spine Crawler	micro-trick: alternating fire
3s.vs.3z	3 Stalkers	3 Zealots	micro-trick: kiting
3s.vs.4z	3 Stalkers	4 Zealots	micro-trick: kiting
3s.vs.5z	3 Stalkers	5 Zealots	micro-trick: kiting
6h.vs.8z	6 Hydralisks	8 Zealots	micro-trick: focus fire
bane.vs.bane	20 Zerglings & 4 Banelings	20 Zerglings & 4 Banelings	micro-trick: positioning
so_many_banelings	7 Zealots	32 Banelings	micro-trick: positioning
2c.vs.64zg	2 Colossi	64 Zerglings	micro-trick: positioning
corridor	6 Zealots	24 Zerglings	micro-trick: wall off

# Basic Spec(cont.)

- **Specifying scenario**

- 주어진 시나리오들을 기존의 알고리즘의 승률에 따라 3단계로 구분
- 기존의 알고리즘 중 가장 강력했던 QMIX 알고리즘의 median 값 +  $\alpha$  를 목표치로 설정

- **Goal**

- Easy 의 경우 +5%를 목표
- Hard 의 경우 +3%를 목표
- Super hard 의 경우 우선 문제 해결을 목표(10 mil try 내 승률 5% 달성)

- **시간이 남는다면...**

- 기존의 QMIX 알고리즘과 개발한 알고리즘 간의 모의 전투 시행 및 분석

# Schedule

## 일정

주차	일정	분업	특징
4월 2째주	시나리오 1 (3m, 8m, 2s3z, MMM), 시나리오3 (2m vs 1z, 2s vs 1sc, 3s vs 3z, bane vs bane, so many panelings) Median 100% 달성	이호준 - (3m, 8m, 2s3z) 이승석 - (MMM, 2m vs 1z, 2s vs 1sc) 박충훈 - (3s vs 3z, bane vs bane, so many panelings)	
4월 3째주	시나리오 1 (25m) 기존 알고리즘의 Median값 + 5% 달성, 현재 : 72% -> 목표 : 77%		
4월 4째주	시나리오 2 (5m vs 6m) 기존 알고리즘의 Median값 + 5% 달성, 현재 : 58% -> 목표 : 63%		
4월 5째주 / 5월 1째주	시나리오 2 (8m vs 9m) 기존 알고리즘의 Median값 + 5% 달성, 현재 : 81% -> 목표 : 86%		중간 발표
5월 2째주	시나리오 2 (10m vs 11m) 기존 알고리즘의 Median값 + 5% 달성, 현재 : 83% -> 목표 : 88%		
5월 3째주	시나리오 3 (3s vs 4z) 기존 알고리즘의 Median값 + 3% 달성, 현재 : 95% -> 목표 : 98%		
5월 4째주	시나리오 3 (2c vs 64 zg) 기존 알고리즘의 Median값 + 3% 달성, 현재 : 62% -> 목표 : 67%		
5월 5째주 / 6월 1째주	시나리오 3 (3s vs 5z) 기존 알고리즘의 Median값 + 3% 달성, 현재 : 31% -> 목표 : 36%		
6월 2째주	hard task (3s5z, 27m vs 30m, MMM2)		
6월 3째주	challenging task (3s5z vs 3s6z, 6h vs 8z, corridor)		기말 발표



# Current Status & Further Plan

- **Current**

- 실험 환경 setting
- 플랫폼 document 및 관련 논문 조사

- **Plan**

- 플랫폼의 실행과 결과 녹화 및 분석
- 승률을 높이기 위한 알고리즘 수정
- Rewards 시스템 변경을 통해 평가방법 다변화
- (if possible)알고리즘 간 비교를 위한 대전 시스템 개발