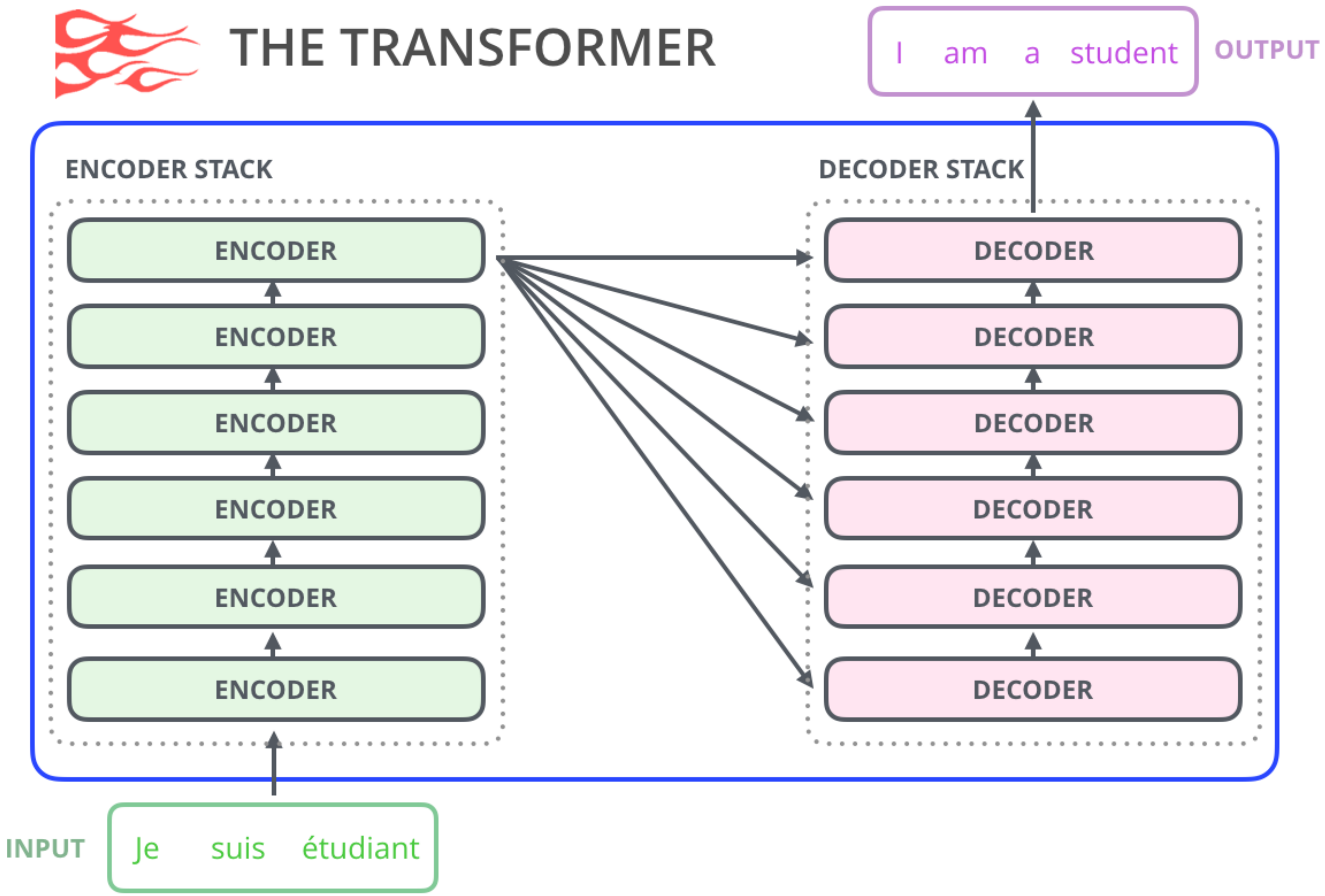
Spatially-Aware Transformer for Embodied Agents

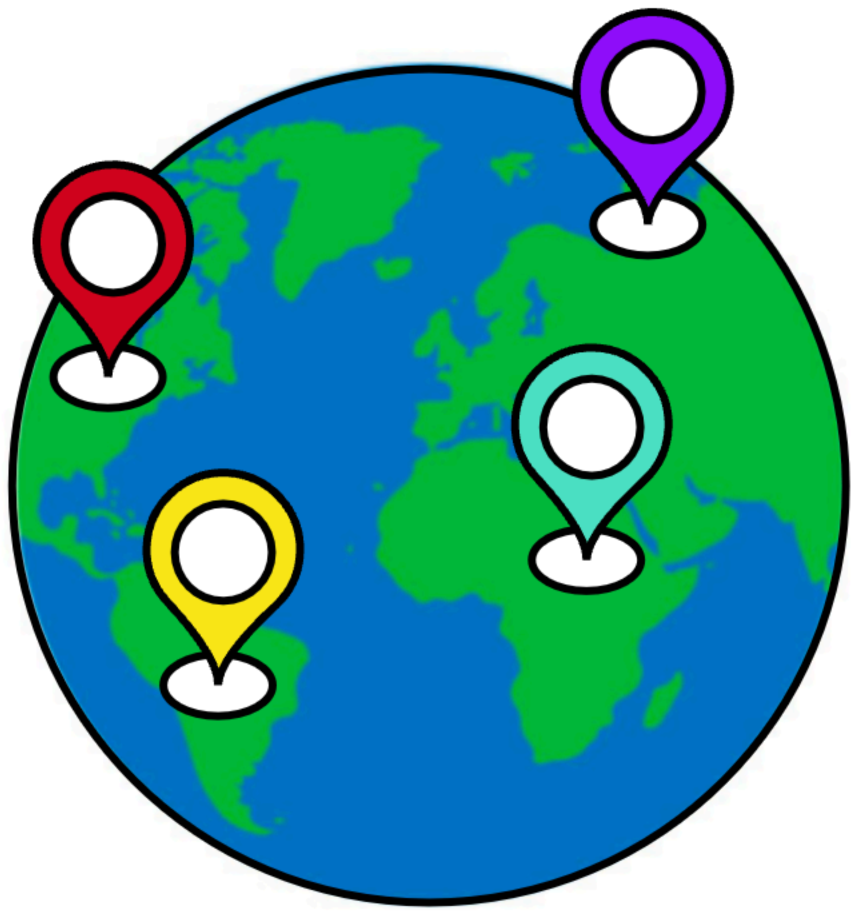
Junmo Cho[1]*, Jaesik Yoon[1,2]*, Sungjin Ahn[1]

[1]KAIST & [2]SAP

* denotes equal contribution.

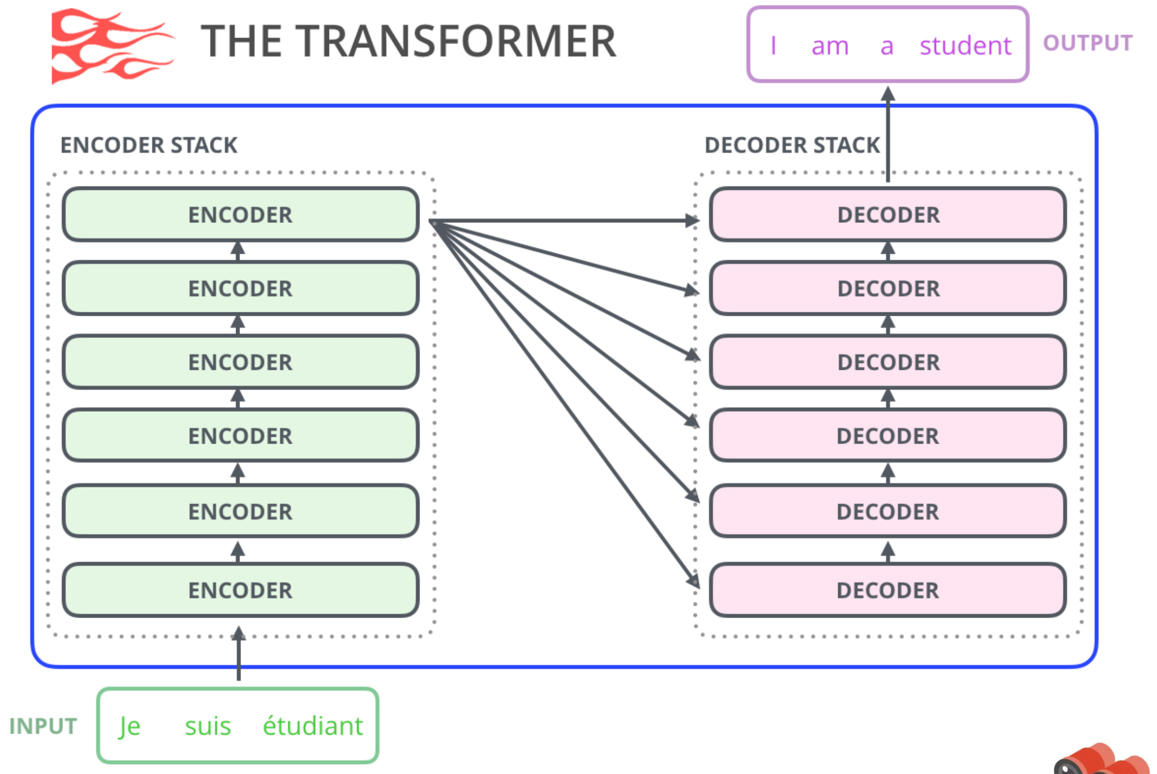
Transformer and Space

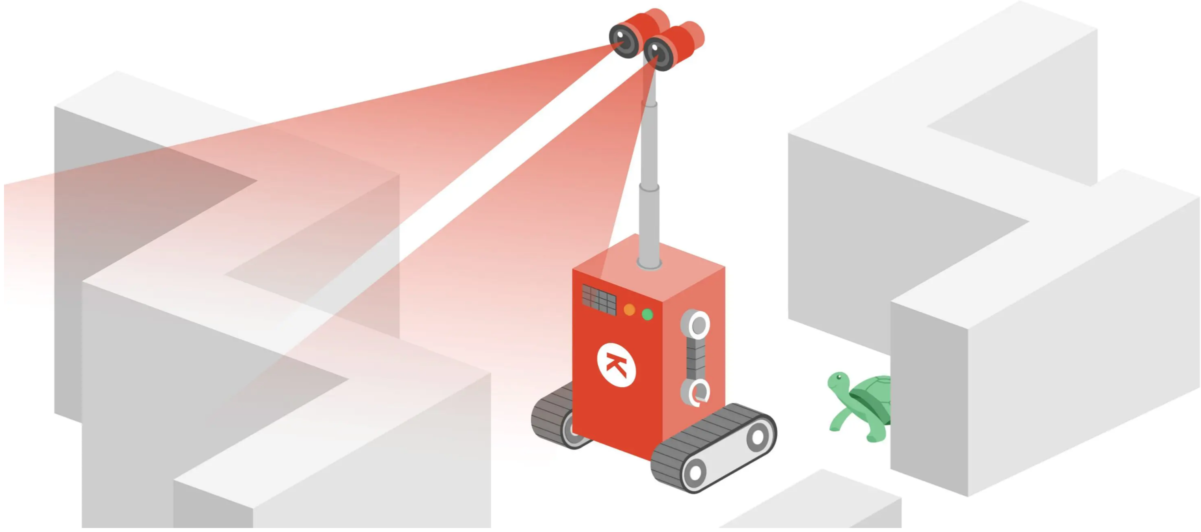




what if spatial information is available to transformers
in the same way as temporal order information?

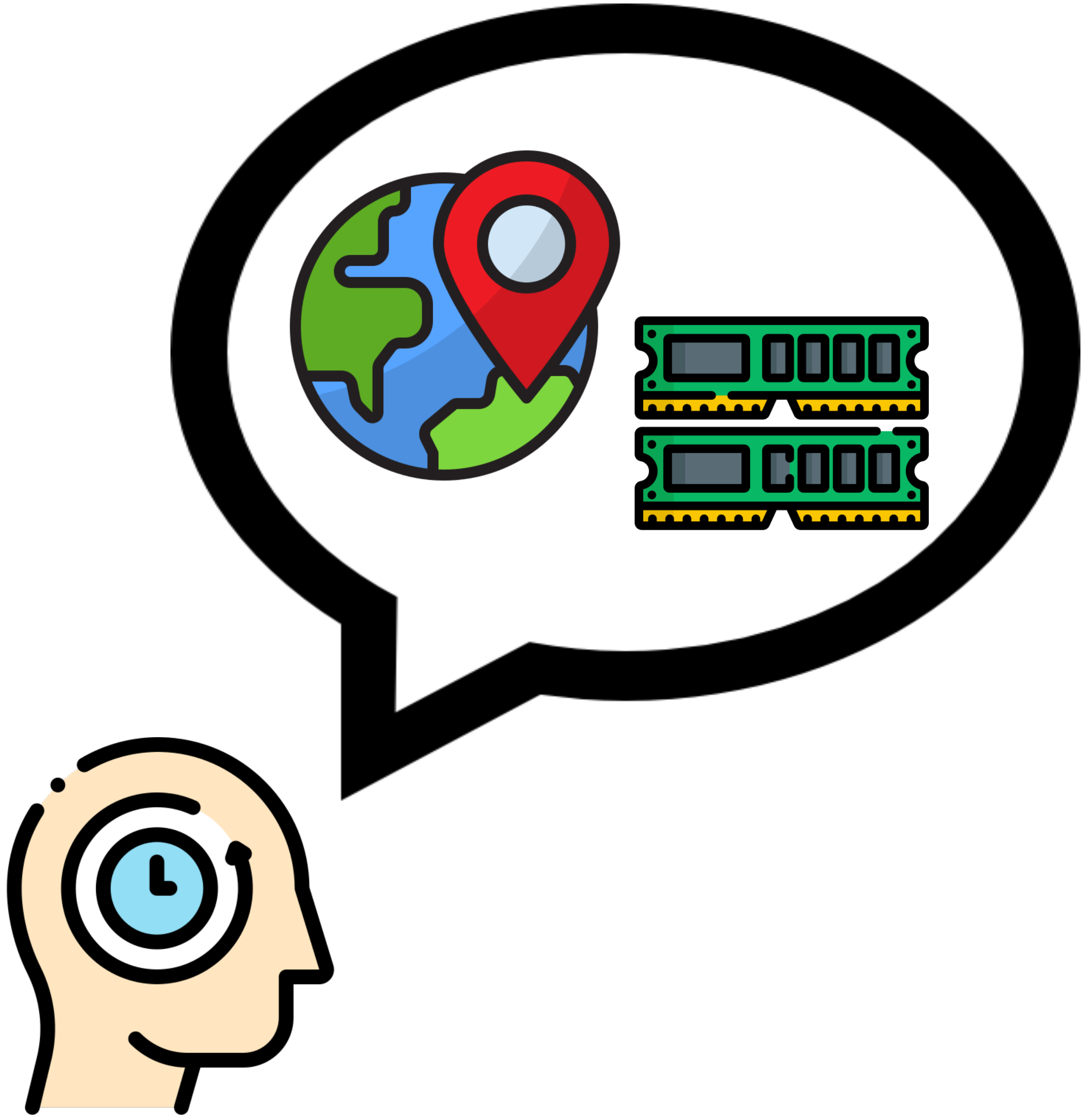
Why space is overlooked?





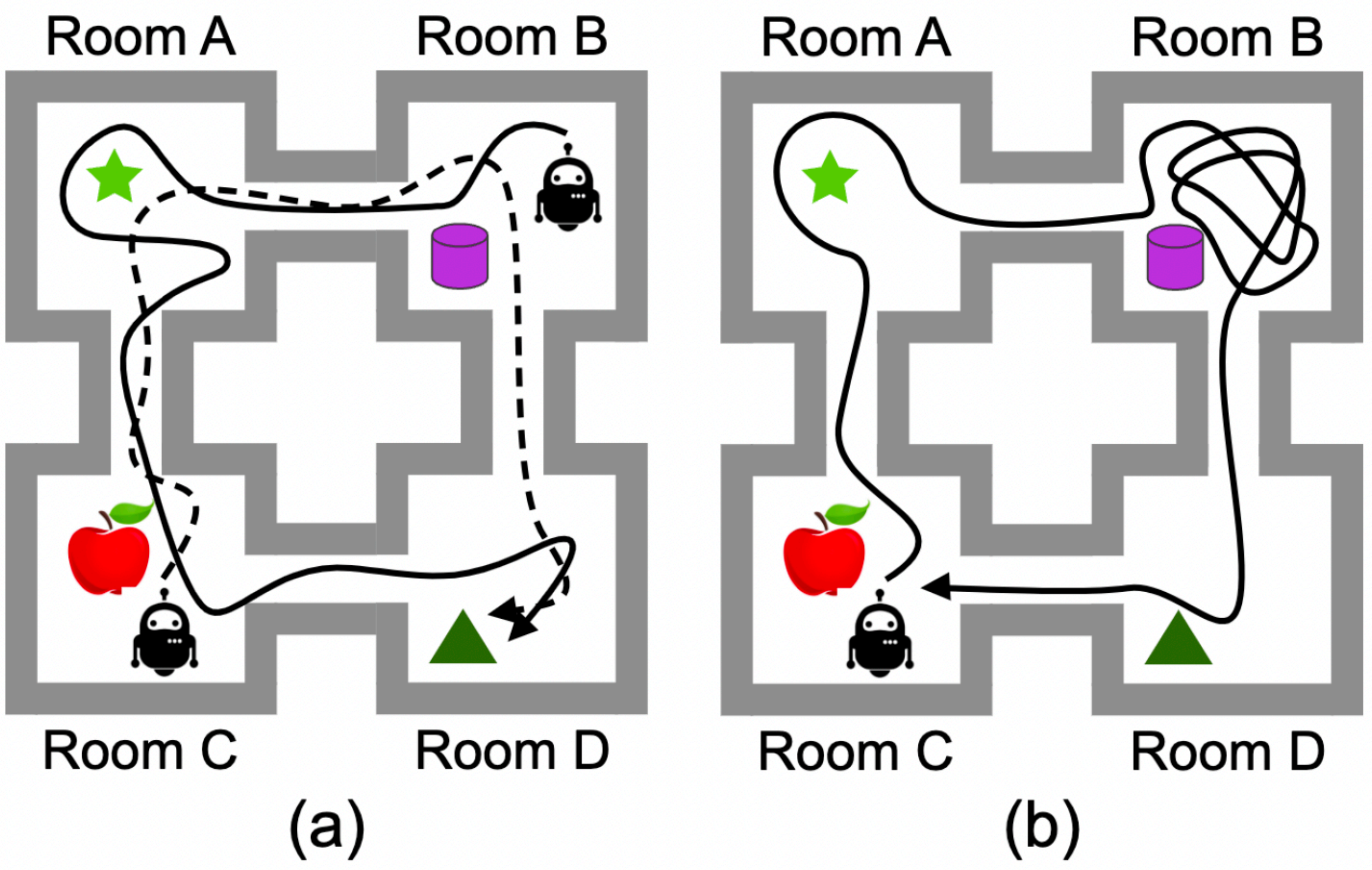
Origin of Transformer
& Access to spatial info.

How it benefits?



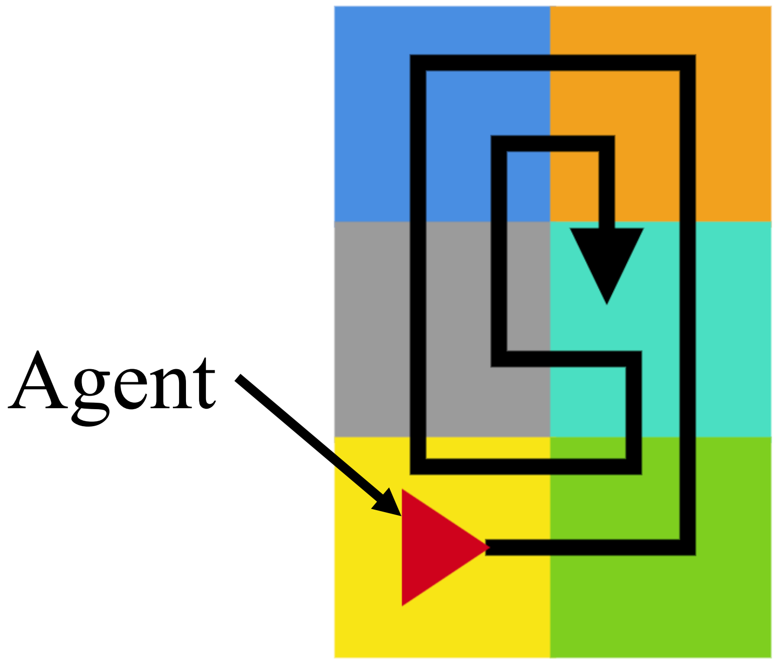
Spatial Reasoning and
Episodic Memory for Agent

Home Robot Thought Experiments

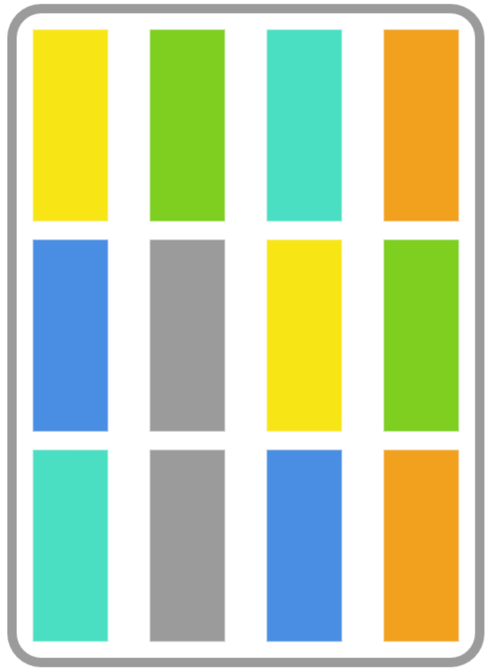


SAT with FIFO Memory

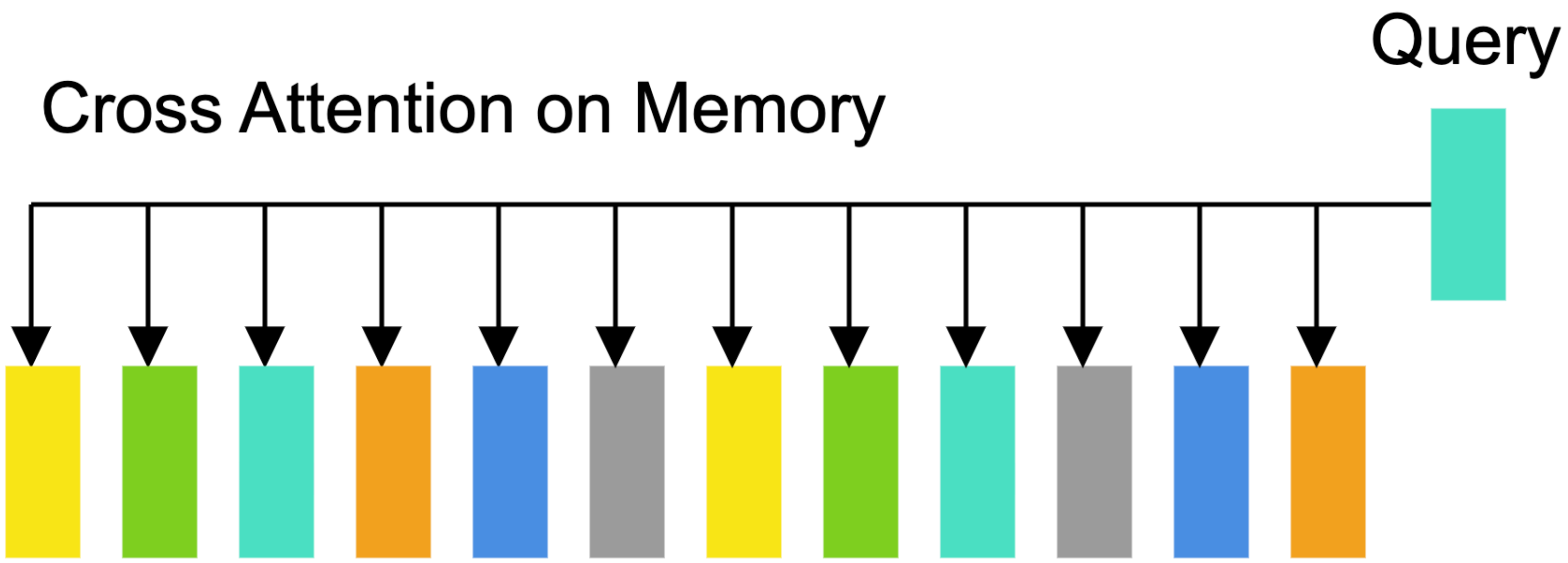
- Assume transformer-based agent in spatial environment
- At time step t, we have observation e_t^{obs}, time e_t^{time}, and location e_t^{loc}
- Experience frame $x_t = \text{sum\_embed}(e_t^{loc}, e_t^{time}, e_t^{obs})$



Spatial environment



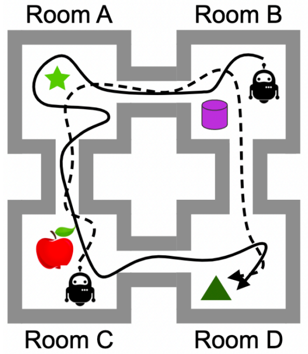
Experience Frames



Attention in SAT-FIFO

SAT with Place Memory

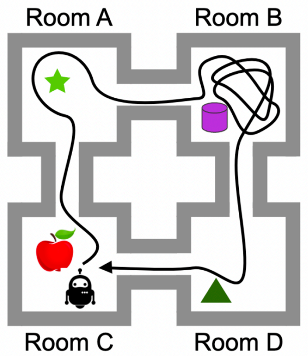
- With SAT-FIFO, we can solve spatial reasoning task in Home Robot Experiment!



(a)

Solved!

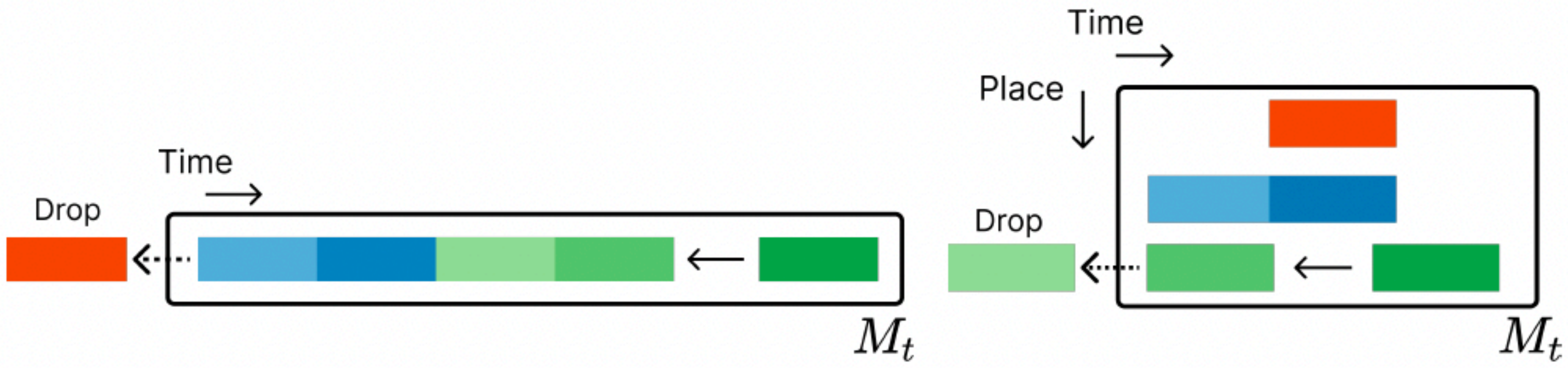
- However, naïve FIFO memory removes the oldest experience



(b)

How?

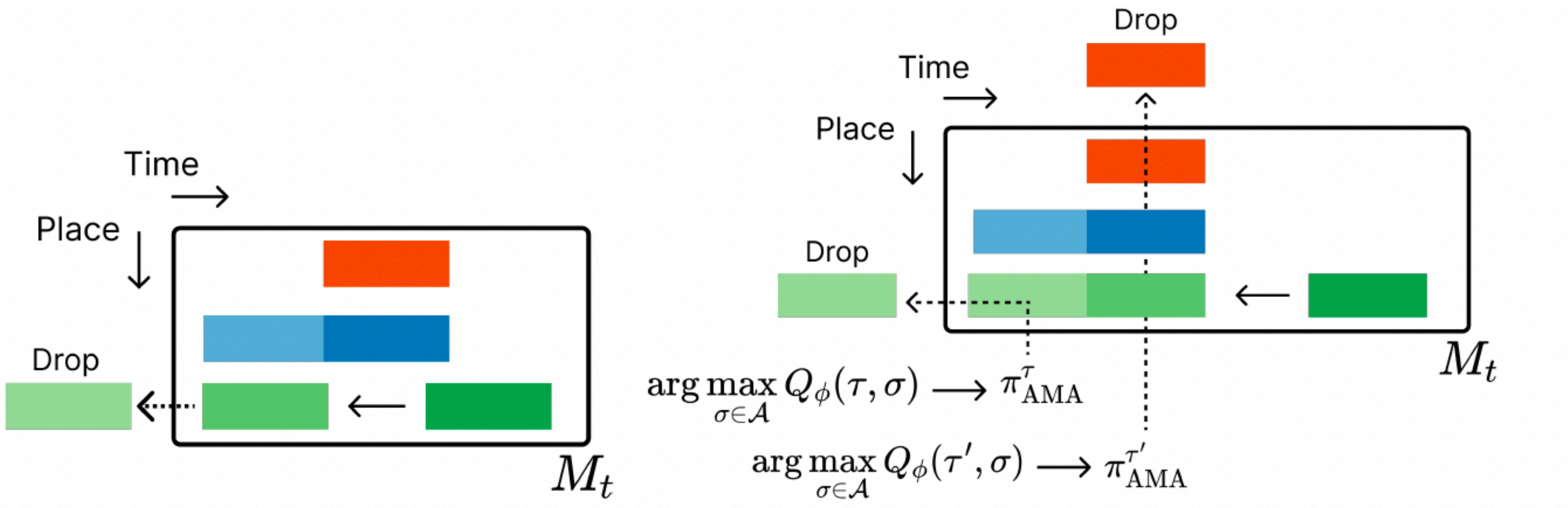
- We introduce SAT with place memory (SAT-PM) that allocates memory slots for each place



(a) FIFO Memory **(b)** Place Memory

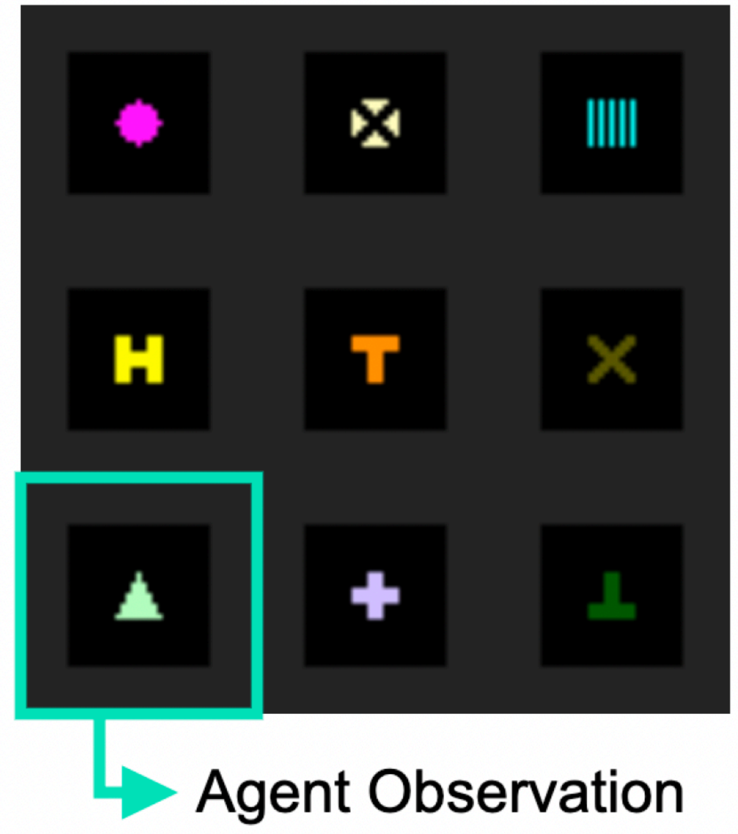
SAT with Adaptive Memory Allocator

- What if we need memory at the beginning of the episode? Or end of the episode? Or at some place?
- To address this issue, we propose Adaptive Memory Allocator (AMA) which is a learnable policy that chooses memory management strategy based on the task type

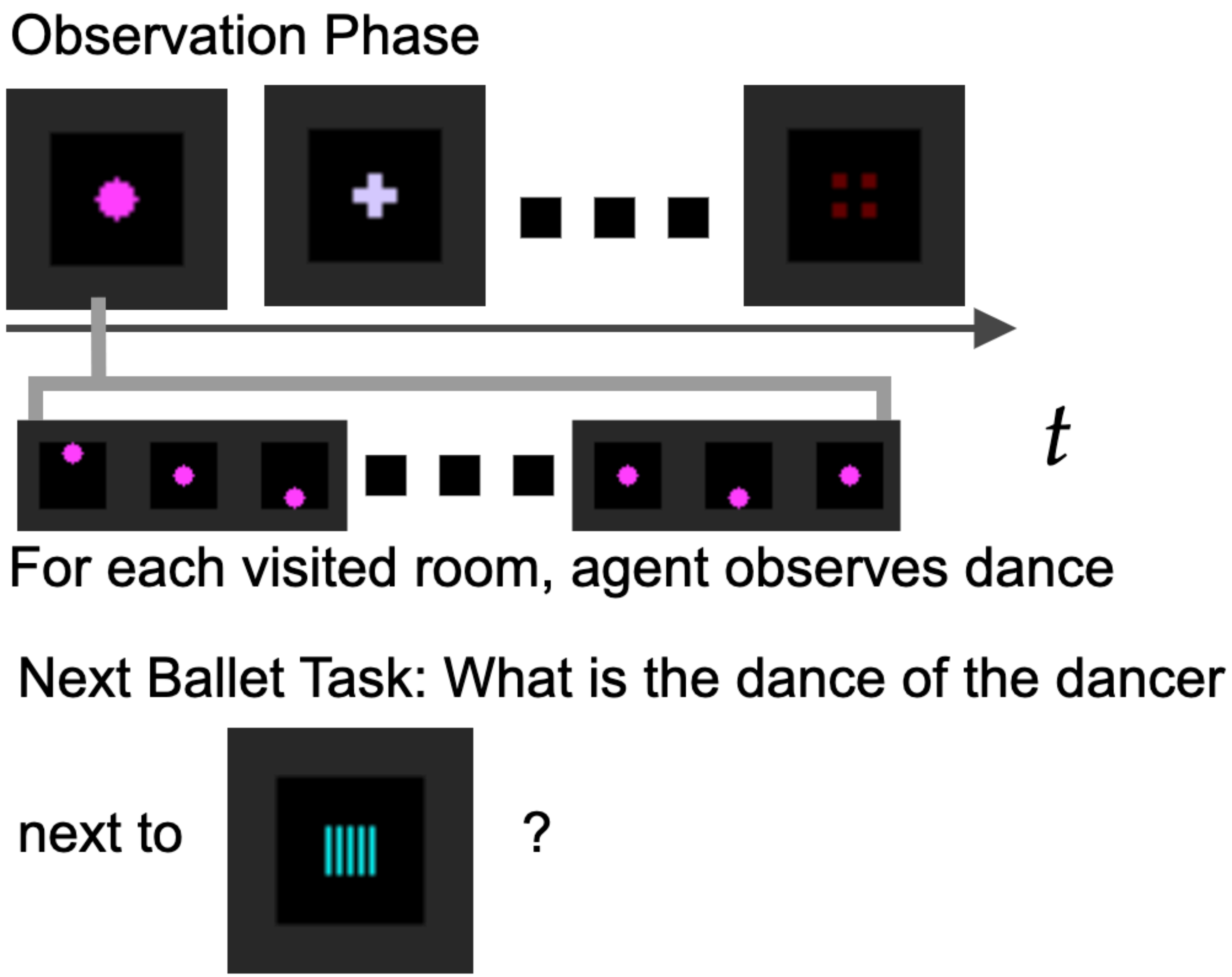


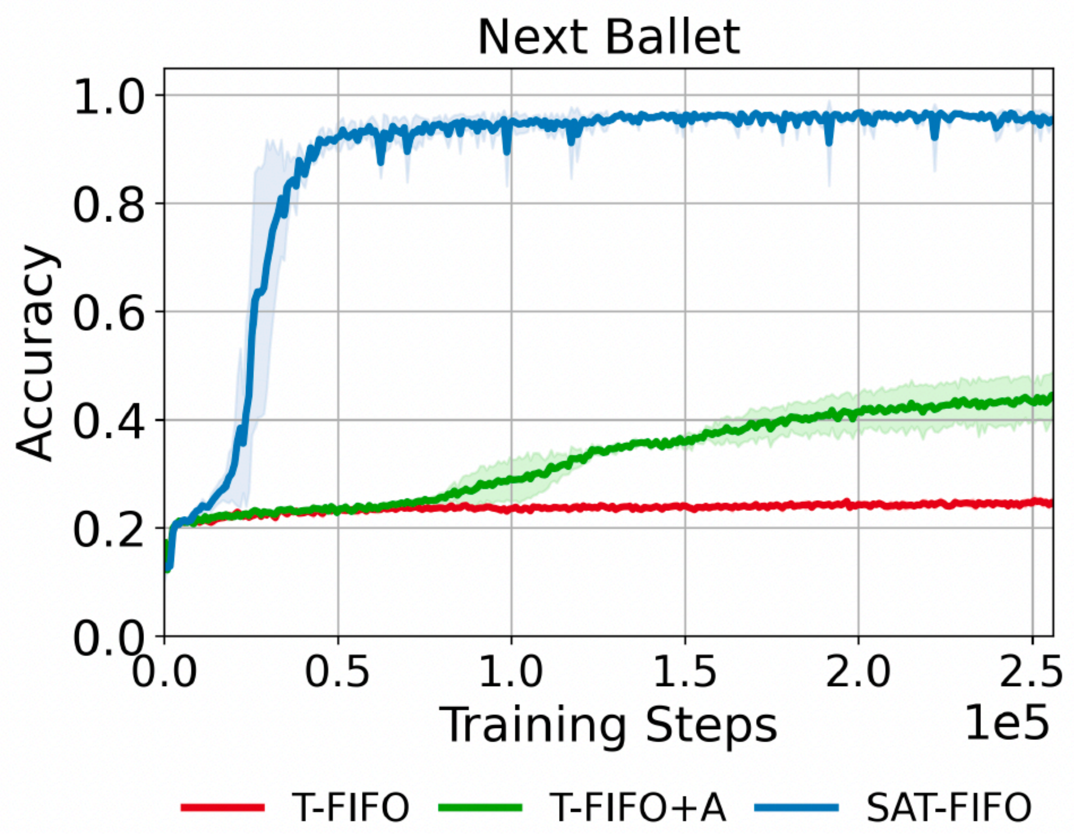
(b) Place Memory **(c)** Adaptive Memory Allocator (AMA)

Experiment 1. Implicit derivation of spatial information

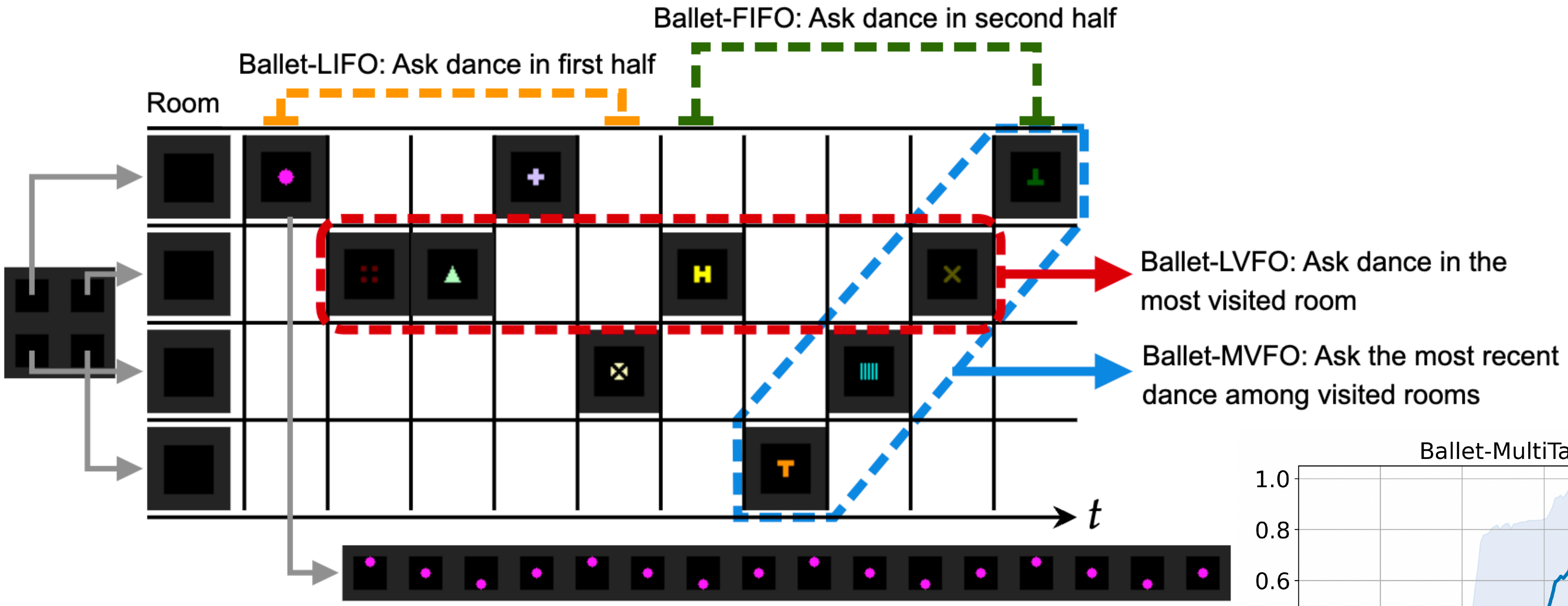


Room Ballet Environment

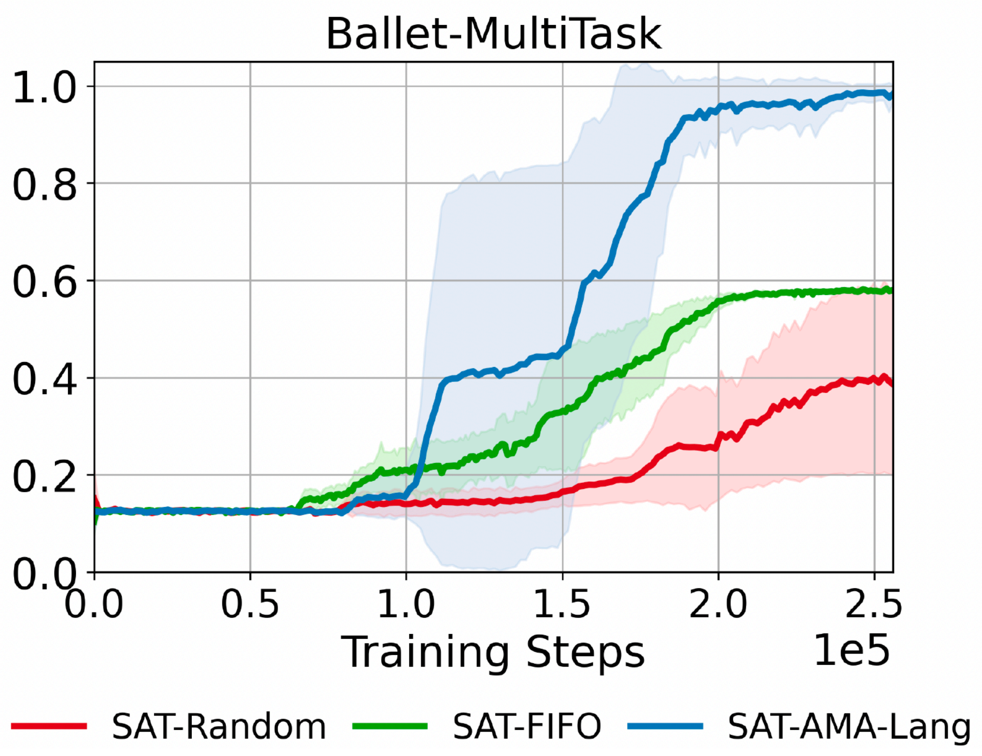




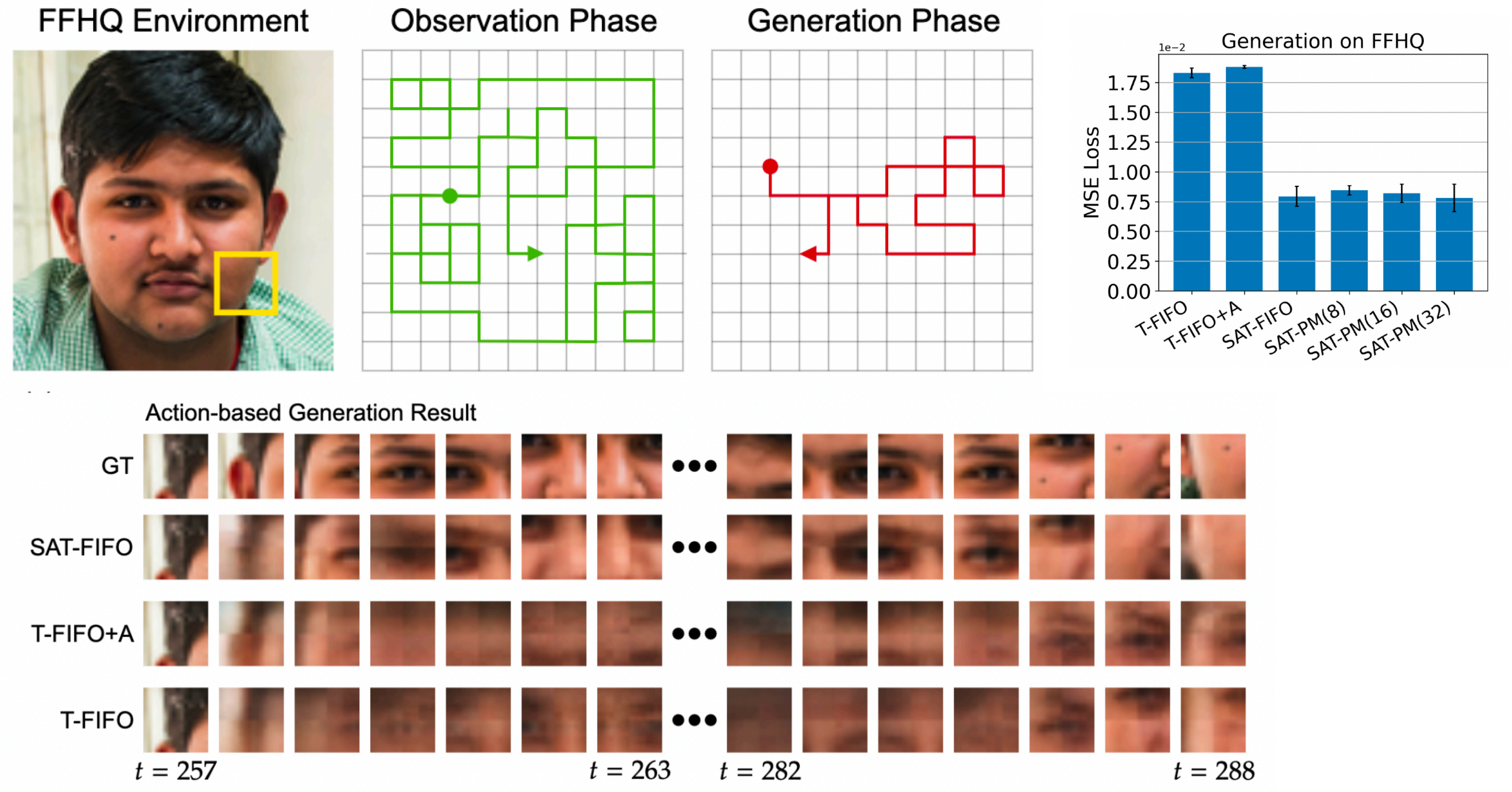
Experiment 2. Learning to select memory allocation strategy



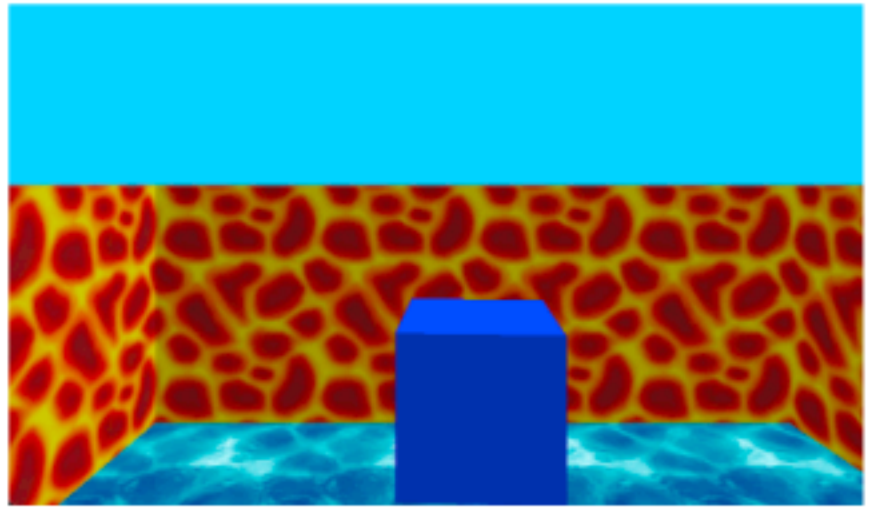
For each visited room, the agent observes a dance of the dancer while there is a 16-step delay before dancing (Total 32 steps)



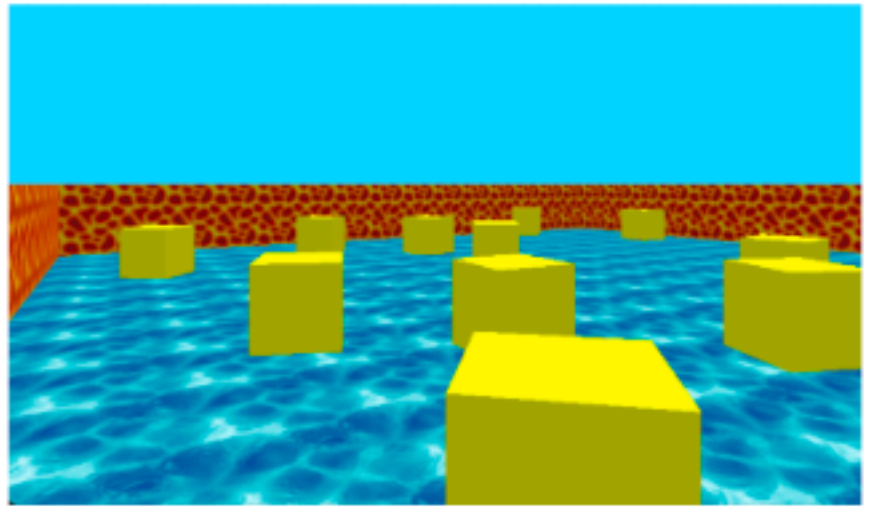
Experiment 3. Action-conditioned Generation



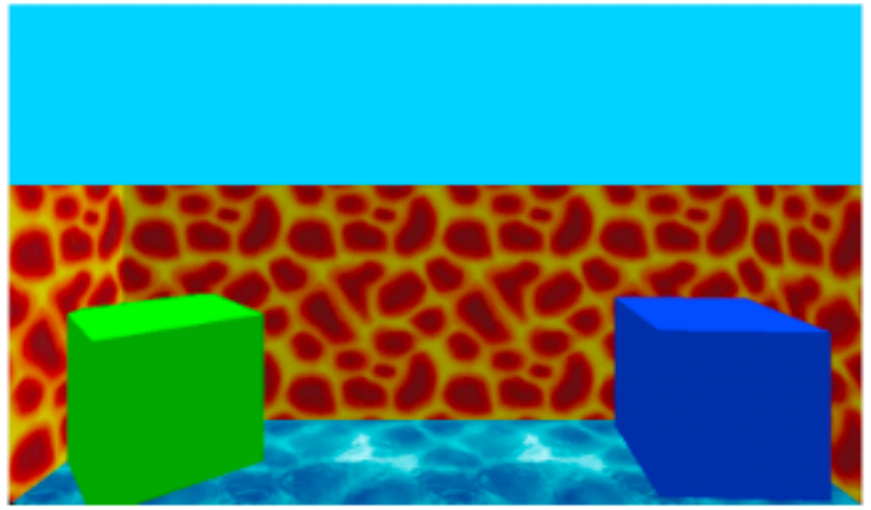
Experiment 4. Reinforcement Learning in MiniWorld



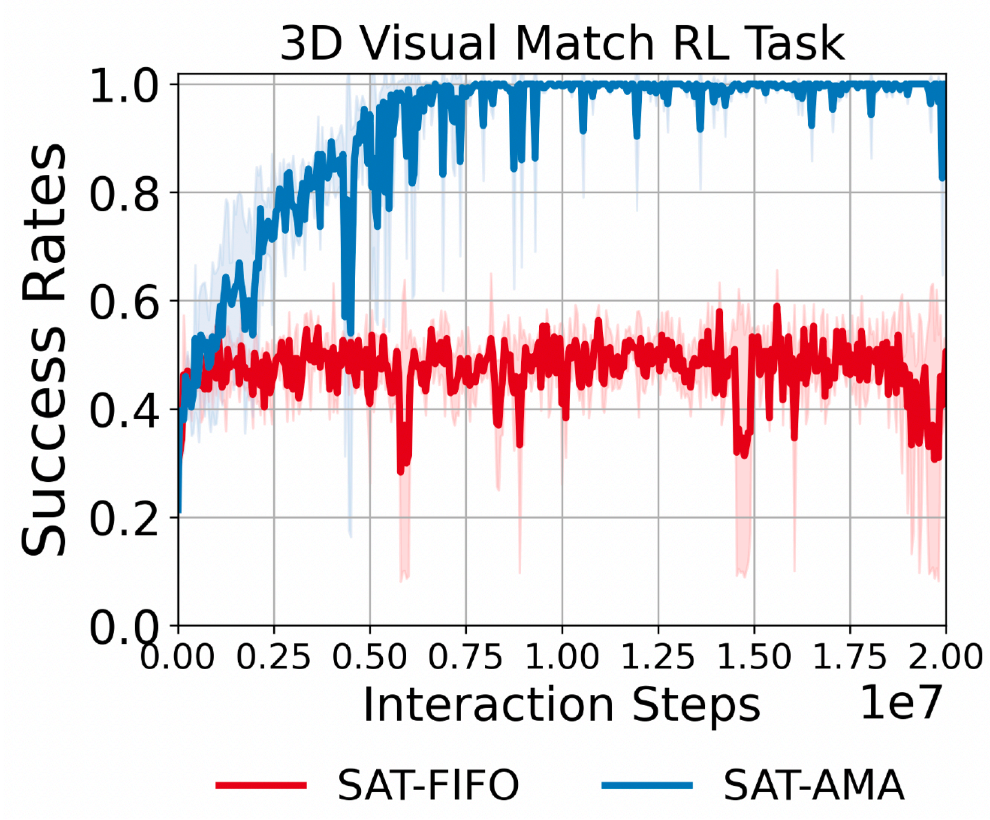
Room 1 (Step 1-5)



Room 2 (Step 6-85)



Room 1 (Step 86-105)



Conclusion

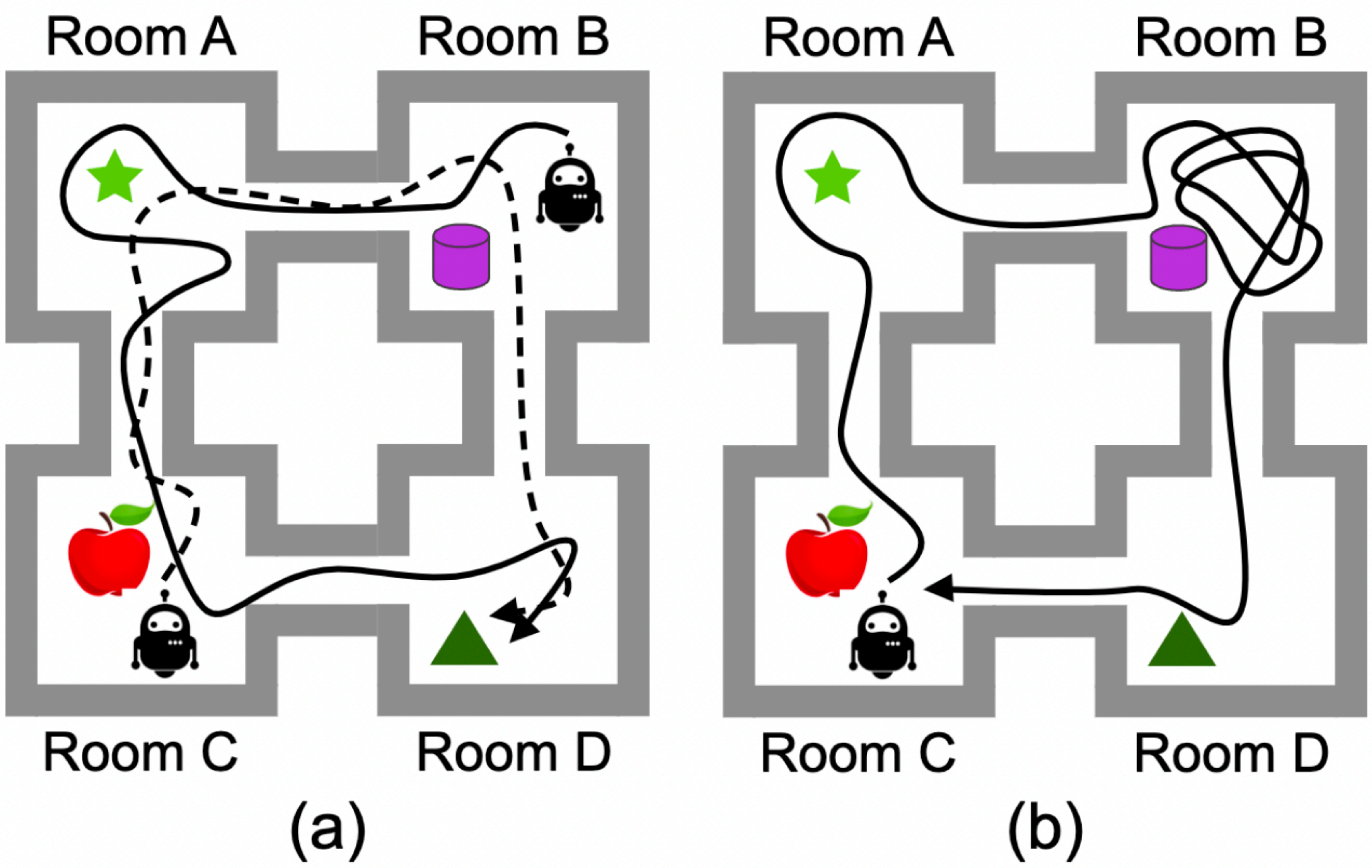
- We introduce SAT for embodied agents, integrating spatial dimension into episodic memory
- We develop SAT-AMA for flexible memory management
- We demonstrate SAT and SAT-AMA applications in various tasks and environments

Thank you!

See you at Halle B

May 9, 2024; 10:45 a.m. ~12:45 p.m.

Home Robot Thought Experiments

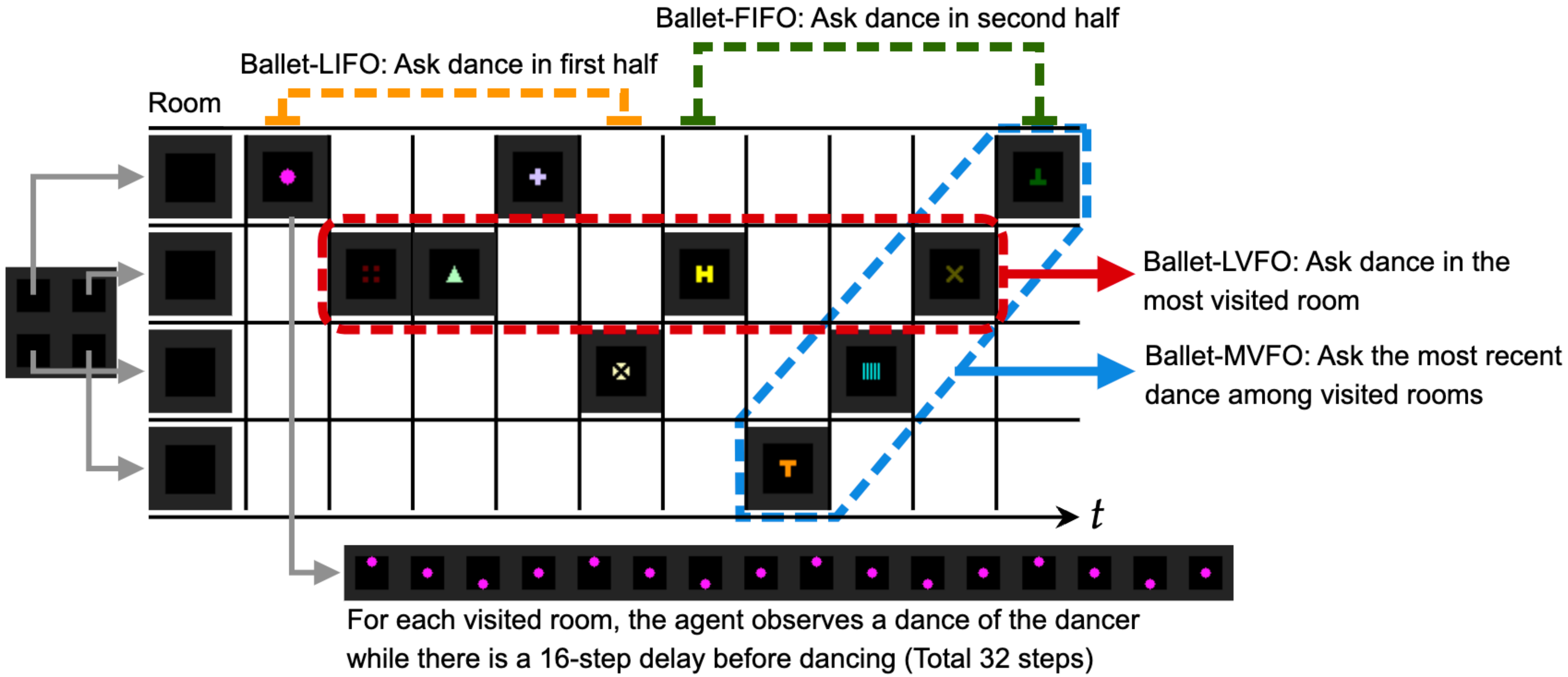


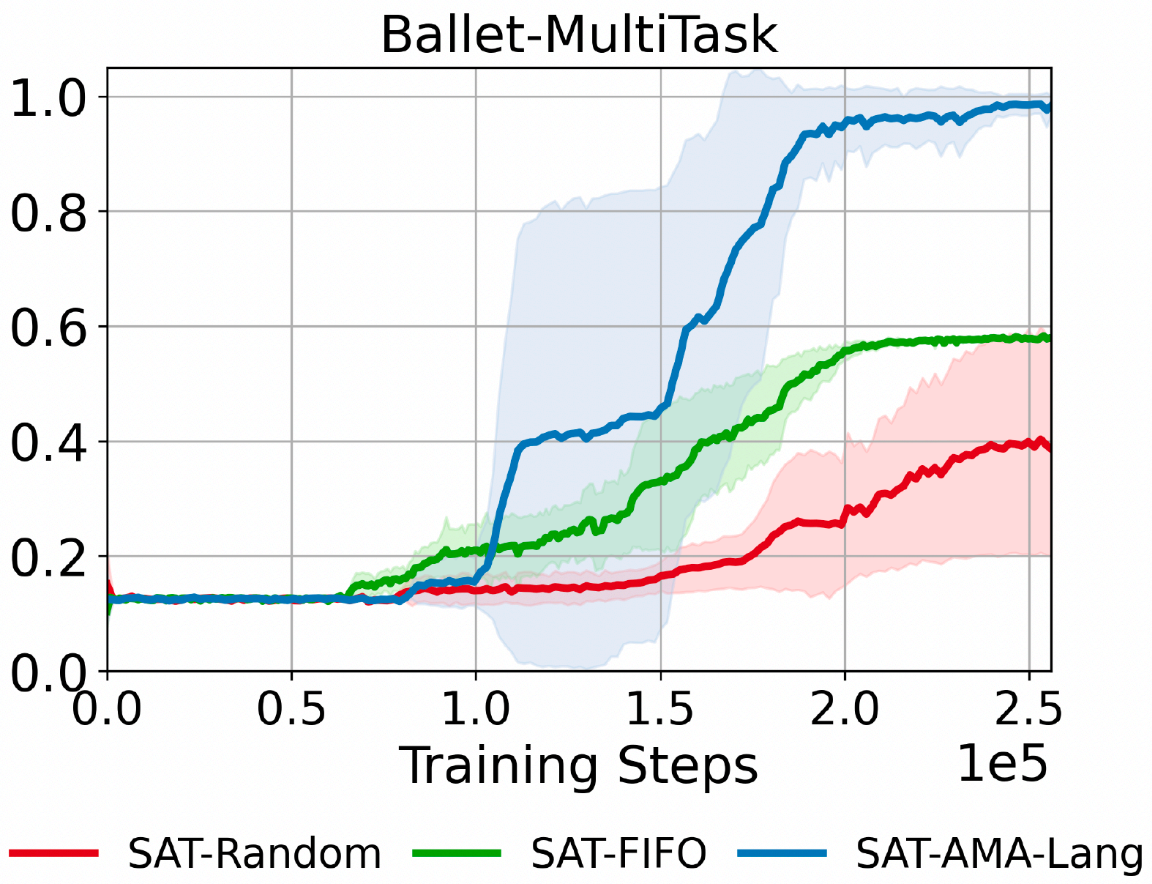
Importance of spatial context in embodied AI

- Episodic memory plays a crucial role in various cognitive processes
- While cognitive science emphasizes the significance of spatial context in episodic memory, current AI system such as transformer dismisses those properties
- It is unclear how to incorporate the spatial axis beyond temporal order alone
- To address this, we explore the use of Spatially-Aware Transformer and investigate the benefits in various tasks

Experiment 2. Learning to select memory allocation strategy

- To validate SAT-AMA, we introduce Ballet-MultiTask
- Each task requires different allocation strategy (4 different strategies available)
- FIFO, LIFO, LVFO (Least Visited First Out), MVFO (Most Visited Frist Out)

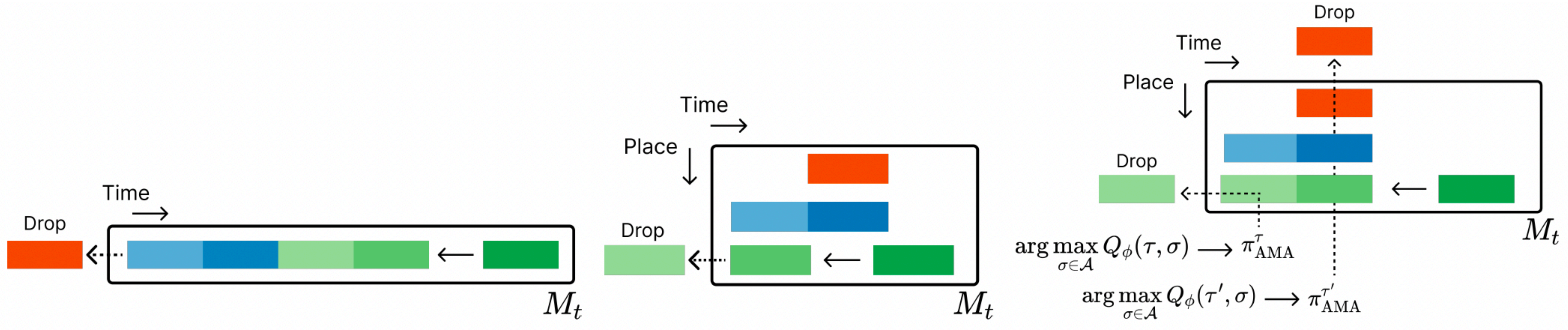




(c)

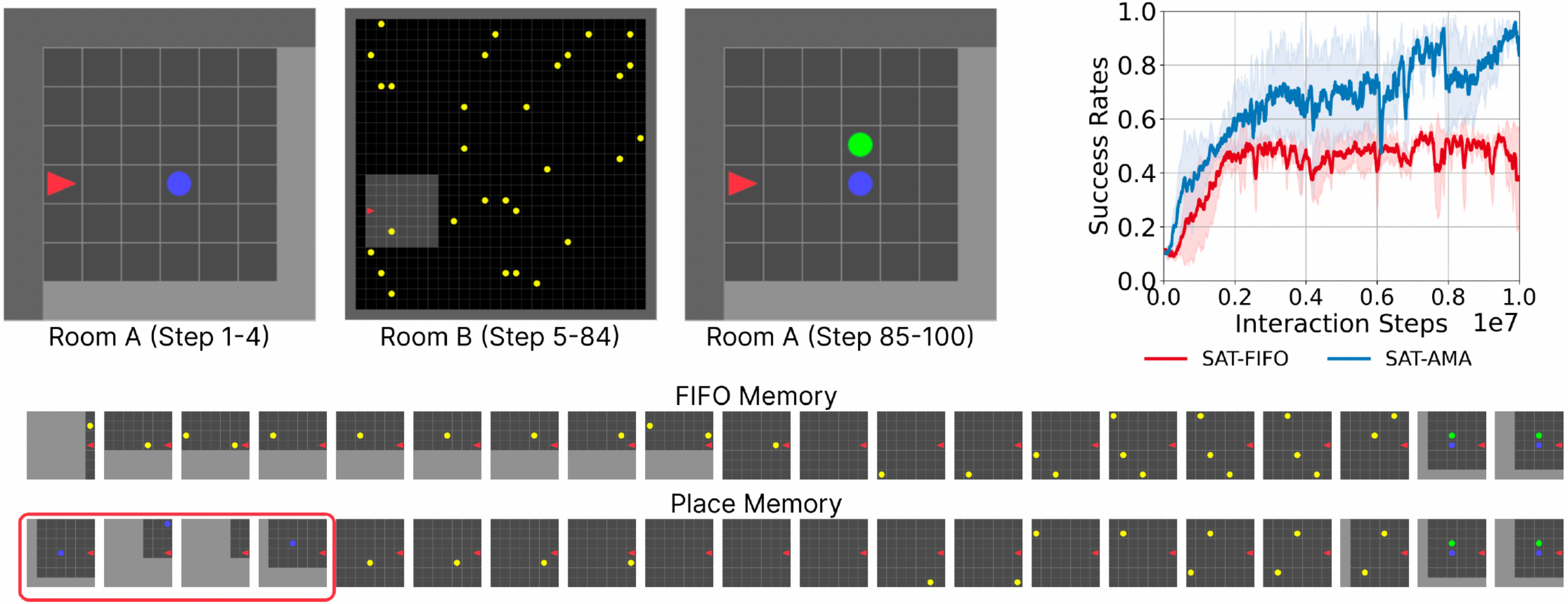
SAT with Adaptive Memory Allocator

- So far, we assumed the experience frames are added to each place memory in the FIFO order
- What if we need memory at the beginning of the episode? Or end of the episode? Or at some place?
- This means we require different memory storing strategies depending on the tasks
- To address this issue, we propose Adaptive Memory Allocator (AMA) which is a learnable policy that chooses memory management strategy based on the task type



(a) FIFO Memory (b) Place Memory (c) Adaptive Memory Allocator (AMA)

Experiment 4. Reinforcement Learning in MiniGrid

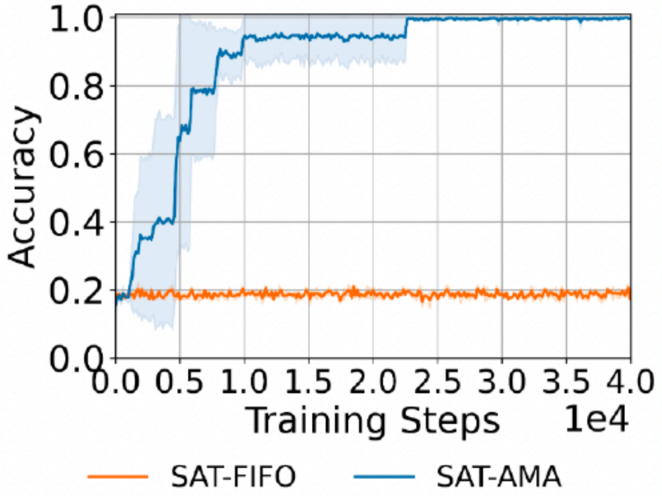


Experiment 6. SAT-AMA in Visual Complex Environment

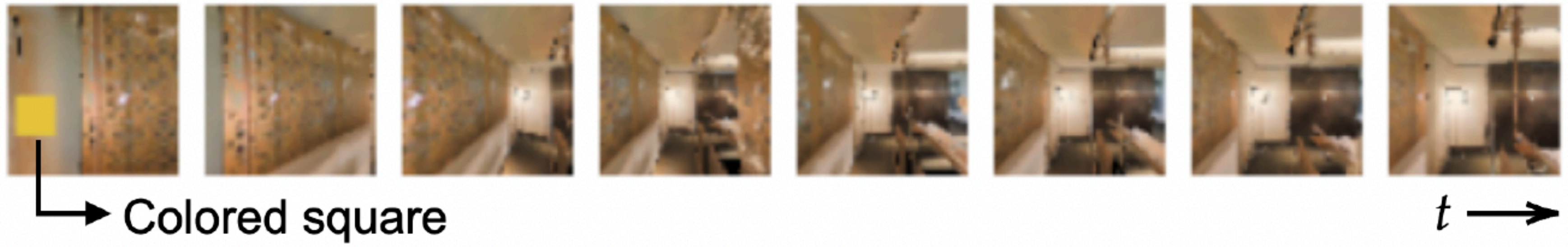
Supervised Prediction

(a) Top-down view and trajectory

(b) Learning curve

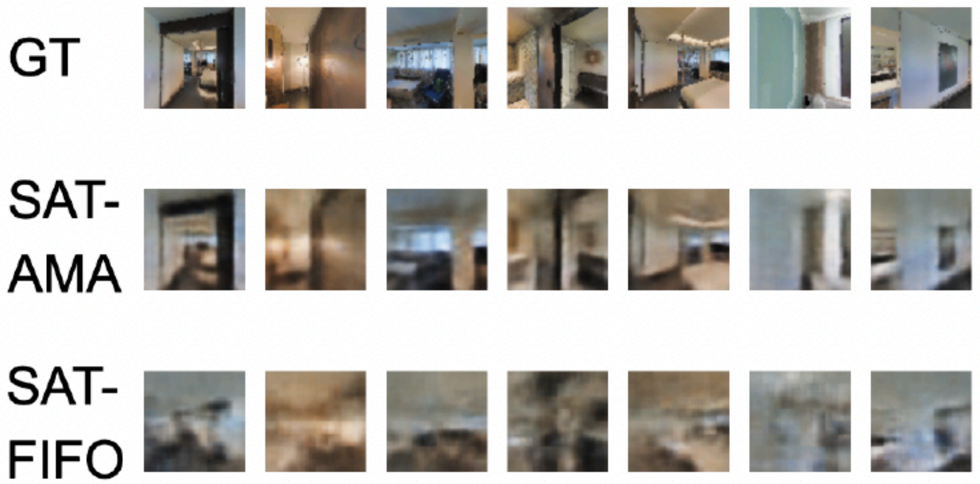


(c) Sample trajectory



Generation

(a) Generation result



(b) Learning curve