FACEBOOK AI

Physical Reasoning Using Dynamics-Aware Models

Eltayeb Ahmed
Anton Bakhtin
Laurens van der Maaten
Rohit Girdhar

Facebook AI Research

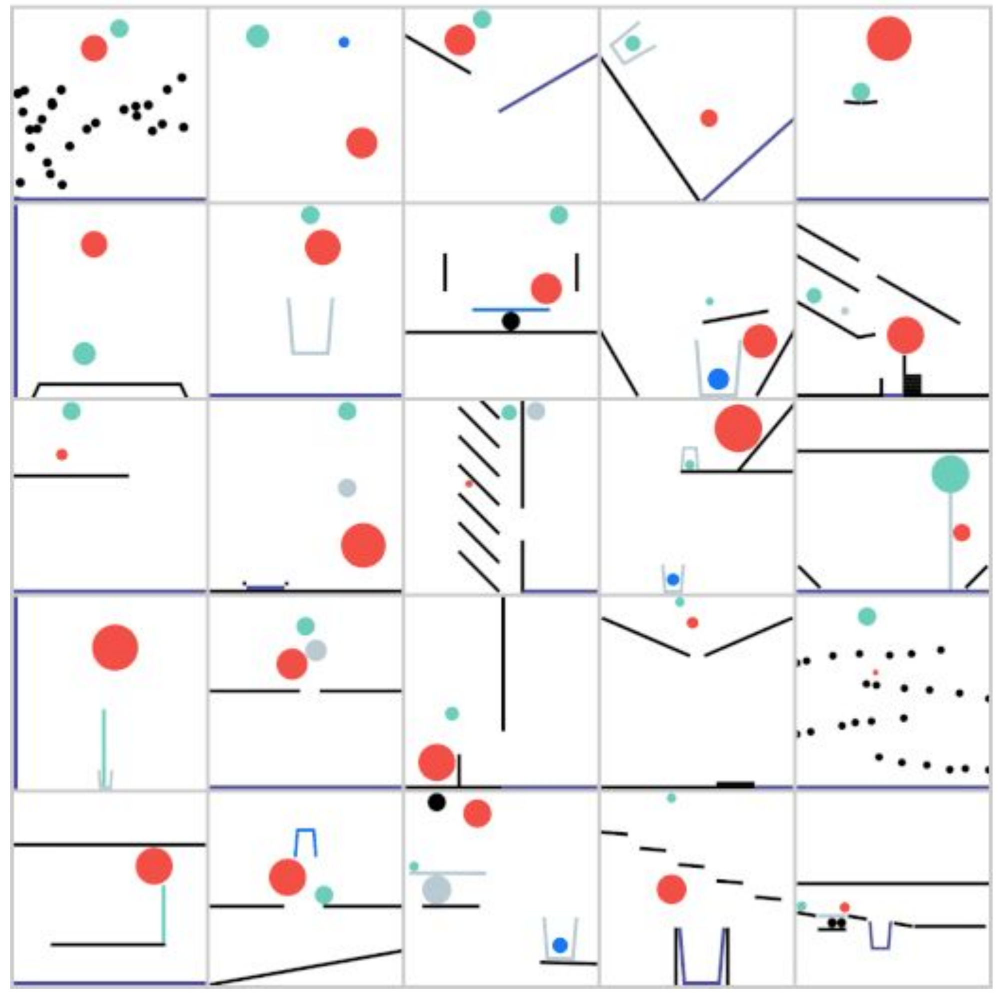
We Use Contrastive Learning To Learn Dynamics-Aware Embeddings that are Useful for Downstream Tasks

A common approach to solving physical-reasoning tasks is to train a value learner on example tasks. A limitation of such an approach is that it requires learning about object dynamics solely from reward values assigned to the final state of a rollout of the environment. We address this limitation by augmenting the reward value with self-supervised signals about object dynamics. Specifically, we train the model to characterize the similarity of two environment rollouts, jointly with predicting the outcome of the reasoning task. We experiment using both a hand-crafted distance measure between rollouts and a learned measure directly from pixels using a contrastive formulation. Empirically, we find that this approach leads to substantial performance improvements on the PHYRE benchmark for physical reasoning establishing a new state-of-the-art.

The PHYRE Benchmark

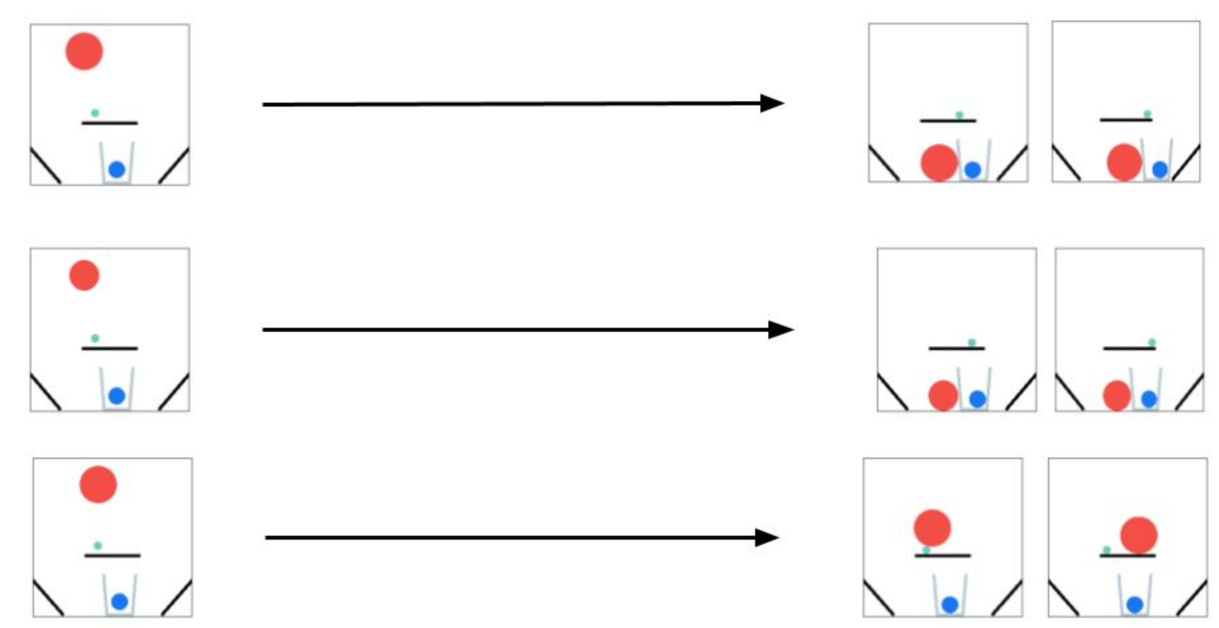
PHYsical REasoning

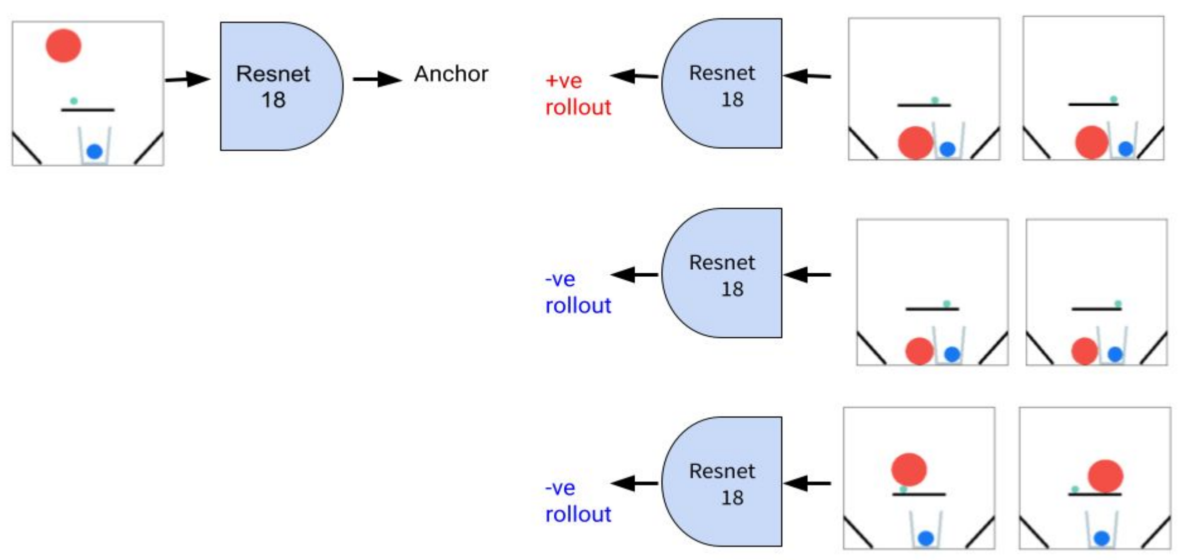
The PHYRE benchmark is a 2-D physical reasoning puzzle. It consists of many tasks. Each task has a unique initial scene with multiple physical objects in it. The agent must attempt to solve the task by inserting one or two red balls into the scene. A physical simulation in then rolled out and the agent is deemed to have solved the task if the green and blue objects touch



Some of the tasks from the PHYRE benchmark

The Contrastive Task



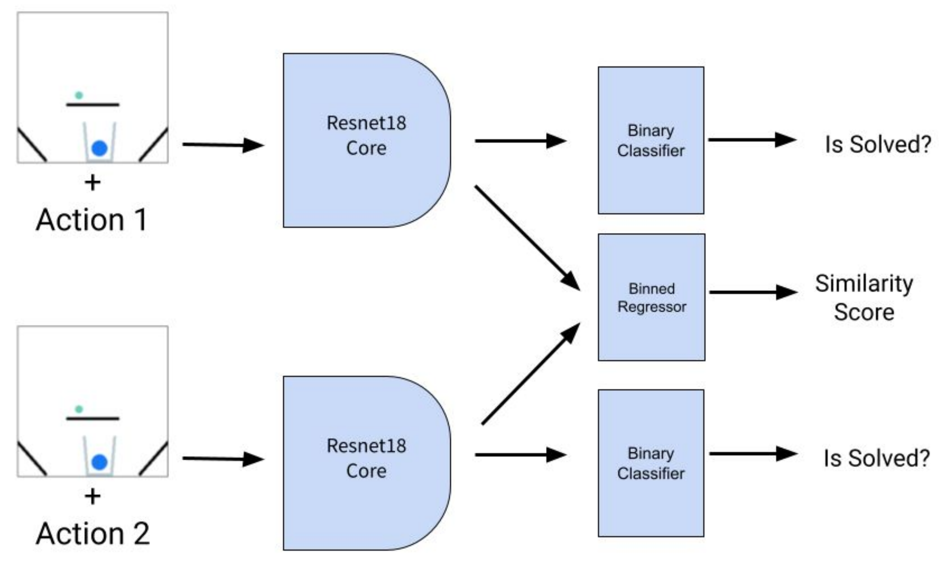


The embeddings used in the InfoNCE loss

In the above Image for each action (shown on the left) we sample two consecutive images from the resulting rollout. The agent is tasked with matching each action to the resulting rollout in contrastive fashion using the InfoNCE loss as an auxiliary task.

The Hand-Crafted Task

Using additional supervision from the simulator we define a metric of similarity between any two action rollouts and then at train-time use the agent network in a siamese fashion to the predict the similarity score.



Emperical Results

	1B-Within	1B-Cross	2B-Within	2B-Cross
DQN Baseline	77.6 ± 1.1	36.8 ± 9.7	67.8 ± 1.5	23.2 ± 9.1
Dec[Joint]	80.0 ± 1.2	40.2 ± 8.0	–	–
RPIN	85.2 ± 0.7	**42.2** ± 7.1	–	–
Ours (Hand Crafted)	85.0 ± 1.1	38.6 ± 8.4	74.0 ± 1.5	23.2 ± 8.8
Ours (Self-Supervised)	**86.2** ± 0.9	41.9 ± 8.8	**77.6** ± 1.4	**24.3** ± 10.0

References

Bakhtin, A., van der Maaten, L., Johnson, J., Gustafson, L., and Girshick, R. PHYRE: A new benchmark for physical reasoning. In NeurIPS, 2019

Qi, H., Wang, X., Pathak, D., Ma, Y., and Malik, J. Learning long-term visual dynamics with region proposal interaction networks. In ICLR, 2021.

Bakhtin, A., van der Maaten, L., Johnson, J., Gustafson, L., and Girshick, R. PHYRE: A new benchmark for physical reasoning. In NeurIPS, 2019

Take a photo to learn more:

FACEBOOK AI

Title of research poster in Optimistic Display 60pt should not exceed two lines

Author Name, Author Name,
Author Name, Author Name,
Author Name

Facebook AI Research, Facebook Reality Labs, Georgia Institute of Technology, Simon Fraser University, Intel Labs, UC Berkeley

Abstract

Text simplification is crucial to making content accessible to everyone.

In 1964 Peter Higgs published his second paper in Physical Review Letters describing Higgs mechanism which predicted a new massive spin-zero boson. In 1964 Peter Higg wrote his paper explaining Higgs mechanism. Higgs mechanism predicted a new elementary particle.

Different audiences have different simplification needs:

1. We adapt simplification models to target audiences by controlling length, paraphrasing, lexical and syntactic complexity.
2. We analyze the effectiveness of the control parameters and their influence on the simplified text
3. Our model advances the state of the art to 41.87 SARI on the WikiLarge test set

Method

Concatenate a plain text special token to the source sentence. The model associates this token to the desired attribute.

- Train time:
  — The Special token provides the model with the actual value of the parameter on the ground truth simplification.
    - Source: <NbChars_15> This is an extremely complex sentence.
    - Ground truth: This is simple.
- Test time:
  — The Special token is set to the desired value
    - Source: <NbChars_?> This is a hard sentence to be simplified.

Control Parameters

We cover 4 attributes of simplification using 4 parameters

- Length: NbChars
  — Length in number of characters.
- Paraphrasing: LevSim
  — Levenshtein similarity between source and target.
- Lexical Complexity: WordRank
  — Aggregation of frequency table ranks (word rarity).
- Syntactic Complexity: DepTreeDepth
  — Depth of the dependency tree (structure complexity).

Overall Performance

State-of-the-art performance of 41.87 SARI. SARI compares system simplifications with the ground truths (similar to BLEU).

Sample		5	10	20	50
Sample	Sample	3542.3	2286.9	1685.9	1281.7
	Sample	0.024	0.059	0.087	0.140
Sample	Sample	205.9	179.4	95.3	92.8
	Sample	0.517	0.503	0.563	0.566
Sample	Sample	4.9	4.02	3.84	3.98
	Sample	0.823	0.851	0.855	0.86

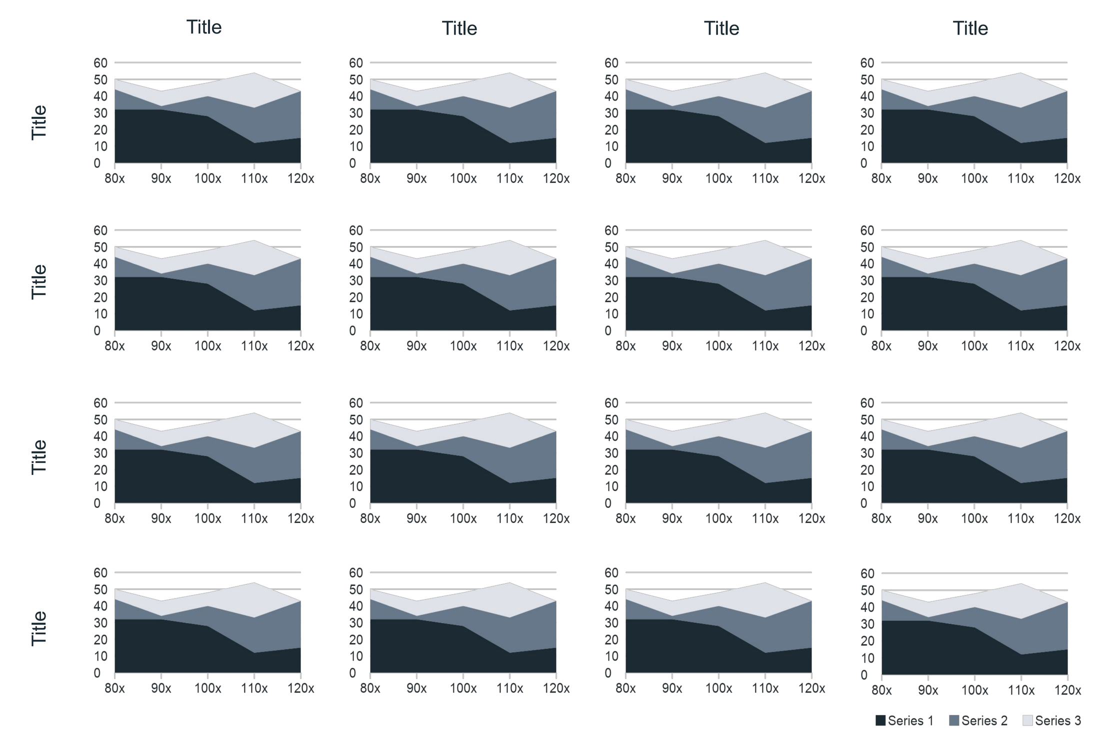
Ablation Studies

Parameter each increase the SARI simplification score.

Sample		5	10	20	50
Sample	Sample	3542.3	2286.9	1685.9	1281.7
	Sample	0.024	0.059	0.087	0.140
Sample	Sample	205.9	179.4	95.3	92.8
	Sample	0.517	0.503	0.563	0.566
Sample	Sample	4.9	4.02	3.84	3.98
	Sample	0.823	0.851	0.855	0.86

Influence of the Parameters

Do the parameters actually control their respective attributes?



Yes they do! Although not with the same effectiveness.

Examples

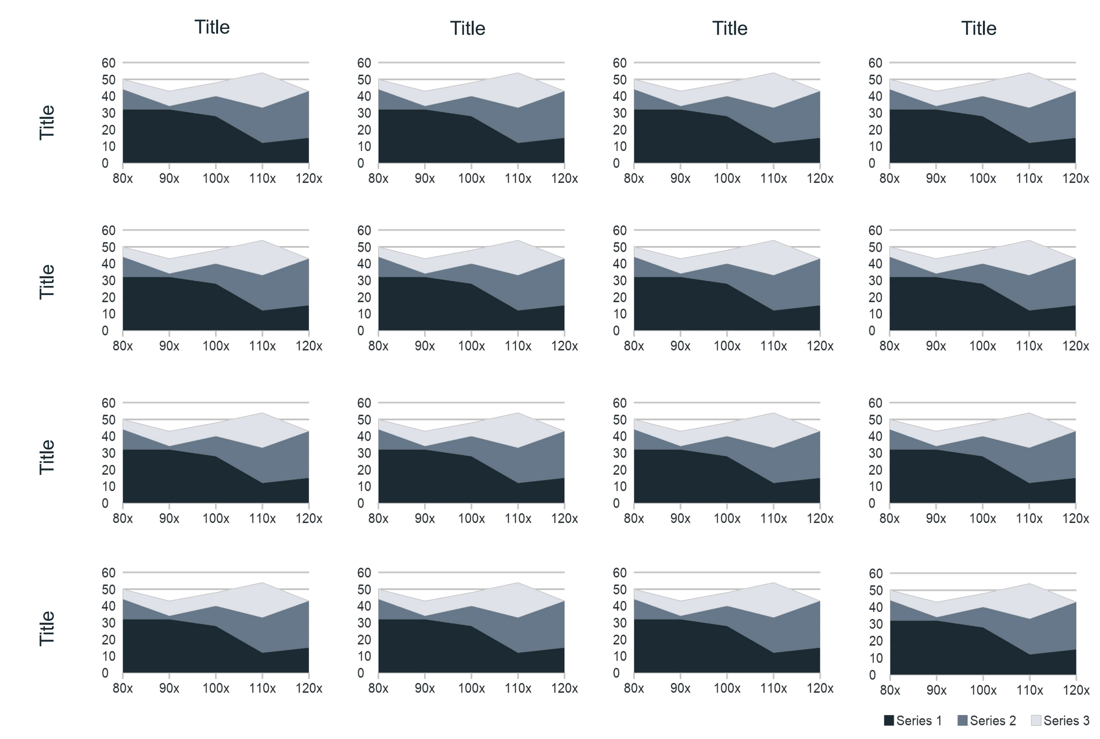
Sample		5	10	20	50
Sample	Sample	3542.3	2286.9	1685.9	1281.7
	Sample	0.024	0.059	0.087	0.140
Sample	Sample	205.9	179.4	95.3	92.8
	Sample	0.517	0.503	0.563	0.566
Sample	Sample	4.9	4.02	3.84	3.98
	Sample	0.823	0.851	0.855	0.86
Sample		5	10	20	50
Sample	Sample	3542.3	2286.9	1685.9	1281.7
	Sample	0.024	0.059	0.087	0.140
Sample	Sample	205.9	179.4	95.3	92.8
	Sample	0.517	0.503	0.563	0.566
Sample	Sample	4.9	4.02	3.84	3.98
	Sample	0.823	0.851	0.855	0.86

Acknowledgements

This research is part of a collaboration between Facebook AI Research Inria, UNAPEI (national association for people with cognitive disabilities) and the French Government.

Influence of the Parameters

Do the parameters actually control their respective attributes?



Yes they do! Although not with the same effectiveness.

Examples

Sample		5	10	20	50
Sample	Sample	3542.3	2286.9	1685.9	1281.7
	Sample	0.024	0.059	0.087	0.140
Sample	Sample	205.9	179.4	95.3	92.8
	Sample	0.517	0.503	0.563	0.566
Sample	Sample	4.9	4.02	3.84	3.98
	Sample	0.823	0.851	0.855	0.86
Sample		5	10	20	50
Sample	Sample	3542.3	2286.9	1685.9	1281.7
	Sample	0.024	0.059	0.087	0.140
Sample	Sample	205.9	179.4	95.3	92.8
	Sample	0.517	0.503	0.563	0.566
Sample	Sample	4.9	4.02	3.84	3.98
	Sample	0.823	0.851	0.855	0.86

Acknowledgements

This research is part of a collaboration between **Facebook AI Research Inria, UNAPEI** (national association for people with cognitive disabilities) and the **French Government**.

Only use this layout if you have an excess of information!
We recommend using pages 1–2, with larger type and 3 columns

Take a photo to learn more:


Resnet 18
Resnet 18