DeepMind

Proving Theorems using Incremental Learning and Hindsight Experience Replay

Eser Aygün, Ankit Anand, Laurent Orseau,
Xavier Glorot, Stephen McAleer, Vlad Firoiu,
Lei Zhang, Doina Precup, Shibl Mourad

Introduction

Motivation

Automated theorem proving did not benefit as much from the deep learning revolution. In first-order logic, heuristic approaches are still the state-of-the-art.

Contribution

We propose a simple method that addresses the main challenges of the first-order automated theorem proving domain. We show that our method can reach state-of-the-art performance without depending on hand-crafted heuristics.

Challenges

Reward is sparse

Existing machine learning systems get a meaningful signal only when a proof is found. Failed attempts are wasted.

There is no curriculum

Without a curriculum, crossing gaps in difficulty can be impossible. Synthetic data does not help enough.

Solution

Hindsight experience replay (HER)

Do not teach the network how to reach a proof. Teach it how to reach any given clause. Thousands of clauses are generated in each attempt, regardless of success. HER allows us to take advantage of all of them. This means faster training and continuous improvement.

Incremental learning (IL)

Do not split the dataset into train and test. Start with a certain budget and solve as many theorems as possible within that budget. We use a smart budget allocation strategy that allows our prover to improve faster. IL also allows us to be more fair towards existing heuristic methods.

Experiments

E basic

State-of-the-art heuristic prover with clause scoring heuristics turned off.

TRAIL (Abdelaziz et al., 2022)

Recent work that uses reinforcement learning.

E best

Best run of E out of various modes and time limits.

Ours

Incremental learning with hindsight experience replay.
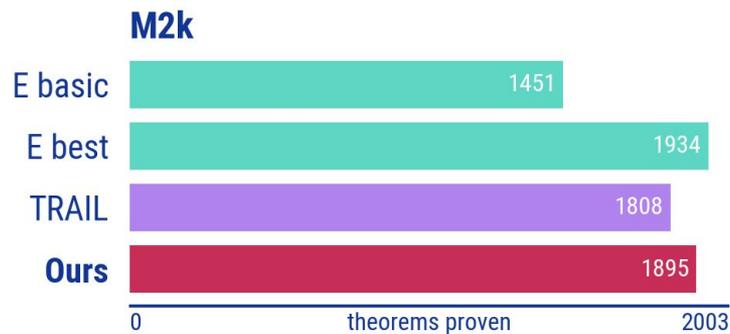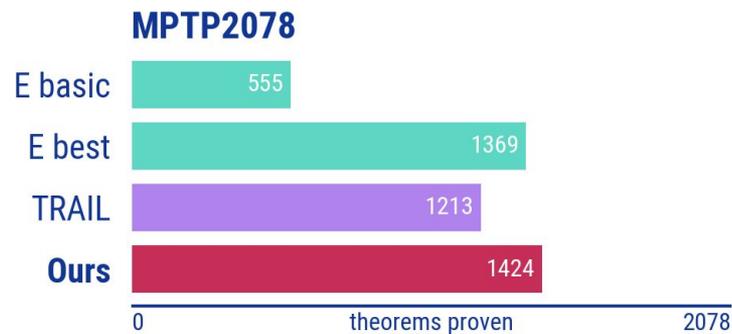
Setup

We ran E on each theorem separately with time limits ranging from 100 seconds to seven days.

We did incremental learning for seven days where each attempt was limited to 100 seconds.

TRAIL was trained for 20 iterations where each attempt was limited to 100 seconds.
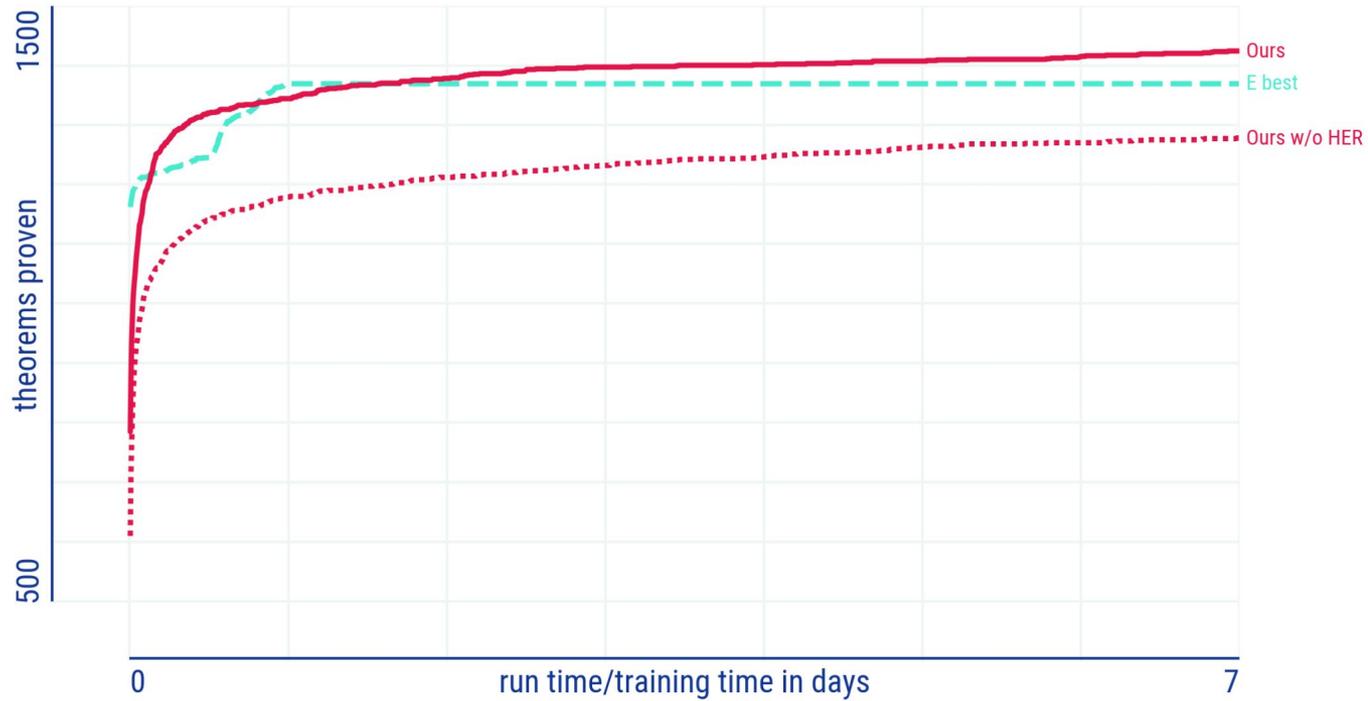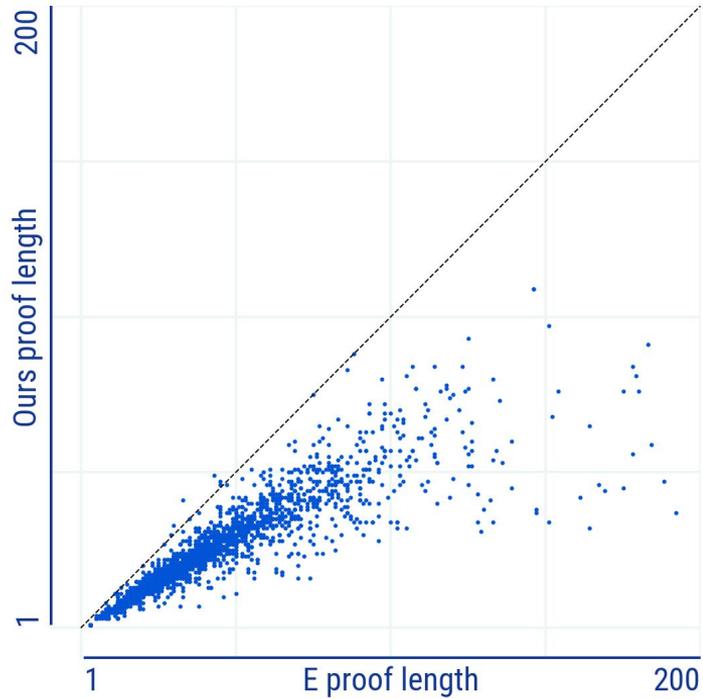
Results: Theorems Proven

MPTP2078

E basic	555
E best	1369
TRAIL	1213
Ours	1424

0   theorems proven   2078

M2k

E basic	1451
E best	1934
TRAIL	1808
Ours	1895

0   theorems proven   2003

Mizar40

E best	23173
Ours	24363

0   theorems proven   57880

Results: Timeline

MPTP2078

y-axis: theorems proven (500, 1500)

x-axis: run time/training time in days (0 to 7)

Lines labeled: Ours, E best, Ours w/o HER

Results: Proof Lengths

MPTP2078 + M2k

Thank You!

Please check out our paper for more details
and reach out to us for your questions.

E-mail addresses:

eser@deepmind.com
anandank@deepmind.com
lorseau@deepmind.com