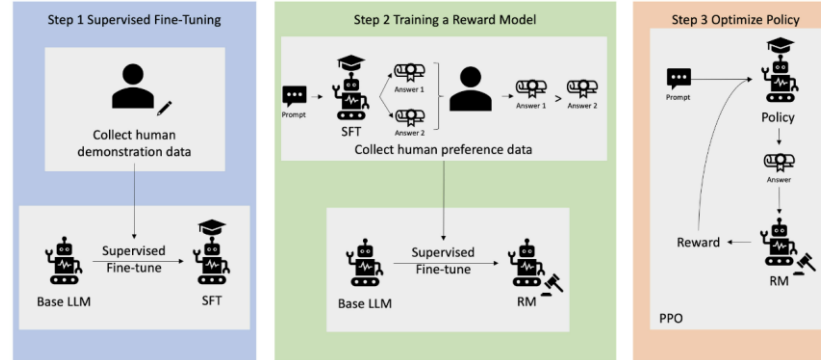


RLHF를 이용한 협동 로봇 제어 프로그램 개발

Reinforcement Learning with Human Feedback (RLHF)

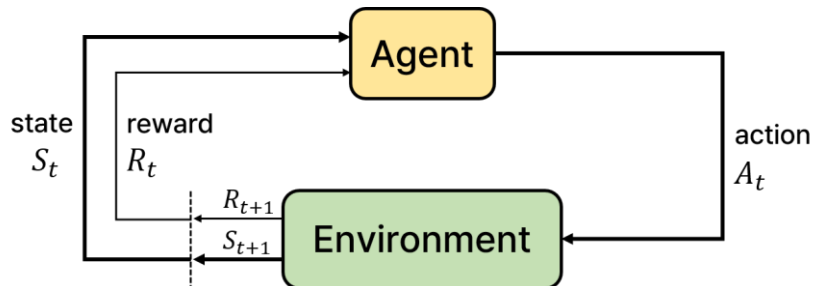
Reinforcement Learning with Human Feedback (RLHF)



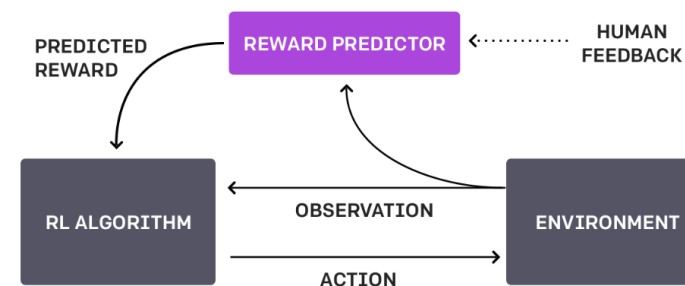
⇒ 원래 **RLHF**는 **LLM 모델**을 **Fine-tuning** 하기 위해 사용하는 방법이지만, 기존의 보상 함수를 설계하는 대신 **RLHF**를 사용한다면 환경이 복잡하고 어려워질 수록 보상 함수를 설계하기 어렵다는 **강화 학습의 한계**를 **극복**해줄 수 있지 않을까?

기존의 RL과 RLHF의 차이

• Reinforcement Learning (RL)



• Reinforcement Learning with Human Feedback (RLHF)



협동 로봇 사고 사례



테슬라 공장 로봇, 근로자 공격..."벽으로 밀고 집게발로 찢러"



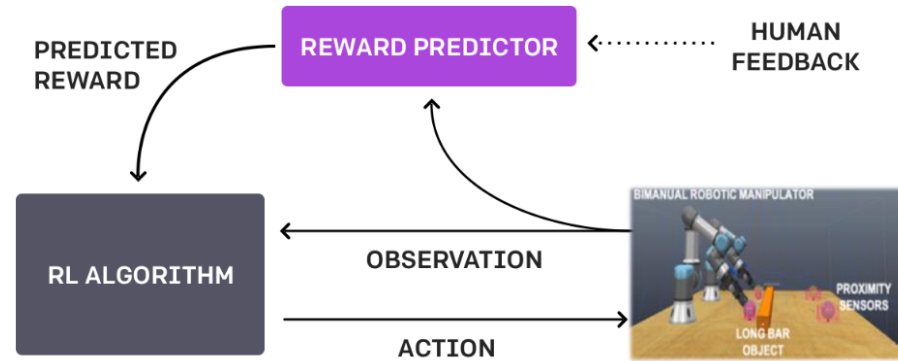
'사람을 상자' 오인한 작업 로봇...노동자 또 사망



인간과 밀접하게 연결되어 작동하는 협동 로봇은 다양한 상황에 대처할 수 있어야 한다.

→ 그러나 현재 자율 동작 지능이 부재한 수동적인 로봇이 대부분

최종 목표 및 결과물 (CoRLHF)

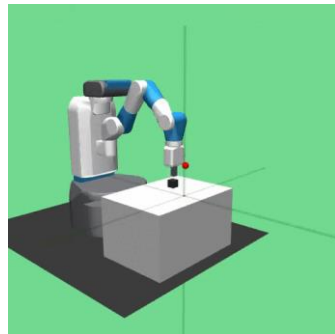


CoRLHF (Collaborative Reinforcement Learning with Human Feedback)

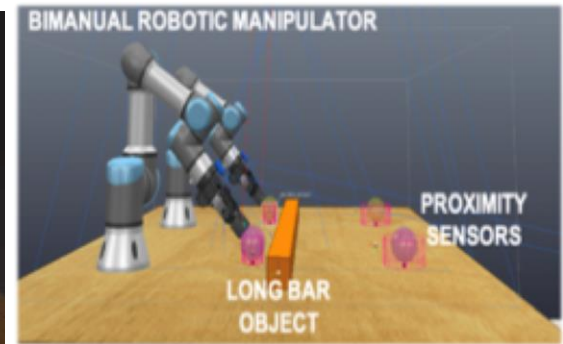
가상환경 내 로봇 팔을 RLHF 알고리즘을 이용하여 제어할 수 있다.

Push 또는 Pick and Place와 같은 다양한 작업을 인간과 유사한 방식으로 제어할 수 있다.

1학기



2학기





프로젝트 기간별 목표

1학기 목표

시뮬레이터 환경(단일 에이전트 환경)에 RLHF 알고리즘을 적용, 다양한 과제를 수행할 수 있는지 확인하고 평가

2학기 목표

협동 로봇(멀티 에이전트 환경)에서 RLHF 알고리즘을 적용, 다양한 과제를 수행할 수 있는지 확인하고 평가

월	활동 내용
3월	기초 연구 및 기술 검토: 문헌 검토, 기술 분석, 요구사항 정의
4월	가상환경 비교 및 선택, 시스템 구조 설계, RLHF 알고리즘 분석
5월	가상환경 내 RLHF 알고리즘 적용 1
6월	가상환경 내 RLHF 알고리즘 적용 2, 작업 수행 평가
9월	협동로봇 환경에서 RLHF 알고리즘 적용 1
10월	협동로봇 환경에서 RLHF 알고리즘 적용 2, 작업 수행 평가 1
11월	작업 수행 평가 2
12월	PoC (최종 테스트 및 평가)

필요한 요구 사항

강화 학습을 위한 시뮬레이터를 사용하기 위해선
Ubuntu 운영체제 기반의 GPU desktop이 필요하다.

- 학교 실습실의 서버를 사용할 예정
- GUI가 필요한 작업은 연구실의 서버를 사용해 진행할 예정

Role & Responsibility (R&R)

RLHF 알고리즘을 각 시뮬레이터에 적용하기 위해선 RLHF 알고리즘에 환경에 맞춰서 MDP를 정의하는 작업을 진행해야한다.
전체 팀원이 모든 전체 작업을 다 함께 진행할 예정이며, MDP를 지정하는 작업에서 코드를 수정하는 작업은 파트를 나눠서 진행할 예정이다.