

Using Insights from Human Language for Multi-Agent Decision Making



Roma Patel

DIMACS, Rutgers University, October 2023

How do humans collaborate to solve tasks?

How do humans collaborate to solve tasks?



Communication is **beneficial** to the team to maximise efficiency and distribution of resources to solve the problem

How do humans collaborate to solve tasks?



Communication is **beneficial** to the team to maximise efficiency and distribution of resources to solve the problem



Communication is **necessary** when there is obstructed or partial observability and agents can talk to, but not view others

Collaboration is adaptable when teams speak the same language



Communication is maximally beneficial when **different components of the solution** can understand one another.

When **team of agents** with partial observability vs. overhead agents to **coach or advise** teams speak the same language

Collaboration is adaptable when teams speak the same language



Communication is maximally beneficial when **different components of the solution** can understand one another.

When **team of agents** with partial observability vs. overhead agents to **coach or advise** teams speak the same language

An example from the real world

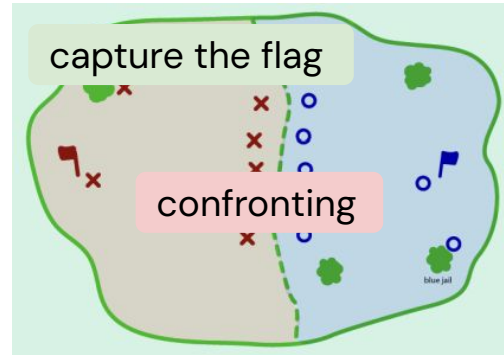


Real world scenario

Some standard multi-agent tasks



Real world scenario



Existing multi-agent tasks



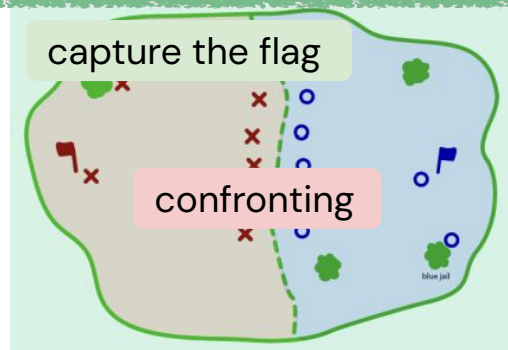
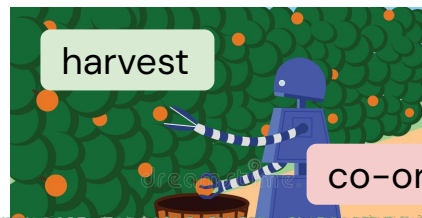
How do we collaborate to solve tasks the way humans do?



learning from or with humans



use natural language



Real world scenario



Existing multi-agent tasks

Modelling language is complex

Modelling language is complex

understanding natural language

- intended meaning

"Walk to the leftmost tree"



$(0, 0), (0, 1), (0, 2), (1, 2), (2, 2) ..$

- fluent generation



"There are two apples left so I will pick only one."



learning transferable representations

- using information from large-scale pretraining



"Apples and oranges are both fruits."

- tying language to images and control policies



la pomme rouge

But finally...

But finally...

Language Processing, Machine Learning

How Large Language Models Will Transform Science, Society, and AI

Scholars in computer science, linguistics, and philosophy explore the pains and promises of GPT-3.

Better Language Models and Their Implications

We've trained a large-scale unsupervised language model which generates coherent paragraphs of text, achieving state-of-the-art performance on many language model

May 25, 2021

Researchers Evaluate Neural Language Models, Find XLNet Excellent

Oliver Peckham



Lexical substitution is essentially the process that a thesaurus helps humans to perform: replacing words in a sentence without changing the meaning. Now, researchers from the Skolkovo Institute of Science and Technology in Moscow – [Skoltech](#), for short – have completed a groundbreaking, large-scale study to examine how the most advanced neural language models perform when handling lexical substitution tasks.

While simple on paper, humans, of course,

MIT News

ON CAMPUS AND AROUND THE WORLD

 [SUBSCRIBE](#)

Teaching language models grammar really does make them smarter

Leverage existing language models for important MA problems

understanding natural language

- intended meaning

Tools to understand and generate language

- fluent generation



"There are two apples left so I will pick only one."

learning transferable representations

- using information from

Large pre-trained language models

- tying language to images and control policies



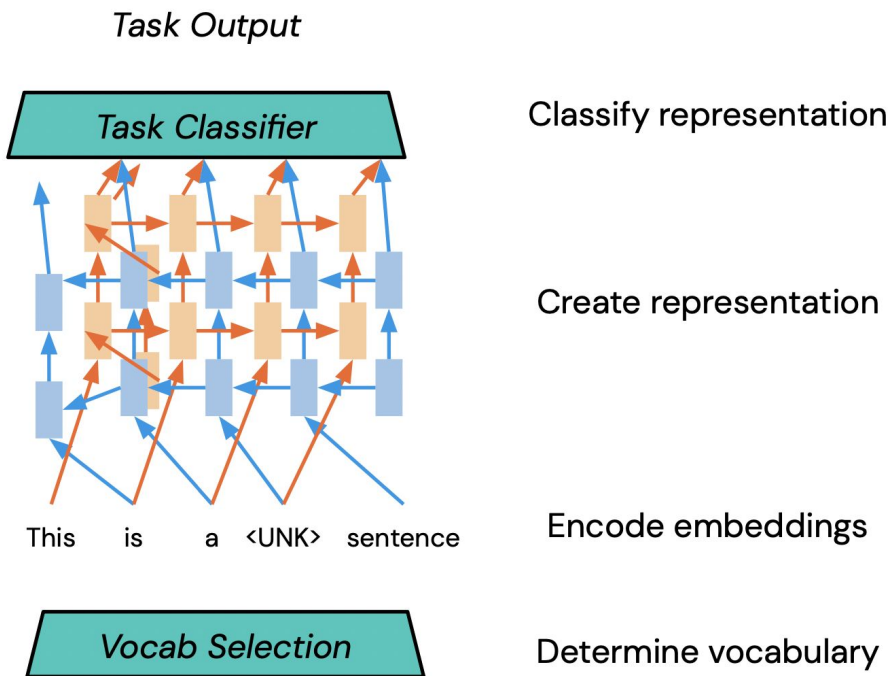
la pomme rouge

How can game theory help language models?

A standard language task

The food was pretty good, but I won't bother coming back.

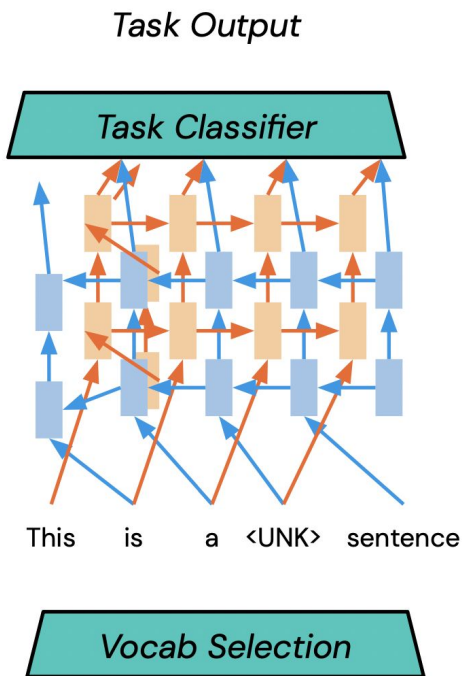
Classify sentiment (+ve or -ve)



A standard language task

The food was pretty good, but I won't bother coming back.

Classify sentiment (+ve or -ve)



Classify representation

Create representation

Encode embeddings

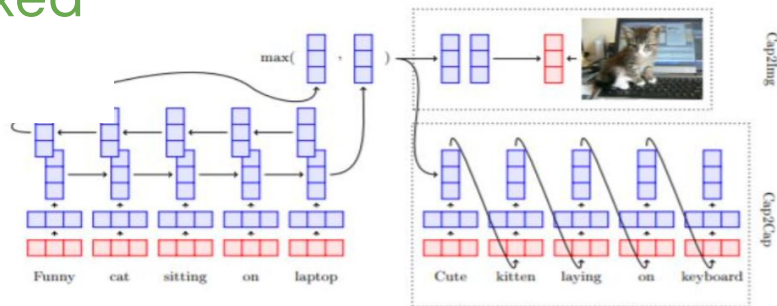
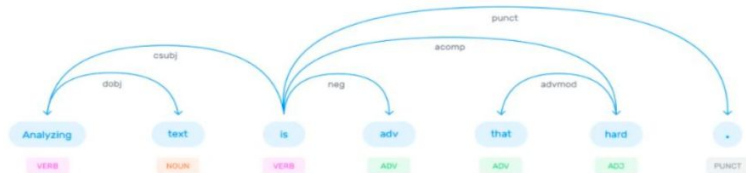
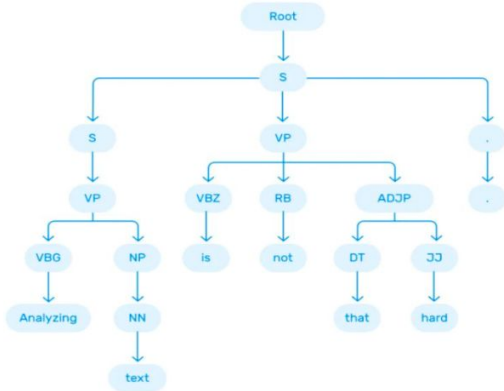
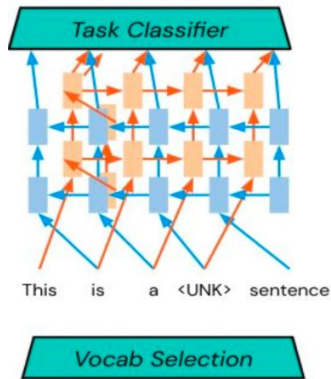
Determine vocabulary

For large amounts of data and large vocabularies makes encoding all words computationally intractable

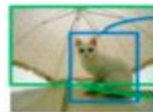
Frequency, information-theoretic measures (TF-IDF) to cut off some words

A standard language task

Nearly all language tasks need a fixed input vocabulary



A cat is on the ground.



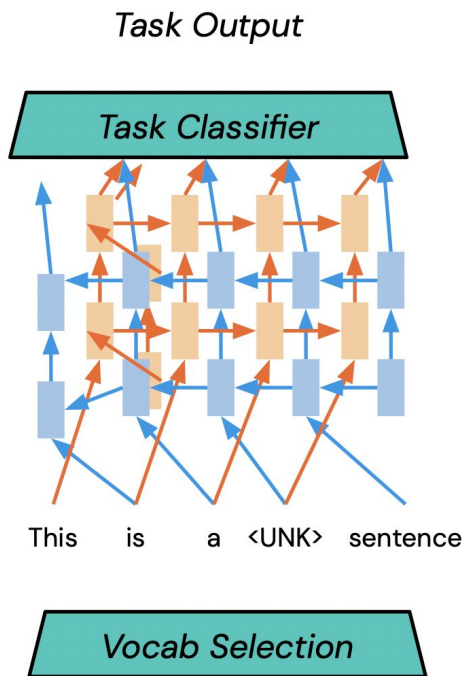
A cat stands under an umbrella.



A dog sits under an umbrella.



Can we use game theoretic power indices to rank words?



Classify representation

Create representation

Encode embeddings

Determine vocabulary

Classify sentiment (+ve or -ve)

For large amounts of data and large vocabularies makes encoding all words computationally intractable

Frequency, information-theoretic measures (TF-IDF) to cut off some words

Shapley index?

Setting up language tasks game-theoretically

Game theory

Language task

Game	=	Classification task
Players	=	Words
Coalitions	=	Sets of words
Contribution	=	Gain in task performance
Valuation fn	=	Model accuracy
Shapley value	=	Average contribution across coalitions

Computing Shapley Values: Algorithm

Algorithm 1 Algorithm for Shapley Vocabulary Selection

```
1: Inputs: NLP dataset  $D$  with full vocabulary  $V$ 
2: for each word  $w$  in  $V$  do
3:    $\phi_w \leftarrow 0$  (initialise Shapley value estimate)
4:   for  $i=1$  to  $S$  (number of sampled permutations) do
5:      $\pi \leftarrow \text{Random-Permutation}(V)$ 
6:      $C_1 \leftarrow b(\pi, w)$  (predecessors of  $w$ )
7:      $C_2 \leftarrow C_1 \cup \{w\}$  (predecessors including  $w$ )
8:      $f_1 \leftarrow \text{TrainModel}(C_1)$  (Train with vocabulary  $C_1$ )
9:      $f_2 \leftarrow \text{TrainModel}(C_2)$  (Train with vocabulary  $C_2$ )
10:     $m(\pi, a) \leftarrow (\text{Performance}(f_2) - \text{Performance}(f_1))$ 
11:     $\phi_w \leftarrow \phi_w + m(\pi, a)$ 
12:   end for
13:    $\phi_w \leftarrow \frac{1}{S} \phi_w$  (average marginal contributions)
14: end for
15: Rank words in  $V$  based on Shapley estimates  $\pi_w$ 
16: Return top  $k$  words in ranking
```

Create classifier model

Train model with coalition (V) including word
Train model with coalition (V) not including word

Average contributions over lots of samples of
coalitions

Rank words by Shapley estimate

Computing Shapley Values: Algorithm

Cost of computing Shapley Values

1. For each word (> 10,000) estimate a value
2. Estimating the value involves:
 - a. Sampling a coalition (> 10,000!)
 - i. Training a model on coalition + word
 - ii. Training a model on coalition
 - b. $\wedge$ doing this a considerable number of times to get a good estimate
3. Parallelise
 - a. Perform each computation on a separate CPU
 - b. Store all valuation function outputs and average at the end
 - c. Write model in Jax for faster computation (e.g., for bigger models like LSTM/Transformer)
 - d. Use a simpler valuation function (logistic regression classifier)

Evaluations

Question 1. Given a task & dataset, do we get better test-time performance with our Shapley vocabulary selection algorithm compared to other vocabulary-selection methods?

Question 2. How many samples do we need in order to estimate a (good) approximate Shapley value? Does performance steadily increase and/or asymptote at some point?

Question 3. Do we see differences in performance/estimation of values for different vocabulary sizes i.e., increasing the population size or number of total players in the game?

Evaluations

Question 4. Do we have insights about the differences in power indices and which works better/why i.e., Shapley vs. Banzhaff?

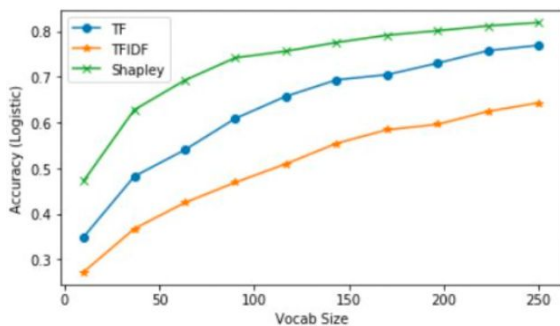
Question 5. Do we have insights about the valuation function used? I.e., do we see differences in performance across a range of classification model architectures?

Question 6. Do we see robust performance across a range of different task structures (single sentence, documents) and datasets?

Results

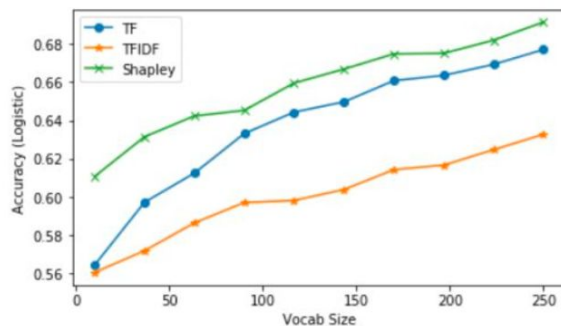
Three different tasks: documentation classification, single sentence classification, question-answer classification

Vocab size vs Performance



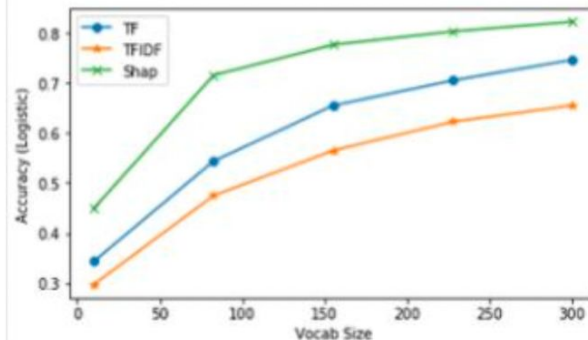
Document classification

Vocab size vs Performance



Single sentence classification

Vocab size vs Performance



Question-answer classification

Results across task structures

Task & Dataset	Method	Vocab	Acc
SST-2 (Socher et al., 2013)	TF-IDF	17,539	80.2
	Frequency		80.3
	Banzhaf		81.7
	Shapley		81.9
COLA (Warstadt et al., 2019)	TF-IDF	9007	63.5
	Frequency		63.7
	Banzhaf		63.9
	Shapley		64.2
SNLI (Bowman et al., 2015b)	TF-IDF	42,392	83.9
	Frequency		83.9
	Banzhaf		84.1
	Shapley		84.3
QQP (Wang et al., 2018)	TF-IDF	117,303	80.8
	Frequency		81.2
	Banzhaf		81.9
	Shapley		81.9
AG-NEWS (Zhang et al., 2015)	TF-IDF	159,697	79.6
	Frequency		78.5
	Banzhaf		79.9
	Shapley		80.2
YELP (Zhang et al., 2015)	TF-IDF	458,705	84.5
	Frequency		83.9
	Banzhaf		86.7
	Shapley		87

Single sentence classification

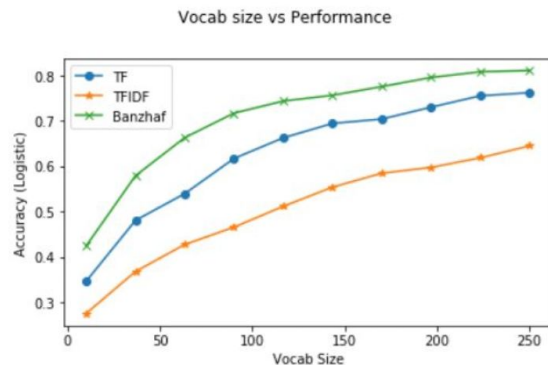
Question-answer classification

Document classification

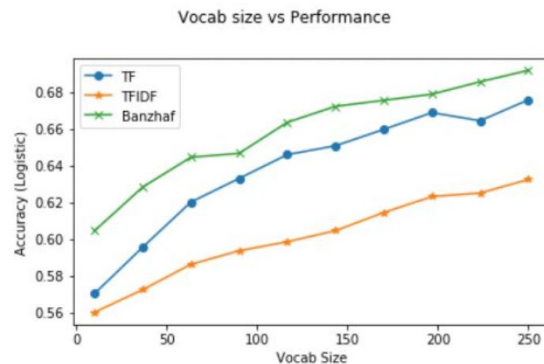
Table 1: Performance of vocabulary selection methods across datasets and tasks, at a target vocabulary size of $|V'| = 750$ words (column 3 is initial vocabulary size). Note performance is lower than state-of-the-art methods, as results are based on a significantly reduced vocabulary size (and using a simple LSTM architecture, with no hyperparameter tuning).

Does Banzhaf also improve over baselines?

Different power indices: Banzhaf (in comparison to baseline frequency and TFIDF)



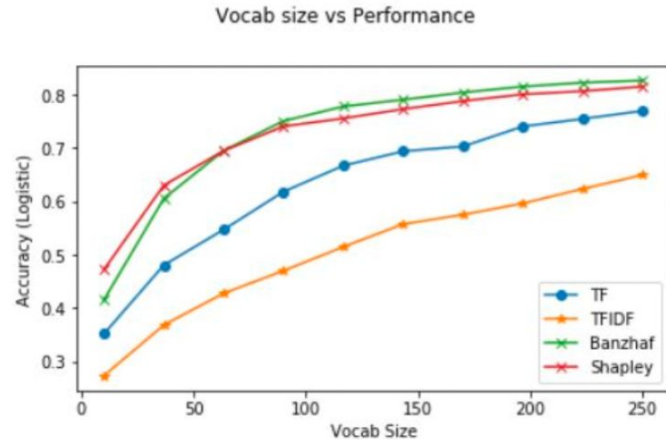
Banzhaf (document)



Banzhaf (single sentence)

How do Shapley and Banzhaf empirically compare to one another?

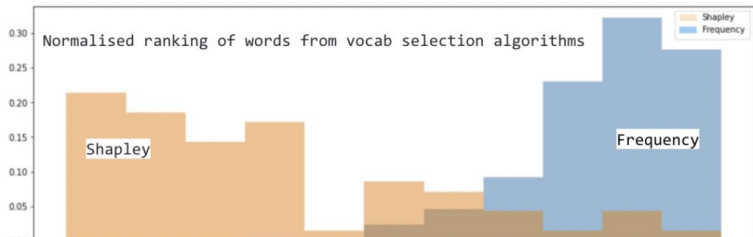
Different power indices: Shapley vs. Banzhaf (in comparison to baselines frequency and TFIDF)



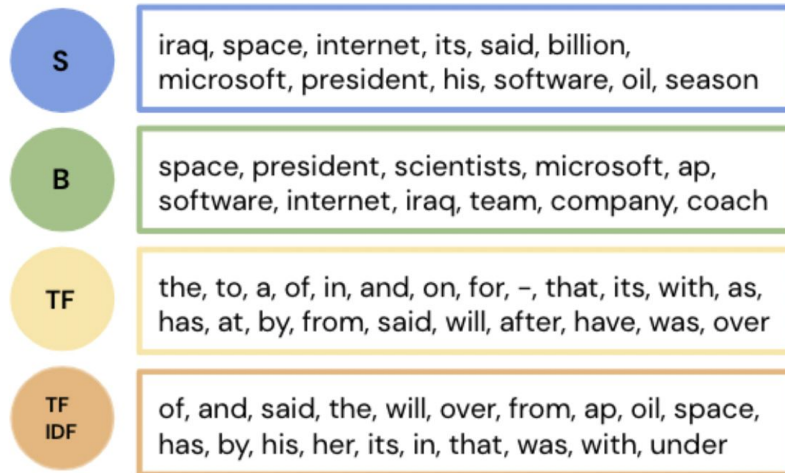
Banzhaf vs. Shapley

Qualitative analysis

superbly photographed and staged by mendes with a series of riveting set pieces the likes of which mainstream audiences have rarely seen .



Distribution of scores in a sentence



Top weighted scores for each word

Thank you! Questions?

How can we use language for multi-agent systems?

Language for communication



Language as supervision



Understanding and improving models



RLang: A Declarative Language for Expressing Prior Knowledge about MDPs
Roma Patel*, Rafael Rodriguez-Sanchez*, Stefanie Tellex, George Konidaris.
ICML 2022 (in submission)



Mapping Language Models to Grounded Conceptual Spaces.
Roma Patel and Ellie Pavlick. ICLR 2022



Does linguistic bias affect generative language models?
Roma Patel and Ellie Pavlick. EMNLP 2021.



Game-theoretic Vocabulary Selection for Text Classification Tasks
Roma Patel, Marta Garnelo, Ian Gemp and Yoram Bachrach. NAACL 2021.



Room-Across-Room: Multilingual Vision-and Language Navigation with Dense Spatiotemporal Grounding.
Alexander Ku*, Peter Anderson*, **Roma Patel**, Eugene Ie, Jason Baldridge. EMNLP 2020.



On the Relationship Between Structure in Natural Language and Models of Sequential Decision Processes
Roma Patel*, Rafael Rodriguez-Sanchez*, George Konidaris. LAREL Workshop, ICML 2020.



Robot Object Retrieval with Contextual Natural Language Queries.
Thao Nguyen, Nakul Gopalan, **Roma Patel**, Matthew Corsaro, Ellie Pavlick, Stefanie Tellex. RSS 2020.



Planning with State Abstractions for Non-Markovian Task Specifications
Yoonseon Oh, **Roma Patel**, Thao Nguyen, Baichuan Huang, Ellie Pavlick, Stefanie Tellex. RSS 2019.



Grounding Language to Non-Markovian Tasks with No Supervision of Task Specifications.
Roma Patel, Ellie Pavlick, Stefanie Tellex. RSS 2020.



Learning Visually Grounded Meaning Representations with Sketches
Roma Patel, Stephen Bach and Ellie Pavlick. How2 Workshop, ICML 2019.

Communication in multi-agent systems



Signalling (non-verbal)

Cooperative
game



State



Communication
protocol

Symbolic communication channel

Communication in multi-agent systems



Signalling (non-verbal)

Cooperative
game



State



Communication
protocol

Symbolic communication channel

Agents can pass messages to one another → meaning is attributed to messages
[have a communication policy along with action policy]

Communication in multi-agent systems



Signalling (non-verbal)

Cooperative
game

Zero-sum
games



State



Communication
protocol

Symbolic communication channel

Agents can pass messages to one another → meaning is attributed to messages

Communication in multi-agent systems



Signalling (non-verbal)



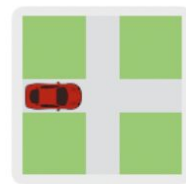
Symbolic communication channel

Having a communication channel can help agents' performance

Communication in multi-agent systems



Signalling (non-verbal)



State



Communication
protocol

Symbolic communication channel



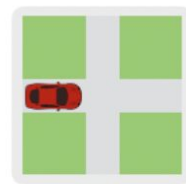
We cannot generalise to humans

Agents converge on a language that is different from human language. Human-agent communication channel does not exist!

Communication in multi-agent systems



Signalling (non-verbal)



State



Communication
protocol

Symbolic communication channel



*We cannot generalise to humans $\Rightarrow$
protocols are not human-interpretable*

Agent language is uninterpretable to
humans

Communication in multi-agent systems



Signalling (non-verbal)



State



Communication
protocol

Symbolic communication channel



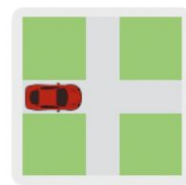
*We cannot generalise to ~~humans~~ any agent
not trained to converge to that protocol*

Agents trained in different pairs or
populations converge to different
languages

Communication in multi-agent systems



Signalling (non-verbal)



State



Communication
protocol

Symbolic communication channel

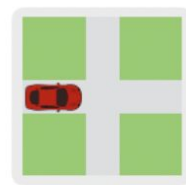
 *Protocols are hard to learn; don't have properties of human language (compositionality, Zipf's law)*

Language learned is less efficient or abstractable

Alleviate these concerns: use natural language



Signalling (non-verbal)



State



Communication
protocol

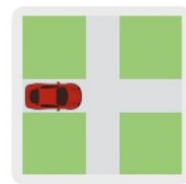
Symbolic communication channel

*Train agents to communicate in **natural language***

Alleviate these concerns: use natural language



Signalling (non-verbal)



State



Communication
protocol

Symbolic communication channel

*Train agents to communicate in **natural language***

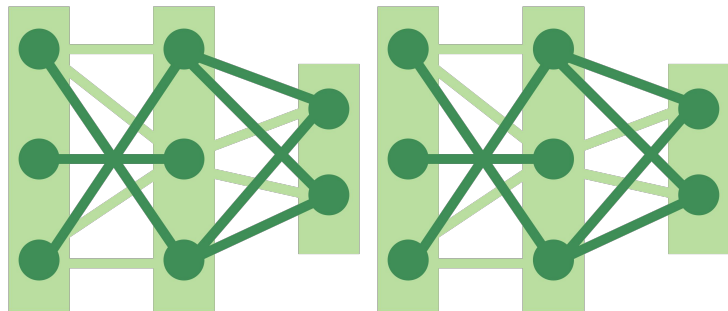
(because LLMs
work now!)

Towards natural language communication for MARL

*Train agents to communicate in **natural language***

Agents start from a vocabulary with **prior semantics attributed to symbols**, then allow **multi-agent self-play** to **shift the semantics**

Towards natural language communication for MARL

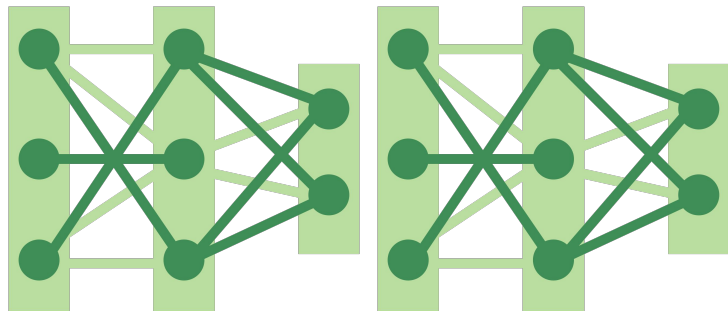


Pretrained language model
prior semantics & word meaning

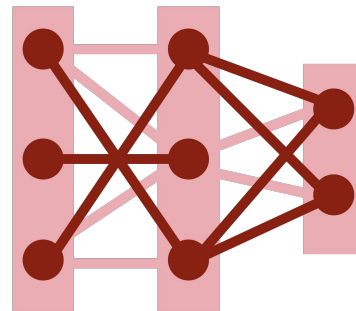
*Train agents to communicate in **natural language***

Agents start from a vocabulary with **prior semantics attributed to symbols**, then allow **multi-agent self-play** to **shift the semantics**

Towards natural language communication for MARL



Pretrained language model
prior semantics & word meaning

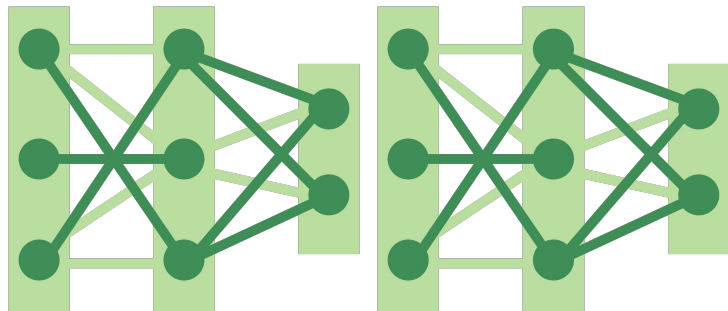


Fine-tune in environment
shift semantics to be task-specific

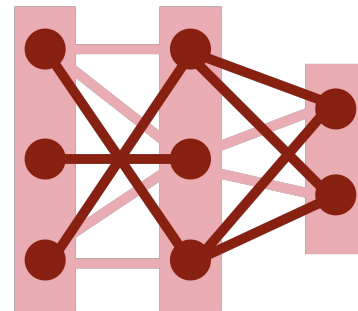
*Train agents to communicate in **natural language***

Agents start from a vocabulary with **prior semantics attributed to symbols**, then allow **multi-agent self-play** to **shift the semantics**

Why is this hard?



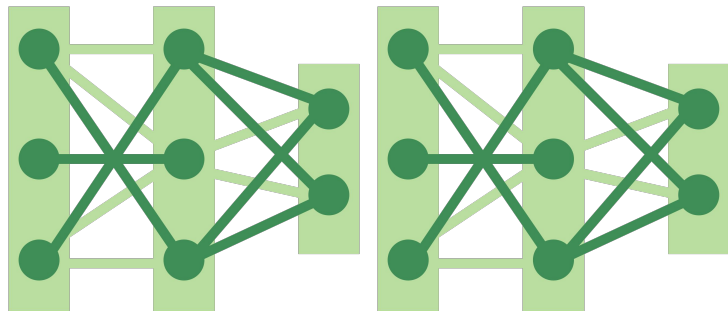
Pretrained language model



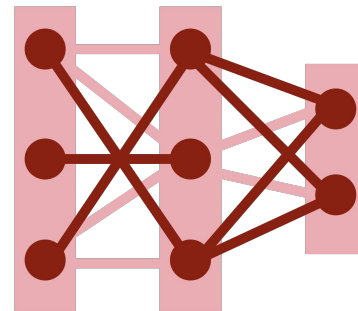
Fine-tune in environment

General-purpose training results in a good generative **model of language** and knowledge of **facts about the world** and **relations between words**

Adapt models to be task specific

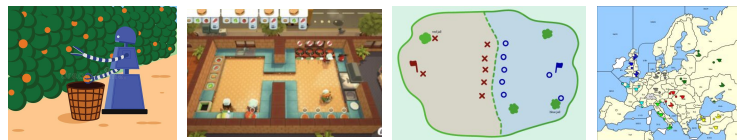


Pretrained language model



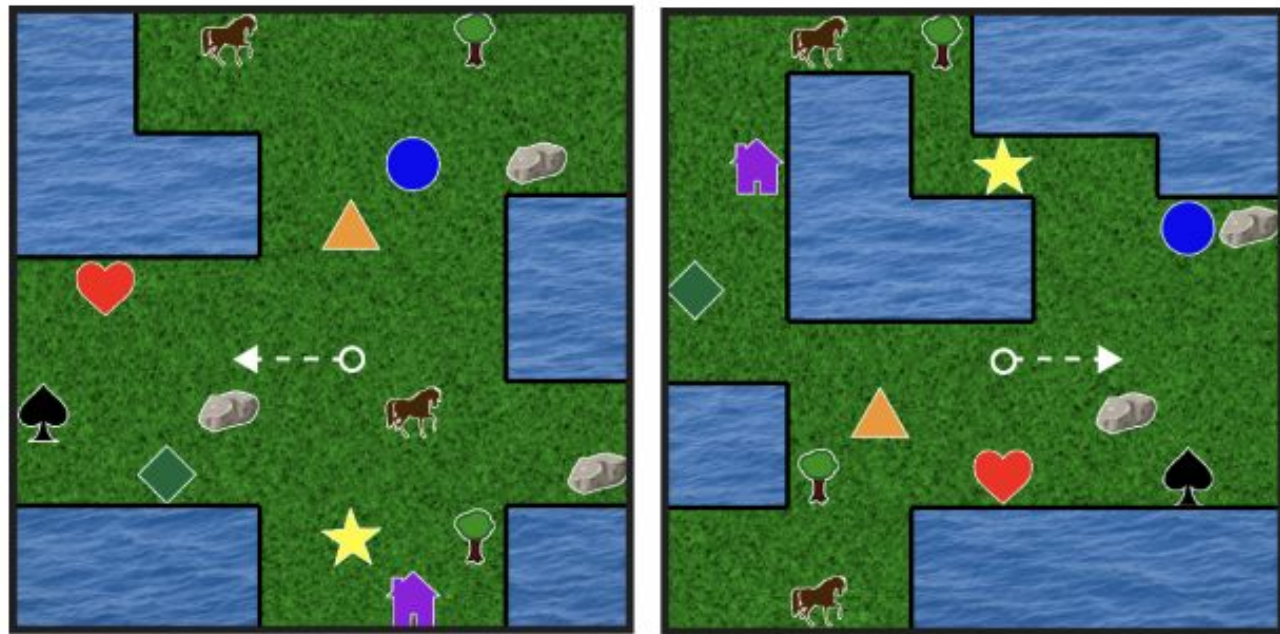
Fine-tune in environment

General-purpose training results in
a good generative **model of
language** and knowledge of **facts
about the world** and **relations
between words**



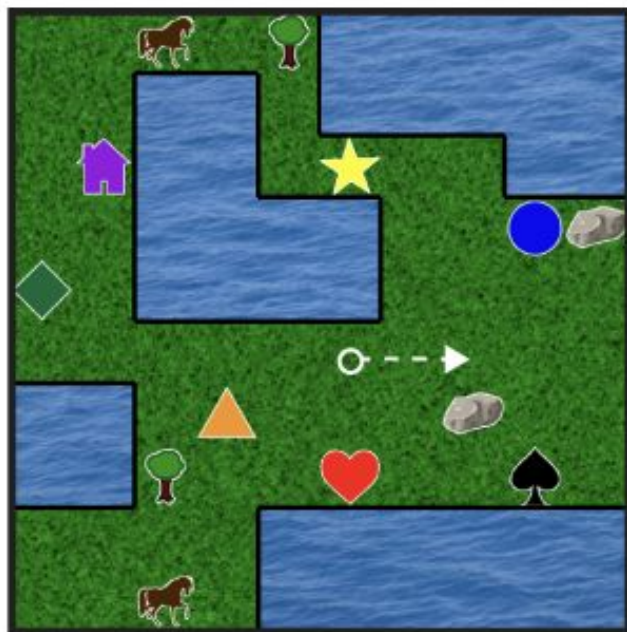
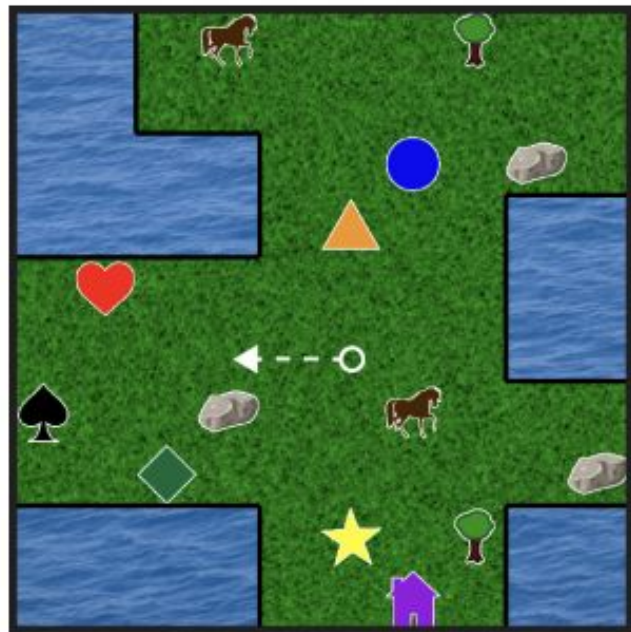
Task-specific RL fine-tuning can adapt
agents to cooperate in new
environments

Testbed: a complex, sequential environment



Reach the cell above the westernmost rock

Testbed: a complex, sequential environment

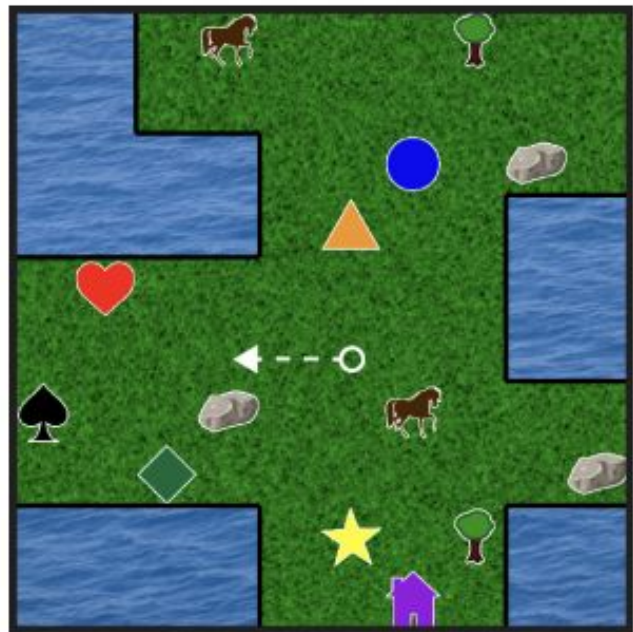


A **speaker agent** must learn to instruct a **listener agent** to cooperatively solve the task.

Speaker model is rewarded only if the listener model correctly reaches the goal.

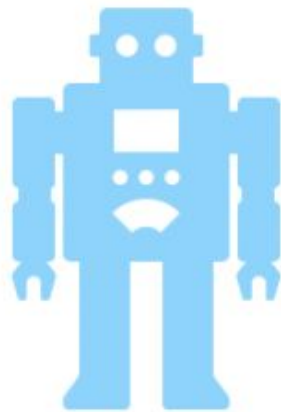
Reach the cell above the westernmost rock

Communicative task



2D environment: 2 **terrains**, many objects w properties differ in **shape** and **colour**

Listener agent



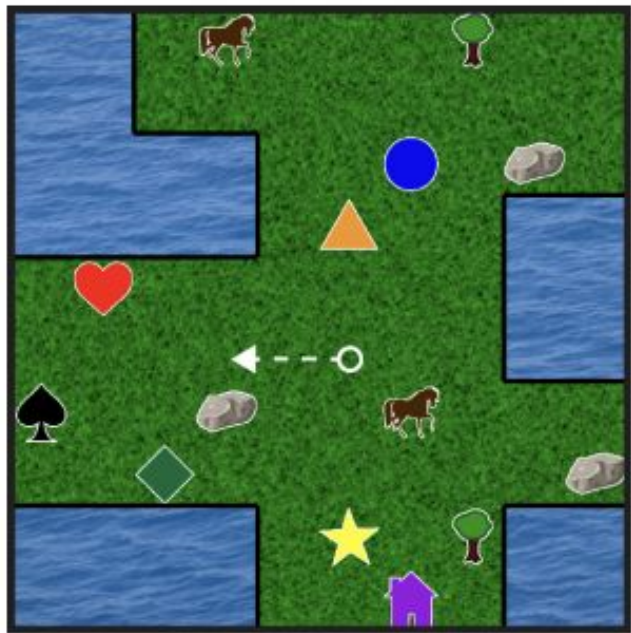
Take in natural language advice and learn to navigate to goal

Speaker agent



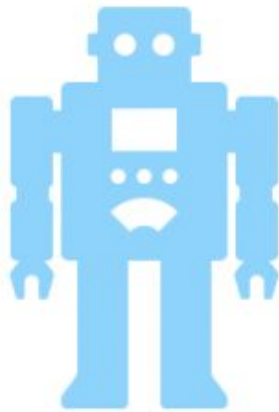
Take in state of the world and learn to generate instructions for listener

Communicative task



2D environment: 2 **terrains**, many objects w properties differ in **shape** and **colour**

Listener agent



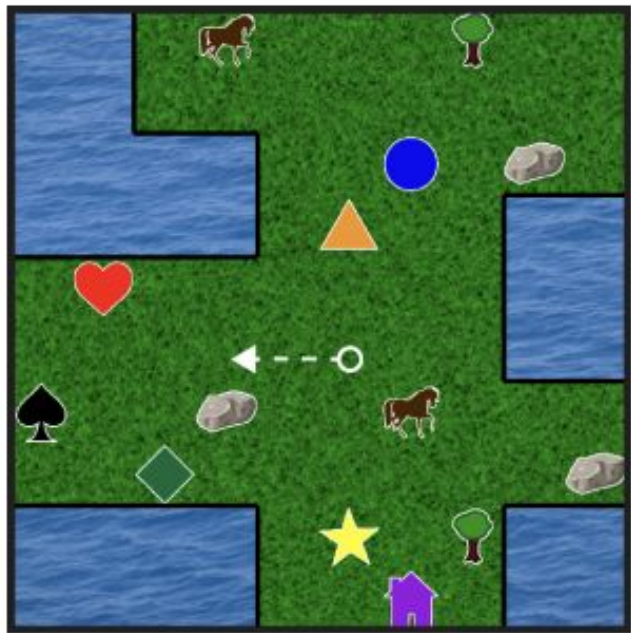
Add an **instruction encoder** into standard **A2C** and **VMPO** agent architectures. Agent can input observations + language and take actions.

Speaker agent



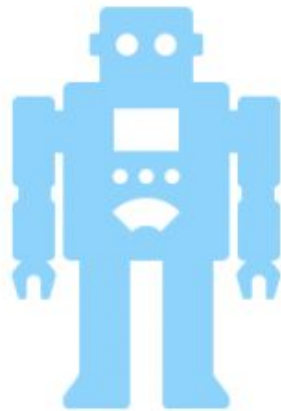
Add a **cooperative reward** mechanism into a standard **transformer** architecture. Agent can input observations and emit language.

Communicative task



2D environment: 2 **terrains**, many objects w properties differ in **shape** and **colour**

Listener agent



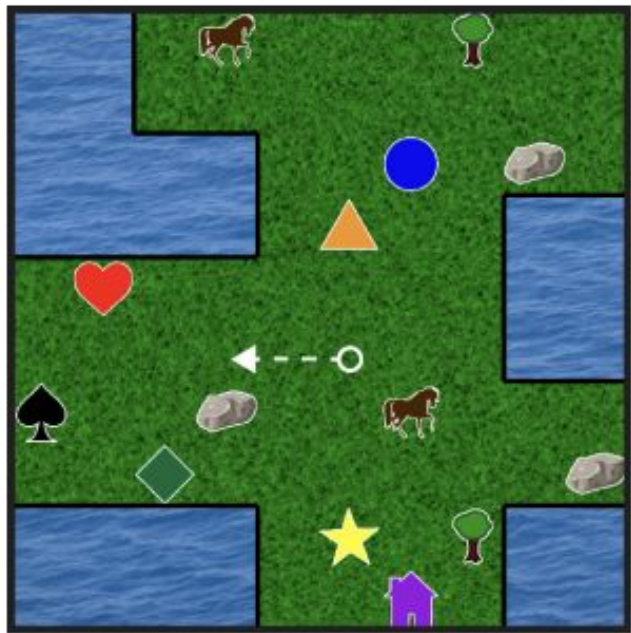
Add an **instruction encoder** into standard **A2C** and **VMPO** agent architectures. Agent can input observations + language and take actions

Speaker agent



Add a **cooperative reward** mechanism into a standard **transformer** architecture. Agent can input observations and emit language.

Communicative task



2D environment: 2 **terrains**, many objects w properties differ in **shape** and **colour**

Listener agent

$$\pi_{\theta_L}(a_t | s_t, m)$$

Estimate listener's action policy by taking in instruction and rewarding on reaching goal.

Add an **instruction encoder** into standard **A2C** and **VMPO** agent architectures. Agent can input observations + language and take actions.

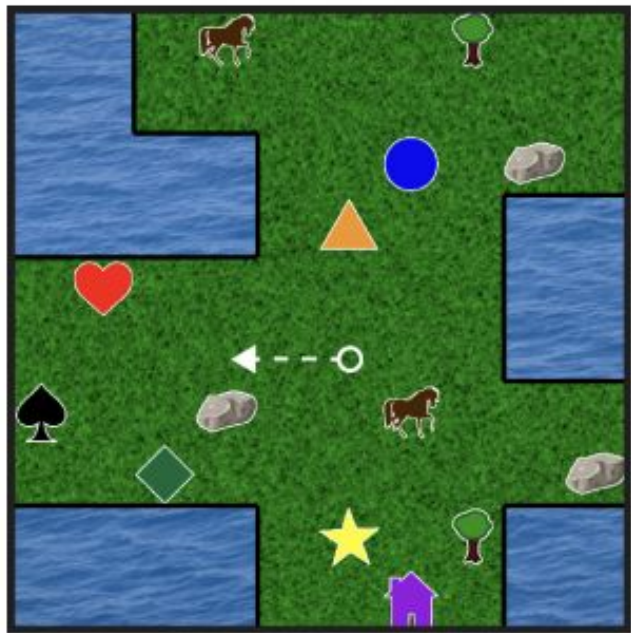
Speaker agent

$$\pi_{\theta_S}(m | s_0)$$

Estimate speaker's generative model by rewarding with listener's reward.

Add a **cooperative reward** mechanism into a standard **transformer** architecture. Agent can input observations and emit language.

Communicative task



2D environment: 2 **terrains**, many objects w properties differ in **shape** and **colour**

Listener agent

$$\pi_{\theta_L}(a_t | s_t, m)$$

Estimate listener's action policy by taking in instruction and rewarding on reaching goal.

Add an **instruction encoder** into standard **A2C** and **VMPO** agent architectures. Agent can input observations + language and take actions.

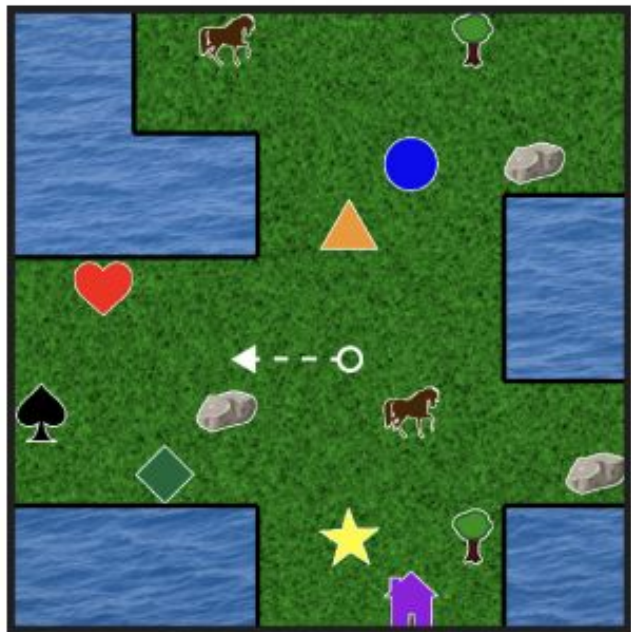
Speaker agent

$$\pi_{\theta_S}(m | s_0)$$

Estimate speaker's generative model by rewarding with listener's reward.

Add a **cooperative reward** mechanism into a standard **transformer** architecture. Agent can input observations and emit language.

Communicative task



2D environment: 2 **terrains**, many objects w properties differ in **shape** and **colour**

Listener agent

$$\pi_{\theta_L}(a_t | s_t, m)$$

Estimate listener's action policy by taking in instruction and rewarding on reaching goal.

Speaker agent

$$\pi_{\theta_S}(m | s_0)$$

Estimate speaker's generative model by rewarding with listener's reward.

Both agents need to cooperate and allow the other to perform well in order to succeed at their individual tasks

How do we do this correctly?

Synergies between learning from data and learning with self-play

When learning from data, a speaker agent learns a good model of language i.e., it maximises a reward that prioritises structure/grammar of language

How do we do this correctly?

Synergies between learning from data and learning with self-play

When learning from data, a speaker agent learns a good model of language i.e., it maximises a reward that prioritises structure/grammar of language

When learning with self-play, a speaker agent only needs to maximise cooperative reward i.e., it can relearn new semantics over the same symbols

How do we do this correctly?

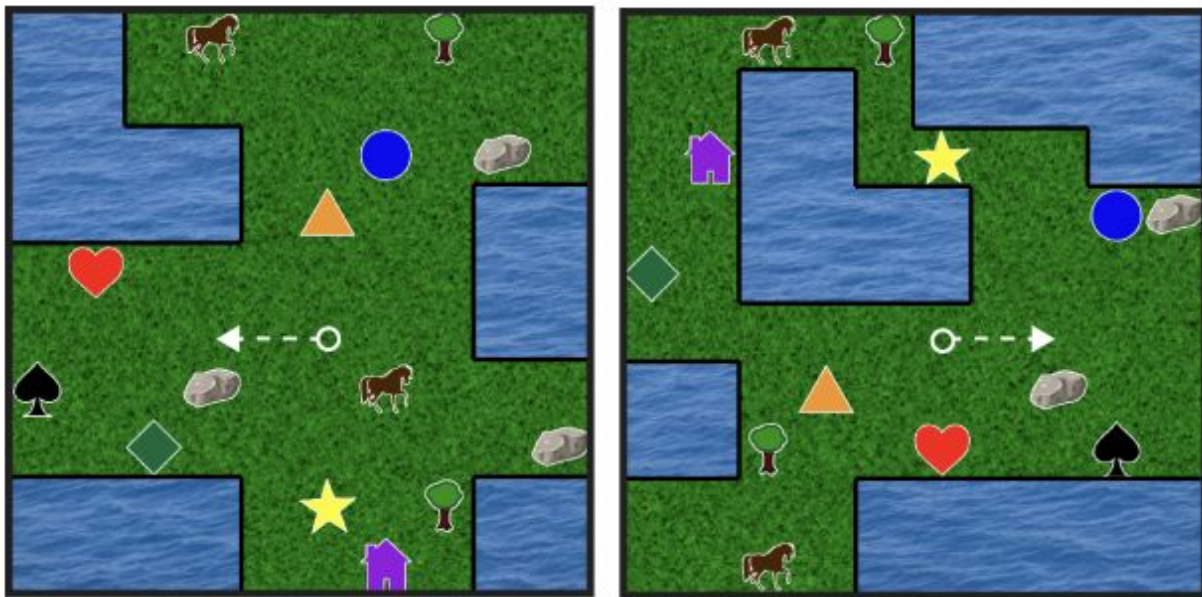
Synergies between learning from data and learning with self-play

When learning from data, a speaker agent learns a good model of language i.e., it maximises a reward that prioritises structure/grammar of language

When learning with self-play, a speaker agent only needs to maximise cooperative reward i.e., it can relearn new semantics over the same symbols

Question 1. How do we ensure the language learned during pre-training doesn't get distorted during self-play?

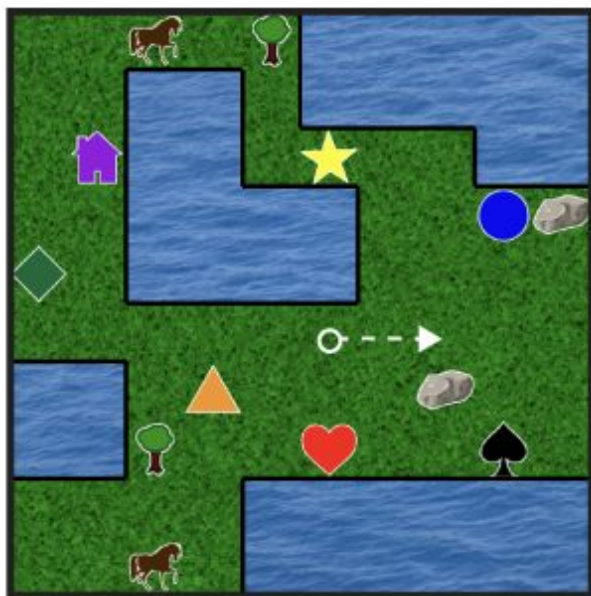
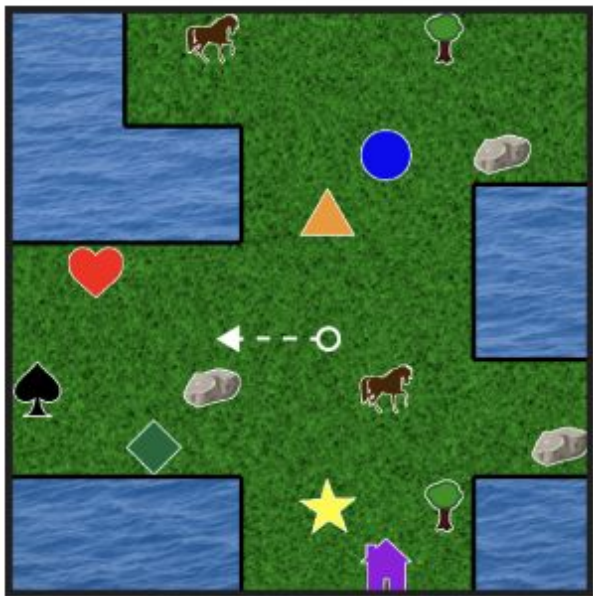
Using a small amount of in-domain human data



We can combine self-play with supervised learning in the speaker's loss function

Reach the cell above the westernmost rock

Using a small amount of in-domain human data

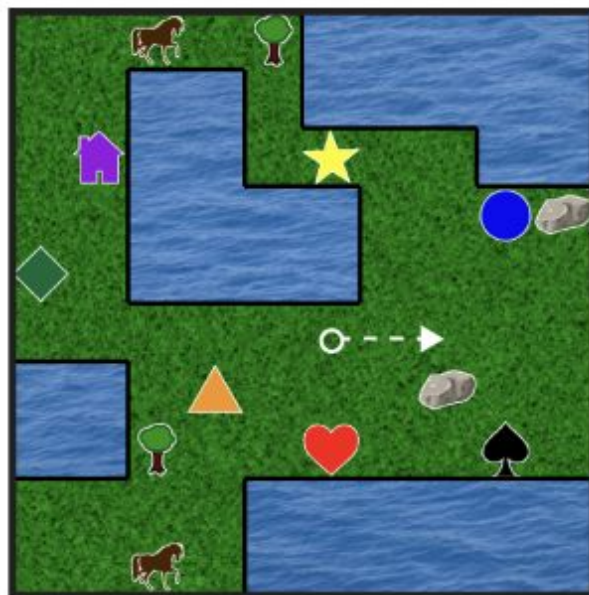
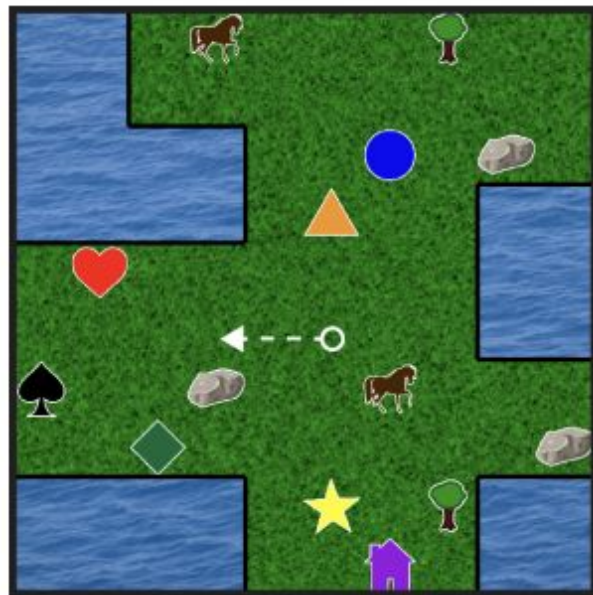


We can combine self-play with supervised learning in the speaker's loss function

First learn weights of the speaker model with a small amount of data, then optimise with cooperative self-play reward.

Reach the cell above the westernmost rock

Using a small amount of in-domain human data



We see that doing this can allow you to shift an LLMs distribution to a new task with a cooperative reward mechanism.

The speaker generates task-specific instructions 73% of the time (as opposed to out-of-domain natural language sentences)

Reach the cell above the westernmost rock

How do we do this correctly?

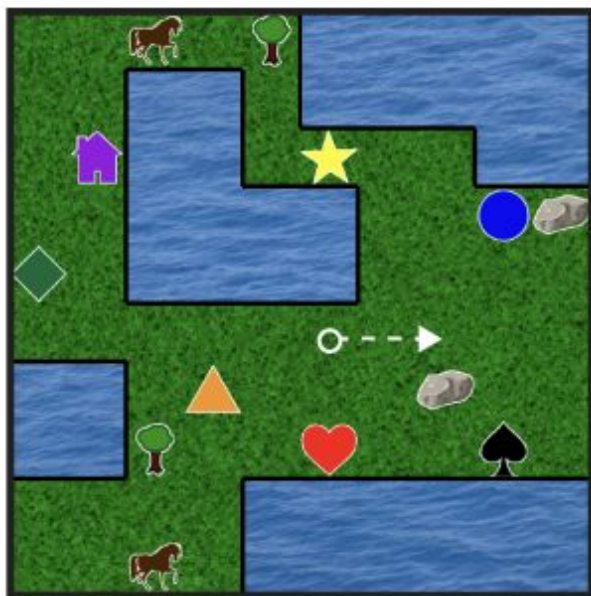
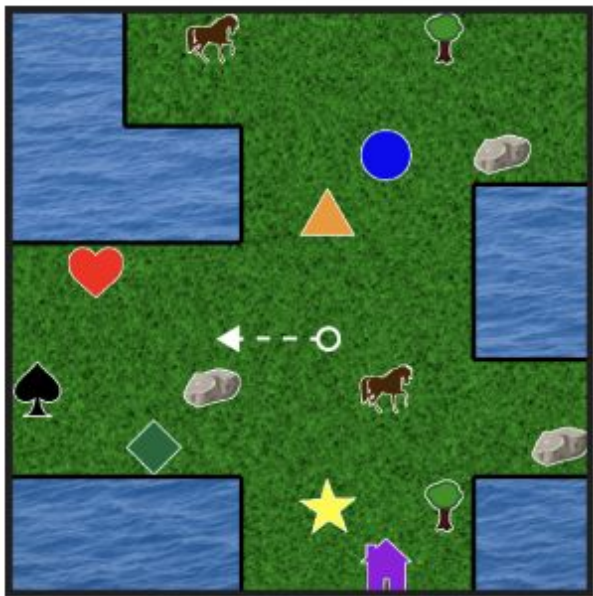
Synergies between learning from data and learning with self-play

When learning from data, a speaker agent learns a good model of language i.e., it maximises a reward that prioritises structure/grammar of language

When learning with self-play, a speaker agent only needs to maximise cooperative reward i.e., it can relearn new semantics over the same symbols

Question 2. What is the minimal amount of data we can use in the supervised phase to still prevent language drift during self-play?

Using a small amount of in-domain human data



We experiment with re-training the speaker with different amounts of held-out data (25%, 50%..)

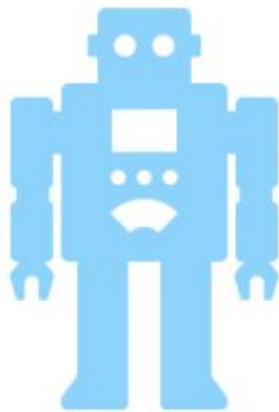
We see that 25% (~800) instructions are enough to achieve a good-enough listener task performance

Reach the cell above the westernmost rock

Towards natural language communication for MARL

Agents can fluently
communicate (with
any agent and any
human)

Listener agent



Add an **instruction encoder** into standard **A2C** and **VMPO** agent architectures. Agent can input observations + language and take actions.

Speaker agent



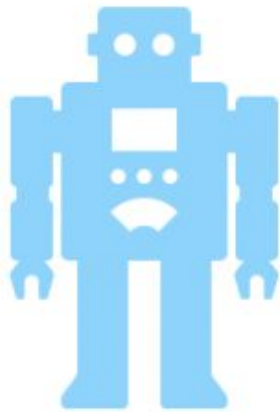
Add a **cooperative reward** mechanism into a standard **transformer** architecture. Agent can input observations and emit language.

Towards natural language communication for MARL

Agents can fluently
communicate (with
any agent and any
human)

Advice-giver can
instruct populations of
agents

Listener agent



Add an **instruction encoder** into standard **A2C** and **VMPO** agent architectures. Agent can input observations + language and take actions.

Speaker agent



Add a **cooperative reward** mechanism into a standard **transformer** architecture. Agent can input observations and emit language.

Multilingual instruction following

Multilingual instruction following

Speaker



P (instr | route)

Our starting point is in a living room, we're facing towards a long beige sofa, and in front of the sofa there are three glass coffee tables, turn around and exit through the doorway that's in front of you, walk pass the bed that's on your right and then turn left, we're now facing towards another living room, and on the left there's an open door, walk towards that open door enter the bathroom that's in front of you, turn towards the right into the shower area, and that's your destination.



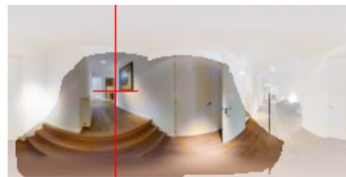
Room-across-Room

Follower

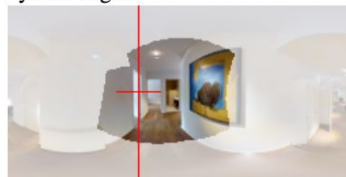


P (route | instr)

Guide Alignment



Now you are standing in-front of a closed door, turn to your left, you can see two wooden steps, climb the steps and walk forward by crossing a...

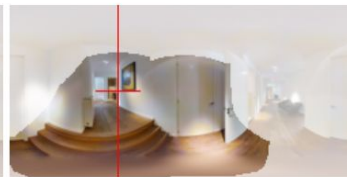


...crossing a wall painting which is to your right side, you can see open door enter...

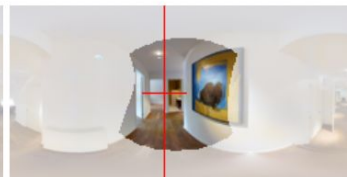


...enter into it. This is a gym room, move forward, walk...

Follower Alignment



Now you are standing in-front of a closed door, turn to your left, you can see two wooden steps, climb the steps and walk forward by...



...by crossing a wall painting which is to your right..



...right side, you can see open door enter into it. This is a gym room, move forward, walk

Multilingual instruction following

Speaker



P (instr | route)

Our starting point is in a living room, we're facing towards a long beige sofa, and in front of the sofa there are three glass coffee tables, turn around and exit through the doorway that's in front of you, walk pass the bed that's on your right and then turn left, we're now facing towards another living room, and on the left there's an open door, walk towards that open door enter the bathroom that's in front of you, turn towards the right into the shower area, and that's your destination.



Room-across-Room

आप सामने वाले छोटे भूरे रंग के अलमारी के ओर तक जाकर वहा से बाए ओर मुड़े और सामने वाले

ఇప్పుడు మీరు ఉన్న దేగ్గెర నుండి కొద్దిగా కుడి పంపు మీ కుడి వైపు మళ్ళి తెరిచి ఉన్న గాజు ద్వారము

Follower



P (route | instr)

Train agents to solve tasks given in languages apart from just English

Summary

Language for
communication



- Agents in MARL environments can be augmented with language understanding modules to take in advice
- Pretrained language models can be adapted to advise, instruct and communicate with teams of agents
- We can train agents in any language (not just English) and evaluate cultural evolution

Summary

Language for
communication



Understanding and
improving models



- Apart from using language to help multi-agent systems, we can also use multi-agent/game-theory techniques to improve language models!
- e.g., better vocabulary selection using game-theoretic power indices (Shapley, Banzhaff), better feature identification for mechanistic interpretability methods of LLMs

Thank you! Questions?

How can we use language for multi-agent systems?

Language for communication



Language as supervision



Understanding and improving models



RLang: A Declarative Language for Expressing Prior Knowledge about MDPs
Roma Patel*, Rafael Rodriguez-Sanchez*, Stefanie Tellex, George Konidaris.
ICML 2022 (in submission)



Mapping Language Models to Grounded Conceptual Spaces.
Roma Patel and Ellie Pavlick. ICLR 2022



Does linguistic bias affect generative language models?
Roma Patel and Ellie Pavlick. EMNLP 2021.



Game-theoretic Vocabulary Selection for Text Classification Tasks
Roma Patel, Marta Garnelo, Ian Gemp and Yoram Bachrach. NAACL 2021.



Room-Across-Room: Multilingual Vision-and Language Navigation with Dense Spatiotemporal Grounding.
Alexander Ku*, Peter Anderson*, **Roma Patel**, Eugene Ie, Jason Baldridge. EMNLP 2020.



On the Relationship Between Structure in Natural Language and Models of Sequential Decision Processes
Roma Patel*, Rafael Rodriguez-Sanchez*, George Konidaris. LAREL Workshop, ICML 2020.



Robot Object Retrieval with Contextual Natural Language Queries.
Thao Nguyen, Nakul Gopalan, **Roma Patel**, Matthew Corsaro, Ellie Pavlick, Stefanie Tellex. RSS 2020.



Planning with State Abstractions for Non-Markovian Task Specifications
Yoonseon Oh, **Roma Patel**, Thao Nguyen, Baichuan Huang, Ellie Pavlick, Stefanie Tellex. RSS 2019.



Grounding Language to Non-Markovian Tasks with No Supervision of Task Specifications.
Roma Patel, Ellie Pavlick, Stefanie Tellex. RSS 2020.



Learning Visually Grounded Meaning Representations with Sketches
Roma Patel, Stephen Bach and Ellie Pavlick. How2 Workshop, ICML 2019.

How can we use language for multi-agent systems?

Language for communication



Language as supervision



Understanding and improving models



RLang: A Declarative Language for Expressing Prior Knowledge about MDPs
Roma Patel*, Rafael Rodriguez-Sanchez*, Stefanie Tellex, George Konidaris.
ICML 2022 (in submission)



Mapping Language Models to Grounded Conceptual Spaces.
Roma Patel and Ellie Pavlick. ICLR 2022



Does linguistic bias affect generative language models?
Roma Patel and Ellie Pavlick. EMNLP 2021.



Game-theoretic Vocabulary Selection for Text Classification Tasks
Roma Patel, Marta Garnelo, Ian Gemp and Yoram Bachrach. NAACL 2021.



Room-Across-Room: Multilingual Vision-and Language Navigation with Dense Spatiotemporal Grounding.
Alexander Ku*, Peter Anderson*, **Roma Patel**, Eugene Ie, Jason Baldridge. EMNLP 2020.



On the Relationship Between Structure in Natural Language and Models of Sequential Decision Processes
Roma Patel*, Rafael Rodriguez-Sanchez*, George Konidaris. LAREL Workshop, ICML 2020.



Robot Object Retrieval with Contextual Natural Language Queries.
Thao Nguyen, Nakul Gopalan, **Roma Patel**, Matthew Corsaro, Ellie Pavlick, Stefanie Tellex. RSS 2020.



Planning with State Abstractions for Non-Markovian Task Specifications
Yoonseon Oh, **Roma Patel**, Thao Nguyen, Baichuan Huang, Ellie Pavlick, Stefanie Tellex. RSS 2019.



Grounding Language to Non-Markovian Tasks with No Supervision of Task Specifications.
Roma Patel, Ellie Pavlick, Stefanie Tellex. RSS 2020.



Learning Visually Grounded Meaning Representations with Sketches
Roma Patel, Stephen Bach and Ellie Pavlick. How2 Workshop, ICML 2019.

Thank you! Questions?

Summary

Understanding and improving models



- Game theory can help understand and improve general machine learning models and algorithms
- Tools for understanding agent's behaviour are the key to building intelligent and interpretable agents

How can we use language for multi-agent systems?

Language for communication



Language as supervision



Understanding and improving models



Room-Across-Room: Multilingual Vision-and Language Navigation with Dense Spatiotemporal Grounding.

Alexander Ku*, Peter Anderson*, **Roma Patel**, Eugene Ie, Jason Baldridge. EMNLP 2020.



Mapping Language Models to Grounded Conceptual Spaces.

Roma Patel and Ellie Pavlick. ICLR 2022



Does linguistic bias affect generative language models?

Roma Patel and Ellie Pavlick. EMNLP 2021.



Game-theoretic Vocabulary Selection for Text Classification Tasks

Roma Patel, Marta Garnelo, Ian Gemp and Yoram Bachrach. NAACL 2021.



RLang: A Declarative Language for Expressing Prior Knowledge about MDPs

Rafael Rodriguez-Sanchez*, **Roma Patel***, Stefanie Tellex, George Konidaris. ICML 2022 (in submission)



On the Relationship Between Structure in Natural Language and Models of Sequential Decision Processes

Rafael Rodriguez-Sanchez*, **Roma Patel**, George Konidaris. LAREL Workshop, ICML 2020.



Robot Object Retrieval with Contextual Natural Language Queries.

Thao Nguyen, Nakul Gopalan, **Roma Patel**, Matthew Corsaro, Ellie Pavlick, Stefanie Tellex. RSS 2020.



Planning with State Abstractions for Non-Markovian Task Specifications

Yoonseon Oh, **Roma Patel**, Thao Nguyen, Baichuan Huang, Ellie Pavlick, Stefanie Tellex. RSS 2019.



Grounding Language to Non-Markovian Tasks with No Supervision of Task Specifications.

Roma Patel, Ellie Pavlick, Stefanie Tellex. RSS 2020.



Learning Visually Grounded Meaning Representations with Sketches

Roma Patel, Stephen Bach and Ellie Pavlick. How2 Workshop, ICML 2019.

What next?

Language for communication



Language as supervision



Understanding and improving models



Room-Across-Room: Multilingual Vision-and Language Navigation with Dense Spatiotemporal Grounding.

Alexander Ku*, Peter Anderson*, **Roma Patel**, Eugene Ie, Jason Baldridge. EMNLP 2020.



Mapping Language Models to Grounded Conceptual Spaces.

Roma Patel and Ellie Pavlick. ICLR 2022



Does linguistic bias affect generative language models?

Roma Patel and Ellie Pavlick. EMNLP 2021.



Game-theoretic Vocabulary Selection for Text Classification Tasks

Roma Patel, Marta Garnelo, Ian Gemp and Yoram Bachrach. NAACL 2021.



RLang: A Declarative Language for Expressing Prior Knowledge about MDPs

Rafael Rodriguez-Sanchez*, **Roma Patel***, Stefanie Tellex, George Konidaris. ICML 2022 (in submission)



On the Relationship Between Structure in Natural Language and Models of Sequential Decision Processes

Rafael Rodriguez-Sanchez*, **Roma Patel**, George Konidaris. LAREL Workshop, ICM 2020.



Robot Object Retrieval with Contextual Natural Language Queries.

Thao Nguyen, Nakul Gopalan, **Roma Patel**, Matthew Corsaro, Ellie Pavlick, Stefanie Tellex. RSS 2020.



Planning with State Abstractions for Non-Markovian Task Specifications

Yoonseon Oh, **Roma Patel**, Thao Nguyen, Baichuan Huang, Ellie Pavlick, Stefanie Tellex. RSS 2019.



Grounding Language to Non-Markovian Tasks with No Supervision of Task Specifications.

Roma Patel, Ellie Pavlick, Stefanie Tellex. RSS 2020.



Learning Visually Grounded Meaning Representations with Sketches

Roma Patel, Stephen Bach and Ellie Pavlick. How2 Workshop, ICML 2019.

Incorporating advice into multi-agent team players



Wolves can use sights and sounds to effectively coordinate to attack prey



Does this suffice when we increase the level of sophistication of tasks?

Real-world tasks could benefit from real-world knowledge

Real-world tasks could benefit from real-world knowledge

If agents are equipped
with a language
understanding mechanism
(joint policy)

Real-world tasks could benefit from real-world knowledge

If agents are equipped
with a language
understanding mechanism
(joint policy)



Understand instructions
from humans

Real-world tasks could benefit from real-world knowledge

If agents are equipped
with a language
understanding mechanism
(joint policy)

Understand and collect
language from the world!

Understand instructions
from humans



Real-world tasks could benefit from real-world knowledge

If agents are equipped
with a language
understanding mechanism
(joint policy)

Understand and collect
language from the world!



Manuals

Understand instructions
from humans

Real-world tasks could benefit from real-world knowledge

If agents are equipped
with a language
understanding mechanism
(joint policy)

Understand and collect
language from the world!

Understand instructions
from humans



Manuals



A black and white dog carries
a tennis ball in its mouth

Image captioning
mechanisms

Real-world tasks could benefit from real-world knowledge

If agents are equipped
with a language
understanding mechanism
(joint policy)

Understand and collect
language from the world!

Understand instructions
from humans

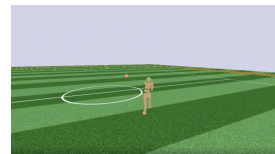


Manuals



A black and white dog carries
a tennis ball in its mouth

Image captioning
mechanisms



State information

Real-world tasks could benefit from real-world knowledge

If agents are equipped
with a language
understanding mechanism
(joint policy)

Understand and collect
language from the world!

Understand instructions
from humans

Generalise to completely
new scenarios with shared
language structure

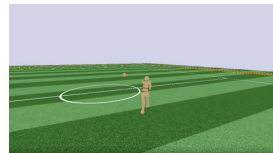


Manuals



A black and white dog carries
a tennis ball in its mouth

Image captioning
mechanisms



State information

Complex environments & interpreting & evaluating agents

Building environments that
reflect complex real-world
scenarios

Interpretability mechanisms to
understand agents' behaviour

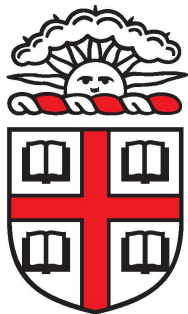


Thank you!

Ellie Pavlick, Stefanie Tellex,
George Konidakis, Michael
Littman, Amy Greenwald,
Rafael Rodriguez-Sanchez

Jason Baldridge, Peter
Anderson, Alexander Ku

Yoram Bachrach, Ian Gemp,
Angeliki Lazaridou, Chris Dyer,
Laura Weidinger



BROWN

Google Research

