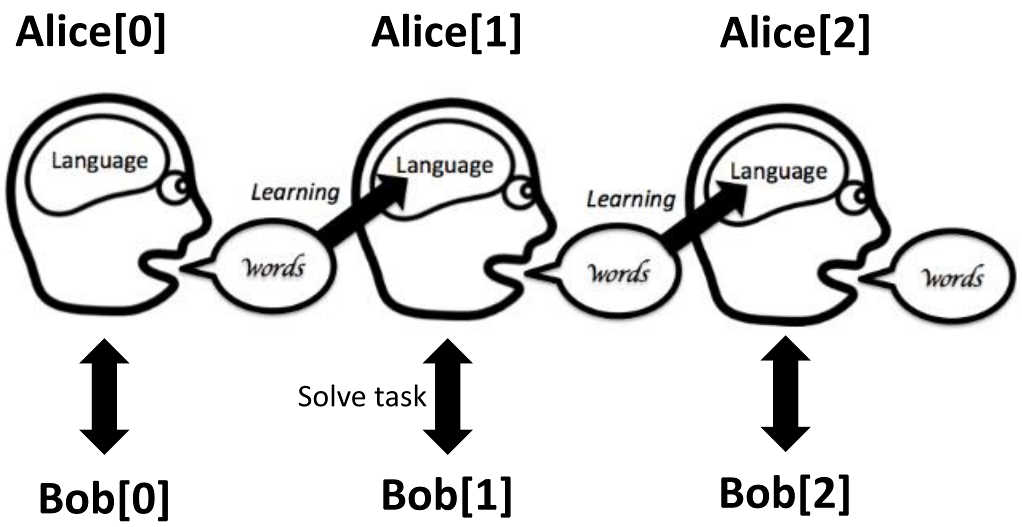
Neural Iterated Learning: Applications and Understandings



Yi (Joshua) Ren, UBC

OUTLINES

- Part 1: Introduce IL by some examples
- Part 2: Extending IL to deep learning
- Part 3: Overview of IL's evolution

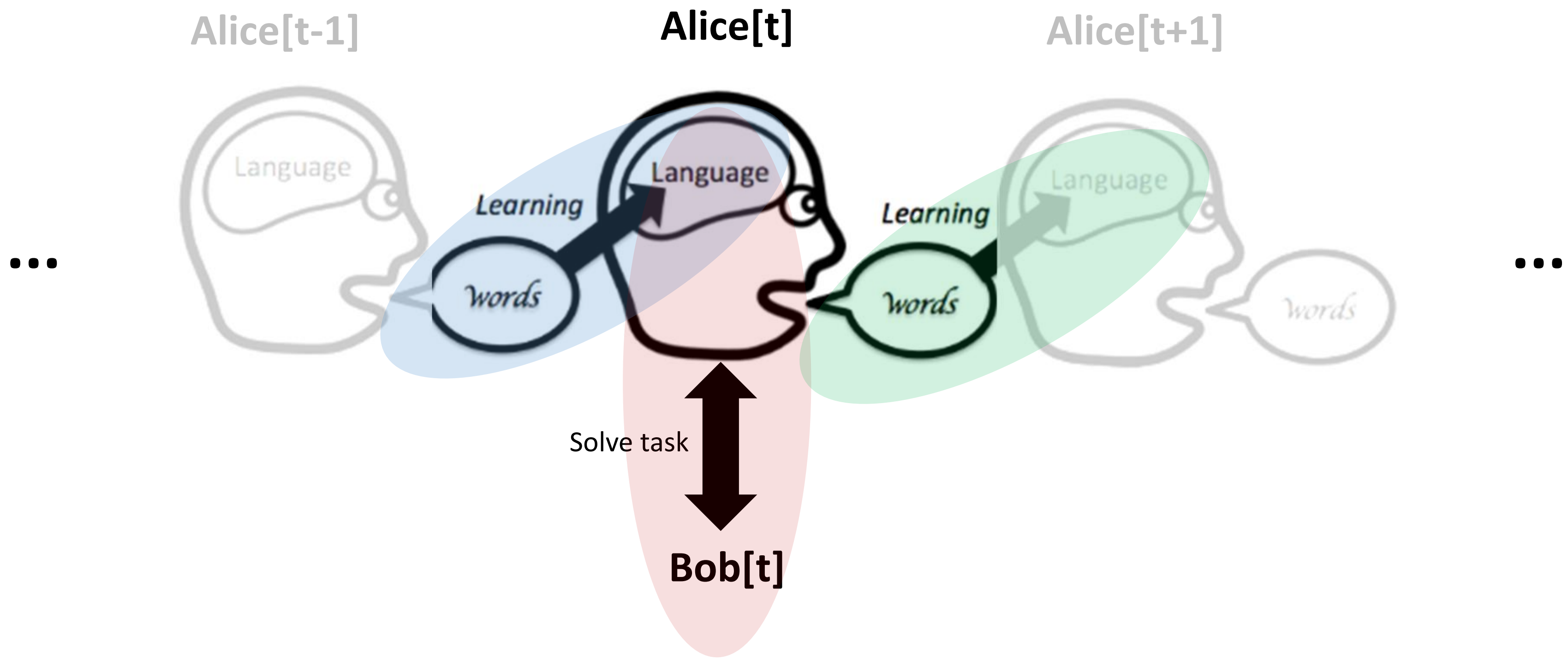
OUTLINES

- Part 1: Introduce IL by some examples
- Part 2: Extending IL to deep learning
- Part 3: Overview of IL's evolution

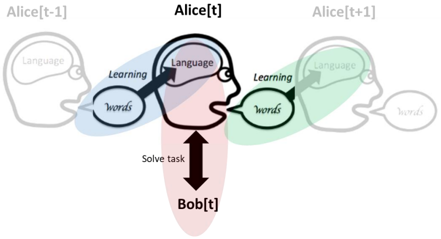
What is Iterated Learning?

IL describes a chain of learning procedure:

1. ***Imitation:*** An innocent agent learn language from its predecessors
2. ***Interaction:*** This agent use learned language to accomplish tasks
3. ***Transmission:*** This agent transfer language to the next generation.



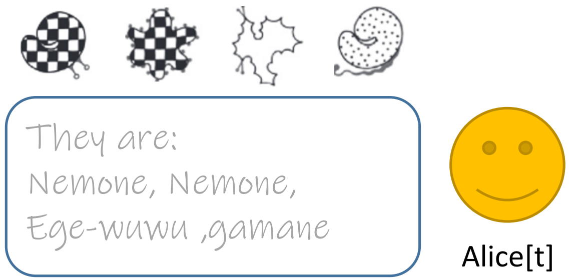
What is Iterated Learning?



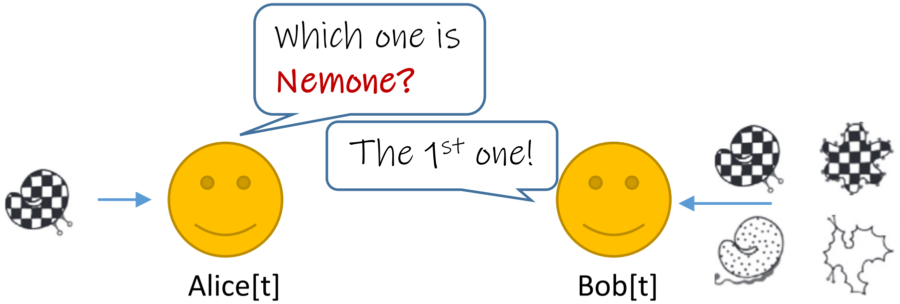
Exp 1: simulating the emergence of ***compositionality*** *in human language (Kirby-2008)*

Task: create names for each icons and use that to accomplish a game.

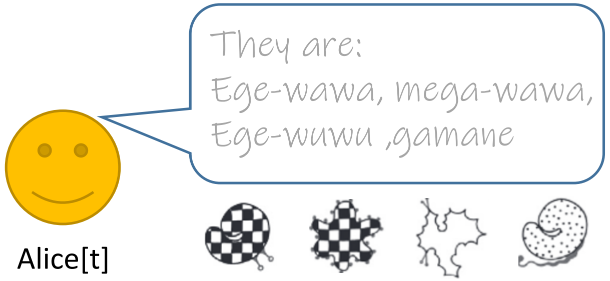
- **Imitation phase:**



- **Interaction phase – Lewis Language Game:**



- **Transmission phase:**



	pihino		kapa		newhomo
	nemone		gakho		kamone
	piga		wuwele		gaku
	kawake		nepi		hokako

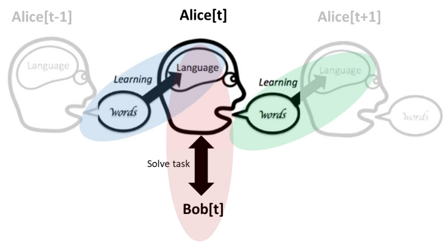
0th Gen

	ege-wawu		mega		gamene-wawu
	ege-wawa		mega-wawa		gamene-wawa
	ege-wuwu		mega-wuwu		gamene-wuwu
	ege		wulagi		gamane

9th Gen

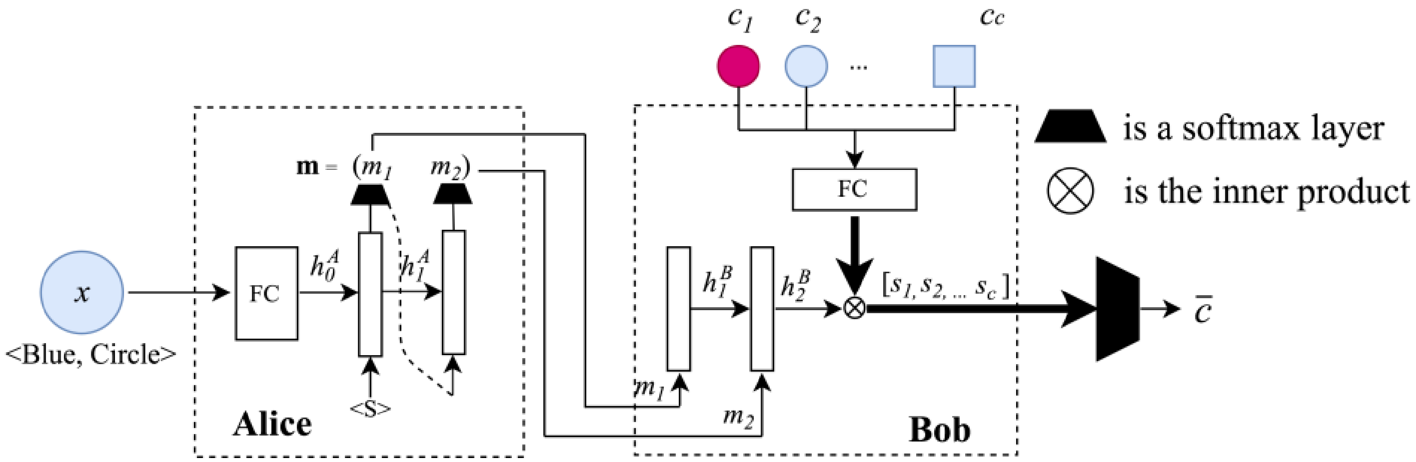
Kirby, Simon, Hannah Cornish, and Kenny Smith. "Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language." Proceedings of the National Academy of Sciences 2008

What is Iterated Learning?



*Exp 2: improve the compositionality of the **neural representation** (Ren-2020)*

Almost the same setting with exp1, but with neural network agents.



$x \in \mathcal{X} \triangleq [\text{Color}, \text{Shape}]$

$m \in \mathcal{M} \triangleq [\mathrm{m}_1, \mathrm{m}_2]$

$h \in \mathcal{H}: \mathcal{X} \rightarrow \mathcal{M}$

0th

	blue	green	cyan	brown	red	black	yellow	white
box	aa	fh	af	hh	cg	fc	ha	hf
circle	da	df	hb	db	fa	da	dh	fb
triangle	gc	ff	ge	gf	gg	fg	ge	he
square	ae	fb	be	bb	bg	fb	gb	ba
star	ad	fd	de	db	dg	fd	ce	hc
diamond	ac	dd	dc	db	dg	fd	dc	dd
pentagon	ad	fe	ef	bd	eg	fc	ee	ed
capsule	aa	dd	de	db	dg	gd	de	fh

9th

	blue	green	cyan	brown	red	black	yellow	white
box	aa	ea	ba	ga	da	ca	ha	fa
circle	ab	eb	bb	gb	db	cb	hb	fb
triangle	ae	**eb**	be	ge	de	ce	he	fe
square	af	ef	bf	gf	df	cf	hf	ff
star	ac	ec	bc	gc	dc	cc	**dh**	fc
diamond	ad	ed	bd	gd	dd	cd	hd	fd
pentagon	ag	eg	bg	gg	dg	cg	hg	fg
capsule	ah	eh	bh	gh	**hc**	ch	hh	fh

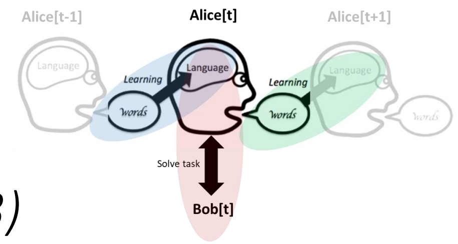
- **Imitation phase:** cross-entropy loss between Alice's prediction and data
- **Interaction phase – Lewis Game:** REINFORCE update for agents

$$\nabla_{\theta_A} J = \mathbb{E}\left[R(\bar{c}, x)\nabla \log p_A(\mathbf{m}|x)\right] + \lambda_A \nabla H[p_A(\mathbf{m}|x)]$$
$$\nabla_{\theta_B} J = \mathbb{E}\left[R(\bar{c}, x)\nabla \log p_B(\bar{c}|\mathbf{m}, c_1, \ldots, c_c)\right] + \lambda_B \nabla H[p_B(\bar{c}|\mathbf{m}, c_1, \ldots, c_c)],$$

- **Transmission phase:** random sample $m \sim p_A(m|x)$

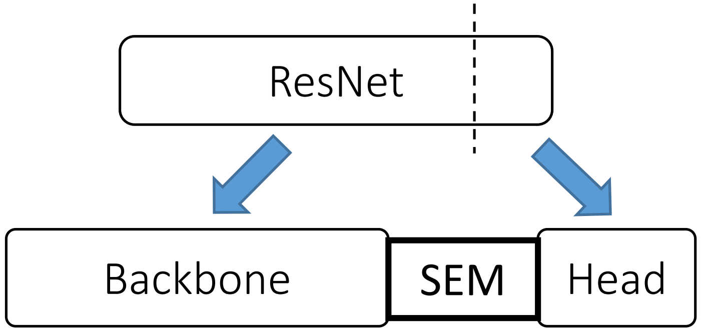
Ren, Yi, et al. "Compositional languages emerge in a neural iterated learning model." ICLR 2020

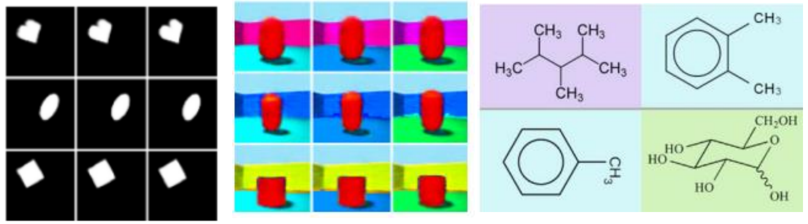
What is Iterated Learning?



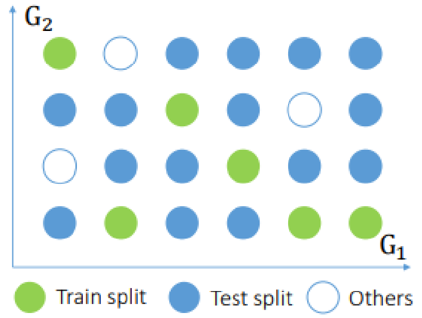
Exp 3: improve the compositional ***generalization*** *in general representation learning (Ren-2023)*

Where is Alice and Bob in a ***general supervised learning*** *system?*



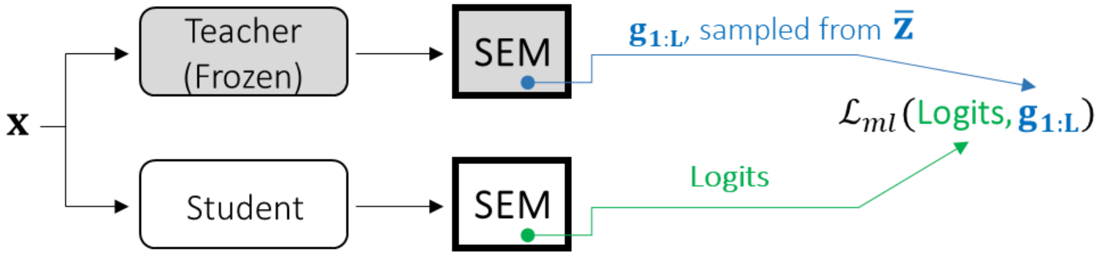


x: Vision or Graph y: Target loss



Comp-gen: non-overlapping split

- **Imitation phase:** knowledge distillation through SEM block



- **Interaction phase:** directly use downstream loss

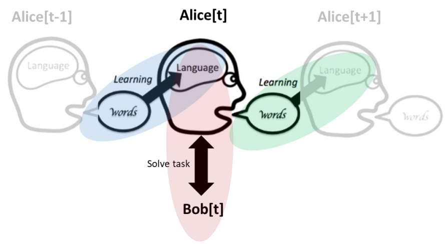
- **Transmission phase:** set student as teacher for next geneneration

Comp-gen ability improved!

Model and Algorithm		molhiv (AUROC ↑)			
		Valid-full	Test-full	Valid-half	Test-half
GCN	Baseline	82.41±1.14	76.25±0.38	75.65±0.91	72.31±1.86
	Baseline+	81.61±0.63	75.58±1.00	73.23±0.75	72.17±1.02
	SEM-only	84.00±1.10	78.40±0.67	74.84±1.57	72.81±2.32
	SEM-IL	**84.89±0.68**	**79.09±0.67**	**78.48±0.67**	**74.02±0.78**
GIN	Baseline	81.76±1.04	76.99±1.42	76.95±1.40	71.63±2.21
	Baseline+	81.55±0.72	77.01±0.94	74.77±1.62	69.75±3.10
	SEM-only	83.05±0.90	78.21±0.78	76.29±2.06	72.70±4.94
	SEM-IL	**83.32±1.51**	**78.61±0.73**	**78.06±1.24**	**72.89±0.48**

Ren, Yi, et al. "Improving compositional generalization using iterated learning and simplicial embeddings." NeurIPS 2023

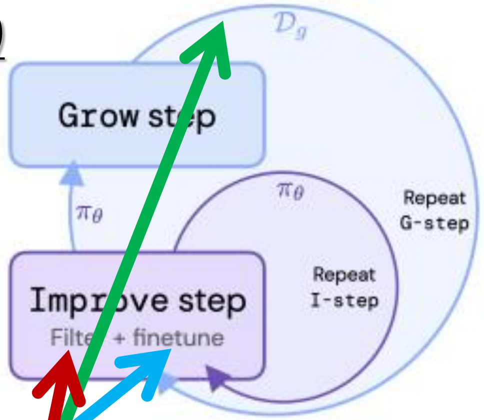
What is Iterated Learning?

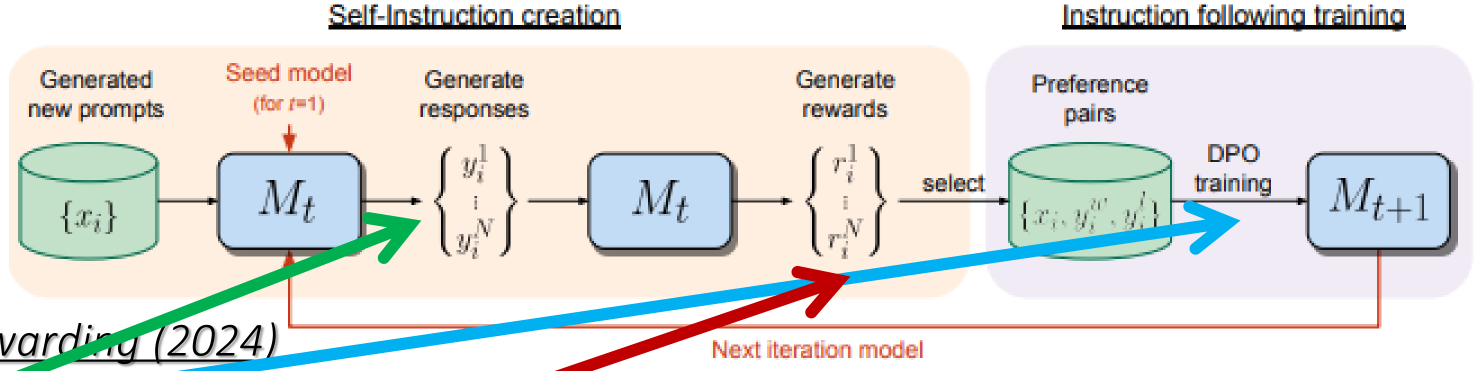


*Exp 4: amplifying **hidden bias** in LLM-agents (Many **Self-XXX algorithms,** Ren-2024)*

Using IL to explain Self-xxx algorithm, treating them as intelligent agents

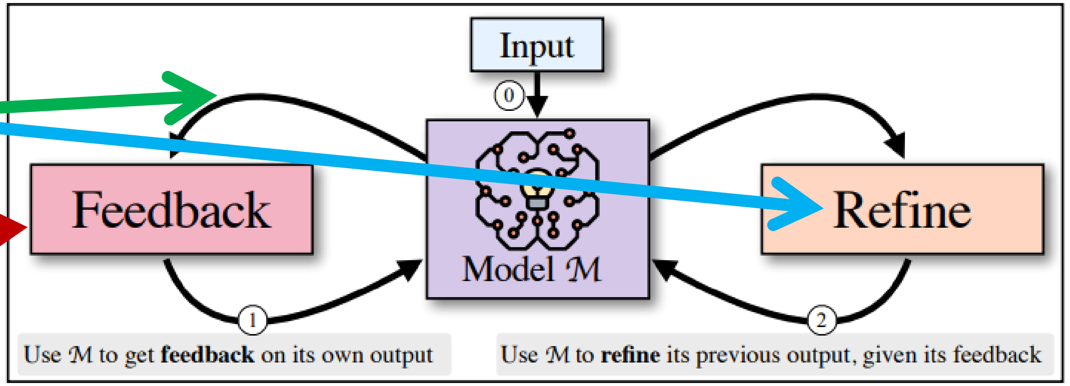
ReST (2023)





Self-rewarding (2024)

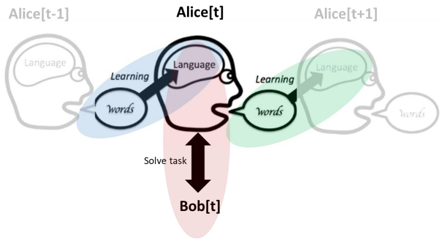
Self-refine (2023)



- **Imitation phase:** finetune, or ICL on generated new data
- **Transmission phase:** sample new data for next-gen
- **Interaction phase:** ranking the generated data by rewarding

Gulcehre, Caglar, et al. "Reinforced self-training (ReST) for language modeling." arXiv 2023.
Yuan, Weizhe, et al. "Self-rewarding language models." arXiv preprint arXiv 2024.
Madaan, Aman, et al. "Self-refine: Iterative refinement with self-feedback." NeurIPS 2023

What is Iterated Learning – Short Summarization



- Known facts:

Exp 1: human prefer compositionality → compositional language is achieved

Exp 2: simple NN prefer compositionality → compositional mapping is achieved

Exp 3: complex NN prefer systematicness→ systematical generalization is improved

Exp 4: LLM have different biases→ the bias (good and bad) are amplified

(More details in Ren-2024)

- What IL brings?

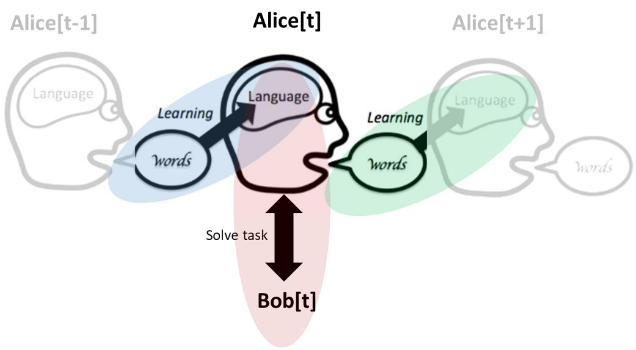
Iterated learning can gradually amplify the hidden bias of the intellegent agent.

(This amplifying effect is hard to achieved by explicit regularizers, more details in Ren-2023)

OUTLINES

- Part 1: Introduce IL by some examples
- Part 2: Extending IL to deep learning
- Part 3: Overview of IL's evolution

Extending IL to Deep Learning – Two Buildingblocks



Recall:

Iterated learning can gradually **amplify** the **hidden bias** of the intellegent agent.

Q1: How could IL amplify bias?
A1: Bayesian-IL framework

Q2: Where the bias comes from?
A2: Depends on the agent
(data, model structure, learning)

A1. Bayesian-IL framework:

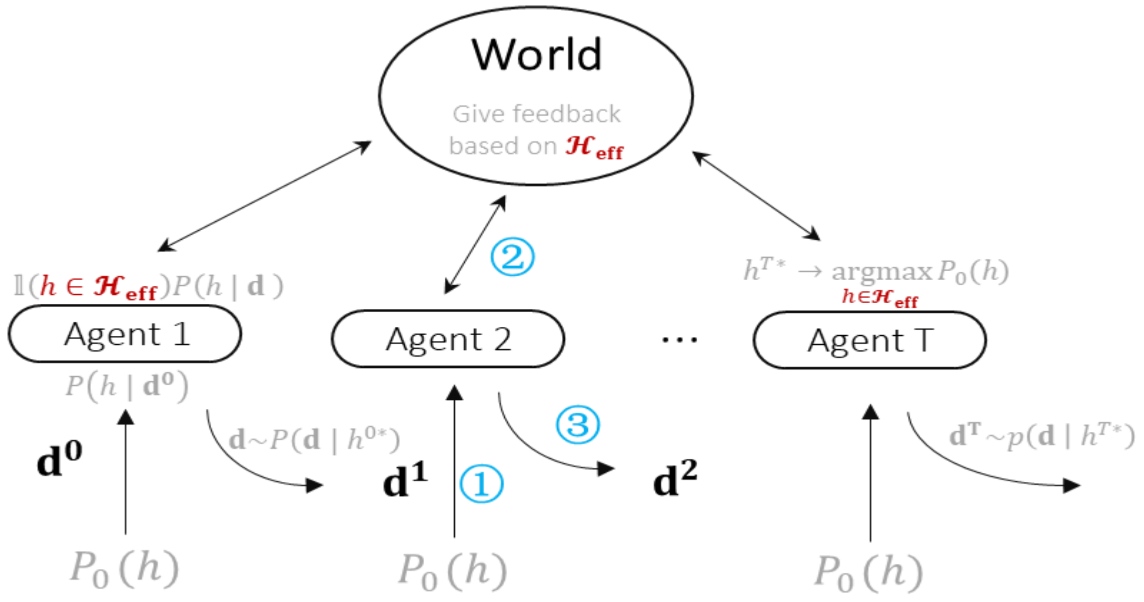
Object: $x \in \mathcal{X}$

Message: $m \in \mathcal{M}$

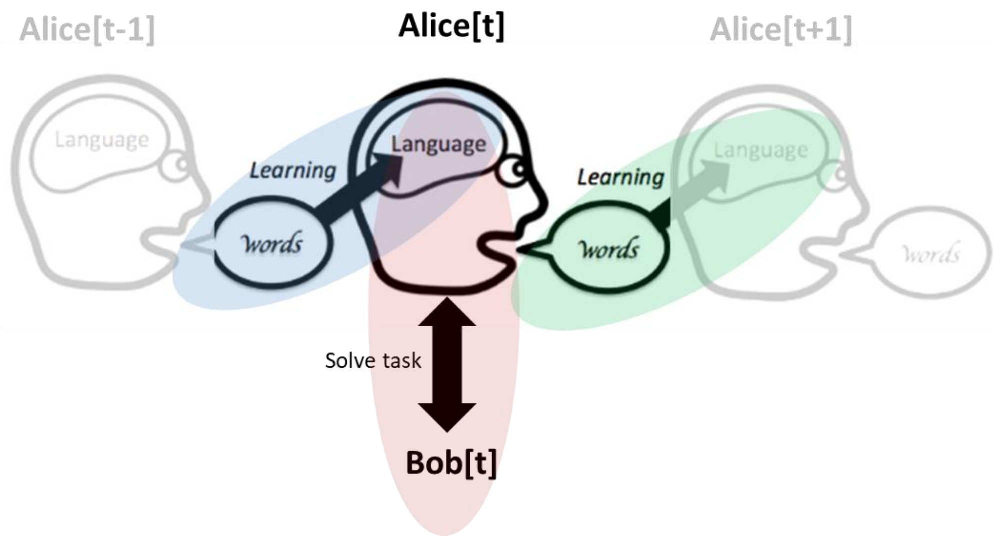
Data pair: $d = (x, m)$

Hypothesis: $h \in \mathcal{H}: \mathcal{X} \to \mathcal{M}$

Prior: $\mathrm{P}_0(h)$



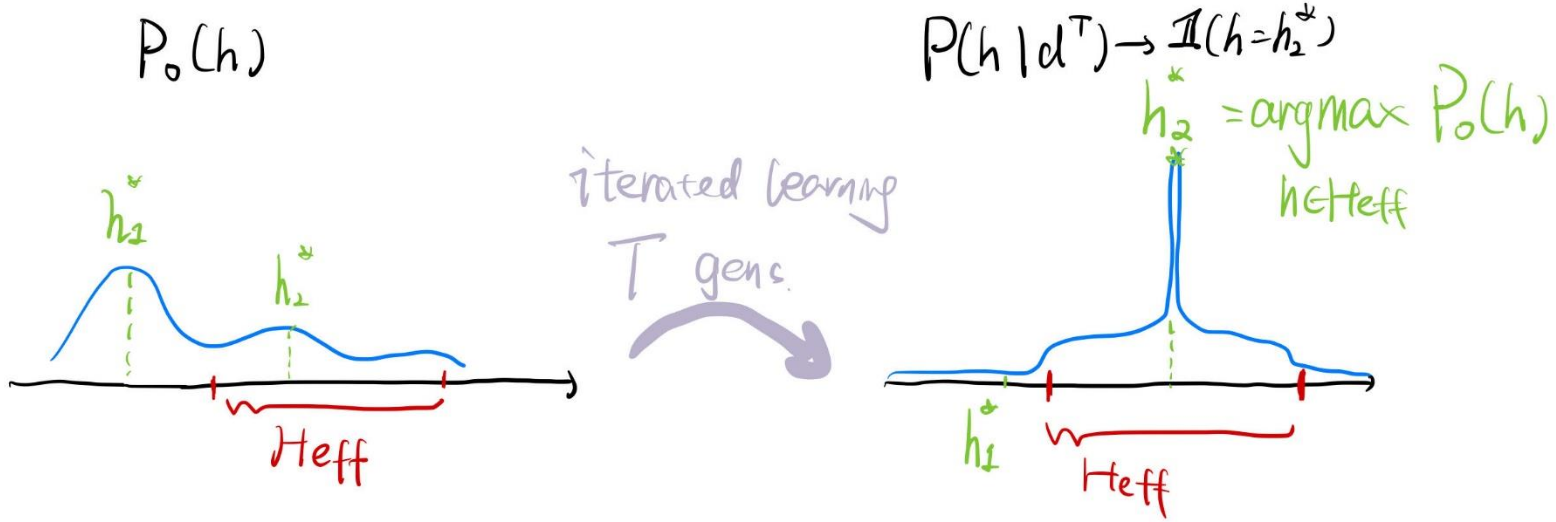
- **Imitation phase:**
  agent start from $\mathrm{P}_0(h)$, learn from $\mathbf{d}^{\mathbf{t-1}}$, becomes $\mathrm{P}\left(h \mid \mathbf{d}^{\mathbf{t-1}}\right)$
- **Interaction phase:**
  conduct task, and have $\mathbb{I}(h \in \mathcal{H}_{\mathrm{eff}})\mathrm{P}\left(h \mid \mathbf{d}^{\mathbf{t-1}}\right)$
- **Transmission phase:**
  sample $\mathbf{d}^{\mathbf{t}} \sim \mathrm{P}(d \mid h^*)$, where $h^* = \operatorname{argmax}_{h \in \mathcal{H}_{\mathrm{eff}}} \mathrm{P}\left(h \mid \mathbf{d}^{\mathbf{t-1}}\right)$
- **Theoretical guarantee:**
  $\mathrm{P}\left(h \mid \mathbf{d}^{T}\right) \to \mathbb{I}(h = h^{T*})$, where $h^{T*} = \operatorname{argmax}_{h \in \mathcal{H}_{\mathrm{eff}}} \mathrm{P}_0(\mathrm{h})$

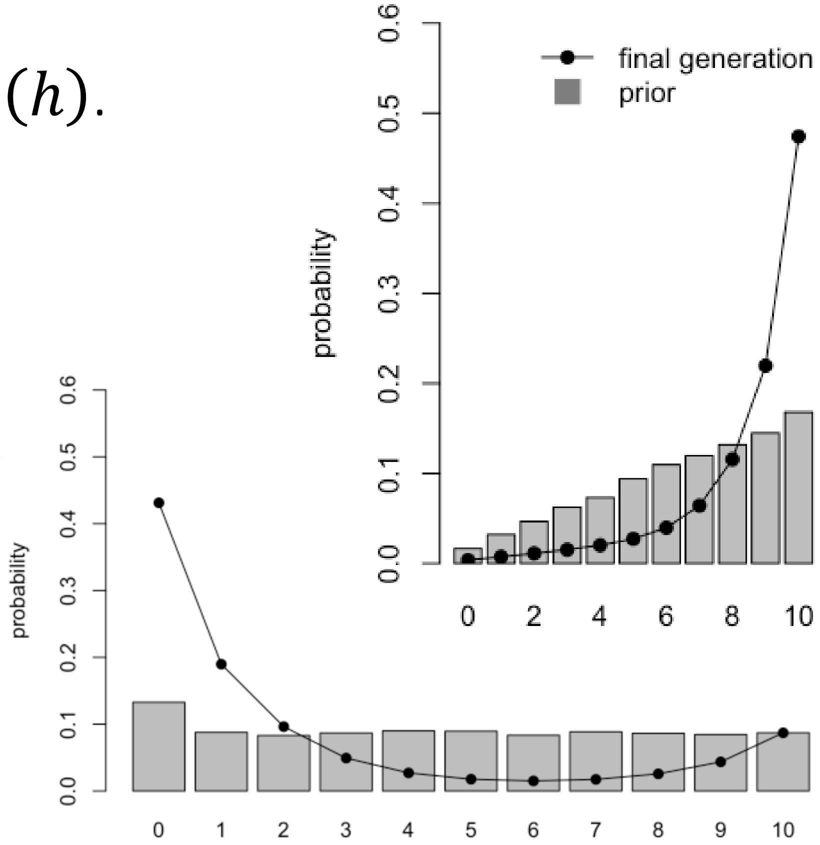


More details: Ren Yi, et.al "Language Model Evolution: An Iterated Learning Perspective", submitted 2024

A1. Bayesian-IL framework:

Iterated learning can amplify bias in model's prior $\mathrm{P}_0(h)$.
Interaction phase further guide the evolution.





Q2: Then, what is a typical bias in $\mathrm{P}_0(h)$, and where it comes from?

Q2: what is a typical bias in prior and where it comes from?

- Known facts:

  Exp 1: **human prefer** compositionality → compositional language is achieved

  Exp 2: simple NN prefer **compositionality** → compositional mapping is achieved

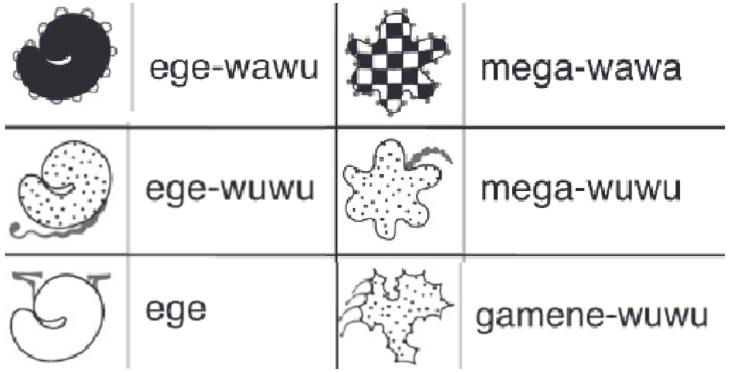
  Exp 3: complex NN prefer **systematicness** → systematical generalization is improved

  Exp 4: LLM have **different inborn biases** → the bias (good and bad) are amplified

- Bias can be arbitrary, but let's start from compositionality in exp 2&3.

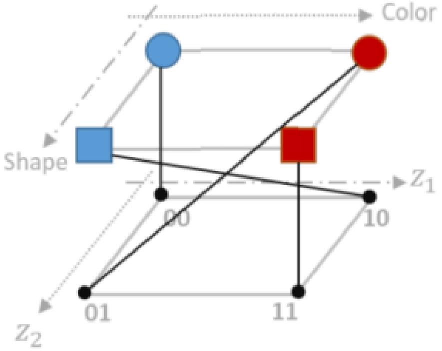
A2. Bias is learning speed advantage:

- Exp 1 assumes human's cognition system is good at **finding patterns**.

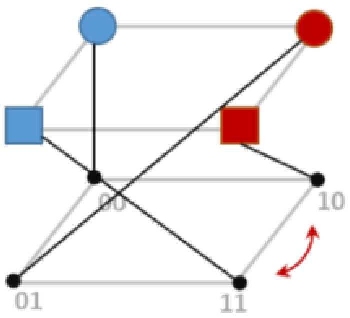


ege/mega → shape?
wuwu/wawa → texture?

- But this is not so obvious for neural network,
  because <u>**mutual information**</u> cannot separate the following two mappings:



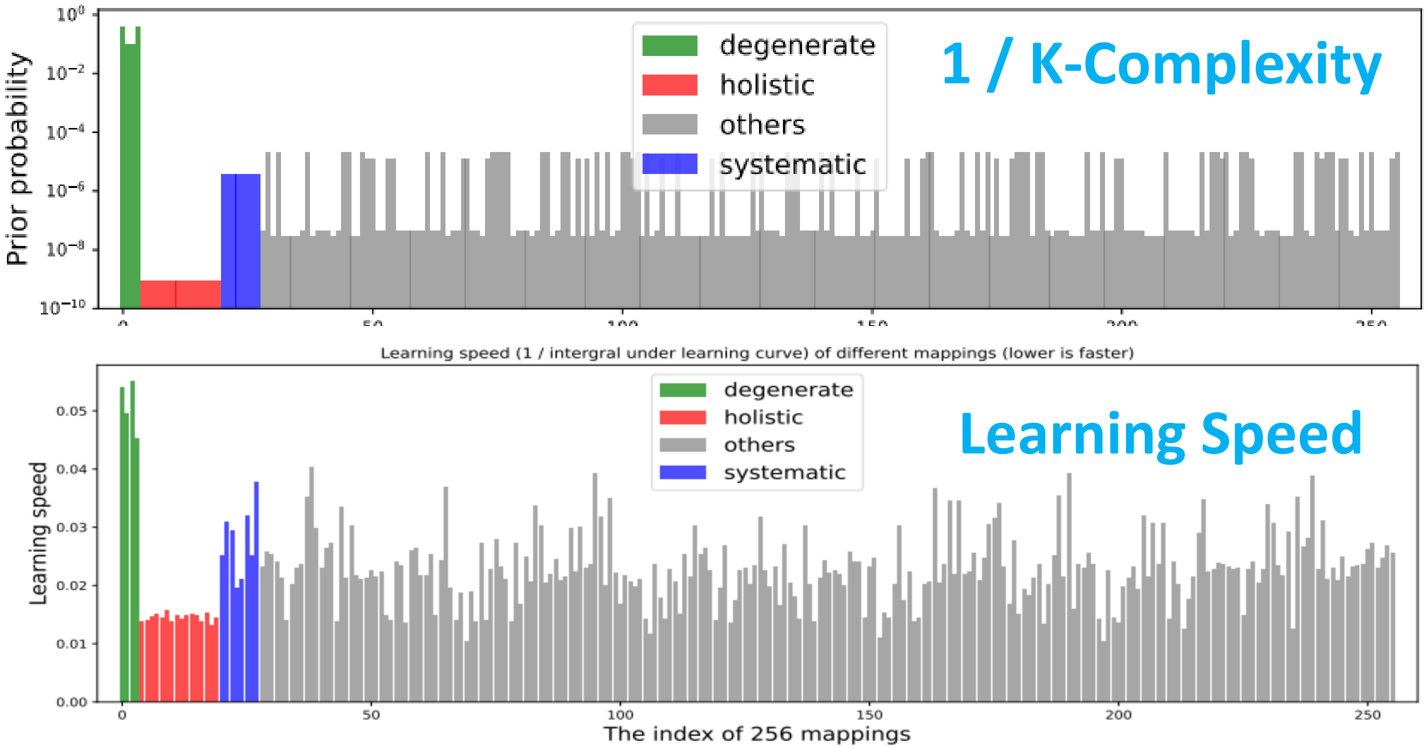
A compositional mapping 5 rules, $\alpha = 43$			
S		$\rightarrow$	z2, z1
z2:	0	$\rightarrow$	blue
z2:	1	$\rightarrow$	red
z1:	0	$\rightarrow$	circle
z1:	1	$\rightarrow$	box

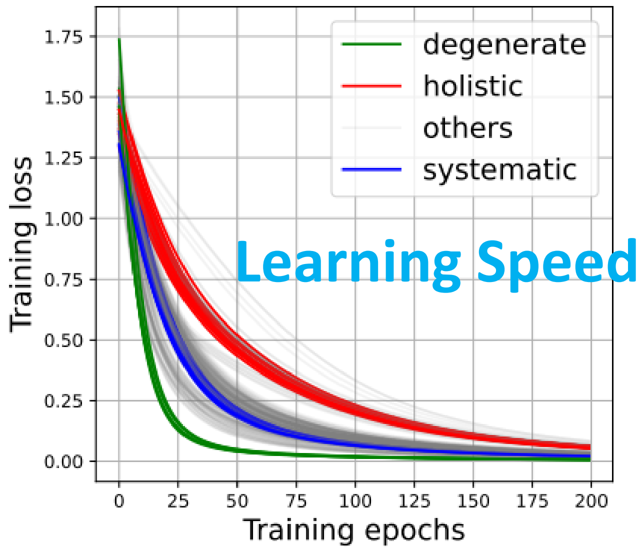


A holistic mapping 4 rules, $\alpha = 56$		
S:	00 $\rightarrow$	blue circle
S:	01 $\rightarrow$	red circle
S:	10 $\rightarrow$	red box
S:	11 $\rightarrow$	blue box

A2. Bias is learning speed advantage:

- We find this bias is embodied in model's learning speed: compositional mapping learns faster!
  - For the 2 color 2 shape problem, we have 256 different mappings
  - We draw their prior probablity based on their coding length: $P(l;\alpha) \propto 2^{-\frac{l}{\alpha}}$
  - We let a MLP learn these 256 mappings seperately, and observe their learning speed (Defined as the integral under learning curve, similar to that mentioned in Jack Rae's talk)

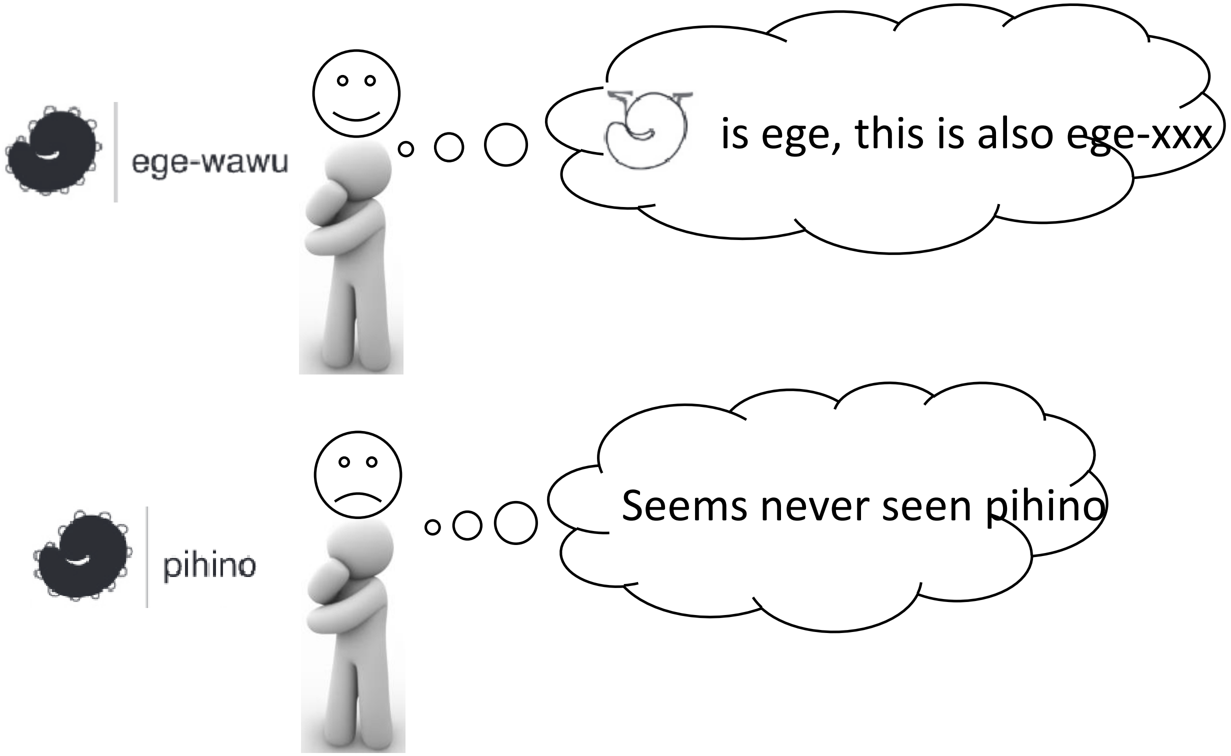




They are highly correlated!

A2. Where the learning speed advantage comes from?

- Explanation:
  - ✓ Propose A: from **learning dynamics** (based on our work of Ren-2022)
  - ✓ Propose B: from group theory and **kolmogorov complexity** (based on Ren-2023)



Yi Ren, et.al "Better supervisory signals by observing learning paths." ICLR 2022
Yi Ren, et al. "Improving compositional generalization using iterated learning and simplicial embeddings." NeurIPS 2023

OUTLINES

- Part 1: Introduce IL by some examples
- Part 2: Extending IL to deep learning
- Part 3: Overview of IL's evolution

Evolution of Iterated Learning

- Experiments and applications (imcomplete list):

[1] Language Evolution
MH Christiansen, S Kirby
Initial idea about 2-stages training

2003

[Many other related works]

[3] Lab experiments, S Kiryb
Conduct experiments on human, propose 2 pressures for IL

2015

Simplicity bias is inborn in human cog-system

[5] Neural agents, Yi Ren
Show IL also works for simple NN on a RL setting

2020

Simplicity bias also exist when training NN using SGD

Different applications on various fields:
[6] Machine Translation, Yuchen Lu
Apply on MT task to counter language drift, enc-dec
[7] Visual Question Answering, Ankit Vani
Improve VQA performance using IL, enc-dec
[8] Multilabel classification, S Rajeswar
Improve multi-label results using IL on IN1K, enc-only
[9] Color naming system, Emil Carlsson
Verify simplicit bias both exist in NN and human, enc-dec
[10] Systematic generalization, Yi Ren
Extend to a general representation learning, backbone-head
...

Other phenomenon might related to IL:
[11] Fortuituous forgetting, Hattie Zhou
Repeatly reset part of the network to create knowledge gap
[12] Primacy bias in RL, E Nikishin
Repeatly reset part of the policy net helps a lot
[13] Iterated learning on CLIP, Ankit Vani,
Applying it on self-supervised learning framework
...

- Theoretical Understandings (imcomplete list):

[2] Bayesian Agents, Tom Griffiths
Proves the convergence of **imitation-only** IL using a Bayesian framework

2007

[Many other related works]

[4] Continuous hypothesis, D J. Navarro
Extend the Bayesian framework to heterogeneous agents case

2018

[14] A patching for interaction phase, Yi Ren
Add a small patching on interaction phase.
Verify the behavior of LLM on different phases

2024

Evolution of Iterated Learning – Citation list

[1] Language Evolution: *Christiansen, Morten H., and Simon Kirby, eds. Language evolution. OUP Oxford, 2003.*
[2] Bayesian Agents: *Griffiths, Thomas L., et.al. "Language evolution by iterated learning with Bayesian agents." Cognitive science 2007*
[3] Lab experiments: *Kirby, Simon, et al. "Compression and communication in the cultural evolution of linguistic structure." Cognition 2015*
[4] Continuous hypothesis: *Navarro, D. J., et.al*
"When extremists win: Cultural transmission via iterated learning when populations are heterogeneous." Cognitive Science 2018
[5] Neural agents: *Yi Ren, et.al, "Compositional languages emerge in a neural iterated learning model" ICLR 2020*
[6] Machine Translation: *Lu, Yuchen, et al. "Countering language drift with seeded iterated learning." ICML 2020*
[7] Visual Question Answering: *Vani, Ankit, et al. "Iterated learning for emergent systematicity in vqa." ICLR 2021*
[8] Multilabel Classification: *Rajeswar, Sai, et al. "Multi-label iterated learning for image classification with label ambiguity." CVPR 2022*
[9] Color naming system: *Carlsson Emil, et.al "Iterated learning and communication jointly explain efficient color naming systems." arXiv 2023*
[10] Sys-gen: *Yi Ren, et al. "Improving compositional generalization using iterated learning and simplicial embeddings." NeurIPS 2023*
[11] Fortuitous forgetting: *Zhou, Hattie, et al. "Fortuitous forgetting in connectionist networks." ICLR 2022*
[12] Primacy bias in RL: *Nikishin, Evgenii, et al. "The primacy bias in deep reinforcement learning." ICML 2022*
[13] Iterated learning on CLIP: *Vani, Ankit, et al. "Iterated Learning Visual Programming" submitted to CVPR 2024*
[14] A patching for interaction phase: *Ren Yi, et.al "Language Model Evolution: An Iterated Learning Perspective", submitted 2024*
[15] Knowledge Distillation: *Furlanello, Tommaso, et al. "Born again neural networks." ICML, 2018*

SUMMARY

Understanding:

Iterated learning can gradually amplify the hidden bias of the intellegent agent.

Q1: How could IL amplify bias?
A1: Bayesian-IL framework

Q2: Where the bias comes from?
A2: Learning speed advantage of compositional mapping.

Applications:

From cognitive science to deep learning; from compositionality to more general bias.

Exp 1: human prefer compositionality → compositional language is achieved

Exp 2: simple NN prefer compositionality → compositional mapping is achieved

Exp 3: complex NN prefer systematicness→ systematical generalization is improved

Exp 4: LLM have different biases→ the bias (good and bad) are amplified

Future work:

More understanding, more applications, more efficient algorithm design.

Thank you for your time.

Q&A