Learning Memory Mechanisms for Decision Making through Demonstrations

William Yue[1], Bo Liu[1], Peter Stone[1,2]
The University of Texas at Austin[1], Sony AI[2]
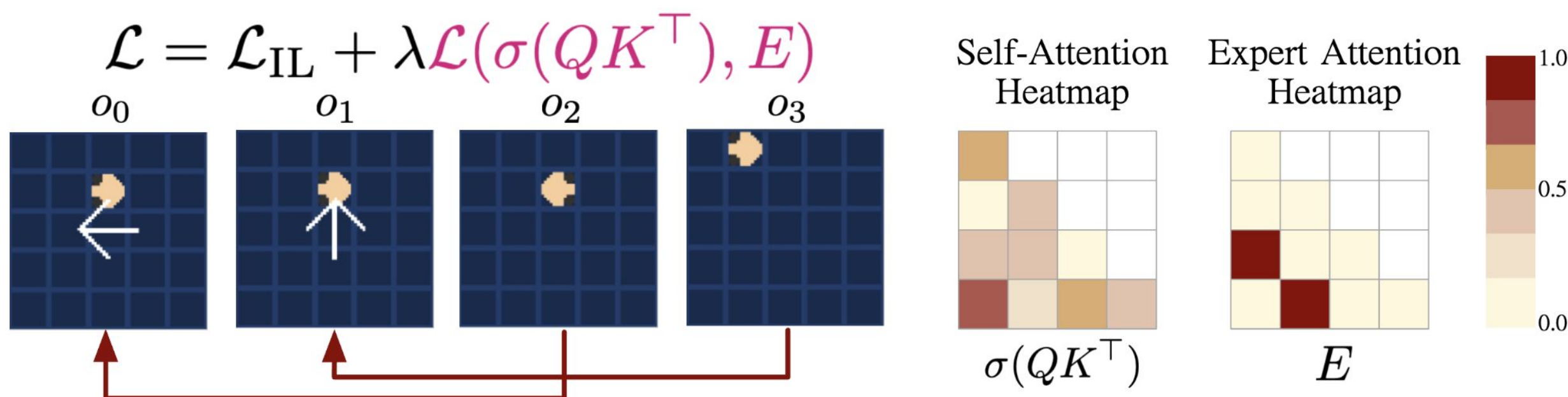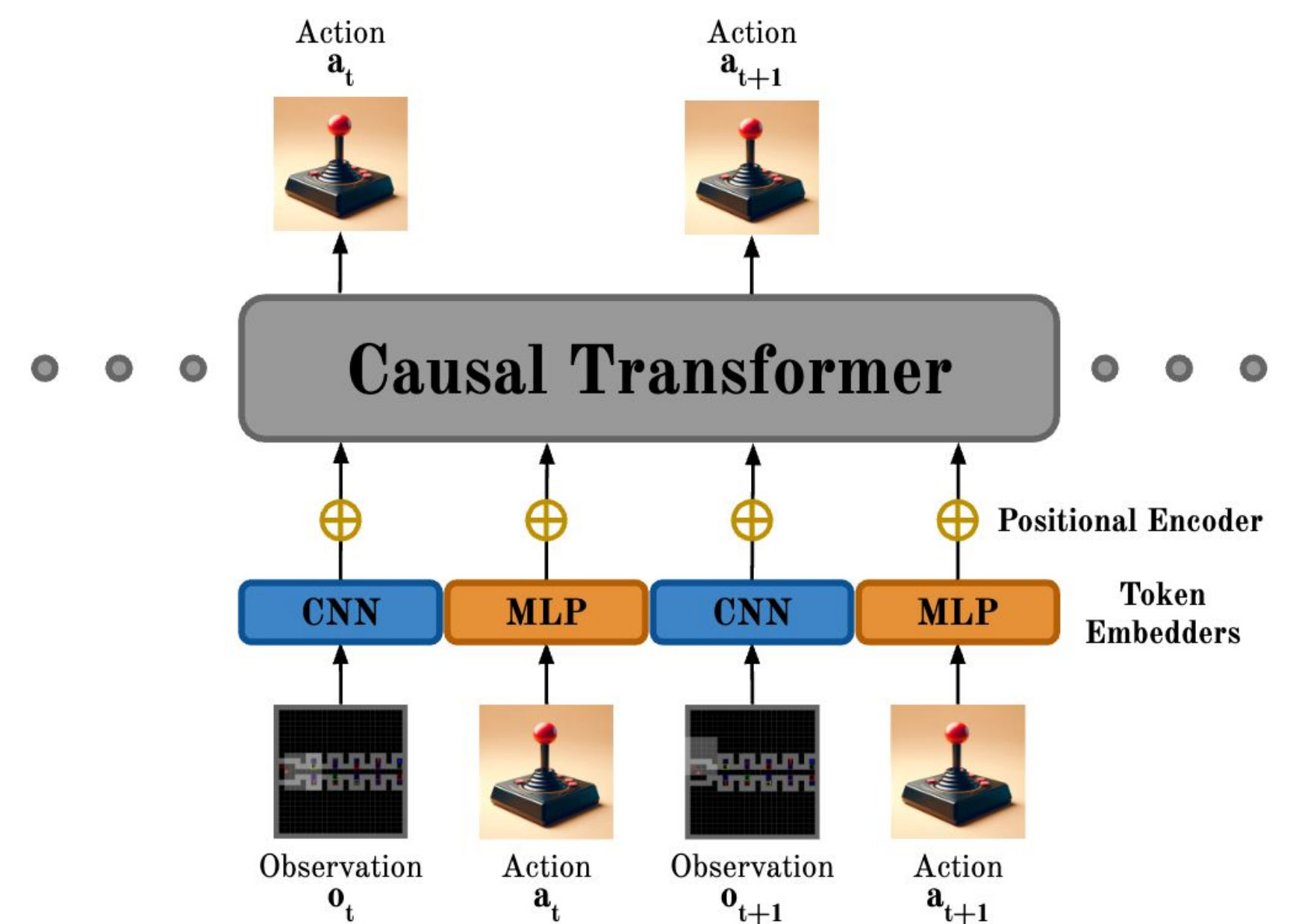
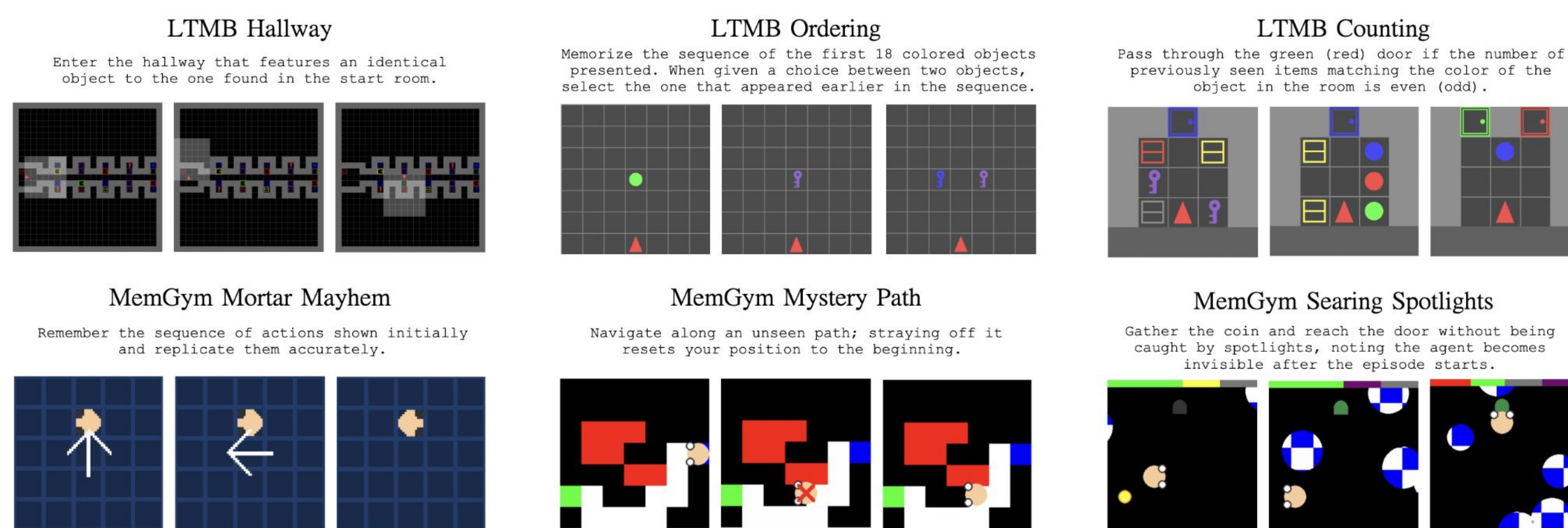# Teaching a Transformer how to use memory leads to much faster learning than letting it learn on its own.
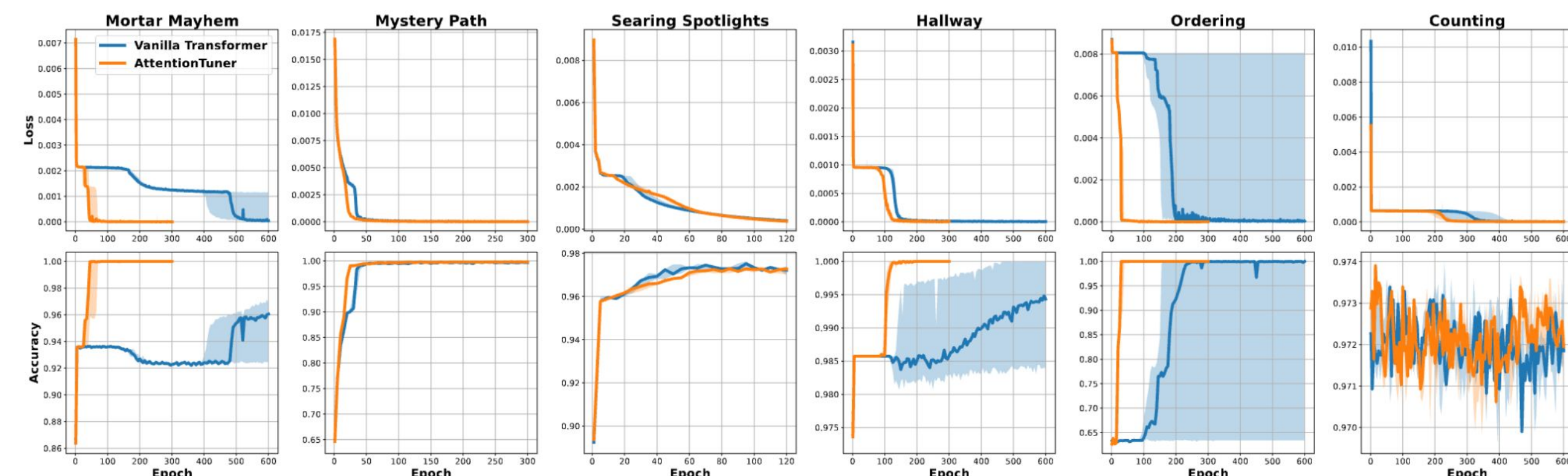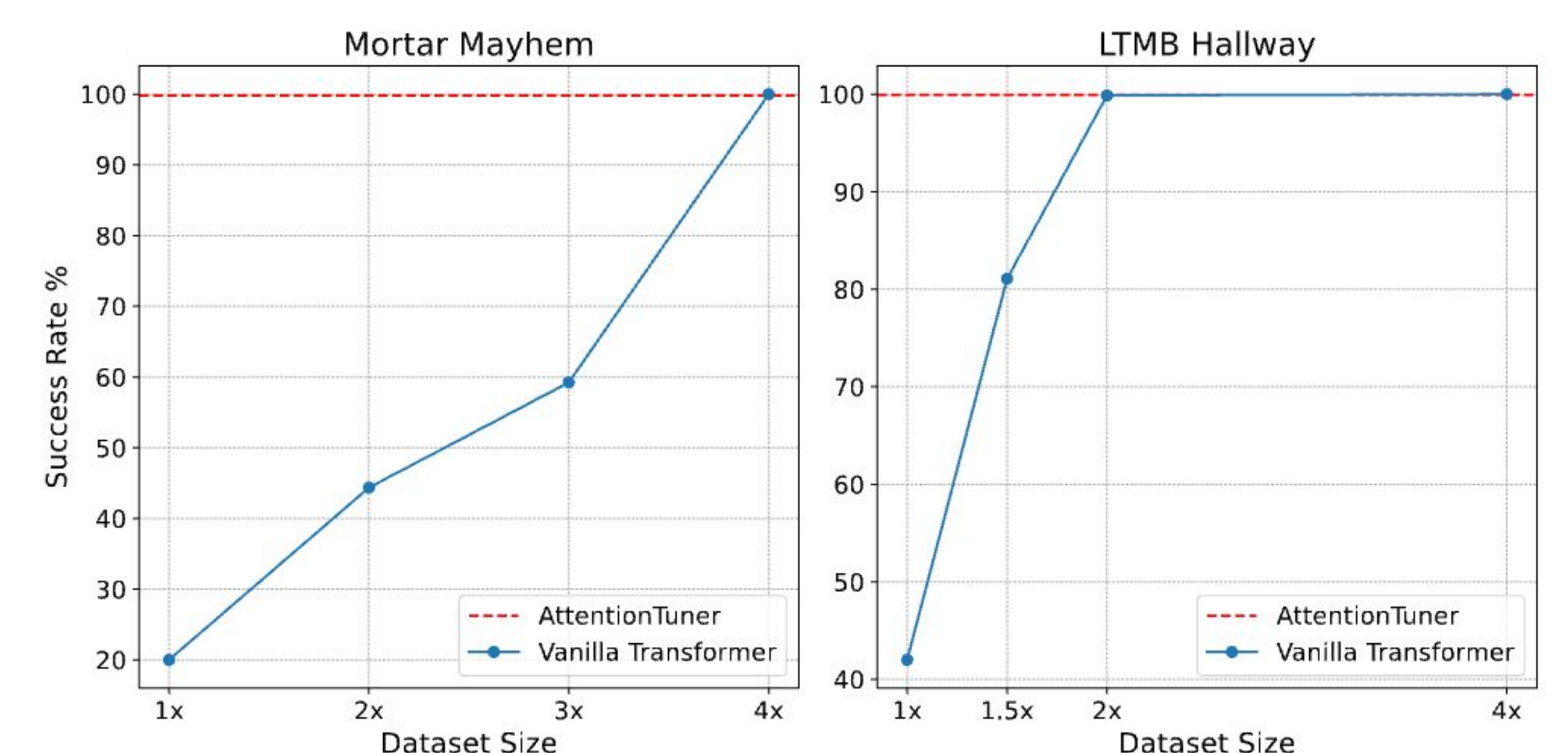
Method

$$\mathcal{L} = \mathcal{L}_{\text{IL}} + \lambda \mathcal{L}(\sigma(QK^\top), E)$$

$o_0 \quad o_1 \quad o_2 \quad o_3$

Self-Attention Heatmap   Expert Attention Heatmap

$\sigma(QK^\top) \qquad E$

Architecture

Action $\mathbf{a}_t$ Action $\mathbf{a}_{t+1}$

Causal Transformer

Positional Encoder

CNN   MLP   CNN   MLP   Token Embedders

Observation $\mathbf{o}_t$ Action $\mathbf{a}_t$ Observation $\mathbf{o}_{t+1}$ Action $\mathbf{a}_{t+1}$

Experiments

LTMB Hallway
Enter the hallway that features an identical object to the one found in the start room.

LTMB Ordering
Memorize the sequence of the first 18 colored objects presented. When given a choice between two objects, select the one that appeared earlier in the sequence.

LTMB Counting
Pass through the green (red) door if the number of previously seen items matching the color of the object in the room is even (odd).

MemGym Mortar Mayhem
Remember the sequence of actions shown initially and replicate them accurately.

MemGym Mystery Path
Navigate along an unseen path; straying off it resets your position to the beginning.

MemGym Searing Spotlights
Gather the coin and reach the door without being caught by spotlights, noting the agent becomes invisible after the episode starts.

Scaling

Results

METHODS	MORTAR MAYHEM	MYSTERY PATH	SEARING SPOTLIGHTS	HALLWAY	ORDERING	COUNTING
VANILLA TRANSFORMER	20.8 ± 42.2	97.3 ± 0.7	62.2 ± 5.5	53.2 ± 28.3	59.4 ± 22.9	6 ± 0.7
ATTENTIONTUNER (OURS)	$\mathbf{99.8 \pm 0.4}$	$\mathbf{98.7 \pm 0.4}$	$\mathbf{64.2 \pm 3.4}$	$\mathbf{99.9 \pm 0.1}$	$\mathbf{99.9 \pm 0.3}$	$\mathbf{6.5 \pm 0.4}$

Table 1: Average success rates and 90% confidence intervals for two different methods—Vanilla Transformer and AttentionTuner (our approach)—across various tasks in the Memory Gym and Long-term Memory Benchmark.

Data Efficiency

Success Rate by Task and Annotation Percentage