Institut
des algorithmes
d'apprentissage
de Montréal

MILA

Countering Language Drift with Seeded Iterated Learning

Yuchen Lu

Content

Language Drift Problem

Iterated Learning for Language Evolution

Seeded Iterated Learning

Future Work

Introduction

In the past few years, great progress in many NLP tasks.

However supervised learning only maximize linguistic objective. It does not measure model's effectiveness, e.g., failing to achieve the tasks.

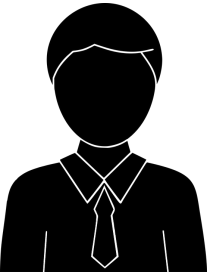
Responses for "Why do you like the ocean?"	**Score**
1. Why don't you like the ocean?	-1.70
2. I don't like the ocean. I love the ocean.	-2.66
3. I don't like the ocean. I love it.	-2.78
10. I don't like the ocean. I love the ocean. There's a difference.	-3.31

Supervised learning for pretraining, and finetune through interactions in a simulator

The Problem of Language Drift

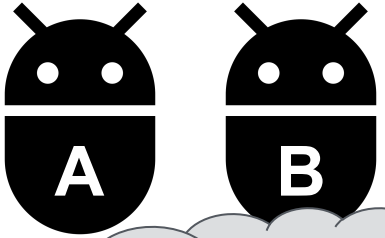
MILA

Step1: Collect Human Corpus



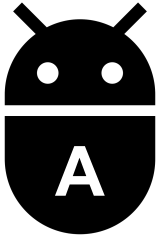
<Goal: Montreal, 7pm>
A: I need a ticket to Montreal.
B: What time?
A: 7 pm
B: Deal.
<Action: Book(Montreal, 7pm)>

Step2: Supervised Learning

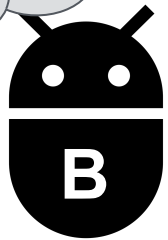


Step3: Interactive Learning (Self-Play)

Language Drift

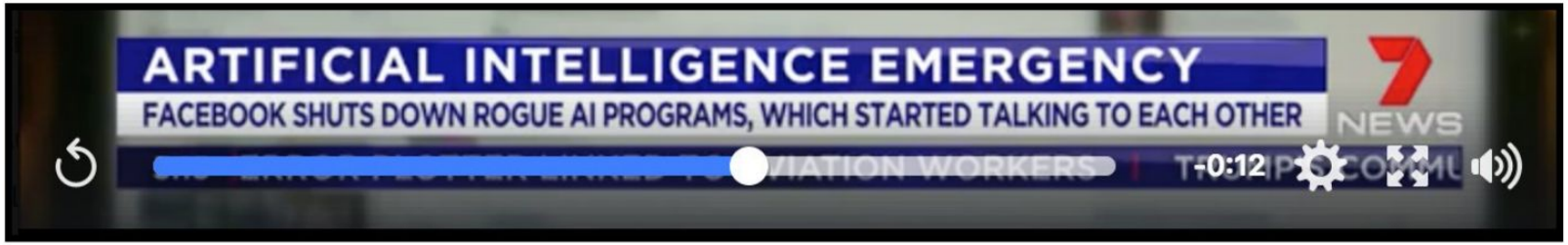


<Goal: Montreal, 7pm> A: I need a ticket to Paris. B: Wha time? A: pm 7 7 7 pm B: Deal. <Action: Book(Montreal, 7pm)>	<Goal: Toronto, 5am> A: I need need 5 am ticket B: Where A: Montreal B: Deal. <Action: Book(Toronto, 5am)>



Good Task Performance but **Poor Language**

Drift happens



Popular Latest **The Atlantic** Sign In

TECHNOLOGY

An Artificial Intelligence Developed Its Own Non-Human Language

When Facebook designed chatbots to negotiate with one another, the bots made up their own way of communicating.

ADRIENNE LAFRANCE JUNE 15, 2017

ROBOSTOP **Facebook shuts off AI experiment after two robots begin speaking in their OWN language only they can understand**

Experts have called the incident exciting but also incredibly scary

By James Beal and Andy Jehring
1st August 2017, 12:03 am | Updated: 2nd August 2017, 4:56 am

NEWS

Facebook AI project halted after bots invent new language

By Malek Murison - August 1, 2017

Existing Strategies: Reward Engineering

Use external labeled data to change the reward in addition to task completion

E.g., Visual Grounding (Lee et al. EMNLP 2019)

$$R_k^{\mathrm{G}} = \log p_B(\mathrm{De}_k|\overline{\mathrm{En}_k}) + \beta_G \log p_G(\mathrm{Img}_k|\overline{\mathrm{En}_k}).$$

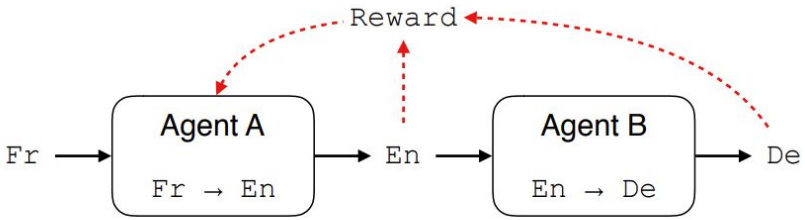


Figure 1: Diagram of our communication game.

Conclusion: The method is task-specific ✗

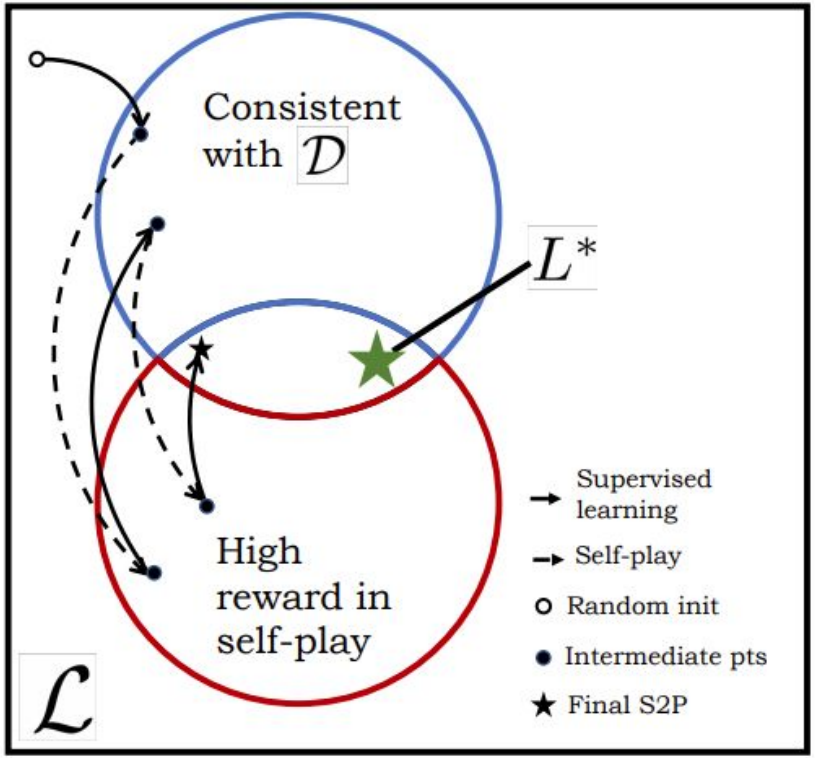
Existing Strategies: Population Based Methods

Community Regularization (Agarwal et al. 2019): For each interactive training steps, sample a pair of agents from the populations.

- Slower drift, but drift together
- Slower convergence of task progress with larger population size

Existing Strategies: Supervised-Selfplay (S2P)

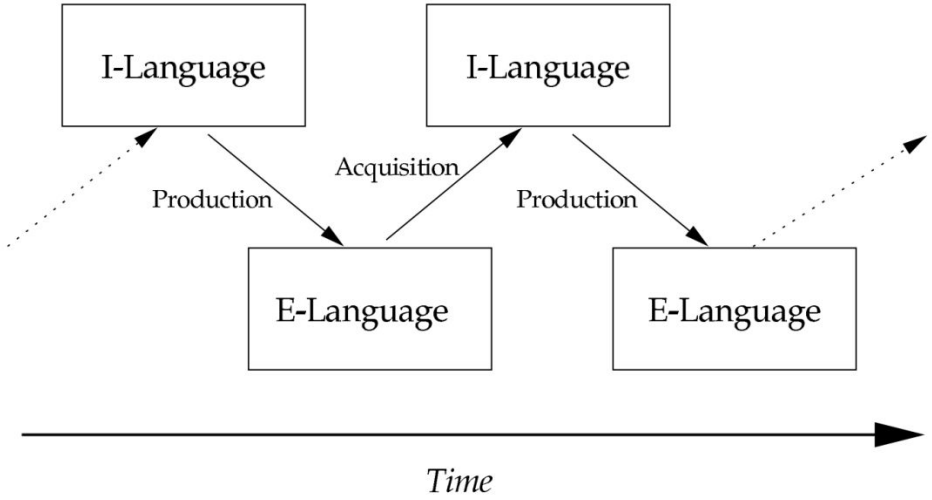
Mix supervised pretraining steps in interactive learning (Gupta & Lowe et al. 2019)



Current SOTA. Trade-off between task performance and language preservation

Content

Iterated Learning Model (ILM)



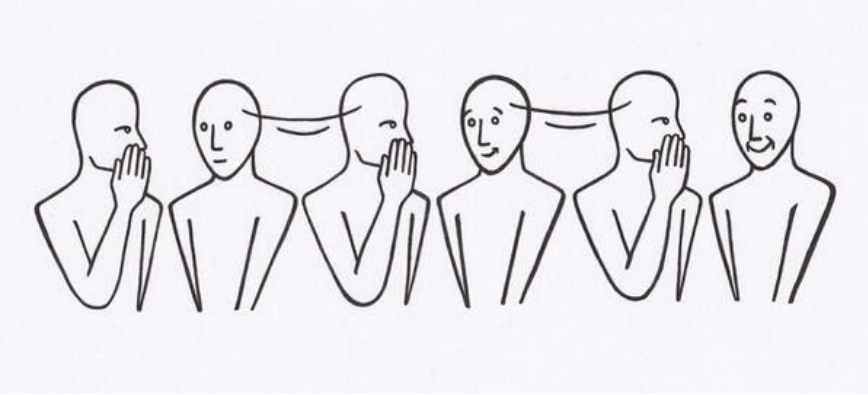
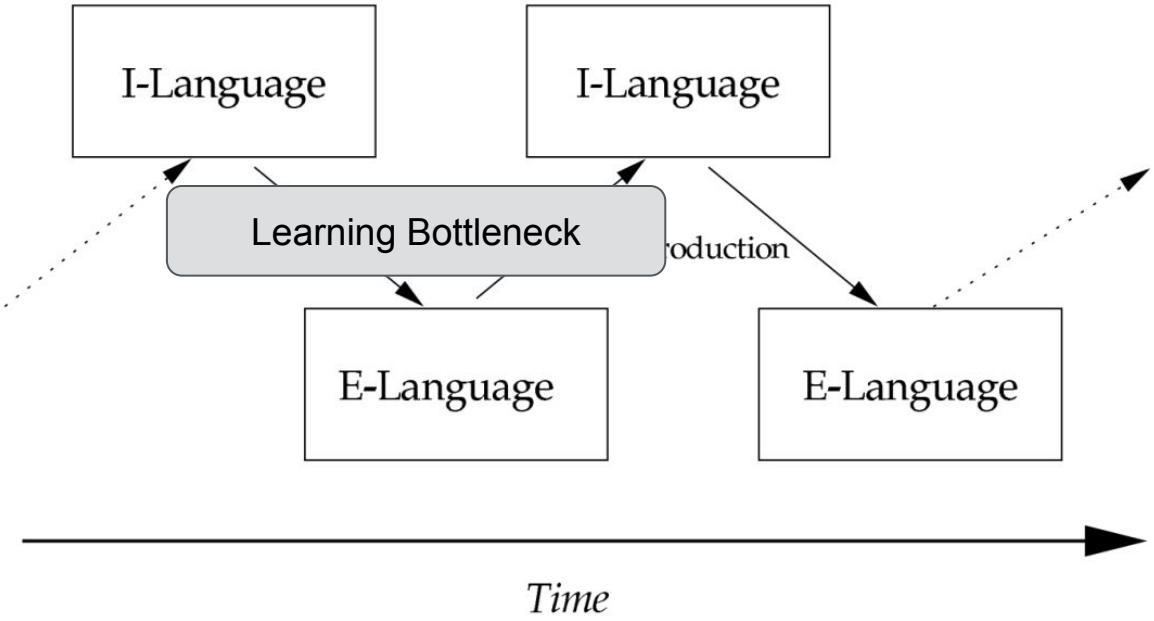


FIGURE 2.1 – Language transmission over time [32]. I-language is the internal language knowledge, while E-language is the external language like utterances.

Learning Bottleneck, aka *The Poverty of Stimulus*

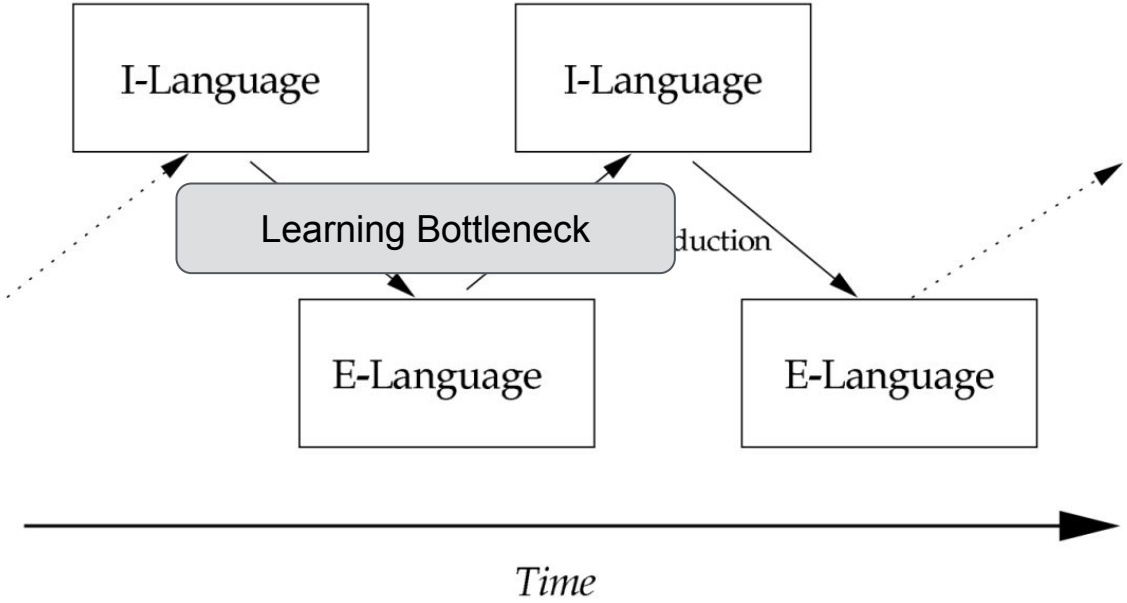
language learners must attempt to learn a **infinitely expressive linguistic system** on the basis of a relatively **small set of linguistic data**



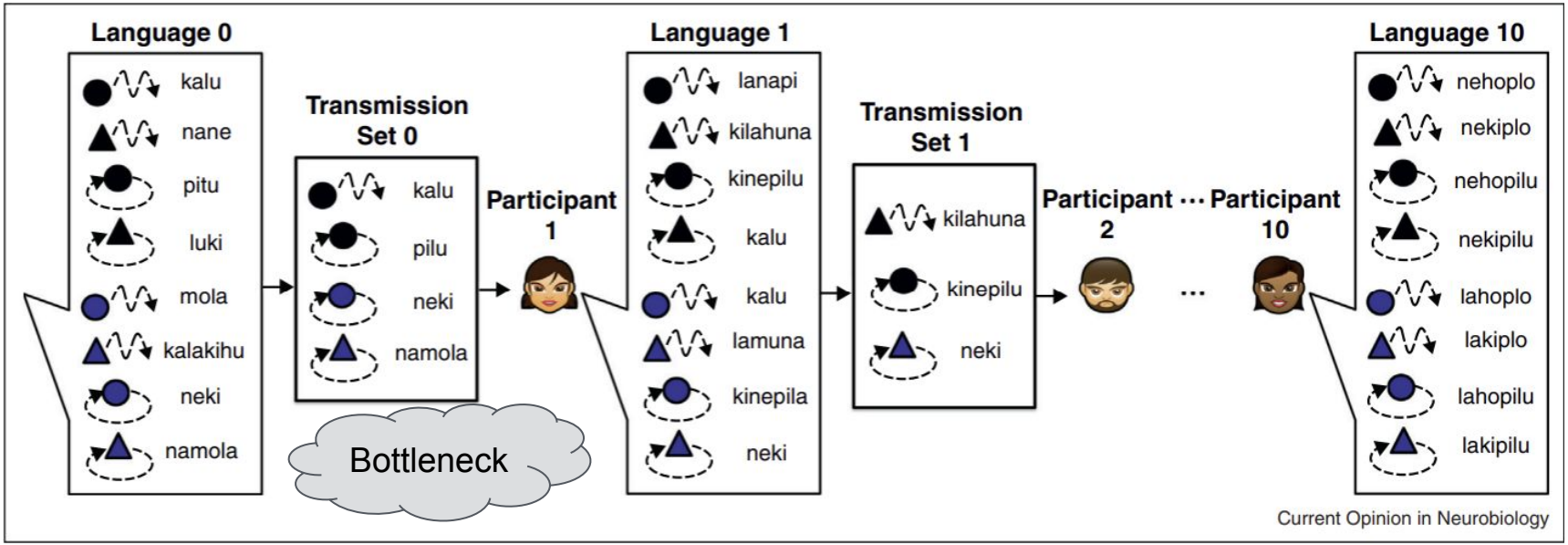
ILM predicts structured language

If a language survives such transmission process (I-Language converges), then I-language should be easy to learn even with a few samples of E-language.

ILM hypothesis: language structure is the adaptation to language transmission with bottleneck.



Iterated Learning: Human experiments



(Kirby et al. 2008 PNAS)

Generation 10: Somewhat compositional.

- ne- for black, la- for blue
- -ho- for circle, -ki- for triangle
- -plo for bouncing, -pilu for looping

Iterated Learning to Counter Language Drift?

ILM hypothesis: language structure is the adaptation to language transmission with bottleneck.

Maybe we can do the same during interactive training to regularize the language drift?

How should we properly implement the “Learning Bottleneck”?

Content

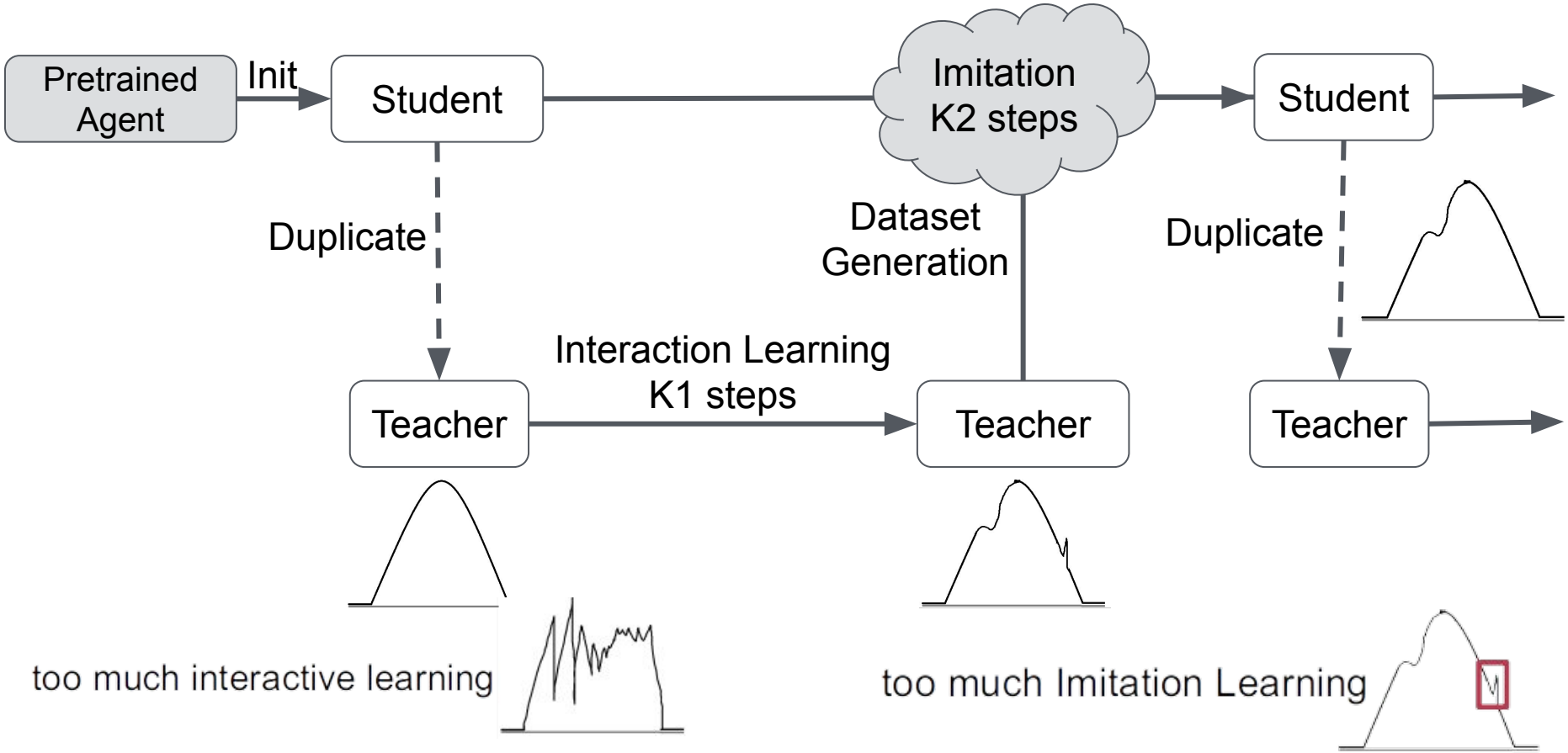
Language Drift Problem

Iterated Learning

Seeded Iterated Learning

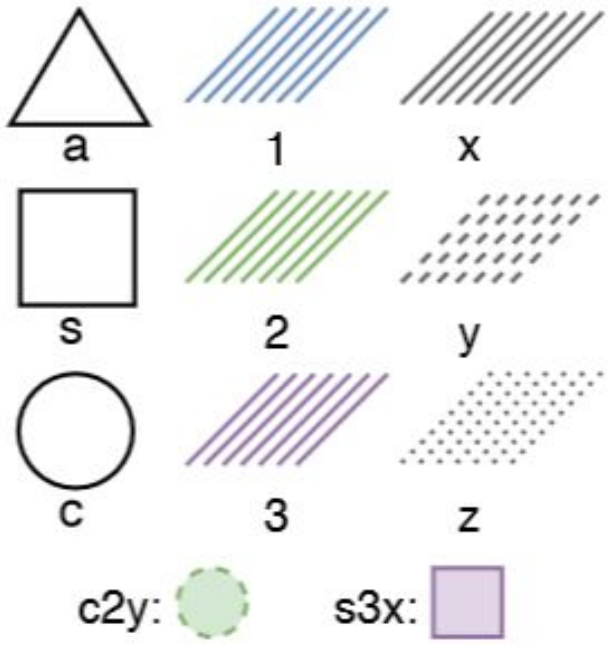
Future Work

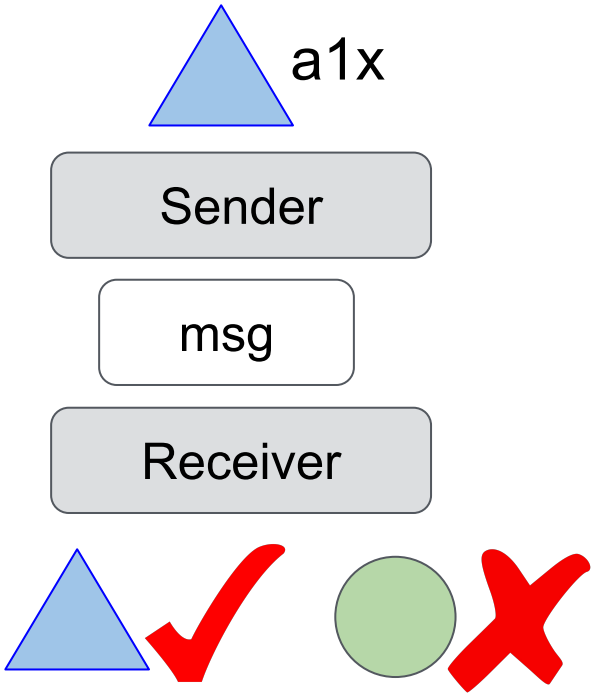
Seeded Iterated Learning (SIL)



Lewis Game: Setup

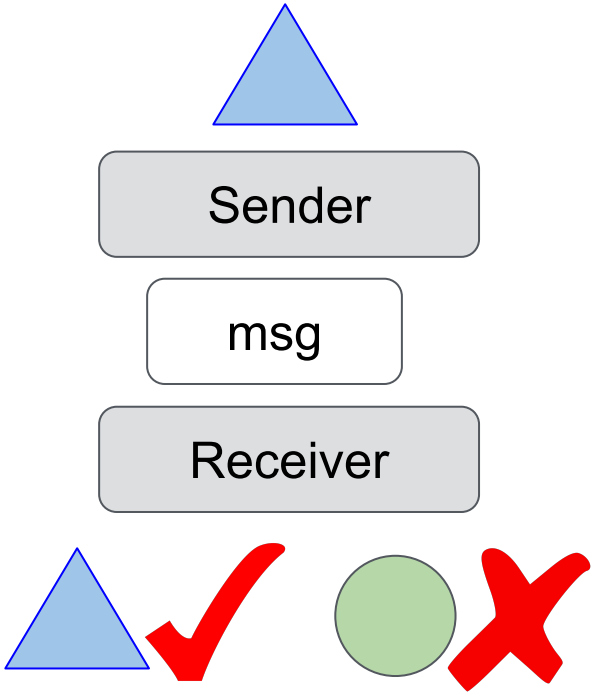
(Lewis, 1969 and Gupta & Lowe et al. 2019)



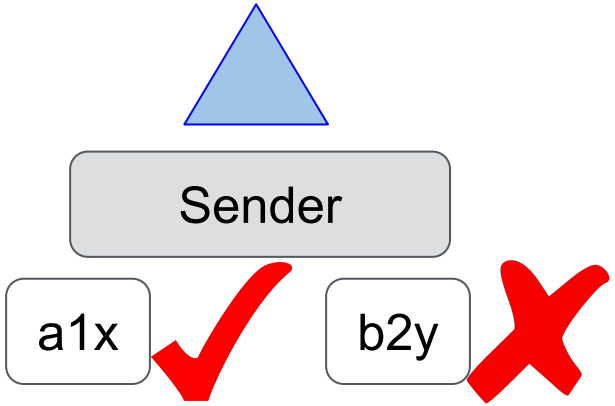


Lewis Game: Setup

(Lewis, 1969 and Gupta & Lowe et al. 2019)



Task Score

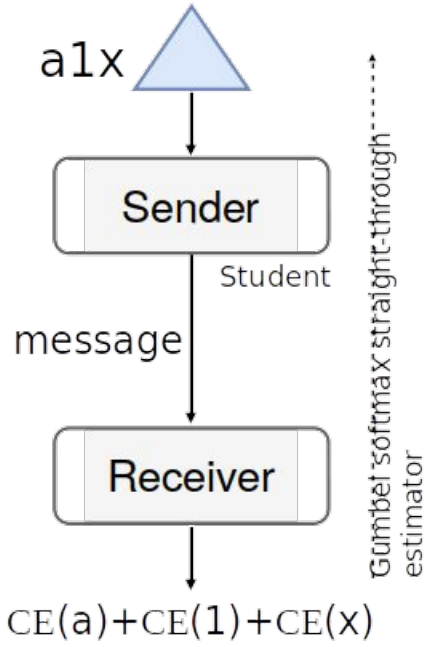


Language Score

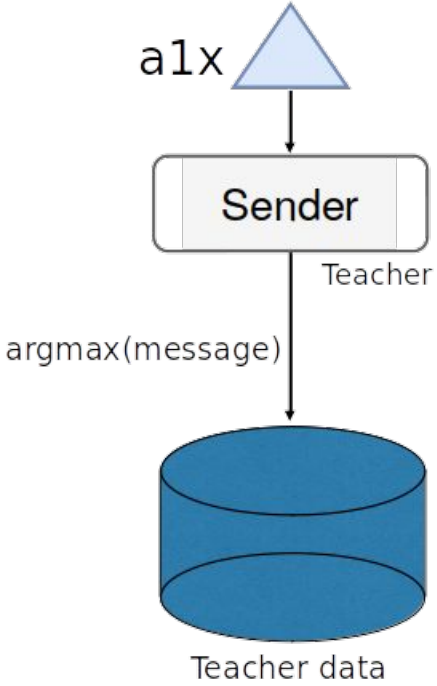
Evaluated on Objects unseen in interactive learning

SIL for Lewis Game

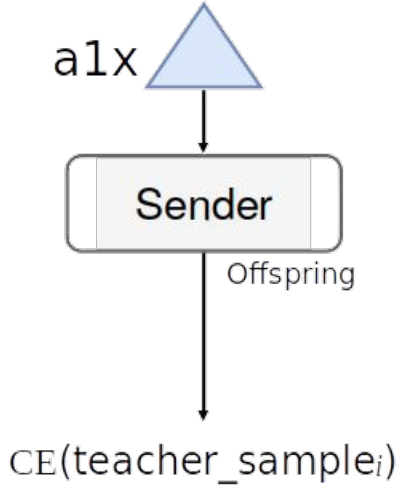
(Lewis, 1969 and Gupta & Lowe et al. 2019)



Interactive Learning Phase



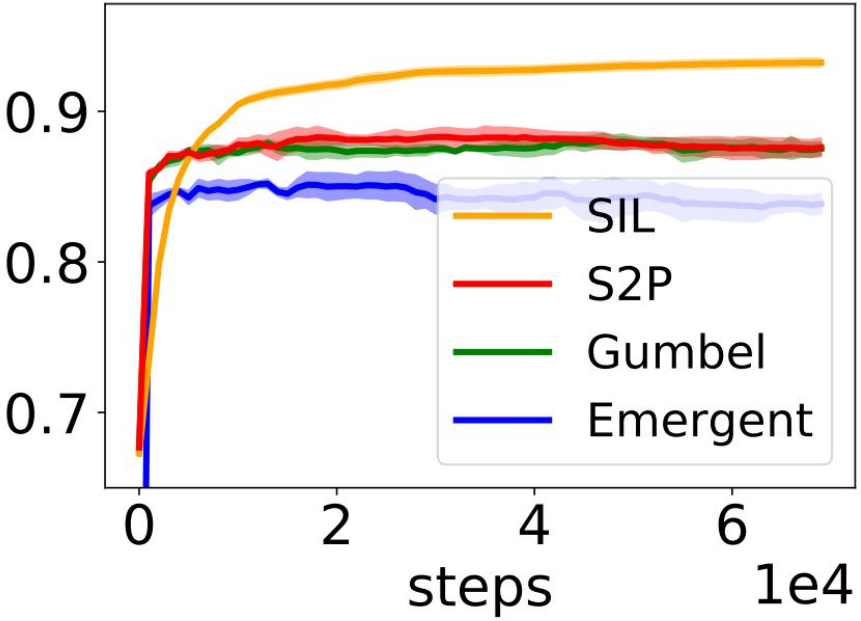
Data Generation



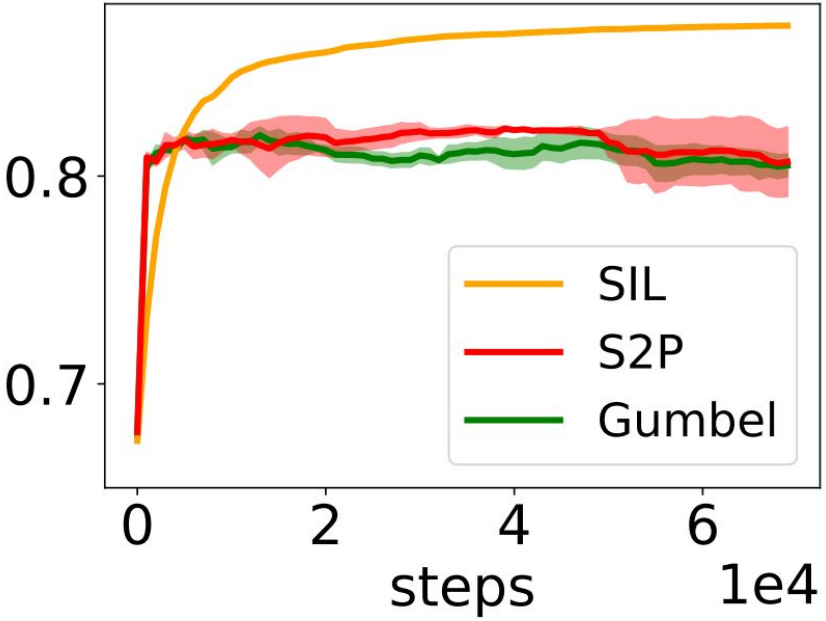
Imitation Learning Phase

Lewis Game: Results

X axis is the number of interactive training steps
Pretrain Task/Language score: 65~70%

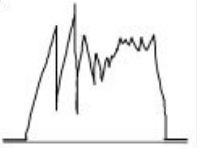


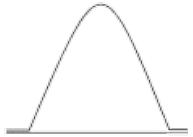
(a) Task Score

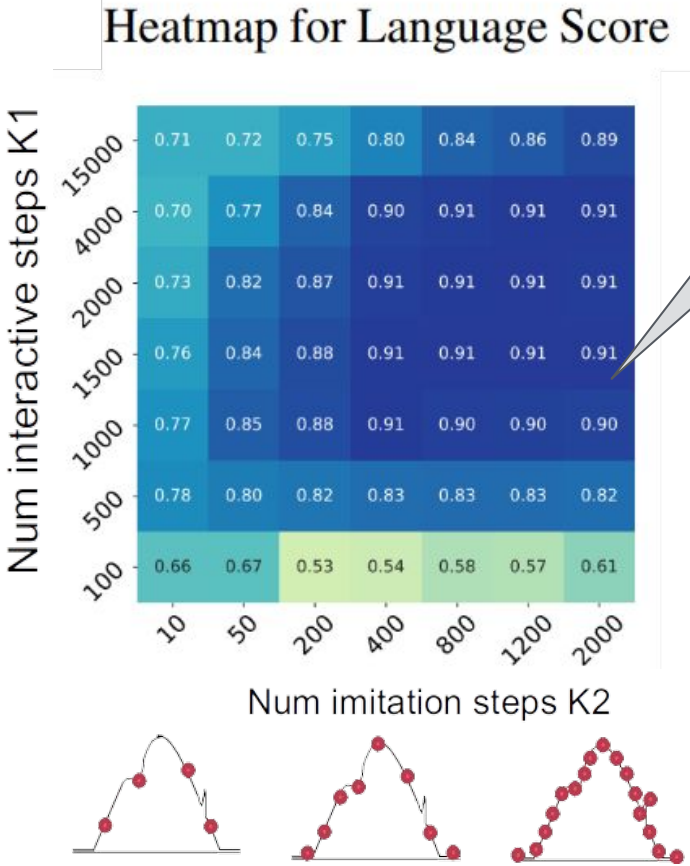


(b) Sender Language Score

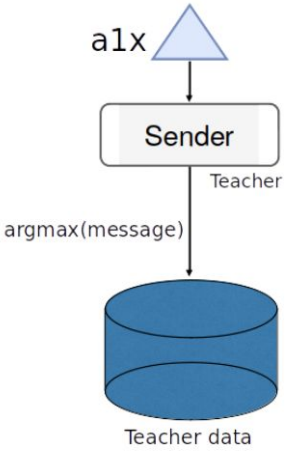
Lewis Game: K1/K2 Heatmap

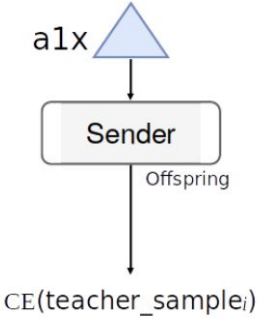






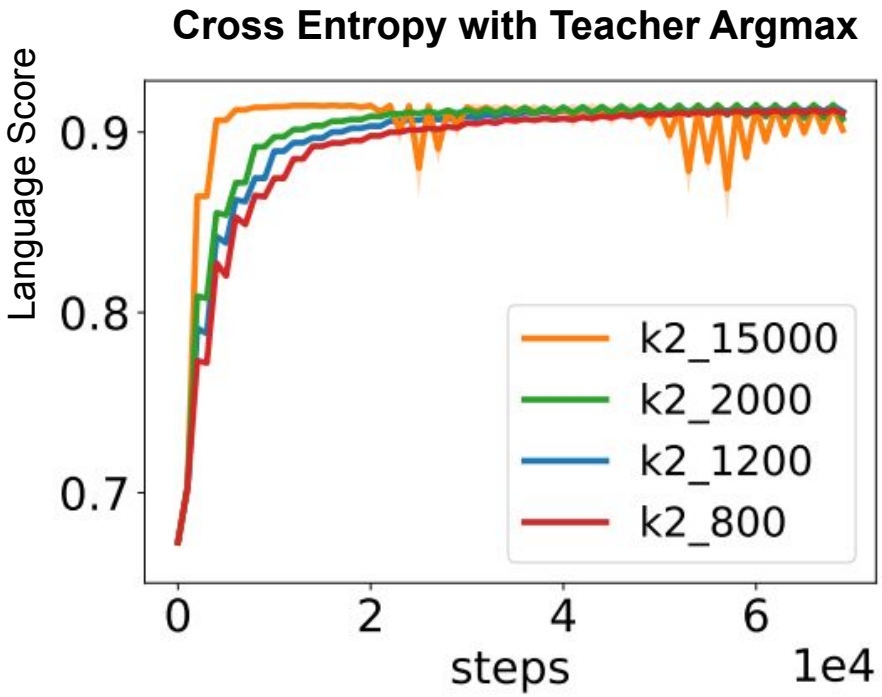
No Overfitting?

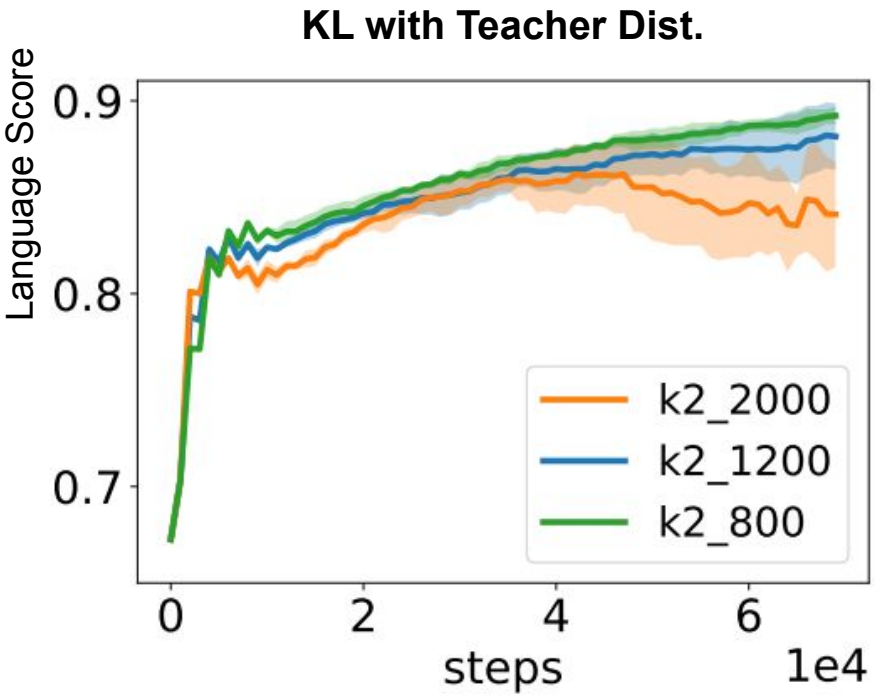




Lewis Game: Results

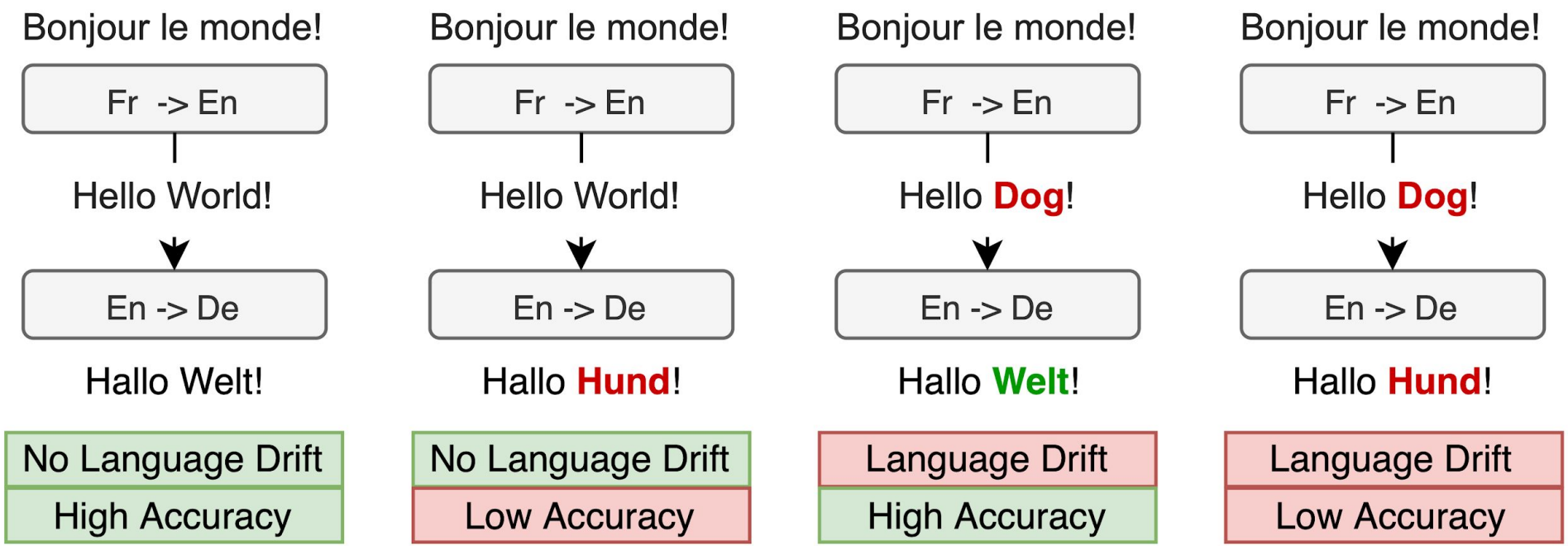
Data production is part of the “Learning Bottleneck”





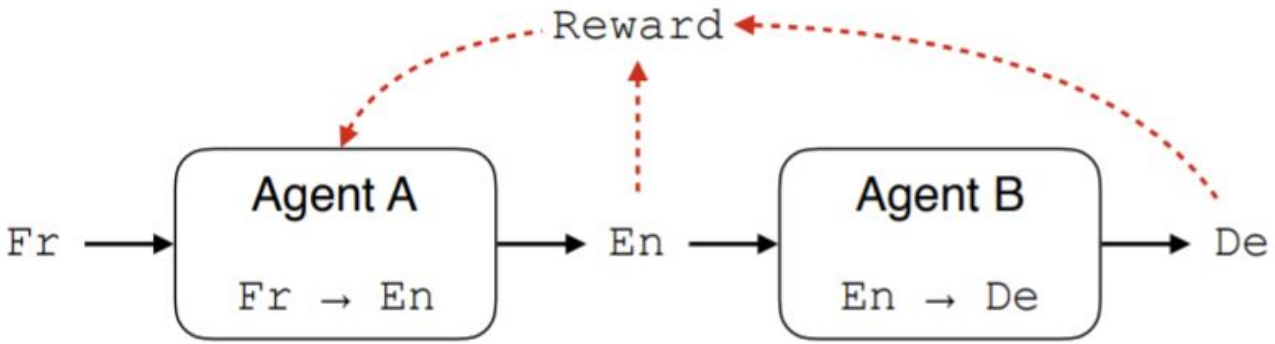
Translation Game: Setup

Lee et al. EMNLP 2019



Translation Game: Setup

Lee et al. EMNLP 2019



Task Score

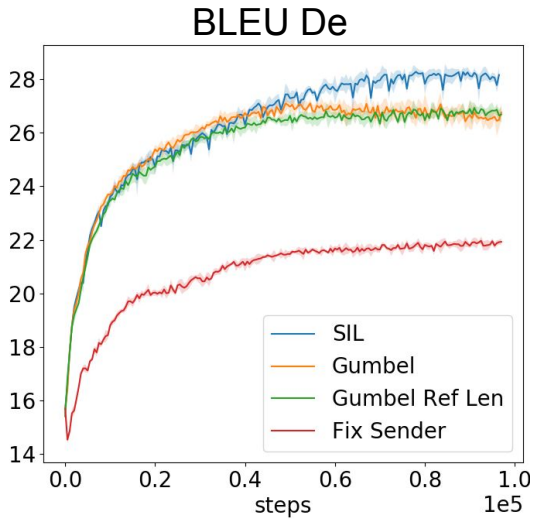
- BLEU DE (German BLEU score)

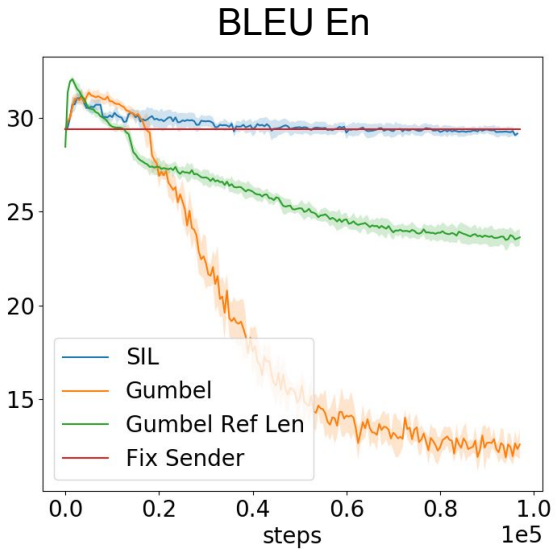
Language Score

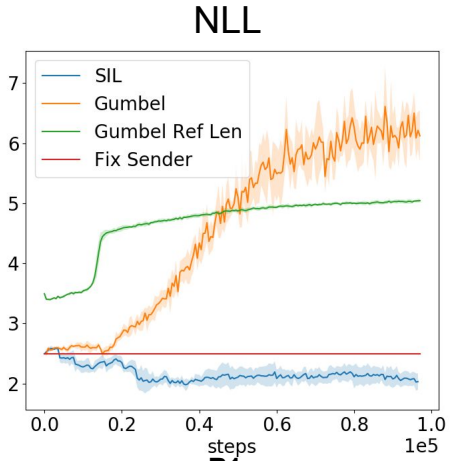
- BLEU EN (English BLEU score)
- English NLL of generated language a pretrained language model.
- R1 (Image retrieval accuracy from sender generated language)

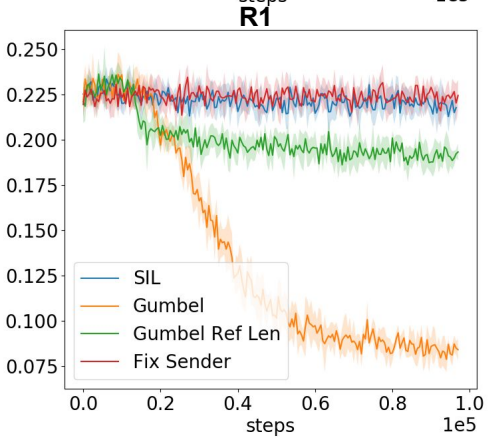
Translation Game: Baselines

NLL

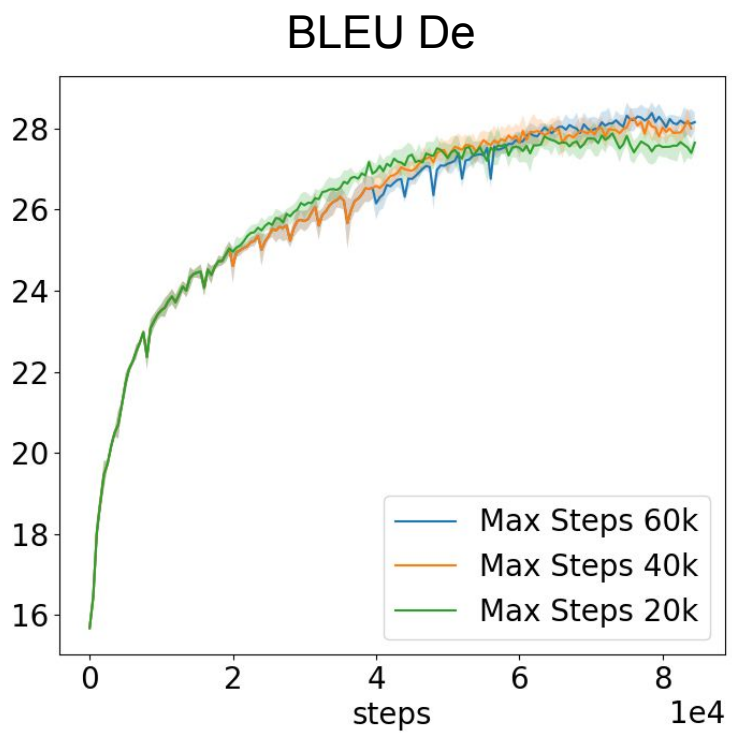


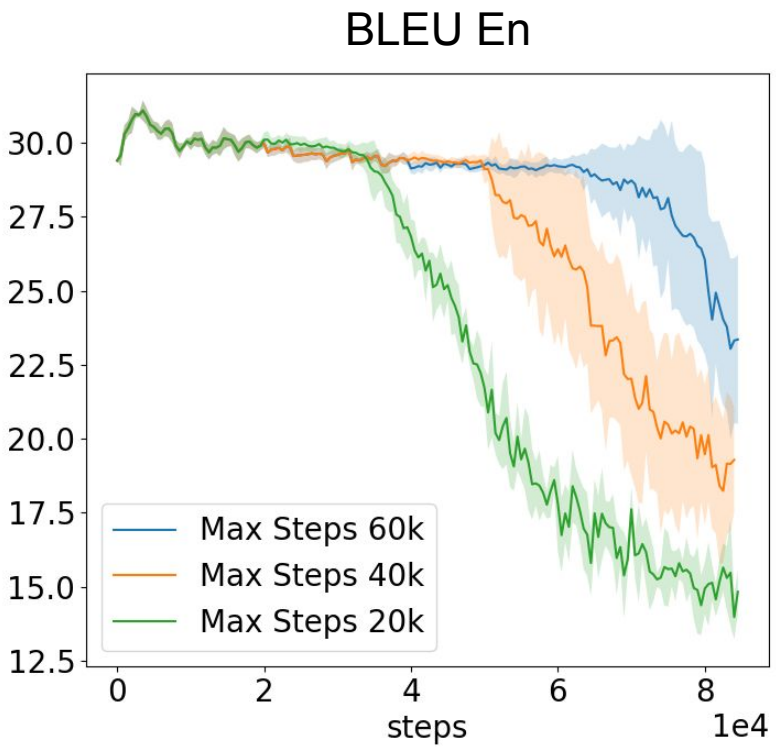




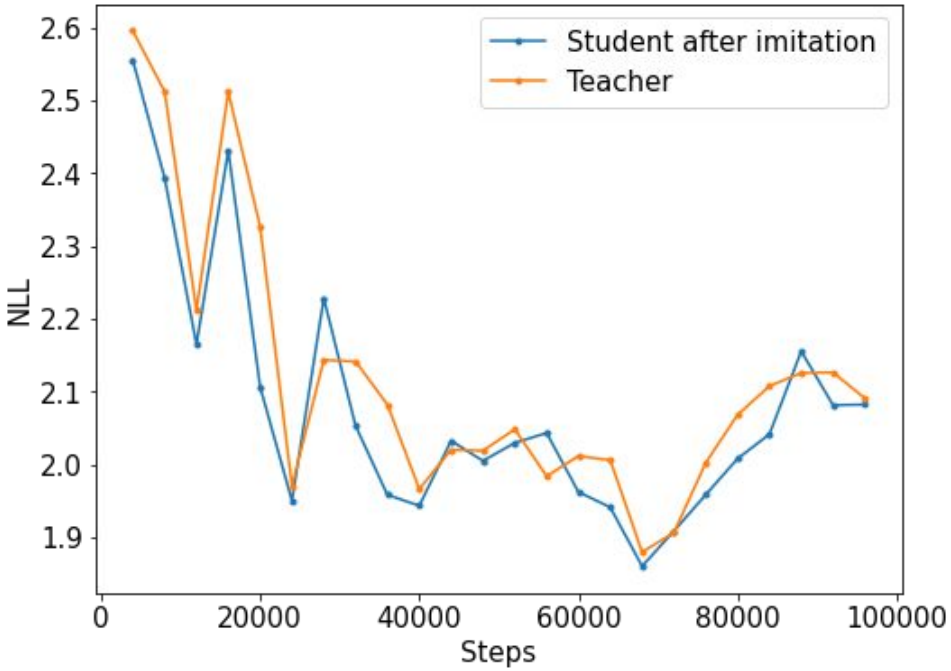


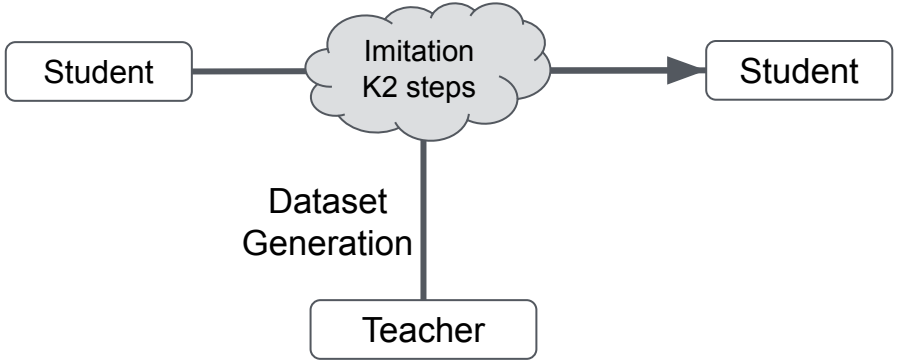
Translation Game: Effects of SIL

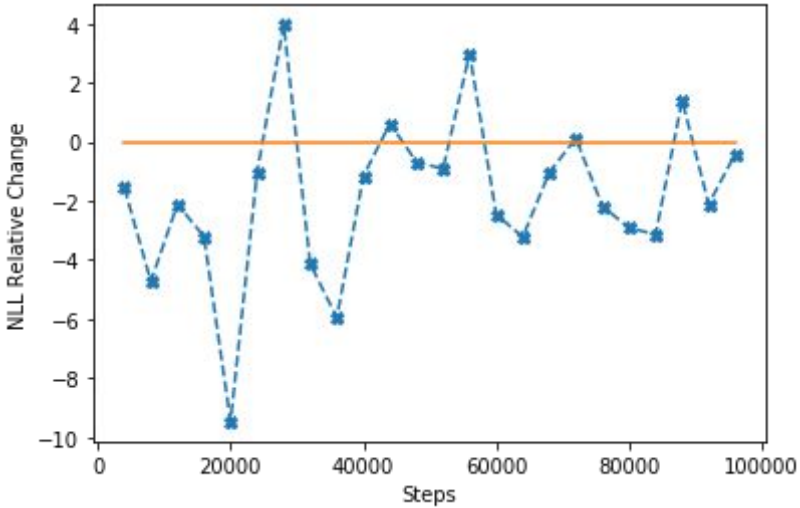




Effect of Imitation Learning

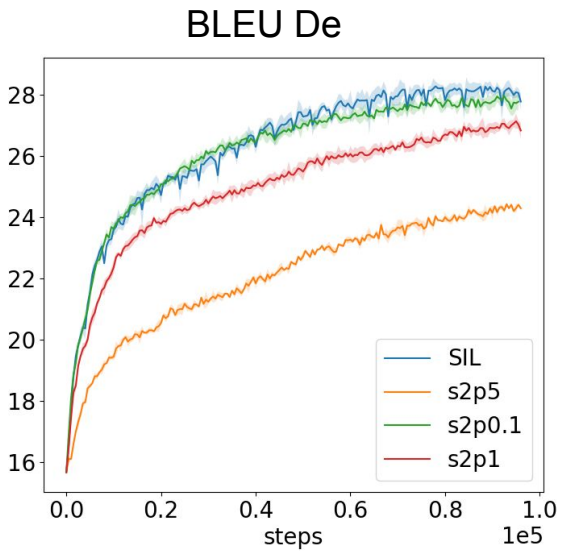


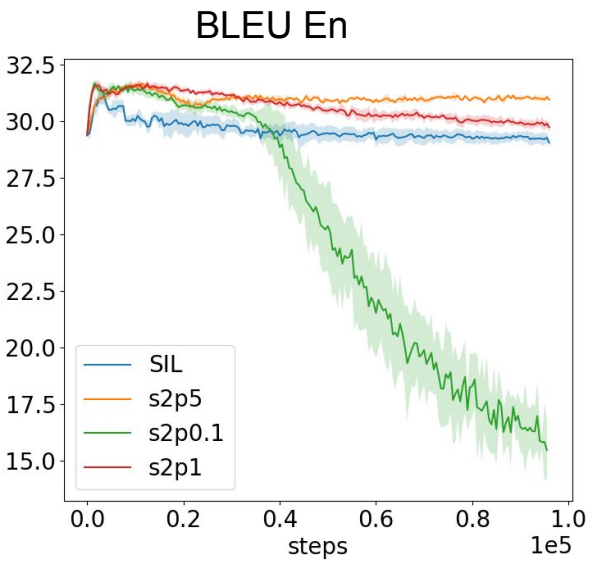


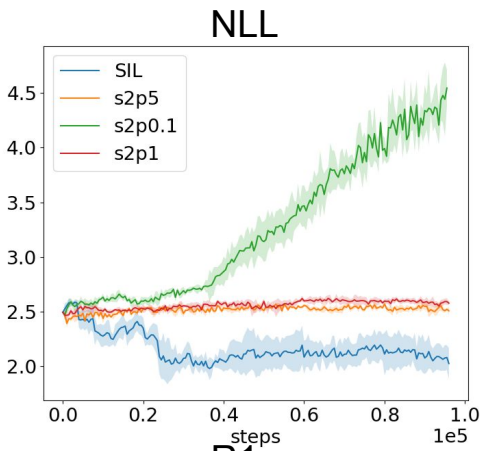


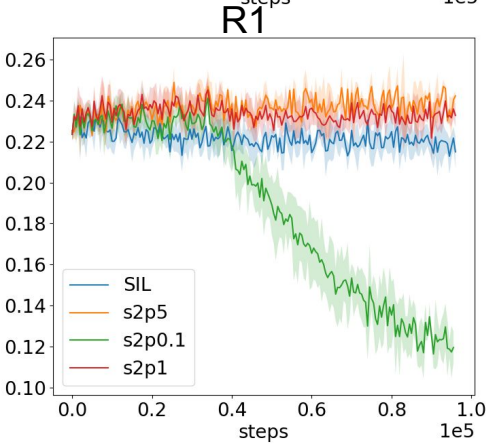
Mostly imitation learning brings the agent more favoured by pretrained language models

Translation Game: S2P





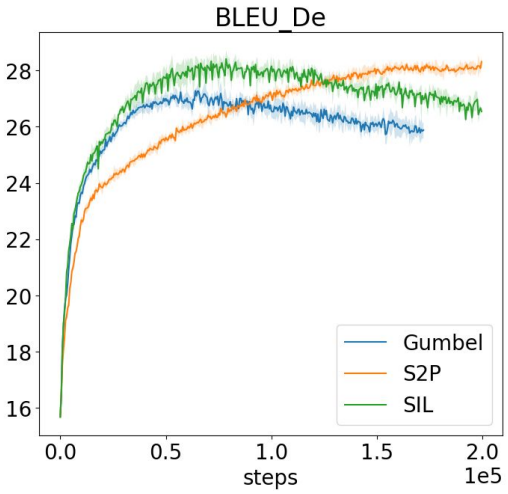


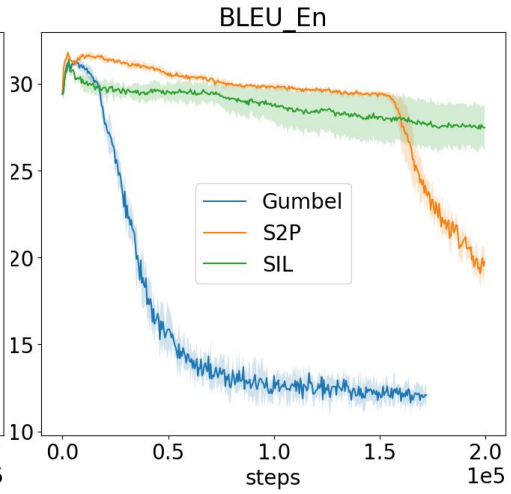


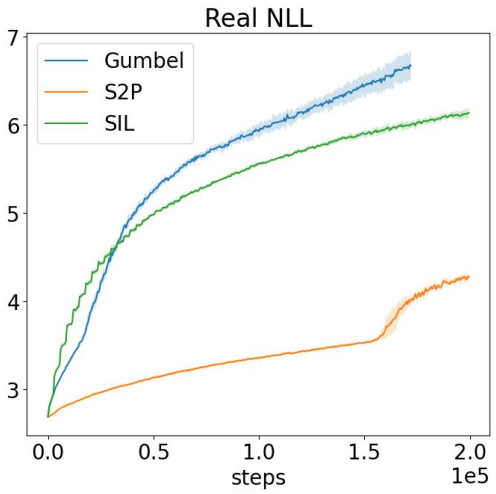
$$L_{S2P} = L_{Gumbel} + \alpha L_{supervised}$$

More on S2P and SIL...

After running for really long time...







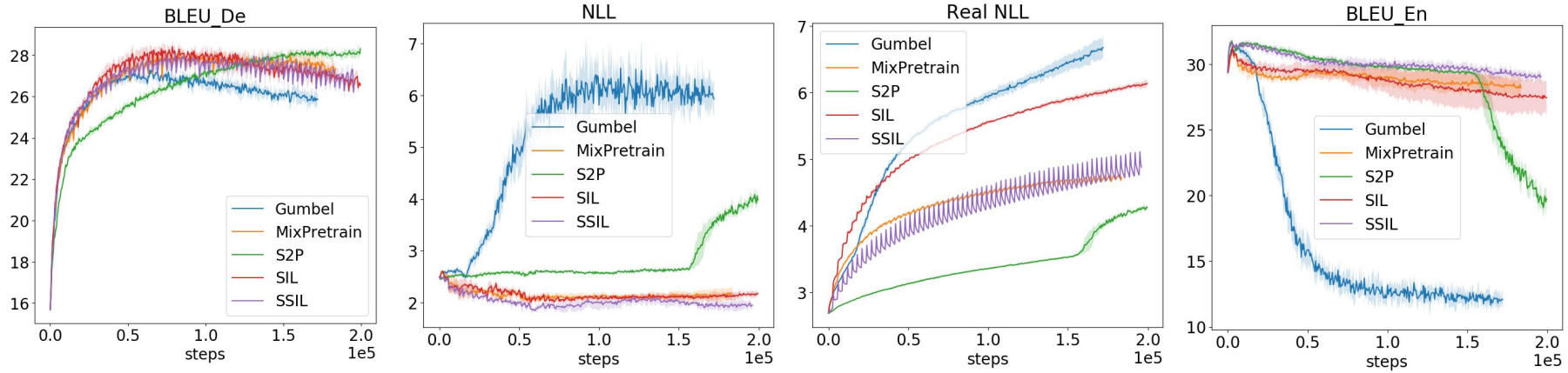
The NLL of the human language under the model. The lower the better

SIL and Gumbel reach the maximum task score and start overfitting, but S2P is very slow on task progress

S2P has a late stage collapse of language score (See BLEU En).

SIL is not able to model human data as good as S2P, which is trained to do so

SSIL: Combining S2P and SIL

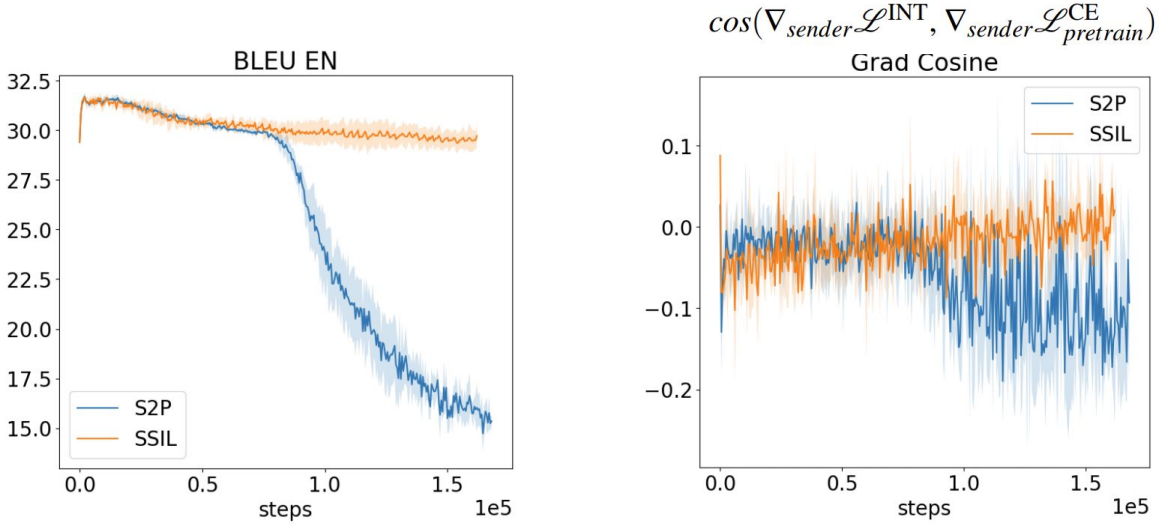


Finetuning Methods	*Training Losses*
Gumbel	$\mathcal{L}^{\text{INT}}$
S2P	$\mathcal{L}^{\text{INT}} + \alpha \mathcal{L}^{\text{CE}}_{pretrain}$
SIL (teacher) SIL (student)	$\mathcal{L}^{\text{INT}}$ $\mathcal{L}^{\text{CE}}_{teacher}$
SSIL (teacher) SSIL (student)	$\mathcal{L}^{\text{INT}} + \alpha \mathcal{L}^{\text{CE}}_{pretrain}$ $\mathcal{L}^{\text{CE}}_{teacher}$

SSIL is able to get best of both world.

MixPretrain is our another attempt by mixing human data and teacher data, but it is very sensitive to hyper-parameters with no extra benefits

Why late stage collapse?



(a) Bleu En

(b) Cosine Similarity

FIGURE 6.2 – Cosine similarity between the gradients issued from $\mathscr{L}^{\text{INT}}$ and $\mathscr{L}^{\text{CE}}_{pretrain}$. The collapse of the BLEU En matches the negative cosine similarity.

After adding iterated learning, reward maximizing is aligned to modelling human data

Summary

It is necessary to train in a simulator for goal-driven language learning.

Simulator training leads to language drift.

Seeded Iterated Learning (SIL) provides a “surprising” new method to counter language drift.

Content

Language Drift Problem

Iterated Learning

Seeded Iterated Learning

Future Work

Applications: Dialogue Tasks

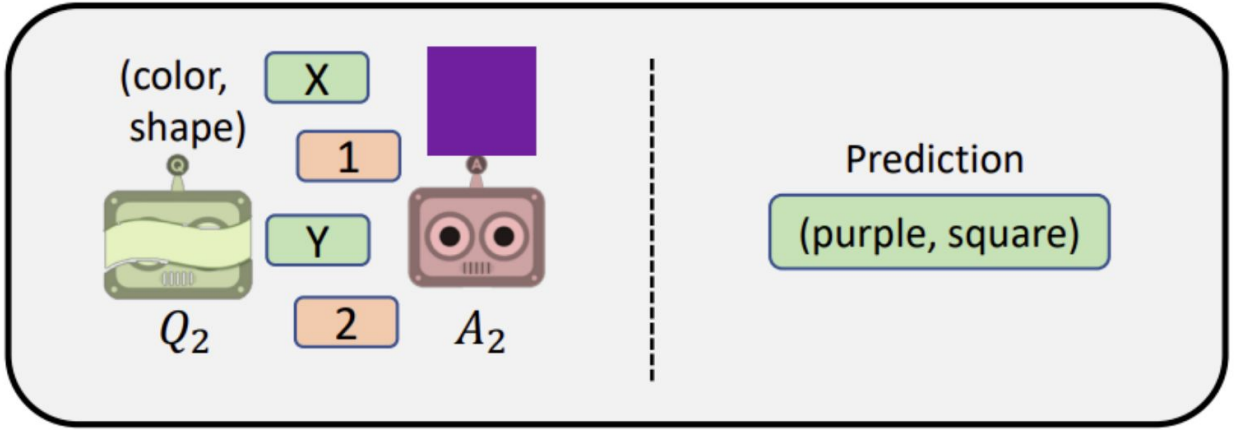


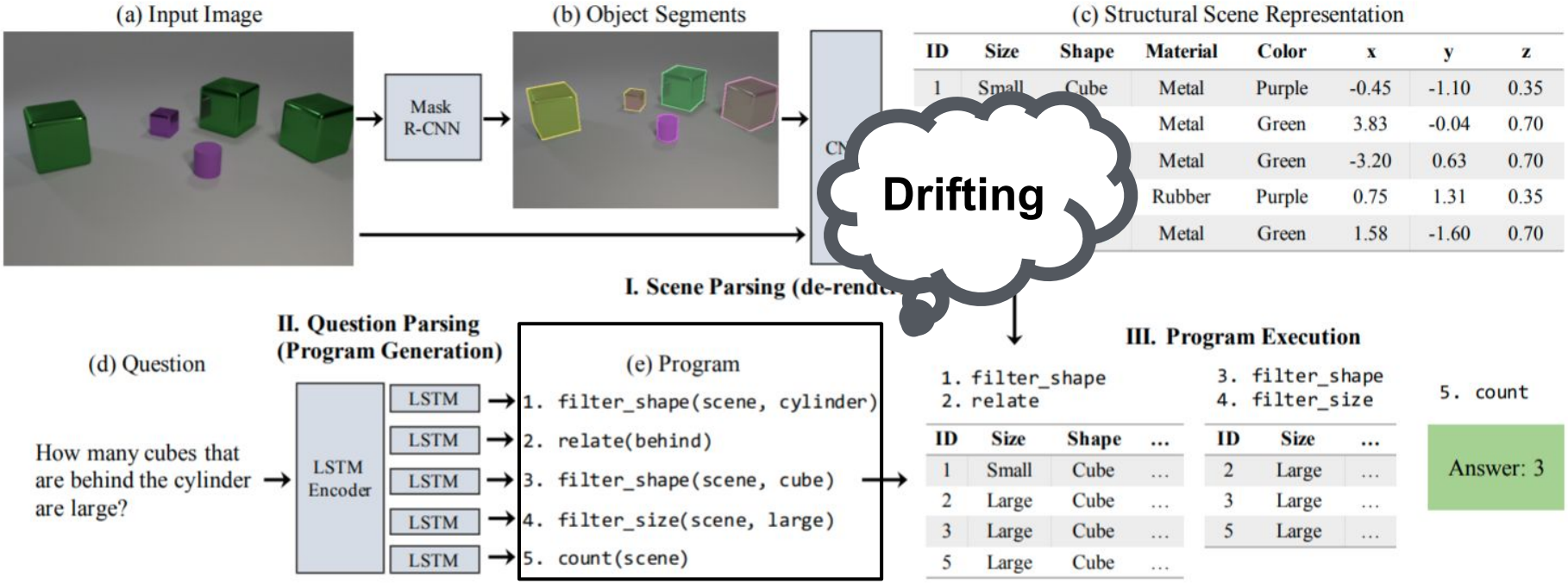
FIGURE 7.1 – Lewis toy dialog [13] might be suitable for studying language drift in goal-oriented dialogue settings.

Changing the student would induce a change of the dialogue context

More advanced imitation learning algorithm (e.g., DAGGER)

Applications: Beyond Natural Language

Neural Symbolic VQA (Yi, Kexin, et al. 2018)

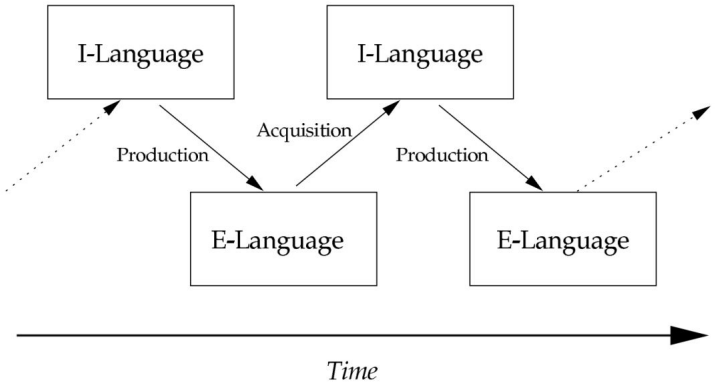


Iterated Learning for Representation Learning

Language survives transmission process

ILM Hypothesis

Language is structured



A representation survives transmission process

ILM for representation?

The representation is structured

Iterated Learning for Representation Learning

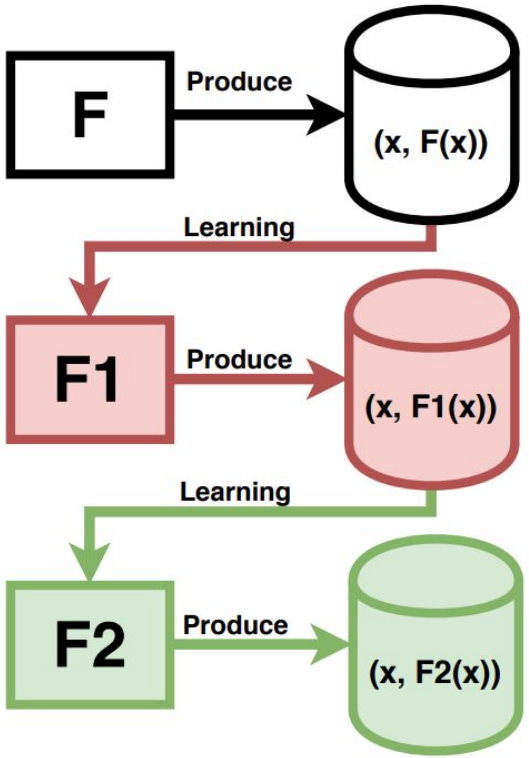


FIGURE 7.4 – Transmission process of representation

Each representation is a function f, mapping an input x into a representation f(x).

Construct a transmission process for n iteration. Each time a student learn on the dataset (x_train, f_i(x_train)) and become f_{i+1}. Repeat for n times.

Define representation **structureness** as the convergence after this chain

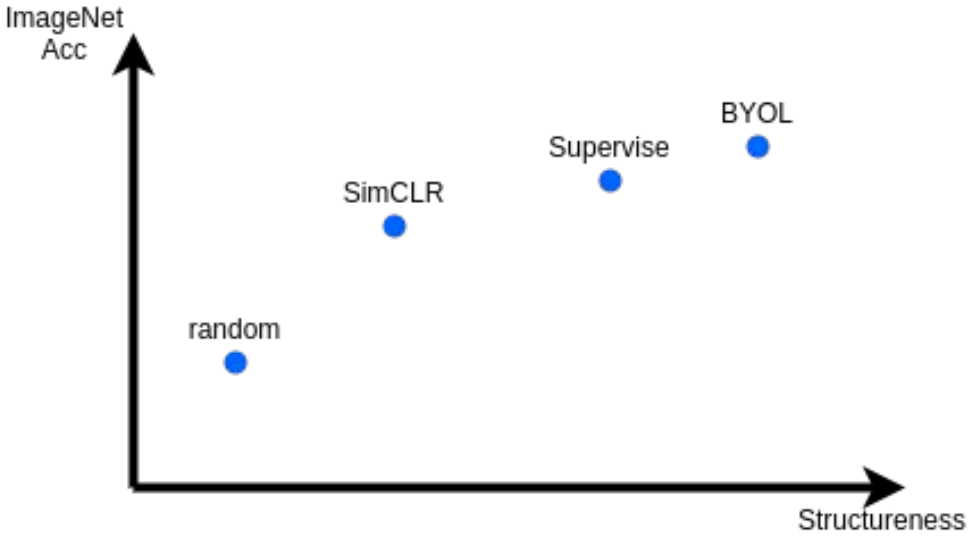
$$|f(x_{val}) - f_n(x_{val})|_2$$

Iterated Learning for Representation Learning

Define **structureness** as the convergence after this chain

$$|f(x_{val}) - f_n(x_{val})|_2$$

Hypothesis: Structureness correlates with the downstream task performance?



Co-Evolution of Language and Agents

Successful Iterated learning requires students to generalize from limited teacher data.

Whether the upper bound of this algorithm is related to the student architecture?

If yes, how should we address it?

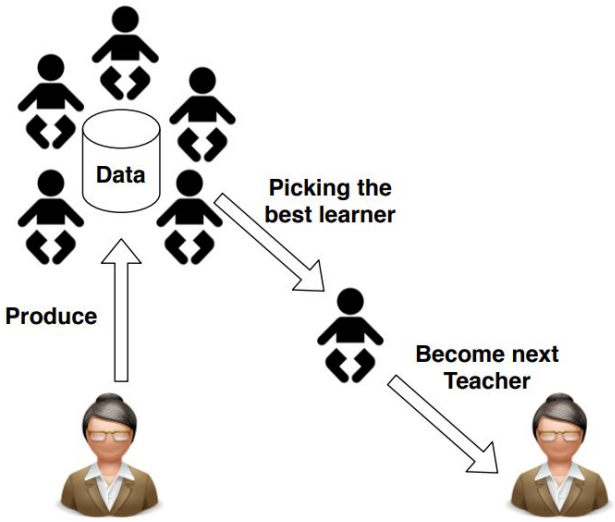


FIGURE 7.3 – Augmenting agent evolution to iterated learning framework. At the beginning of each imitation learning phase, the teacher generate the dataset and a population of agents attempt to learn on this dataset. We the pick the learner that has best generalization to become the next teacher. Different learners might have different architecture, so there is a architecture search loop in each imitation learning phase.

Summary

Iterated Learning provides future research directions on both applications and fundamentals for machine learning

Thanks!

"Human children appear preadapted to guess the rules of syntax correctly, precisely because languages evolve so as to embody in their syntax the most frequently guessed patterns. The brain has co-evolved with respect to language, but languages have done most of the adapting."

-Deacon, T. W. (1997). *The symbolic species*

Translation Game: Samples

SIL can remain close to the valid pretrained models

there are construction workers working hard on a project
there are workers working hard work on a project.
there are construction working hard on a project .....
there are workers working hard working on a project ..
there are workers working hard on a project .

SIL/S2P still drift when facing rare word occurrences (shaped lollipop)

a closeup of a child's face eating a blue , heart shaped lollipop.
a big one 's face plan a blue box.
a big face of a child eating a blue th-acof of of of chearts.....
a big face plan of eating a blue of the kind of hearts.
a big plan of a child eating a blue datadof the datadof the datadof the data@@

Translation Game: Human Evaluation (in progress)

Table 3. The Win-Ratio Results. The number in row X and column Y is the probability that method X beats method Y. Last column is a summation of previous columns, which can be viewed as the ranking scores of the methods.

Table 4. With French Sentences

	Gumbel	Pretrain	S2P	SIL	ref	SUM
Gumbel	0	0.27	0.16	0.12	0	0.55
Pretrain	0.73	0	0.5	0.35	0.15	1.73
S2P	0.84	0.5	0	0.45	0.23	2.02
SIL	0.88	0.65	0.55	0	0.22	2.31
Ref	1	0.84	0.77	0.77	0	3.39
Ranking	Ref, SIL, S2P, Pretrain, Gumbel					

Table 5. Without French Sentences

	Gumbel	Pretrain	S2P	SIL	ref	SUM
Gumbel	0	0.12	0.05	0.09	0	0.26
Pretrain	0.88	0	0.73	0.57	0.15	2.33
S2P	0.95	0.26	0	0.36	0.05	1.62
SIL	0.91	0.43	0.64	0	0.02	2.00
Ref	1	0.85	0.95	0.98	0	3.78
Ranking	Ref, Pretrain, SIL, S2P, Gumbel					

Translation Game: Samples

SIL successfully prevent language drift

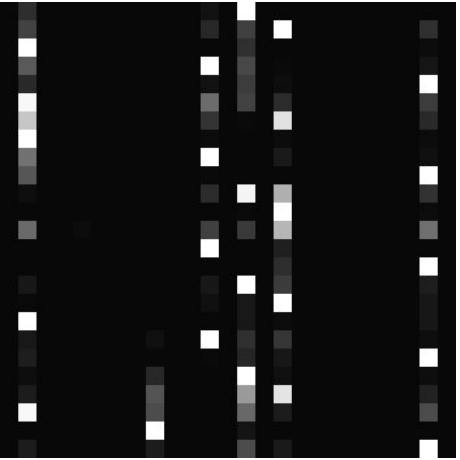
Human	**two men, one in blue and one in red, compete in a boxing match.**
Pretrain	two men, one in blue and the other in red, fight in a headaching game
Gumbel	two men one of one in blue and the other in red cfighting in a acacgame.....
S2P	two men, one in blue and the other in red, fighting in a kind of a kind.
SIL	two men, one in blue and the other in red, fighting in a game.

SIL partially recovers the sentence without drifting

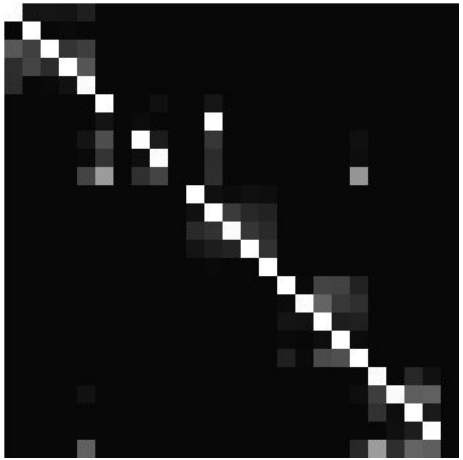
Human	**a group of friends lay sprawled out on the floor enjoying their time together.**
Pretrain	a group of friends on the floor of fun together.
Gumbel	a group of defriends comadeof on the floor together of of of of of together.....
S2P	a group of friends of their commodities on the floor of fun together.
SIL	a group of friends that are going on the floor together.

Lewis Game: Sender Visualization

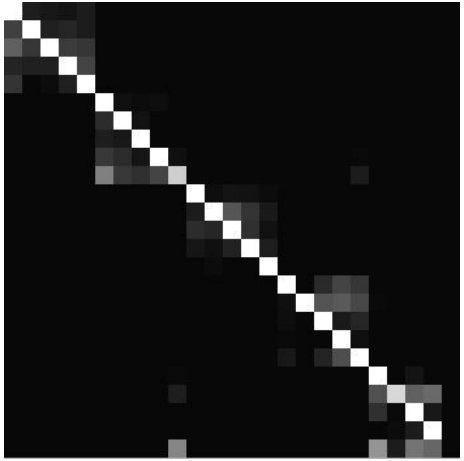
Row: Property Values
Col: Words



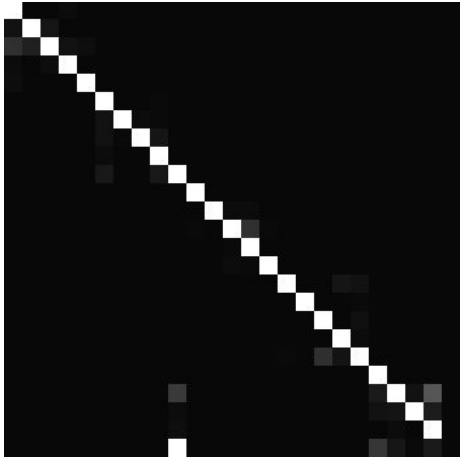
Emergent Communication



Std. Interactive Learning



S2P



SIL

Iterated Learning in Emergent Communication

Li, Fushan, and Michael Bowling. "Ease-of-teaching and language structure from emergent communication." *Advances in Neural Information Processing Systems*. 2019.

Guo, Shangmin, et al. "The Emergence of Compositional Languages for Numeric Concepts Through Iterated Learning in Neural Agents." *arXiv preprint arXiv:1910.05291* (2019).

Ren, Yi, et al. "Compositional Languages Emerge in a Neural Iterated Learning Model." *arXiv preprint arXiv:2002.01365* (2020).

Introduction

Agents that can converse **intelligibly** and **intelligently** with humans is a long standing goal.

On specific narrowly scoped applications, progress has been good.

… But on more open ended tasks where it is difficult to constrain the natural language interaction, progress has been less good.

Not Limited in Natural Language

Neural Module Networks for QA (Gupta, Nitish, et al. 2019)