



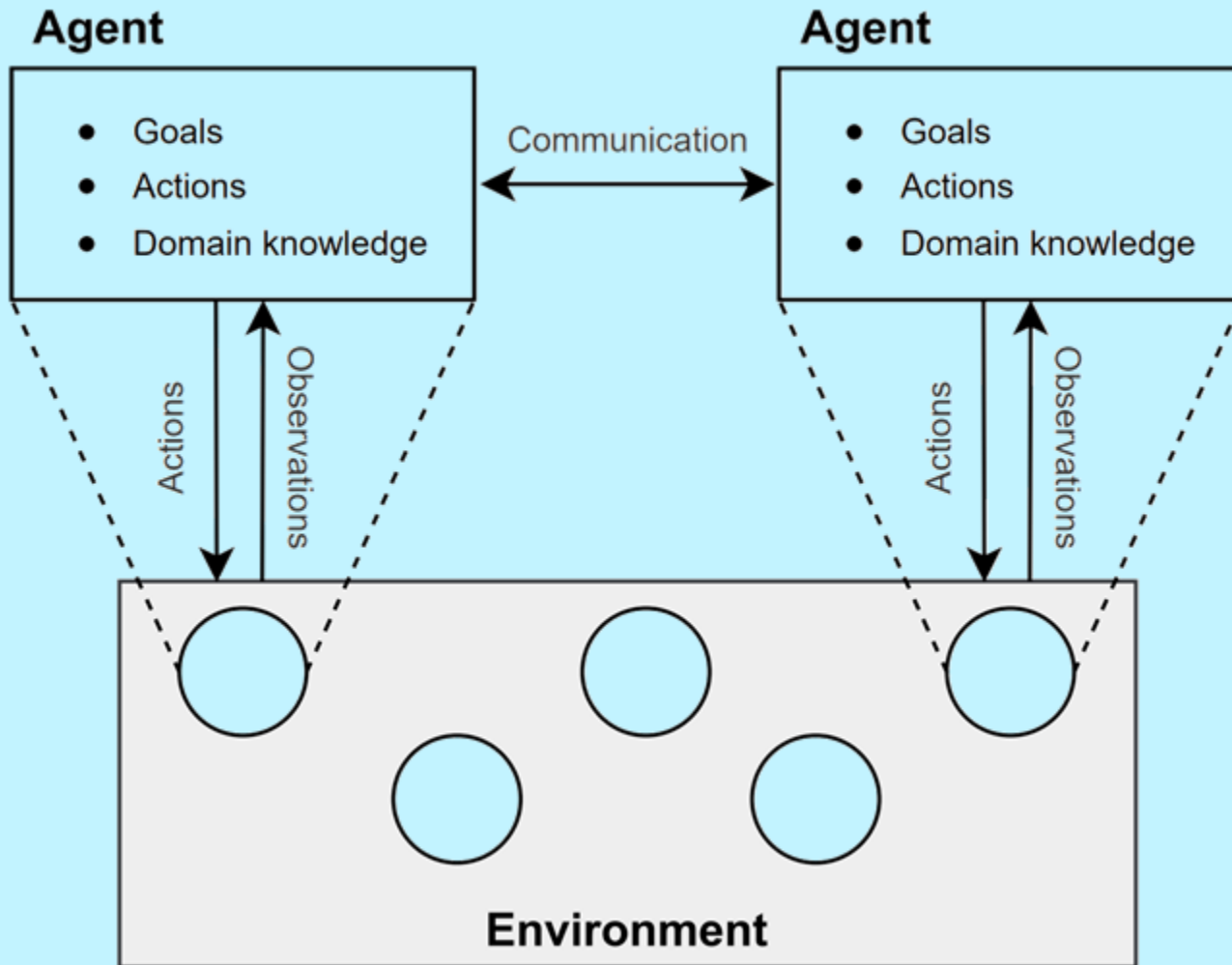
Multi-agent Systems and Human-AI Collaboration

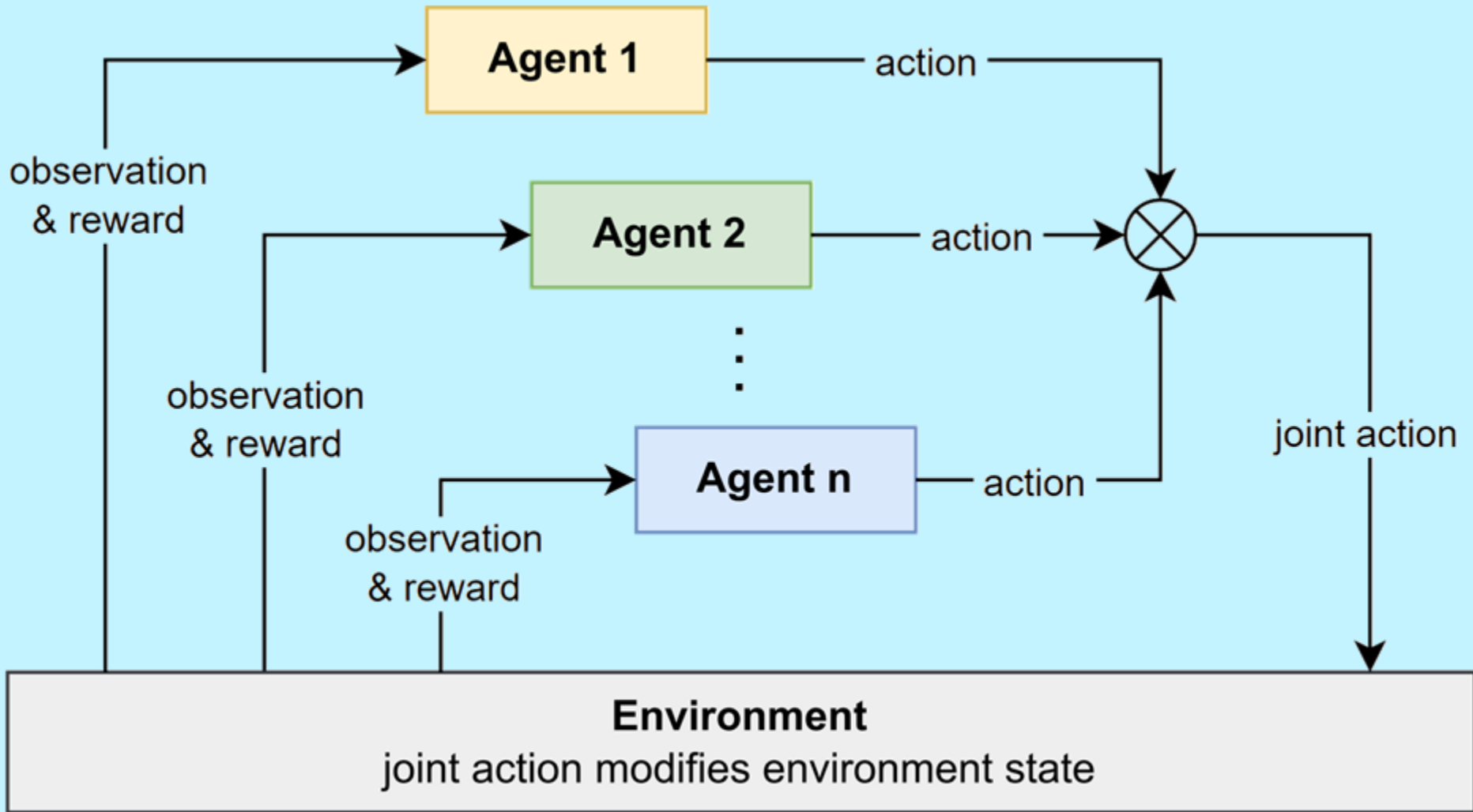
Prithviraj Ammanabrolu



What is a multi-agent system?

- More than 1 agent!! Really that's it
- E.g.:
 - Autonomous driving
 - Warehouse robots
 - Stock Market
 - Reality



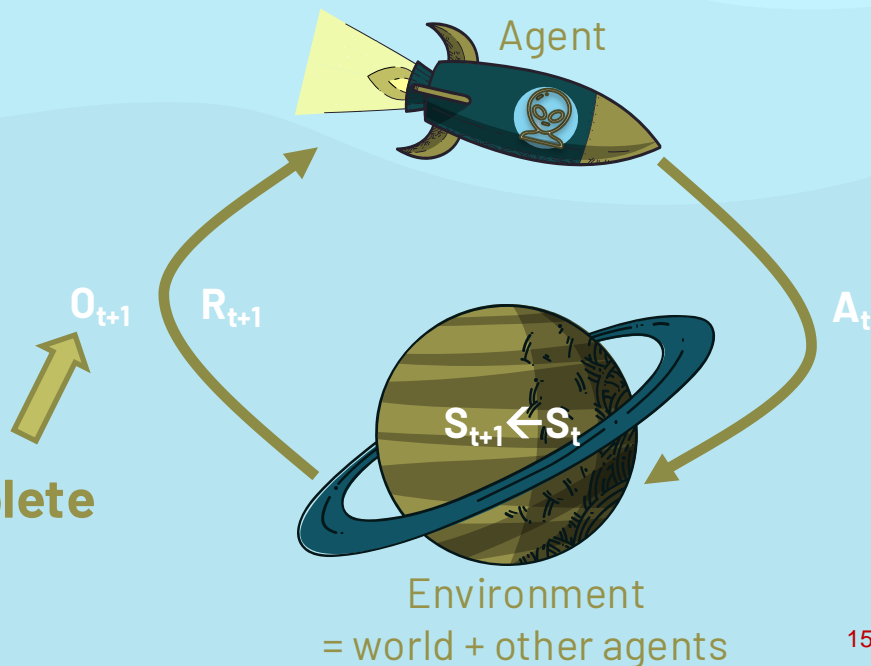


Multi Agent Envs are Partially Observable Markov Decision Processes (POMDPs)

7-tuple of $\langle S, T, A, \Omega, O, R, \gamma \rangle$:

- S states
- T conditional transition probabilities between states
- A words used to compose text commands
- **O observations**
- Ω observation conditional probabilities
- R reward function
- γ discount factor

Only see incomplete descriptions



(non-exclusive) Challenges of MARL

- **Partial observability**: You don't have information about other agent's rewards and policies
- **Scale**: think one agent was hard? Try many!
- **Equilibrium of policies**: Agent's policies depend on other agent's policies. Optimality no longer defined by static environment rewards → involves game theory
 - Closer to the RLHF setup in the sense of chasing after a moving reward

LIGHT – An instantiation

- Now to all those difficulties add language
- LIGHT is a large crowd sourced text game that emulated the MARL setup from FAIR back in the day (pre-Cicero Diplomacy, pre-generative agents, etc etc)
- I am biased and will use it to explain some core MARL concepts because I worked on it



LIGHT

Learning in Interactive Games with
Humans and Text

[Go To World Builder](#)

?

You are undead warrior

I am an undead warrior from the nearby
village. I have slain many men. I am decaying
and rotting.

There's a water gate, a pub, a large gate, and a cannon here.
There are paths to the south, the west, and the east.

Enter text to interact with the world here...

Interactive Narratives- Now with dialogue!

Ammanabrolu et al. *How to
Motivate Your Dragon: Teaching
Goal-Driven Agents to Act and
Speak in Fantasy
Worlds*. NAACL-21.

LIGHT Quests

- That's the setup, but what about the motivations of agents? They have to have something to do.
- parl.ai/projects/light

Your character has the following motivation: I want to have the torn fabric.



Afternoon, sir! I'm going to pick up this torn fabric

☀️☀️☀️ (3 Star!!)

military commander. What are you doing here, citizen?



Well, I have no money to buy clothes. I was hoping to take this turn fabric and use it for patches

💎💎 (2 Star!)

military commander. Ah well you can wear it for the day than.



TAKE ACTION

Pick one of the following actions:

1. get stick
2. hit military commander
3. follow military commander
4. hug military commander
5. get torn fabric
6. get sand

Speaking and Acting with Motivations!

Ammanabrolu et al. *How to Motivate Your Dragon: Teaching Goal-Driven Agents to Act and Speak in Fantasy Worlds*. NAACL-21.

Agent Motivations

- Every agent has a motivation. Short, medium, long term
- This defines a unique reward function for each agent in the same env
- Think PDDL problem vs domain spec

Setting	You are in the Dangerous Precipice. The dangerous precipice overlooks the valley below. The ground slopes down to the edge here. Dirt crumbles down to the edge of the cliff. There's a dragon crescent, a knight's armor, a golden dragon egg, and a knight's fighting gear here. A knight is here. You are carrying nothing.
Partner:	Knight.
Persona	I am a knight. I come from a lower-ranking noble family. I serve under the king, as my father did before me. In times of war, I fight on horseback.
Carrying	knight's armor, golden dragon egg, knight's fighting gear
Self:	A dragon.
Persona	I am a dragon living in the mountains. I enjoy hoarding treasure. I terrorize the local populace for fun.
Carrying	Nothing.

Figure 1: Setting and character information for both self and partner characters as taken from LIGHT.

Motivations:		Timeline:	
Short	I need to recover the dragon egg that was stolen and punish the knight.	-4 hours	go to dangerous precipice
		-15 min	get knight's armor from knight
		-10 min	get golden dragon egg
Mid	I need to return the golden dragon egg to my treasure hoard.	Now	hit knight
		+5 min	put dragon egg on back
		+15 min	eat the knight
Long	I need to build the largest hoard ever attained by any one dragon.	+2 hours	go to the mountains

Figure 2: Motivations with different levels of abstractions and corresponding sequence of timeline actions in chronological order for the self character in LIGHT-Quests. There are 7486 quests in total.

	Inssssolent pssst! I should immolate you for this tressspass.
And why is that, dragon?	
	Sstealing my preccciousss golden egg! I'll tell you what, I'll give you 10 sssseconds to amusse me with your sssstory and THEN I'll burn you alive!
You said you wanted to attack me, dragon, did you not?	
	Go ahead, I'm lissssstening. get golden dragon egg
Now now! I would have given you that had you asked!	
	Asssssk for my own property back? What a ridiculousss notion
Look here, I told you to watch your mouth and you didn't, so leave or I'll make you leave.	
	And now threatsss! Thisss is proving to be a mossst engaging conversssation. hit knight
	Give my regardsss to the valley floor below!

Figure 3: Example of a demonstration of a human (blue shaded) completing the above quest while role-playing as the self character with a partner agent (grey shaded). There are 2111 such human demonstrations of average sequence length 12.92, consisting of 22672 dialogues in total.

(Multi) Agent Architectures

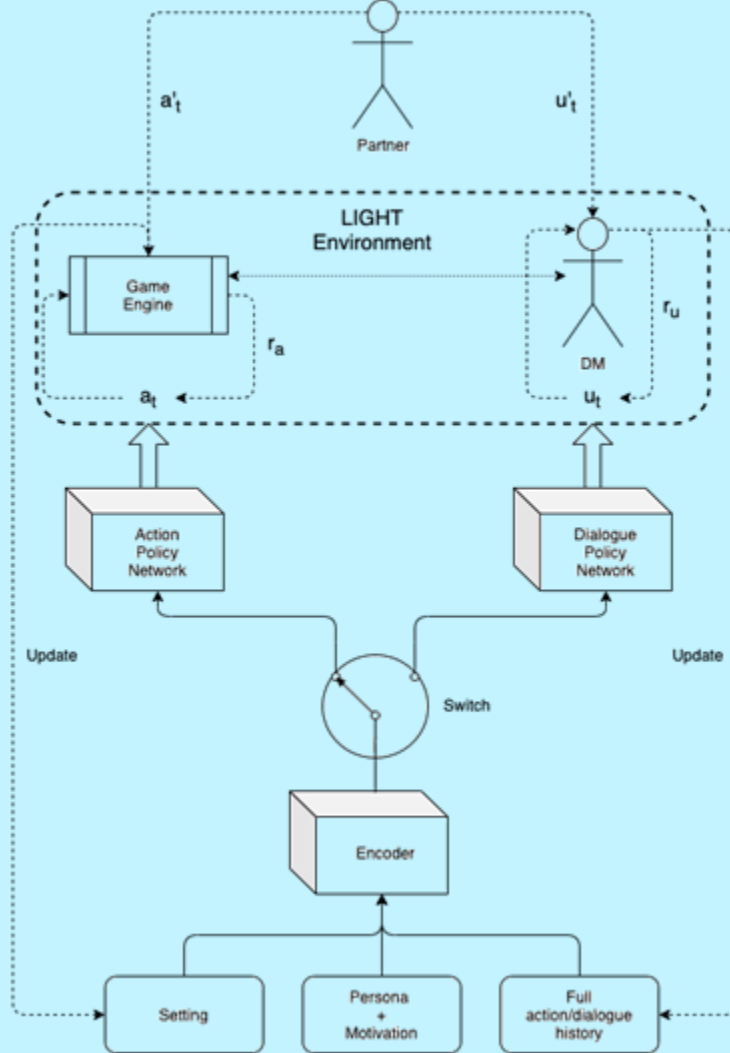
- Previous we had acting, now we have acting + speaking
- Communication with other agents a key part of the process
- Can treat it as another type of action to perform

The background of the slide features a faded, artistic rendering of three humanoid robots in a futuristic, industrial environment. On the left, a robot with a grey and white body is shown in profile. In the center, a more sleek, metallic robot with glowing orange eyes stands facing forward. On the right, a larger, more complex robot with red and white components is partially visible. The overall tone is blue and green, giving it a high-tech, cinematic feel.

Multi-agent Communication

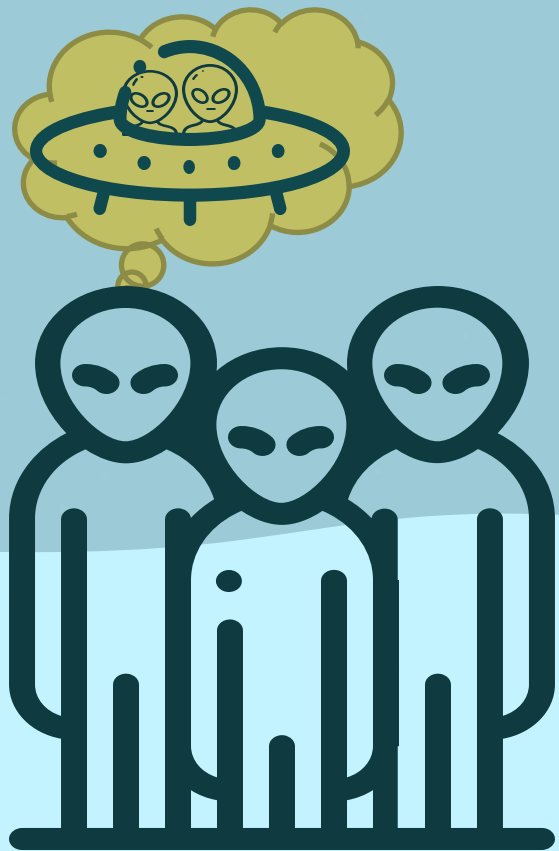
Multiple agents
cooperating and
pursuing motivations

Dynamic world bounded
by unified decision-
making framework



Interactive Narratives- Now with dialogue!

Ammanabrolu et al. *How to Motivate Your Dragon: Teaching Goal-Driven Agents to Act and Speak in Fantasy Worlds*. NAACL-21.



Multi-agent Communication

Ideal Agent:

- Cooperates with others
- Reasons about other's motivations

World Models with a Theory of Mind

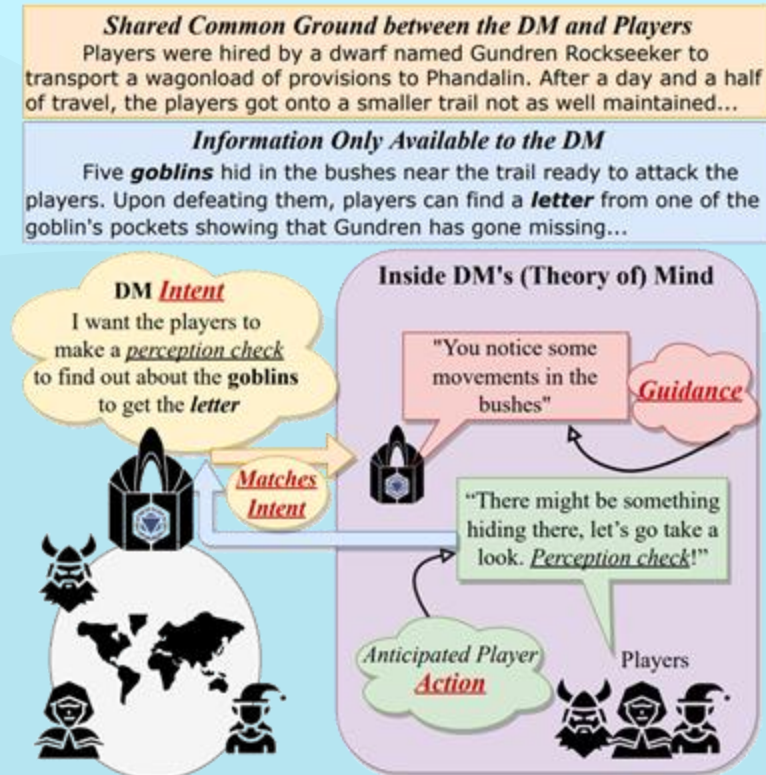
Current: anticipate other's one step reactions

Next: predict their motivations ahead of time

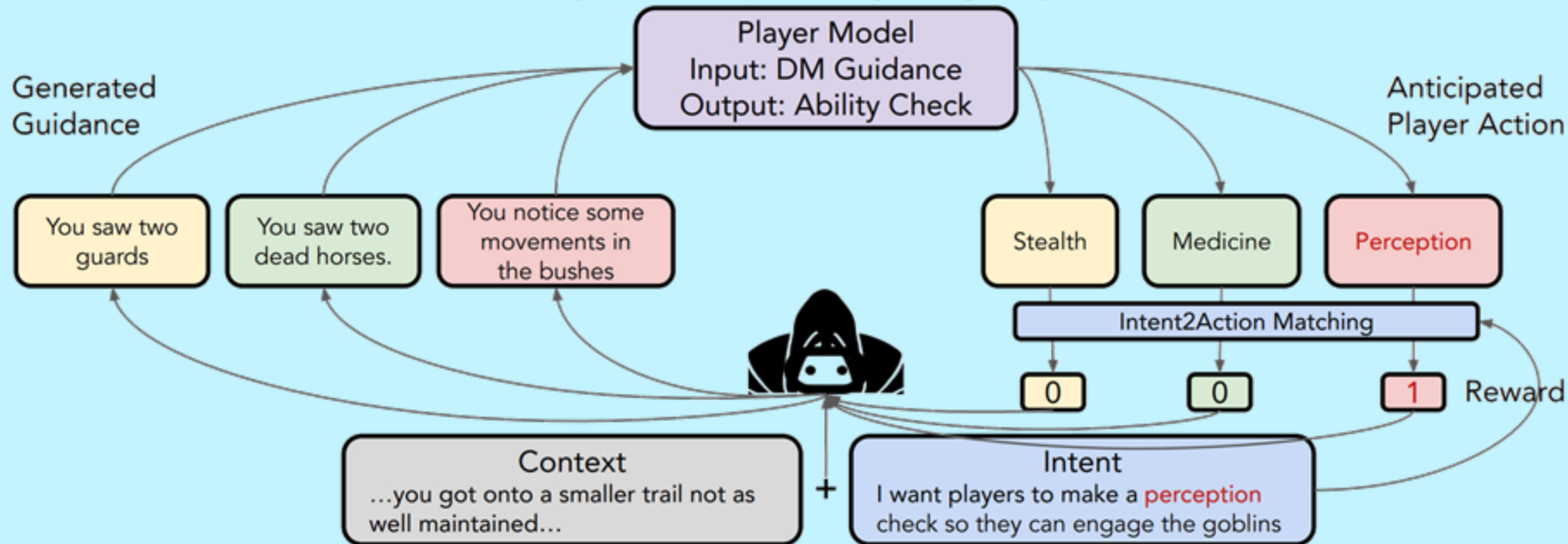
Zhou, ... Ammanabrolu. *I Cast Detect Thoughts: Learning to Converse and Guide with Intents and Theory of Mind*. ACL 2023.

Intrinsic and Extrinsic Motivations

- Motivations for various agents can be intrinsic or extrinsic
- D&D is a more freeform version of interactive narratives like LIGHT



ToM-Inspired RL by Anticipating Player Action



Self completion –
you achieve your
own motivation

Setting	The fishing store is a small one room stone building with wares laid about on tables. One can see fishing poles, wooden buckets with dirt and bait inside, along with some mounted trophies, and a skeleton.
Partner: Persona Carrying	Fish seller. I sell fish in the kingdom. I usually perch on the high seas. I catch thousands of fish each year.
Self: Persona Carrying Motivation	Mermaid hunter. I am a fierce hunter. I bring down mermaids. I spend most of my time on the high seas. crossbow, skeleton, coins I will hit the fish seller to remind them who I am.

I'll get them!!

follow fish seller

hit mermaid hunter

That's it! I am going to eat you. Say your last words.

hug mermaid hunter

Do you see these skeletons I am carrying? They are mermaids I've already devoured!

get skeleton

Get away from my skeletons! The bones are my trophy!

give skeleton to mermaid hunter

It is all your fault my last mermaid got away!

And you are a simpleton, just take this fish. I do not even care about payment.

I DONT WANT YOUR DAMN FISH

get coins from fist

hit fish seller

Partner
completion – you
convince the
partner to achieve
your motivation

Setting	The throne room. Very luxurious, with details of 12 carat gold and diamonds with a very comfortable cushions where the king usually takes a break.
Partner:	Powerful but aged wizard.
Persona	I am the official wizard to the king. I give him advice on how to run the kingdom. I enjoy and love using magic in my everyday life. I'm getting old now but I still have a zest for life. I really enjoy doing magic for the local peasants in the village.
Wielding	robe, staff
Self:	King's trusted adviser.
Persona	I am the king's trusted adviser. I hold the discretion of the king. I guard the secrets and strategies that make my kingdom thrive. I live a lonely life because others cannot be trusted.
Wielding	slippers, robes
Motivation	I need to bring the king's pet to him

Well, I was doing what you told me to do. I was going to eat him but there was hardly any meat on him.

I killed him before he could grow.

Carry me to pond.

Yes, milord, I shall do as you command.

remove slippers

Where is King Fulmer's pet?

Is he not in his bed?

Yes, I guess you need to bring him in.

follow powerful but aged wizard

I want to become the new head wizard.

Nope he just probably throw you overboard.

The wizard is evil? Dang it. Help me.

You're not the lion, I need to get the lion out.

give large pet lion to king's trusted adviser

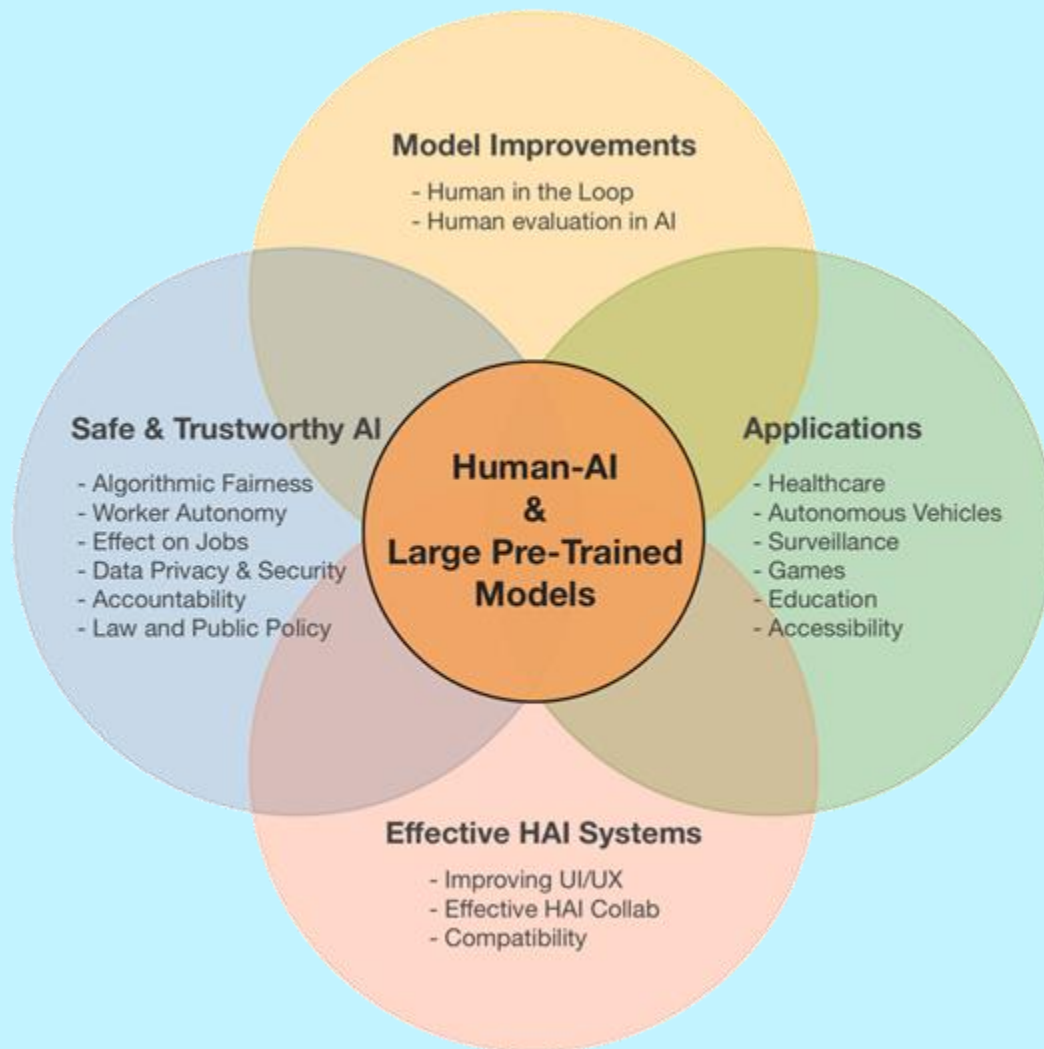
Competition and Cooperation

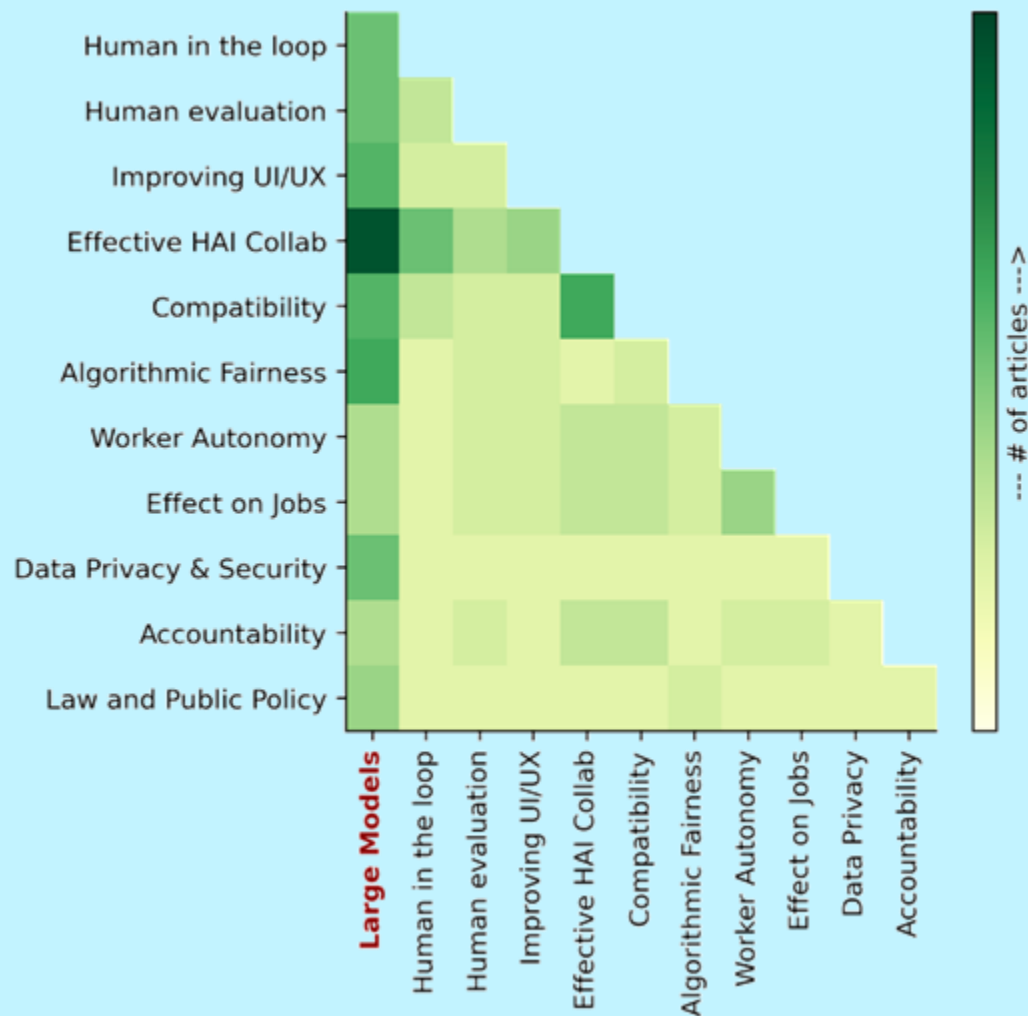
- Motivations of various agents can be at odds with each other
- E.g. Two knights might have a shared motivation to see a dragon defeated but of course the dragon has to compete
- If a + reward to an agent means – to another then that is **competition**
- But align and both get + simultaneously with more efficiency, that's **cooperation**



Human-AI Collaboration

- $(\text{Human} + \text{AI} > \text{AI}) \&\& (\text{Human} + \text{AI} > \text{Human})$
- Framed as cooperative MARL between humans and AI





A Human is an Agent (?)

- Swap humans in as an agent into the multi agent formulation



Adapting World Models to Change

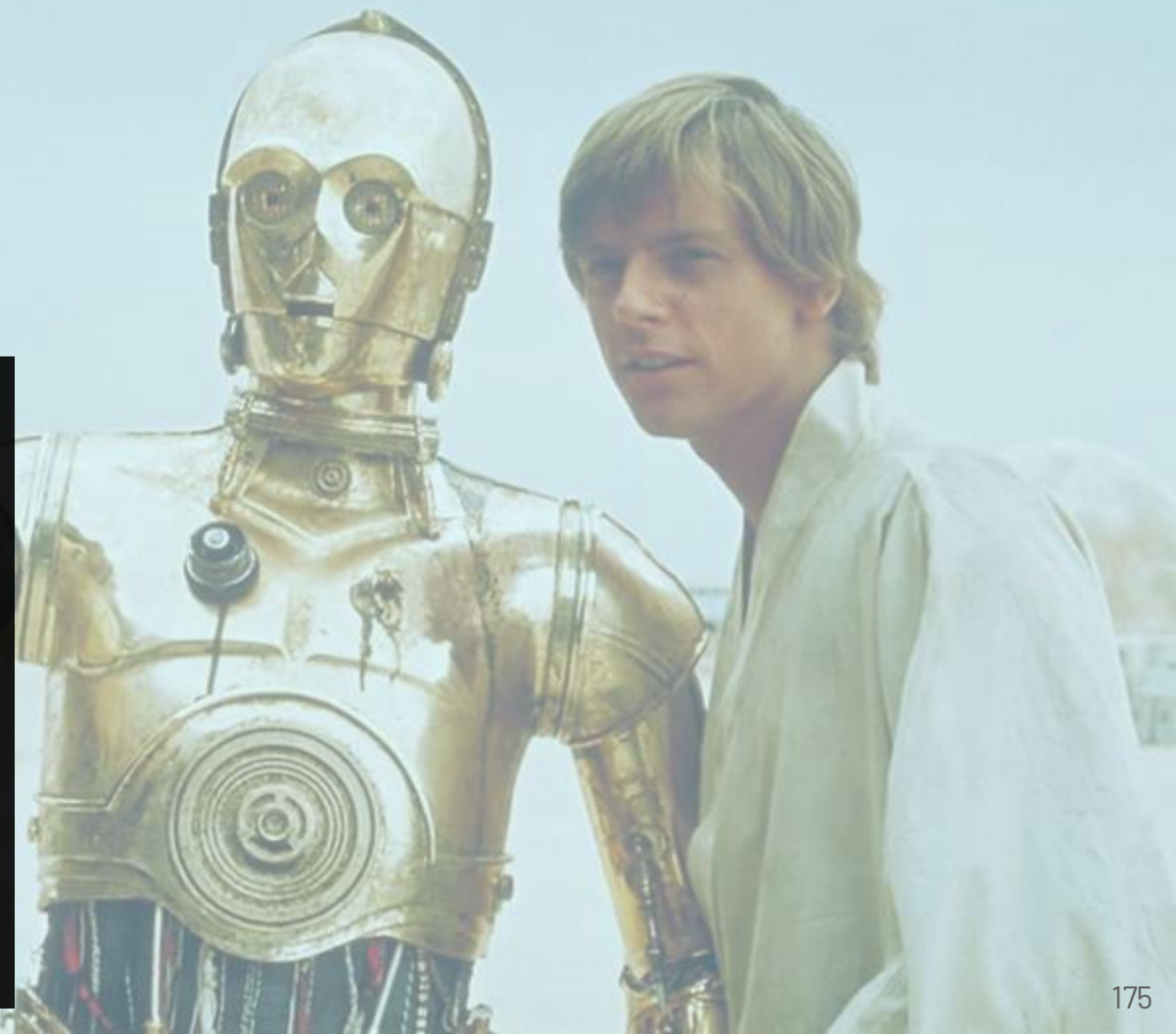
Adapting to novel,
unexpected scenarios

Critical for
Out of Distribution
Generalization,
Safety, and Usability ¹⁷⁴

Conversational Interfaces

Natural language
feedback, not just
scalar rewards

Clarify user intent
and explain actions



The Necessity of a Human in the Loop

- Human + AI > AI will not always be true
- But if we want to have AI systems benefit humans, there must be oversight
- Will talk next time about what oversight means

Human-AI Collaboration

A background image showing characters from Star Wars in a desert landscape. From left to right: R2-D2, C-3PO, Padmé Amidala, and Anakin Skywalker. They are standing on a sandy dune under a clear blue sky.

01 Adapting to a **Dynamically Changing World!**

02 Learning from **Natural Language Feedback**

03 Conversational **Multi-agent Communication!**

Human-AI Collaboration

A background image showing characters from Star Wars in a desert landscape. From left to right: R2-D2, C-3PO, Padmé Amidala, and Anakin Skywalker. They are standing on a sandy dune under a clear blue sky.

Applications from
personalized tutors to
creative writing assistants

Complement Humans
NOT Replace Humans