CMPT 983

Grounded Natural Language Understanding

April 01, 2021

Interactive language learning

What is interactive language learning?

How do people learn language?

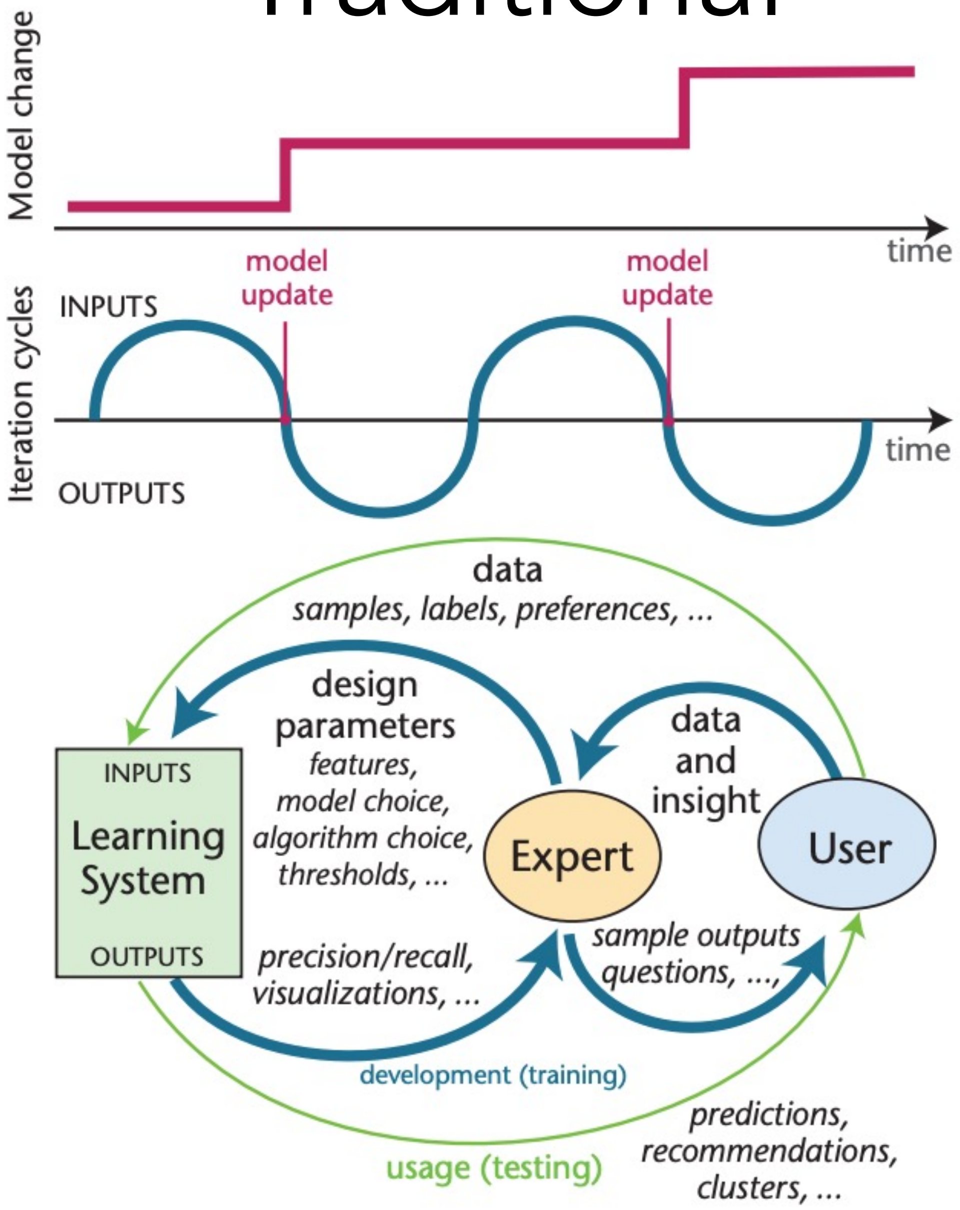
- Not just with static training pairs

- By interaction, through others

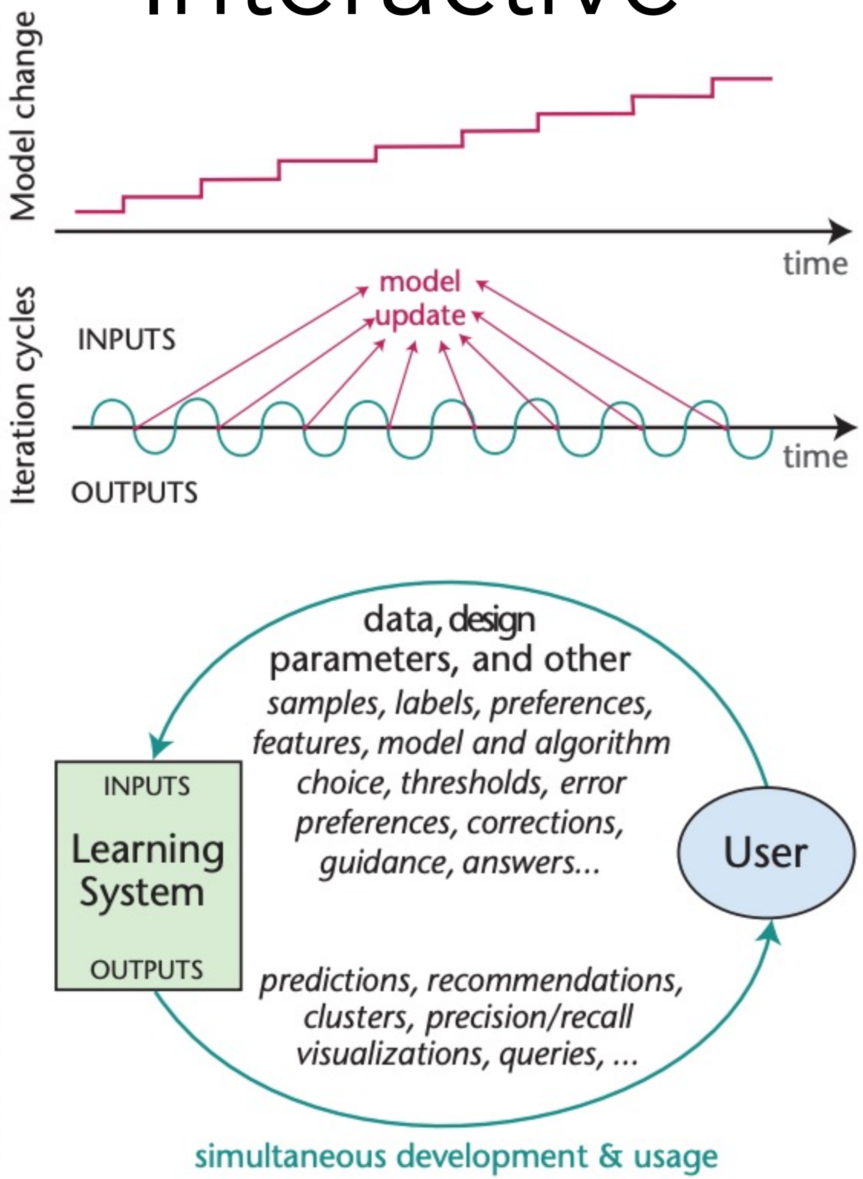
What is interactive machine learning?

- People provide feedback to the computer
- Also known as ``Human-in-the-loop''
- The world is constantly changing, learned models also need to evolve.

Traditional



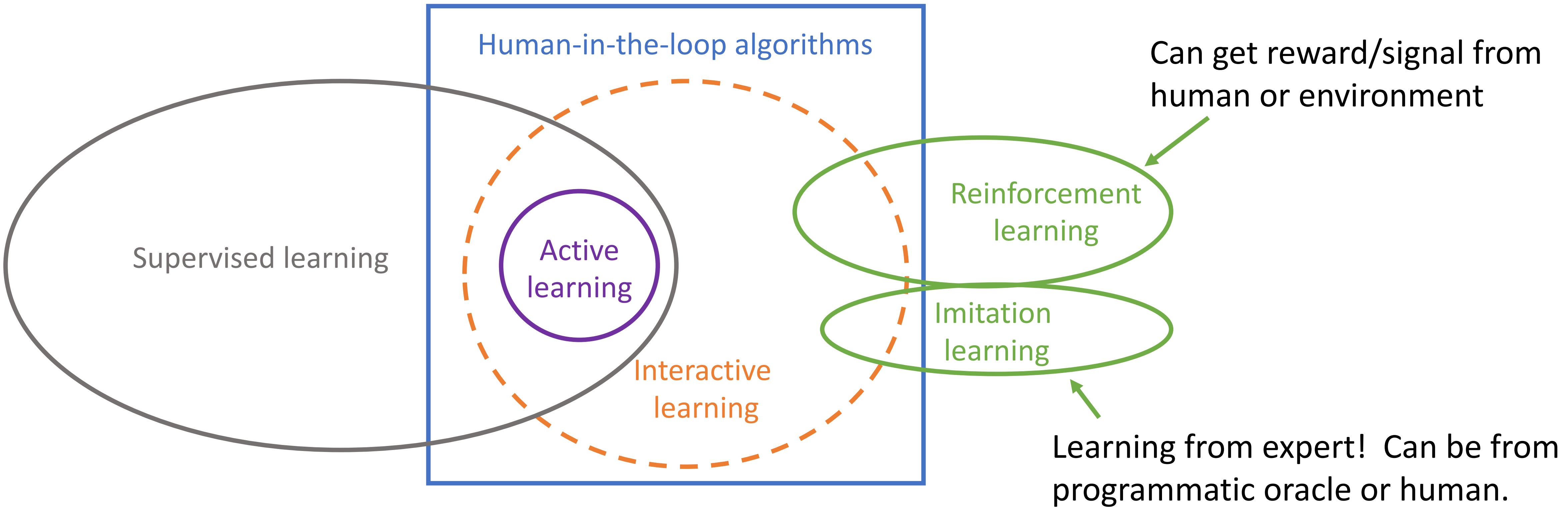
Interactive



Power to the People: The Role of Humans in Interactive Machine Learning, Amershi et al, AI Magazine, 2014

ML taxonomy

- Active learning: identify set of examples that should be labeled
  - Ideal setting: Interactively query user for labels
  - Often in papers: start with fully labeled set, assume that labels are not known for a part of the data, and then study what strategy to select a set to label will be best



https://www.trivedigaurav.com/blog/on-interactive-machine-learning/

What is interactive language learning?

Agent (model) learns language interactively either through

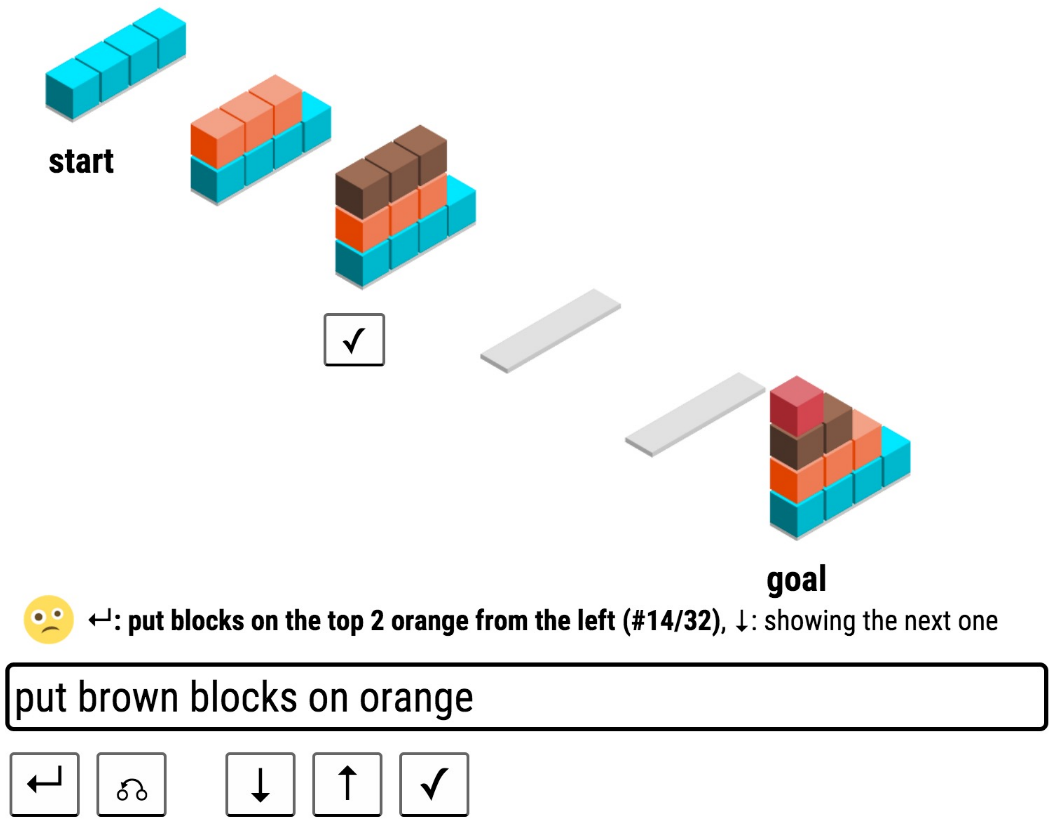
- Human feedback (like in interactive machine learning)
- Environment feedback (more traditionally known as grounded/situated language learning)

Commonality:

- interactive feedback indicating if an action or response is correct
- updating of model (weights) based on feedback
- ideally happens in real time, but practically there is still the train, test, deploy cycle

Interactive language learning (with human feedback)

Learning language through interaction



Human
- Has a goal, cannot perform action
- Can use language, provide feedback

Computer
- Does not know goal, can perform action
- Does not understand language

Game:
- Computer predicts an action (initially just random)
- Human provide feedback on correct or not

Can we teach the computer to understand language through interactions?

Learning Language Games through Interaction
http://shrdlurn.sidaw.xyz/
Wang et al, ACL, 2016

Learning language through interaction

Model

Log-linear model with L1 loss, optimized using AdaGrad

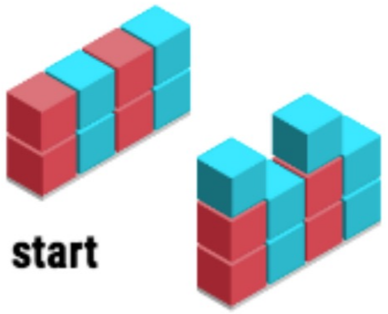
$$p_\theta(z \mid x) \propto \exp(\phi(x, z) \cdot \theta)$$

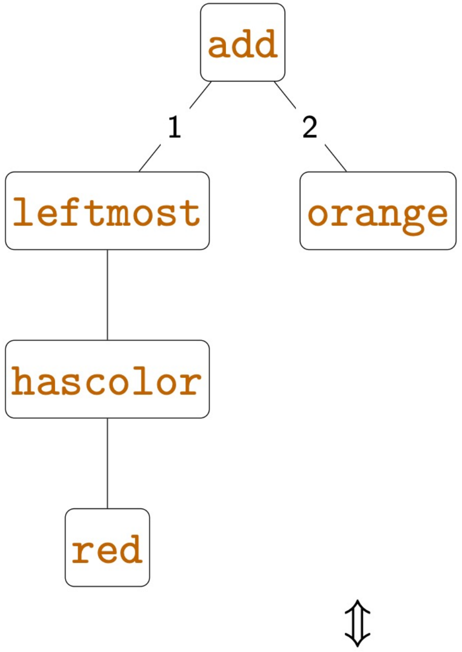
$$p_\theta(y \mid x) = \sum_{z:\mathrm{Exec}(z)=y} p_\theta(z \mid x)$$

x : *add a cyan block to red blocks*

z : add(hascolor(red), cyan)

y :





put orange on the very left red block

uni-, bi-, skip- grams

put, orange, on, the

put orange, orange on, ...,

*put * on, orange * the, ...,*

tree-grams

`add(leftmost(*), orange)`

`leftmost(hascolor(*))`

$\lambda c.$`(hascolor(c))`

cross product features

(*put*,`add(*,*)`)

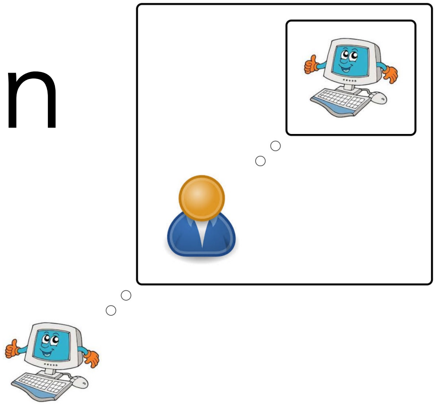
(*put orange*,`add(*,orange)`)

(*put*,`orange`)

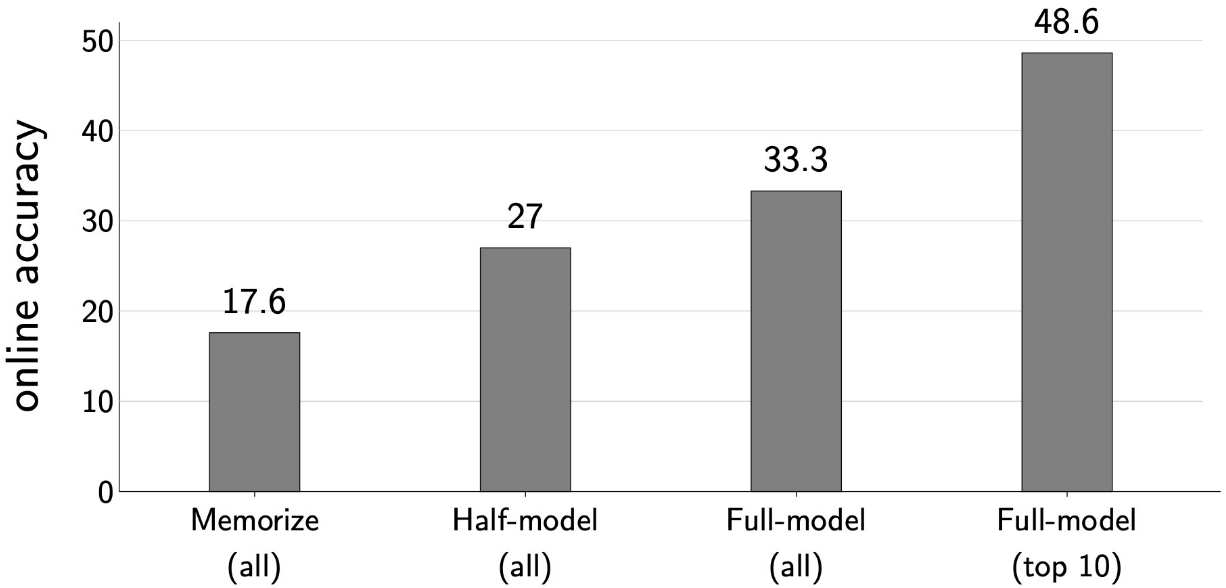
http://shrdlurn.sidaw.xyz/, Wang et al, ACL, 2016

Learning language through interaction

Results

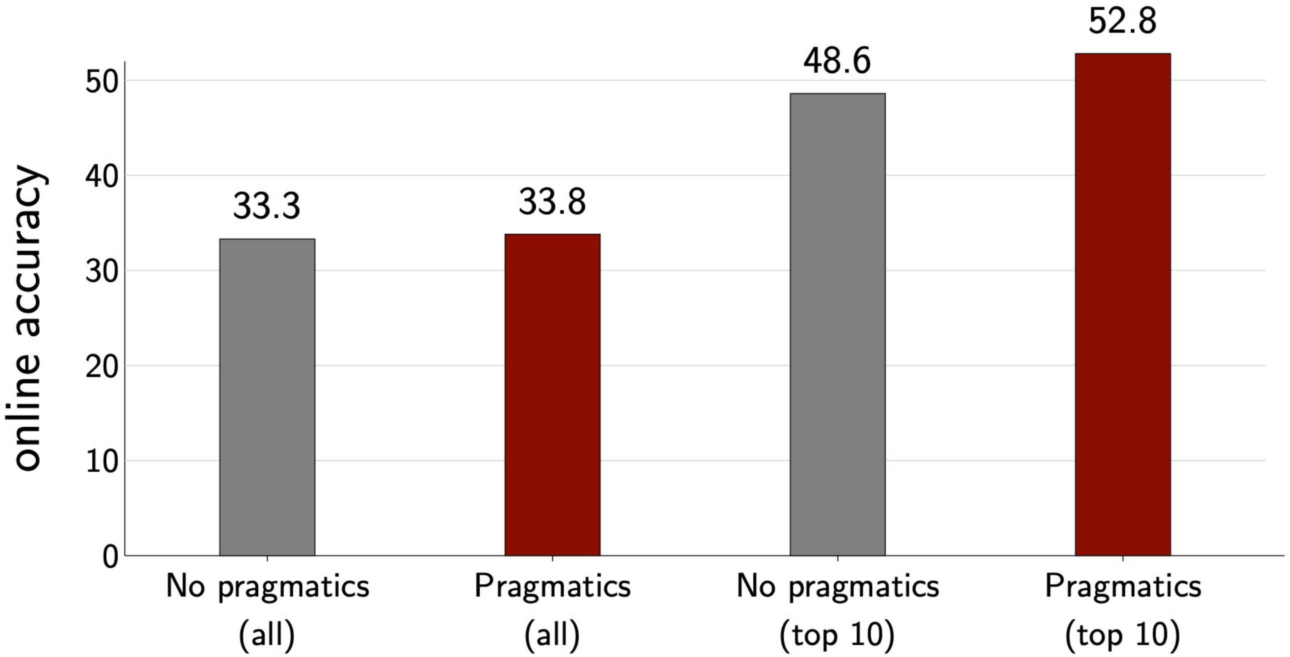


Non-pragmatic model



Learning works fairly well, especially for top players

Pragmatic model



pragmatics helps top (cooperative, rational) players

http://shrdlurn.sidaw.xyz/, Wang et al, ACL, 2016

Results: top players (rank 1-20)

precise and consistent:

(3.01)

rem cy pos 1
stack or blk pos 4
rem blk pos 2 thru 5
rem blk pos 2 thru 4
stack bn blk pos 1 thru 2
fill bn blk
stack or blk pos 2 thru 6
rem cy blk pos 2 fill rd blk

(2.72)

Remove the center block
Remove the red block
Remove all red blocks
Remove the first orange block
Put a brown block on the first brown block
Add blue block on first blue block

(2.78)

remove the brown block
remove all orange blocks
put brown block on orange blocks
put orange blocks on all blocks
put blue block on leftmost blue block in top row

http://shrdlurn.sidaw.xyz/, Wang et al, ACL, 2016

Results: average players (rank 21-50)

inconsistent or mismatches computer capability:

(9.17)

reinsert pink
take brown
put in pink
remove two pink from second layer
Add two red to second layer in odd intervals
Add five pink to second layer
Remove one blue and one brown from bottom layer

(7.18)

move second cube
double red with blue
double first red with red
triple second and fourth with orange
add red
remove orange on row two
add blue to column two
add brown on first and third

(8.37)

remove red
remove 1 red
remove 2 4 orange
add 2 red
add 1 2 3 4 blue
emove 1 3 5 orange
add 2 4 orange
add 2 orange
remove 2 3 brown
add 1 2 3 4 5 red
remove 2 3 4 5 6
remove 2
add 1 2 3 4 6 red

http://shrdlurn.sidaw.xyz/, Wang et al, ACL, 2016

Results: worst players (rank 51-100)

spammy, vague, did not tokenize:

(12.6)

'add red cubes on center left
center right
far left and far right'
'remove blue blocks on row two column two
row two column four'
remove red blocks in center left and center right on second row

(14.32)

laugh with me
red blocks with one aqua
aqua red alternate
brown red red orange aqua orange
red brown red brown red brown
space red orange red
second level red space red space red space

(14.15)

holdleftmost
holdbrown
holdleftmost
blueonblue
brownonblue1
blueonorange
holdblue
holdorange2
blueonred2
holdends1
holdrightend
hold2
orangeonorangerightmost

http://shrdlurn.sidaw.xyz/, Wang et al, ACL, 2016

Results: interesting players

(Polish)

usuń brązowe klocki
usuń niebieski klocek
usuń pomarańczowe klocki
usuń czerwony klocek
postaw brązowy klocek na pierwszym klocku
postaw czerwony klocek na pierwszym klocku
postaw pomarańczowe klocki na brązowych
postaw czerwone klocki
usuń ostatni brązowy klocek
usuń wszystkie klocki oprócz ostatniego
postaw niebieski klocek na czerwonym
postaw brązowy klocek na pierwszym klocku

(Polish notation)

rm scat + 1 c
+ 1 c
rm sh
+ 1 2 4 sh
+ 1 c
- 4 o
rm 1 r
+ 1 3 o
full fill c
rm o
full fill sh
- 1 3
full fill sh
rm sh
rm r
+ 2 3 r
rm o
+ 3 sh
+ 2 3 sh

http://shrdlurn.sidaw.xyz/, Wang et al, ACL, 2016

- Data from June 2016 - May 2017
  - 26k+ labeled examples, 1599 games

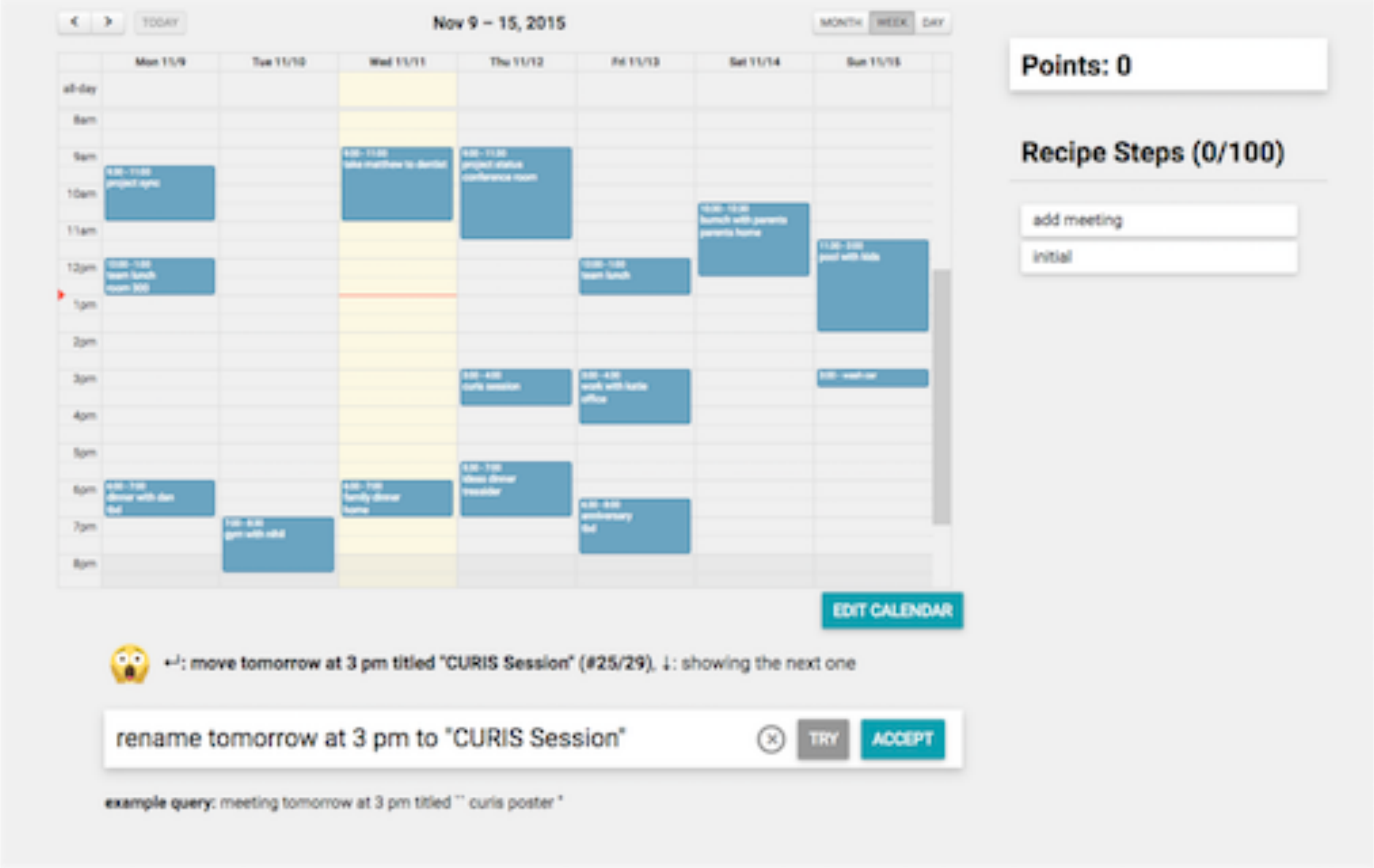
(NLPers?)

add brown on the top unless the rightmost
not(red)
pick up blue blocks
+ 1 2 3 4 5 r
Not the brown block!
The orange block!
છોડો વાદળી 0 1
બધા વાદળી દૂર
છોડો નારંગી 1 4
add blo 1 bro
rem ora blo
add blo 6 pin
add blo 134 bl
去掉最后一个块
在蓝色块上面加一层橙色块
smaz 1 a 3 jednou
retire les blocs bleus

(NLPers?)

move all blocks but middle
- 1 br - 4 br - 6 br
一番奥にオレンジを置く
一番右の赤を消す
add red one on the first
lift 1 3 5
add one orange block on top of each orange
去掉 蓝色 方块
smaz 1 a 2 a 3 a 5
quita el bloque marrón
quita el primer bloque por la derecha
drop orange not left not right
add brown on all blue in line 2 in line 3
Add x x o x o x red block
只保留桔黄色的方块
quitar cubo rojo
quitar ultimo cubo rojo

Can be adapted to real world applications

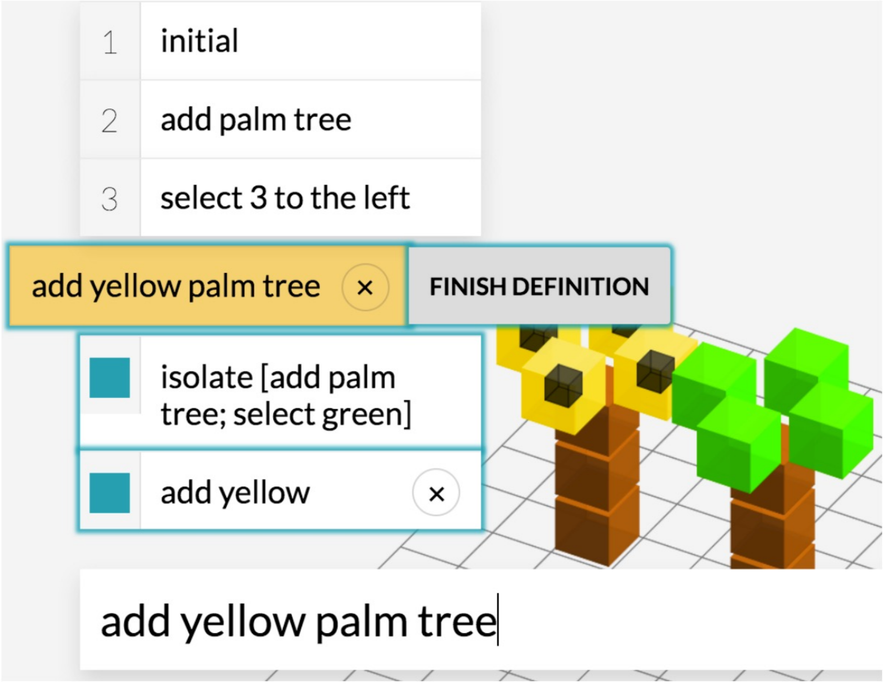


https://nlp.stanford.edu/blog/interactive-language-learning/

Follow-up work

- Let users write programs using natural language

- define new things in terms of what's already defined
- trace back to the core language



add palm tree:

add brown trunk height 3:

add brown top 3 times:

repeat 3 [add brown top]

go to top:

select very top of all

add leaves here:

select left or right or front or back; add green

Naturalizing a programming language via interactive learning
https://arxiv.org/pdf/1704.06956.pdf
Wang et al, ACL, 2017

Interactive language acquisition

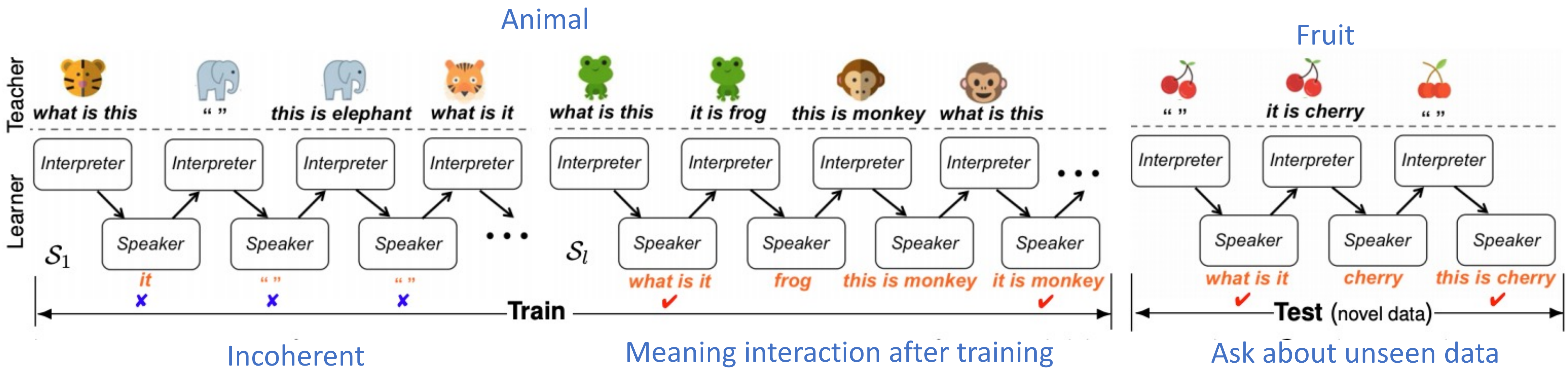
- Learn about new objects with
  - a conversational game between teacher and learner
- Learner needs to:
  - Extract + remember important information (Interpreter)
  - Ask question (Speaker)
  - Name novel objects (Speaker)

Interactive Language Acquisition with One-shot Visual Concept Learning through a Conversational Game
https://arxiv.org/pdf/1805.00462.pdf
Zhang et al, ACL, 2018

Interactive language acquisition

Teacher (select random object for session):
- Ask question
- Answer question / Make statement
- Say nothing

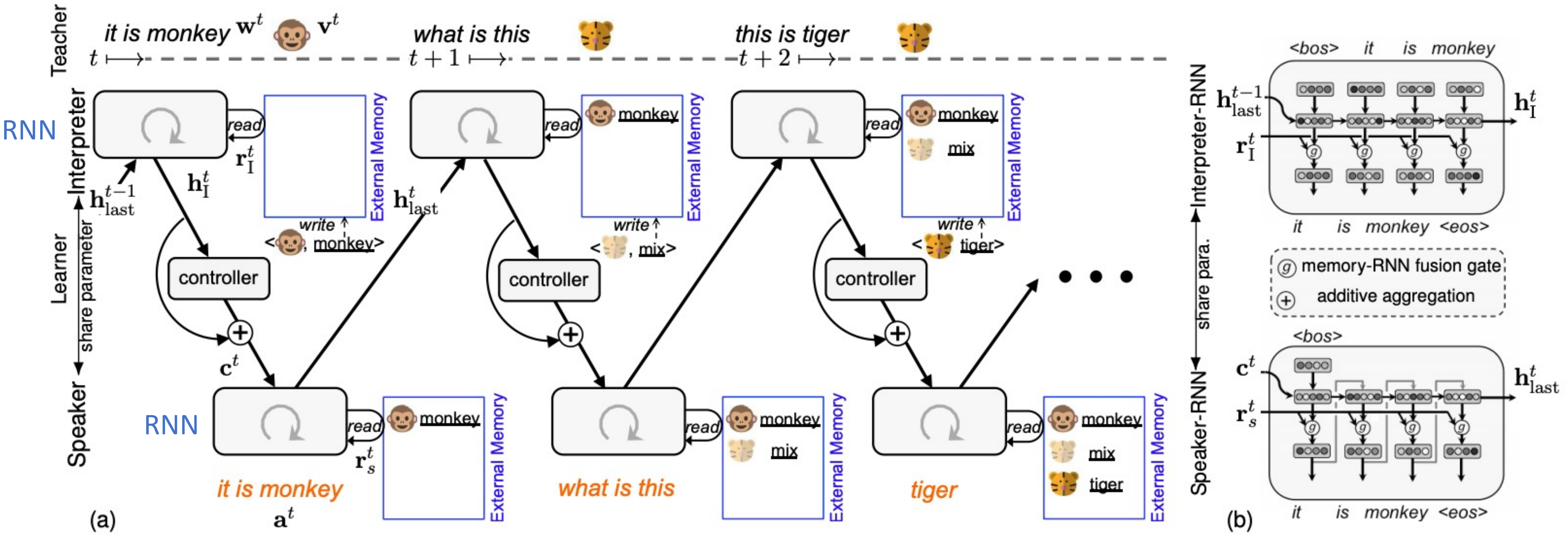
Learner is rewarded for
- Asking question: +0.1
- Providing correct answer: +1 (session terminates)
- Incorrect answer: -1



https://arxiv.org/pdf/1805.00462.pdf, Zhang et al, ACL, 2018

Interactive language acquisition

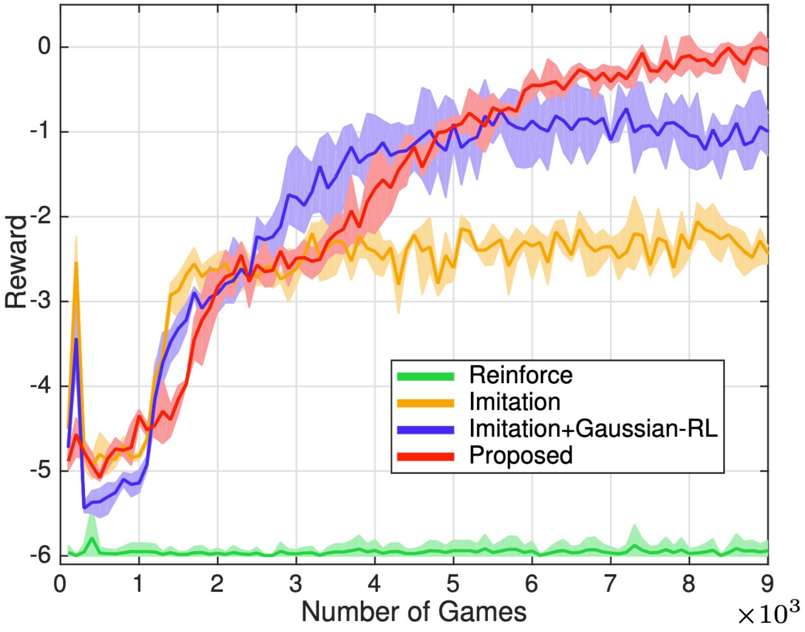
- Model: RNN with external memory

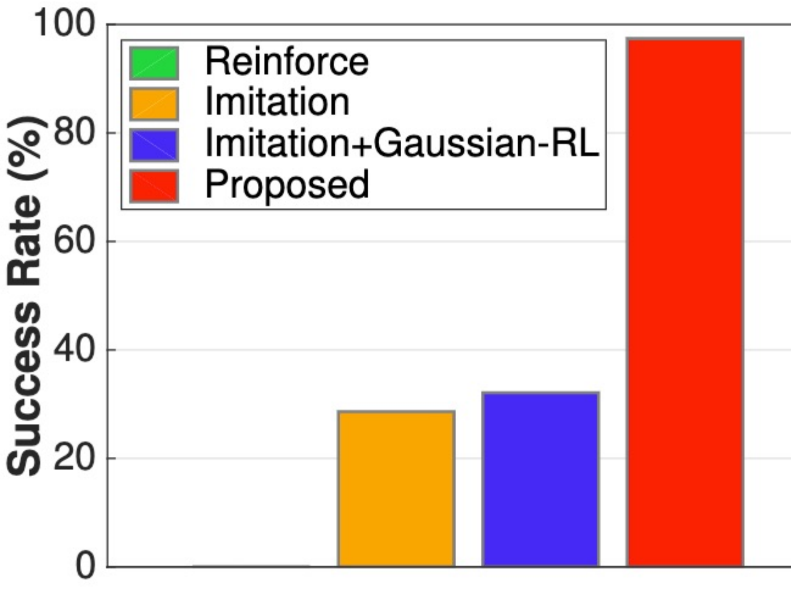


https://arxiv.org/pdf/1805.00462.pdf, Zhang et al, ACL, 2018

Results

- Compared proposed training with combined imitation and reinforce learning loses against baselines
- Compared impact of image variations (dashed) against not using image variations (solid)





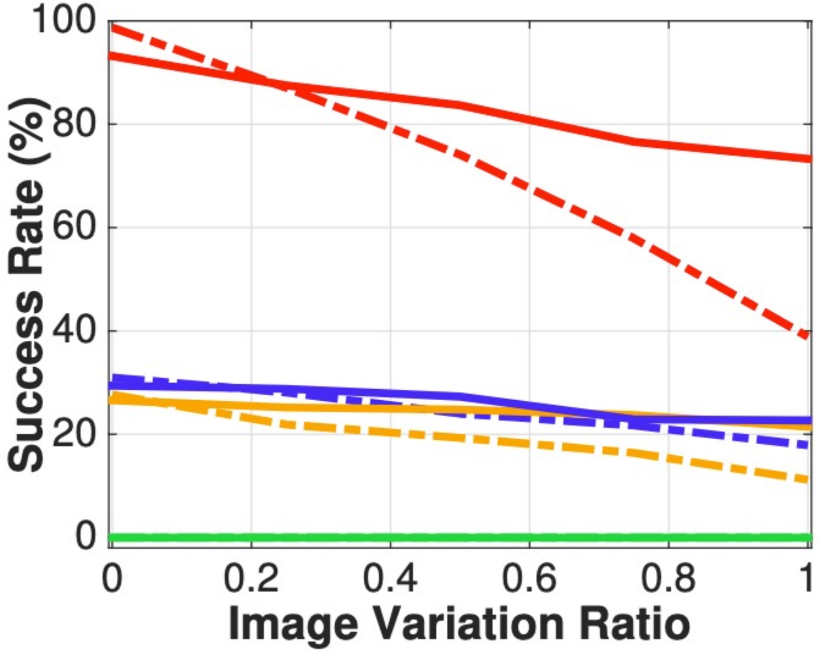


Image variations

Grounded/Situated

Interactive language learning (with environment feedback)

Grounded natural language learning

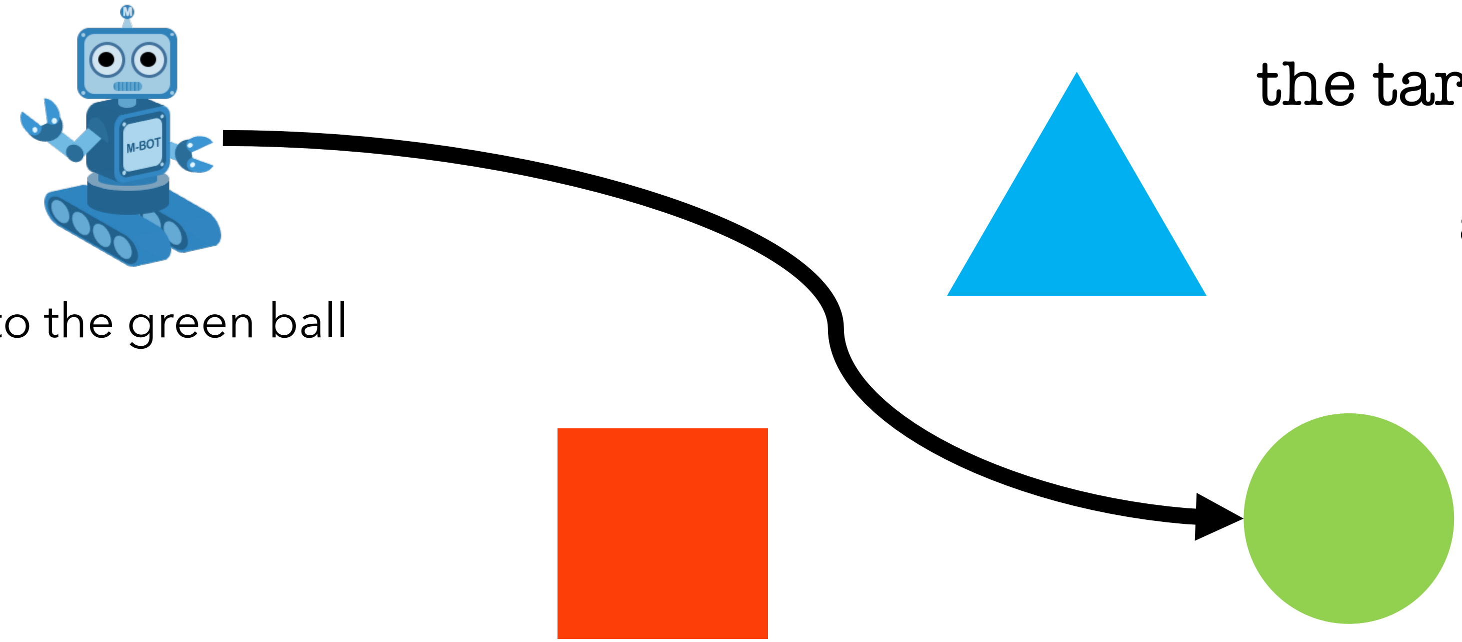
Learning natural language by interacting with an environment

Goal specified as an attributed object

go to the small red object

the target is left of the hair dryer

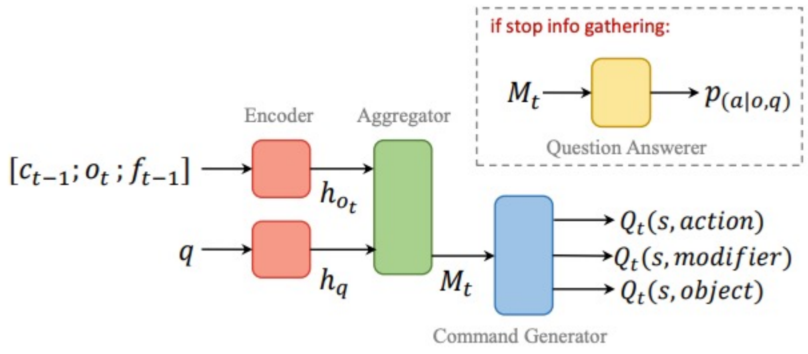
go to any green object



Focus is on language learning – often study generalization to compositionally novel instances

Interactive language learning in TextWorld

- Text adventure game
- Information seeking actions
- Question answering actions



	edible	drinkable	portable	openable	cuttable	sharp	heat_source	cookable	holder
Butter knife			✓			✓			
Oven				✓			✓		✓
Raw chicken			✓		✓			✓	
Fried chicken	✓		✓		✓			✓	

Game	You find yourself in a backyard. You make out a patio table. You see a patio chair. It is stylish. In a corner you can see a BBQ. There is an open screen door leading south.
Question	Is ghargh edible?
Agent	> `Qapla'`
Game	That's not a verb I recognize.
Agent	> `eat ghargh`
Game	You can't see any such thing.
Agent	> `go south`
Game	You've just entered a kitchen. You can see an open fridge here. It is empty! You see a closed oven. Look over there, it's a counter. On the counter you can see a diced potato, a red ghargh and a knife. There is an open screen door leading north.
Agent	> `eat ghargh`
Game	You need to take the red ghargh first.
Agent	> `take red ghargh`
Game	You take the red ghargh from the counter.
Agent	> `eat ghargh`
Game	You eat the red ghargh. Not bad.
Agent	> `wait`
Agent	> `answer: yes`

Interactive Language Learning by Question Answering, https://arxiv.org/pdf/1908.10909.pdf, Yuan et al, EMNLP 2019

Grounded Language Learning

- Grounded Language Learning in a Simulated 3D World arxiv.org/abs/1706.06551
- Understanding Grounded Language Learning Agents arxiv.org/abs/1710.09867

Slide credit: Stefan Lee

Grounded Language Learning

- Gated-Attention Architectures for Task-Oriented Language Grounding
  arxiv.org/abs/1706.07230

Slide credit: Stefan Lee

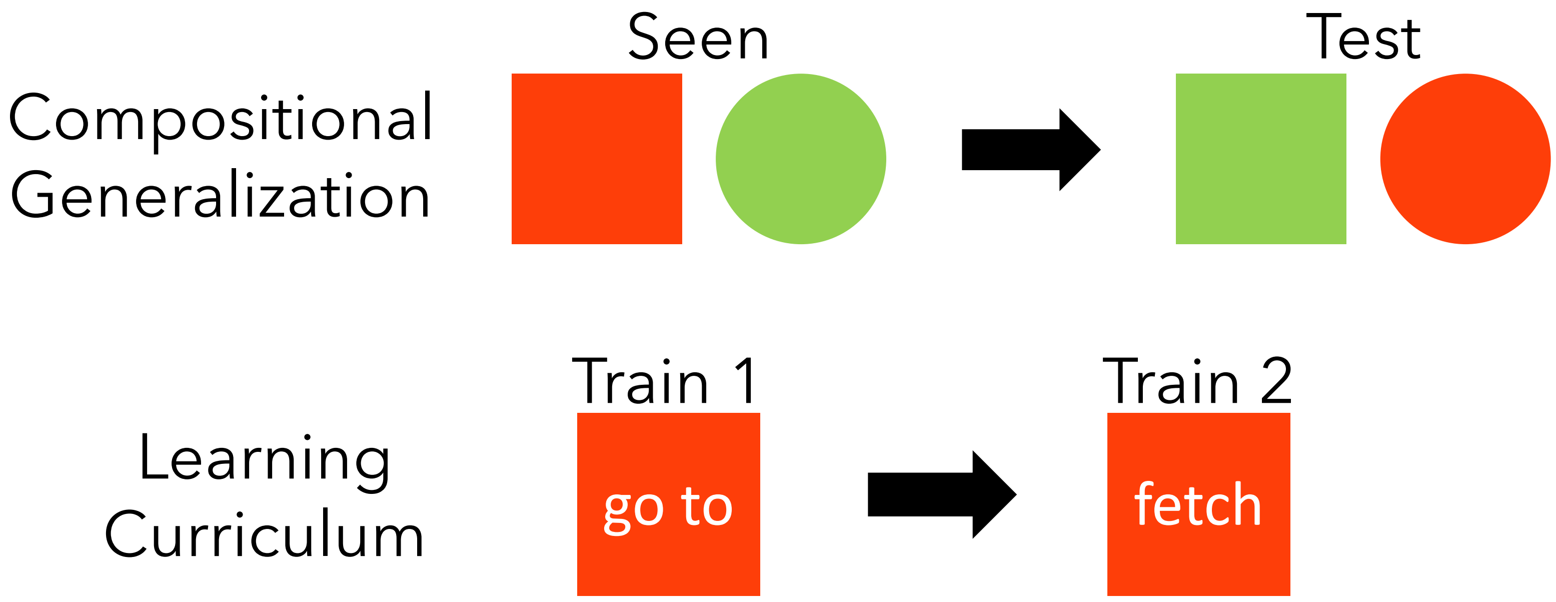
Grounded Language Learning

- Guided Feature Transformation (GFT): A Neural Language Grounding Module for Embodied Agents arxiv.org/abs/1805.08329

Slide credit: Stefan Lee

What is the difference from instruction following?

- Focus is less on measuring whether the agent can understand language and follow instructions correctly but on whether the agent can learn language
- Controlled settings to study specific aspects of language learning (measure what is learned)



Gated-Attention Architectures for Task-Oriented Language Grounding

Environment:

Observation: Egocentric RGB Frame

Actions: turn_left, turn_right, forward

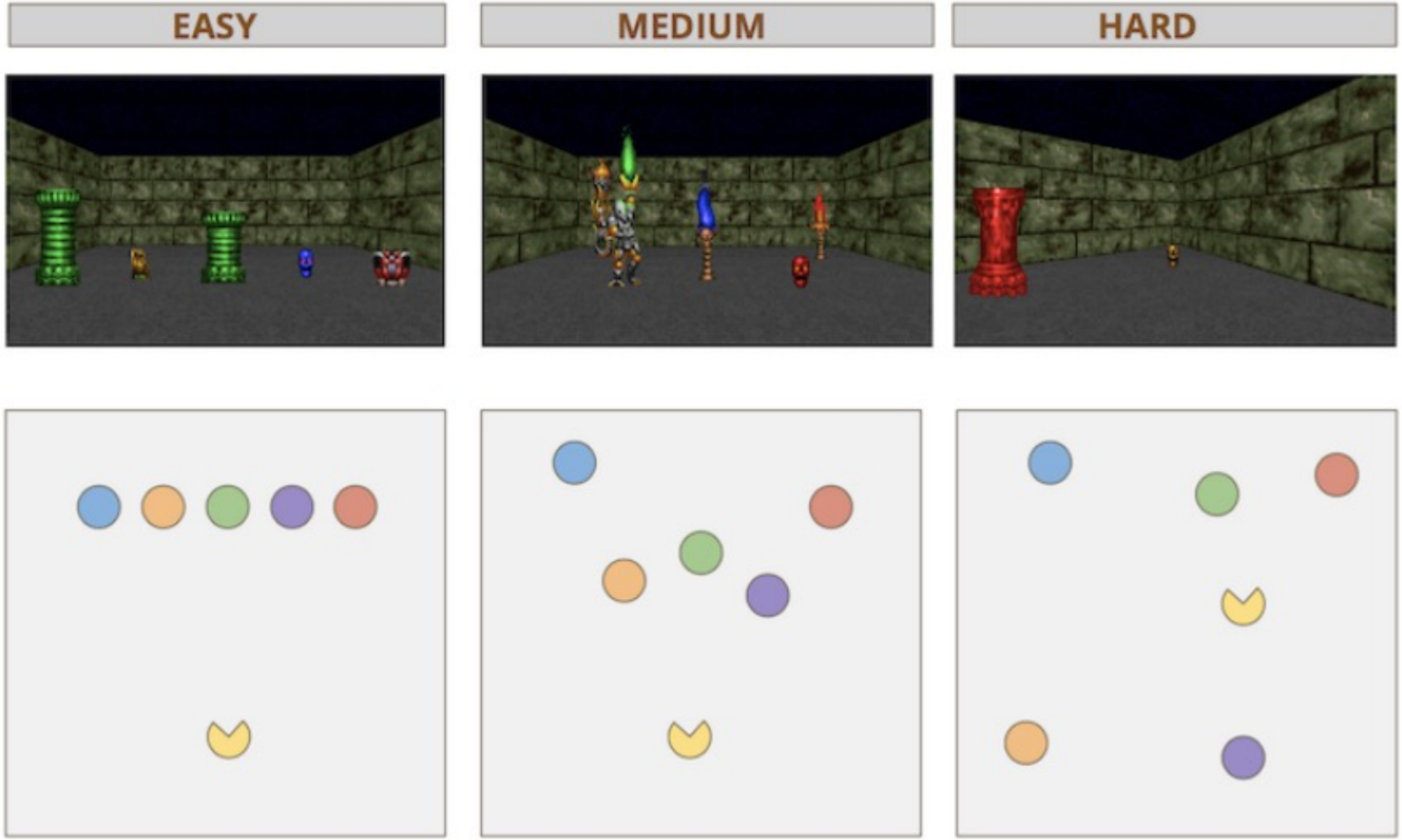
Goal Specification: Templated directions ``go to the red torch"

Slide credit: Stefan Lee

arxiv.org/pdf/1706.07230.pdf

Gated-Attention Architectures for Task-Oriented Language Grounding

Experimental Setting:



Slide credit: Stefan Lee

arxiv.org/pdf/1706.07230.pdf

Gated-Attention Architectures for Task-Oriented Language Grounding

Experimental Setting:

Slide credit: Stefan Lee

arxiv.org/pdf/1706.07230.pdf

Gated-Attention Architectures for Task-Oriented Language Grounding

Experimental Setting:

70 possible instructions (object / attribute combinations)

Instruction Type	Instruction
Size + Color	tall green torch, short red object, short red pillar, short red torch, tall red object, tall blue object, tall green object, tall red pillar, tall green pillar, short blue torch, tall red torch, short green torch, short green object, short blue object, tall blue torch, short green pillar
Color + Size	red short object, green tall torch, red short pillar, red short torch, red tall object, green tall object, blue tall object, red tall pillar, green tall pillar, red tall torch, blue tall torch, green short object, green short torch, blue short object, green short pillar, blue short torch
Color	blue torch, red torch, green torch, yellow object, green armor, tall object, red skullkey, red object, green object blue object, red pillar, green pillar, red keycard, red armor, blue skullkey, blue keycard, yellow keycard, yellow skullkey
Object Type	torch, keycard, skullkey, pillar, armor
SuperlativeSize+Color	smallest yellow object, smallest blue object, smallest green object, largest blue object, largest red object, largest green object, largest yellow object, smallest red object
SuperlativeSize	largest object, smallest object
Size	short torch, tall torch ,tall pillar ,short pillar ,short object, tall object

Slide credit: Stefan Lee

arxiv.org/pdf/1706.07230.pdf

Gated-Attention Architectures for Task-Oriented Language Grounding

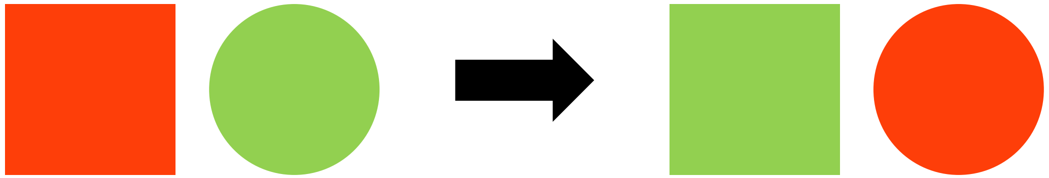
Experimental Setting:

70 possible instructions (object / attribute combinations)

55 used in training, 15 for test

Compositional Generalization

Seen Test



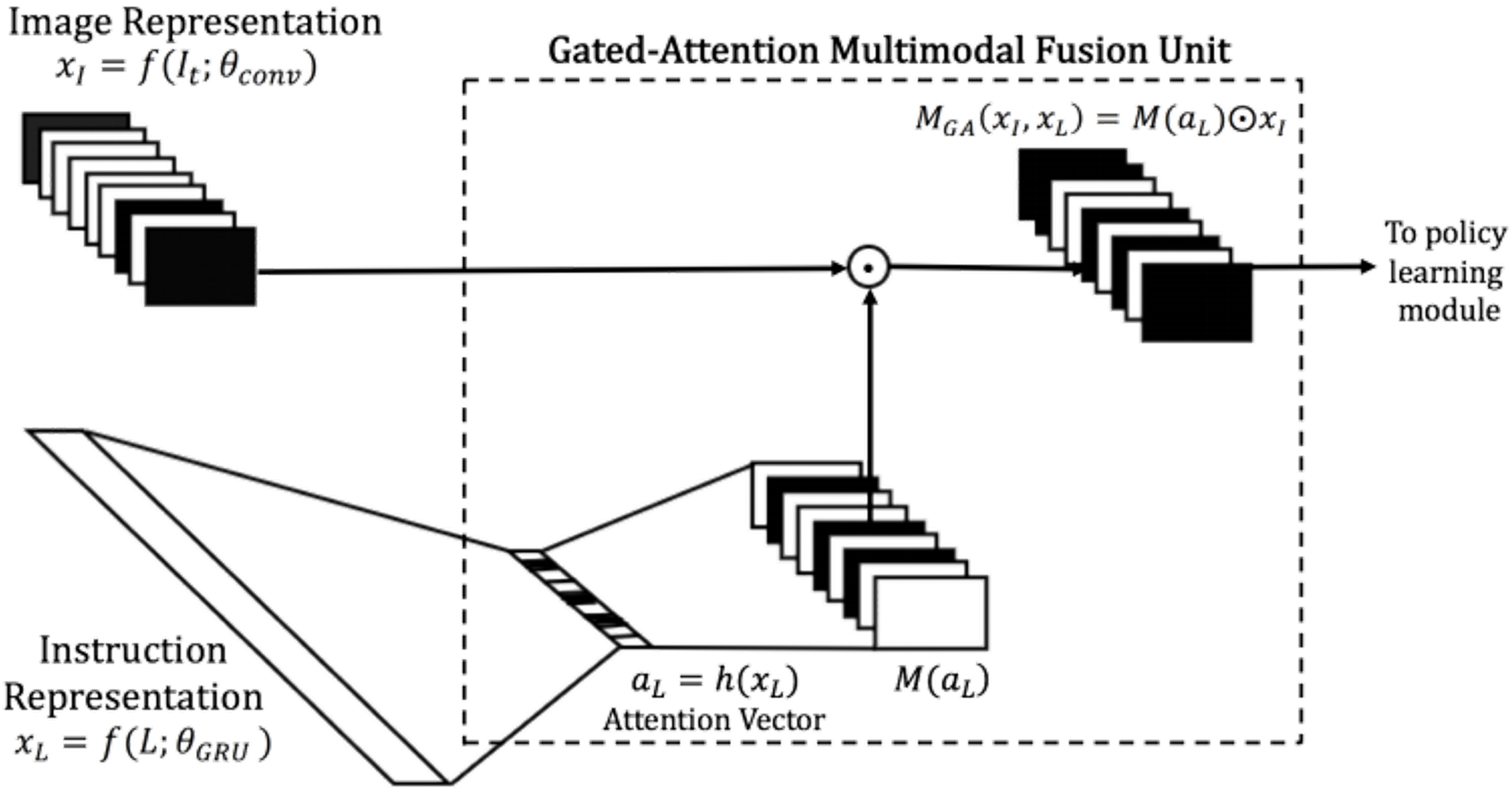
Episodes end on contact with any object or after 50 steps.

Slide credit: Stefan Lee

arxiv.org/pdf/1706.07230.pdf

Gated-Attention Architectures for Task-Oriented Language Grounding

Model: Representation

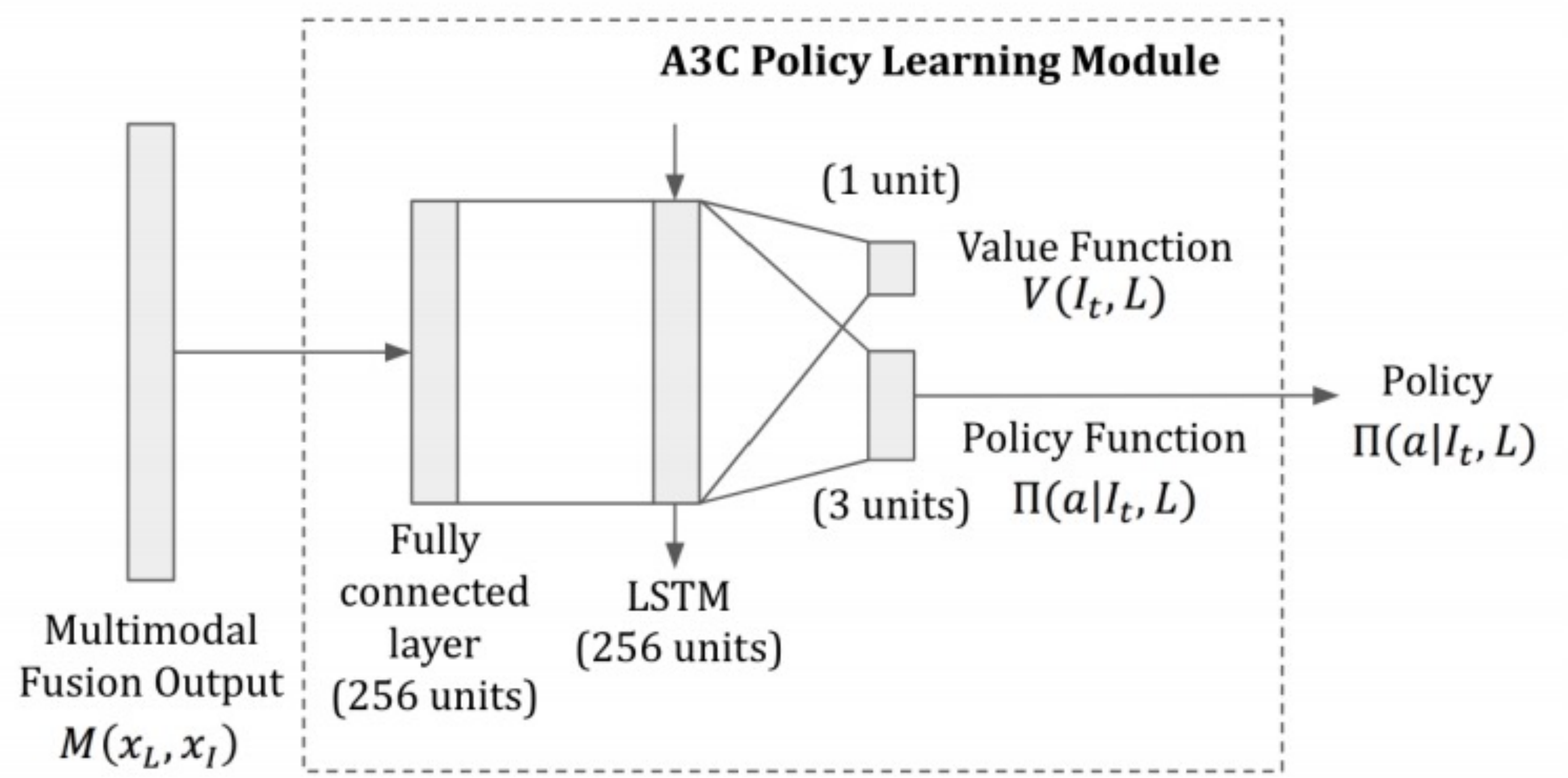


Slide credit: Stefan Lee

arxiv.org/pdf/1706.07230.pdf

Gated-Attention Architectures for Task-Oriented Language Grounding

Model: Policy



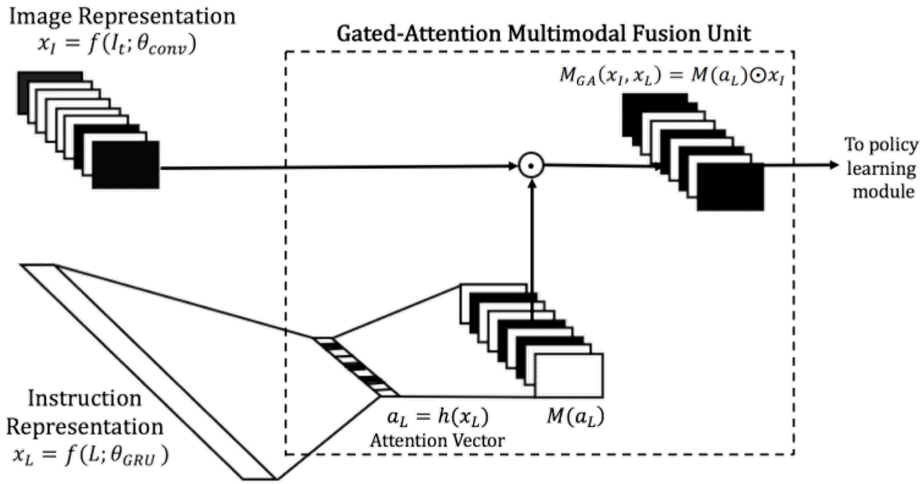
Slide credit: Stefan Lee

arxiv.org/pdf/1706.07230.pdf

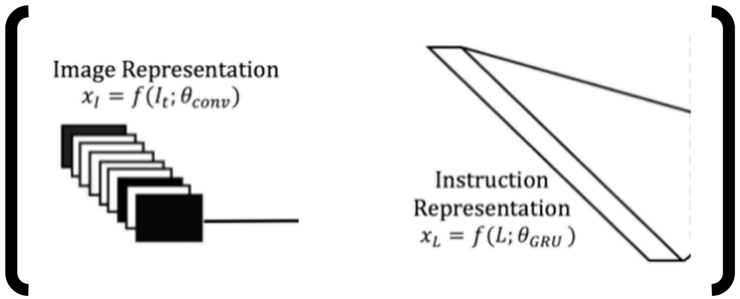
Gated-Attention Architectures for Task-Oriented Language Grounding

Results

Model		Parameters	Easy MT	Easy ZSL	Medium MT	Medium ZSL	Hard MT	Hard ZSL
Imitation Learning	BC Concat	5.21M	0.86	0.71	0.23	0.15	0.20	0.15
	BC GA	5.09M	**0.97**	0.81	0.30	0.23	**0.36**	0.29
	DAgger Concat	5.21M	0.92	0.73	0.45	0.23	0.19	0.13
	DAgger GA	5.09M	0.94	**0.85**	**0.55**	**0.40**	0.29	**0.30**
Reinforcement Learning	A3C Concat	3.44M	1.00	0.80	0.80	0.54	0.24	0.12
	A3C GA	3.39M	1.00	**0.81**	**0.89**	**0.75**	**0.83**	**0.73**



vs.



Slide credit: Stefan Lee

arxiv.org/pdf/1706.07230.pdf

Gated-Attention Architectures for Task-Oriented Language Grounding

Results

MT = Seen instructions, same room, new combination of objects

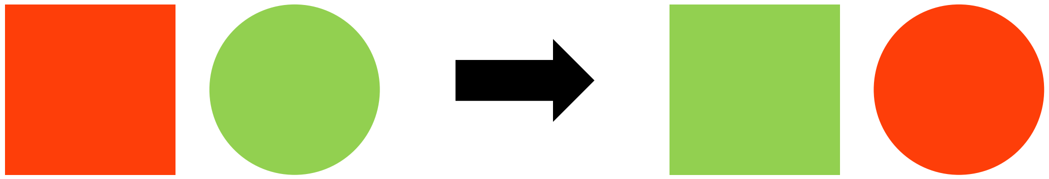
ZSL = New instructions, same room, new combination of objects

Model		Parameters	Easy		Medium		Hard	
			MT	ZSL	MT	ZSL	MT	ZSL
Imitation Learning	BC Concat	5.21M	0.86	0.71	0.23	0.15	0.20	0.15
	BC GA	5.09M	**0.97**	0.81	0.30	0.23	**0.36**	0.29
	DAgger Concat	5.21M	0.92	0.73	0.45	0.23	0.19	0.13
	DAgger GA	5.09M	0.94	**0.85**	**0.55**	**0.40**	0.29	**0.30**
Reinforcement Learning	A3C Concat	3.44M	1.00	0.80	0.80	0.54	0.24	0.12
	A3C GA	3.39M	1.00	**0.81**	**0.89**	**0.75**	**0.83**	**0.73**

Compositional Generalization

Seen

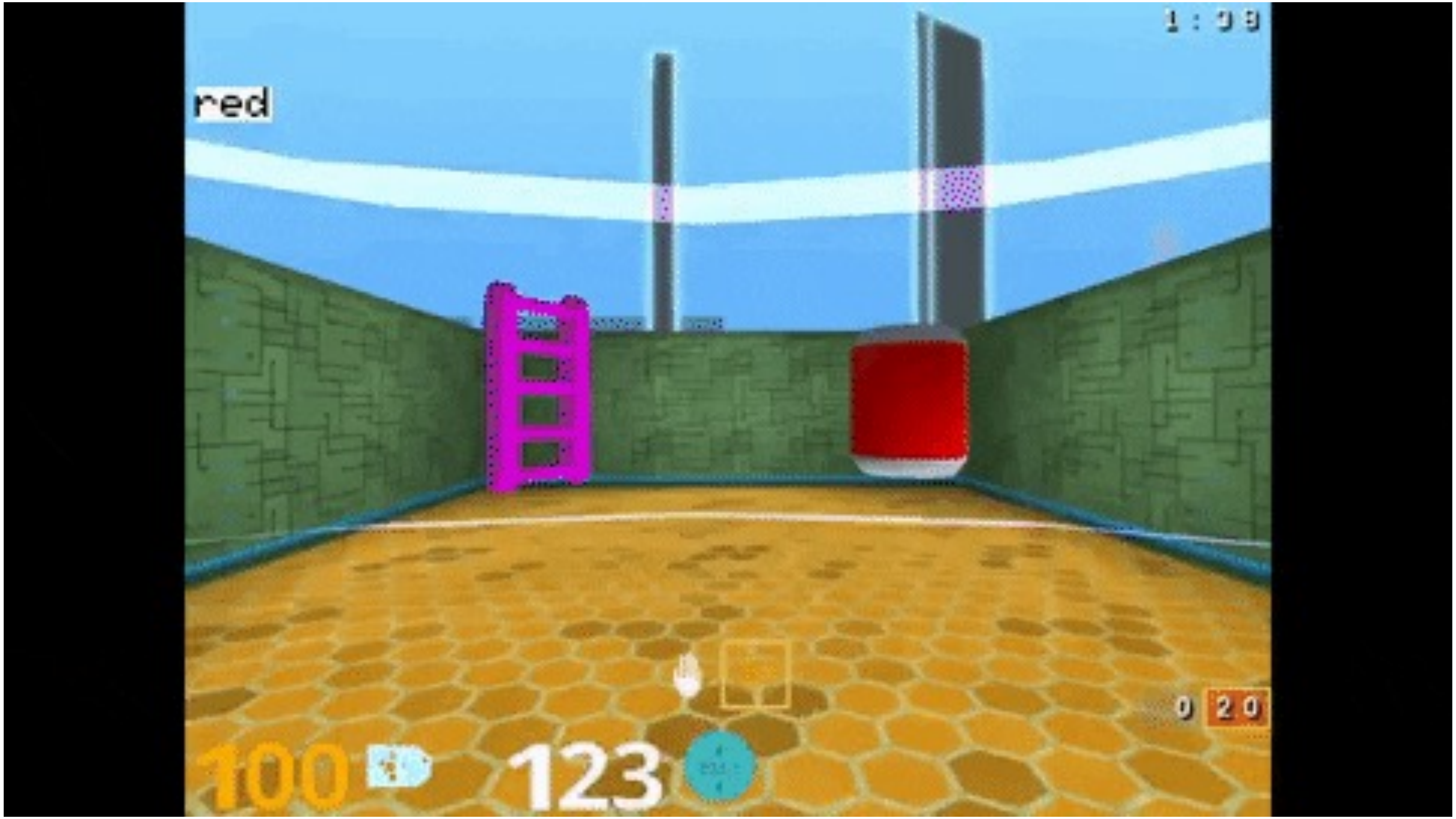
Test



Slide credit: Stefan Lee

arxiv.org/pdf/1706.07230.pdf

Understanding Early Word Learning in Situated Artificial Agents

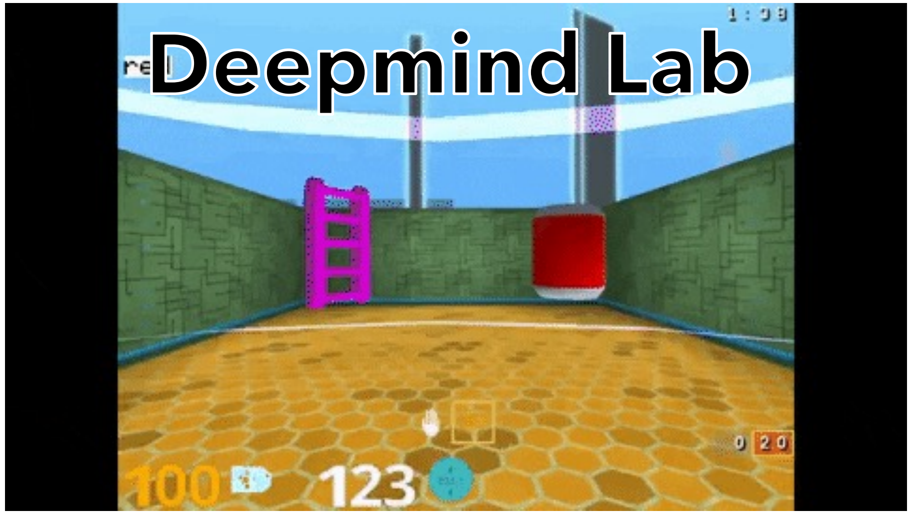


Slide credit: Stefan Lee

arxiv.org/pdf/1710.09867.pdf

Understanding Early Word Learning in Situated Artificial Agents

Environment:



Observation: Egocentric RGB Frame

Actions: move-forward, move-back, move-left, move-right, look-left, look-right, strafe-left, strafe-right

Goal Specification: Single word descriptor

Slide credit: Stefan Lee arxiv.org/pdf/1710.09867.pdf

Understanding Early Word Learning in Situated Artificial Agents

Goal Specification: Single word descriptor

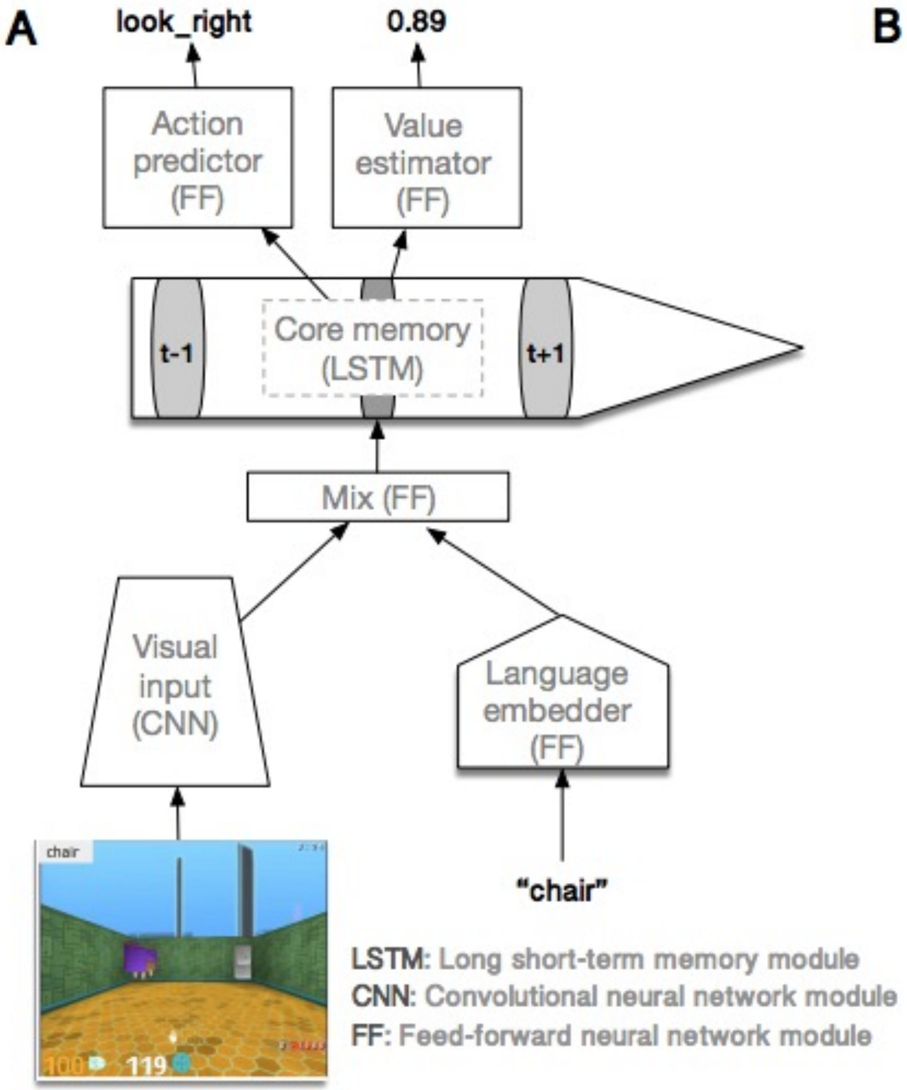
Word class (class size)	**Example**	**Instruction meaning (in this setting)**
shapes (40)	"*pencil*"	Find and bump into a pencil.
colors (10)	"*blue*"	Find and bump into any blue object.
patterns (2)	"*striped* "	Find and bump into any striped object.
relative shades (2)	"*darker*"	Find and bump into the darker of the two objects in front of you.
directions (2)	"*left*"	Find and bump into the object furthest to the left as you look.

Slide credit: Stefan Lee

arxiv.org/pdf/1710.09867.pdf

Understanding Early Word Learning in Situated Artificial Agents

Model:



Slide credit: Stefan Lee

arxiv.org/pdf/1710.09867.pdf

Understanding Early Word Learning in Situated Artificial Agents

Experimental Setting:

Fixed room
Fixed spawn
Fixed object locations
Randomized objects

No notion of generalization

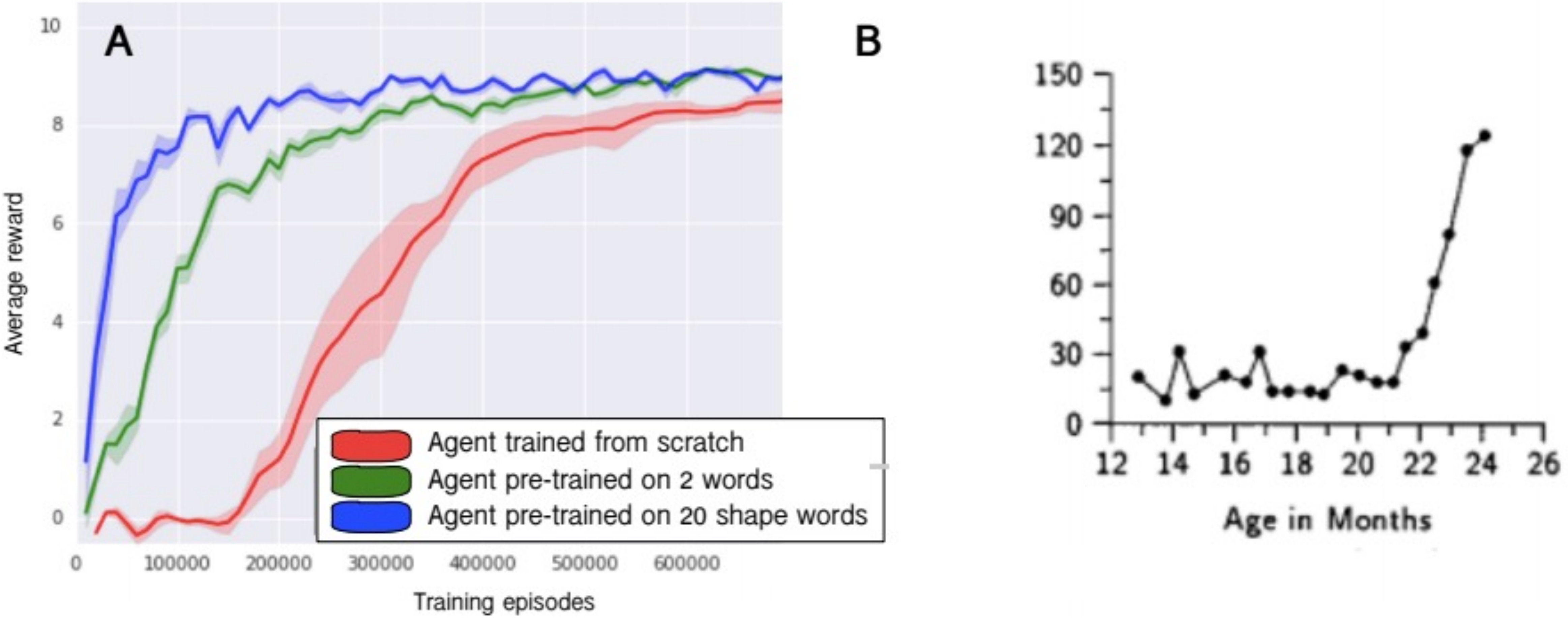
Interested in dynamics of learning instead.

Slide credit: Stefan Lee

arxiv.org/pdf/1710.09867.pdf

Understanding Early Word Learning in Situated Artificial Agents

Results

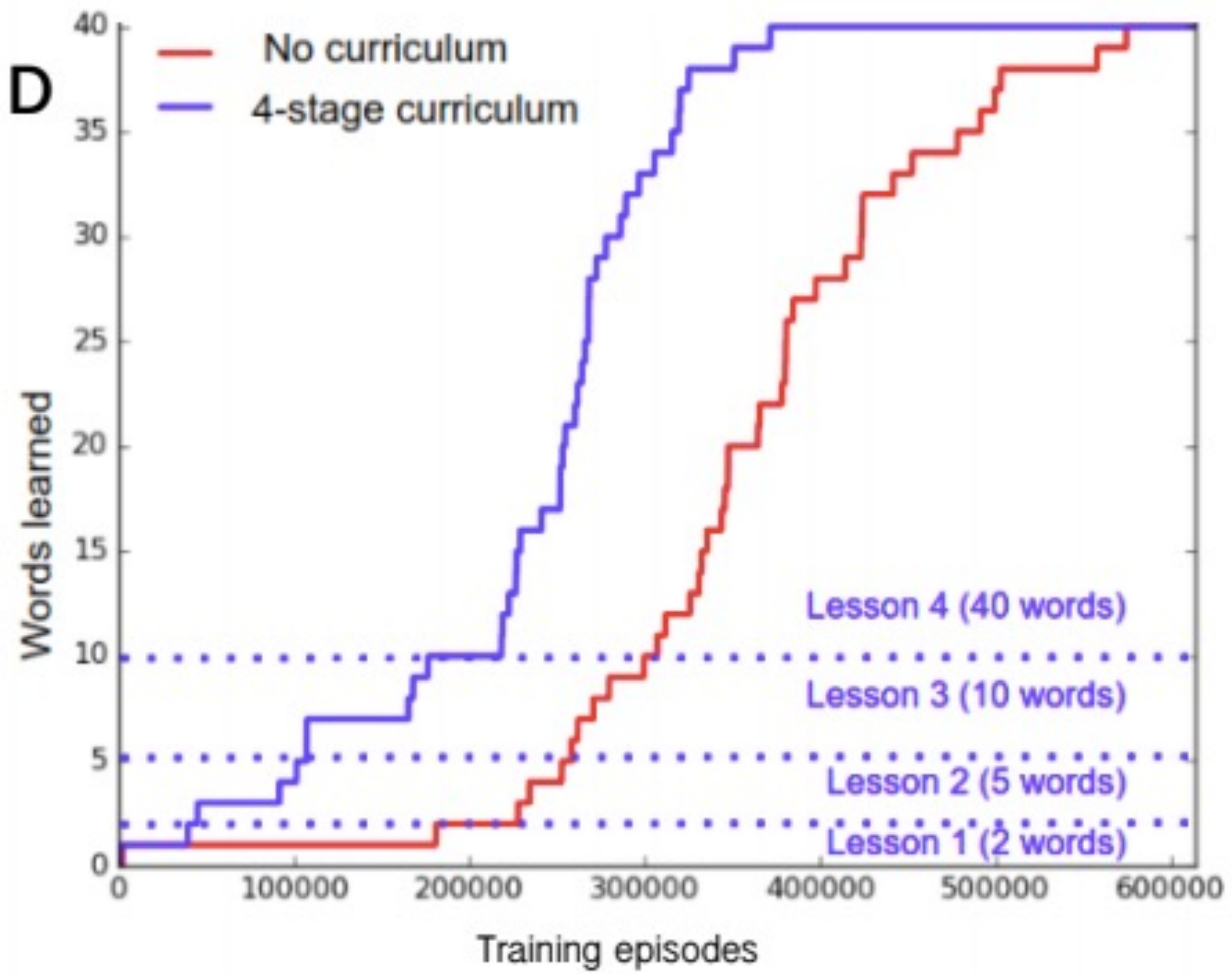


Do agents have 'language bursts' like infants?

Slide credit: Stefan Lee arxiv.org/pdf/1710.09867.pdf

Understanding Early Word Learning in Situated Artificial Agents

Results



Does curriculum help? Some work suggests it does in humans.

Slide credit: Stefan Lee

arxiv.org/pdf/1710.09867.pdf

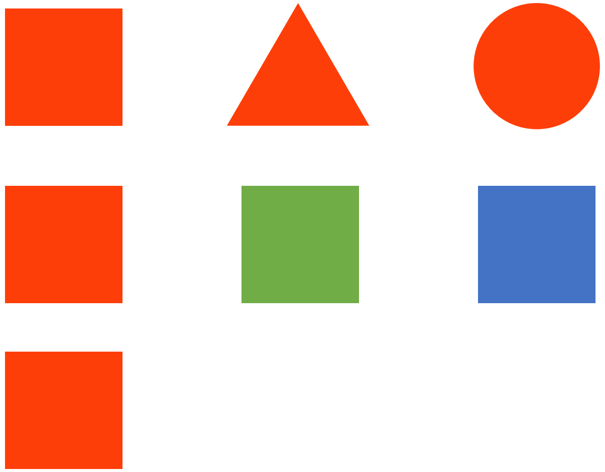
Understanding Early Word Learning in Situated Artificial Agents

Results

red

square

dax



What happens now when the agent see this?

dax

Slide credit: Stefan Lee

arxiv.org/pdf/1710.09867.pdf

Understanding Early Word Learning in Situated Artificial Agents

Results

Humans assume shape words. Agent leans towards color.

Slide credit: Stefan Lee

arxiv.org/pdf/1710.09867.pdf

Next time / end of term

- Easter Holiday (4/5) – let’s take a break!
- Optional readings (extra credit)
  - Learning Language Games through Interaction
  - Learning Adaptive Language Interfaces through Decomposition
- Thursday (4/8): Conditional generative models from text
- Monday (4/12): Paper discussion or more on generating stuff from text
- Thursday (4/15): Last day – project discussion and conclusion