

Pokemon_RL

https://github.com/IMJeongseon/pokemon_RL

피카츄 / 120250351 조정선

■ 주제 선정 배경:

- 포켓몬 배틀은 턴제 전략 게임으로, 불완전한 정보(상대방의 남은 기술, 노력치 등)와 방대한 State Space를 가지고 있어 강화학습을 적용하기에 매우 도전적이고 흥미로운 환경임.
- pokemon-env와 pokemon-showdown을 활용하여 새로운 데이터와 환경을 이용한 강화학습 프레임워크 디자인 을 목표로 함.
- Showdown의 BSS Reg I 규칙을 기준으로 해서 배틀 에이전트를 생성 (2마리의 전설의 포켓몬 사용 가능)
이는 실제 포켓몬 게임에서 현 시각 (2025.12.06) 기준 공식으로 사용하는 룰임

■ 최적화 목표:

- 에이전트가 포켓몬 배틀에서 승률을 극대화하는 최적의 전략(Policy)을 학습하는 것
- 단순한 데미지 계산을 넘어, 교체 플레이나 상태 이상 활용 등 전략적 판단을 수행하도록 유도함.



■ 환경(Environment)

- **Simulator:** Pokemon-Showdown (오픈소스 포켓몬 배틀 시뮬레이터)
- **Interface:** pokemon-env (Gymnasium/Gym interface 호환)
- pokemon-env: <https://github.com/hsahovic/poke-env>
- pokemon-showdown: <https://github.com/smogon/pokemon-showdown>

■ 데이터 전처리 (Preprocessing):

- **입력 데이터:** 배틀 로그(Text/JSON) 및 현재 화면 정보.
 - **Feature Engineering:**
 - 아군/적군 포켓몬의 정보 (HP, Type, Status Condition, Stat boosts).
 - 날씨(Weather) 및 필드(Terrain) 상태.
 - 기술(Move) 정보 (Type, Power, Accuracy).
-

▪ Active Pkm

- 나의 행동 + 상대방의 행동 (중요 정보: HP 비율, HP/Atk/Def/SpA/SpD/Spe, 랭크상태 등)
- On-hot encoding으로 인한 다른 것들 (도구 소유 여부, 테라스탈, 기절, 졸음, 상태이상 등)

▪ Move Information

- 4가지 기술 × 4가지 정보 (기술 위력, 명중, PP, 타입상성배수)

▪ Team & Bench Information

- Team Summary: 생존포켓몬/기절포켓몬 비율 + 아군/적군 수적 우위 등
- Bench Details: 벤치에 있는 포켓몬의 HP, 상태이상, 기절여부 등

▪ Global Information

- 현재 턴 수, 날씨, 필드, 함정, 벽, 순풍 등

▪ Goal

- 나와 상대방의 상태에 대한 정보를 최대한 많이 반영하도록
 - 현재 턴의 배틀 상황 뿐만 아니라 벤치상황, 날씨, 장판 등을 모두 포함하여 Long-term Strategy 수립 가능
-

▪ Action Space

- Type: Discrete Action Space (Size = 10)
- Moves (0~3): 기본 기술 사용
- Tera + Moves (4~7): 테라스탈 발동 후 기술 사용
- Switching (8~9): bench의 포켓몬으로 교체 (Slot 0, 1)

▪ Action Masking

- 기본 규칙 마스킹:
 - 현재 사용이 불가능한 기술 (구매 도구 소유, PP 0, 봉인, 트집 등)
 - 테라스탈을 이미 발동함 → 더이상의 테라스탈은 불가능 : 테라스탈 마스킹
 - 교체 마스킹: 포켓몬이 쓰러졌을 경우
- 휴리스틱 마스킹:
 - 중복 설치 방지 (스텔스룩이 깔려있는데 또 스텔스룩을 쓰지 않도록)
 - 중복 상태이상 방지 (상대방이 마비 상태인데 상태이상을 또다시 걸지 않도록)
→ Search Space를 효율적으로 줄여 불필요한 행동을 하지 않도록 방지함

■ Core Objectives

- Terminal Reward: Win +100 / Lose -100
- 상대 포켓몬 KO 시: +5, 공격 시 데미지의 $\times 1.5$, 피격시 $\times 1.0$ → 방어보단 공격을 유도하도록

■ Strategic Incentives

- 장판기 설치시 reward 보상 → 장판플레이 유도
- Rank 상황에 따라서 reward를 보상. 단 랭크 상태에 따라서 효율이 달라지기 단계적으로 reward를 부여
 - 장판기, 랭크에 대해서 reward를 따로 관리 → 장판 초기화 및 랭크 초기화 시 장판/랭크에 대한 reward를 쉽게 초기화 할 수 있음
- 무효화 reward: 교체 및 테라스탈로 상대의 기술을 무효로 받아낼 시 reward 부여 → 교체나 테라스탈의 빈도 증가
- 하품 방지: “졸음”상태에서 교체시 reward 부여 → 하품 맞고 잠드는 것 방지

■ Penalties

- Turn Costs: 의미 없는 턴 끝기를 위한 매 턴 -0.05의 reward
- 효과가 없는 기술 사용 시 -2 penalty 부여 → 위력이 없는 기술을 사용하지 않도록
- Risky Rank boost: 체력이 없는 상황에서선 무리한 랭크업/다운보다는 공격을 유도
- Yawn sleep: 하품을 맞고 잠들었을 시 패널티 부여

$$R_{total} = R_{basic} + \sum (R_{strategic} + R_{tactic} + R_{penalty})$$

▪ Maskable PPO

- Library: sb3_contrib
- 포켓몬 배틀은 상황에 따라 불가능한 행동(기절한 포켓몬 교체, PP 없는 기술 등)이 빈번함
- 일반 PPO는 불가능한 행동을 선택하고 패널티를 받으며 배우지만,
Maskable PPO는 신경망 단계에서 invalid 행동의 확률을 0 으로 만들어 **Valid Action만 학습**하도록 보장함.

▪ Network

- Input Layer: 164차원 Observation Vector
- Shared Hidden Layers: MLP
 - Structure: [256, 256] neurons
 - Activation Functions: tanh (SB3 Default)
- Output Heads: 10차원 discrete actions (with Masking)

▪ Training Strategy

- Self-Play Mechanism:
 - 고정된 상대가 아닌 “과거의 나”를 상대로 학습
 - Update Frequency: 매 10,000 step마다 상대 모델을 최신 버전으로 갱신
 - Effect: 학습이 진행될수록 상대도 강해지기에, 에이전트가 점진적으로 고도화된 전략을 습득함
-







Hyper Parameter


8

Parameter	Value	Description
Total Timesteps	500,000	총 학습 스텝 수
Learning Rate	2e-4	학습률
Batch Size	64	업데이트 배치 크기
n_steps (Rollout)	2,048	한 번의 업데이트를 위해 수집하는 step의 길이
Gamma	0.99	할인율 (미래 보상의 가치)
Entropy Coef	0.01	탐험(Exploration) 유도를 위한 엔트로피 계수


9

- **My Team**


<div> <div>Nickname</div> <div>Koraidon</div> </div> <div>  <div> <div>Pokémon</div> <div>Koraidon</div> </div> </div>	<div> <div>Copy</div> <div>Import/Export</div> <div>Move</div> <div>Delete</div> </div> <div> <div>Details</div> <div> <div>Level</div> <div>50</div> <div>Gender</div> <div>—</div> <div>Shiny</div> <div>No</div> <div>Tera Type</div> <div>Fire</div> </div> <div> <div>Item</div> <div>Choice Band</div> </div> <div> <div>Ability</div> <div>Orichalcum Pulse</div> </div> </div> <div> <div>Moves</div> <div> <div>Flare Blitz</div> <div>Close Combat</div> <div>Flame Charge</div> <div>U-turn</div> </div> </div> <div> <div>Stats</div> <div> <div>HP</div> <div>Atk</div> <div>Def</div> <div>SpA</div> <div>SpD</div> <div>Spe</div> <div>EV</div> <div>4</div> <div>252</div> <div>—</div> <div>252+</div> </div> </div>	<div> <div>Nickname</div> <div>Chien-Pao</div> </div> <div>  <div> <div>Pokémon</div> <div>Chien-Pao</div> </div> </div>	<div> <div>Copy</div> <div>Import/Export</div> <div>Move</div> <div>Delete</div> </div> <div> <div>Details</div> <div> <div>Level</div> <div>50</div> <div>Gender</div> <div>—</div> <div>Shiny</div> <div>No</div> <div>Tera Type</div> <div>Electric</div> </div> <div> <div>Item</div> <div>Focus Sash</div> </div> <div> <div>Ability</div> <div>Sword of Ruin</div> </div> </div> <div> <div>Moves</div> <div> <div>Icicle Crash</div> <div>Swords Dance</div> <div>Sucker Punch</div> <div>Tera Blast</div> </div> </div> <div> <div>Stats</div> <div> <div>HP</div> <div>Atk</div> <div>Def</div> <div>SpA</div> <div>SpD</div> <div>Spe</div> <div>EV</div> <div>—</div> <div>252</div> <div>4</div> <div>—</div> <div>252+</div> </div> </div>
<div> <div>Nickname</div> <div>Calyrex</div> </div> <div>  <div> <div>Pokémon</div> <div>Calyrex-Shadow</div> </div> </div>	<div> <div>Copy</div> <div>Import/Export</div> <div>Move</div> <div>Delete</div> </div> <div> <div>Details</div> <div> <div>Level</div> <div>50</div> <div>Gender</div> <div>—</div> <div>Shiny</div> <div>No</div> <div>Tera Type</div> <div>Fairy</div> </div> <div> <div>Item</div> <div>Choice Scarf</div> </div> <div> <div>Ability</div> <div>As One (Spectrier)</div> </div> </div> <div> <div>Moves</div> <div> <div>Astral Barrage</div> <div>Trick</div> <div>Psychic</div> <div>Draining Kiss</div> </div> </div> <div> <div>Stats</div> <div> <div>HP</div> <div>Atk</div> <div>Def</div> <div>SpA</div> <div>SpD</div> <div>Spe</div> <div>EV</div> <div>140</div> <div>—</div> <div>12</div> <div>196</div> <div>4</div> <div>156+</div> </div> </div>	<div> <div>Nickname</div> <div>Dondozo</div> </div> <div>  <div> <div>Pokémon</div> <div>Dondozo</div> </div> </div>	<div> <div>Copy</div> <div>Import/Export</div> <div>Move</div> <div>Delete</div> </div> <div> <div>Details</div> <div> <div>Level</div> <div>50</div> <div>Gender</div> <div>Female</div> <div>Shiny</div> <div>No</div> <div>Tera Type</div> <div>Fairy</div> </div> <div> <div>Item</div> <div>Covert Cloak</div> </div> <div> <div>Ability</div> <div>Unaware</div> </div> </div> <div> <div>Moves</div> <div> <div>Curse</div> <div>Rest</div> <div>Wave Crash</div> <div>Fissure</div> </div> </div> <div> <div>Stats</div> <div> <div>HP</div> <div>Atk</div> <div>Def</div> <div>SpA</div> <div>SpD</div> <div>Spe</div> <div>EV</div> <div>212</div> <div>—</div> <div>252+</div> <div>—</div> <div>44</div> </div> </div>
<div> <div>Nickname</div> <div>Flutter Mane</div> </div> <div>  <div> <div>Pokémon</div> <div>Flutter Mane</div> </div> </div>	<div> <div>Copy</div> <div>Import/Export</div> <div>Move</div> <div>Delete</div> </div> <div> <div>Details</div> <div> <div>Level</div> <div>50</div> <div>Gender</div> <div>—</div> <div>Shiny</div> <div>No</div> <div>Tera Type</div> <div>Normal</div> </div> <div> <div>Item</div> <div>Choice Specs</div> </div> <div> <div>Ability</div> <div>Protosynthesis</div> </div> </div> <div> <div>Moves</div> <div> <div>Moonblast</div> <div>Perish Song</div> <div>Shadow Ball</div> <div>Mystical Fire</div> </div> </div> <div> <div>Stats</div> <div> <div>HP</div> <div>Atk</div> <div>Def</div> <div>SpA</div> <div>SpD</div> <div>Spe</div> <div>EV</div> <div>—</div> <div>—</div> <div>4</div> <div>252</div> <div>—</div> <div>252+</div> </div> </div>	<div> <div>Nickname</div> <div>Ting-Lu</div> </div> <div>  <div> <div>Pokémon</div> <div>Ting-Lu</div> </div> </div>	<div> <div>Copy</div> <div>Import/Export</div> <div>Move</div> <div>Delete</div> </div> <div> <div>Details</div> <div> <div>Level</div> <div>50</div> <div>Gender</div> <div>—</div> <div>Shiny</div> <div>No</div> <div>Tera Type</div> <div>Fire</div> </div> <div> <div>Item</div> <div>Assault Vest</div> </div> <div> <div>Ability</div> <div>Vessel of Ruin</div> </div> </div> <div> <div>Moves</div> <div> <div>Earthquake</div> <div>Ruination</div> <div>Heavy Slam</div> <div>Payback</div> </div> </div> <div> <div>Stats</div> <div> <div>HP</div> <div>Atk</div> <div>Def</div> <div>SpA</div> <div>SpD</div> <div>Spe</div> <div>EV</div> <div>36</div> <div>252+</div> <div>—</div> <div>220</div> <div>—</div> </div> </div>




Koraidon




Calyrex




Flutter Mane



Chien-Pao



Dondozo









Ting-Lu

Experiments

10





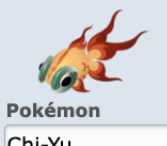
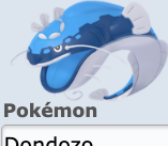
▪ Enemy Team1

<div><div>Nickname</div><div>Miraidon</div></div> <div><div></div><div>Pokémon</div></div> <div><div>Miraidon</div></div>	<div><div>Copy</div><div>Import/Export</div><div>Move</div><div>Delete</div></div> <div><div>Details</div><div>Level 50Gender —Shiny NoTera Type Fairy</div><div>Item Choice SpecsAbility Hadron Engine</div></div> <div><div>Moves</div><div>Electro Drift</div><div>Draco Meteor</div><div>Dazzling Gleam</div><div>Volt Switch</div></div> <div><div>Stats</div><div>HP 140</div><div>Atk —</div><div>Def 4</div><div>SpA 236</div><div>SpD 4</div><div>Spe 124+</div></div>	<div><div>Nickname</div><div>Iron Treads</div></div> <div><div></div><div>Pokémon</div></div> <div><div>Iron Treads</div></div>	<div><div>Copy</div><div>Import/Export</div><div>Move</div><div>Delete</div></div> <div><div>Details</div><div>Level 50Gender —Shiny NoTera Type Normal</div><div>Item Choice BandAbility Quark Drive</div></div> <div><div>Moves</div><div>Earthquake</div><div>Iron Head</div><div>Ice Spinner</div><div>Supercell Slam</div></div> <div><div>Stats</div><div>HP 60</div><div>Atk 196</div><div>Def 20</div><div>SpA —</div><div>SpD 92</div><div>Spe 140+</div></div>
<div><div>Nickname</div><div>Ho-Oh</div></div> <div><div></div><div>Pokémon</div></div> <div><div>Ho-Oh</div></div>	<div><div>Copy</div><div>Import/Export</div><div>Move</div><div>Delete</div></div> <div><div>Details</div><div>Level 50Gender —Shiny NoTera Type Normal</div><div>Item Heavy-Duty BootsAbility Regenerator</div></div> <div><div>Moves</div><div>Sacred Fire</div><div>Tailwind</div><div>Whirlwind</div><div>Recover</div></div> <div><div>Stats</div><div>HP 252</div><div>Atk —</div><div>Def 236+</div><div>SpA —</div><div>SpD —</div><div>Spe 20</div></div>	<div><div>Nickname</div><div>Archaludon</div></div> <div><div></div><div>Pokémon</div></div> <div><div>Archaludon</div></div>	<div><div>Copy</div><div>Import/Export</div><div>Move</div><div>Delete</div></div> <div><div>Details</div><div>Level 50Gender —Shiny NoTera Type Fairy</div><div>Item Citrus BerryAbility Stamina</div></div> <div><div>Moves</div><div>Draco Meteor</div><div>Body Press</div><div>Foul Play</div><div>Thunder Wave</div></div> <div><div>Stats</div><div>HP 220</div><div>Atk —</div><div>Def 236+</div><div>SpA —</div><div>SpD —</div><div>Spe 52</div></div>
<div><div>Nickname</div><div>Ting-Lu</div></div> <div><div></div><div>Pokémon</div></div> <div><div>Ting-Lu</div></div>	<div><div>Copy</div><div>Import/Export</div><div>Move</div><div>Delete</div></div> <div><div>Details</div><div>Level 50Gender —Shiny NoTera Type Fairy</div><div>Item LeftoversAbility Vessel of Ruin</div></div> <div><div>Moves</div><div>Earthquake</div><div>Spikes</div><div>Rest</div><div>Whirlwind</div></div> <div><div>Stats</div><div>HP 244</div><div>Atk 4</div><div>Def 244+</div><div>SpA —</div><div>SpD 4</div><div>Spe 12</div></div>	<div><div>Nickname</div><div>Chien-Pao</div></div> <div><div></div><div>Pokémon</div></div> <div><div>Chien-Pao</div></div>	<div><div>Copy</div><div>Import/Export</div><div>Move</div><div>Delete</div></div> <div><div>Details</div><div>Level 50Gender —Shiny NoTera Type Dark</div><div>Item Focus SashAbility Sword of Ruin</div></div> <div><div>Moves</div><div>Icicle Crash</div><div>Crunch</div><div>Taunt</div><div>Sucker Punch</div></div> <div><div>Stats</div><div>HP —</div><div>Atk 252+</div><div>Def 196</div><div>SpA —</div><div>SpD —</div><div>Spe 60</div></div>

Experiments

11





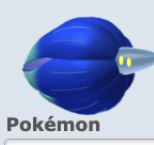
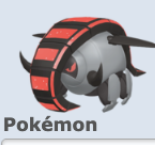
▪ Enemy Team2

<div><div>Nickname</div><div>Lunala</div></div> <div><div></div><div>Pokémon</div></div>	<div><div>Copy</div><div>Import/Export</div><div>Move</div><div>Delete</div></div> <div><div>Details</div><div>Level 50Gender —Shiny NoTera Type Normal</div></div> <div><div>Moves</div><div>Moongeist Beam</div><div>Ice Beam</div><div>Meteor Beam</div><div>Trick Room</div></div> <div><div>Stats</div><div>HP 244</div><div>Atk —</div><div>Def —</div><div>SpA 244+</div><div>SpD —</div><div>Spe 20</div></div> <div><div>Item</div><div>Power Herb</div></div> <div><div>Ability</div><div>Shadow Shield</div></div>	<div><div>Nickname</div><div>Ting-Lu</div></div> <div><div></div><div>Pokémon</div></div>	<div><div>Copy</div><div>Import/Export</div><div>Move</div><div>Delete</div></div> <div><div>Details</div><div>Level 50Gender —Shiny NoTera Type Steel</div></div> <div><div>Moves</div><div>Earthquake</div><div>Ruination</div><div>Heavy Slam</div><div>Payback</div></div> <div><div>Stats</div><div>HP 164</div><div>Atk 76+</div><div>Def 4</div><div>SpA —</div><div>SpD 252</div><div>Spe 12</div></div> <div><div>Item</div><div>Assault Vest</div></div> <div><div>Ability</div><div>Vessel of Ruin</div></div>
<div><div>Nickname</div><div>Koraidon</div></div> <div><div></div><div>Pokémon</div></div>	<div><div>Copy</div><div>Import/Export</div><div>Move</div><div>Delete</div></div> <div><div>Details</div><div>Level 50Gender —Shiny NoTera Type Fairy</div></div> <div><div>Moves</div><div>Flare Blitz</div><div>Close Combat</div><div>Outrage</div><div>U-turn</div></div> <div><div>Stats</div><div>HP 220</div><div>Atk 196+</div><div>Def 4</div><div>SpA —</div><div>SpD 28</div><div>Spe 60</div></div> <div><div>Item</div><div>Choice Band</div></div> <div><div>Ability</div><div>Orichalcum Pulse</div></div>	<div><div>Nickname</div><div>Gliscor</div></div> <div><div></div><div>Pokémon</div></div>	<div><div>Copy</div><div>Import/Export</div><div>Move</div><div>Delete</div></div> <div><div>Details</div><div>Level 50Gender —Shiny NoTera Type Steel</div></div> <div><div>Moves</div><div>Earthquake</div><div>Toxic</div><div>Protect</div><div>Substitute</div></div> <div><div>Stats</div><div>HP 228</div><div>Atk —</div><div>Def 28</div><div>SpA —</div><div>SpD 252+</div><div>Spe —</div></div> <div><div>Item</div><div>Toxic Orb</div></div> <div><div>Ability</div><div>Poison Heal</div></div>
<div><div>Nickname</div><div>Chi-Yu</div></div> <div><div></div><div>Pokémon</div></div>	<div><div>Copy</div><div>Import/Export</div><div>Move</div><div>Delete</div></div> <div><div>Details</div><div>Level 50Gender —Shiny NoTera Type Fire</div></div> <div><div>Moves</div><div>Overheat</div><div>Flamethrower</div><div>Dark Pulse</div><div>Psychic</div></div> <div><div>Stats</div><div>HP 228</div><div>Atk —</div><div>Def 12</div><div>SpA 116+</div><div>SpD 44</div><div>Spe 108</div></div> <div><div>Item</div><div>Choice Specs</div></div> <div><div>Ability</div><div>Beads of Ruin</div></div>	<div><div>Nickname</div><div>Dondozo</div></div> <div><div></div><div>Pokémon</div></div>	<div><div>Copy</div><div>Import/Export</div><div>Move</div><div>Delete</div></div> <div><div>Details</div><div>Level 50Gender —Shiny NoTera Type Fairy</div></div> <div><div>Moves</div><div>Wave Crash</div><div>Yawn</div><div>Protect</div><div>Rest</div></div> <div><div>Stats</div><div>HP 252</div><div>Atk —</div><div>Def 252+</div><div>SpA —</div><div>SpD 4</div><div>Spe —</div></div> <div><div>Item</div><div>Leftovers</div></div> <div><div>Ability</div><div>Unaware</div></div>

Experiments

12

▪ Enemy Team3

<div><div>Nickname</div><div>Miraidon</div></div> <div><div></div><div>Pokémon</div></div> <div><div>Miraidon</div></div>	<div><div>Copy</div><div>Import/Export</div><div>Move</div><div>Delete</div></div> <div><div>Details</div><div>Level 50Gender —Shiny NoTera Type Electric</div><div>Item Assault VestAbility Hadron Engine</div></div> <div><div>Moves</div><div>Electro Drift</div><div>Draco Meteor</div><div>Mirror Coat</div><div>U-turn</div></div> <div><div>Stats</div><div>HP 252</div><div>Atk —</div><div>Def 4</div><div>SpA 60+</div><div>SpD 116</div><div>Spe 76</div></div>	<div><div>Nickname</div><div>Chien-Pao</div></div> <div><div></div><div>Pokémon</div></div> <div><div>Chien-Pao</div></div>	<div><div>Copy</div><div>Import/Export</div><div>Move</div><div>Delete</div></div> <div><div>Details</div><div>Level 50Gender —Shiny NoTera Type Ghost</div><div>Item Focus SashAbility Sword of Ruin</div></div> <div><div>Moves</div><div>Icicle Crash</div><div>Ice Shard</div><div>Sucker Punch</div><div>Swords Dance</div></div> <div><div>Stats</div><div>HP —</div><div>Atk 252</div><div>Def 4</div><div>SpA —</div><div>SpD —</div><div>Spe 252+</div></div>
<div><div>Nickname</div><div>Dragonite</div></div> <div><div></div><div>Pokémon</div></div> <div><div>Dragonite</div></div>	<div><div>Copy</div><div>Import/Export</div><div>Move</div><div>Delete</div></div> <div><div>Details</div><div>Level 50Gender —Shiny NoTera Type Fairy</div><div>Item Electric SeedAbility Multiscale</div></div> <div><div>Moves</div><div>Breaking Swipe</div><div>Earthquake</div><div>Dragon Dance</div><div>Roost</div></div> <div><div>Stats</div><div>HP 244</div><div>Atk 20+</div><div>Def 60</div><div>SpA —</div><div>SpD 68</div><div>Spe 116</div></div>	<div><div>Nickname</div><div>Urshifu</div></div> <div><div></div><div>Pokémon</div></div> <div><div>Urshifu-Rapid-Str</div></div>	<div><div>Copy</div><div>Import/Export</div><div>Move</div><div>Delete</div></div> <div><div>Details</div><div>Level 50Gender —Shiny NoTera Type Steel</div><div>Item Punching GloveAbility Unseen Fist</div></div> <div><div>Moves</div><div>Surging Strikes</div><div>Drain Punch</div><div>Swords Dance</div><div>Aqua Jet</div></div> <div><div>Stats</div><div>HP 156</div><div>Atk 252+</div><div>Def 100</div><div>SpA —</div><div>SpD —</div><div>Spe —</div></div>
<div><div>Nickname</div><div>Glimmora</div></div> <div><div></div><div>Pokémon</div></div> <div><div>Glimmora</div></div>	<div><div>Copy</div><div>Import/Export</div><div>Move</div><div>Delete</div></div> <div><div>Details</div><div>Level 50Gender —Shiny NoTera Type Fairy</div><div>Item Red CardAbility Toxic Debris</div></div> <div><div>Moves</div><div>Power Gem</div><div>Dazzling Gleam</div><div>Mortal Spin</div><div>Endure</div></div> <div><div>Stats</div><div>HP 244</div><div>Atk —</div><div>Def 228+</div><div>SpA 4</div><div>SpD 20</div><div>Spe 12</div></div>	<div><div>Nickname</div><div>Iron Treads</div></div> <div><div></div><div>Pokémon</div></div> <div><div>Iron Treads</div></div>	<div><div>Copy</div><div>Import/Export</div><div>Move</div><div>Delete</div></div> <div><div>Details</div><div>Level 50Gender —Shiny NoTera Type Ground</div><div>Item Choice BandAbility Quark Drive</div></div> <div><div>Moves</div><div>Earthquake</div><div>Iron Head</div><div>Knock Off</div><div>Supercell Slam</div></div> <div><div>Stats</div><div>HP 60</div><div>Atk 196</div><div>Def 20</div><div>SpA —</div><div>SpD 92</div><div>Spe 140+</div></div>

■ Results

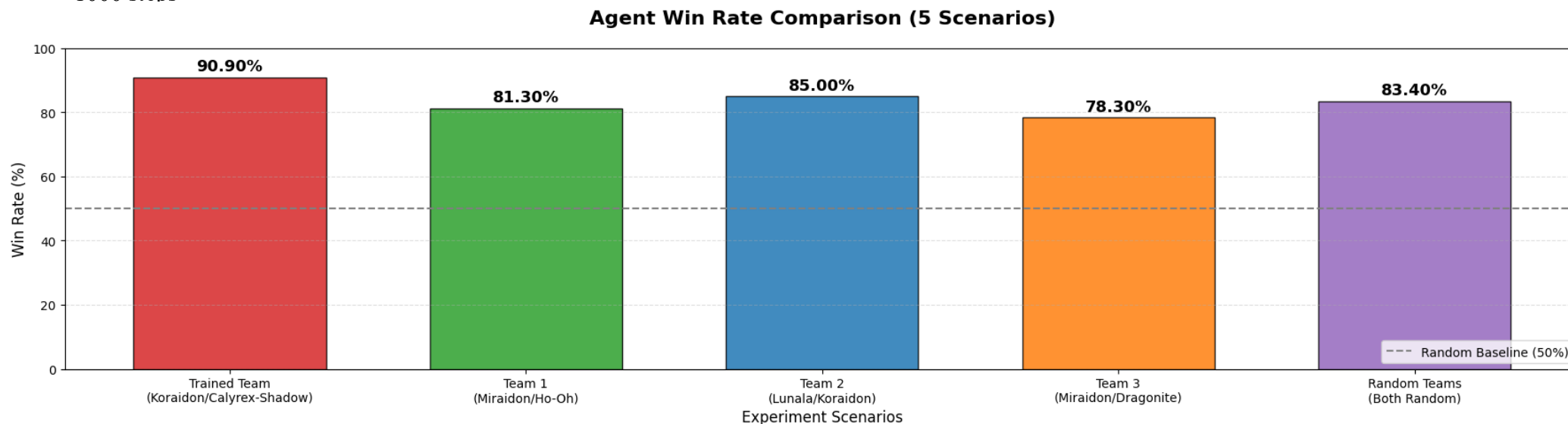
- 학습된 파티 그대로 + 상대 에이전트는 학습 X (Random 행동)
- 상대방의 파티는 Enemy 1, 2, 3 중 단 하나
- 1000번 중 909승을 달성하며 90.9%의 승률

EVALUATION RESULTS

```
=====
EVALUATION RESULTS
=====
Total Battles:      1000
Wins:               909
Losses:             91
Win Rate:           90.90%
Avg Reward:         18.29 ± 9.02
Avg Episode Len:    18.9 steps
Avg Alive Diff:     +9.43 Pokémon
Reward Range:       [-23.14, 31.75]
=====
```

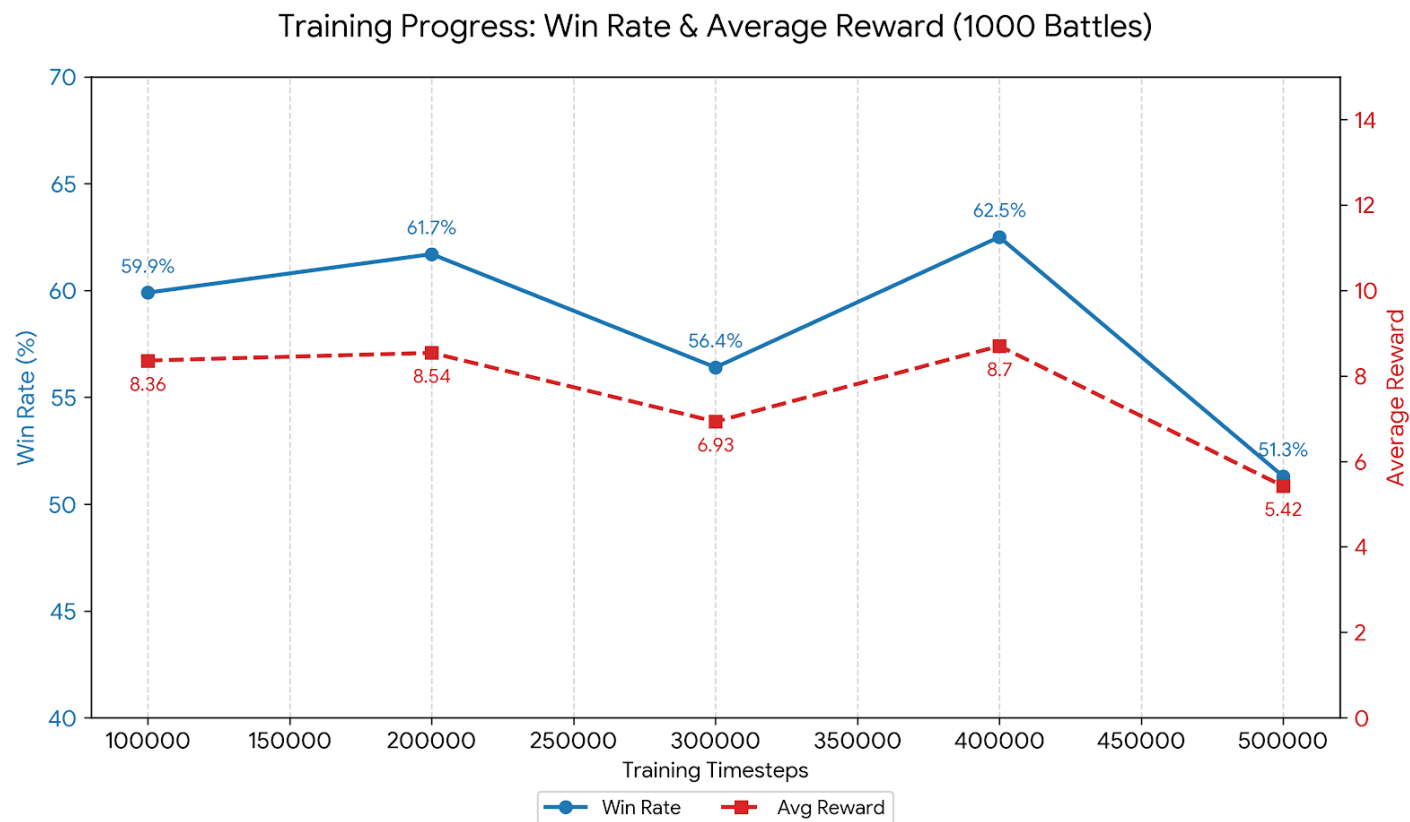
■ Ablation Results

- 학습된 팀이 아닌 다른 팀에서도 어느정도의 승률을 가지는지 평가
- 각각 Enemy team 1, 2, 3로 My Team으로 학습한 전략을 사용
- 각각의 팀들은 다음과 같은 성격을 가짐
 - Enemy Team1: 학습한 팀과 유사한 성격의 화력으로 밀어붙이는 조합
 - Enemy Team2: Gliscor과 Dondozo의 회복 능력을 바탕으로, 눌러 앉는 조합
 - Enemy Team3: Glimmora가 장판을 설치한 후, 장판기를 힘을 바탕으로 다른 포켓몬이 활약하는 조합
- My Team이 Enemy Team 1, 2, 3 중 Random으로 결정
- 1000 steps



■ Mirror Battle (1000 episodes)

- 이전의 step들보다는 강력한 마지막의 final checkpoint
but final vs final에서는 **51.3%**의 Win rate를 기록 (거의 반반)



▪ Performance Summary

- 제안된 **Action Masking** 및 **Dense Reward Shaping** 기법을 통해, 500,000 Step 학습만으로 **90.9%의 승률**을 달성함
- 특정 팀에 국한되지 않고, 각기 다른 상대로도 **평균 80% 이상의 승률**을 기록하여 일반화 성능을 입증함

▪ Strategic Behavior

- 단순 데미지 딜링뿐만 아니라, **교체(Switching)** 및 **테라스탈(Terastallization)** 을 활용한 방어적 플레이가 관측됨

▪ Limitations & Future Work

- **상대 모델의 단순함:**
현재 평가는 Random/Simple Heuristic 상대를 기반으로 함. 향후 더 강력한 **Pre-trained Agent**나 사람과의 대전을 통해 검증 필요.
 - **방어 상성 학습:**
현재 State에는 방어 상성 정보가 잠재적(Latent)으로만 존재함. 이를 명시적 정보로 추가하면 학습 효율이 개선될 것으로 예상됨
 - **파티 선택 학습:**
현재는 6마리 중 맨 앞의 3마리만 선출되고 있는 상황으로 builder를 업데이트 해야 파티 중 어떤 포켓몬을 선출할지 학습함
 - **더 많은 상태 학습:**
멸망의 노래, 선공기, 트릭룸에 따라서 state와 reward를 재정의해서 바르게 대처할 수 있도록 업그레이드 해야함
-