

카카오모빌리티 자율주행시스템개발 부서 과제

휴먼지능정보공학과 201910803 박채희

우선 자율 주행 자동차의 trolley problem의 예시는 다음과 같다. 자율 주행 자동차에 운전자가 탄 상태인데, 그 앞에 5명의 사람이 서있다고 가정한다. 이때 차는 정지할 수 없는 상태이며 사고를 피하는 유일한 방법은 방향을 틀어서 골목과 충돌하는 것인데, 그렇게 된다면 운전자가 다치게 된다.

이때, 운전자의 예상 피해수준이 신체가 크게 다치지 않아 몇일 정도 뒤에 일상생활에 복귀가 가능한 수준이라면, 나는 골목과 충돌하는 방법을 택할 것이다. 하지만 그 차에 나와 가족들이나 친구들이 타고 있다면, 나는 5명의 사람을 칠 것 같다. 왜냐면 나 혼자 다치는 것은 내가 골목에 충돌하겠다는 내 판단이지만, 다른 사람과 합승하였을 때는 나만의 판단만 존재하지 않기 때문에 이때는 5명의 사람을 치는 판단을 할 것 같다. 또한, 만약 피해 수준이 엄청 커서 사망 또는 일상생활에 바로 회복이 어려운 상황이 된다면 이때는 나 혼자 있던 여러 명이 합승해 있던 상관없이 5명의 사람을 판단을 할 것 같다. 나는 내 목숨이 더 중요하기 때문이다.

물론 내가 결정한 판단이 사회적으로 옳은 지 아닌지는 모르겠다. 운전자와 보행자 어느 사람의 목숨에 더 가치를 두어야 하는지 생각해 보아야 한다. 사실, 이런 paradox를 만들지 않는 정확한 인공지능을 개발하는 것이 가장 이상적이라고 생각하지만, 모든 인공지능이 항상 올바른 판단을 내리는 것은 어렵다고 생각할 때, 내 생각은 다음과 같다. 만약 자율 주행 자동차를 이용하는 사람이 보행자보다 많아진다면, 그때는 운전자를 보호하는 인공지능 시스템에 중점을 두어야 한다고 생각하고, 아니라면 보행자를 보호하는 인공지능 시스템에 중점을 두어야 한다고 생각한다.