

배달원 배치 action, state, reward 설정

- **ACTION** : 배달원 배치
- **STATE** : 배달원 위치(x,y), 배달원 근무시간(on,off), 배달 건 수, 신뢰도, 음식점 위치(x,y)
- **Reward** : 배달 건 수의 표준편차를 작게 / 신뢰도(시간 내 도착)를 높게
- **목적** : 배달원에게 배달을 고르게 분배하기 위해

Reward에서 고려하는 사항

표준편차 : 배달원의 배달 건수에 대한 표준편차

- 라이더(a)가 배달을 했을 때, 전체 표준편차가 작아진다면(0에 가까워져 골고루 배달 된다면)
→ $+(\text{원래 표준편차} + (\text{원래 표준편차} - \text{변화된 표준편차}))$
- 라이더(a)가 배달을 했을 때, 전체 표준편차가 커진다면
→ $-(\text{원래 표준편차} + (\text{원래 표준편차} - \text{변화된 표준편차}))$

신뢰도 : 시간 내에 배달을 완료하였을 때 부여

- 배달원이 60분 내에 배달을 완료해야 한다고 가정했을 때, 단위는 분(min)
→ $+60 - (\text{배달원이 음식점까지 걸린 시간})$