

# Preliminary Examination 2021

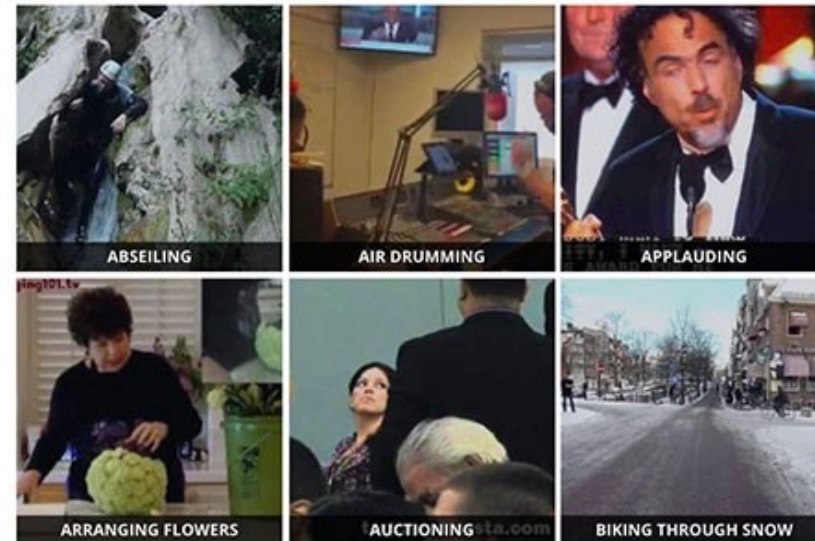
Unnat Jain

<https://unnat.github.io/>

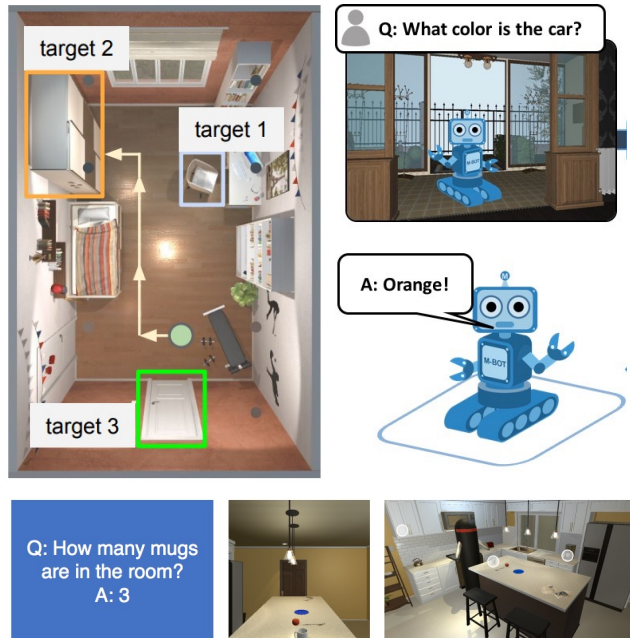


Chairs: Alexander Schwing and Svetlana Lazebnik  
Committee: Derek Hoiem, Kristen Grauman, Nan Jiang

# Visual Embodied Agents



# Visual Embodied Agents

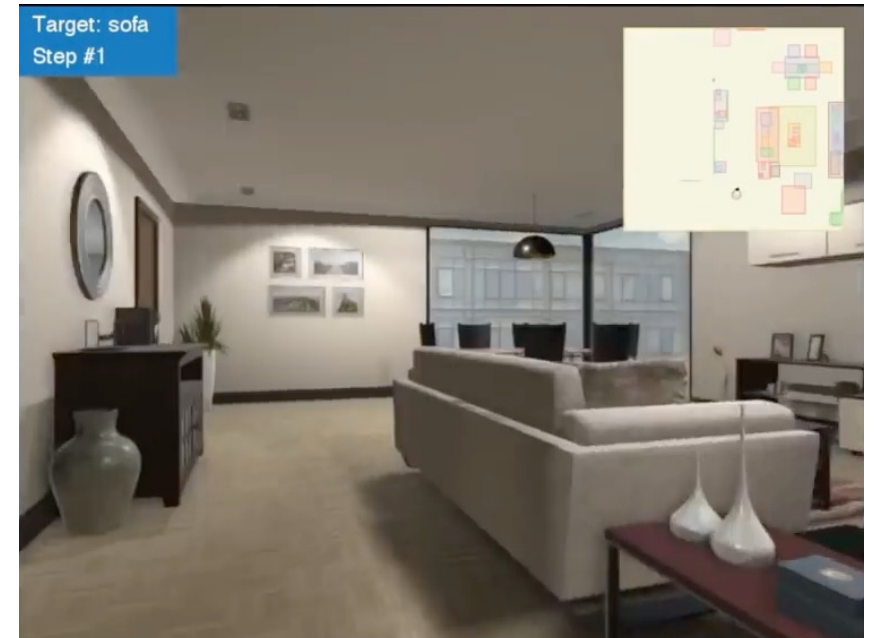


Visual Navigation

Task completion

Question Answering

Instruction Following



# Visual Embodied Agents

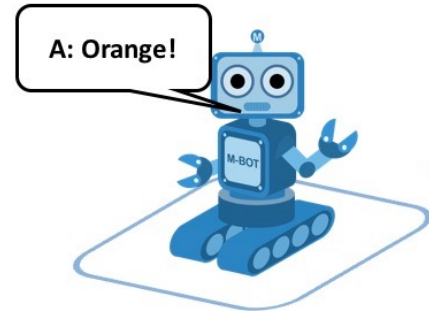
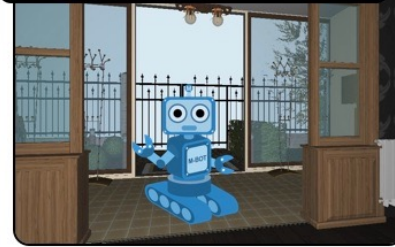


What is the mustache made of?

Visual Question Answering



Q: What color is the car?



Embodied Question Answering



# Collaborative Embodied Agents

AI Agents that can collaborate  
in virtual visual worlds





# Collaborative Embodied Agents

## Two Body Problem

CVPR 2019 (oral)

## SYNC Policies

ECCV 2020 (spotlight)

## GRIDTOPIX

(ongoing work)



# Collaborative Embodied Agents

## Two Body Problem

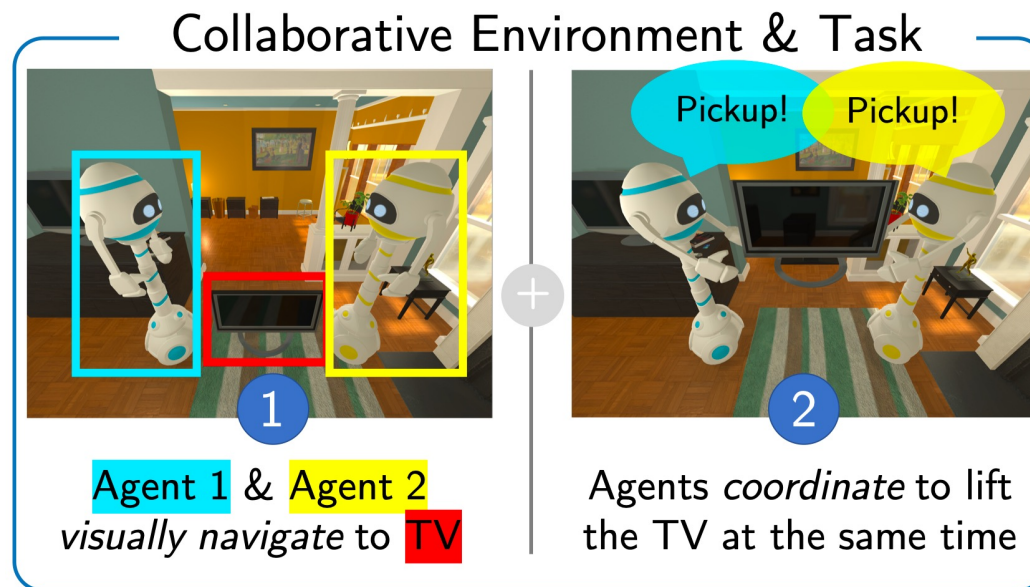
CVPR 2019 (oral)

## SYNC Policies

ECCV 2020 (spotlight)

## GRIDTOPIX

(ongoing work)



1. First collaborative embodied task - FurnLift





# Collaborative Embodied Agents

## Two Body Problem

CVPR 2019 (oral)

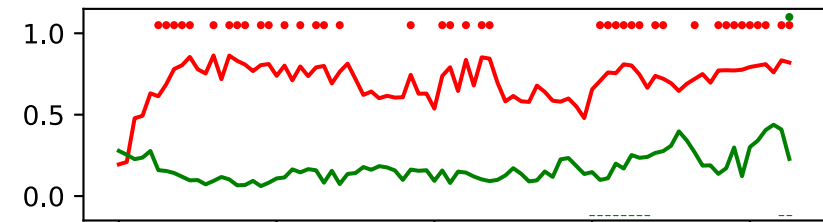
## SYNC Policies

ECCV 2020 (spotlight)

## GRIDTOPIX

(ongoing work)

*Talk (Agent 1)*  
*Talk (Agent 2)*



“I am near TV!”

2. Interpretation of emergent communication



# Collaborative Embodied Agents

## Two Body Problem

CVPR 2019 (oral)

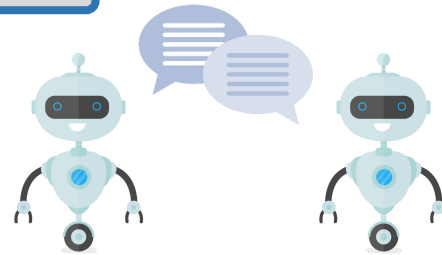
## SYNC Policies

ECCV 2020 (spotlight)

## GRIDTOPIX

(ongoing work)

Explicit



Explicitly sending messages to communicate

Implicit



Visibility of other agent communicates information

## 3. Effect of communication



# Collaborative Embodied Agents

## Two Body Problem

CVPR 2019 (oral)

## SYNC Policies

ECCV 2020 (spotlight)

## GRIDTOPIX

(ongoing work)



4. Intricately coordinated embodied task - FurnMove

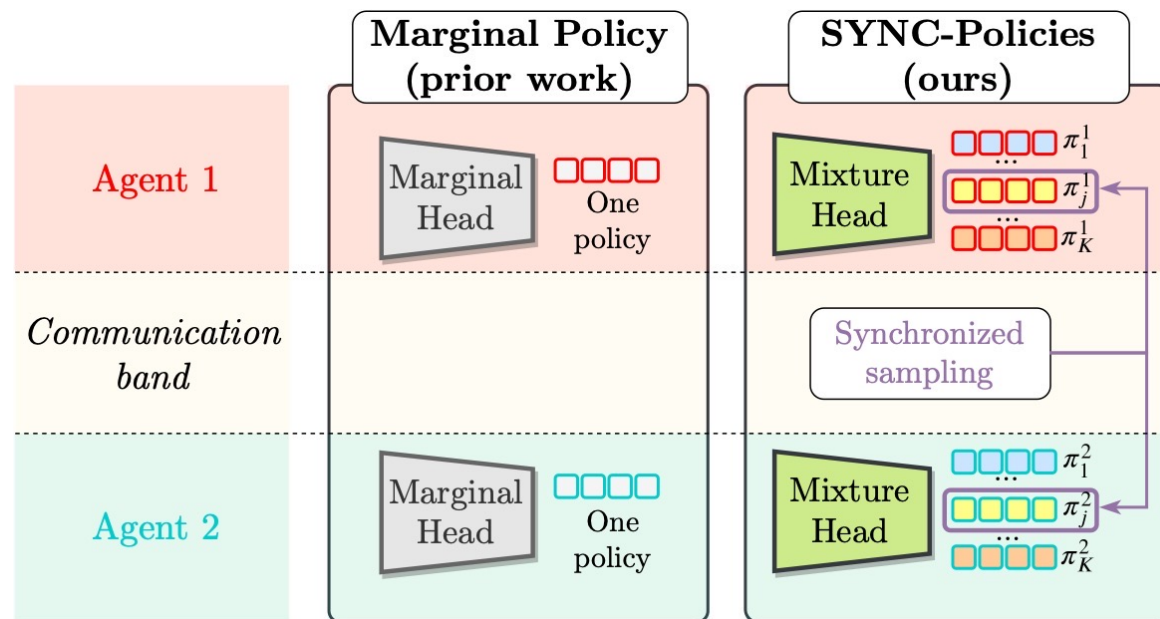


# Collaborative Embodied Agents

Two Body Problem  
CVPR 2019 (oral)

**SYNC Policies**  
ECCV 2020 (spotlight)

**GRIDTOPIX**  
(ongoing work)



5. Richer representation of multi-agent policy



# Collaborative Embodied Agents

## Two Body Problem

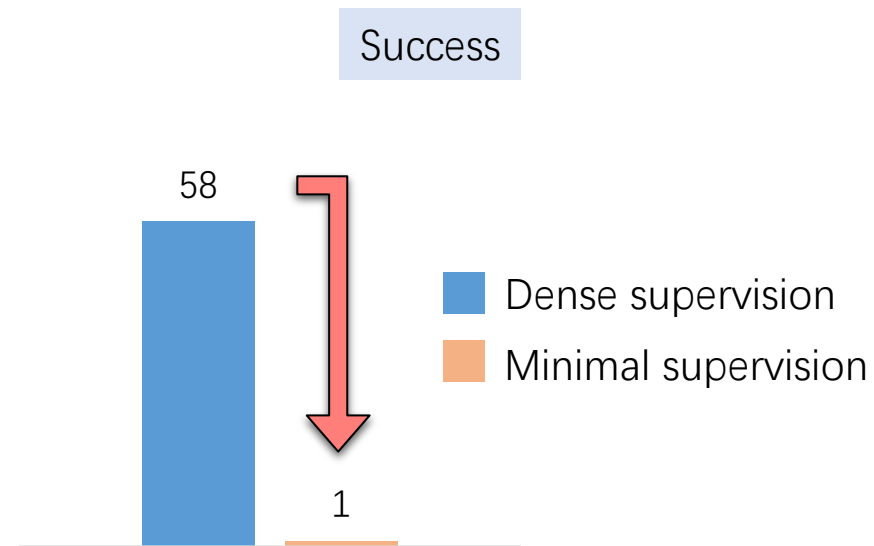
CVPR 2019 (oral)

## SYNC Policies

ECCV 2020 (spotlight)

## GRIDTOPIX

(ongoing work)



6. Learning policies from minimal supervision

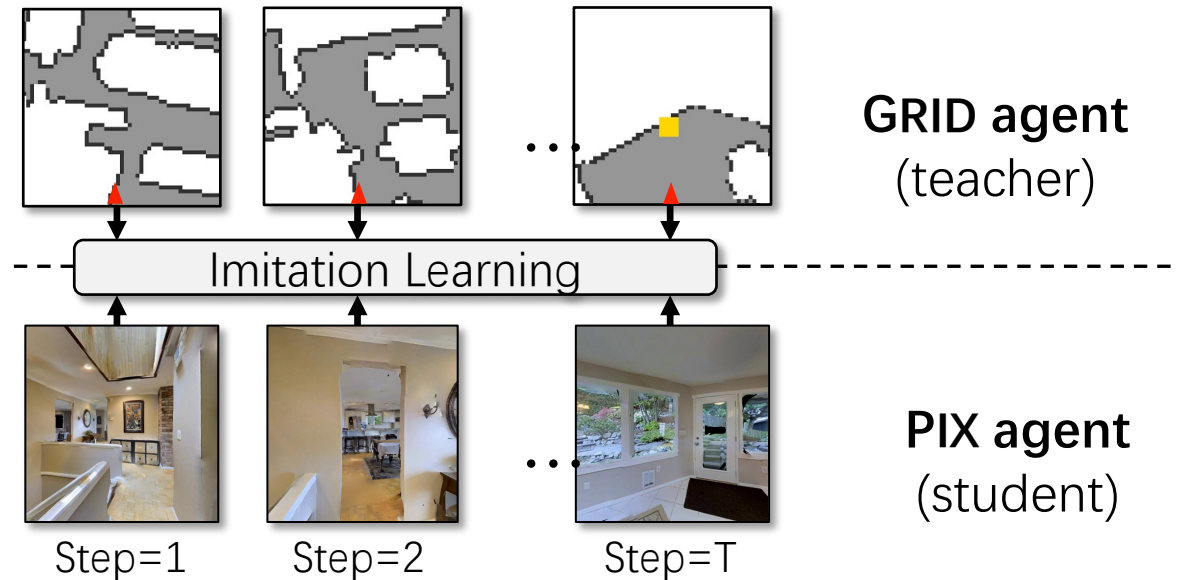


# Collaborative Embodied Agents

**Two Body Problem**  
CVPR 2019 (oral)

**SYNC Policies**  
ECCV 2020 (spotlight)

**GRIDTOPIX**  
(ongoing work)



7. Leveraging perfect-perception gridworlds for training



# Collaborative Embodied Agents

## Two Body Problem

CVPR 2019 (oral)

## SYNC Policies

ECCV 2020 (spotlight)

## GRIDTOPIX

(ongoing work)

1. First collaborative embodied task – FurnLift
2. Interpretation of emergent communication
3. Effect of communication
4. Intricately coordinated embodied task – FurnMove
5. Richer representation of multi-agent policy
6. Learning policies from minimal supervision
7. Leveraging perfect-perception gridworlds for training



# Collaborative Embodied Agents

## Two Body Problem

CVPR 2019 (oral)

## SYNC Policies

ECCV 2020 (spotlight)

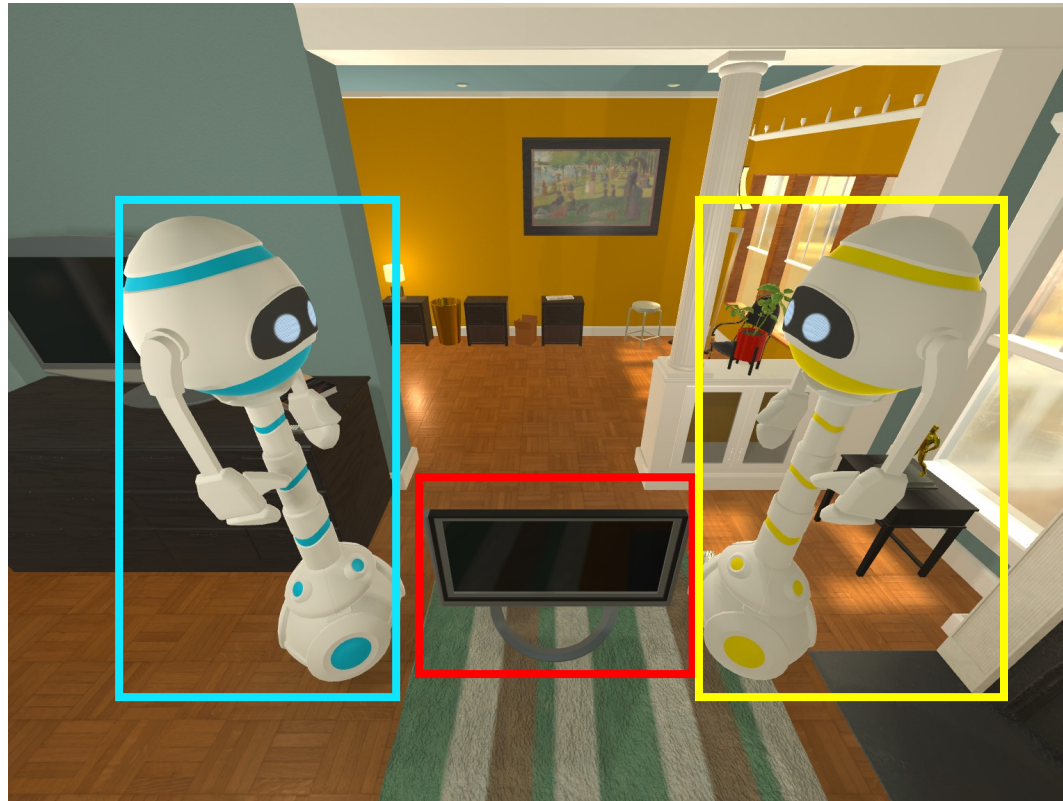
## GRIDTOPIX

(ongoing work)

1. First collaborative embodied task – FurnLift
2. Interpretation of emergent communication
3. Effect of communication
4. Intricately coordinated embodied task – FurnMove
5. Richer representation of multi-agent policy
6. Learning policies from minimal supervision
7. Leveraging perfect-perception gridworlds for training



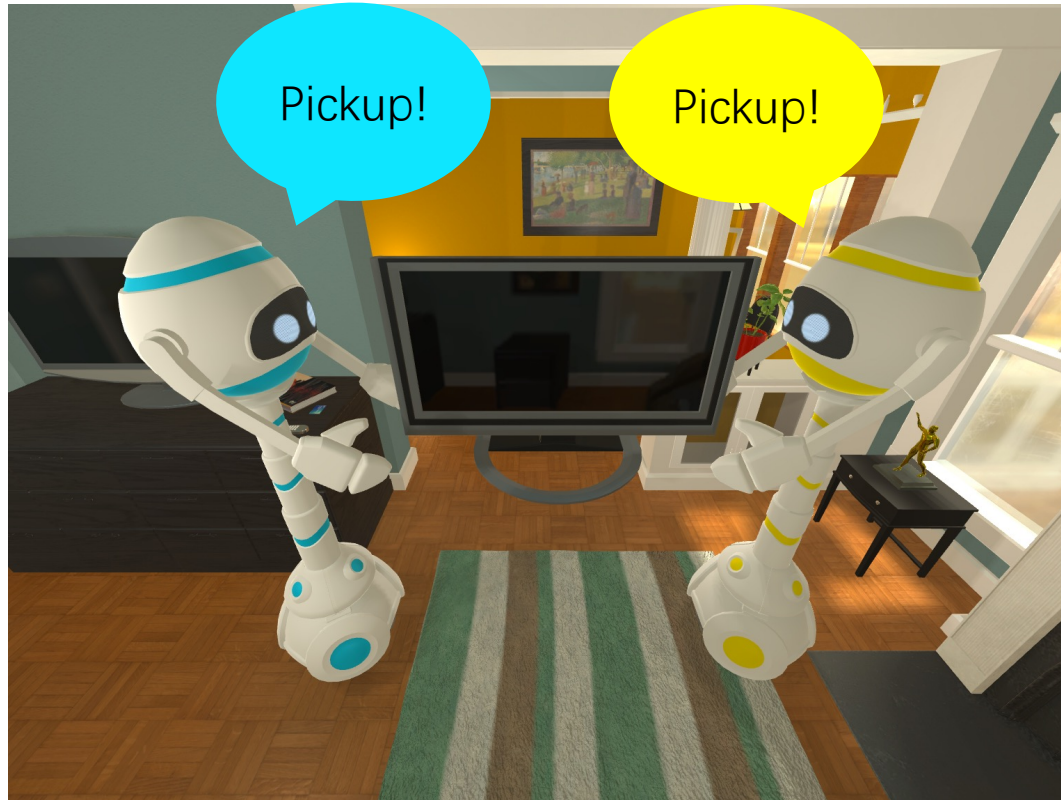
# FurnLift Task



Agent 1 and Agent 2 *visually*  
navigate to the TV

\* Agents have only egocentric visual inputs

# FurnLift Task



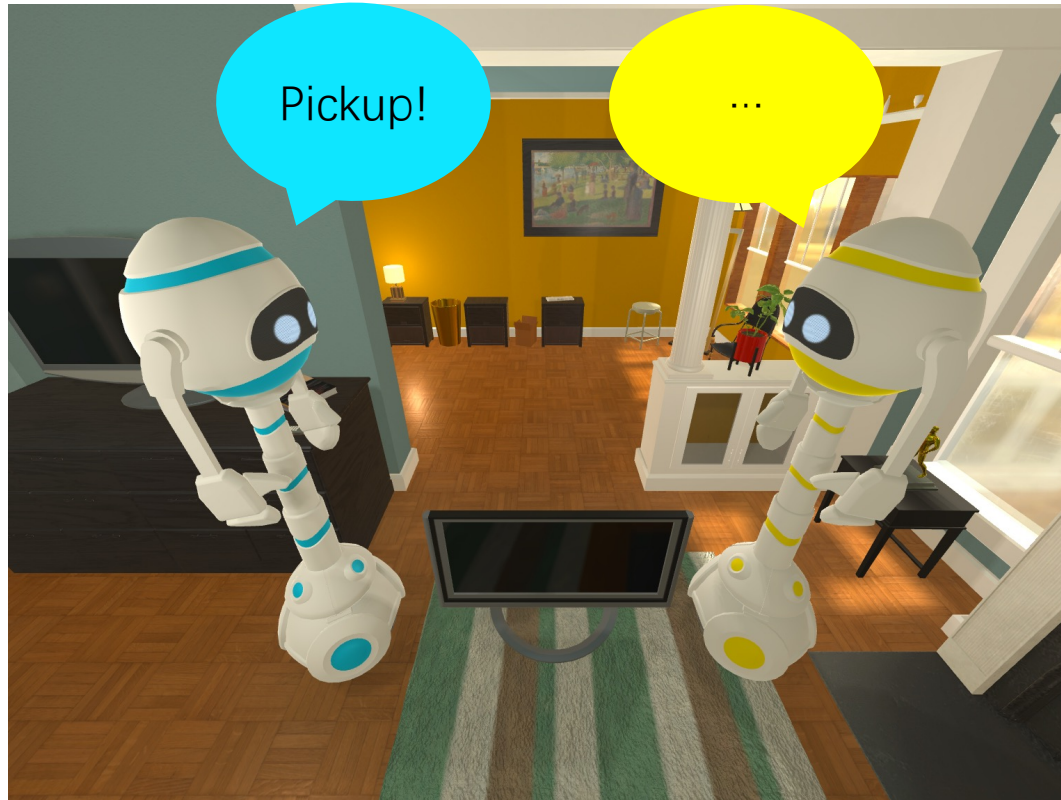
Agent 1 and Agent 2 *visually*  
navigate to the TV

+

Agents *coordinate* to lift/pickup  
the TV at the same time

\* Agents have only egocentric visual inputs

# FurnLift Task



\* Agents have only egocentric visual inputs

Agent 1 and Agent 2 *visually*

*navigate* to the TV

+

Agents *coordinate* to lift/pickup

the TV at the same time

# Agent observations

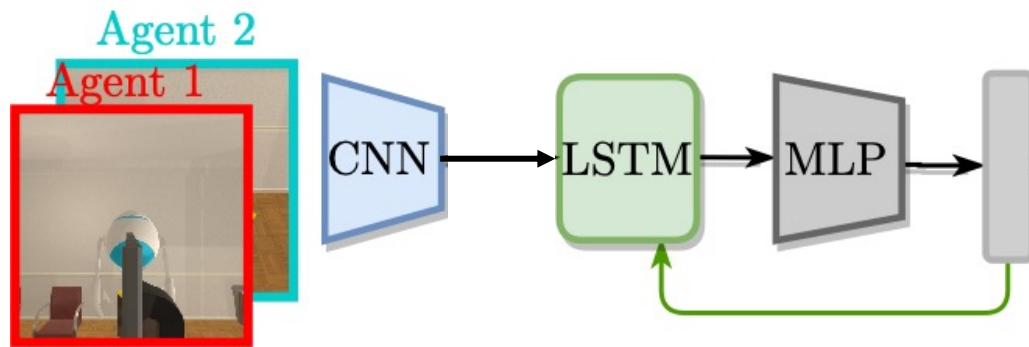


## Top-down view

Not available to the agents

(for illustration only)

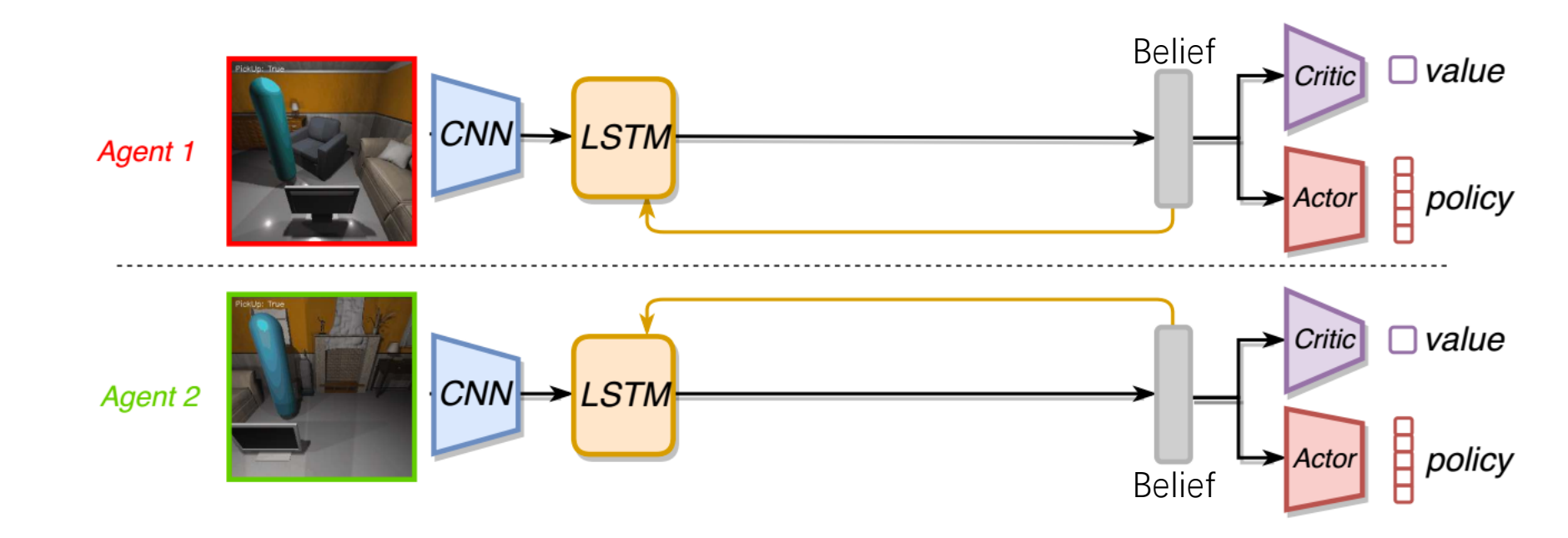
# Agent Policy for FurnLift



Central agent

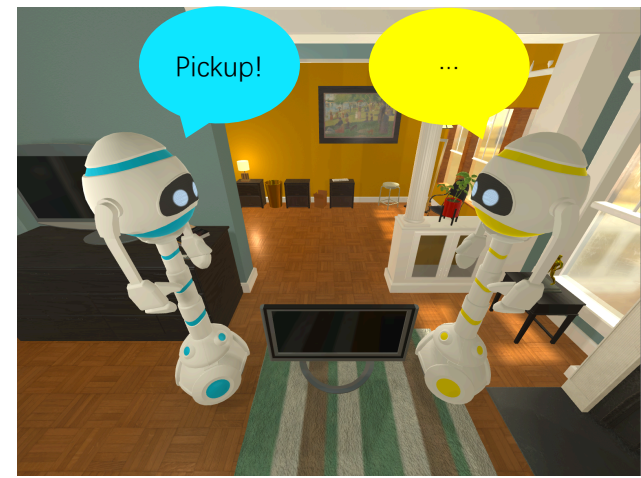
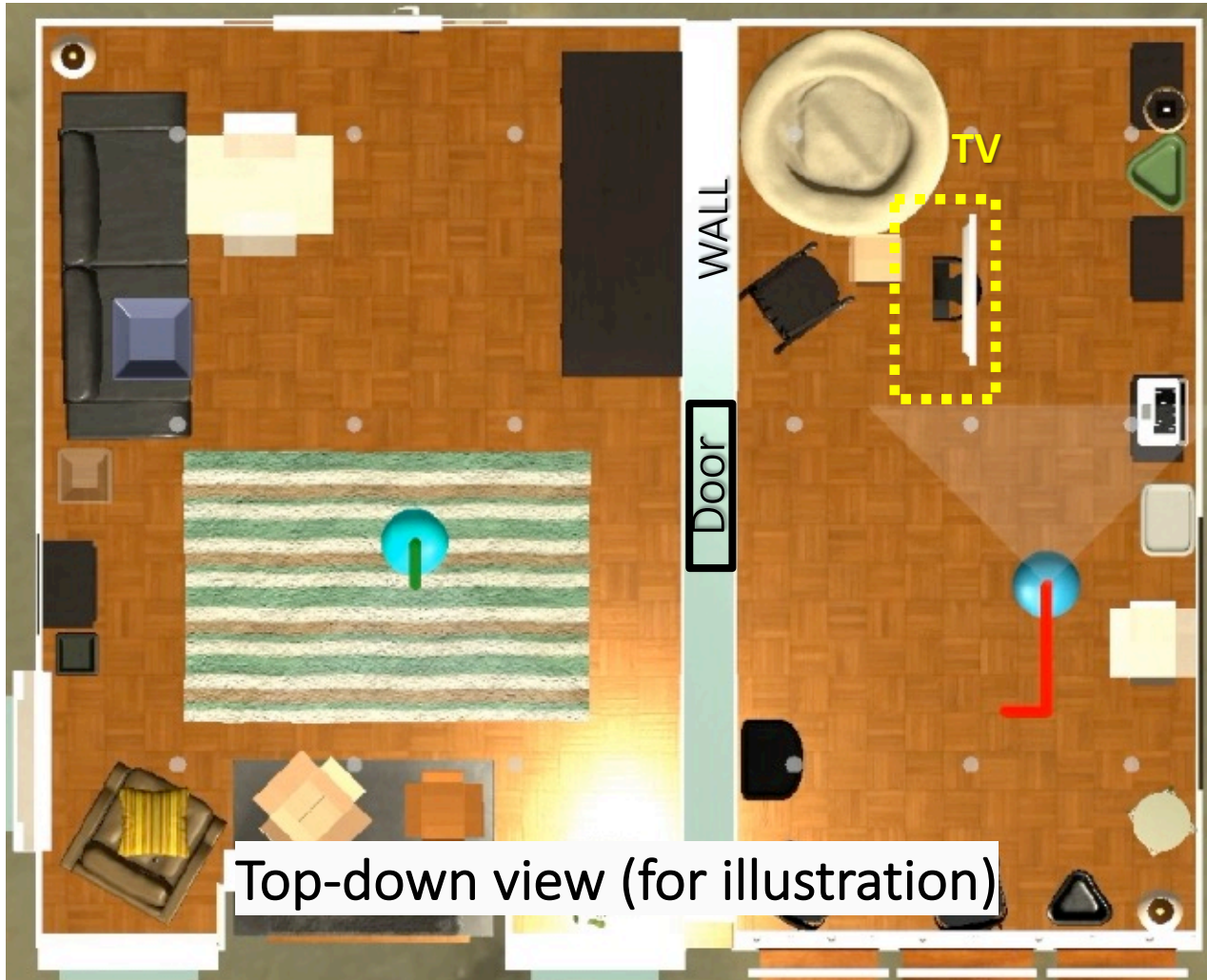
- ✗ Model complexity
- ✗ Policy parameters
- ✗ Comm. bandwidth

# Agent Policy for FurnLift



Decentral agent

# FurnLift Task



**Agent 1** and **Agent 2**

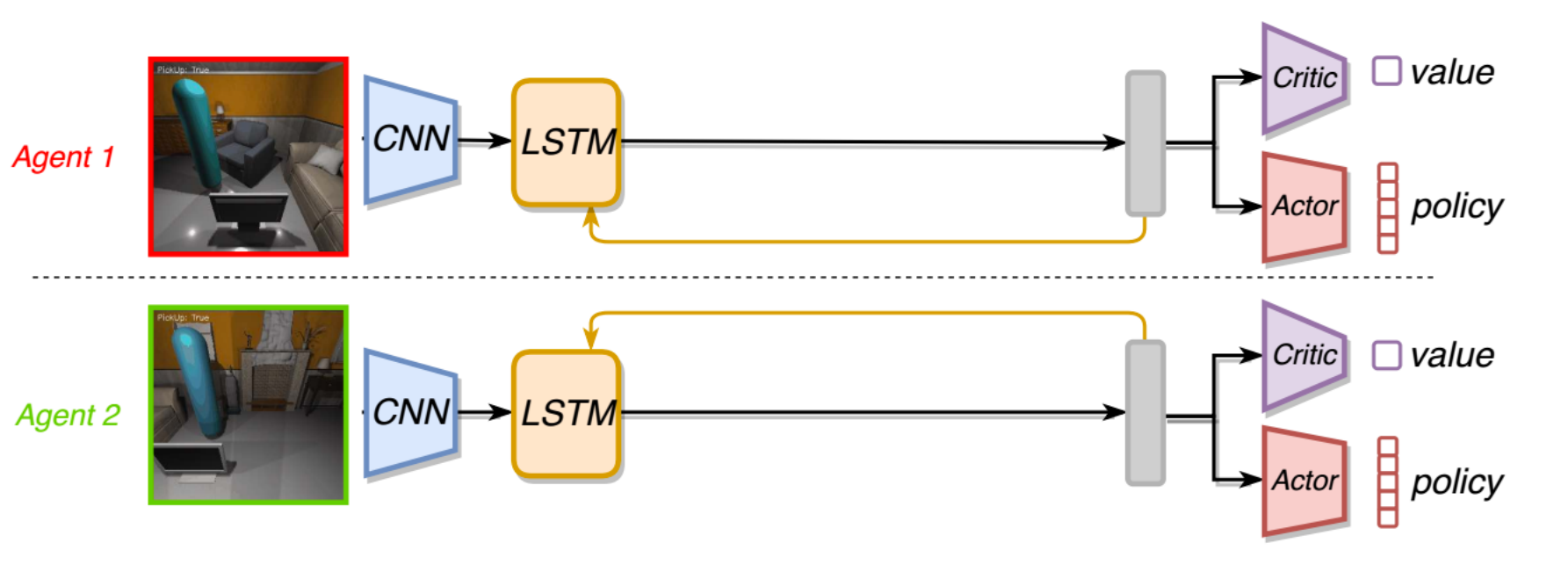
1. Navigate to TV
2. Team Lift

**Agent 1** quickly finds it

**Agent 2** is on the wrong side

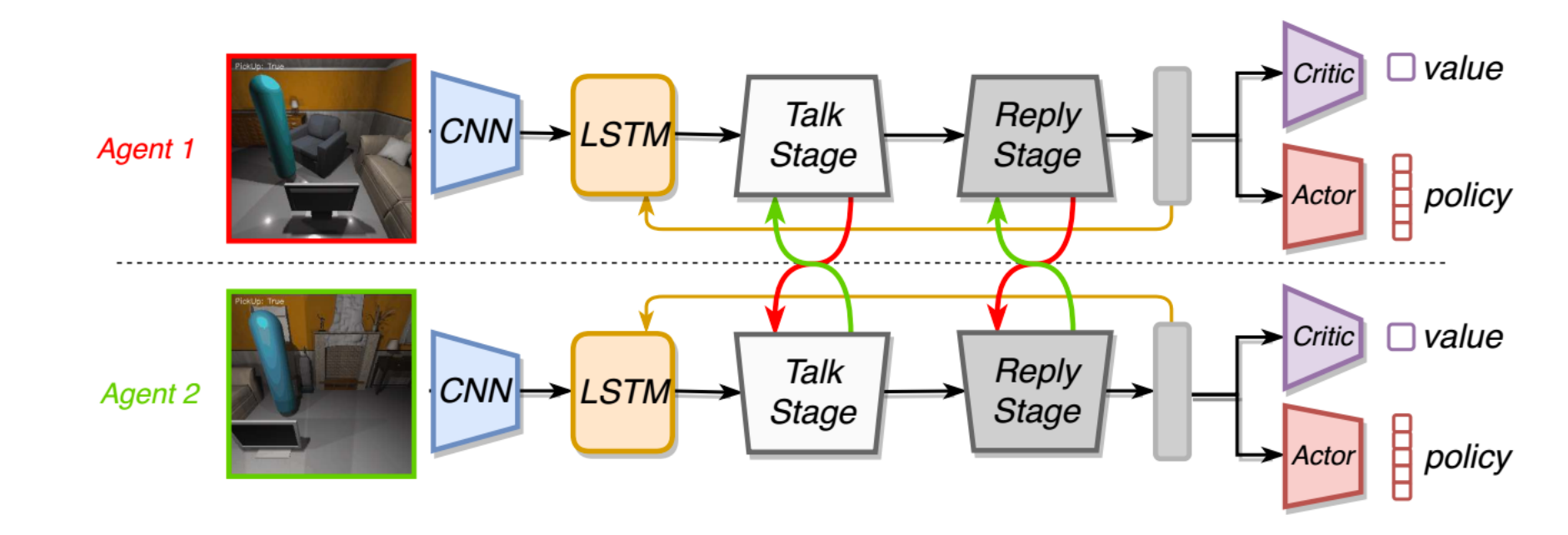
Need for communication

# Two Body Network



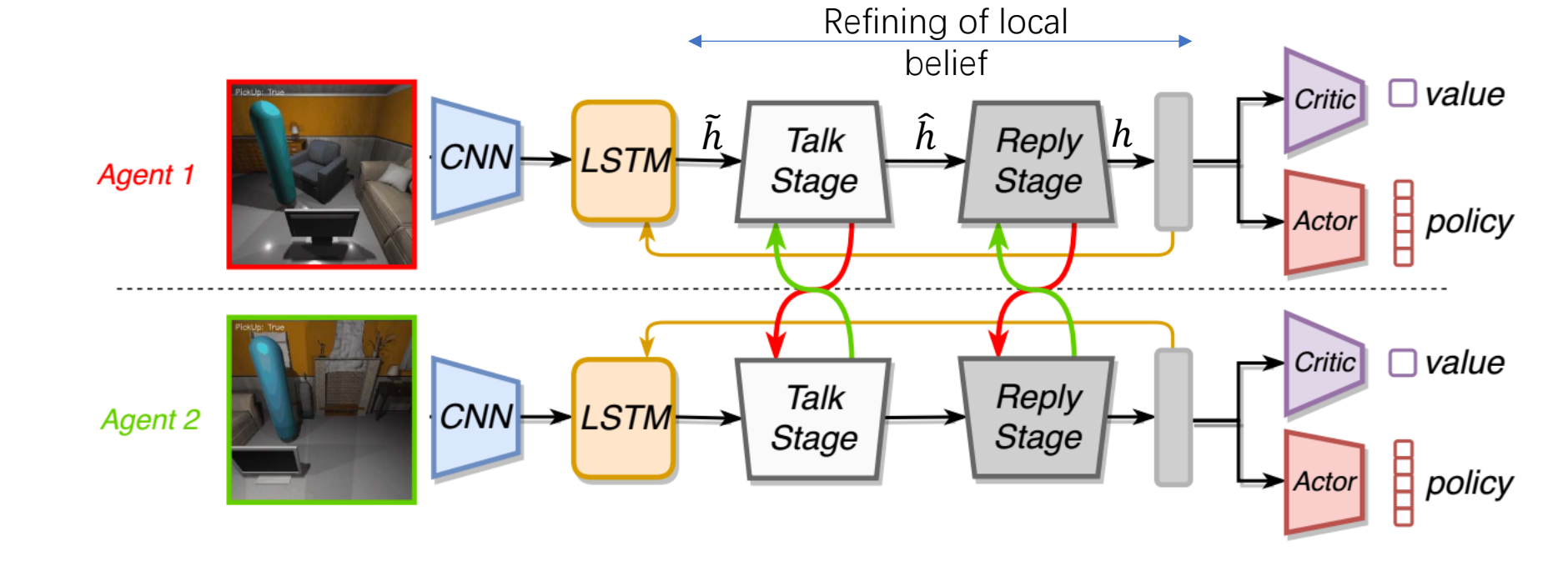


# Two Body Network



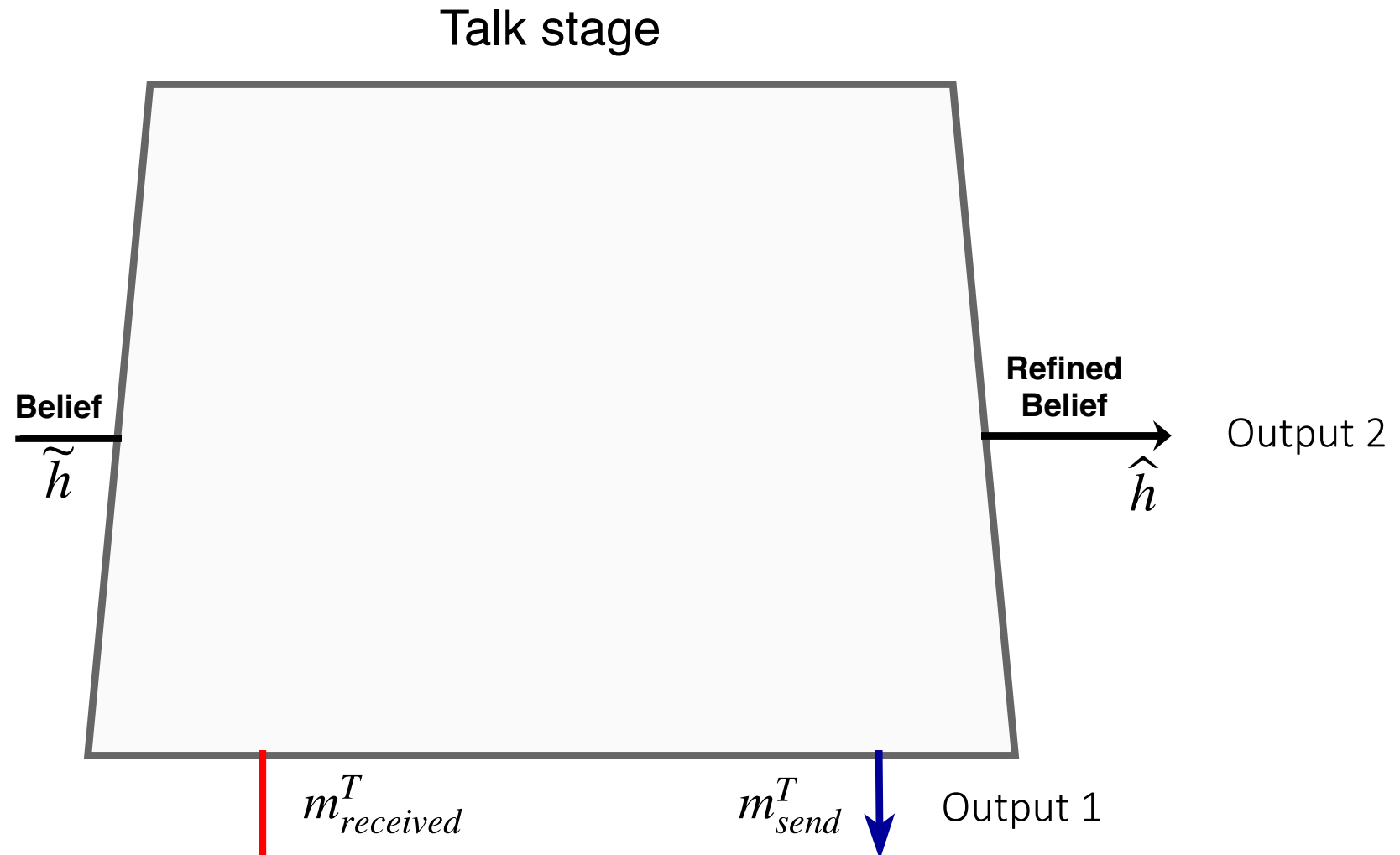
Message-based or 'explicit communication'

# Two Body Network

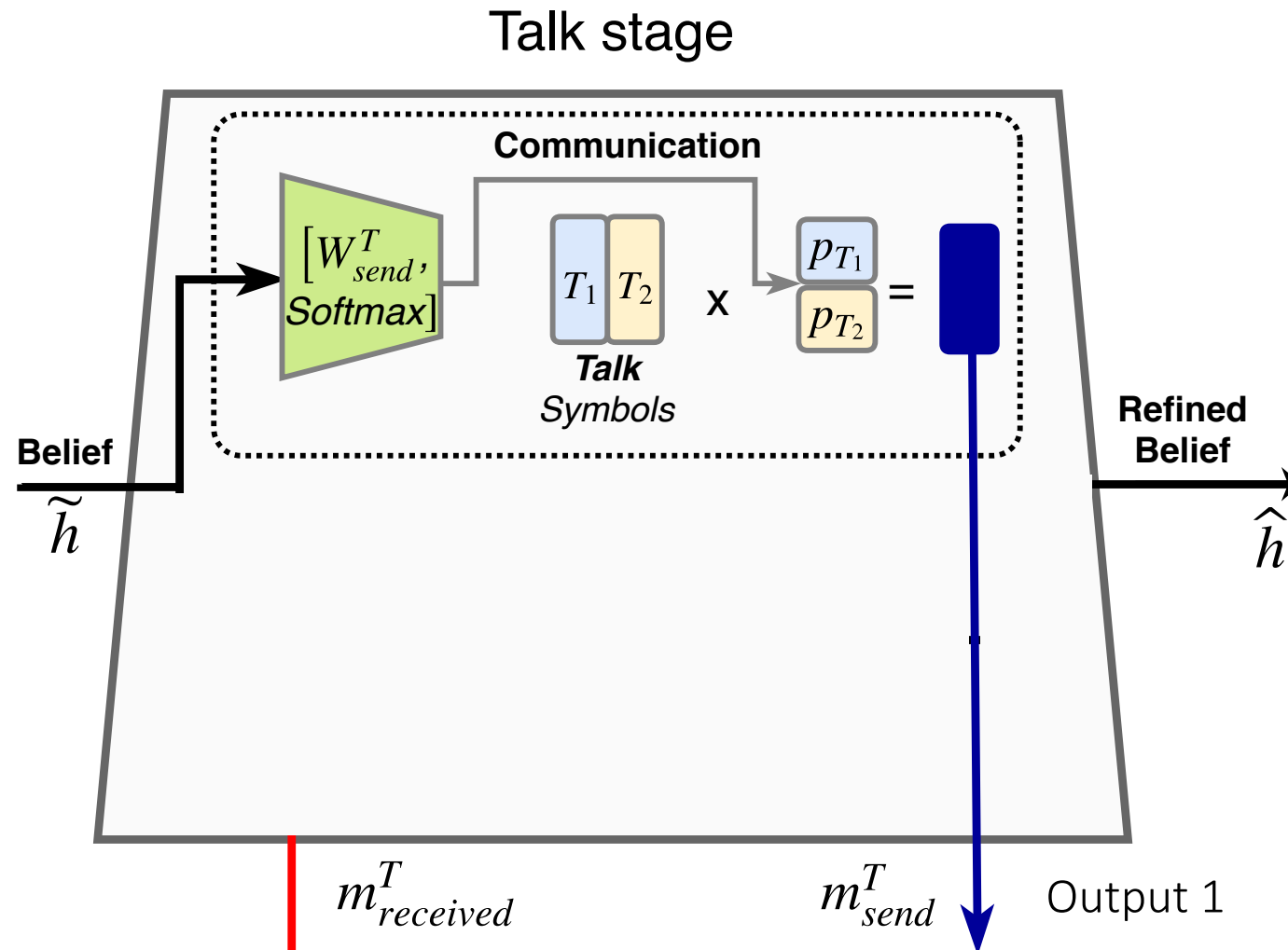


Message-based or 'explicit communication'

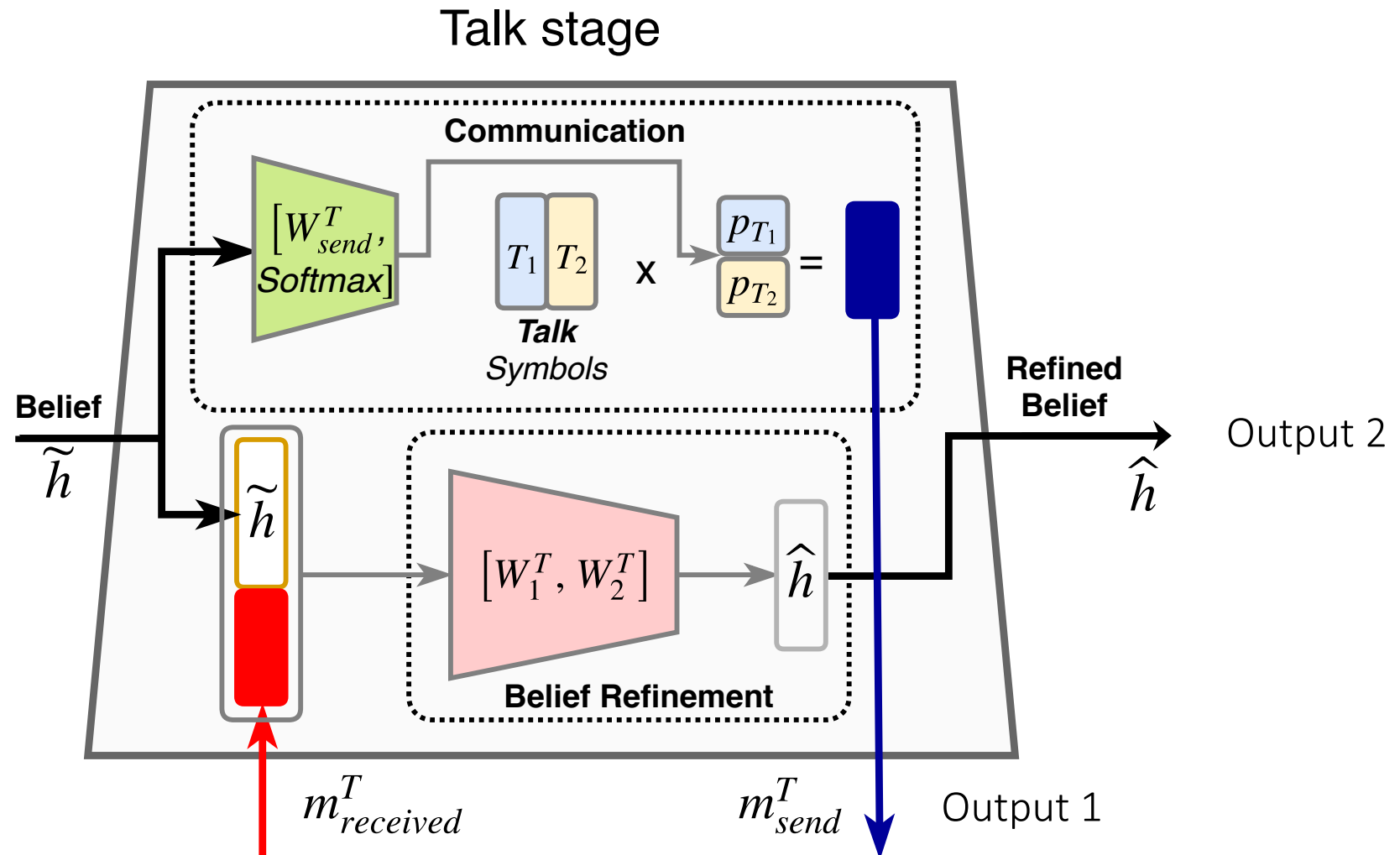
# Communication and Belief Refinement



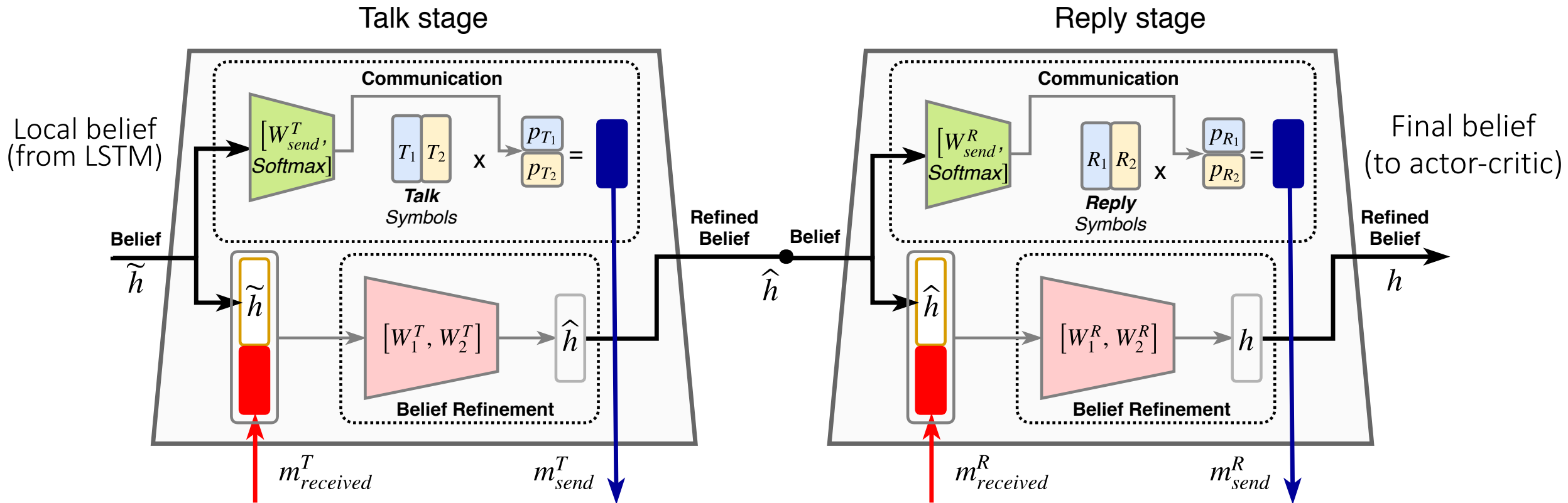
# Communication and Belief Refinement



# Communication and Belief Refinement



# Talk and reply modules



# Explicit Communication Helps

✗ Without  
explicit communication



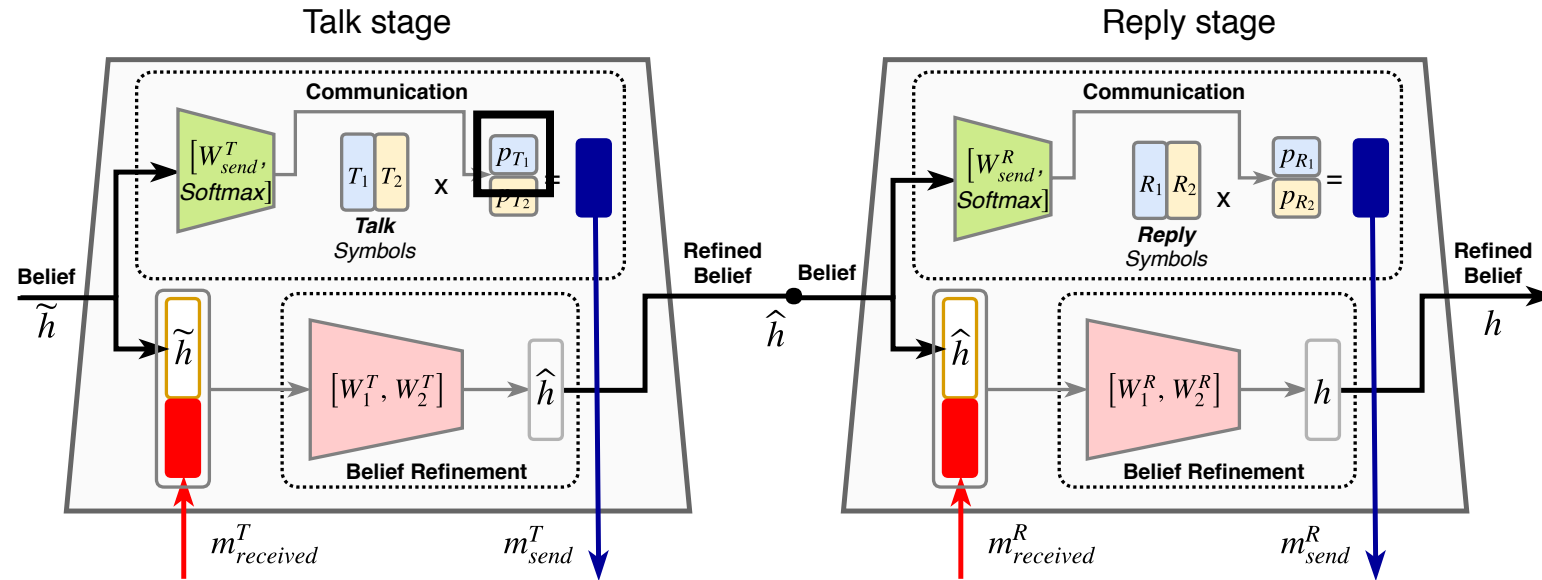
Total steps: 165  
Unsuccessful pickups: 6

✓ With  
explicit communication

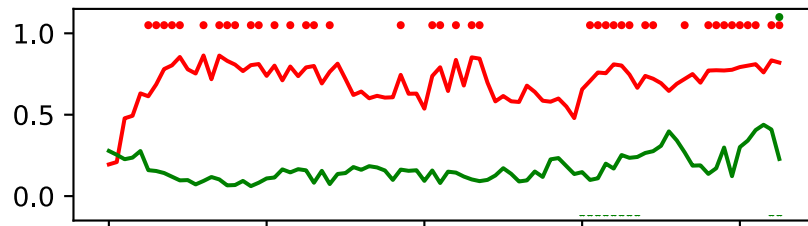


Total steps: 86  
Unsuccessful pickups: 0

# Interpretation of messages

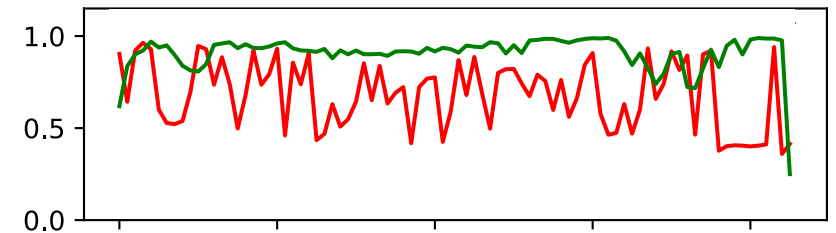


Talk (Agent 1)  
Talk (Agent 2)



“I am near TV!”

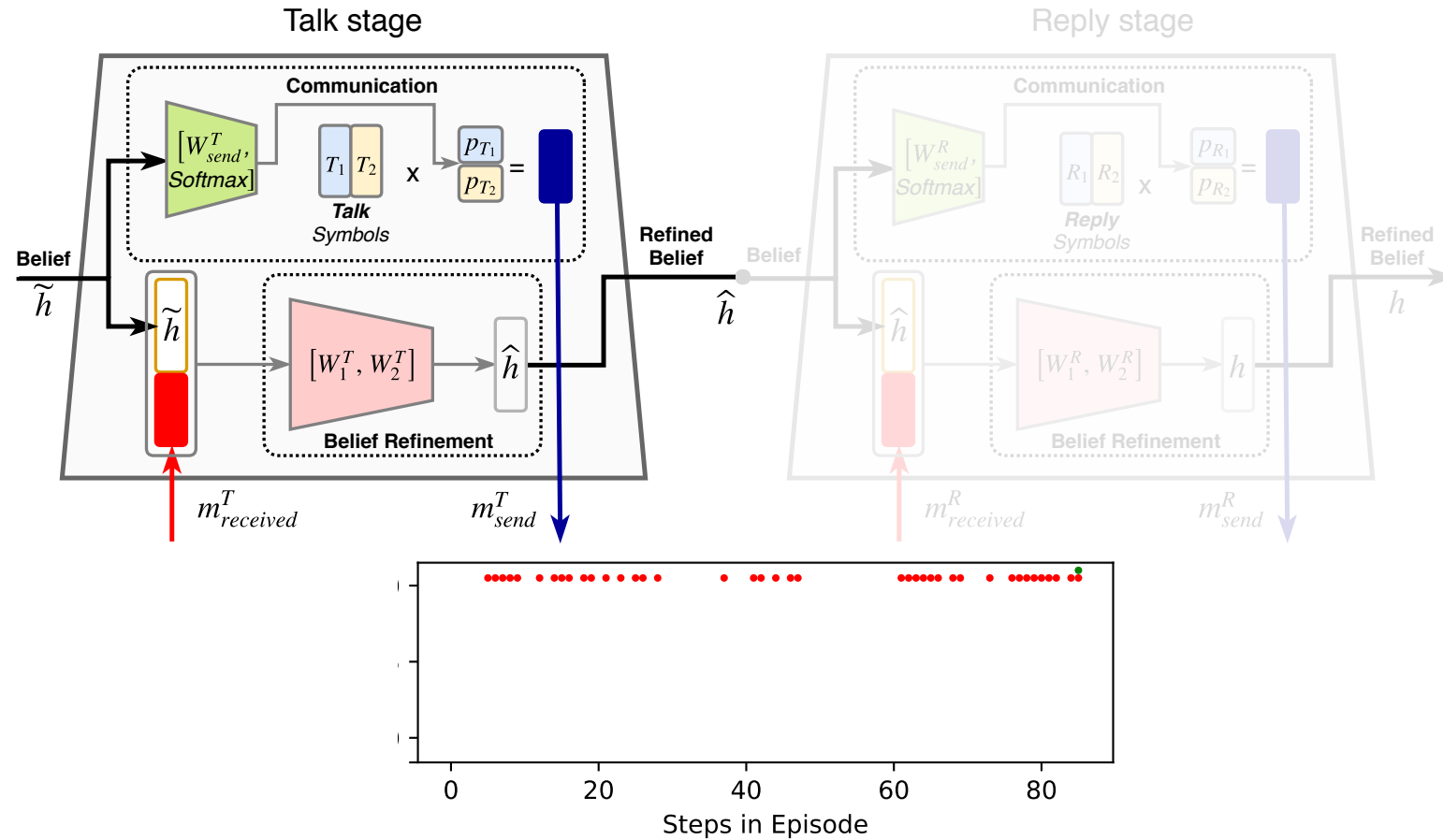
Reply (Agent 1)  
Reply (Agent 2)



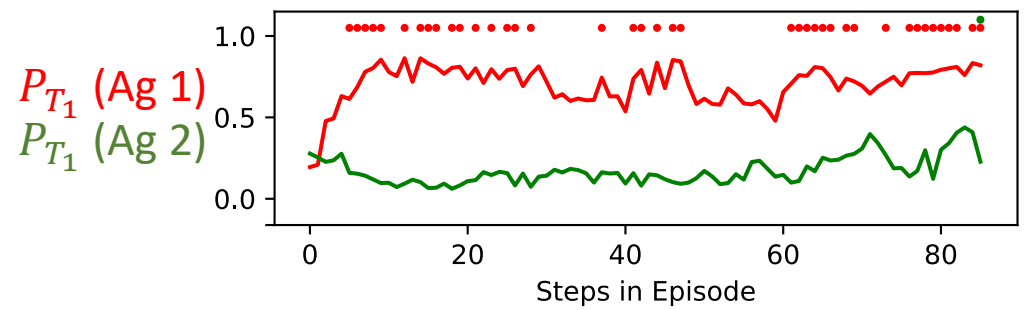
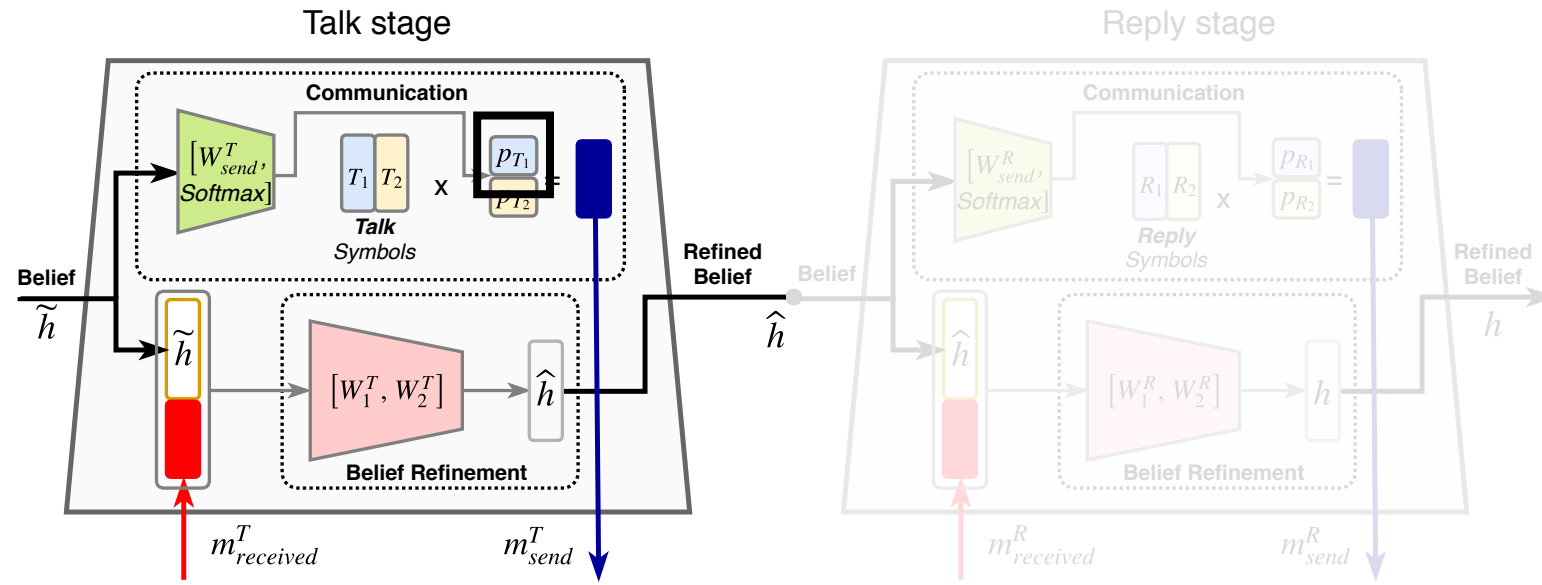
“Let us Pickup!”



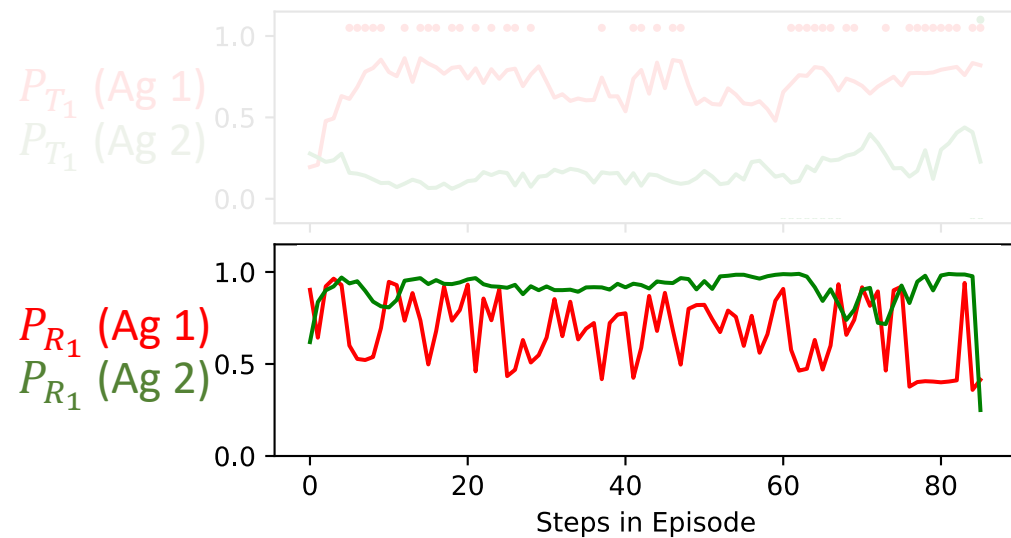
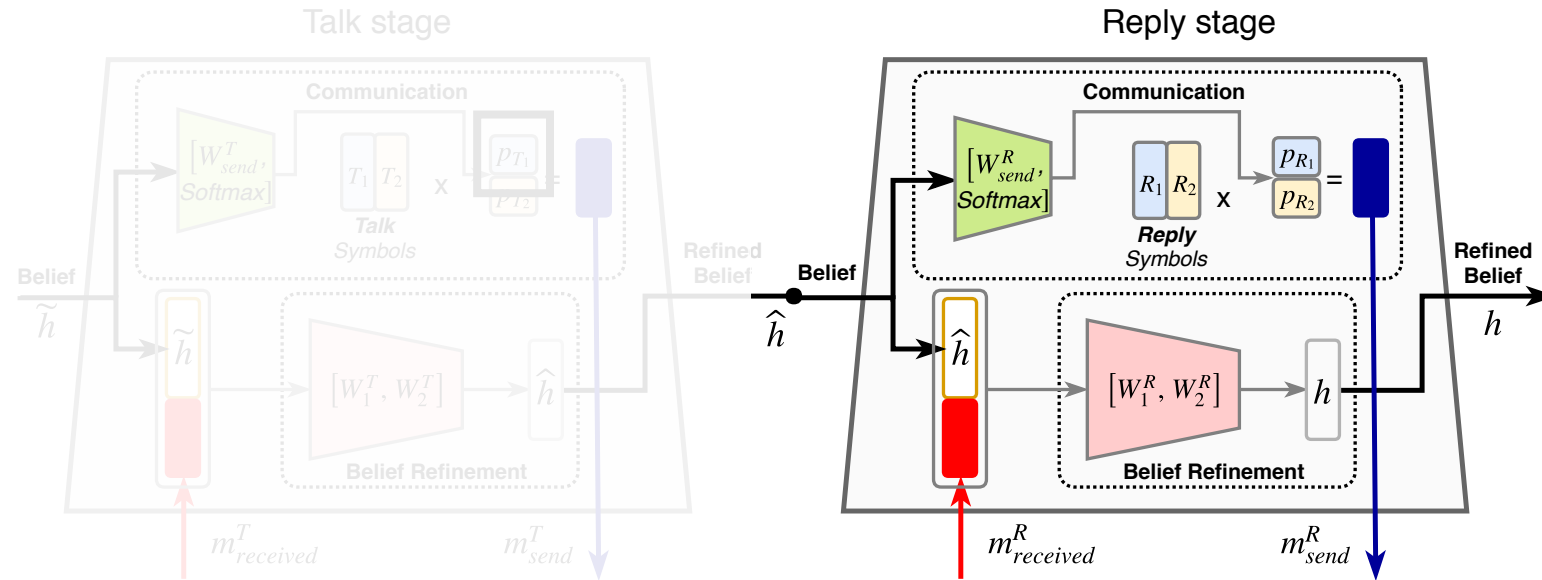
# Interpretation of messages



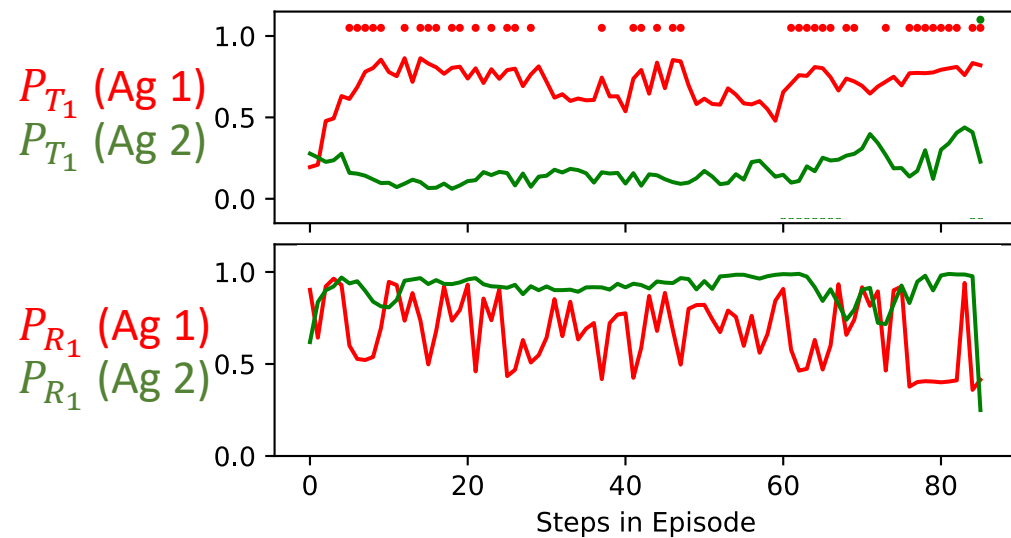
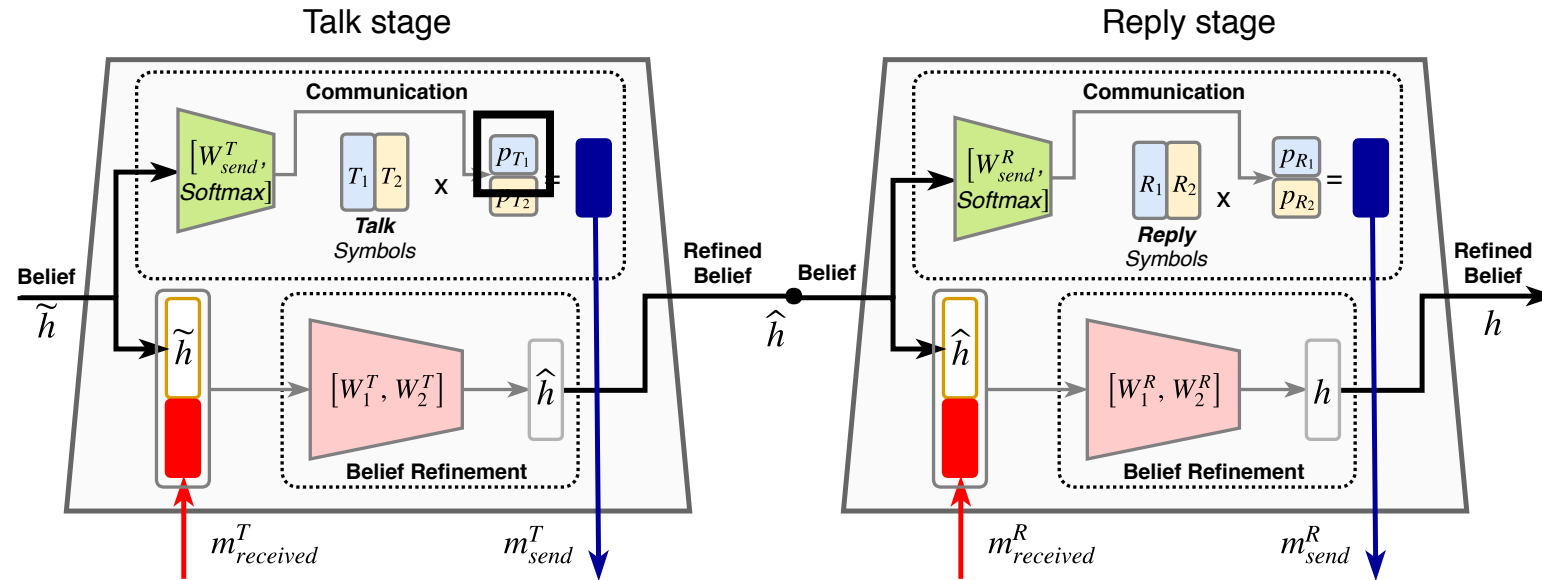
# Interpretation of messages



# Interpretation of messages

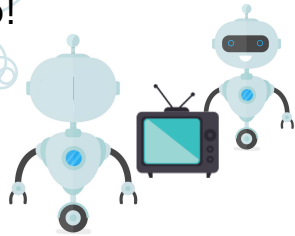


# Interpretation of messages

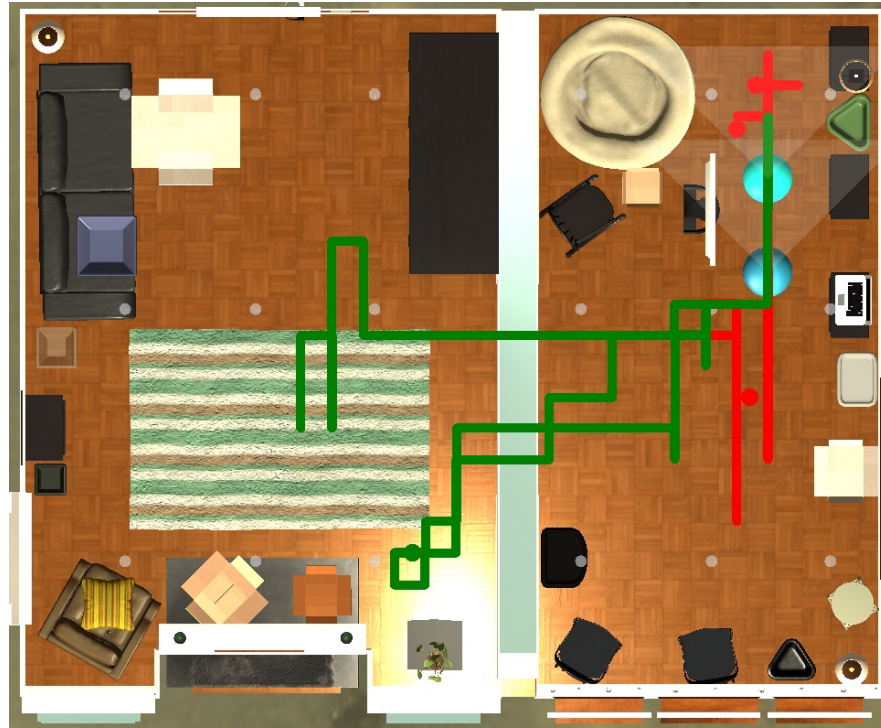


# Implicit Communication Helps

Other agent is on the opposite side of TV.  
So let me try pickup!



Visibility of other agent  
communicates information

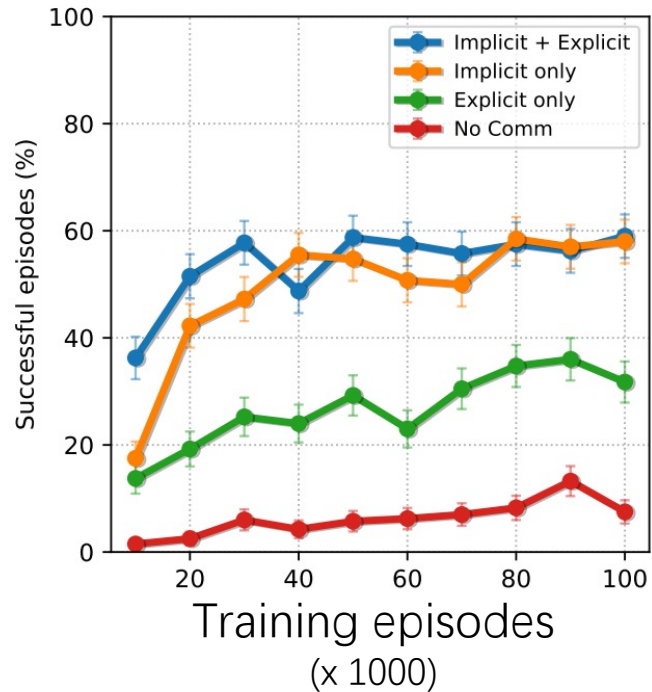


**Without** any  
communication:

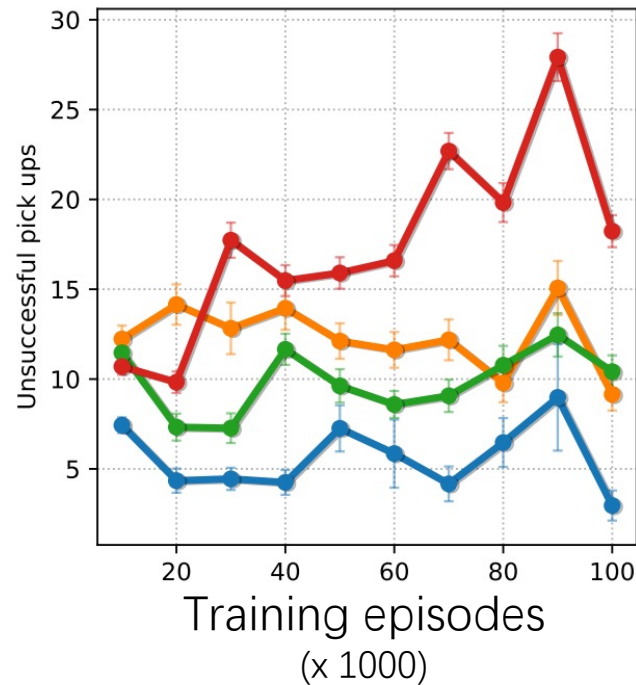
Episode  
Unsuccessful

# Effect of communication

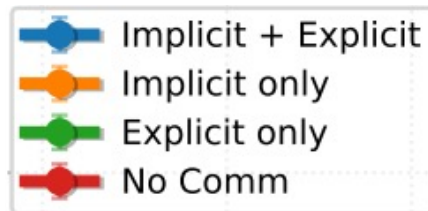
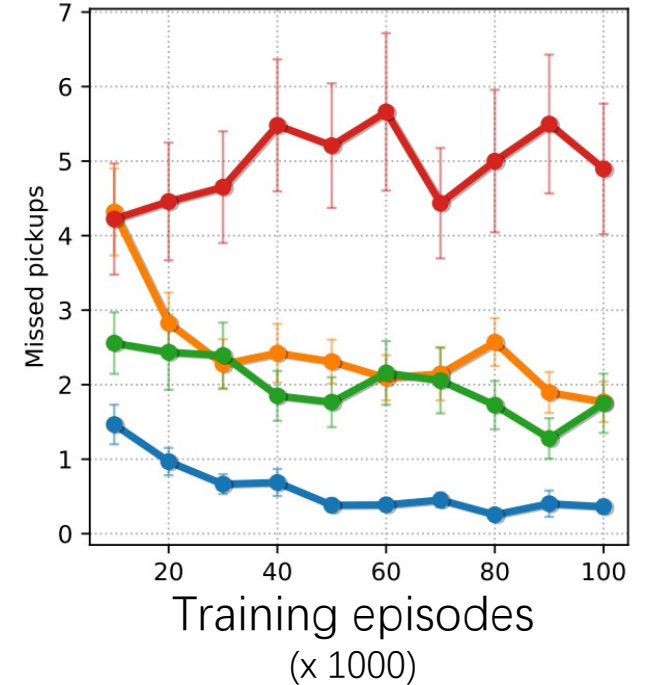
Successful episodes  $\uparrow$



Unsuccessful pickups  $\downarrow$



Missed pickups  $\downarrow$





# Collaborative Embodied Agents

## Two Body Problem

CVPR 2019 (oral)

## SYNC Policies

ECCV 2020 (spotlight)

## GRIDTOPIX

(ongoing work)

## Takeaways

- Study collaborative behavior in visual environments
- Explicit and implicit communication are helpful
- Emergence of human-like communication pattern



# Collaborative Embodied Agents

## Two Body Problem

CVPR 2019 (oral)

## **SYNC Policies**

ECCV 2020 (spotlight)

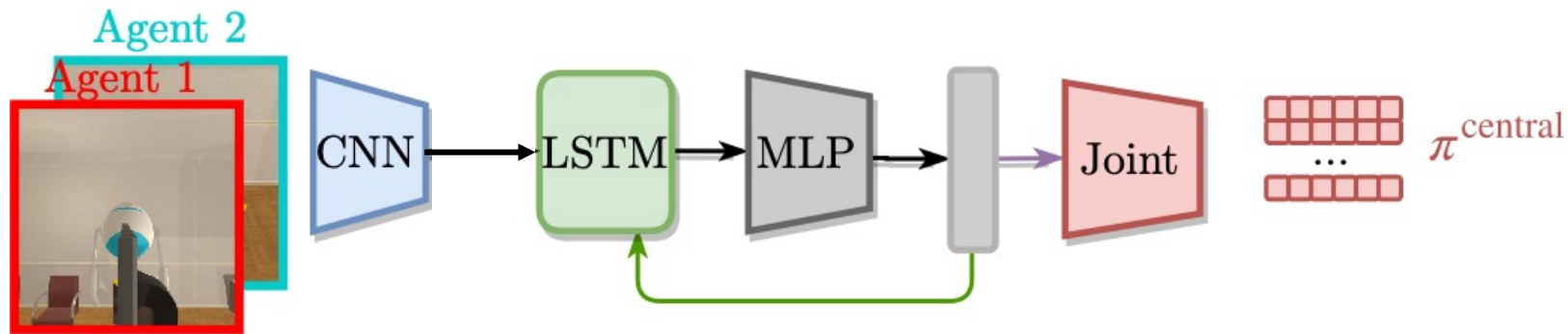
## GRIDTOPIX

(ongoing work)

1. First collaborative embodied task – FurnLift
2. Interpretation of emergent communication
3. Effect of communication
4. Intricately coordinated embodied task – FurnMove
5. Richer representation of multi-agent policy
6. Learning policies from minimal supervision
7. Leveraging perfect-perception gridworlds for training



# Constraints of Decentralized MARL



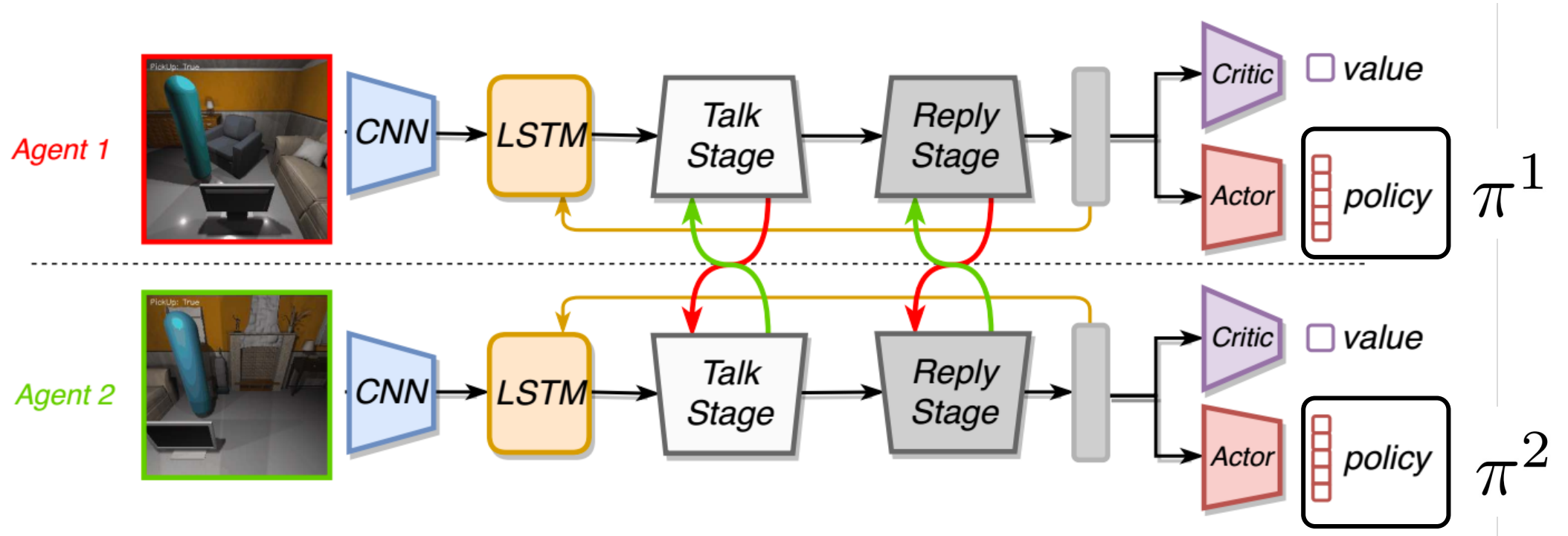
- ✗ Model complexity
- ✗ Policy parameters
- ✗ Comm. bandwidth

Central agent

Representative but disallowed in MARL (scalability!)

# Constraints of Decentralized MARL

Single (marginal) policy per agent



$$\Pi = \pi^1 \otimes \pi^2$$

$$\Pi = \pi^1 \otimes \dots \otimes \pi^N$$

# Constraints of Decentralized MARL

## Optimal Joint

Central agent can represent and sample from the joint.

$\Pi^* =$

0.2	0	0	0
0	0	0.1	0
0	0.6	0	0
0	0	0	0.1

## Marginal Agents

Agents choose their actions by independently sampling.

Agent 1 Policy ( $\pi^1$ )

0.2	0.1	0.6	0.1
-----	-----	-----	-----

Agent 2 Policy ( $\pi^2$ )

0.2	0.6	0.1	0.1
-----	-----	-----	-----

$\pi^1 \otimes \pi^2 =$

Effective Joint Policy

0.04	0.12	0.02	0.02
0.02	0.06	0.01	0.01
0.12	0.36	0.06	0.06
0.02	0.06	0.01	0.01

Rank 1

# Idea: Mixture-of-Marginals

Optimal Joint

$\Pi^* =$

0.05	0	0	0.05
0.05	0	0.15	0
0	0	0	0.15
0	0.4	0.05	0

Mixture of Marginals

Agent 1 Policies

$\pi_1^1$	0	0.3	0	0.7
$\pi_2^1$	0.9	0	0.1	0
$\pi_3^1$	0	0	0	1
$\pi_4^1$	0.33	0.33	0.33	0

Agent 2 Policies

$\pi_1^2$	0	0	0	1
$\pi_2^2$	0.4	0	0	0.6
$\pi_3^2$	0	0.5	0.5	0
$\pi_4^2$	0.8	0.2	0	0

$\alpha$

0.1
0.6
0.2
0.1

Effective Joint Policy

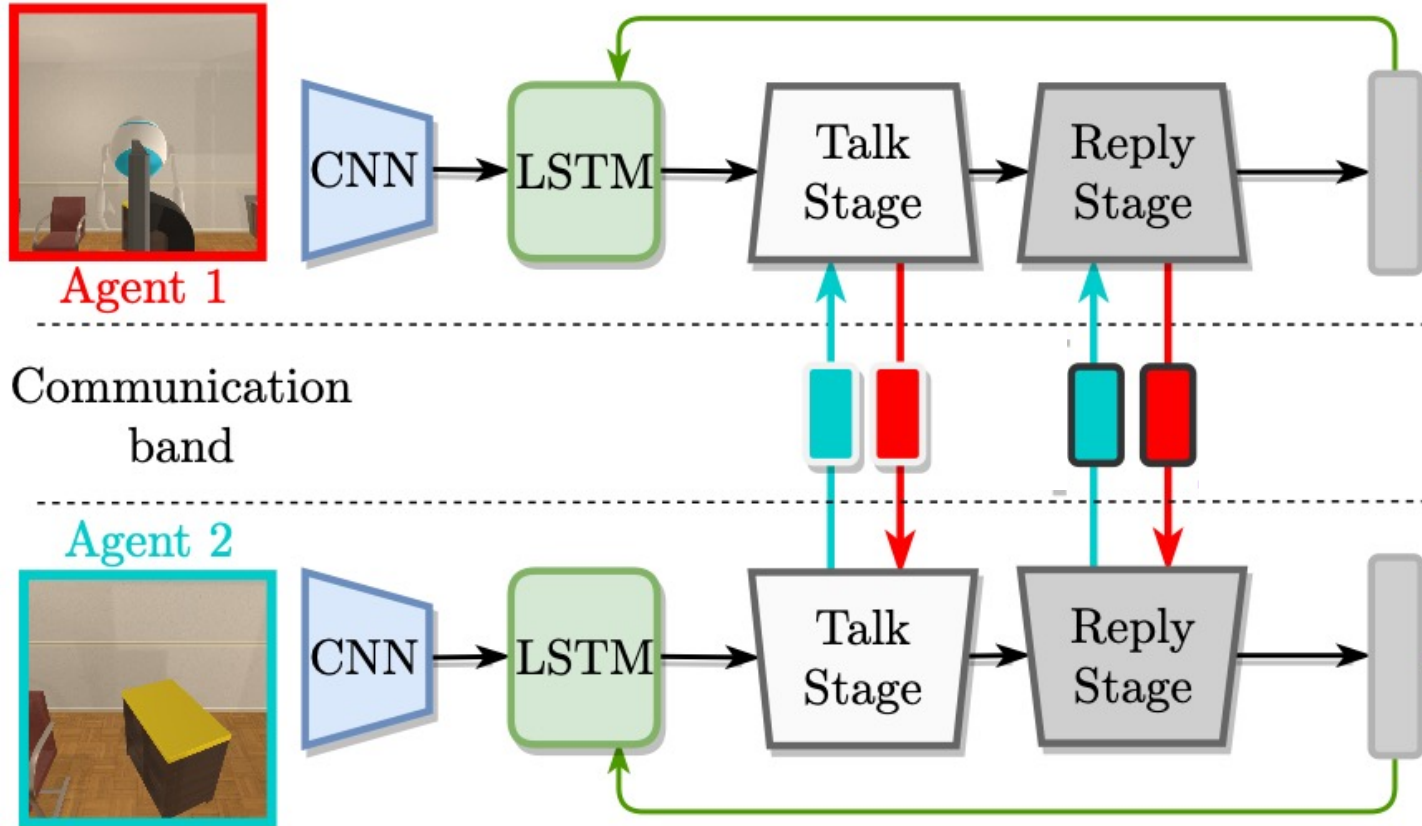
$$\sum_{i=1}^4 \alpha_i \cdot (\pi_i^1 \otimes \pi_i^2) =$$

0.05	0	0	0.05
0.05	0	0.15	0
0	0	0	0.15
0	0.4	0.05	0

# SYNC-Policies

How to sample from  $\sum_{i=1}^K \alpha_i \cdot (\pi_i^1 \otimes \pi_i^2)$  in practice?

1. Compute  $\alpha$  and  $K$  policies per agent.
2. Sample  $1 \leq i \leq K$  with probability  $\alpha_i$ . Use a shared seed so both agents sample the same  $i$ .
3. Sample actions from  $\pi_i^1$  and  $\pi_i^2$  independently.



# Intricately coordinated embodied task

Could we put communicative models to a harder test?

- Lifting furniture requires only one step of action coordination.
- Get agents to coordinate at **every** step.

# FurnMove task



# FurnMove task

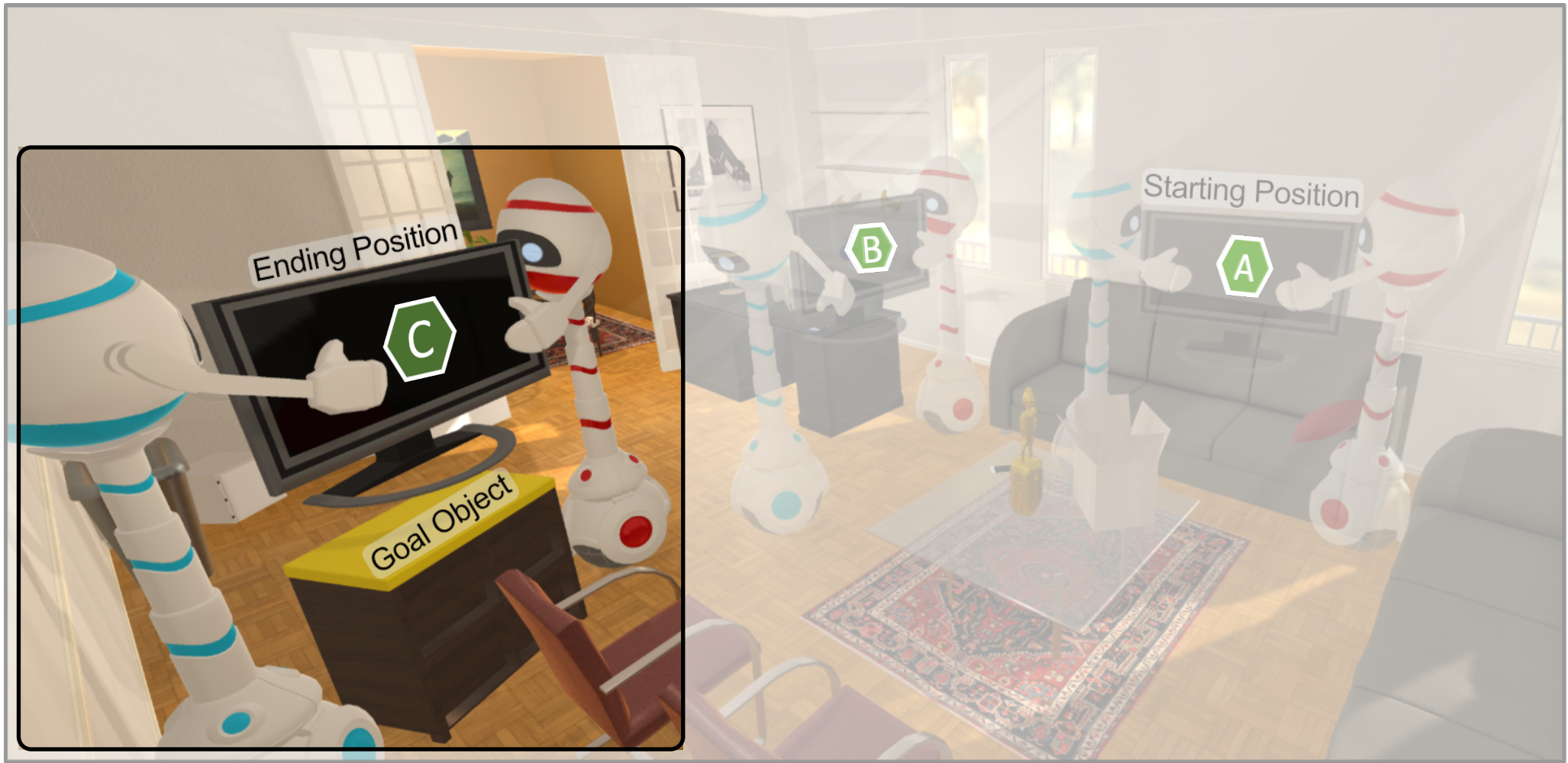




# FurnMove task

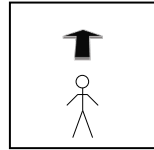


# FurnMove task

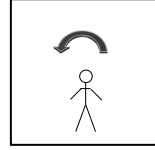


# Action Space / Agent

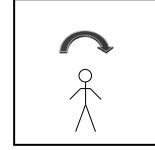
Single-Agent  
Navigation



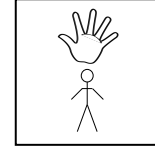
MoveAhead



RotateLeft

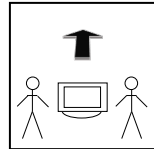


RotateRight

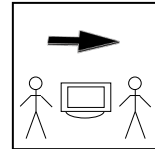


Hold

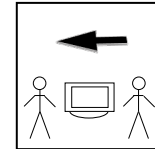
MoveWithObject  
MWO



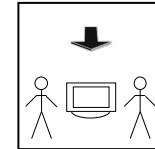
MWOAhead



MWORight

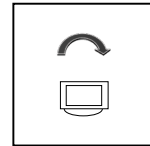


MWOLeft



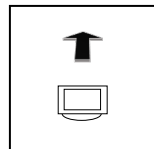
MWOBack

RotateObject  
RO

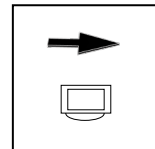


RotateObject  
Right

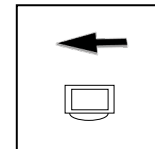
MoveObject  
MO



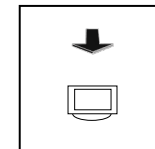
MOAhead



MORight

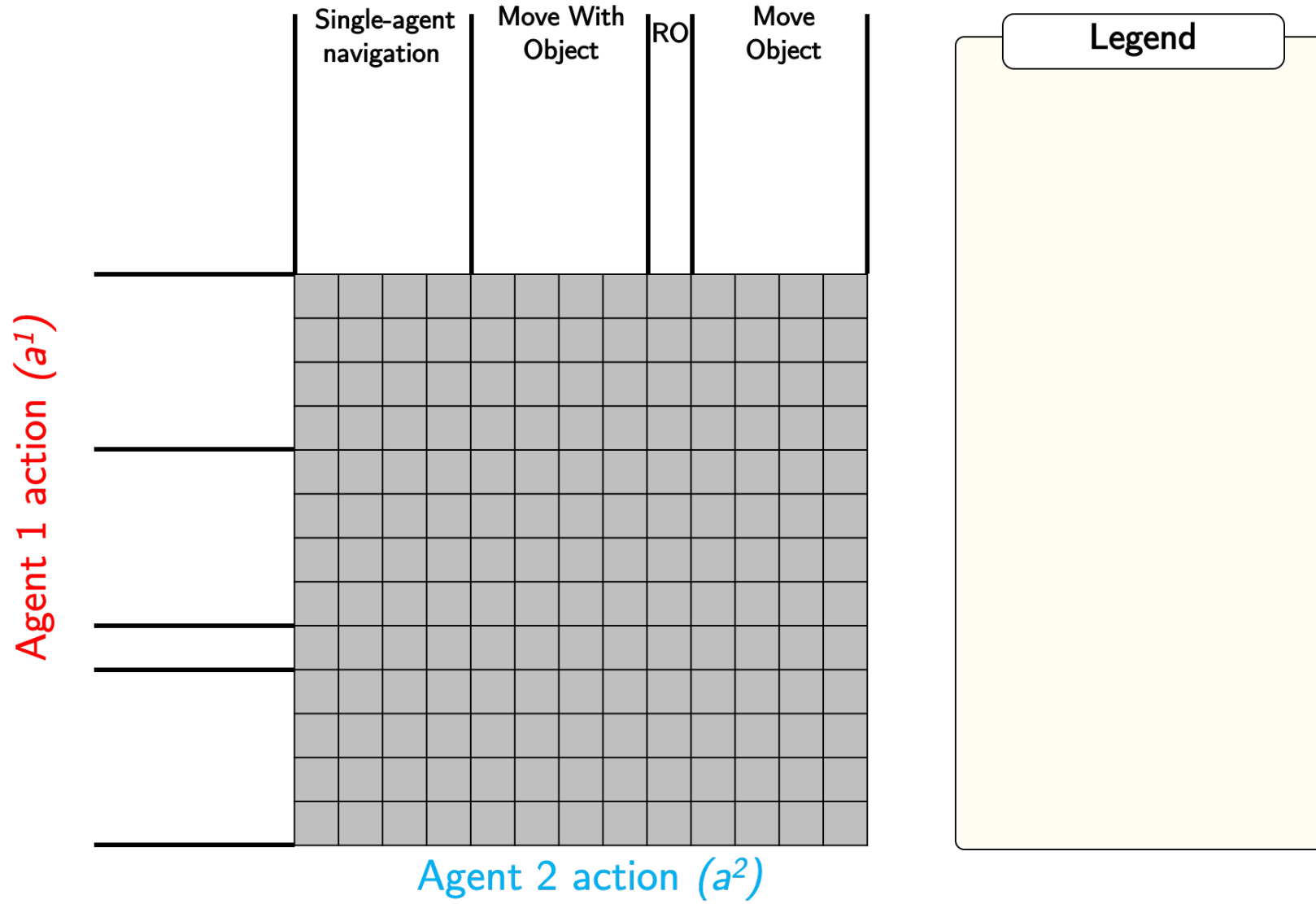


MOLeft



MOBack

# Joint action space



# Joint action space

	Single-agent navigation				Move With Object				RO	Move Object			
	MAhead	RotateLeft	RotateRight	Hold	MWOAhead	MWORight	MWOBack	MWOLeft	RORight	MOAhead	MORight	MOBack	MOLeft
Agent 1 action ( $a^1$ )	MAhead												
RotateLeft													
RotateRight													
Hold													
MWOAhead													
MWORight													
MWOBack													
MWOLeft													
RORight													
MOAhead													
MORight													
MOBack													
MOLeft													

Agent 2 action ( $a^2$ )

Legend

# Joint action space

Agent 1 action ( $a^1$ )

	Single-agent navigation				Move With Object				RO	Move Object			
	MAhead	RotateLeft	RotateRight	Hold	MWOAhead	MWORight	MWOBack	MWOLeft	RORight	MOAhead	MORight	MOBack	MOLeft
MAhead	x	x	x		x	x	x	x	x	x	x	x	x
RotateLeft	x	x	x		x	x	x	x	x	x	x	x	x
RotateRight	x	x	x		x	x	x	x	x	x	x	x	x
Hold					x	x	x	x	x	x	x	x	x
MWOAhead	x	x	x	x	x	x	x	x	x	x	x	x	x
MWORight	x	x	x	x	x	x	x	x	x	x	x	x	x
MWOBack	x	x	x	x	x	x	x	x	x	x	x	x	x
MWOLeft	x	x	x	x	x	x	x	x	x	x	x	x	x
RORight	x	x	x	x	x	x	x	x	x	x	x	x	x
MOAhead	x	x	x	x	x	x	x	x	x	x	x	x	x
MORight	x	x	x	x	x	x	x	x	x	x	x	x	x
MOBack	x	x	x	x	x	x	x	x	x	x	x	x	x
MOLeft	x	x	x	x	x	x	x	x	x	x	x	x	x

Agent 2 action ( $a^2$ )

**Legend**

X - Always Invalid

Always Valid

# Joint action space

Agent 1 action ( $a^1$ )

	Single-agent navigation				Move With Object				RO	Move Object			
	MAhead	RotateLeft	RotateRight	Hold	MWOAhead	MWORight	MWOBack	MWOLeft	RORight	MOAhead	MORight	MOBack	MOLeft
MAhead	x	x	x		x	x	x	x	x	x	x	x	x
RotateLeft	x	x	x		x	x	x	x	x	x	x	x	x
RotateRight	x	x	x		x	x	x	x	x	x	x	x	x
Hold					x	x	x	x	x	x	x	x	x
MWOAhead	x	x	x	x	x	x	x	x	x	x	x	x	x
MWORight	x	x	x	x	x	x	x	x	x	x	x	x	x
MWOBack	x	x	x	x	x	x	x	x	x	x	x	x	x
MWOLeft	x	x	x	x	x	x	x	x	x	x	x	x	x
RORight	x	x	x	x	x	x	x	x		x	x	x	x
MOAhead	x	x	x	x	x	x	x	x	x	x	x	x	x
MORight	x	x	x	x	x	x	x	x	x	x	x	x	x
MOBack	x	x	x	x	x	x	x	x	x	x	x	x	x
MOLeft	x	x	x	x	x	x	x	x	x	x	x	x	x

Agent 2 action ( $a^2$ )

**Legend**

X - Always Invalid

Always Valid

# Joint action space

Agent 1 action ( $a^1$ )

	Single-agent navigation				Move With Object				RO	Move Object			
	MAhead	RotateLeft	RotateRight	Hold	MWOAhead	MWORight	MWOBack	MWOLeft	RORight	MOAhead	MORight	MOBack	MOLeft
MAhead	x	x	x		x	x	x	x	x	x	x	x	x
RotateLeft	x	x	x		x	x	x	x	x	x	x	x	x
RotateRight	x	x	x		x	x	x	x	x	x	x	x	x
Hold					x	x	x	x	x	x	x	x	x
MWOAhead	x	x	x	x	x	x	x	x	x	x	x	x	x
MWORight	x	x	x	x	x	x	x	x	x	x	x	x	x
MWOBack	x	x	x	x	x	x	x	x	x	x	x	x	x
MWOLeft	x	x	x	x	x	x	x	x	x	x	x	x	x
RORight	x	x	x	x	x	x	x	x		x	x	x	x
MOAhead	x	x	x	x	x	x	x	x	x	x	x	x	x
MORight	x	x	x	x	x	x	x	x	x	x	x	x	x
MOBack	x	x	x	x	x	x	x	x	x	x	x	x	x
MOLeft	x	x	x	x	x	x	x	x	x	x	x	x	x

Agent 2 action ( $a^2$ )

**Legend**

X - Always Invalid

Always Valid




# Joint action space

Agent 1 action ( $a^1$ )


	Single-agent navigation				Move With Object				RO	Move Object			
	MAhead	RotateLeft	RotateRight	Hold	MWOAhead	MWORight	MWOBack	MWOLeft	RORight	MOAhead	MORight	MOBack	MOLeft
MAhead	x	x	x		x	x	x	x	x	x	x	x	x
RotateLeft	x	x	x		x	x	x	x	x	x	x	x	x
RotateRight	x	x	x		x	x	x	x	x	x	x	x	x
Hold					x	x	x	x	x	x	x	x	x
MWOAhead	x	x	x	x	x	x	x	x	x	x	x	x	x
MWORight	x	x	x	x	x	x	x	x	x	x	x	x	x
MWOBack	x	x	x	x	x	x	x	x	x	x	x	x	x
MWOLeft	x	x	x	x	x	x	x	x	x	x	x	x	x
RORight	x	x	x	x	x	x	x	x		x	x	x	x
MOAhead	x	x	x	x	x	x	x	x	x	x	x	x	x
MORight	x	x	x	x	x	x	x	x	x	x	x	x	x
MOBack	x	x	x	x	x	x	x	x	x	x	x	x	x
MOLeft	x	x	x	x	x	x	x	x	x	x	x	x	x

Agent 2 action ( $a^2$ )

**Legend**



$A^1$   
angle



$A^2$   
angle

0°

Relative  
orientation

X - Always Invalid

Always Valid

# Joint action space

High Rank



10% actions are valid



Agent 1 action ( $a^1$ )

	Single-agent navigation				Move With Object				RO	Move Object			
	MAhead	RotateLeft	RotateRight	Hold	MWOAhead	MWORight	MWOBack	MWOLeft	RORight	MOAhead	MORight	MOBack	MOLeft
MAhead	x	x	x		x	x	x	x	x	x	x	x	x
RotateLeft	x	x	x		x	x	x	x	x	x	x	x	x
RotateRight	x	x	x		x	x	x	x	x	x	x	x	x
Hold					x	x	x	x	x	x	x	x	x
MWOAhead	x	x	x	x	x	x	x	x	x	x	x	x	x
MWORight	x	x	x	x	x	x	x	x	x	x	x	x	x
MWOBack	x	x	x	x	x	x	x	x	x	x	x	x	x
MWOLeft	x	x	x	x	x	x	x	x	x	x	x	x	x
RORight	x	x	x	x	x	x	x	x		x	x	x	x
MOAhead	x	x	x	x	x	x	x	x	x	x	x	x	x
MORight	x	x	x	x	x	x	x	x	x	x	x	x	x
MOBack	x	x	x	x	x	x	x	x	x	x	x	x	x
MOLeft	x	x	x	x	x	x	x	x	x	x	x	x	x

Agent 2 action ( $a^2$ )

### Legend

0°

90°

$A^1$  angle     $A^2$  angle    Relative orientation

X - Always Invalid

Always Valid

# Joint action space

Agent 1 action ( $a^1$ )

	Single-agent navigation				Move With Object				RO	Move Object			
	MAhead	RotateLeft	RotateRight	Hold	MWOAhead	MWORight	MWOBack	MWOLeft	RORight	MOAhead	MORight	MOBack	MOLeft
MAhead	x	x	x		x	x	x	x	x	x	x	x	x
RotateLeft	x	x	x		x	x	x	x	x	x	x	x	x
RotateRight	x	x	x		x	x	x	x	x	x	x	x	x
Hold					x	x	x	x	x	x	x	x	x
MWOAhead	x	x	x	x	x	x	x		x	x	x	x	x
MWORight	x	x	x	x		x	x	x	x	x	x	x	x
MWOBack	x	x	x	x	x		x	x	x	x	x	x	x
MWOLeft	x	x	x	x	x	x		x	x	x	x	x	x
RORight	x	x	x	x	x	x	x	x		x	x	x	x
MOAhead	x	x	x	x	x	x	x	x	x	x	x		x
MORight	x	x	x	x	x	x	x	x	x		x	x	x
MOBack	x	x	x	x	x	x	x	x	x	x		x	x
MOLeft	x	x	x	x	x	x	x	x	x	x	x		x

Agent 2 action ( $a^2$ )

Legend

$A^1$   
angle

$A^2$   
angle

Relative  
orientation

0°

90°

X - Always Invalid

Always Valid

# Joint action space

Agent 1 action ( $a^1$ )

Agent 2 action ( $a^2$ )

	Single-agent navigation				Move With Object				RO	Move Object			
	MAhead	RotateLeft	RotateRight	Hold	MWOAhead	MWORight	MWOWBack	MWOLeft	RORight	MOAhead	MORight	MOBack	MOLeft
MAhead	x	x	x		x	x	x	x	x	x	x	x	x
RotateLeft	x	x	x		x	x	x	x	x	x	x	x	x
RotateRight	x	x	x		x	x	x	x	x	x	x	x	x
Hold					x	x	x	x	x	x	x	x	x
MWOAhead	x	x	x	x					x	x	x	x	x
MWORight	x	x	x	x					x	x	x	x	x
MWOWBack	x	x	x	x					x	x	x	x	x
MWOLeft	x	x	x	x					x	x	x	x	x
RORight	x	x	x	x	x	x	x	x		x	x	x	x
MOAhead	x	x	x	x	x	x	x	x	x				
MORight	x	x	x	x	x	x	x	x	x				
MOBack	x	x	x	x	x	x	x	x	x				
MOLeft	x	x	x	x	x	x	x	x	x				

### Legend

$A^1$   
angle

$A^2$   
angle

Relative  
orientation

0°

90°

180°

270°

Relative orientation  
defines validity

X - Always Invalid

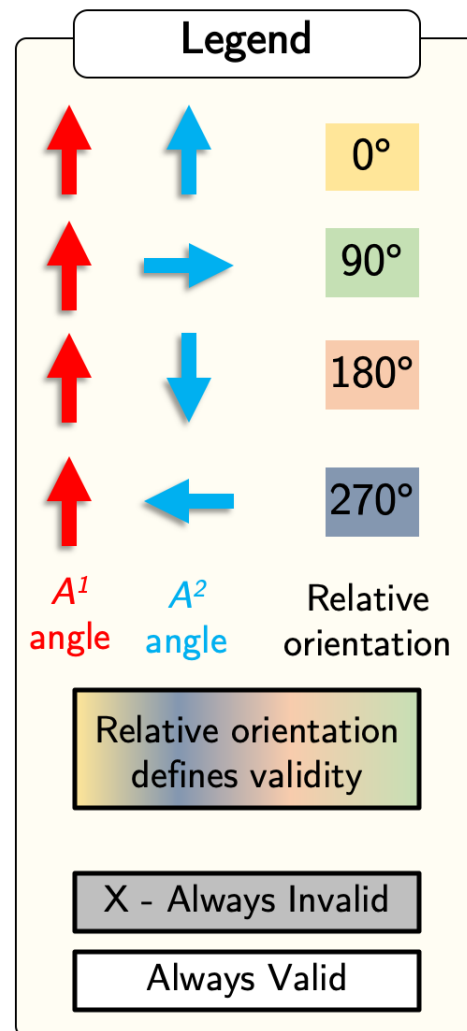
Always Valid

# Joint action space

space		Single-agent navigation				Move With Object				RO	Move Object			
		MAhead	RotateLeft	RotateRight	Hold	MWOAhead	MWORight	MWOWBack	MWOWLeft	RORight	MOAhead	MORight	MOBack	MOLeft
Agent 1 action ( $a^1$ )	MAhead	x	x	x		x	x	x	x	x	x	x	x	x
	RotateLeft	x	x	x		x	x	x	x	x	x	x	x	x
	RotateRight	x	x	x		x	x	x	x	x	x	x	x	x
	Hold					x	x	x	x	x	x	x	x	x
	MWOAhead	x	x	x	x					x	x	x	x	x
	MWORight	x	x	x	x					x	x	x	x	x
	MWOWBack	x	x	x	x					x	x	x	x	x
	MWOWLeft	x	x	x	x					x	x	x	x	x
	RORight	x	x	x	x	x	x	x	x		x	x	x	x
	MOAhead	x	x	x	x	x	x	x	x	x				
	MORight	x	x	x	x	x	x	x	x	x				
	MOBack	x	x	x	x	x	x	x	x	x				
	MOLeft	x	x	x	x	x	x	x	x	x				

Agent 2 action ( $a^2$ )

10% actions are valid



		Single-agent navigation				
		MAhead	RotateLeft	RotateRight	Pass	PickUp
Agent 1 action ( $a^1$ )	MAhead					x
	RotateLeft					x
	RotateRight					x
	Pass					x
	PickUp	x	x	x	x	

Agent 2 action ( $a^2$ )

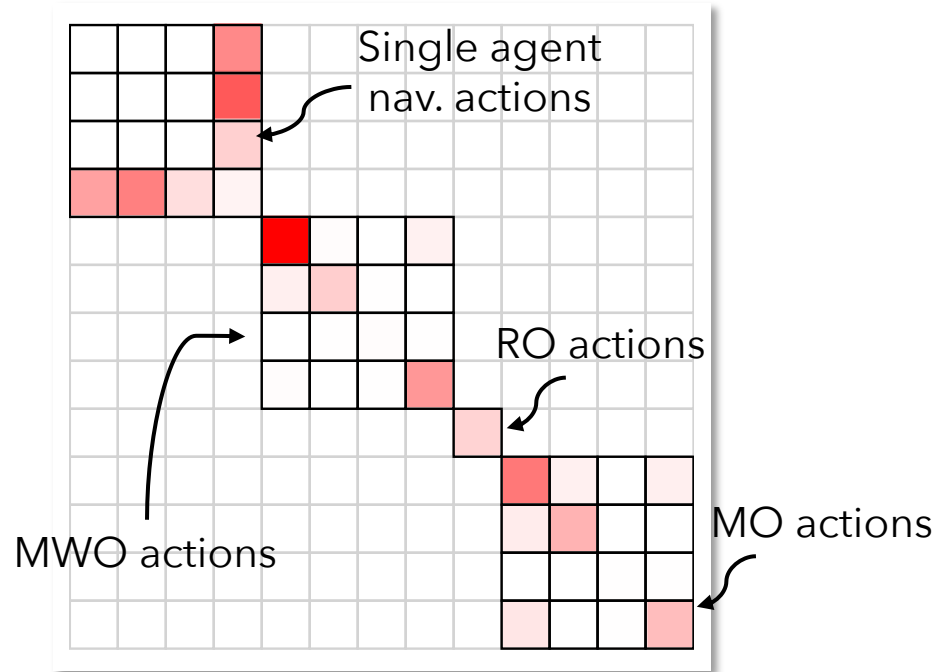
64% actions are valid

# How coordinated is FurnMove?

Central Agent  $\longleftrightarrow$  Marginal Agents

	Success	Failed Pickups
FurnLift	0.6%	5.1 vs. 8.9
FurnMove	32.0%	

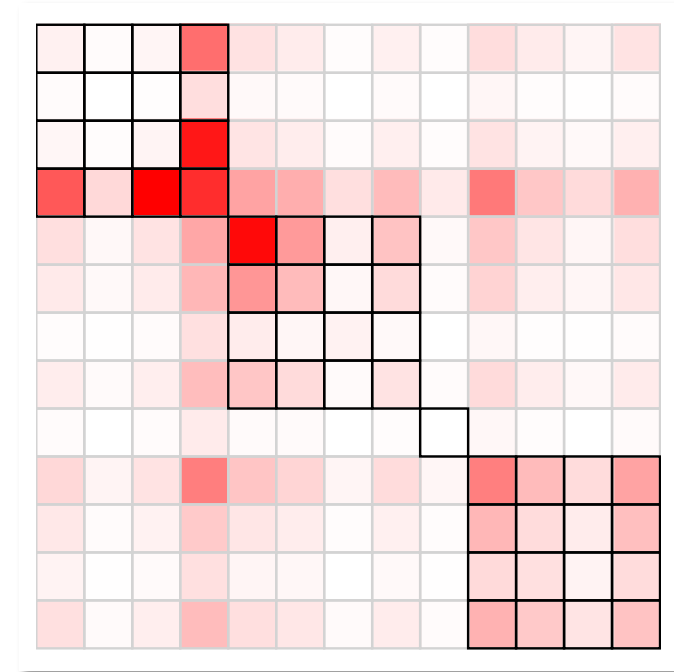
# Joint Policy Summary



Central Model

65% task success  
7% actions fail

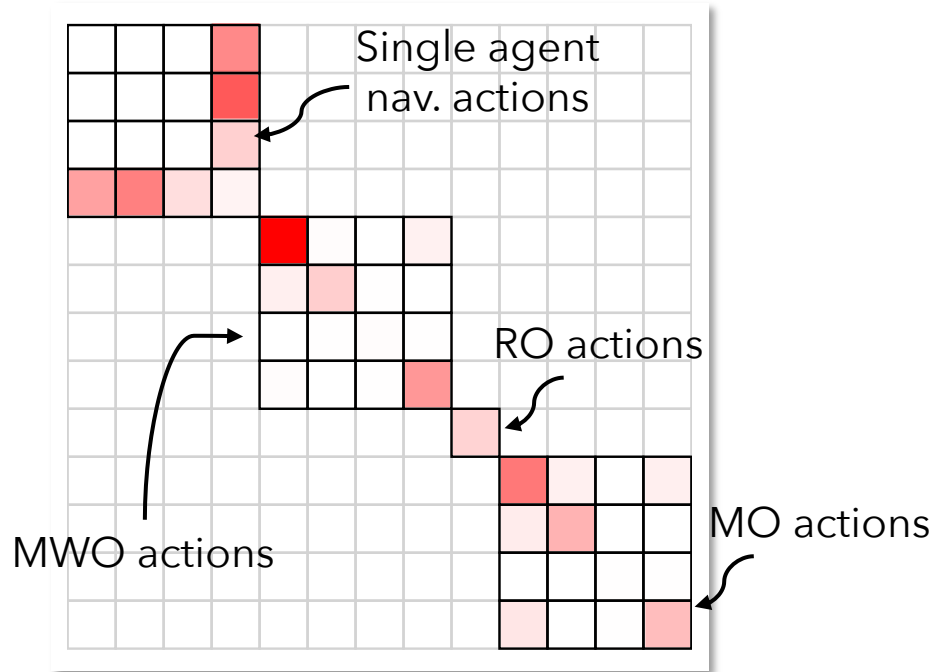
VS



Marginal Model

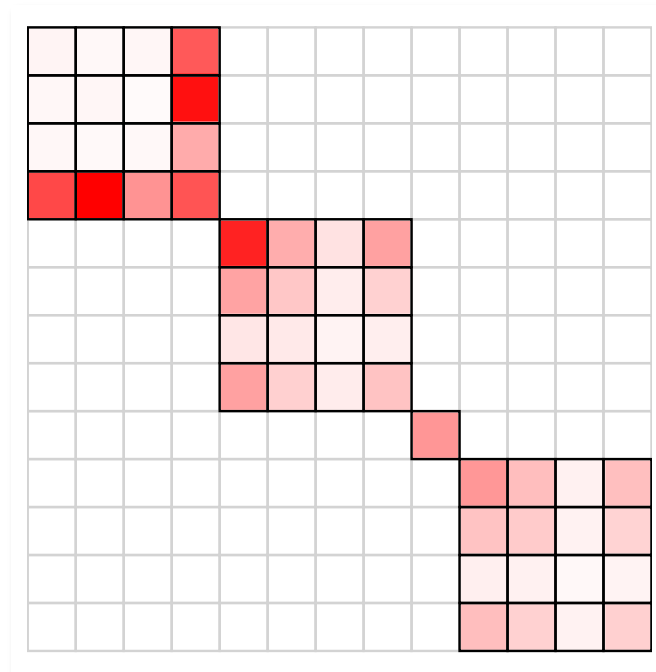
33% task success  
65% actions fail

# Joint Policy Summary



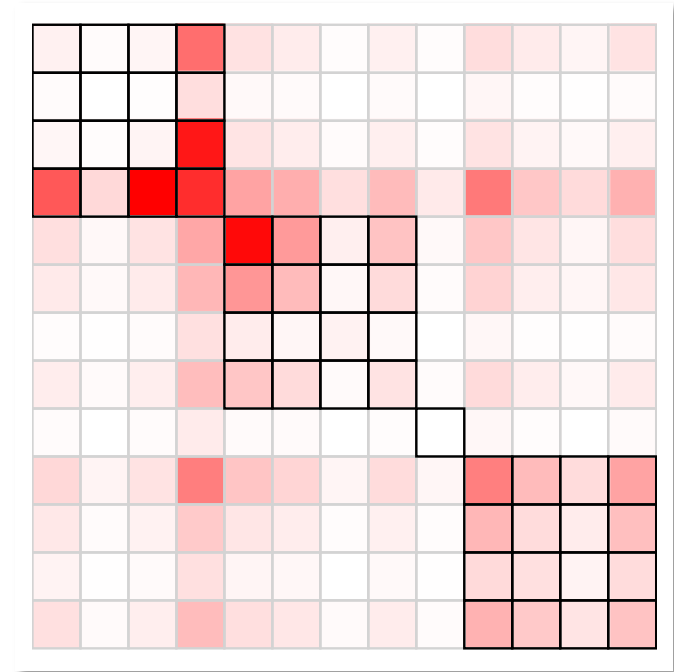
Central Model

65% task success  
7% actions fail



SYNC Model

59% success rate  
31% actions fail



Marginal Model

33% task success  
65% actions fail



# Qualitative runs

Field of view:

Triangles denote field of view & orientation of agents

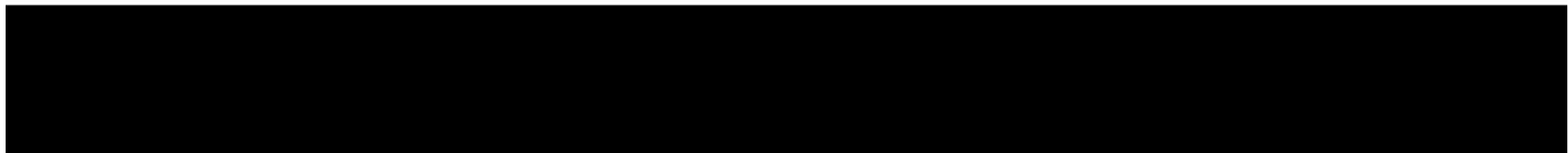
Trajectories:

- Agent 1 trajectory in red
- Agent 2 trajectory in green
- TV trajectory in blue
- Trajectory shades become *lighter* as episode progresses

Top-down view



# Marginal Agents



# SYNC Agents



# How many mixtures components in SYNC?

Diminishing returns from additional mixture components

# Mixture Components	Success ↑	Final Distance ↓
1 component	33	1.83
2 components	50	1.23
4 components	57	<b>1.08</b>
13 components	<b>59</b>	1.15



# Collaborative Embodied Agents

## Two Body Problem

CVPR 2019 (oral)

## **SYNC Policies**

ECCV 2020 (spotlight)

## GRIDTOPIX

(ongoing work)

## Takeaways

- Independent and decentral execution  $\Rightarrow$  Rank-1
- Mixture-of-marginals adapted as SYNC-policies
- Useful for solving high-rank tasks - FurnMove



# Collaborative Embodied Agents

## Two Body Problem

CVPR 2019 (oral)

## SYNC Policies

ECCV 2020 (spotlight)

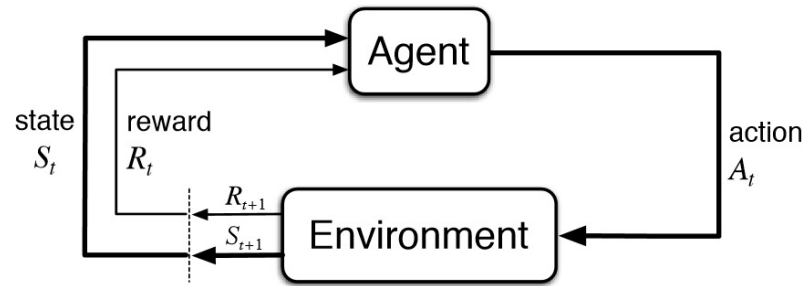
## GRIDTOPIX

(ongoing work)

1. First collaborative embodied task – FurnLift
2. Interpretation of emergent communication
3. Effect of communication
4. Intricately coordinated embodied task – FurnMove
5. Richer representation of multi-agent policy
6. Learning policies from minimal supervision
7. Leveraging perfect-perception gridworlds for training



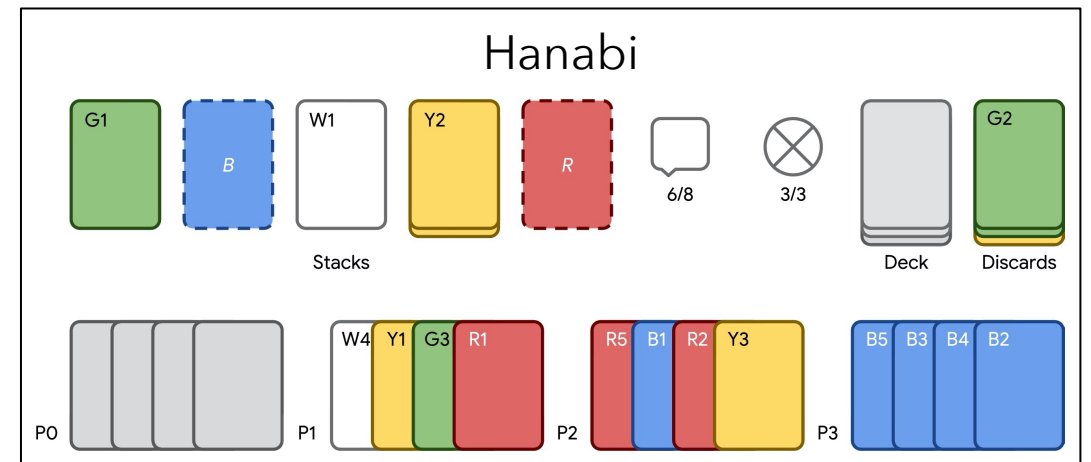
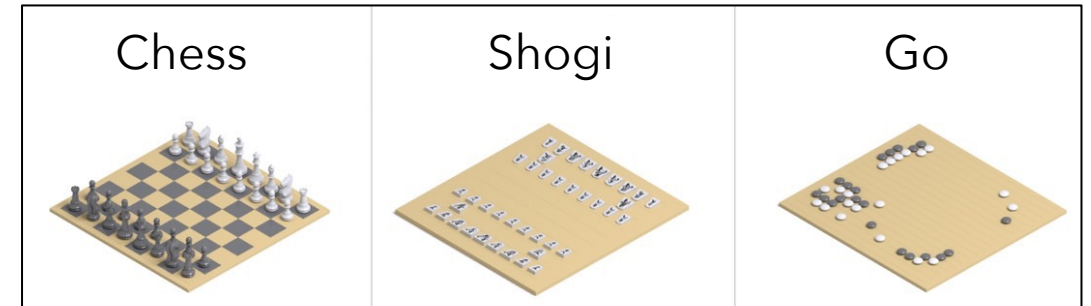
# (1) Terminal Rewards = Minimal Supervision



## 'Terminal rewards'

Goal dependent or success rewards available at termination of episode

Works for (non-visual) RL benchmarks

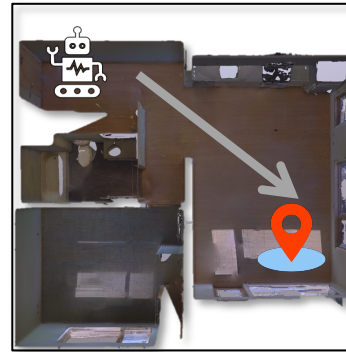


# (2) Visual Agents Need Shaped Rewards

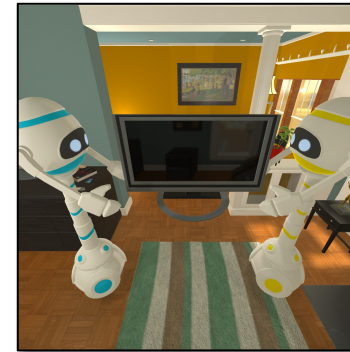
## 'Shaped rewards'

- Dense indicators of success
- Furniture Moving:
  - Furn. moved closer to the goal
- PointGoal Navigation
  - Geodesic distance to goal
- Google Football
  - Checkpoint reward

PointGoal Navigation  
(Habitat+Gibson)



Furniture Moving  
(AI2-THOR)



3 vs. 1 with Keeper  
(Google Football)



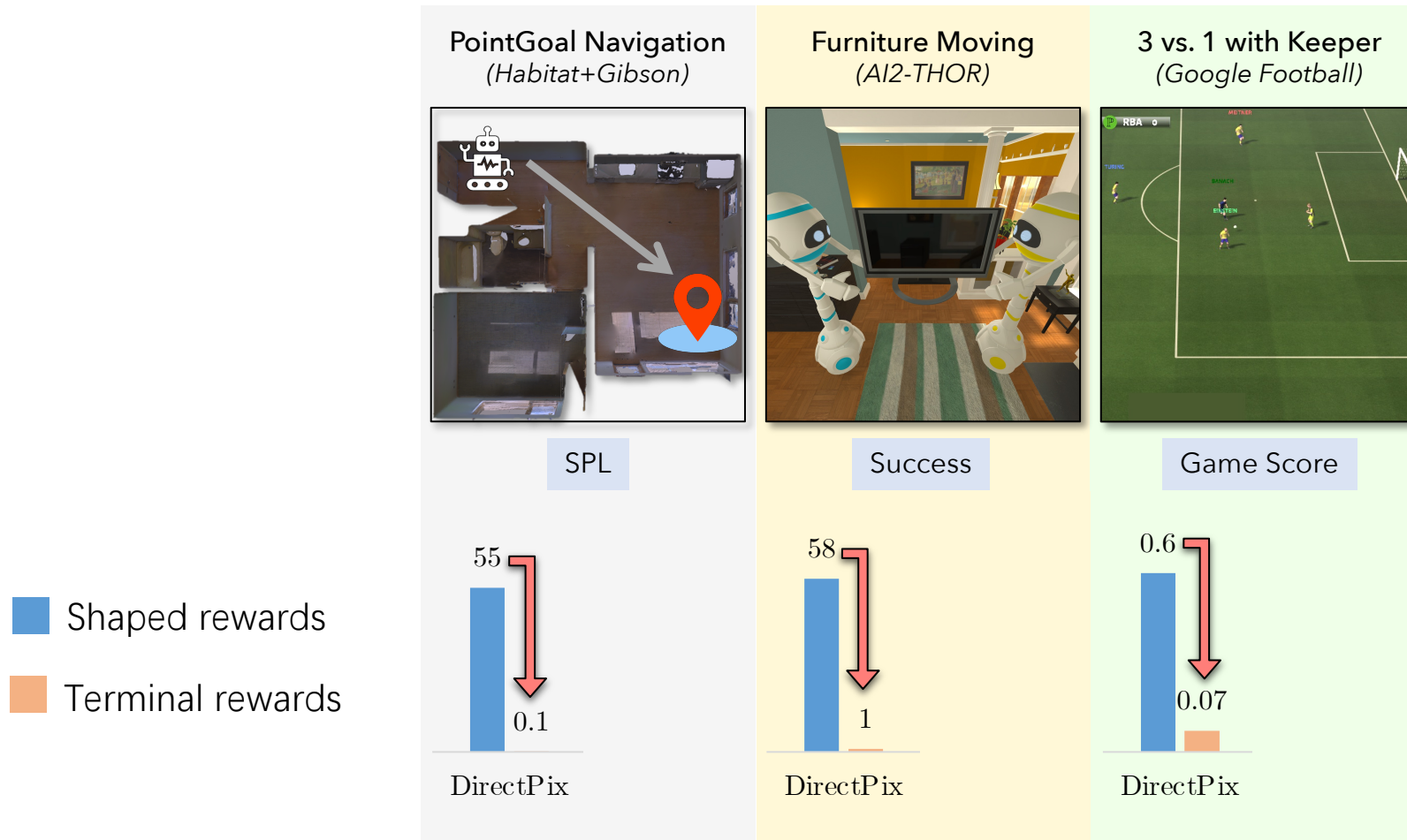


## (2) Visual Agents Need Shaped Rewards

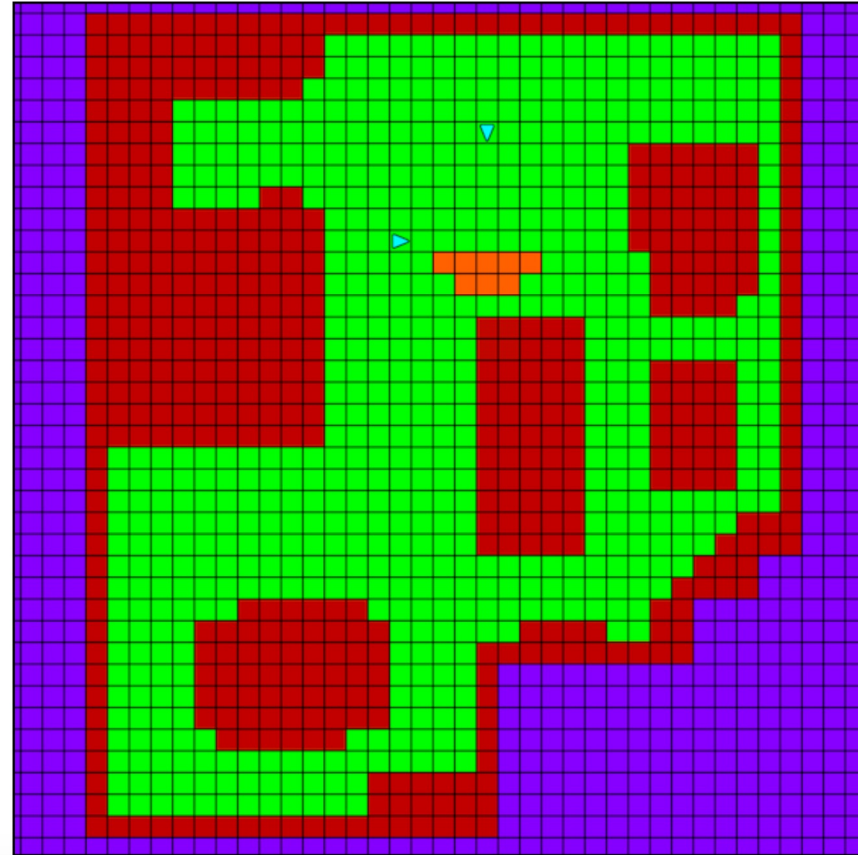
- Shaped rewards
- Terminal rewards



# (2) Visual Agents Need Shaped Rewards



# (3) Terminal Rewards Work in Gridworlds



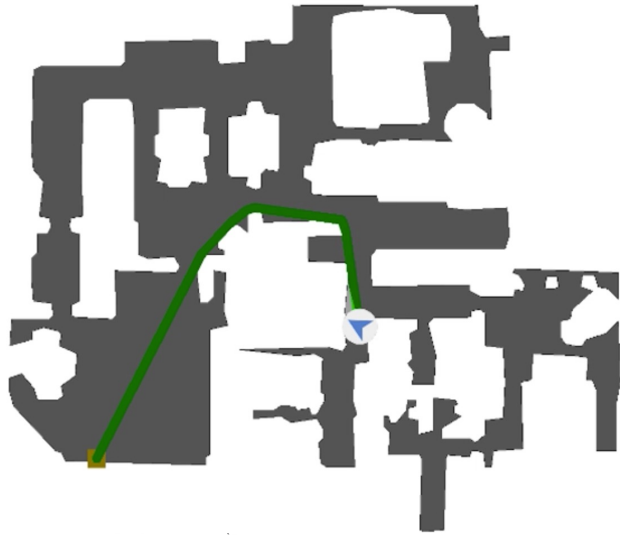
### (3) Terminal Rewards Work in Gridworlds

Observation space	Accuracy $\uparrow$	SPL $\uparrow$
Visual (terminal)	1	0.0
Gridworld (terminal)	<b>56</b>	<b>0.19</b>

Observation space	Accuracy $\uparrow$	SPL $\uparrow$
Visual (shaped)	58	0.11
Gridworld (shaped)	<b>76</b>	<b>0.22</b>



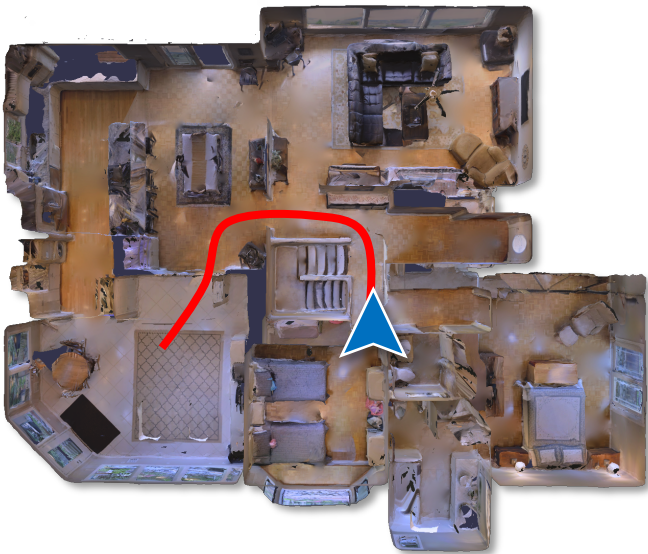
# “How Can We Leverage Gridworlds?”



(1) Terminal Rewards = Minimal Supervision

(2) Visual Agents Need Shaped Rewards

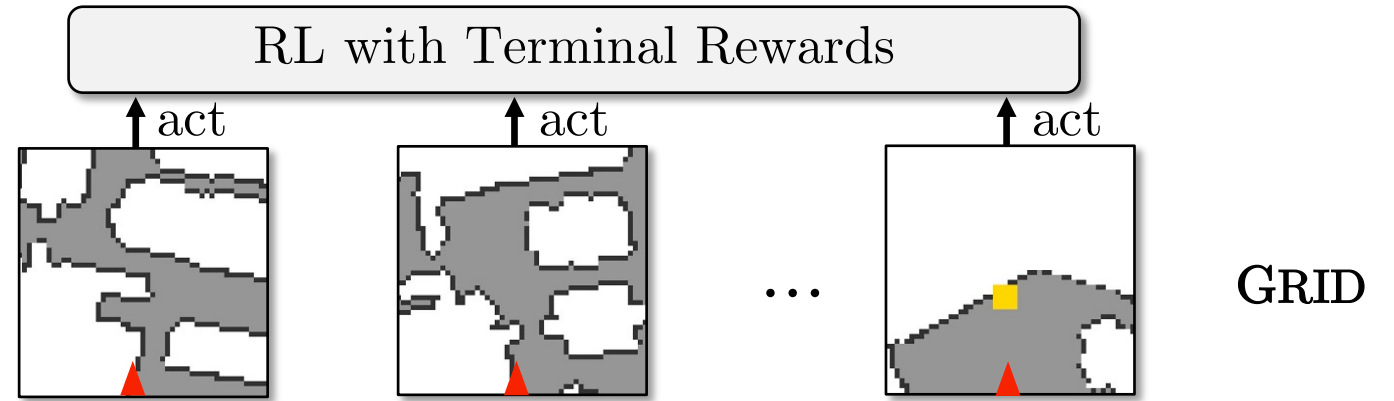
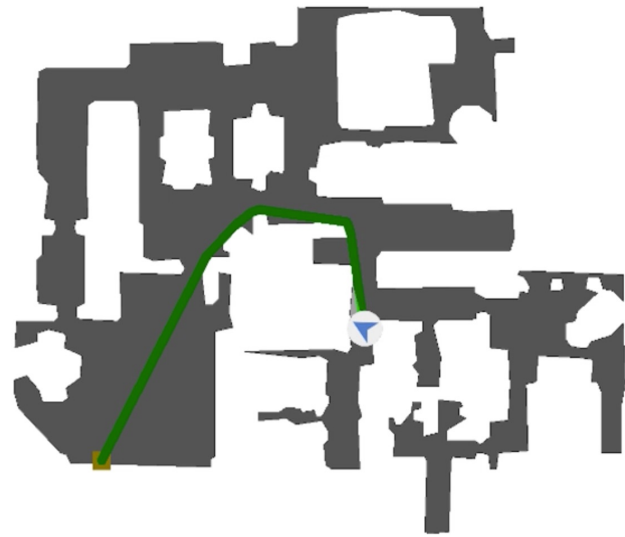
(3) Terminal Rewards Work in Gridworlds



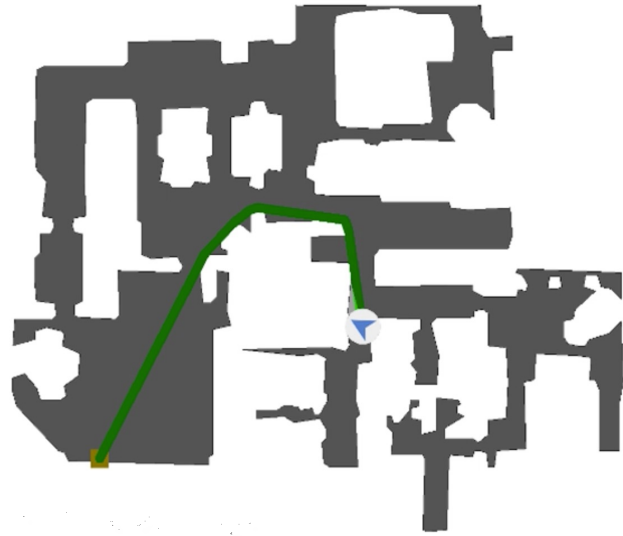
# GRIDTOPIX



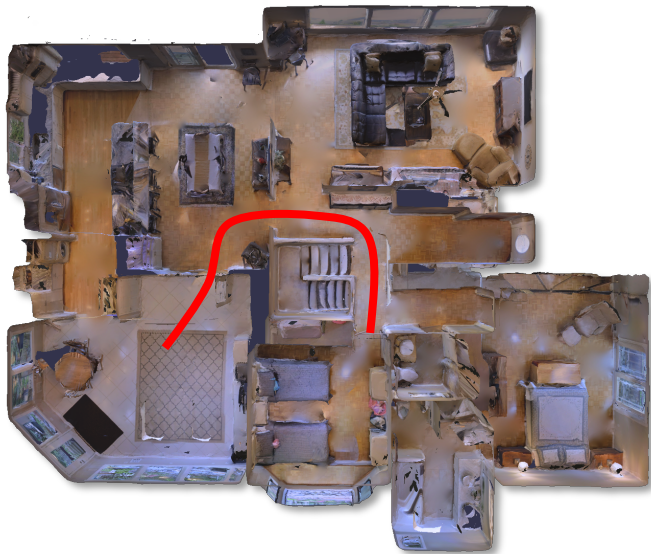
# GRIDTOPIX



# GRIDTOPIX



**GRID**  
(teacher)

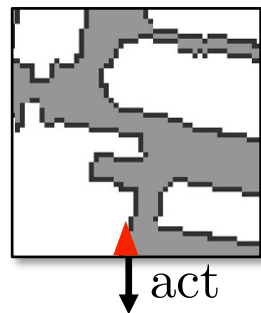
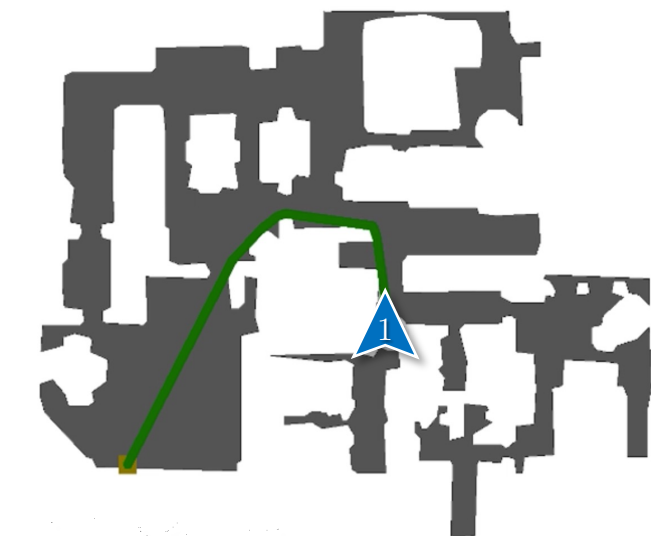


**PIX**  
(student)

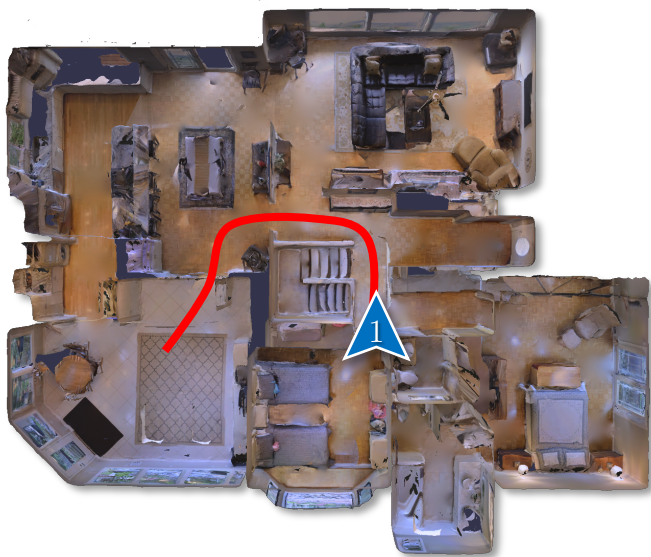




# GRIDTOPIX



GRID  
(teacher)

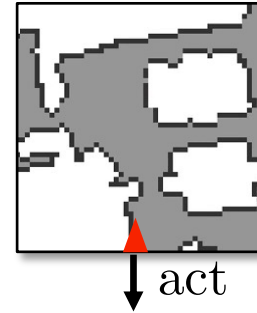
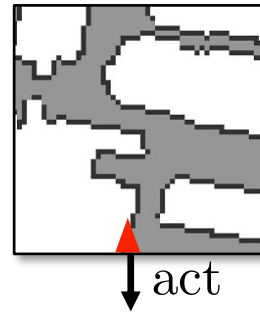
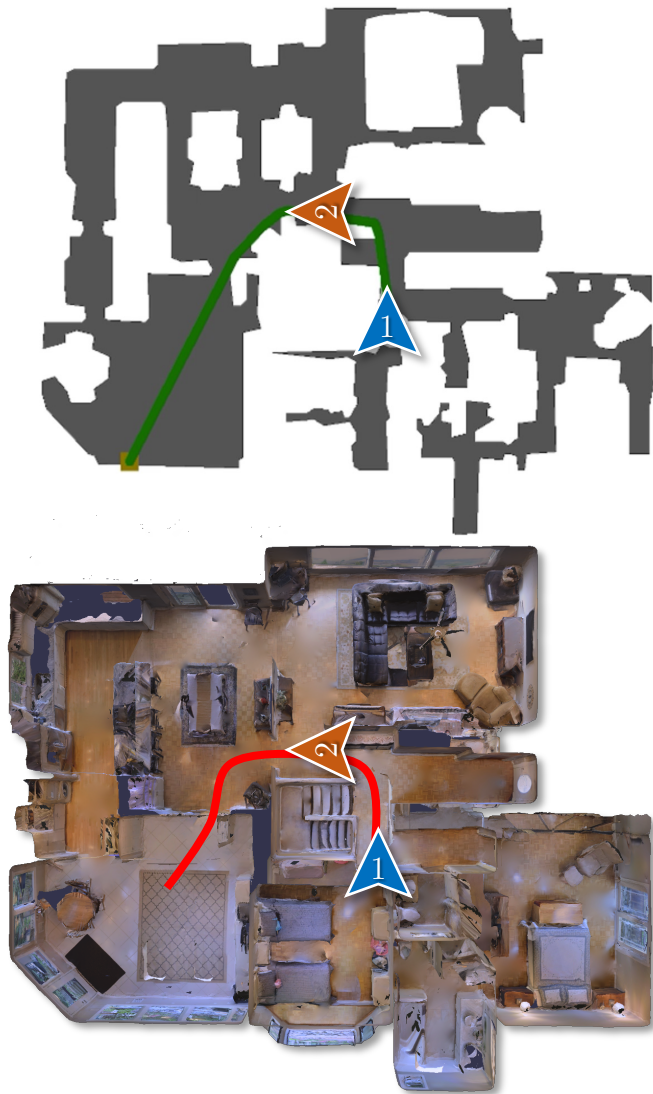


PIX  
(student)

Step=1



# GRIDTOPIX



GRID  
(teacher)



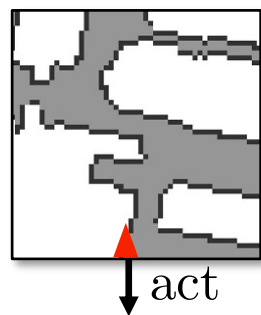
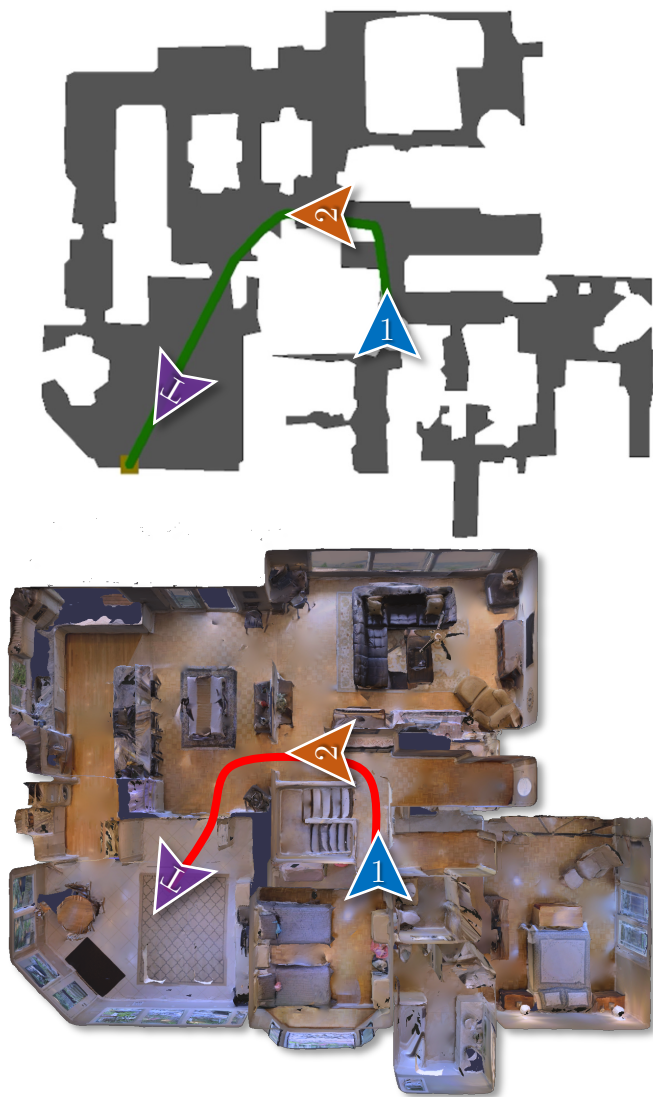
PIX  
(student)

Step=1

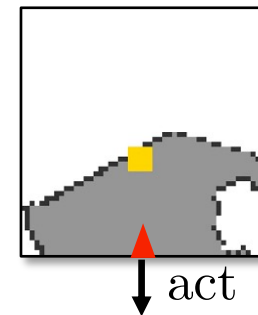
Step=2



# GRIDTOPIX

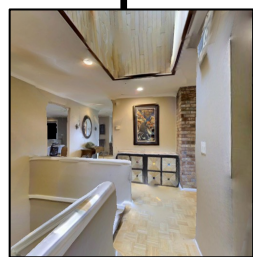


...



GRID  
(teacher)

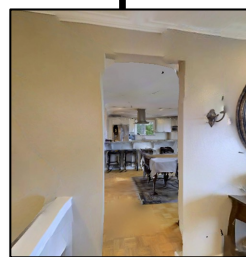
↑ act



Step=1



↑ act

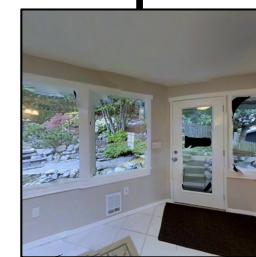


Step=2



...

↑ act



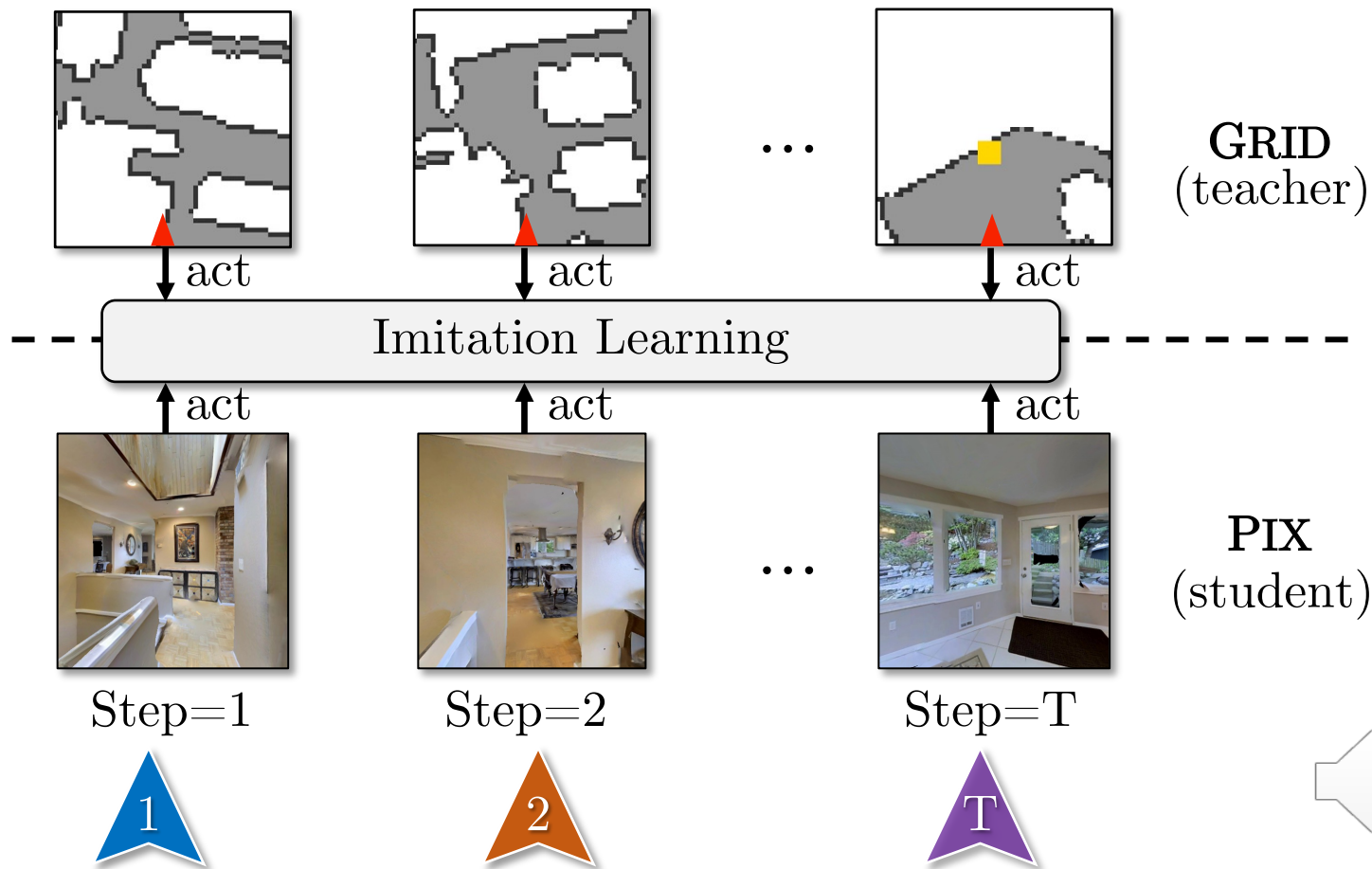
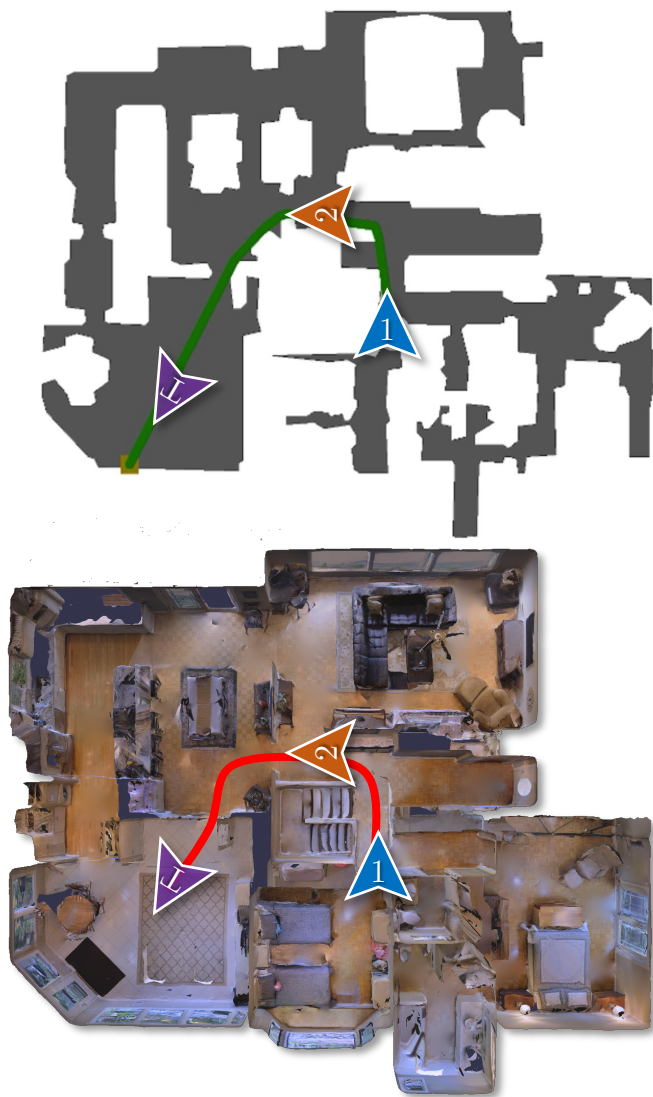
Step=T



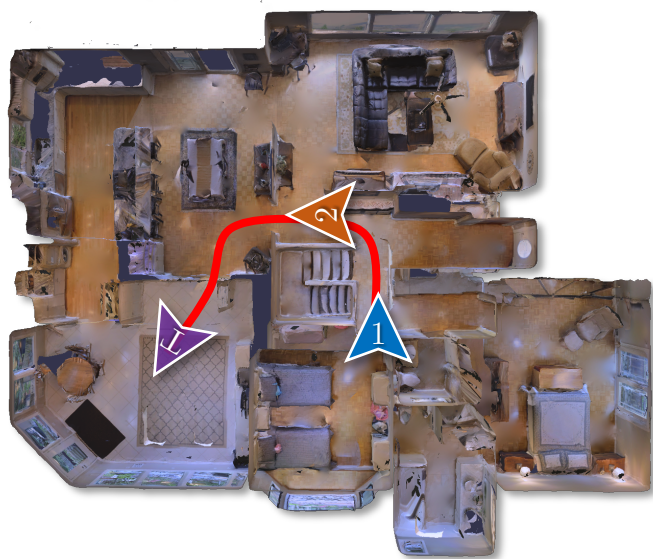
PIX  
(student)



# GRIDTOPIX



# GRIDTOPIX



Step=1



Step=2



...



Step=T



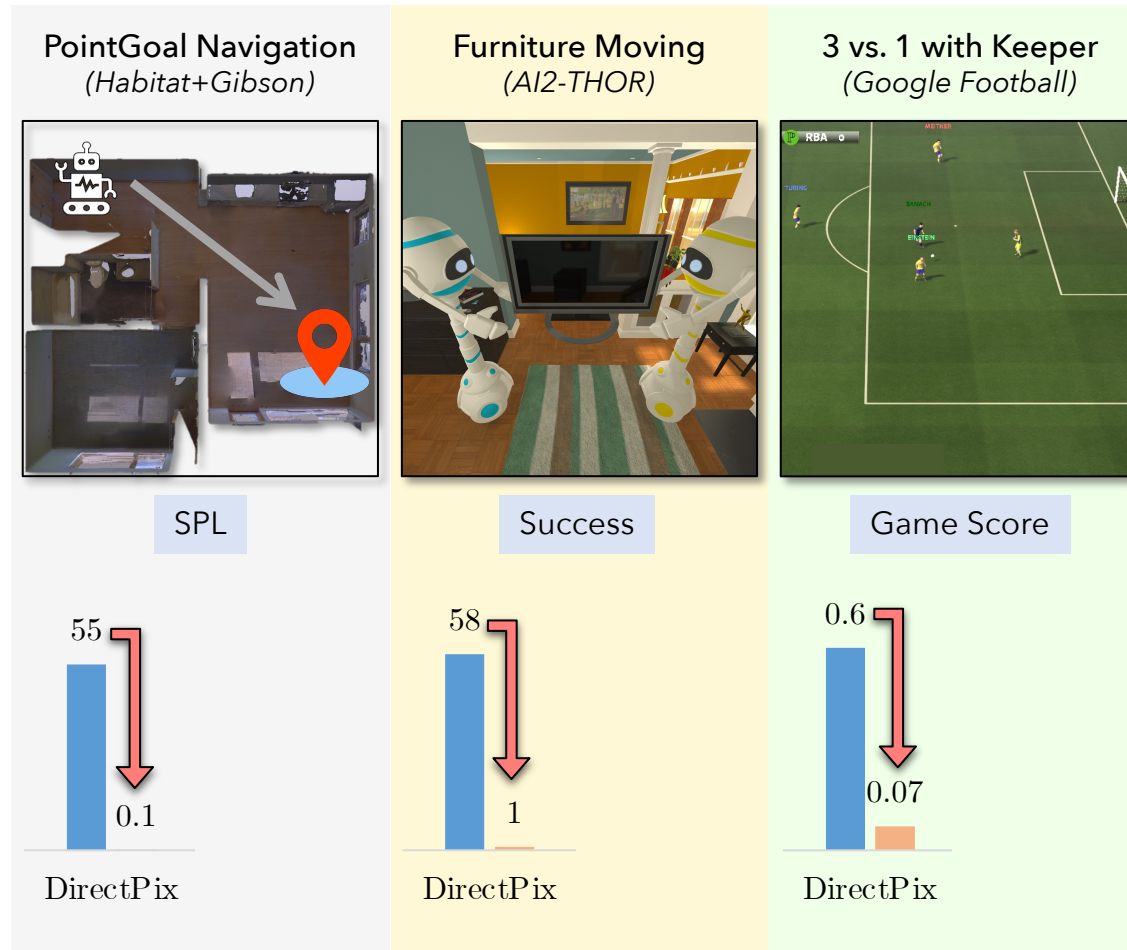
PIX  
(student)



# Preliminary Results

Terminal rewards do not work off-the-shelf.

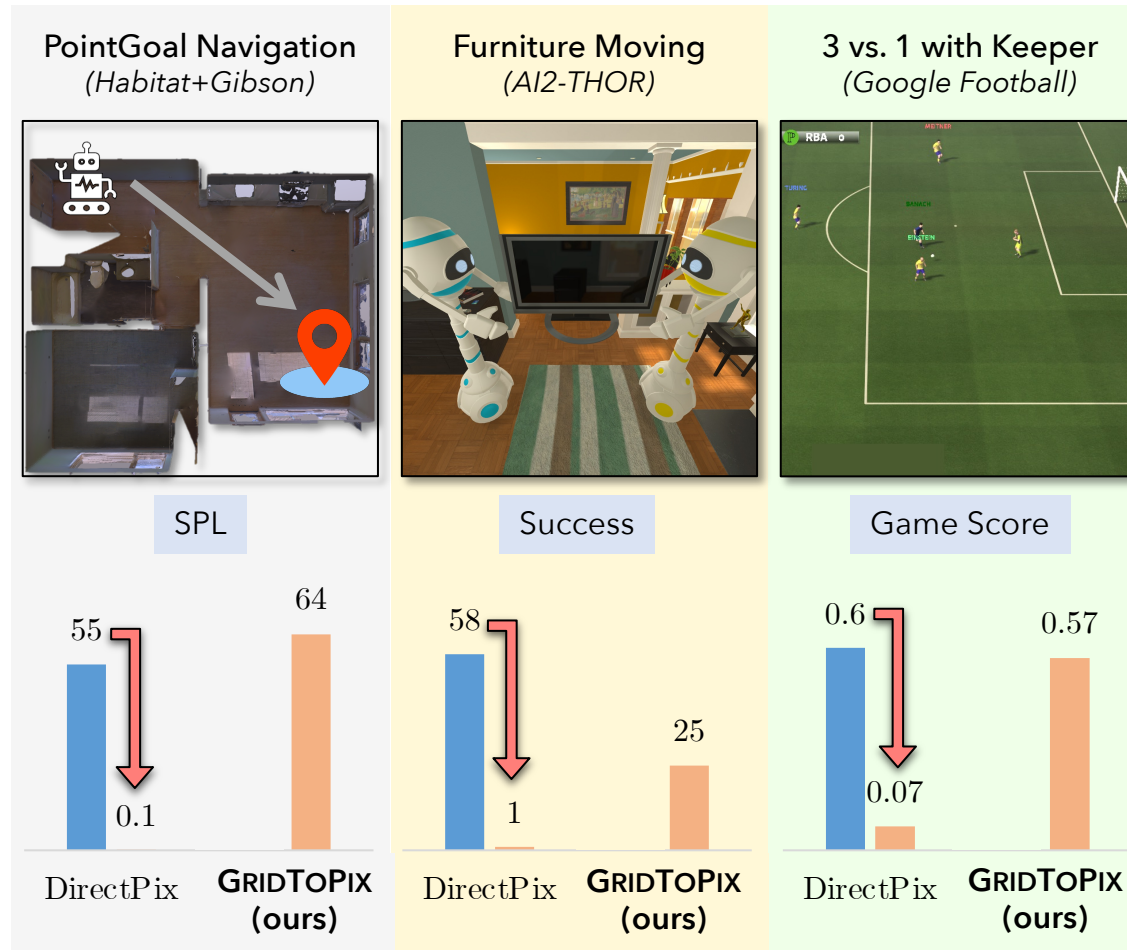
■ Shaped rewards  
■ Terminal rewards



# Preliminary Results

Terminal rewards via  
GRIDTOPIX work well.

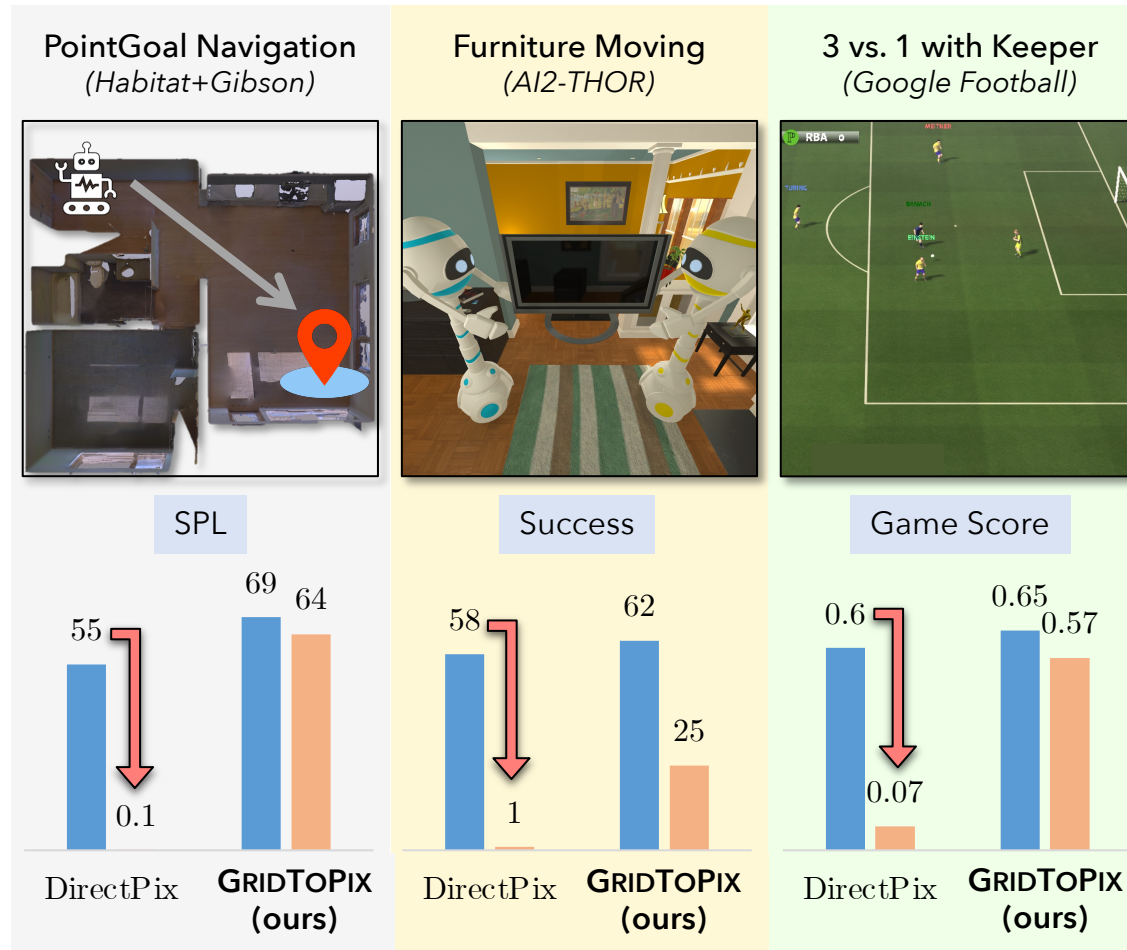
■ Shaped rewards  
■ Terminal rewards



# Preliminary Results

Shaped rewards via  
GRIDTOPIX is better than a  
direct training.

■ Shaped rewards  
■ Terminal rewards







# Collaborative Embodied Agents

## Two Body Problem

CVPR 2019 (oral)

## SYNC Policies

ECCV 2020 (spotlight)

## GRIDTOPIX

(ongoing work)

## Takeaways

- Visual RL agents crave dense and shaped rewards
- GRIDTOPIX leverages gridworlds for free supervision
- Improve results in terminal and shaped reward settings



# Timeline

Steps	Timeline
Submit Imitation Gap work	Spring 2021
Internship at DeepMind	Summer 2021
Further experiments for GRIDTOPIX	Summer 2021 or Fall 2021
Publish GRIDTOPIX	
Publish internship project	Fall 2021 to Spring 2022
Complete dissertation (depends on next step)	Fall 2021 or Spring 2022



# Publications

UNDER REVIEW		<u>Interpretation of Emergent Communication in Heterogeneous Collaborative Embodied Agents</u> S. Wani*, S. Patel*, <b>U. Jain*</b> , A. Schwing, S. Lazebnik, A. X. Chang, M. Savva (Under review at ICCV 2021)
		<u>Cooperative Exploration for Multi-Agent Deep Reinforcement Learning</u> I. Liu, <b>U. Jain*</b> , R. Yeh, A. Schwing (Under review at ICML 2021)
PREPRINTS		<u>AllenAct: A Framework for Embodied AI Research</u> (2020) L. Weihs*, J. Salvador*, K. Kotar*, <b>U. Jain*</b> , K. Zeng, R. Mottaghi, A. Kembhavi <a href="#">[project]</a> <a href="#">[arxiv]</a>
		<u>Bridging the Imitation Gap by Adaptive Insubordination</u> (2020) L. Weihs*, <b>U. Jain*</b> , J. Salvador, S. Lazebnik, A. Kembhavi, A. Schwing <a href="#">[project]</a> <a href="#">[arxiv]</a>
PUBLICATIONS		<u>Multi-ON: Benchmarking Semantic Map Memory using Multi-Object Navigation</u> S. Wani*, S. Patel*, <b>U. Jain*</b> , A. X. Chang, M. Savva <i>Neural Information Processing Systems (NeurIPS), 2020</i> <a href="#">[project]</a> <a href="#">[pdf]</a>
		<u>A Cordial Sync: Going Beyond Marginal Policies for Multi-Agent Embodied Tasks</u> <b>U. Jain*</b> , L. Weihs*, E. Kolve, A. Farhadi, S. Lazebnik, A. Kembhavi, A. Schwing <i>European Conference on Computer Vision (ECCV), 2020 (Spotlight talk)</i> <a href="#">[project]</a> <a href="#">[arxiv]</a>
		<u>SoundSpaces: Audio-Visual Navigation in 3D Environments</u> C. Chen*, <b>U. Jain*</b> , C. Schissler, S. Gari, Z. Al-Halah, V. Ithapu, P. Robinson, K. Grauman <i>European Conference on Computer Vision (ECCV), 2020 (Spotlight talk)</i> <a href="#">[project]</a> <a href="#">[arxiv]</a>
		<u>TABVCR: Tags and Attributes for Visual Commonsense Reasoning</u> J. Lin, <b>U. Jain*</b> , A.G. Schwing <i>Neural Information Processing Systems (NeurIPS), 2019</i> <a href="#">[project]</a> <a href="#">[arxiv]</a>
		<u>Two Body Problem: Collaborative Visual Task Completion</u> <b>U. Jain*</b> , L. Weihs*, E. Kolve, M. Rastegari, S. Lazebnik, A. Farhadi, A. Schwing, A. Kembhavi <i>Computer Vision and Pattern Recognition (CVPR), 2019 (Oral talk)</i> <a href="#">[project]</a> <a href="#">[arxiv]</a>
		<u>Two can play this Game: Visual Dialog with Discriminative Question Generation and Answering</u> <b>U. Jain*</b> , S. Lazebnik and A.G. Schwing <i>Computer Vision and Pattern Recognition (CVPR), 2018</i> <a href="#">[arxiv]</a>
		<u>Creativity: Generating Diverse Questions using Variational Autoencoders</u> <b>U. Jain*</b> , Z. Zhang* and A.G. Schwing <i>Computer Vision and Pattern Recognition (CVPR), 2017 (Spotlight talk)</i> <a href="#">[arxiv]</a>

 Embodied Visual Agents

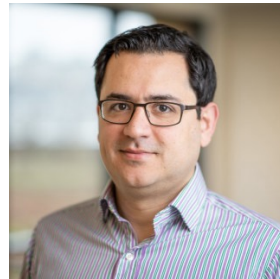
 Communication

 Collaboration

# Thanks!



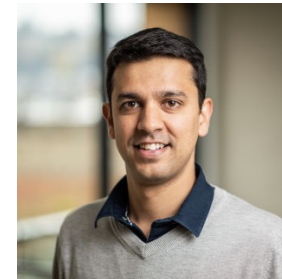
Alexander Schwing



Ali Farhadi  
(UW, Apple)



Angel Chang  
(SFU)



Aniruddha Kembhavi  
(AI2, UW)



Svetlana Lazebnik



Kristen Grauman  
(UT Austin, FAIR)



Luca Weihs  
(AI2)



Manolis Savva  
(SFU)

