

HOList: An Environment for Machine Learning of Higher-Order Theorem Proving

Kshitij Bansal, Sarah M. Loos, Markus N. Rabe,
Christian Szegedy, Stewart Wilcox

Google Research

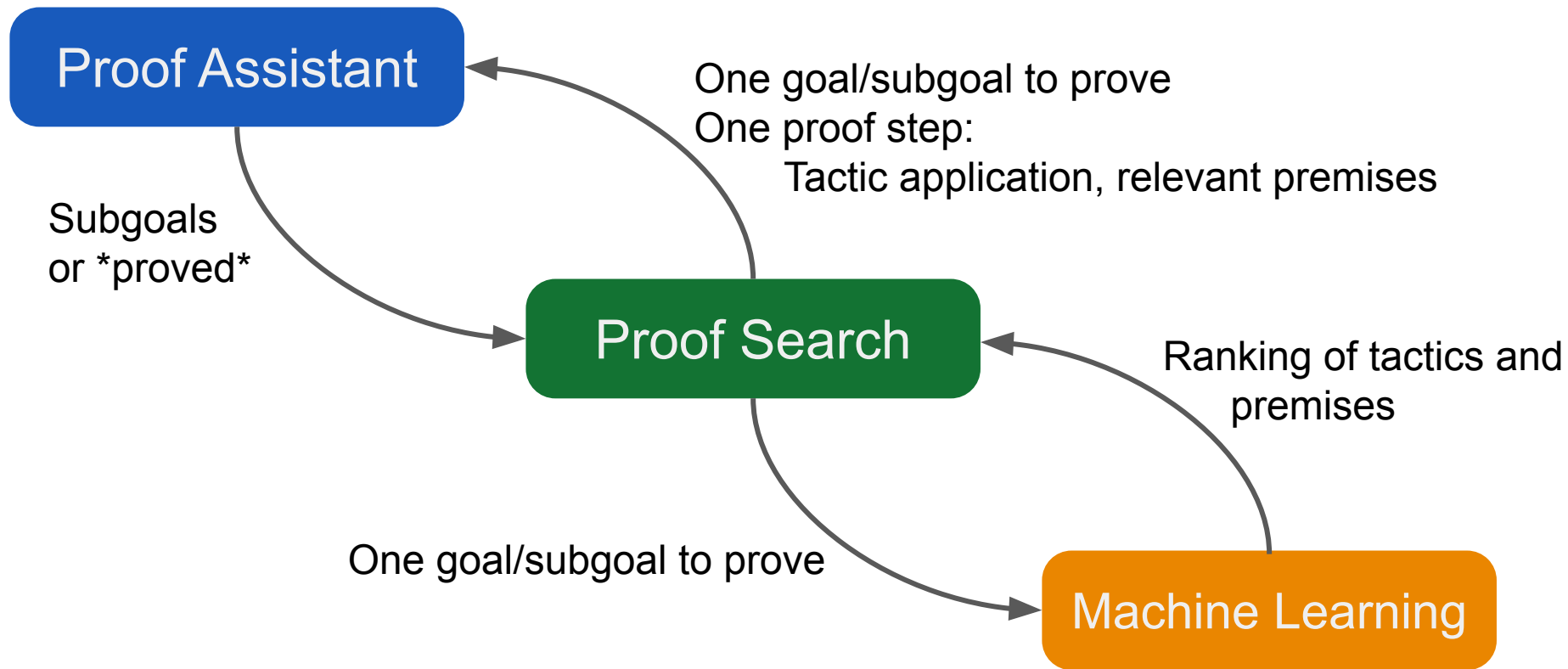
Can we create a human level AI to reason about mathematics?

HOList

An Environment for Machine Learning of Higher-Order Theorem Proving

- HOList provides a simple API for ML researchers and theorem prover developers to experiment with using machine learning for mathematics.
- We use deep networks trained on an existing corpus of human proofs to guide the prover.
- We can improve our results by adding synthetic proofs (generated from supervised models and verified correct by the prover) to the training corpus.

APIs for Theorem Prover Developers and ML Researchers



Results - Supervised Learning on Human Proofs

	Percent of Validation Theorems Closed
Baseline: ASM_MESON_TAC	6.10%
ASM_MESON_TAC + WaveNet premise selection	9.20%
Wavenet	31.72%
Deeper WaveNet	32.65%
Wider WaveNet	27.60%

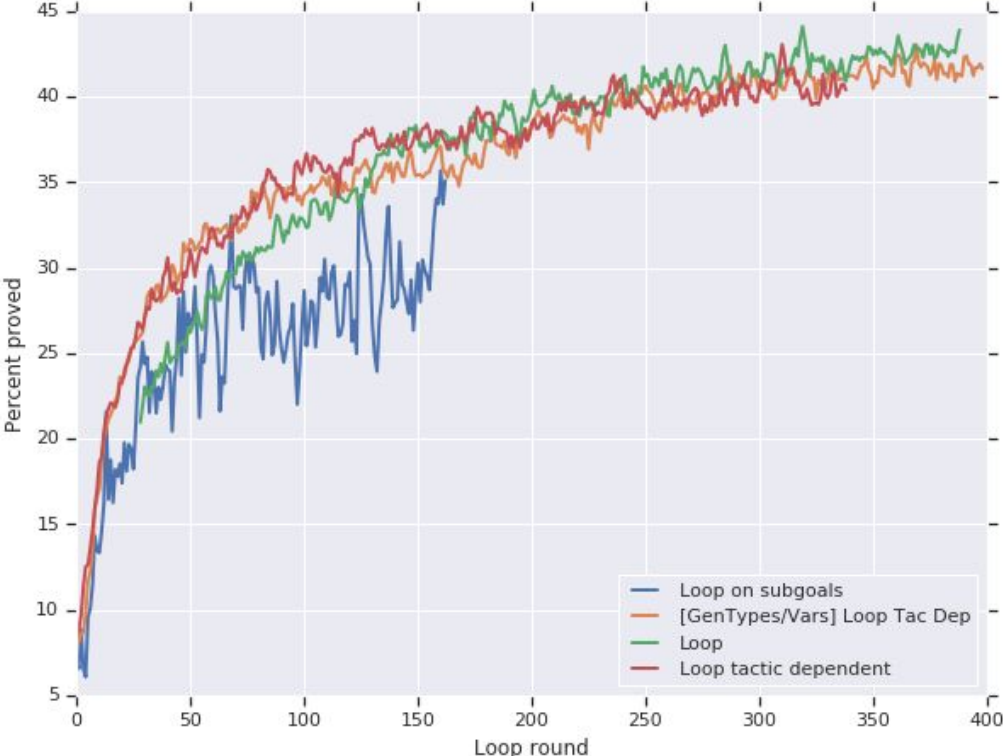
Results - Supervised Learning on Human Proofs

	Percent of Validation Theorems Closed
Baseline: ASM_MESON_TAC	6.10%
ASM_MESON_TAC + WaveNet premise selection	9.20%
Wavenet	31.72%
Deeper WaveNet	32.65%
Wider WaveNet	27.60%

Results - Supervised Learning on Human Proofs

	Percent of Validation Theorems Closed
Baseline: ASM_MESON_TAC	6.10%
ASM_MESON_TAC + WaveNet premise selection	9.20%
Wavenet	31.72%
Deeper WaveNet	32.65%
Wider WaveNet	27.60%

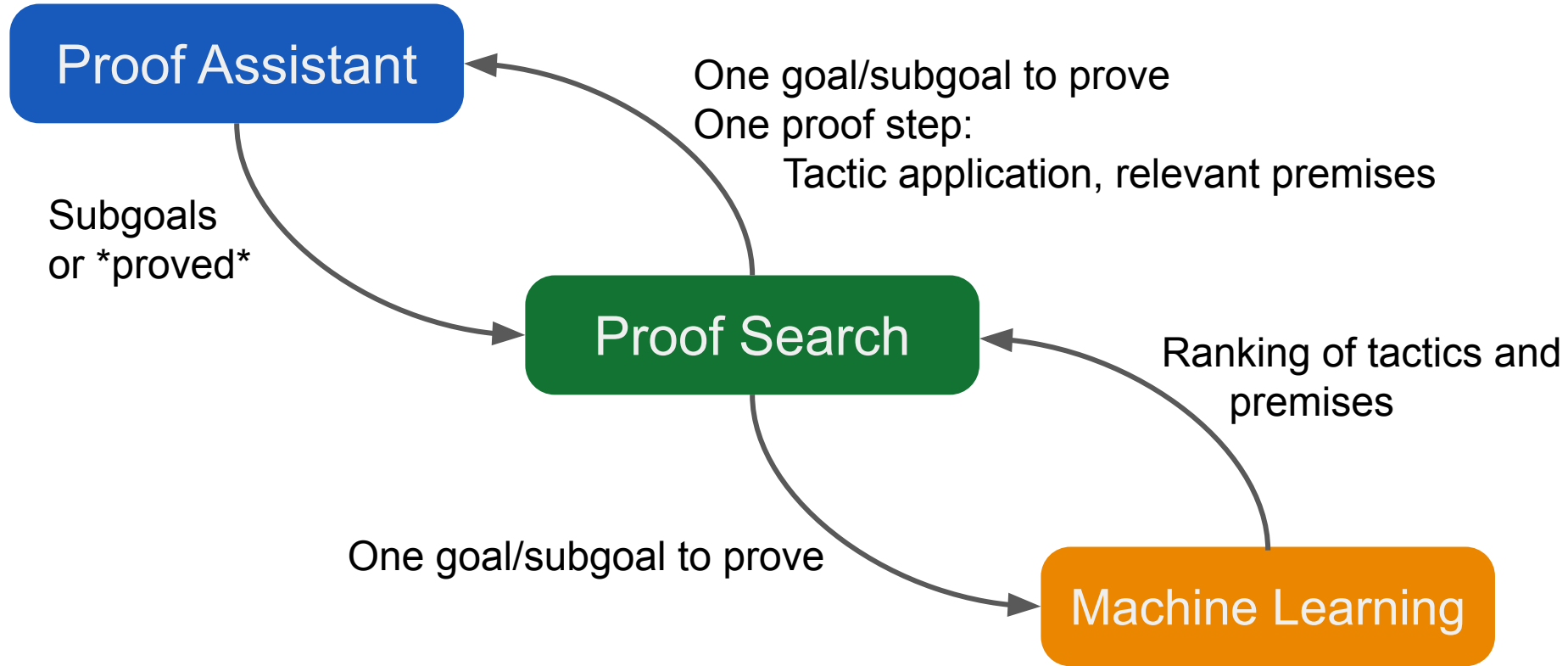
Results - Prover in the loop



	Percent Closed
Wavenet Loop	36.30%
- Trained on loop output	36.80%
Tactic Dependent Loop	38.90%

*

APIs for Theorem Prover Developers and ML Researchers



APIs for Theorem Prover Developers and ML Researchers

Prover

HOL-Light

Input:

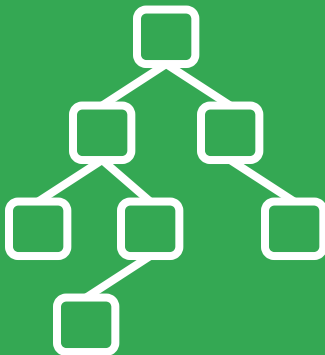
- Load premises
- Apply a tactic to a goal

Output:

- Open goals left to prove

Proof Search

- Manages the state of the proof search tree.
- Allows arbitrary nodes to be explored.



Supervised Learning

Training Data:

TF Examples from Human & Synthetic Proofs

Features:

- Goal (or subgoal)

Labels:

- Tactic applied
- Premises used

deephol.org

- Code is available on GitHub
- Training data
 - 30K theorems and definitions
In the areas of: topology, multivariate calculus, real and complex analysis, geometric algebra, measure theory, etc., as well as the formal proof of the Kepler Conjecture.
 - 375K human proof steps
 - 830K synthesized proof steps
- Trained model checkpoints
- Docker images for the proof assistant and proof search