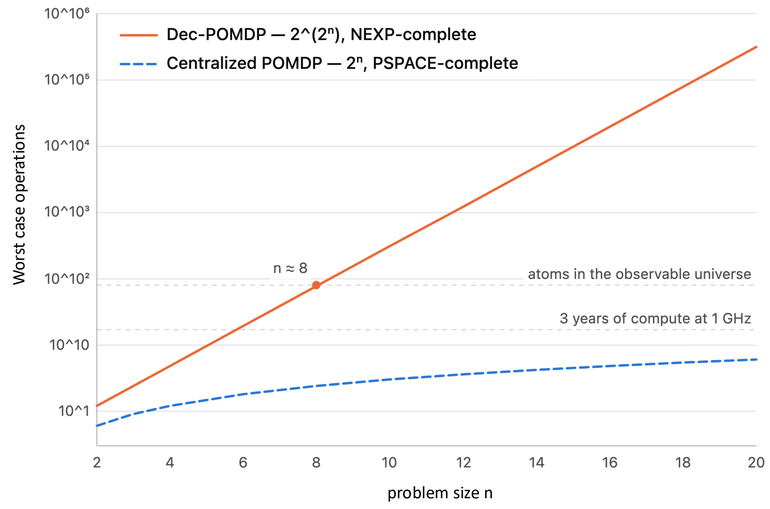
Embodied Multi-Agent Coordination by Aligning World Models Through Dialogue

Vardhan Dongre · Dilek Hakkani-Tür

Siebel School of Computing & Data Science, University of Illinois Urbana-Champaign

01 Motivation

Embodied agents act under **partial observability**: each agent sees only part of the environment and maintains a private world model. In PARTNR, the two agents also have asymmetric capabilities, so neither can solve many tasks alone.



Two-agent decentralized planning is NEXP-complete; a shared channel combines both partial observations into a joint view, reducing the problem toward centralized planning at polynomial space.

RESEARCH QUESTION

Do embodied LLM agents use dialogue to build a shared, grounded model of the world or do they merely appear coordinated?

Existing embodied multi-agent work reports task success without probing whether communication is meaningful. We close that gap with metrics defined directly over each agent's world graph.

02 Problem Setup

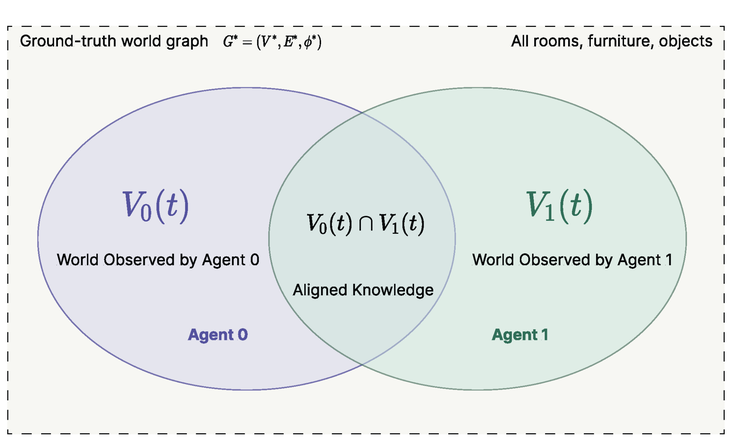
We extend **PARTNR** (household robotics in the Habitat simulator) with a natural-language dialogue channel. Two heterogeneous agents, a quadruped **robot** and a **humanoid**, share one scene under **asymmetric partial observability** .

Private world graphs. Each agent maintains a tree `house → room → furniture → object` , populated incrementally as it explores. Initial graphs are **disjoint**; neither can read the other's.

Asymmetric action spaces. State-change actions (`Clean`, `Fill`, `Pour`, `PowerOn`) are exclusive to the humanoid, so no agent can solve a task alone.

Shared step budget. Every action, motor or communicative, consumes one of $T = 2{,}000$ simulator steps.

Category	Agent 0: Robot	Agent 1: Humanoid
Navigation	`Navigate`, `Explore`	`Navigate`, `Explore`
Manipulation	`Pick`, `Place`, `Rearrange`, `Open`, `Close`	`Pick`, `Place`, `Rearrange`, `Open`, `Close`
State-change	*not available*	`Clean`, `Fill`, `Pour`, `PowerOn/Off`,
Perception	`FindObject`, `FindReceptacle`, `FindRoom`, `FindAgent`	`FindObject`, `FindReceptacle`, `FindRoom`, `FindAgent`
Control	`Wait`, `Done`	`Wait`, `Done`



Alignment as set overlap. Observation convergence is the Jaccard overlap of the two node sets.

OC **Observation Convergence**

Jaccard overlap of directly-observed node sets $V_0(t)$, $V_1(t)$. Do private world models align over time?

BC **Belief Convergence & Δ_{align}**

Extends OC to nodes an agent was *told* about. Δ_{align} = BC − OC isolates what dialogue adds beyond co-exploration.

BSM **Belief-Sensitive Messaging**

Does an utterance target what the partner lacks? Do agents' model what their partner knows?

Two communication architectures:

`SC` (synchronous-costed): `SendMessage` is an action, so speaking means not acting that turn.

`ACF` (asynchronous cost-free): messaging rides along with an action and partner messages auto-inject at the next replan.

SC (synchronous, costed):		
Dialogue	`SendMessage`, `ReadMessages`	`SendMessage`, `ReadMessages`
ACF (asynchronous, free):		
Dialogue	`SendMessage` (dual-output)	`SendMessage` (dual-output)

03 Empirical Findings & Research Gaps

HEADLINE RESULT

LESS CONFLICT. WORSE OUTCOMES. Dialogue makes agents **look coordinated**, but they **solve fewer tasks**. Across all three LLMs, every tested dialogue setup sharply reduces redundant actions while lowering task success. Making messages free or encouraging agents to speak more does not recover performance.

↓ 41–93 pp

ACTION CONFLICT

Agents choose the same action on the same target far less often.

↓ 5–37 pp

TASK SUCCESS

No tested model benefits from dialogue (paired Wilcoxon, $p < 0.01$).

60–69%

UNGROUNDED REFERENCES

Most entity references are hallucinated or forecasted, grounded in neither world graph.

+0.06 to +0.19

GROUNDED-ONLY ALIGNMENT

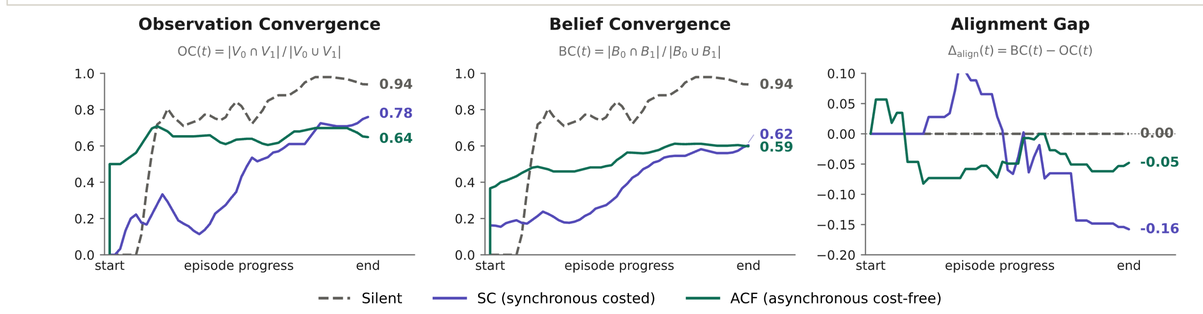
Once ungrounded references are removed, dialogue improves belief alignment in every condition.

The conflict–success tradeoff holds across all three LLMs

SC − SILENT	Δ SUCCESS	Δ PARTIAL CREDIT	Δ CONFLICT
Sonnet	−0.177	−0.039	−0.926
Haiku	−0.368	−0.331	−0.602
Mistral-L	−0.050	−0.012	−0.405

SC versus Silent on paired episodes

CONDITION	SR	PC	R_{CONF}	TURNS/EP
Silent	0.706	0.755	0.941	0.00
SC	0.529	0.716	0.015	3.76
SC*	0.471	0.615	0.000	4.18
ACF	0.412	0.520	0.224	13.35
ACF*	0.392	0.410	0.014	19.15



DIAGNOSIS 1 · GROUNDED MESSAGES ALIGN; UNGROUNDED REFERENCES UNDO THE GAIN

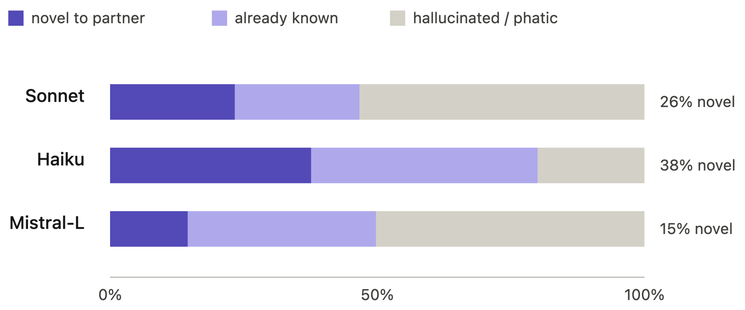
Pooled Δ_{align} is negative in every dialogue condition (−0.15 SC, −0.10 SC*, −0.05 ACF; $p \approx 0.002$), yet $\Delta^{grounded}_{align}$ is positive throughout (+0.19, +0.17, +0.06). **Hallucinated references inflate belief sets erasing the alignment produced by grounded messages.**

CONDITION	OC	BC	Δ_{ALIGN}	Δ^{GRND}
Silent	0.92	0.92	0.00	0.00
SC	0.67	0.51	−0.15	+0.19
SC*	0.72	0.57	−0.15	+0.17
ACF	0.65	0.61	−0.05	+0.06
ACF*	0.75	0.63	−0.07	+0.12

The sign flip is the key diagnosis: *grounded dialogue aligns beliefs, but ungrounded references reverse the net effect.*

What Do Agents Talk About?

Dialogue Composition by Mention Class



Most references fall in the **observed-by-neither** category; only sender-only mentions directly transfer new observations to the partner.

DIAGNOSIS 2 · FREE MESSAGING EXPOSES NEW PATHOLOGIES

Under SC, status messages are the locus of information transfer ($BSM_{STATUS} \approx +0.001$) while plan messages skew toward already-shared entities (−0.155). Under ACF this structure **collapses** : pooled BSM falls to −0.232 and STATUS flips to −0.194. Agents speak far more, but to no audience model in particular.

A typical failure cascade:

`Plan around an unseen entity` → `Confirm progress that never happened` → `Propagate the false belief` → `Stop with the task incomplete`

04 Key Takeaways

Coordination is not alignment

Agents can avoid redundant actions while their world models diverge. Success and conflict metrics show what happened; world-model metrics reveal why.

Grounding is the bottleneck

The grounded-only alignment gap is positive in every condition. Ungrounded and forecasted references reverse that gain, turning potentially useful communication into belief misalignment.

Policies were never trained for coordination via dialogue

Neither cost-free messaging nor stronger prompts recover success. Embodied collaboration needs policies trained for grounded reference, partner-belief tracking, and persistent common ground not simply chat models instructed to speak more.

2026 SIGdial Atlanta, GA, USA

Presented at the 27th Meeting of the Special Interest Group on Discourse and Dialogue