LAB MEETING 24/03/25

Optimal Task Selection in Meta Curriculum Learning

Andrea Burda, Haopeng Chen, Tom Verguts

TABLE OF CONTENTS

01 Theoretical basis

02 Goal of the project

03 Method

04 Results

05 Next objectives

06 References

Theoretical basis

What is learning?

Learning is the **acquisition of knowledge** through **experience**, in an uncertain **environment**

(Silver et al., 2021)

How do we learn?

Guidance of a **teacher**

Complex sequences are **broken down** into simpler tasks

Meaningful order

Based on **feedback**

In humans: **Internal teacher** (autonomous organisation)

(Bengio et al., 2009; Tong et al., 2023)

Why do we learn this way?

REWARD

→ Efficient learning

Reinforcement learning agent : learns how to maximise reward through trial and error exploration

Curiosity and spontaneous exploration

(Forestier et al., 2022; Silver et al., 2021; Sutton & Barto, 1998; Ten et al., 2021)

The curriculum strategy

Curriculum: handpicked tasks

The **teacher learns** how to maximise learning by exploring the values of actions

Curriculum learning

Filtering

Chaining

1 → 2 → 3

(Bengio et al., 2009; Tong et al., 2023)

The curriculum strategy

Curriculum: handpicked tasks

The **teacher learns** how to maximise learning by exploring the values of actions

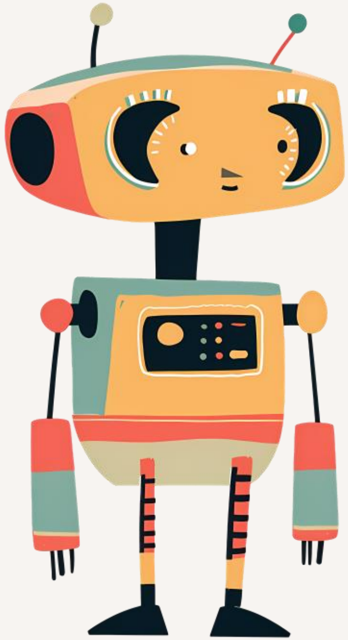
Curriculum learning

Which principles/criteria?

Meta curriculum learning

(Bengio et al., 2009; Tong et al., 2023)

Benefits of a curriculum on machine learning



Faster convergence

Higher **quality** local minima

(Bengio et al., 2009)

Study by Bengio and colleagues (2009):

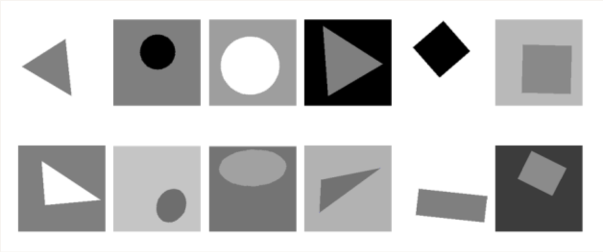


Fig. 1 Basicshapes (top) and Geomshapes (bottom) datasets example

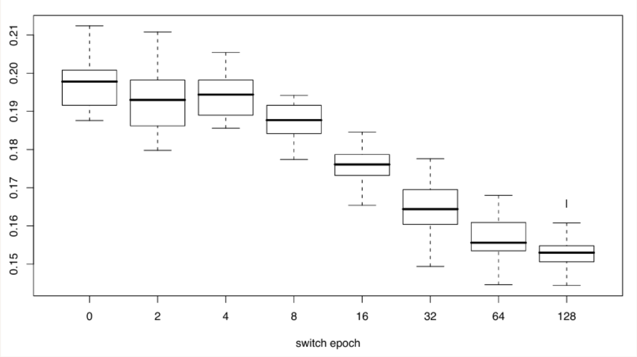


Fig. 2 Test classification error distribution as a function of the switch epoch

Benefits in machines

A comparison by Wu and colleagues (2021):

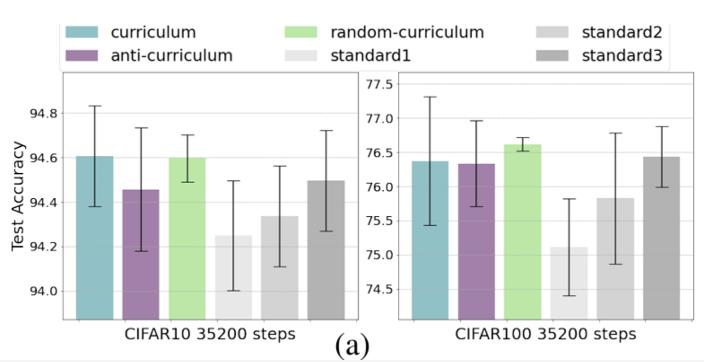


Fig. 1 Curriculum shows little benefit for standard machine learning

(Wu et al., 2021)

Benefits in machines

A comparison by Wu and colleagues (2021):

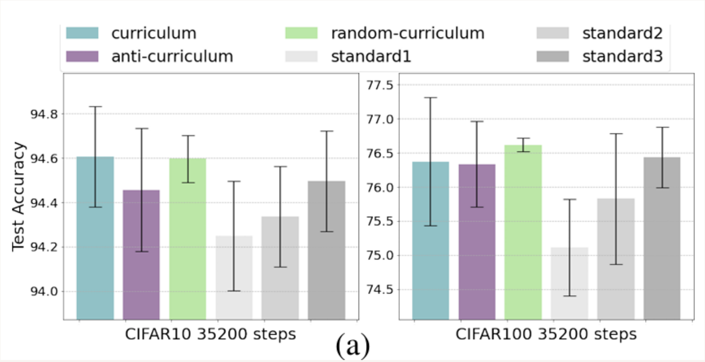


Fig. 1 Curriculum shows little benefit for standard machine learning

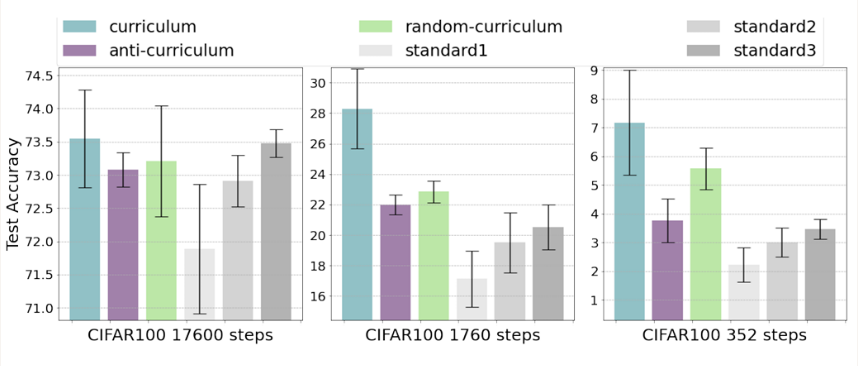


Fig. 2 Curriculum helps when the time budget for training is limited

(Wu et al., 2021)

Benefits in machines

A comparison by Wu and colleagues (2021):

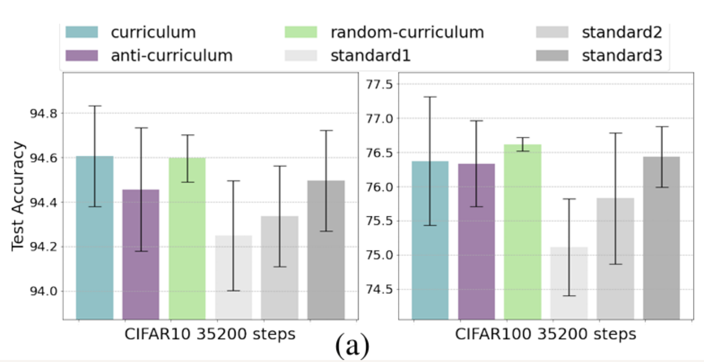


Fig. 1 Curriculum shows little benefit for standard machine learning

Curriculum **benefits are situational!**

(Wu et al., 2021)

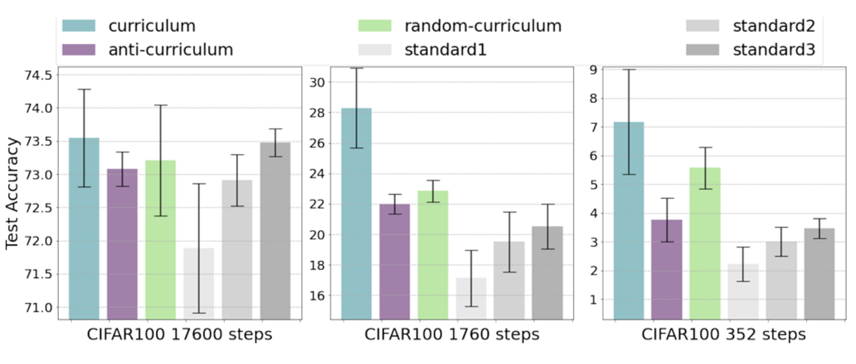


Fig. 2 Curriculum helps when the time budget for training is limited

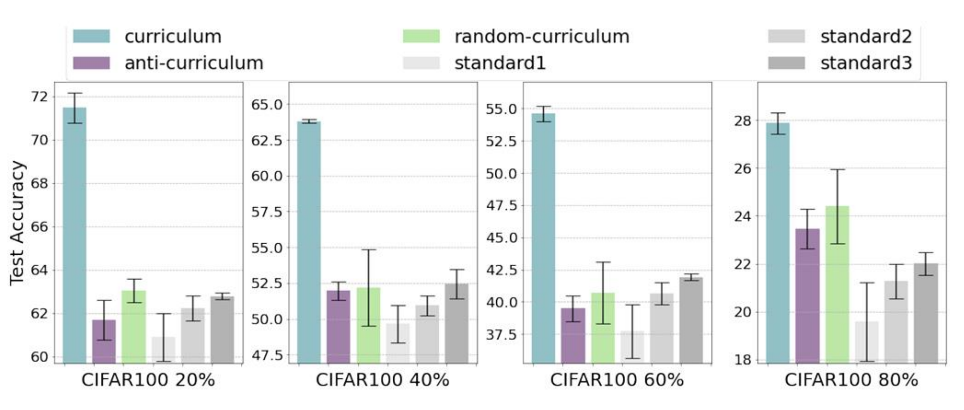


Fig. 3 Curriculum helps when training with noisy data (% of label noise)

Goals of the project

Goals of the project

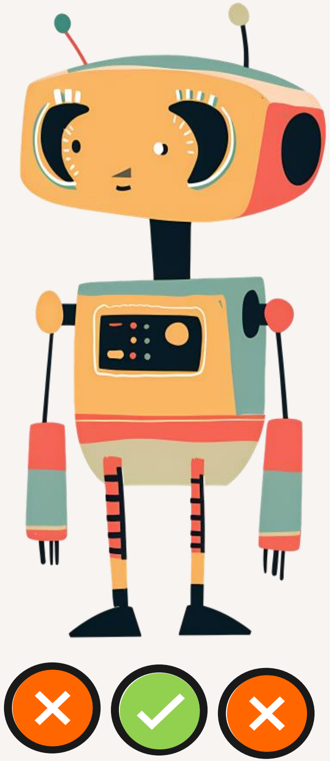
A multi-level computational model

Which task selection strategy allows for optimal learning?

- Reinforcement learning computational model
- Meta-curriculum learning

In the future:

Which task selection strategy do humans use for learning?



Methods

Three Tasks

Name	AND	XOR (Exclusive OR)	Random mapping
Characteristics	Linear logical relationship	Non-linear logical relationship	No relationship
Difficulty	Easy	Hard	Impossible

Three Tasks

AND

x1	x2	y
0	0	0
1	0	0
0	1	0
1	1	1

	0	1
0	0	0
1	0	1

XOR

x1	x2	y
0	0	0
1	0	1
0	1	1
1	1	0

	0	1
0	0	1
1	1	0

Three Tasks

RANDOM

x1	x2	y
0	0	?
1	0	?
0	1	?
1	1	?

	0	1
0	?	?
1	?	?

A multi-level computational model

LEVEL I: The Student

LEVEL II: The Teacher

LEVEL III: The Coordinator

Level 1: The Student

A multi-level computational model

ROLE: Learns the tasks

OBJECTIVE: Minimise loss

MODEL: Supervised learning model

Level 1: The Student

A multi-level computational model

TASK

0	0
1	0
0	1
1	1

CONTEXT

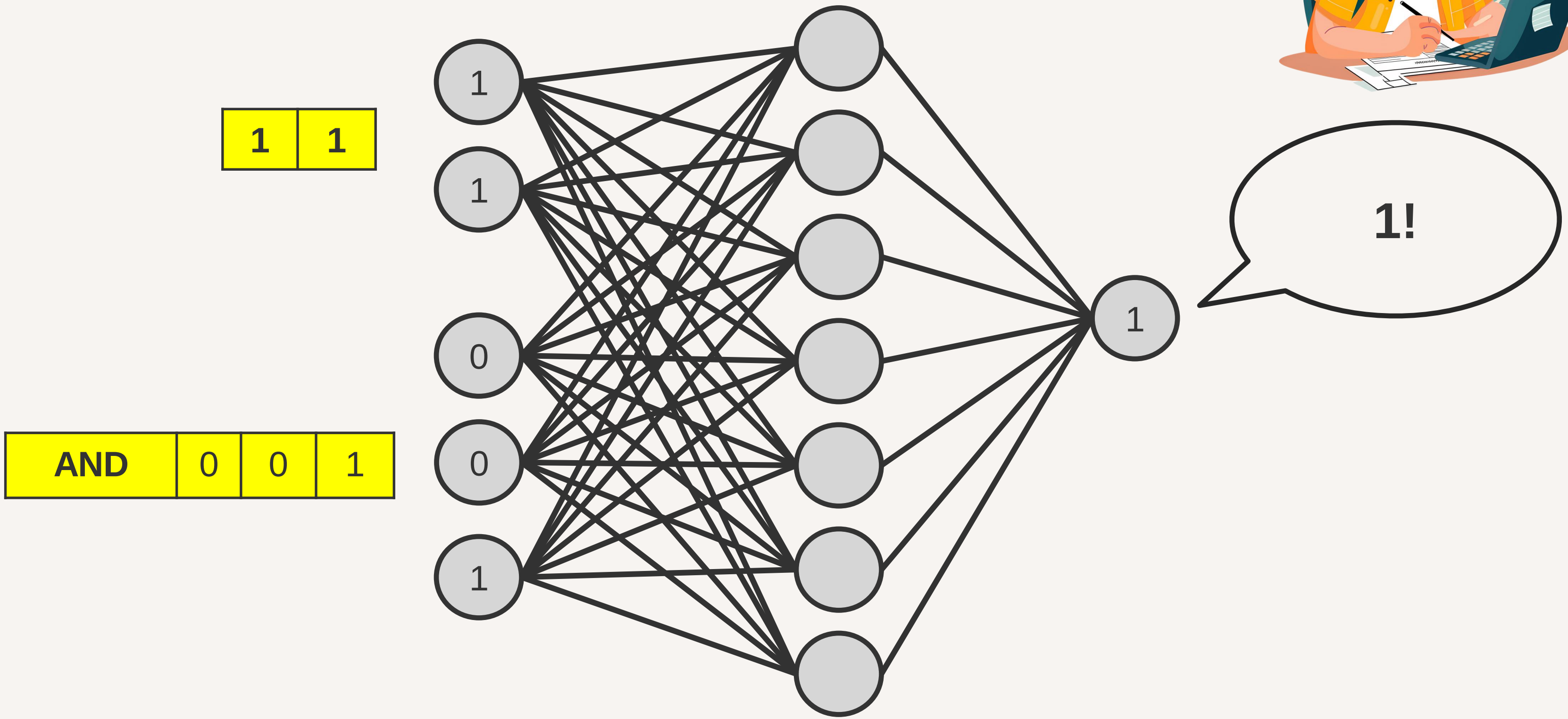
AND	0	0	1
XOR	1	0	0
RANDOM	0	1	0

OUTPUT

1!

Level 1: The Student

A multi-level computational model



Level II: the Teacher(s)

A multi-level computational model

ROLE: Choose and administer the tasks

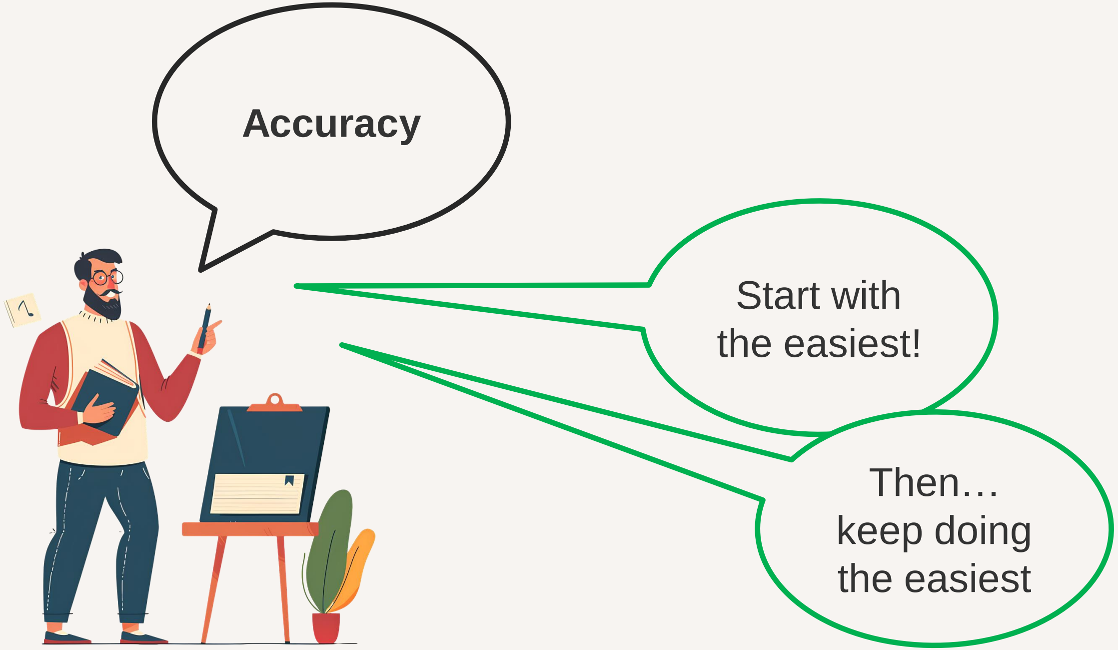
OBJECTIVE: Maximise the criterion

MODEL: Reinforcement learning model

Level II: the Teacher(s)

A multi-level computational model

The **criterion** can be different things:



(Poli et al., 2022; Ten et al., 2021)

Level II: the Teacher(s)

A multi-level computational model

The **criterion** can be different things:

Accuracy

Novelty

Explore what you don't know!

(Poli et al., 2022; Ten et al., 2021)

Level II: the Teacher(s)

A multi-level computational model

The **criterion** can be different things:

Accuracy

Novelty

Signed Learning progress

Keep improving!

(Poli et al., 2022; Ten et al., 2021)

Level II: the Teacher(s)

A multi-level computational model

The **criterion** can be different things:

Accuracy

Novelty

Signed Learning progress

Unsigned Learning progress

Don't forget anything…

(Poli et al., 2022; Ten et al., 2021)

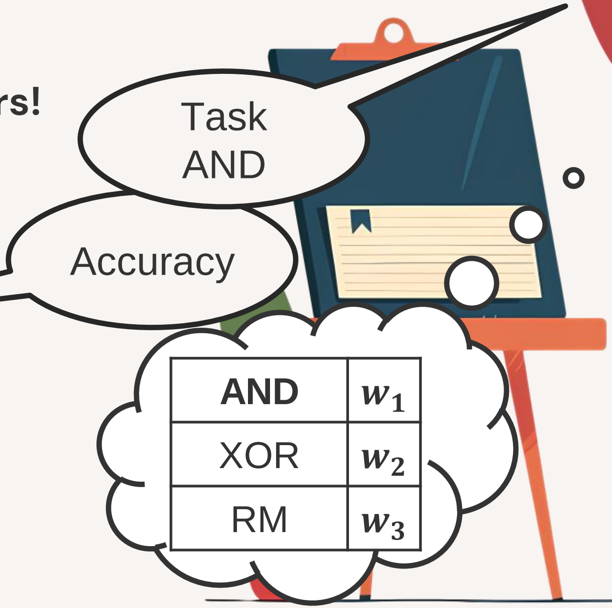
Level II: the Teacher(s)

A multi-level computational model

Uses the Rescorla–Wagner Rule:

$$V_{t\,[option]} = V_{t-1\,[option]} + \alpha\,(R_{t-1} - V_{t-1\,[option]})$$

→ This applies to **all teachers!**



Level III: the Coordinator

A multi-level computational model

ROLE: Gives weights to the teachers (LVL II)

OBJECTIVE: Maximise reward

MODEL: Reinforcement learning model

LVL II:

$$V_{t\,[option]} = \beta_1 G_t + \beta_2 B_t + \beta_3 A_t + \beta_4 H_t$$

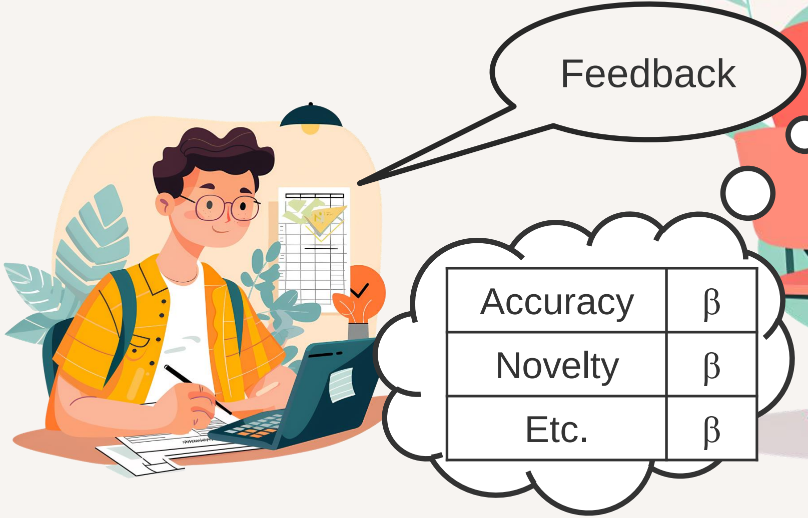
Level III: the Coordinator

A multi-level computational model

LVL III:

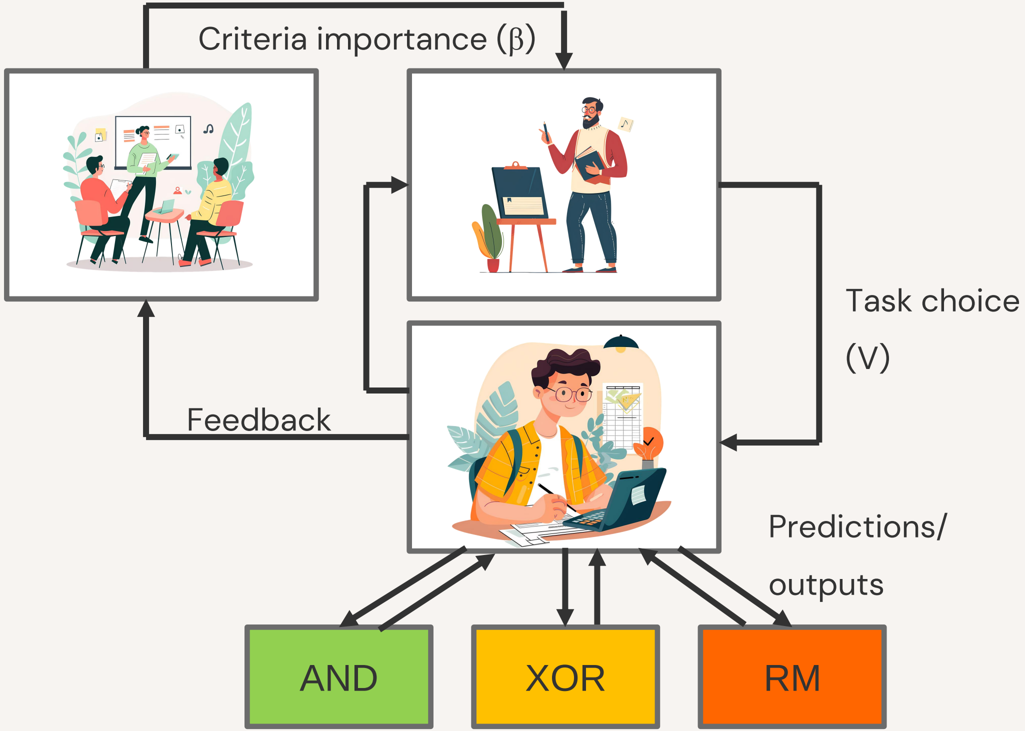
Receives feedback from student

Update each criteria weight (β)



A multi-level computational model

Structure of the levels



Hypotheses

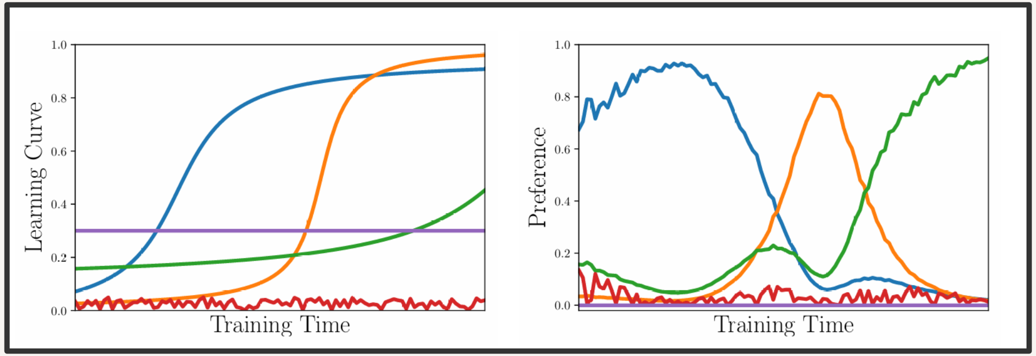
The curriculum model:

- **first** focus on the **easy task**
- then **hard task**
- **disregard** the **impossible task**

→ **faster learning** of both easy and hard task.

The random sampling model:

– **slowed down** by the **noisy task** (RM).



(Forestier et al., 2022)

Results

Results overview

LVL2 progress:

- **Accuracy OK**
- **Signed LP in progress**
- **Unsigned LP in progress**

→ Early testing for each criterion is done separately

General parameters of the models:

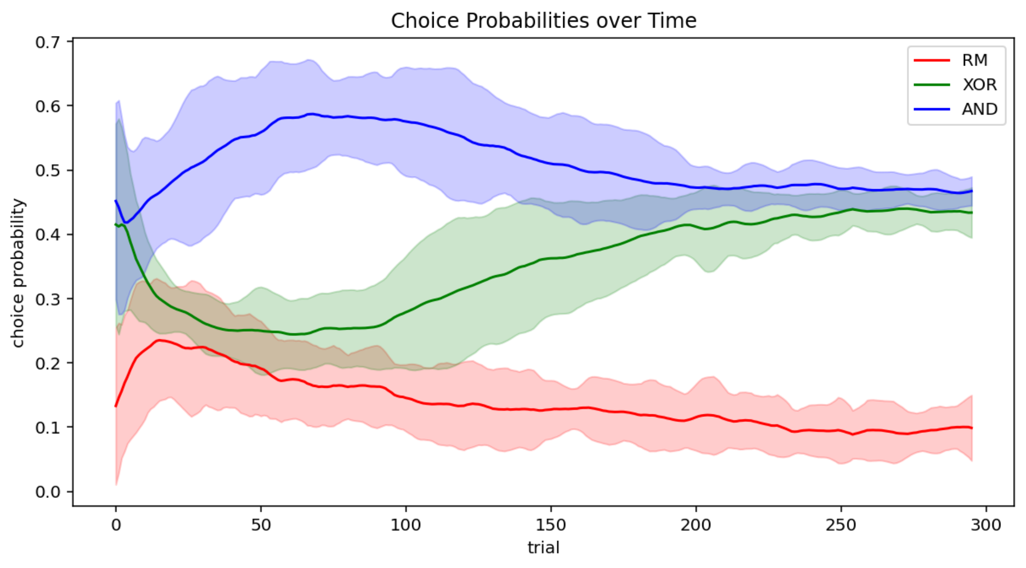
Learning rate of level 1: **0,05**

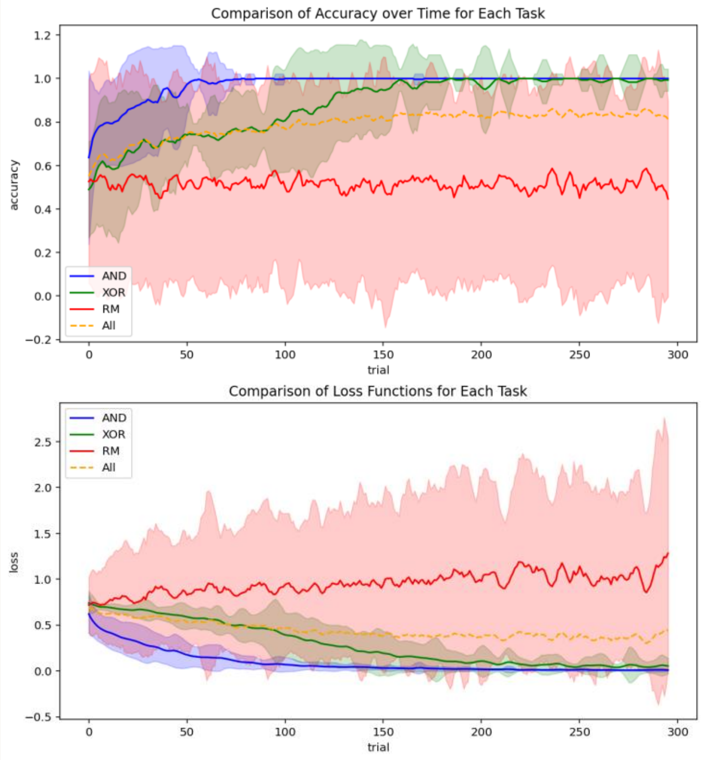
Learning rate of level 2: **0,3**

Inverse temperature: **2**

Number of runs: **15**

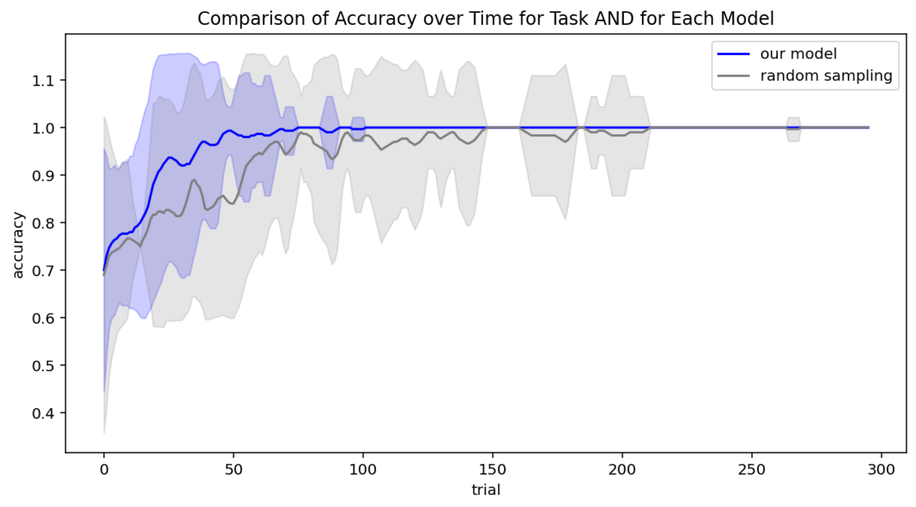
Criterion 1: Accuracy

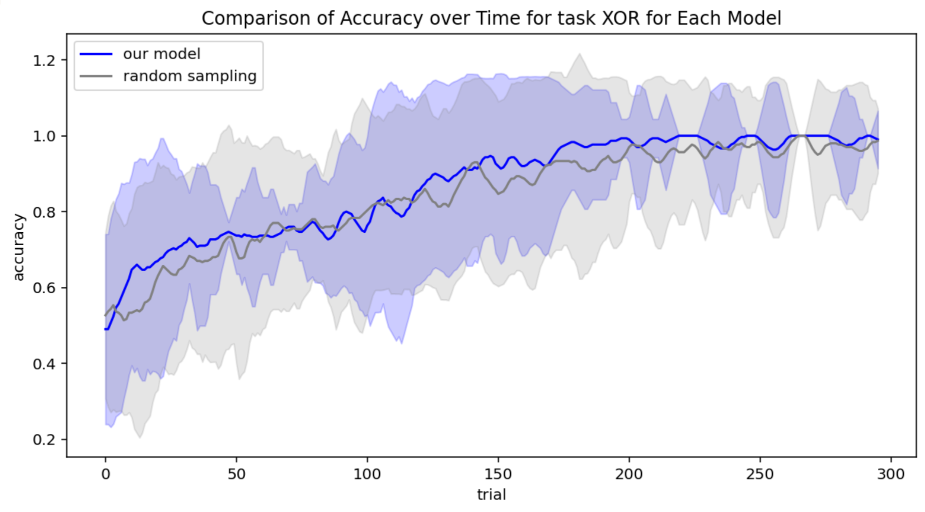




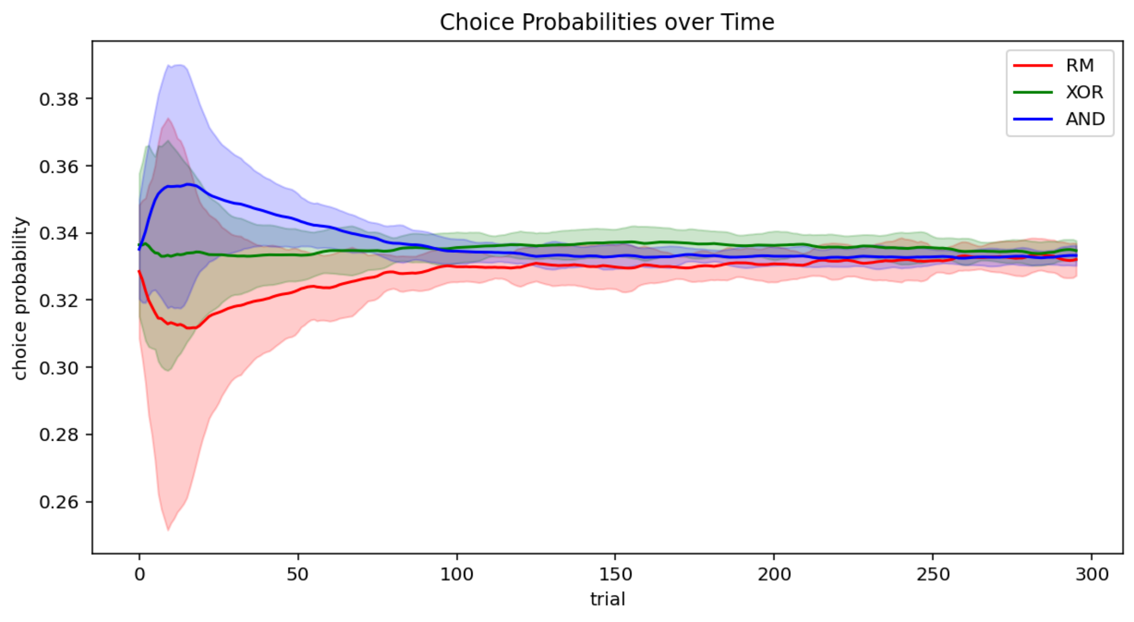
Criterion 1: Accuracy

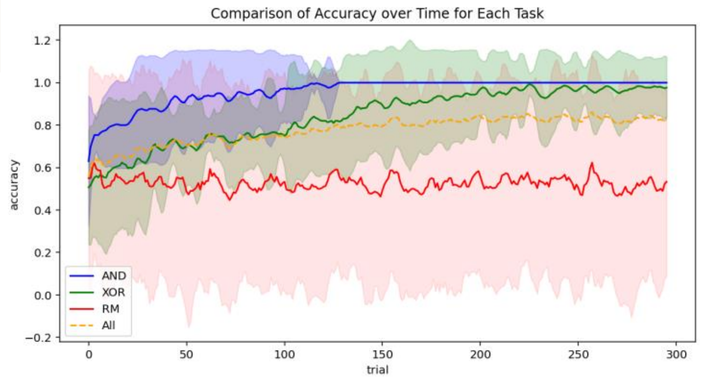
Performance

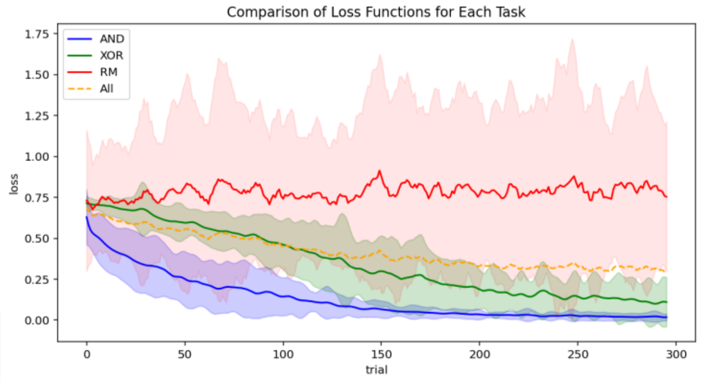




Criterion 2: Signed Learning Progress

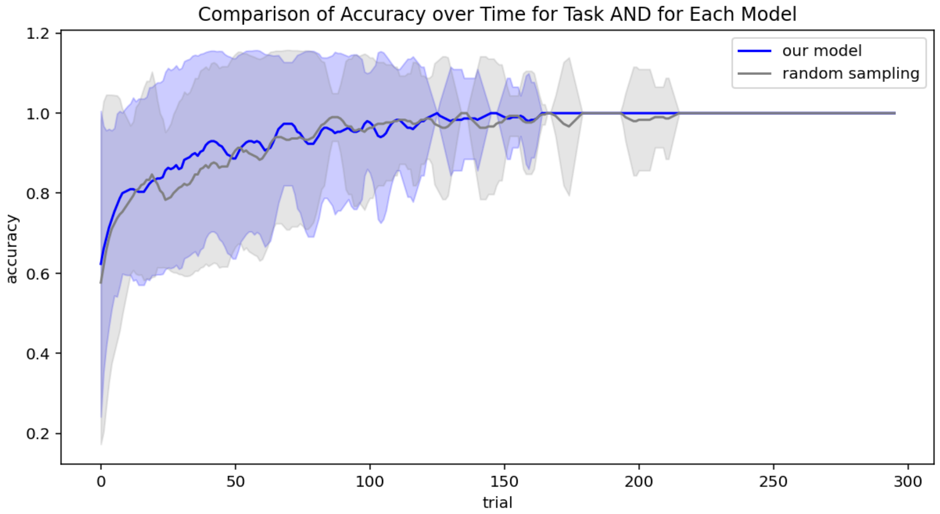


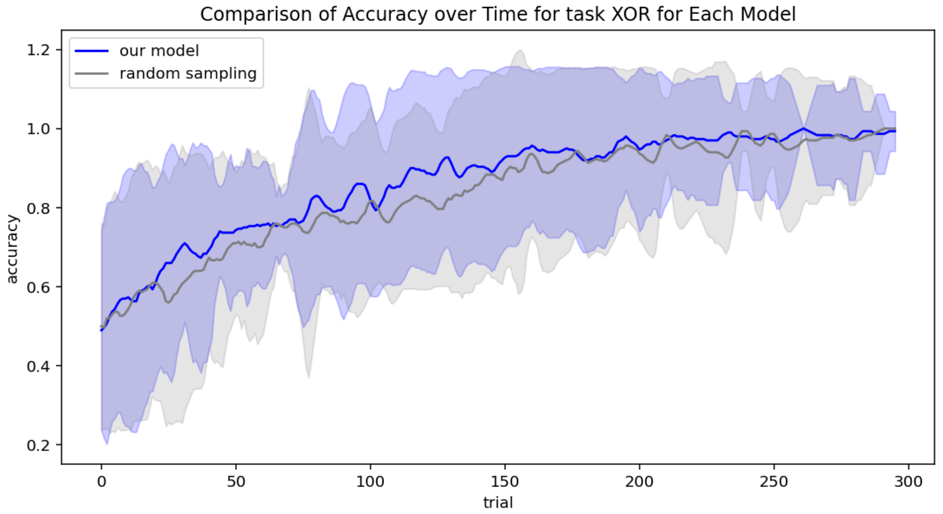




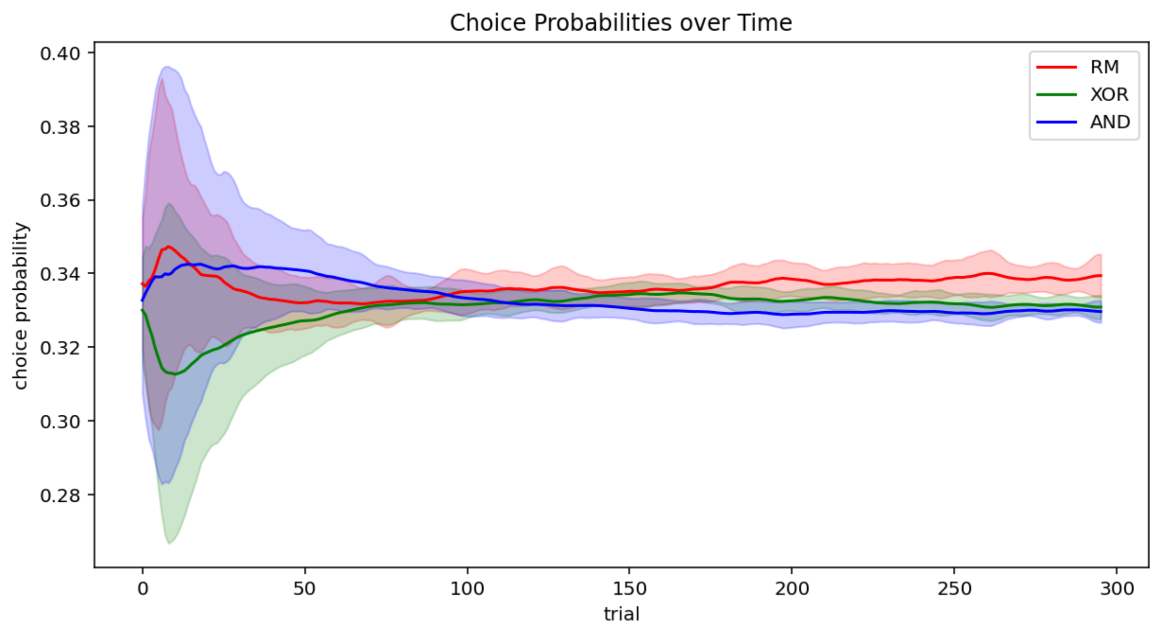
Criterion 2: Signed Learning Progress

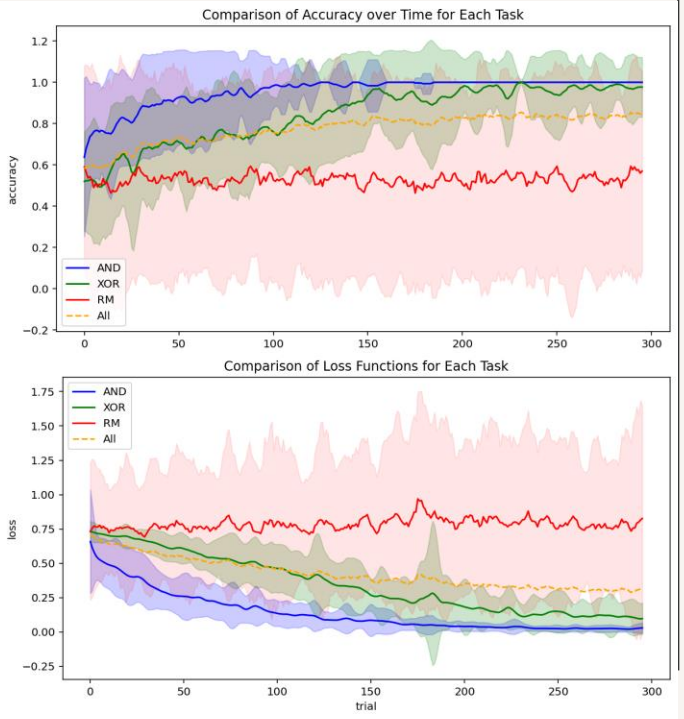
Performance





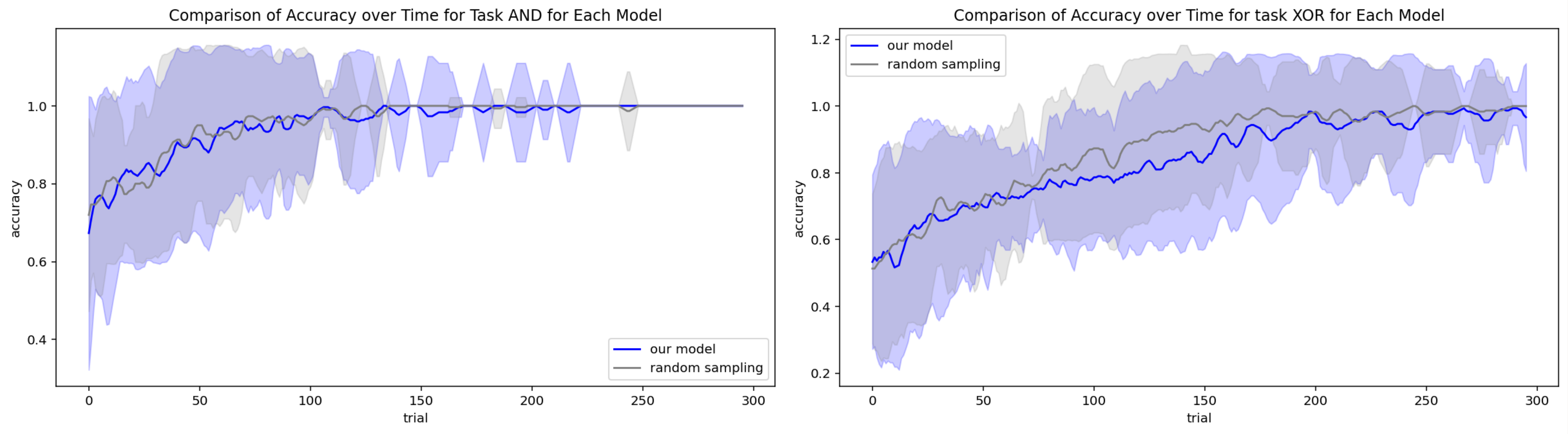
Criterion 3: Unsigned Learning Progress





Criterion 3: Unsigned Learning Progress

Performance



Results Conclusion

Not as sequential as we thought

Some effects are **still too small**

But:

Good **task differentiation**

Slightly **faster learning speed** for the **accuracy teacher.**

Next objectives

Next objective

- Add Novelty
- Set up LVL3, Coordinator

Thank you for listening

References

References

Bengio, Y., Louradour, J., Collobert, R. & Weston, J. (2009). Curriculum learning. in Proceedings of the 26th Annual International Conference on Machine Learning 41–48. ACM, Montreal Quebec Canada. https://doi.org/10.1145/1553374.1553380

Molinaro, G., Colas, C., Oudeyer, P.-Y., Collings, A. G. E. (2024). Latent learning progress drives autonomous goal selection in human reinforcement learning. Advances in Neural Information Processing Systems, 37.

Poli, F., Meyer, M., Mars, R. B., & Hunnius, S. (2022). Contributions of expected learning progress and perceptual novelty to curiosity-driven exploration. Cognition, 225, 105119. https://doi.org/10.1016/j.cognition.2022.105119

Silver, D., Singh, S., Precup, D., & Sutton, R. S. (2021). Reward is enough. Artificial Intelligence, 299, 103535. https://doi.org/10.1016/j.artint.2021.103535

Sutton, R., S., Barto, A., G. (1998). Reinforcement Learning: An Introduction. Cambridge University Press.

Ten, A., Kaushik, P., Oudeyer, P., & Gottlieb, J. (2021). Humans monitor learning progress in curiosity-driven exploration. Nature Communications, 12(1). https://doi.org/10.1038/s41467-021-26196-w

Tong, W. L., Iyer, A., Murthy, V. N., & Reddy, G. (2023). Adaptive algorithms for shaping behavior. bioRxiv (Cold Spring Harbor Laboratory). https://doi.org/10.1101/2023.12.03.569774

Wu, X., Dyer, E. & Neyshabur, B. (2021) When Do Curricula Work? in International Conference on Learning Representations.