통합 다음에는,
협동의 효과를 증명하기

상황이 바뀔 때 **언제 다시 판단하고,**
학습으로 **무엇이 나아지는가?**

협동 게임 AI의 남은 반학기 연구와 응용

1팀 더미 · 2026.10.08

문헌 근거 → 비교 실험 → 응용 시연

출발점 · 10/2 상담 요지

교수님이 요구한 것은 목표와 근거가 연결되는 설명

상담의 핵심 요구	이번 계획에서 보여 줄 것
최종 목표·연구 질문을 먼저 제시	재판단 방식 × 학습 여부의 효과를 비교
단순 기능 연결·JSON 출력 이상의 기여	실패를 재현하는 시나리오와 자동 판정
미리 알려 준 정보와 추론을 구분	관측·관계·행동 후보의 공개 범위 명세
복잡한 협동·실패 상황을 검증	목표 변경, 역할 충돌, 선행 조건 위반
전체 계획 속 이번 주 위치를 설명	통합 완료 → 계측 → 실험 → 결과 분석

내부 근거: 2026-10-02 교수상담 정리 · 타임스탬프와 원문 경로는 발표자 노트·제작 기록에 보관

확인된 상태 · 2026.10.08

통합 기반은 있다.
성능 비교와 사람 플레이 검증이 남았다.

현재 직접 확인한 범위

경민의 퍼즐·상태 → 조정기 →
준영의 Qwen·Unreal 실행기

107

Node 회귀 시험

24

실제 Qwen + Unreal

브라우저 12 · 네이티브 3 · Editor/Game 빌드 통과

다음에 확보할 근거

- 최신 수정 후 Live API 재시험
- 사람 마이크·끼어들기·재접속 플레이
- 주기 / 사건 재판단의 반복 비교
- 행동 모델 학습 전후의 미지 과제 성능

시험 통과 개수는 게임 성공률이나 연구 성능 수치가 아니다.

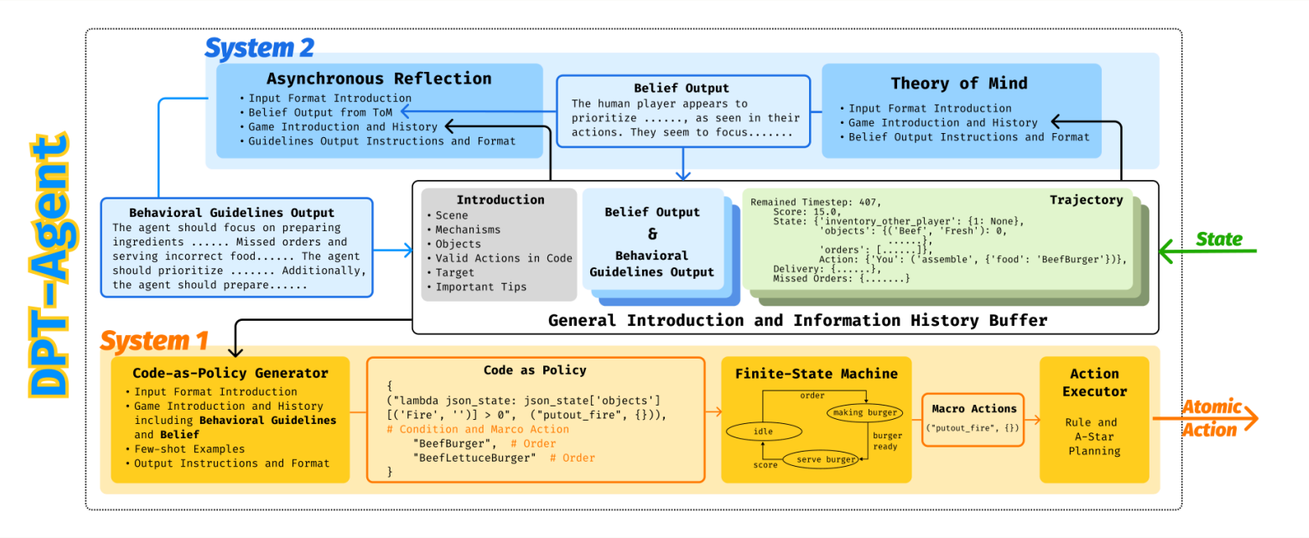
구현 기준: CapstoneDesign main · bd92d2f

우선순위 · 실행 제안

권고: 핵심 비교를 끝내고, 응용은 같은 구조로 하나만 확장

우선순위	남은 작업	판단
1 · 먼저	반복 실행·이벤트 로그·성공/실패 판정	모든 연구 결과의 기반
2 · 핵심	상황 기반 재판단 + 행동 AI LoRA	기존 확정 주제를 검증 가능하게 완성
3 · 보조	정보 공개 수준·모호한 지시의 질문	주 실험과 분리한 작은 추가 비교
4 · 응용	가상 작업 훈련 시나리오 1개	새 플랫폼 없이 활용 가능성 시연
이번 학기 보류	실물 로봇·새 비전 모델·대규모 사용자 연구	장치·데이터·평가 범위가 크게 증가

DPT-Agent: 추론을 기다리는 동안에도 실행이 이어지는 구조를 참고



[R1] Zhang et al. · ACL 2025 · PDF 5쪽, Figure 3

가져올 원리

추론과 실행을 분리하고,
실행 중 관측을 다시 반영한다.

우리의 비교 질문

같은 실행기에서 모델 호출 시점을
바꾸면 성공·지연·비용이 달라질까?

해석의 경계

원 논문도 상황 변화에 대응한다.
이벤트 반응 자체를 새 기법이라 주장하지 않는다.

가설 **H1 · 호출 시점**

재판단 정책을 바꾸면 이 장면에서 무엇이 달라져야 할까?

설명용 가상 예시 · 성능 측정 결과 아님

04 / 04

결과를 비교하는 지점

발판 유지 → 동료 이탈 → 재판단 → 유효한 행동.
전체 경로를 기록한다.

공개 관측 → 정책 판단 → 행동 AI → 실행 결과

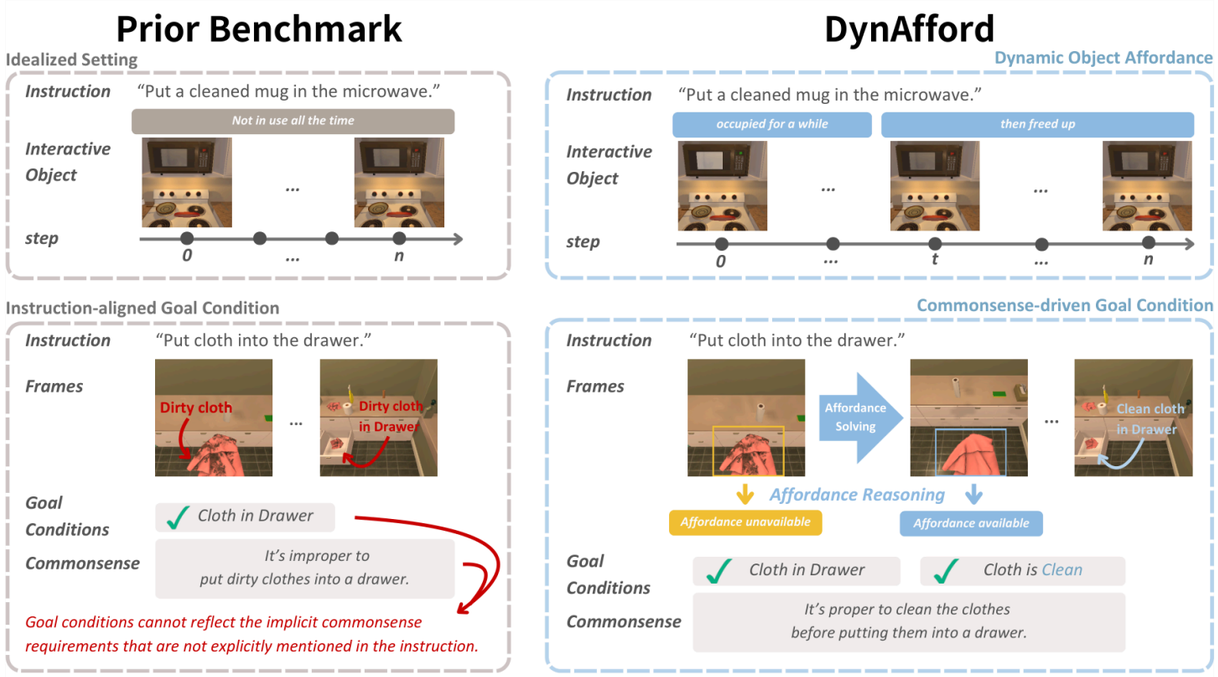
정책	현재 장면의 동작
주기적 호출	변화부터 유효 실행까지의 지연과 호출 수
사건 기반 호출	동일 지표로 비교하고 누락 사건도 기록

두 조건 모두 같은 취소·유효성 검사 적용

1 유지 | 2 변화 | 3 판단 | 4 검증

제안 정책: 의미 있는 변화 + 중복 병합 + 호출 상한 + 최대 대기 시간

ADAPT: 물체가 있다는 것과 지금 쓸 수 있다는 것은 다르다



[R3] Chen et al. · ACL 2026 · PDF 2쪽, Figure 1

논문에서 다룬 문제

물체의 사용 가능 조건이 바뀌고,
지시에 빠진 상식적 제약이 존재한다.

게임에 적용

문·발판의 존재 정보와
현재 점유·잠김·통과 가능 여부를 분리한다.

범위 유지

이 논문은 시각·행동 추론 환경이다.
우리 JSON 관측을 비전 인식으로 부르지 않는다.

정보 계약 · 추가 통제 실험

교수님 질문에 답하려면 모델에 주는 정보를 먼저 명세해야 한다

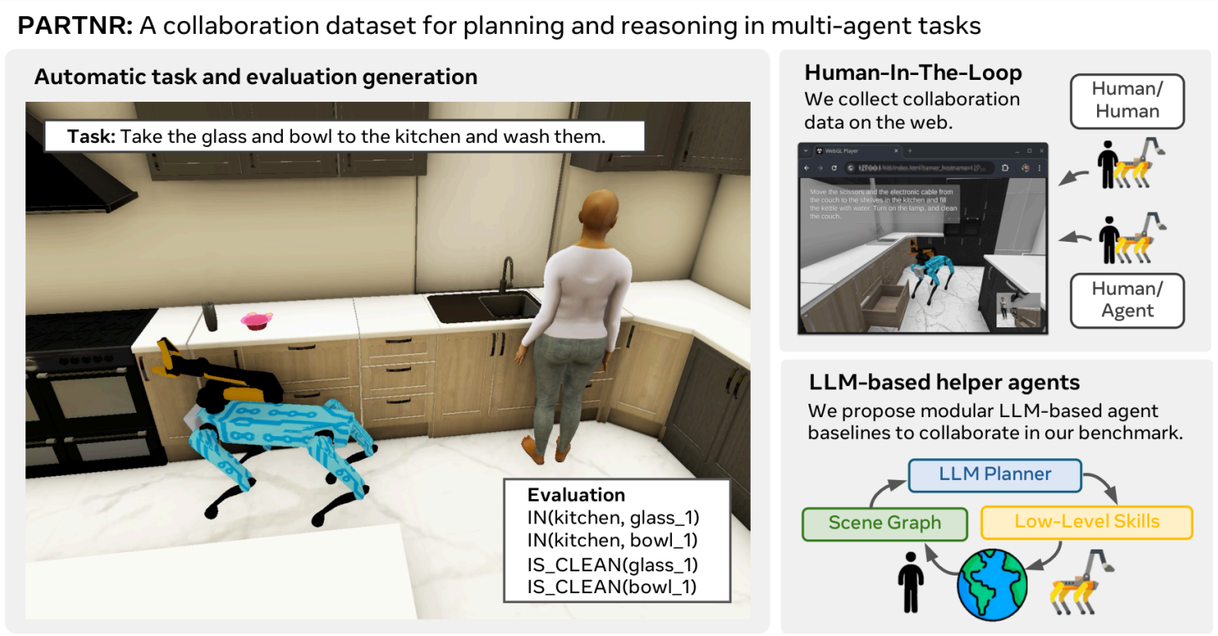
공개 조건 제안	모델이 받는 것	검증할 질문
O · 관측	보이는 객체·점유·문 상태 + 공통 행동 규격	현재 정보만으로 선택 가능한가?
O + R · 관계	O + 작성자가 설정한 연결·선행 관계	관계 정보를 주면 무엇이 달라지는가?
O + R + H · 추가 힌트	위 정보 + 정답 경로에 가까운 단계 힌트	풀이를 미리 알려 준 영향은 얼마나 큰가?

핵심 2×2 실험은 정보 조건 하나를 고정한다.
정보 수준 비교는 별도 실험으로 진행한다.

감사 대상: 후보 행동 필터, exit_id, 관계명, 예시 정답, 메모리, 지연 도착한 관측

[R3] Chen et al., ACL 2026 · [R2] Chang et al., ICLR 2025

PARTNR: 협동의 난도를 규칙과 판정 조건으로 설계한다



[R2] Chang et al. · ICLR 2025 · PDF 2쪽, Figure 1

참고할 요소

공간·순서·역할별 제약이 있는 과제와
실행 가능한 평가 조건.

우리의 작은 벤치마크

맵 수를 늘리기 전에
같은 맵에 실패 원인을 통제해 주입한다.

판정의 분리

모델이 완료라고 말한 것과
실제 게임 목표 달성을 따로 기록한다.

벤치마크 초안 · 작은 범위에서 깊게

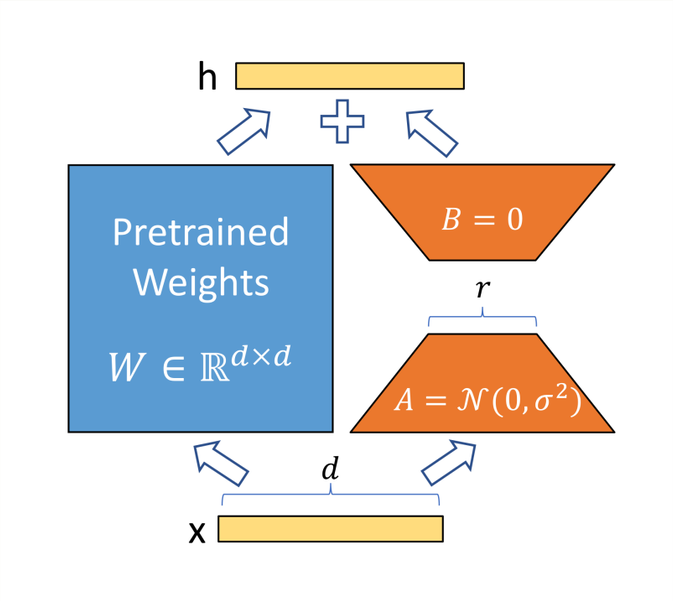
기존 퍼즐에서 시작할 다섯 가지 시험 장면

장면	의도적으로 바꿀 조건	자동 판정의 예
S0 · 기본 협동	기존 발판·문·출구 흐름	실제 문 개방 후 출구 도달
S1 · 목표 변경	유지 중 사람이 다른 목표를 요청	취소된 행동의 재실행 0건
S2 · 전제 붕괴	동료가 발판을 벗어나 문이 닫힘	변화 이후 유효한 행동으로 복구
S3 · 역할·순서	두 행동의 동시 점유 / 선행 조건 추가	충돌·중복·순서 위반 기록
S4 · 지시 모호성	“저쪽 발판”처럼 후보가 둘인 지시	확인 질문 후 의도한 대상 수행

S0는 현재 검증 경로를 재사용한다. S1–S4는 반복 실험용 시나리오·판정기를 보완하는 제안이다.

[R2] Chang et al., ICLR 2025 · [R3] Chen et al., ACL 2026 · [R10] Chen et al., arXiv 2026, preprint

LoRA: 형식보다 상황에 맞는 행동 선택을 학습



[R6] Hu et al. · ICLR 2022 · PDF 1쪽, Figure 1

기법

기존 모델 가중치를 고정하고
작은 저랭크 행렬을 학습한다.

가설 H2

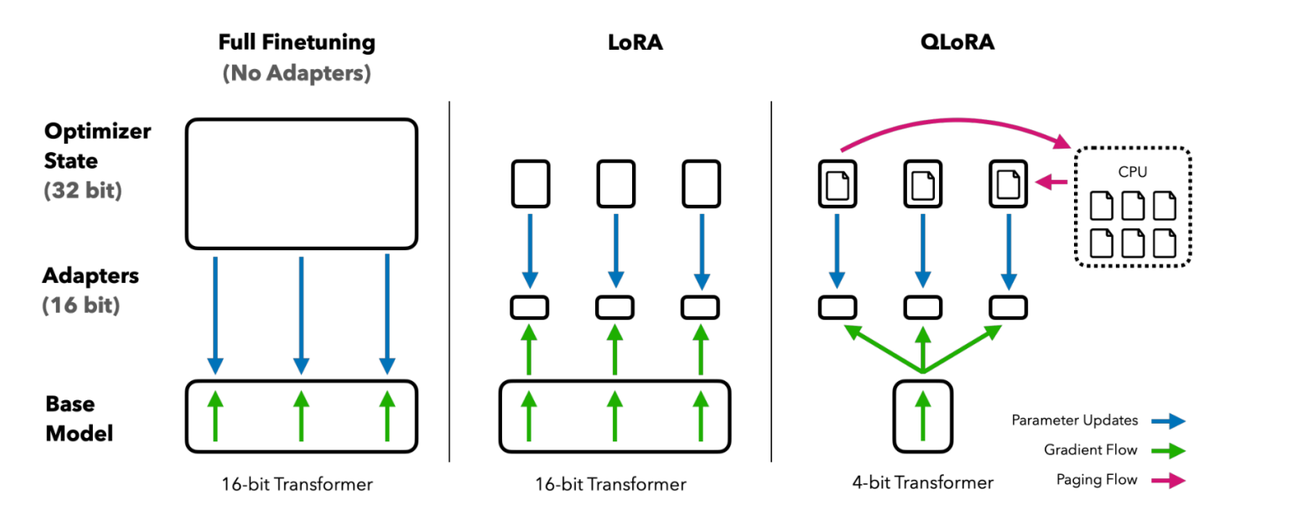
같은 Qwen의 학습 전후를 비교해
잘못된 선택·중복 행동이 줄어드는지 본다.

학습 과제

관측 + 지시 + 실행 피드백 →
행동 / 대기 / 확인 질문 선택

JSON 형식 통과와 의미상 올바른 선택은 별도 지표

QLoRA는 학습 메모리의 선택지. 실행 속도는 따로 측정한다.



[R7] Dettmers et al. · NeurIPS 2023 · PDF 3쪽, Figure 1

원 논문의 초점

양자화된 기본 모델 위에
어댑터를 학습해 메모리를 절약한다.

우리 장비의 첫 관문

현재 모델 계열을 확인한 뒤
짧은 학습으로 VRAM·시간·변환을 점검한다.

추론 조건 통제

같은 정밀도·컨텍스트·디코딩 조건에서
기본 모델과 학습 모델을 평가한다.

[R11] Qwen Team, Official model card

학습 데이터보다 먼저 평가 데이터의 경계를 고정

01

과제군부터 분리

맵·관계 조합·전체 실행 기록 단위로 train / validation / test 분리

02

실패를 라벨링

옛 목표 지속, 중복 행동, 잘못된 대상, 필요한 질문의 누락

03

그다음 학습

기본 모델과 같은 출력 규격.
추가로 선택할 행동만 학습

**같은 장면의 문장만 바꾼 데이터를 학습·평가에 나누면,
미지의 협동 상황에 대한 성능을 평가할 수 없다.**

평가 세트는 잠근다. 프롬프트·주기·트리거·학습 설정은 validation에서만 고른다.

[R6] Hu et al., ICLR 2022 · [R11] Qwen Team, Official model card

주 실험 · 기존 확정 방향의 구체화

네 조건으로 효과를 분리한다

호출 정책 / 행동 모델	기본 모델	같은 모델 + LoRA
고정 주기	**A** 주기 · 기본	**C** 주기 · 학습
사건 + 최대 대기	**B** 사건 · 기본	**D** 사건 · 학습

호출: B−A / D−C · 학습: C−A / D−B · 상호작용: (D−C)−(B−A)

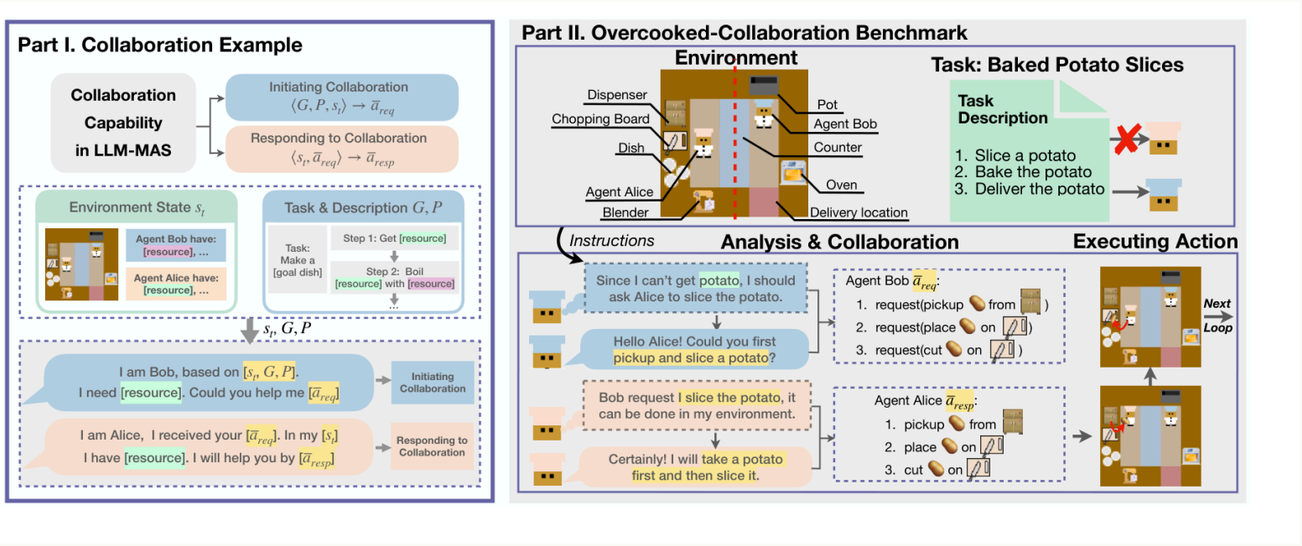
동일 정보·시드·파트너·보호 장치·추론 정밀도·제한 시간. 고정 주기는 여러 후보를 검증한다.

성공률·지연·비용을 같은 실행 기록에서 함께 읽는다

지표	측정 정의	주의할 함정
과제 성공	실제 목표 조건 충족 / 전체 시도	타임아웃·복구 실패를 분모에서 빼지 않기
변화 대응 지연	변화 발생 → 첫 유효 행동의 실행 시작	p50 / p95 + 제한 시간 내 무반응 비율
모델 의미 오류	잘못된 대상·순서·옛 목표 선택	형식 오류·실행 차단 건수와 구분
계산·협동 비용	호출·토큰·벽시계 시간·파트너 대기	성공한 실행만 골라 계산하지 않기

같은 장면·시드의 짝 비교 + 변동 범위 + 대표 실패 기록

Collab-Overcooked:
협동을 먼저 제안하는 능력도 평가



[R4] Sun et al. · EMNLP 2025 · PDF 4쪽, Figure 1

논문의 구분

협동을 시작하는 행동과
받은 요청에 응답하는 행동을 나눈다.

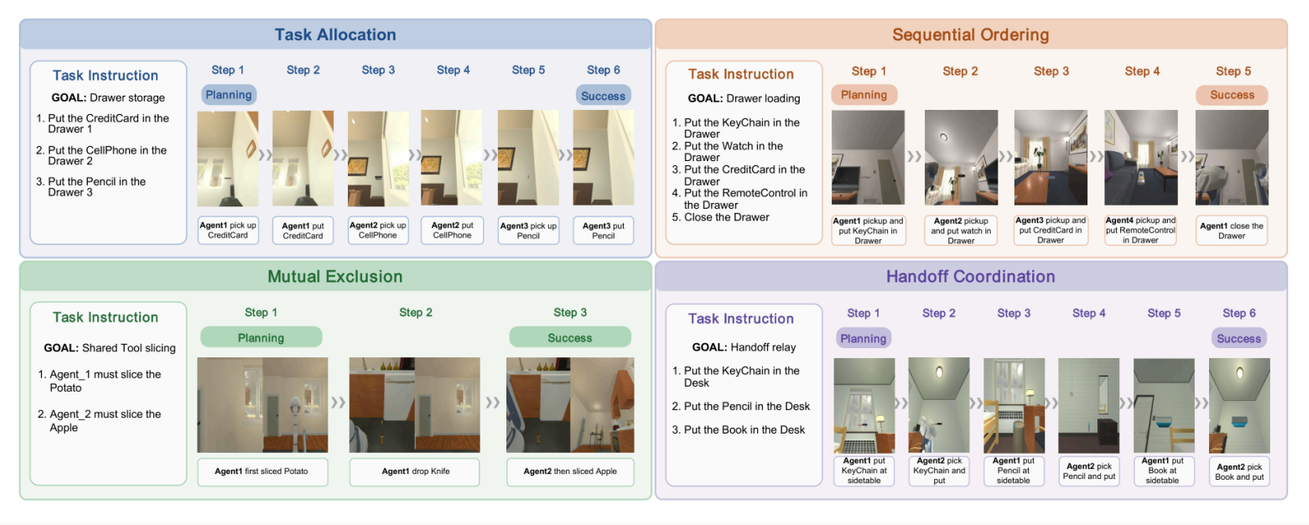
우리의 응용

“잘 따라왔나”뿐 아니라
필요할 때 역할 분담을 제안했는지 기록한다.

작게 적용할 지표

유효한 도움 요청, 중복 요청,
요청 후 실제 작업 진전

CoCoBench: 성공한 실행 안에도 비효율적인 협동이 숨어 있다



[R10] Chen et al. · arXiv 2026, preprint · PDF 5쪽, Figure 4

네 가지 진단 축

작업 분담 · 순서 준수
공유 자원 충돌 · 인계 조율

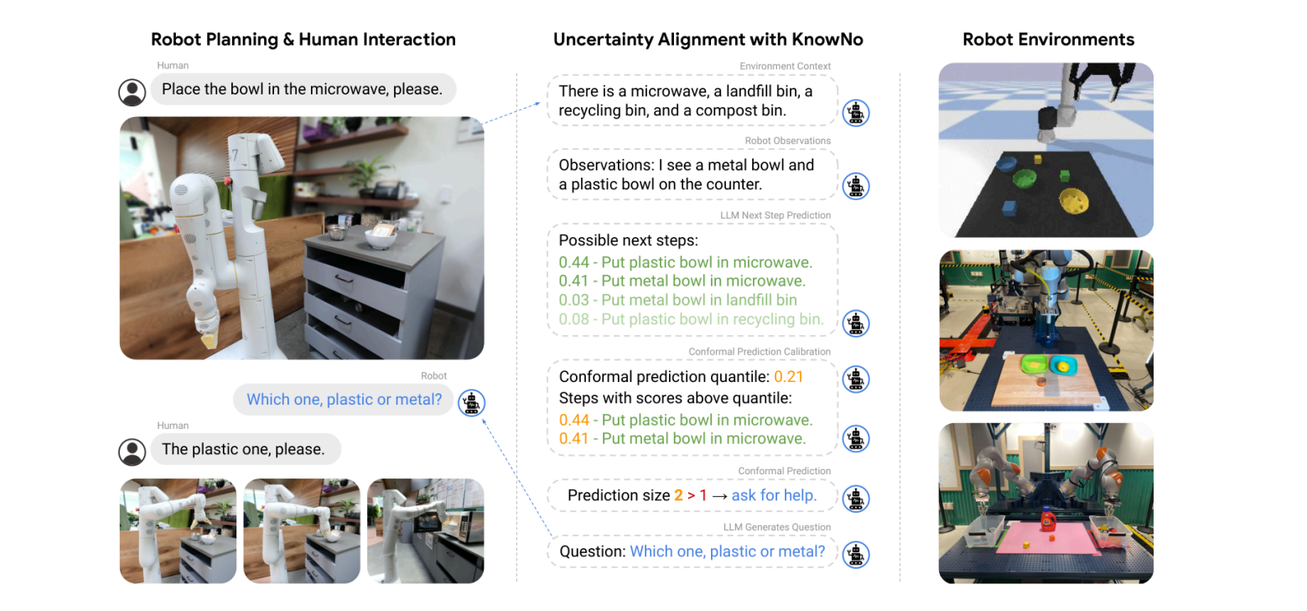
우리에게 유용한 이유

둘 다 같은 발판을 택하거나
상대가 기다리는 이유를 구분할 수 있다.

근거의 수준

2026.08 공개 사전 논문.
지표 아이디어로 참고하고 검증 수준을 구분한다.

KnowNo: 모호할 때 질문하는 정책은 작지만 설득력 있는 확장



[R5] Ren et al. · CoRL 2023 · PDF 2쪽, Figure 1

원 논문의 방법

보정된 예측 집합으로
행동을 정할지 도움을 요청할지 선택한다.

우리의 첫 실험

항상 실행 / 항상 질문 / 선택적 질문의
성공·추가 질문 수·완료 시간을 비교한다.

보장과 휴리스틱 구분

자기보고 confidence만으로
논문의 통계적 보장을 주장할 수 없다.

TEACh: 대화와 행동 연결을 가상 작업 훈련으로 확장할 수 있다



[R8] Padmakumar et al. · AAAI 2022 · PDF 3쪽, Figure 2

문헌에서 확인한 범위

가정 내 과제에서 서로 다른 정보를
가진 두 역할이 대화하며 행동한다.

우리의 응용 제안

“전원 차단 → 부품 이동 → 점검”처럼
절차와 역할이 있는 가상 과제 1개.

주장의 경계

절차 수행 시연은 가능성의 근거.
교육 효과나 실물 안전성은 별도 검증 대상.

응용 후보는 “새로 필요한 것”으로 비교

응용	재사용할 기반	추가 부담 / 이번 학기 판단
가상 절차 훈련	대화 지시·순서 조건·실행 피드백	절차 1개·판정 규칙 → 가장 권고
협동 NPC 도움 기능	기존 퍼즐·취소·질문·역할 분담	질문 정책·작은 사용성 시험 → 가까운 확장
가정 내 보조 시뮬레이션	부분 정보·대화·협동 평가	물체 조작·탐색이 새로 필요 → 후속
실물 로봇 보조	의사결정 인터페이스·로그 설계	센서·조작·장치 안전 검증 → 이번 범위 밖

응용 시연은 기존 연구를 설명하는 하나의 추가 장면으로.
새 플랫폼 개발이 핵심 비교 실험을 밀어내지 않게 한다.

[R8] Padmakumar et al., AAAI 2022 · [R2] Chang et al., ICLR 2025 · [R5] Ren et al., CoRL 2023

사용자 경험도 연구와 연결할 수 있다

AI가 언제 듣고, 무엇을 실행하며, 어떻게 멈추는지 보이게 한다.

설계 원칙의 적용	실제 플레이에서 확인할 것	기록할 근거
현재 상태 표시	듣는 중 / 판단 중 / 실행 중 / 확인 필요	사용자가 상태를 맞게 이해했는가
쉬운 취소와 수정	“멈춰”, “왼쪽 말고 오른쪽”	취소 지연·잘못된 재시작·수정 성공
불확실할 때 범위 축소	후보가 여럿이면 짧은 확인 질문	질문 부담과 오동작의 균형

소규모 사용성 관찰은 탐색적 근거로 보고한다. 만족도만으로 행동 모델 성능을 대체하지 않는다.

[R9] Amershi et al., CHI 2019

일정 제안 · 실제 마감에 맞춰 조정

약 8주를 가정한 실행 순서

기간	완료할 작업	다음 단계로 넘어갈 근거
W1 · 지금	최신 Live·사람 플레이, 구간 계측	입력부터 실제 실행까지 같은 기록으로 추적
W2	S0–S2 반복 실행·판정, 데이터 분리	시드·변화 주입·기준 성공 판정 재현
W3–4	주기/사건 정책 비교 + LoRA 시범 학습	호출 조건 동등성, 학습→추론 왕복 확인
W5–6	잠근 평가 세트로 네 조건 비교	성공·지연·비용과 대표 실패 근거 확보
W7	보조 실험 또는 응용 1개, 결과 정리	핵심 실험 완료 시에만 확장
W8	시연 안정화·보고서·재현 자료	버전·데이터·설정·영상·한계가 함께 남음

역할 유지 · 세부 과제 제안

첫 주에는 역할별 결과물을 하나의 실행 기록으로 합친다

기존 담당	우선 작업 제안	완료 기준
영호 · 조정기·대화	이벤트/요청/취소/모델/실행 타임라인	같은 trial_id로 변화부터 결과까지 연결
경민 · 게임·상태 필터	상태 변화 주입·초기화·실제 성공 판정	같은 시드로 같은 실패 장면 재현
준영 · 행동 AI·실행기	기본 모델 고정·학습/변환 시범 실행	모델·프롬프트·정밀도·설정 기록

학습 관문	W4까지 학습→동일 실행 경로가 안 되면 모델/학습 범위를 재협의
범위 관문	반복 판정이 불안정하면 응용보다 시나리오를 줄여 재현성 확보

최종 결과물은 “된다”를 넘어 “어떤 조건에서 왜”

01 사람과 끝까지 플레이할 수 있는
안정적인 협동 시연

02 같은 장면으로 다시 실행 가능한
작은 벤치마크와 이벤트 로그

03 재판단 × 학습의 효과와 한계를 보여 주는
비교표·변동 범위·실패 사례

04 가상 절차 훈련 또는 NPC 도움 기능의
응용 장면 하나

추천하는 다음 행동: S0–S2 판정과 계측부터 고정한다.

1차 출처 · 확인일 **2026.10.08**

참고문헌 1/2 · 구조·과제·학습

R1 Leveraging Dual Process Theory in Language Agent Framework for Real-time Simultaneous Human-AI Collaboration
Zhang et al. · ACL 2025

R2 PARTNR: A Benchmark for Planning and Reasoning in Embodied Multi-Agent Tasks
Chang et al. · ICLR 2025

R3 ADAPT: Benchmarking Commonsense Planning under Unspecified Affordance Constraints
Chen et al. · ACL 2026

R4 Collab-Overcooked: Benchmarking and Evaluating Large Language Models as Collaborative Agents
Sun et al. · EMNLP 2025

R5 Robots That Ask For Help: Uncertainty Alignment for Large Language Model Planners
Ren et al. · CoRL 2023

R6 LoRA: Low-Rank Adaptation of Large Language Models
Hu et al. · ICLR 2022 · 캡처: arXiv v2

1차 출처 · 확인일 **2026.10.08**

참고문헌 2/2 · 응용·실행 조건

R7 QLoRA: Efficient Finetuning of Quantized LLMs
Dettmers et al. · NeurIPS 2023

R8 TEACh: Task-Driven Embodied Agents That Chat
Padmakumar et al. · AAAI 2022

R9 Guidelines for Human-AI Interaction
Amershi et al. · CHI 2019

R10 CoCoBench: A Cooperative Coordination Benchmark for Embodied Multi-Agent Task Planning
Chen et al. · arXiv 2026, preprint · 사전 공개, 심사 완료로 간주하지 않음

R11 Qwen3-4B-Instruct-2507 model card
Qwen Team · Official model card

원문 PDF의 그림 9개를 직접 캡처했다.
수치 재현·성능 이전을 주장하지 않으며, 우리 적용안은 제안으로 구분한다.