Reusable Slotwise Mechanisms

Trang Nguyen[1,2], Amin Mansouri[1,4], Kanika Madan[1,4], Khuong Nguyen[2], Kartik Ahuja[1], Dianbo Liu[1,3], and Yoshua Bengio[1,4,5]

[1] Mila - Quebec AI Institute, [2] FPT Software AI Center, [3] National University of Singapore, [4] University of Montreal, [5] CIFAR AI Chair

New Orleans, December 2023

NEURAL INFORMATION PROCESSING SYSTEMS

Mila

FPT Software AICenter

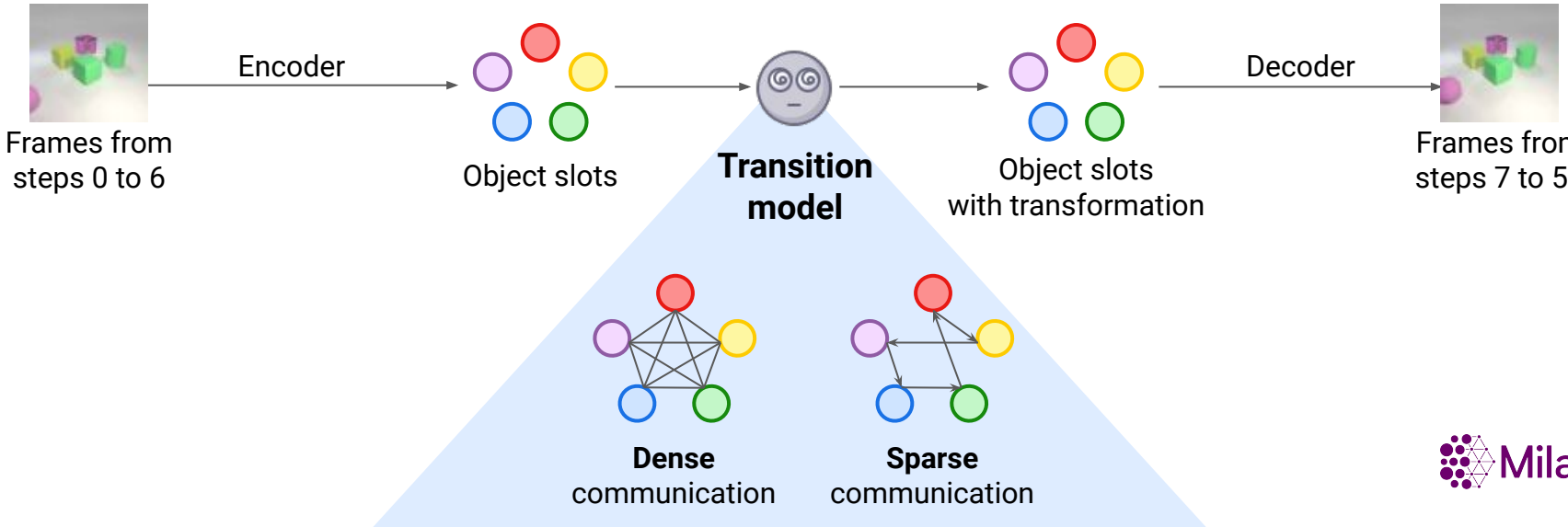
NUS National University of Singapore

Université de Montréal

CIFAR

Background

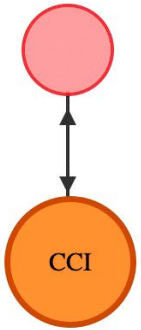
- The ability to reason about the dynamics of objects helps in achieving robustness and generalization when presented with distribution shifts
- It requires representing scenes effectively and understanding the mechanisms that govern the interactions
- Prior works model interactions either through too dense or too sparse form of communication among slots



Mila

Reusable Slotwise Mechanisms (RSM)

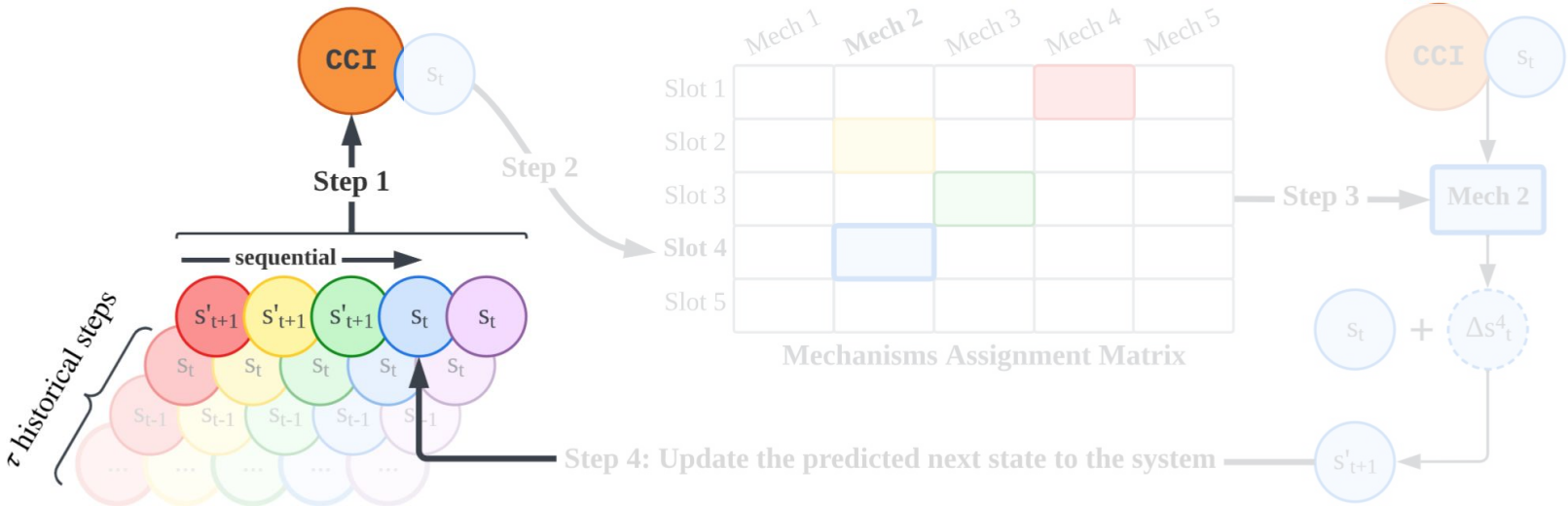
- **Our goals**:
  - Leverage *context-dependent* communication among slots
  - Relaxing the inductive biases over communication density
  - Central Contextual Information (CCI) treats latest slots status as a spatiotemporal context.
- **CCI:**
  - … is computed from latest slots
  - … is used for selecting the relevant mechanism for each slot
  - … provides context for determining the output of the selected mechanism, i.e., is provided along the object slots as input to the mechanism



Mila

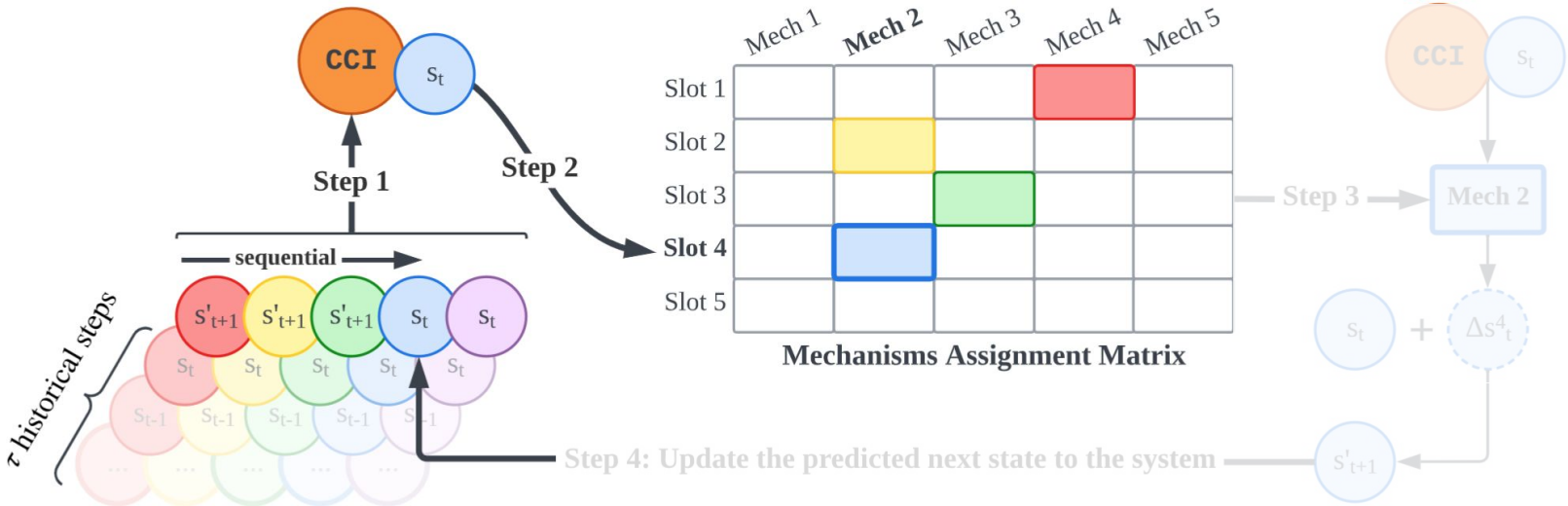
Reusable Slotwise Mechanisms (RSM)

- **Step 1:** Compute the CCI from historical steps



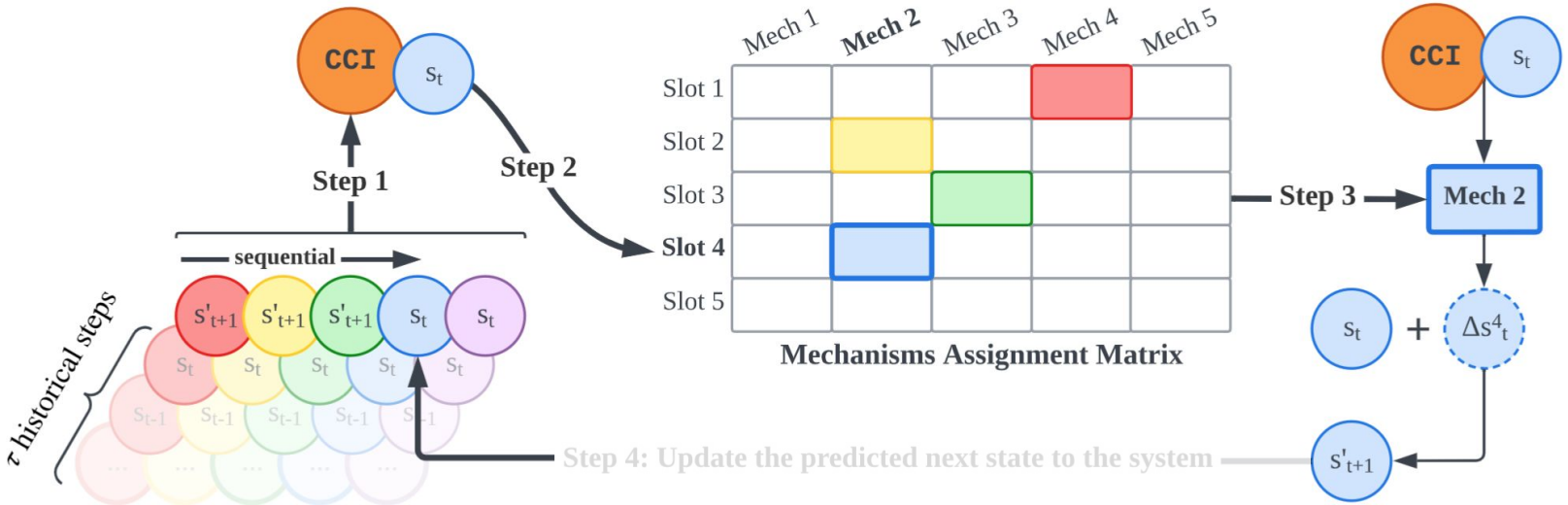
Reusable Slotwise Mechanisms (RSM)

- **Step 2:** Select a mechanism for the slot of interest



Reusable Slotwise Mechanisms (RSM)

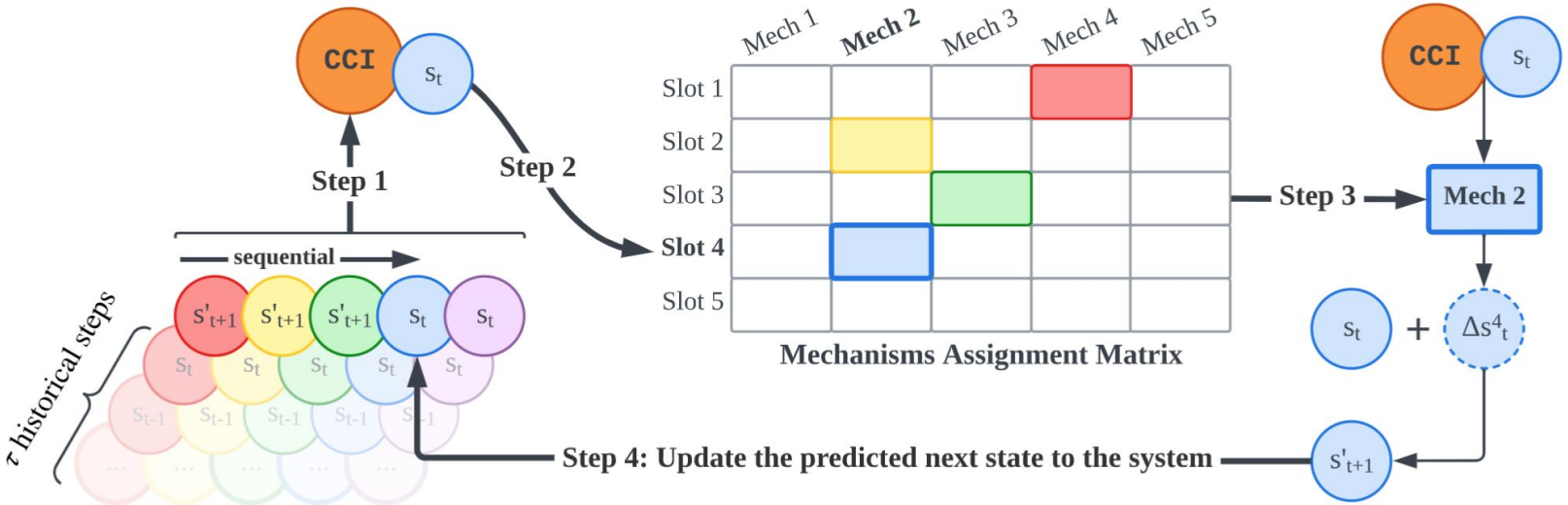
- **Step 3:** Compute the transformation of slot from *t* to *t+1*



Mila

Reusable Slotwise Mechanisms (RSM)

- **Step 4:** Sync the predicted next state of slot to the system



Mila

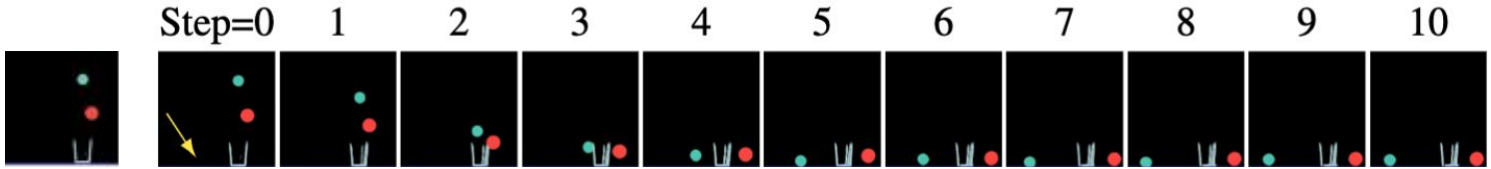
Experiments Setup

	Video Prediction	Visual Question Answering	Action Planning	OOD Evaluation
OBJ3D 	✓			
CLEVRER 	✓	✓		
PHYRE 	✓		✓	✓
Physion 	✓	✓		

Mila

Experiment Results - Action planning

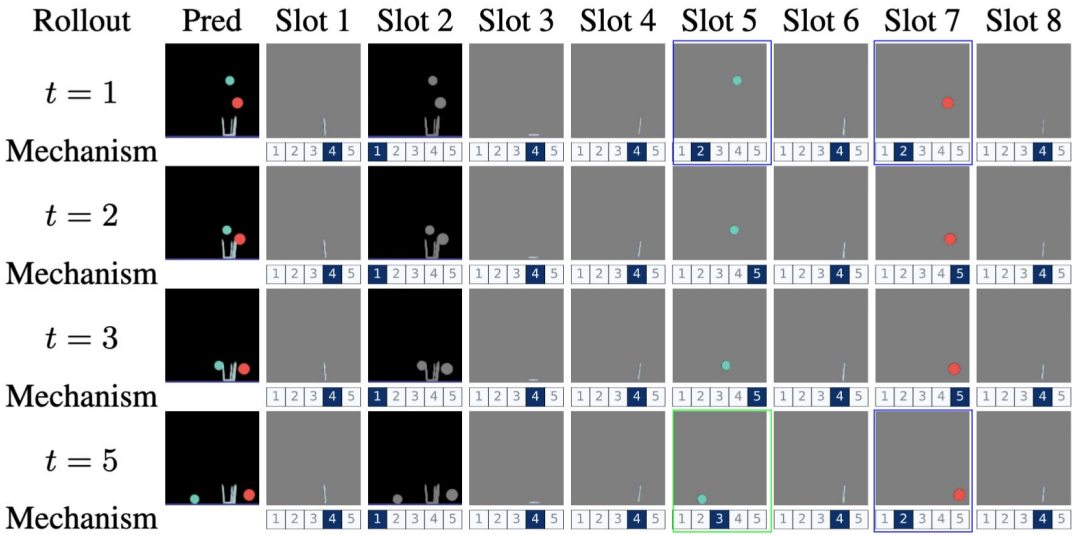
- RSM strategically positions a red ball at step 0 that helps the green ball make contact with the blue floor (indicated by the arrow).



Mila

Experiment Results - Mechanism Selection

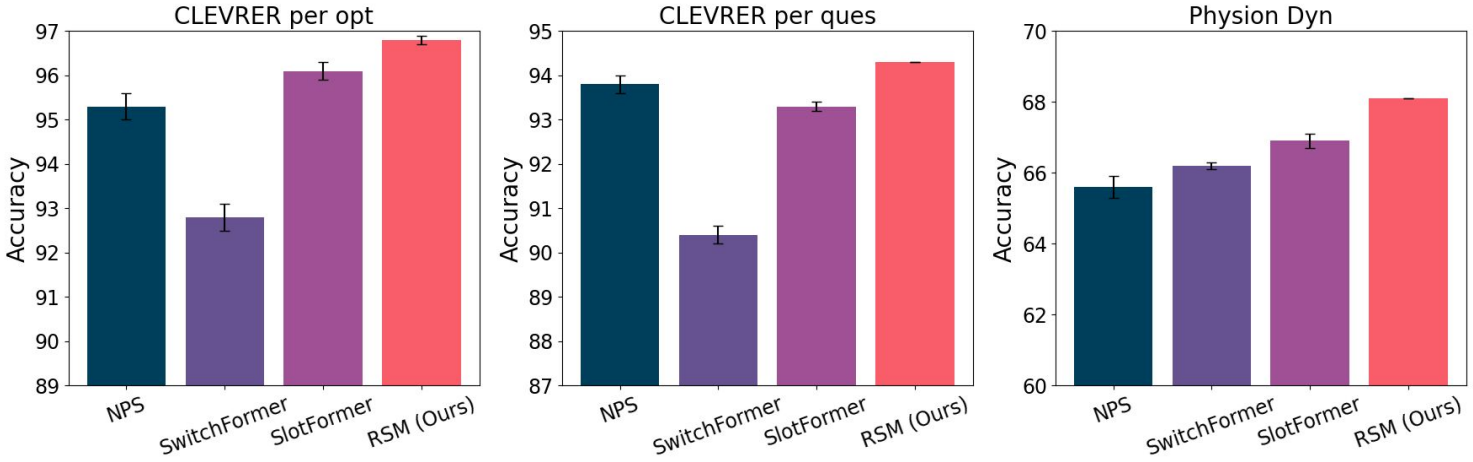
- RSM disentangles objects' dynamics into reusable mechanisms,
- Mechanisms can be expressed as Collision (2), Moving left or right (3), Idle (4), and Falling (5).



Mila

Experiment Results - Visual Question Answering

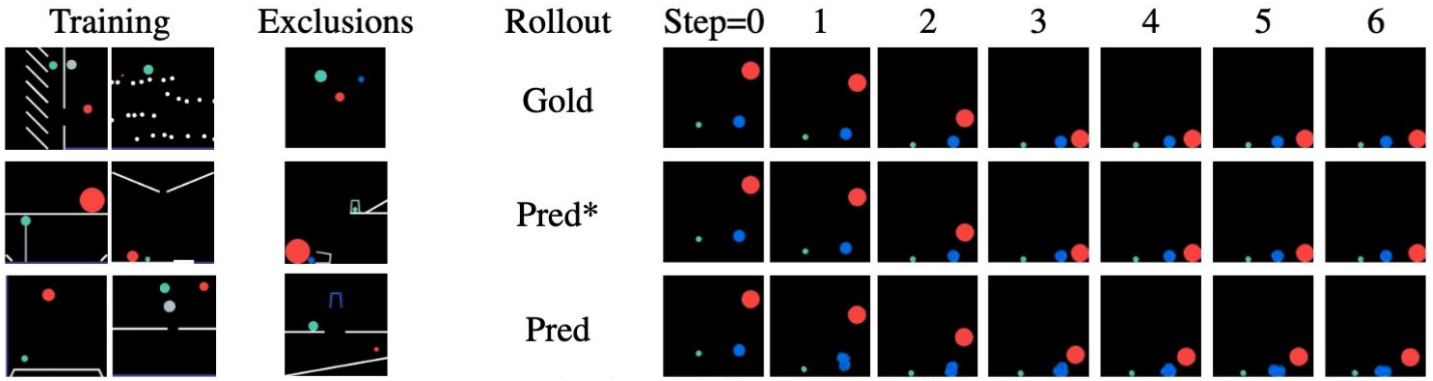
- RSM consistently outperforms all baselines in CLEVRER and Physion



Mila

Experiment Results - OOD Evaluation

- **OOD in Dynamics**
  - The **Blue** moving objects only appear in the test set
  - RSM generalizes the movements of the ball-shaped objects in OOD settings

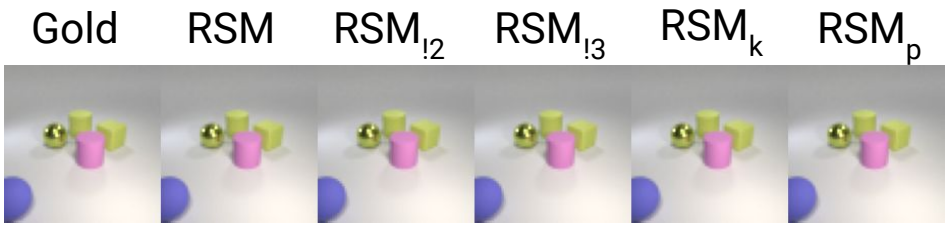


Mila

Ablation Studies

The original RSM always demonstrates accurate future frame predictions compared to the modified versions

- **Gold**: Ground truth frames
- **RSM**: Original design of RSM
- **RSM$_{!2}$**: Omit the CCI in step 2
- **RSM$_{!3}$**: Omit the CCI in step 3
- **RSM$_{k}$**: Randomly selecting the mechanisms
- **RSM$_{p}$**: Slots are updated in parallel



Mila

Summary

- RSM relaxes the inductive biases in communication sparsity among slots using a bottleneck called Central Contextual Information (CCI)
- RSM is advantageous over the baselines in various tasks, both iid and OOD scenarios
- There is a promise for exploring more sophisticated stochastic attention mechanisms for information integration

Mila

Thank you!

Mila