Selective Replay Enhances Learning in Online Continual Analogical Reasoning

Tyler L. Hayes and Christopher Kanan

Rochester Institute of Technology, Rochester NY

Code Available

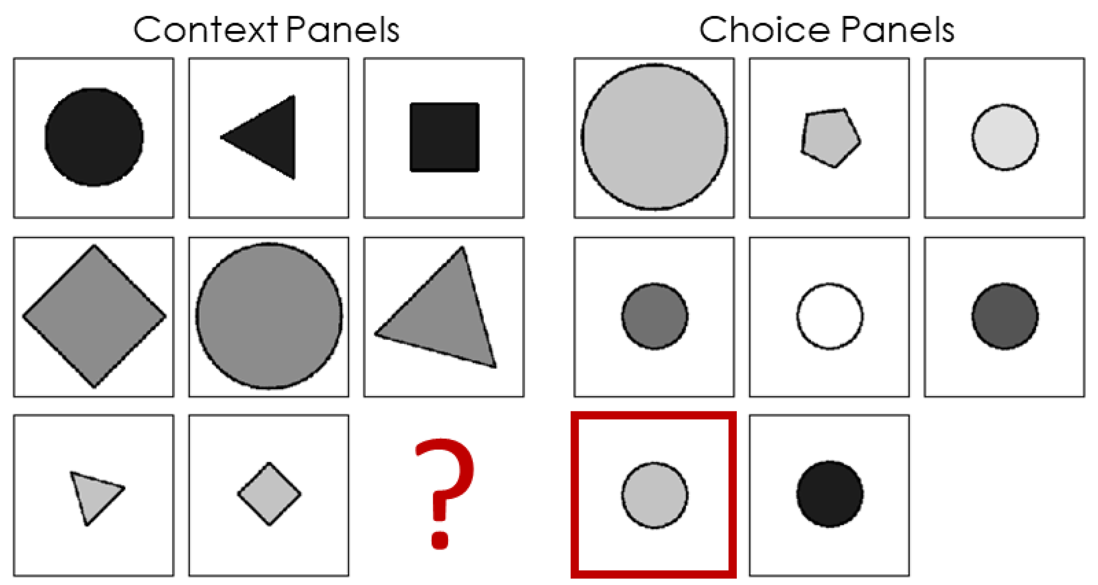
https://github.com/tyler-hayes/Continual-Analogical-Reasoning

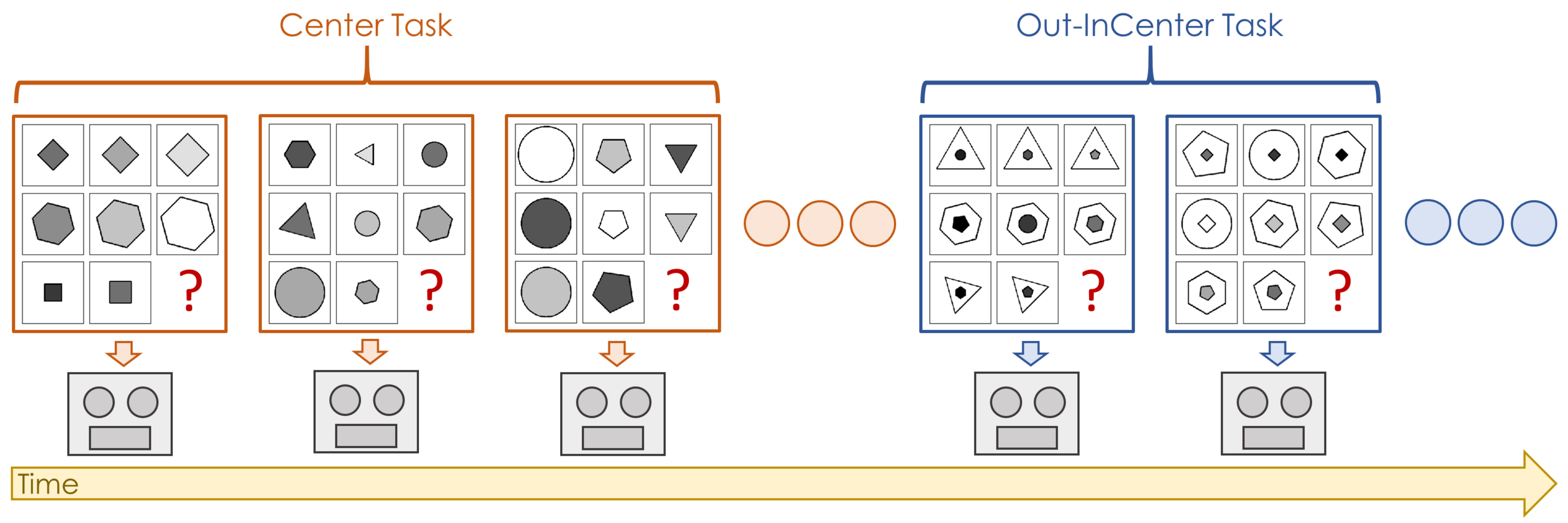
Overview

- In **continual learning**, systems learn from non-stationary data streams or batches without catastrophic forgetting.
- Continual learning has been heavily studied for supervised classification and reinforcement learning, but not yet for abstract reasoning tasks.
- Here, we study continual learning of **analogical reasoning** using **Raven's Progressive Matrices** (RPMs).
- We establish experimental baselines, protocols, and forward and backward transfer metrics to evaluate models.
- We employ **experience replay** to mitigate catastrophic forgetting and demonstrate that selective replay can significantly outperform random selection for the RPM task.

Problem Formulation

RPM Example from RAVEN





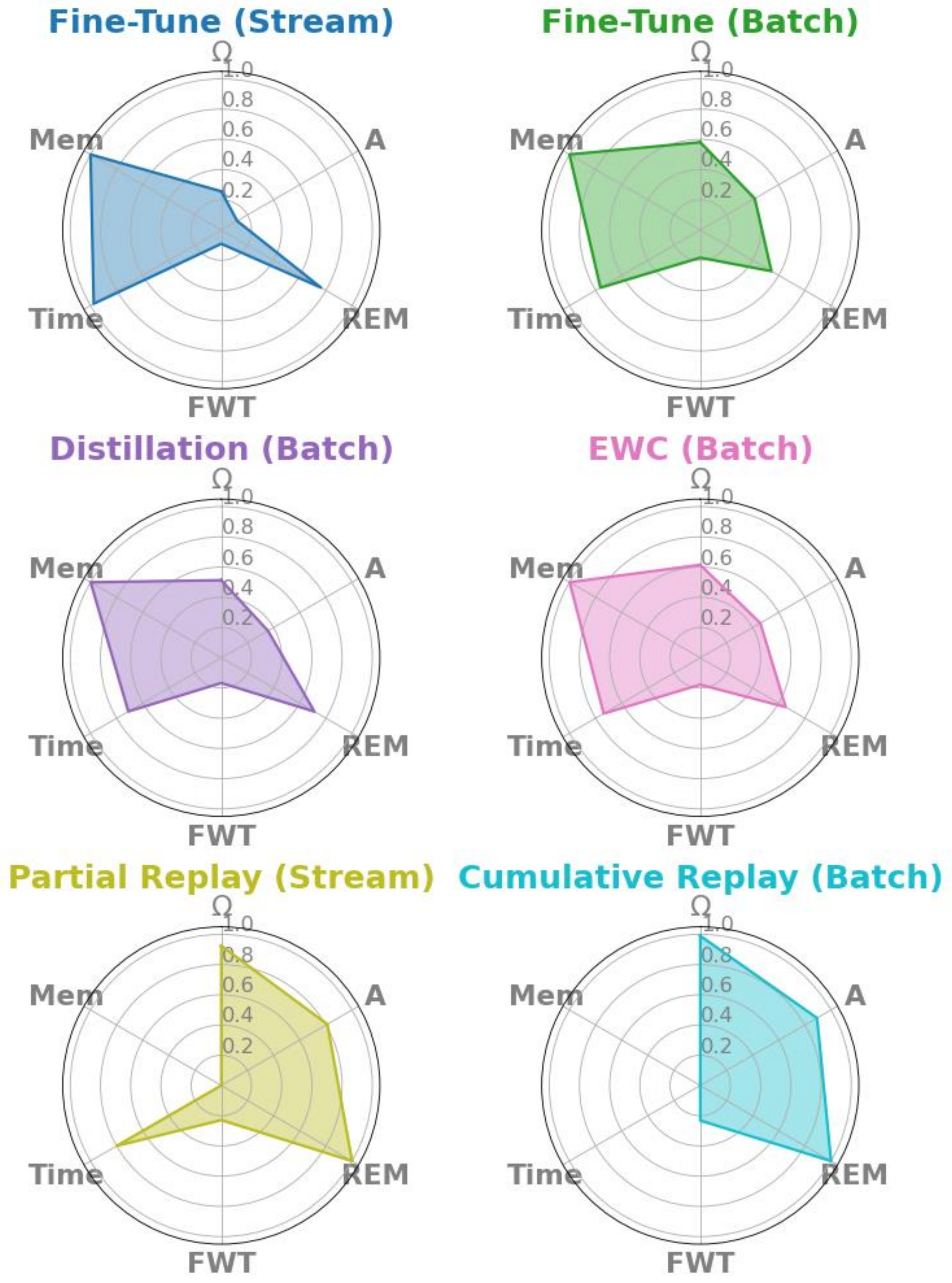
Experimental Setup

- We use the **RAVEN dataset**, which contains 1.12 million images with 70k questions.
- RAVEN is divided into 7 unique tasks. We require models to learn one task at a time either as a whole (**batch**) or one question at a time (**stream**).
- Metrics:
- Ω: performance w.r.t. offline baseline
- A (**white** and cyan): average accuracy
- **REM** (cyan): remembering
- **FWT** (gray): forward transfer

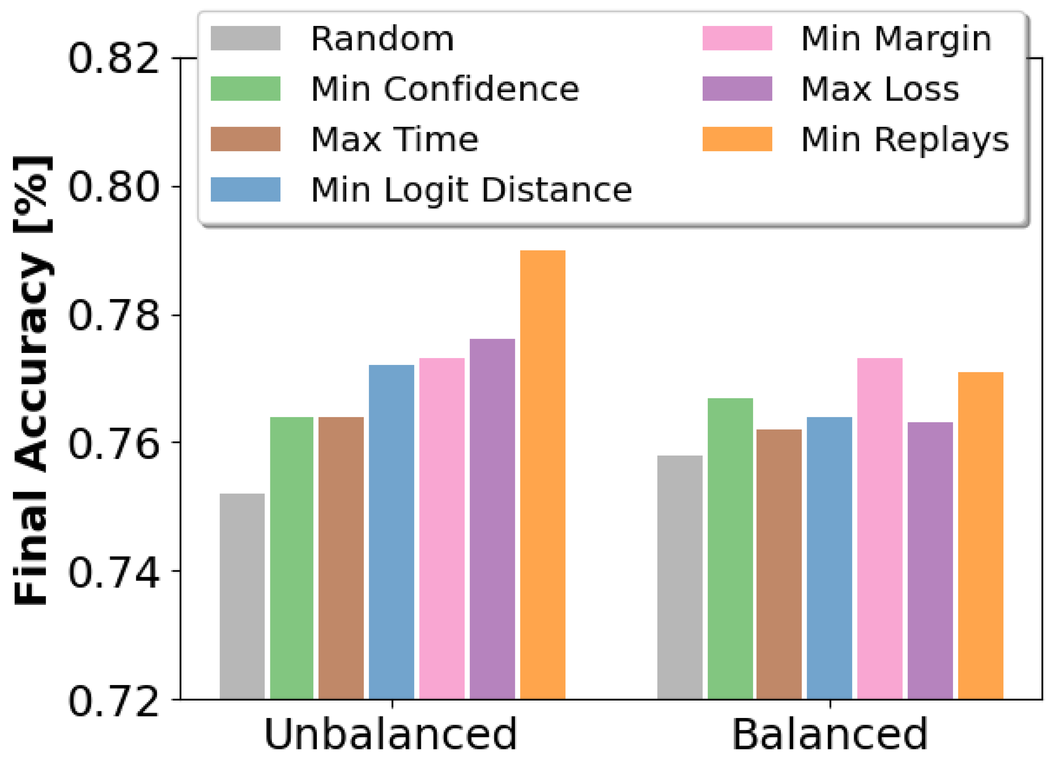
$R_{i,j}$ = accuracy on task j after learning task i

R	Te_1	Te_2	Te_3
Tr_1	R^*	R_{ij}	R_{ij}
Tr_2	R_{ij}	R^*	R_{ij}
Tr_3	R_{ij}	R_{ij}	R^*

Experimental Results



Partial Replay Performance



All Partial Replay selection methods are **statistically significant** from Random and Min Replays at a significance level of 0.1 after Holm-Bonferroni correction.

Models & Baselines

- We extend the **Rel-Base model** to the continual learning setting. Rel-Base processes image frames independently before passing them to a sequence encoder to extract relationships and provide score predictions.
- We study three continual learning models:
- **Distillation** optimizes a classification and distillation loss, where soft targets are computed as the scores of the model from the previous time-step.
- **EWC** uses a quadratic regularization term to encourage weights to remain close to their previous values.
- **Partial Replay** fine-tunes on all new data and a subset of previous data.
  - We study **seven policies** for choosing which samples to replay: random, minimum logit distance, minimum confidence, minimum margin, maximum loss, maximum time since last replay, and minimum replays.
  - We study the policies in both balanced and unbalanced settings.
- We also study three baselines:
- **Fine-Tune** (lower bound) does not use any mechanisms to mitigate catastrophic forgetting.
- **Cumulative Replay** fine-tunes a model with all new and old data.
- **Offline** (upper bound) is a conventional network trained offline on all data.

Summary

- **Replay**-based learners perform the best for continually solving RPM puzzles.
- Selectively choosing which samples to replay can yield **statistically significant performance improvements over uniform random sampling**. This is interesting as selective replay for image classification has yielded marginal benefits.
- Future work consists of designing and testing **more sophisticated network architectures and continual learning strategies** for analogical reasoning.

Acknowledgements:

We thank DARPA L2M, AFOSR, and NSF for financially supporting this research.