

현업에서 활용하는 나만의 StarCraft2

강화학습 Agent

만들기

2020. 09

풀어야 할 문제? 게임의 승리

[State]

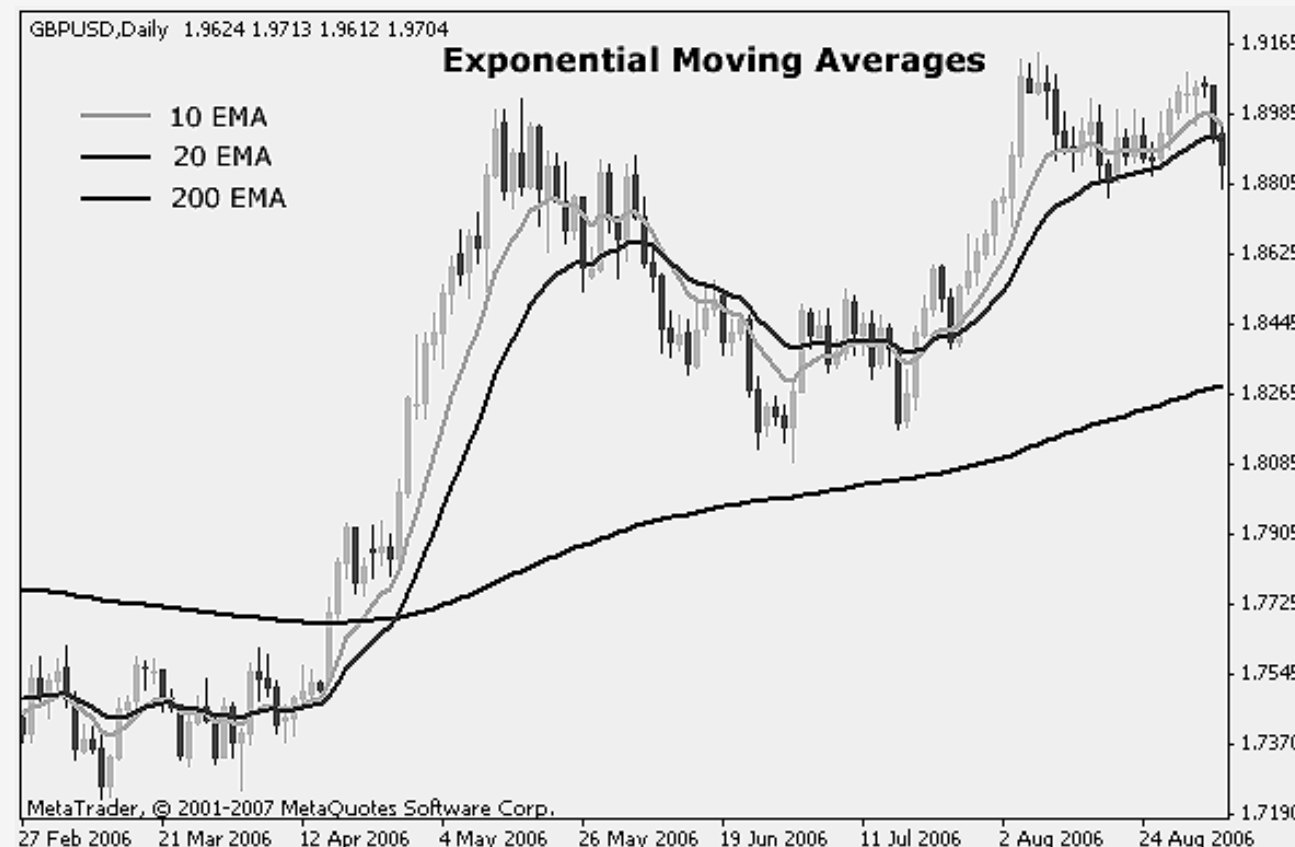
- 1 전장의 상태를 정의한다
- 2 자신의 상태 뿐만 아니라 상대방의 상태도 파악한다
- 3 자원상태, 건물 상태, 유닛 상태 등을 포함한다
- 4 매 Step 마다 State 정보를 최신화 하며 Action을 취한다

[Action]

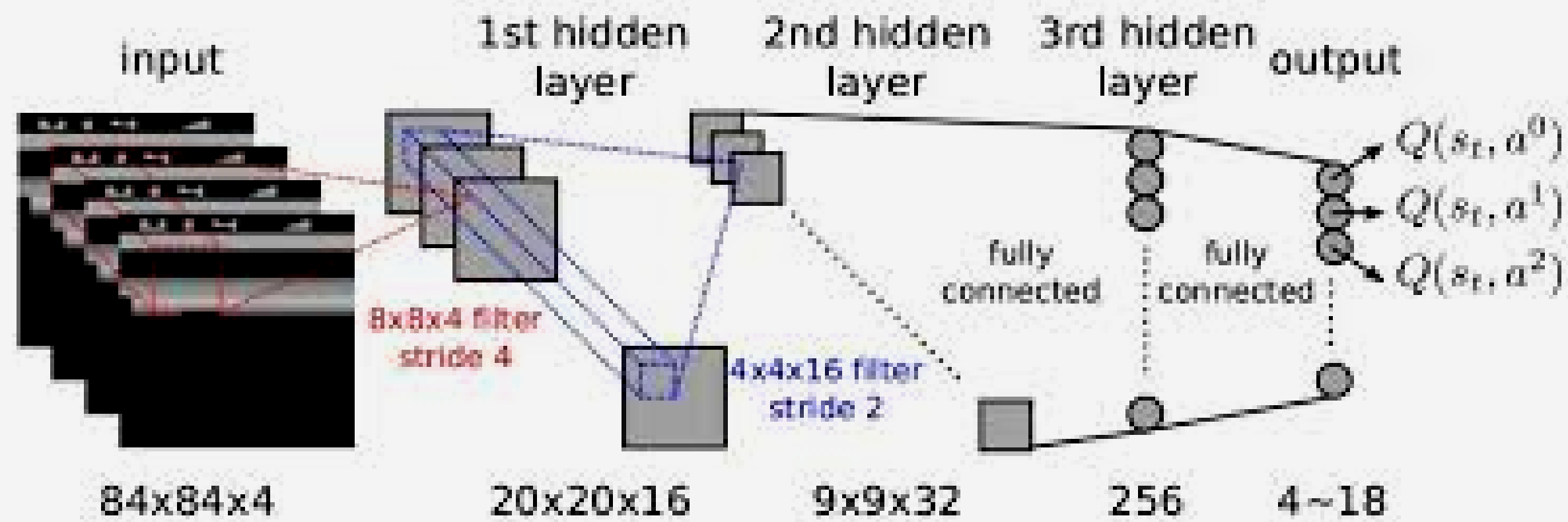
- 1 새로운 건물을 짓는 행동을 추가한다 (Refinery, Factory, Lab)
- 2 새로운 유닛을 생성하는 행동을 추가한다 (Siege Tank)
- 3 새로운 자원(가스)를 채취하는 행동을 추가한다
- 4 다양한 루트로 공격한다 (멀티 기지로 인해 비김현상 방지)

Reward Function

using **EMA** (지수이동평균)



- 1 이기면 1, 비기면 0, 지면 -1의 Reward를 받는다
- ✓ 2 학습의 양상을 보기 위해 '이동평균' 지표를 활용한다
- 3 최근의 추이를 더 큰 비중으로 반영하기 위해 EMA를 활용한다
- 4 과거의 데이터 영향도를 차츰 줄여나가는 장점이 있다
- 5 갑자기 값이 증가하면, 변동성이 커진 것으로 판단하면 된다



Algorithm

Vanilla DQN

- 1 이미지를 사용하지 않으므로 CNN 영역은 제외한다
- 2 Experience Replay를 활용한다
 - 과거의 데이터를 학습에 활용한다
 - Sample간의 시간적 연관성을 해결할 수 있다
- 3 Fixed Q-Target을 활용한다
 - Moving Q-Target 문제를 해결할 수 있다
 - 원하지 않는 State의 Q Table Update를 막을 수 있다

학습 결과

[난이도에 따른 결과]

- 1 Very Easy : 낮은 학습수에도 거의 100%의 승률을 보임
- 2 Easy : 위와 비슷하게 90% 이상의 승률을 보임
- 3 Medium : 거의 이기지 못하다가 30% 대의 승률을 보임

[학습에 따른 결과]

- 1 쉬운 난이도에서는 학습수와 무관하게 거의 이기는 모습을 보였다
다양한 Action으로 고급 유닛 활용이 가능해 쉽게 이긴 것으로 보인다
- 2 Medium 난이도에서, 50 lter 정도까지는 학습수가 유의미하나,
30% 근처 달성 이후, 그 이후 보합을 유지하였다.
- 3 단순 몇몇 Hyperparameter 변경으로는 유의미한 결과가 없었다.
- 4 추후 Rainbow 등을 통해 개선해보고싶다.

Thank
You

