



LeanAgent: Lifelong Learning for Formal Theorem Proving

Adarsh Kumarappan*, Mo Tiwari*, Peiyang Song, Robert Joseph George, Chaowei Xiao, Anima Anandkumar



Summary

- Current LLMs only provide informal, textual proofs of mathematical theorems
- Existing progress in formal mathematical systems like Lean has been scant
- We propose a lifelong learning framework that allows LLMs to learn and re-use easier theorems in the proofs of harder theorems
- LeanAgent successfully generates formal proofs for 155 theorems across 23 diverse Lean repositories where formal proofs were previously missing, many from advanced mathematics

Introduction

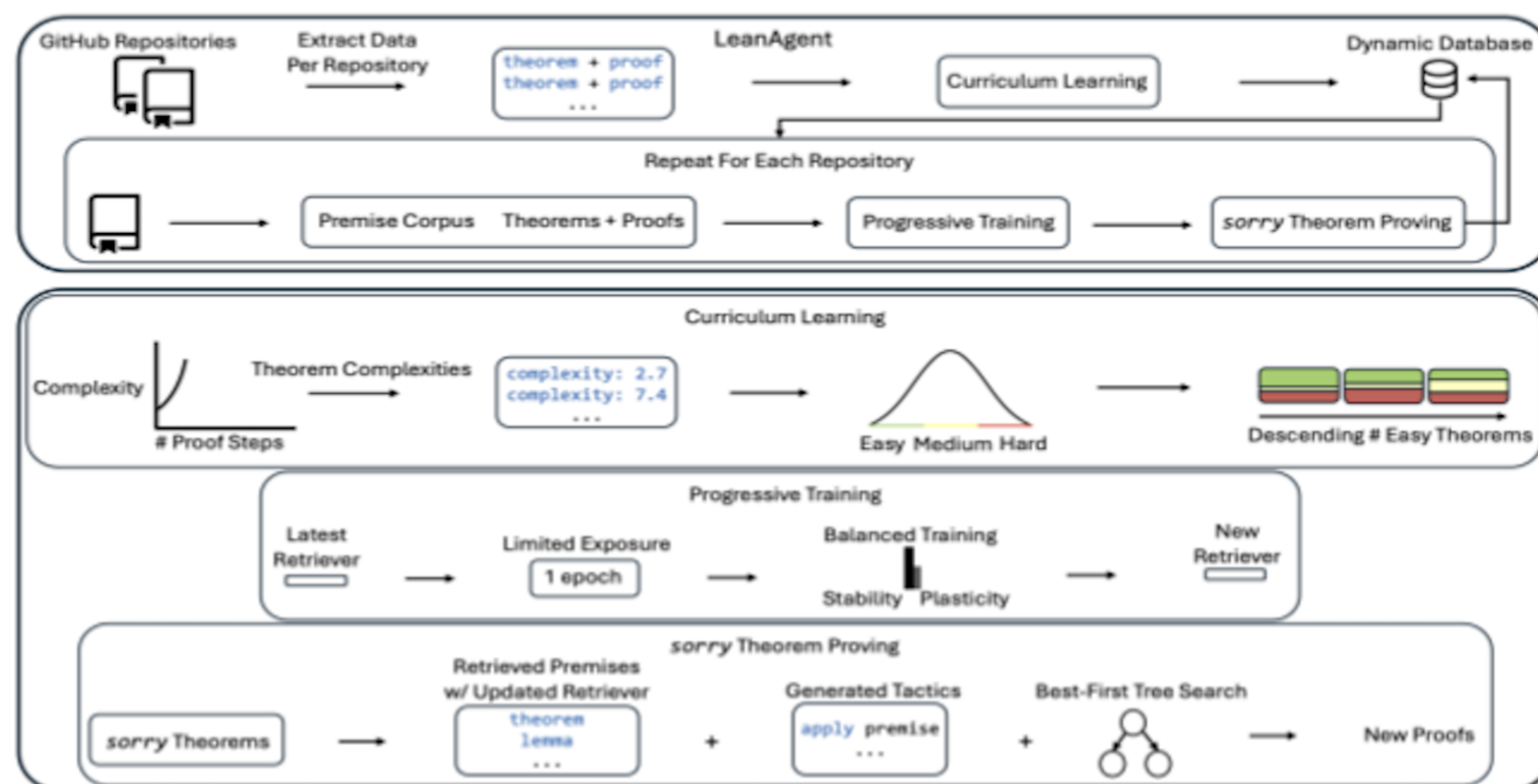
- Current LLMs only provide informal, textual proofs of mathematical theorems
- Lean is a popular framework for formalizing theorems and their proofs

```
theorem re_float (a : Float)
: RCLike.re a = a := by
  exact RCLike.re_eq_self_of_le le_refl
```

- Theorem proofs consist of *tactics* and *premises*
- The key problem is to identify which tactics and premises to use
- Existing models struggle to work across domains and balance stability (performance in one domain) with plasticity (performance across domains and novel learning)
- Fine-tuning often sacrifices cross-domain performance and is markedly different from how human mathematicians learn continually

Methods

- **Goal:** to allow LLMs to use lifelong learning to prove formal theorems in Lean
- Gather data: 23 Lean repositories across Github, including domains such as abstract algebra and algebraic topology
- **Step 1 - Curriculum Learning:** for each repository:
 - Extract all theorems and proofs
 - Measure complexity of each proof as e^S
 - Assign each proof to one of 3 buckets: easy (<33rd percentile globally), medium (between 33rd - 66th percentile globally), and hard (>66th percentile globally)
- **Step 2 - Dynamic Database:** Sort repositories by number of easy theorems, so LeanAgent progresses from more basic to more advanced concepts, and keep track of theorems, proofs, traces, and complexities
- **Step 3 - Progressive Training:** in order of easiness, perform a single training epoch over each repository using contrastive learning objective
- **Step 4 - sorry Theorem Proving:** Try proving *sorry* theorems (theorems with no proof) and add proven theorems to the dynamic database and repeat



Accuracy Results

- LeanAgent outperforms existing methods in theorem proving

Repository	#sorrys	LA Accuracy (%)			ReProver Accuracy (%)	ReProver+ Accuracy (%)
		Total	During	Add. After		
MIL	29	72.4	48.3	24.1	48.3	55.2
MiniF2F	406	24.4	24.4	-	20.9	20.9
Formal Book	29	10.3	6.9	3.4	6.9	10.3
SciLean	294	9.2	7.5	1.7	8.2	8.5
Hairy Ball	14	7.1	0.0	7.1	0.0	7.1
Coxeter	15	6.7	6.7	0.0	0.0	6.7
Carleson	24	4.2	4.2	0.0	4.2	4.2
Lean4 PDL	30	3.3	3.3	0.0	3.3	3.3
PFR	37	2.7	2.7	0.0	0.0	0.0

Lifelong Learning Results

- Compare LeanAgent to several different setups:
 - Setup1: No EWC (elastic weight consolidation), popularity order
 - Setup 2: EWC, Popularity Order
 - Setup 3: EWC, Curriculum Learning
- LeanAgent performs favorably for most lifelong learning metrics:

Metric	Single Repository			
	LeanAgent	Setup 1	Setup 2	Setup 3
WF5 (↓)	0.18	7.60	7.17	0.73
FM (↓)	0.85	6.53	4.04	2.11
CFR (↑)	0.88	0.87	0.88	0.85
EBWT (↑)	1.21	0.51	1.04	0.76
WP5 (↑)	2.47	0.89	1.47	3.42
IP (↑)	1.02	0.36	0.26	1.06

- **Conclusion:** LeanAgent is able to prove more theorems in lean than other baselines and demonstrates better lifelong learning characteristics.
- **Conclusion:** LeanAgent successfully generates formal proofs for 155 new theorems across 23 Lean repositories