HOL4PRS: Proof Recommendation System for HOL4 Theorem Prover

Nour Dekhil, Adnan Rashid, and Sofiène Tahar

Department of Electrical and Computer Engineering

Concordia University, Montreal, QC, Canada

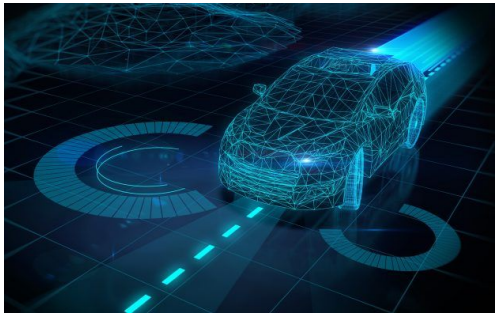
Aug 7, 2024

Why formal verification is so important?

- Ensuring the correct functionality of complex software systems or critical hardware components is crucial.

Aerospace

Healthcare



Automotive

Formal Verification

- Formal verification verifies a system's behavior under all conditions.
- Interactive Theorem Provers (ITPs) help apply mathematical logic to verify these behaviours.

PVS

Challenges of ITPs

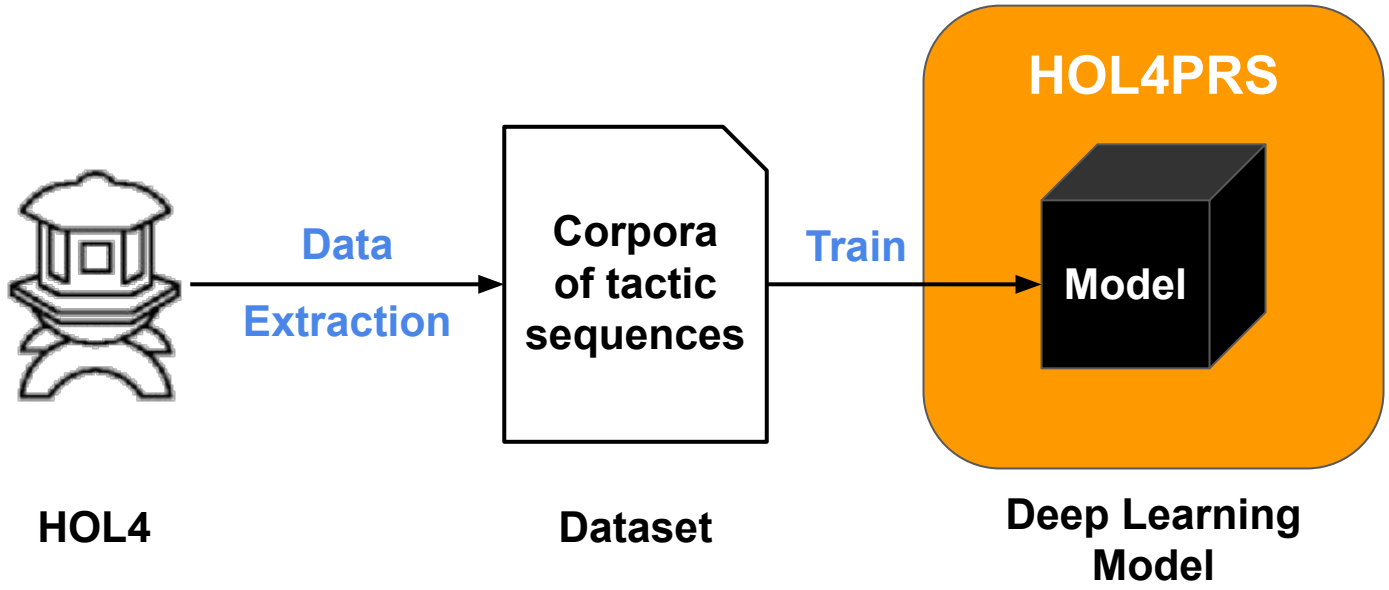
Challenging and labor-intensive !

Smart copilot that guides us through theorem proving.

Related Works

Related Work	Approach	Prover	Success Rate
Gauthier et al., 2020	Machine Learning (k-NN)	HOL4	66.4%
Blaauwbroek et al., 2020	Machine Learning	Coq	23.4%
Luan et al., 2021	Deep Learning (LSTM)	Coq	87%
Yeh et al., 2023	Deep Learning (T5)	PVS	70%

HOL4PRS: Proof Recommendation System for HOL4 Theorem Prover

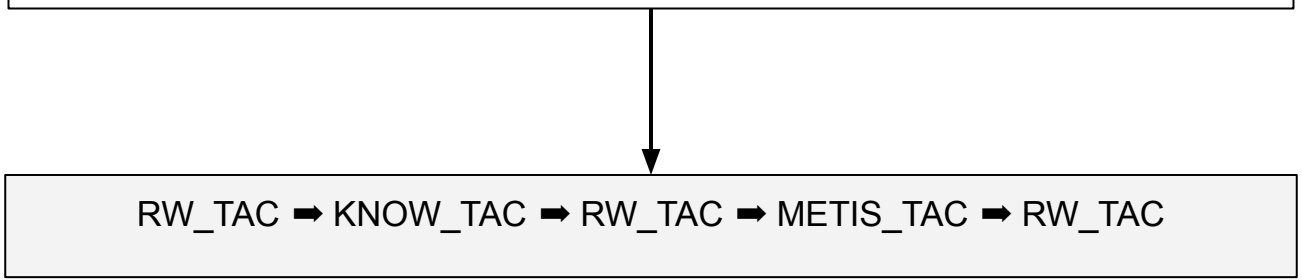


Data Extraction

.*sml* File

```
val rule_18 = store_thm("rule_18",
  ``!X Y. D_AND X Y = D_OR (P_AND X Y)(P_AND Y X)``,

RW_TAC std_ss[D_AND_def, P_AND_def, D_OR_def]
THEN KNOW_TAC(``  !s.( max (X s) (Y s)) =
    (
         min (if X s <= Y s then Y s else PosInf)
           (if Y s <= X s then X s else PosInf))``)
    THEN1 (RW_TAC std_ss[extreal_max_def, extreal_min_def] THEN
METIS_TAC[extreal_lt_def, extreal_le_def, lt_le, le_lt, le_trans,
lt_trans, lt_infty, le_infty])
THEN RW_TAC std_ss[] ) ;
```



Tactic Sequence

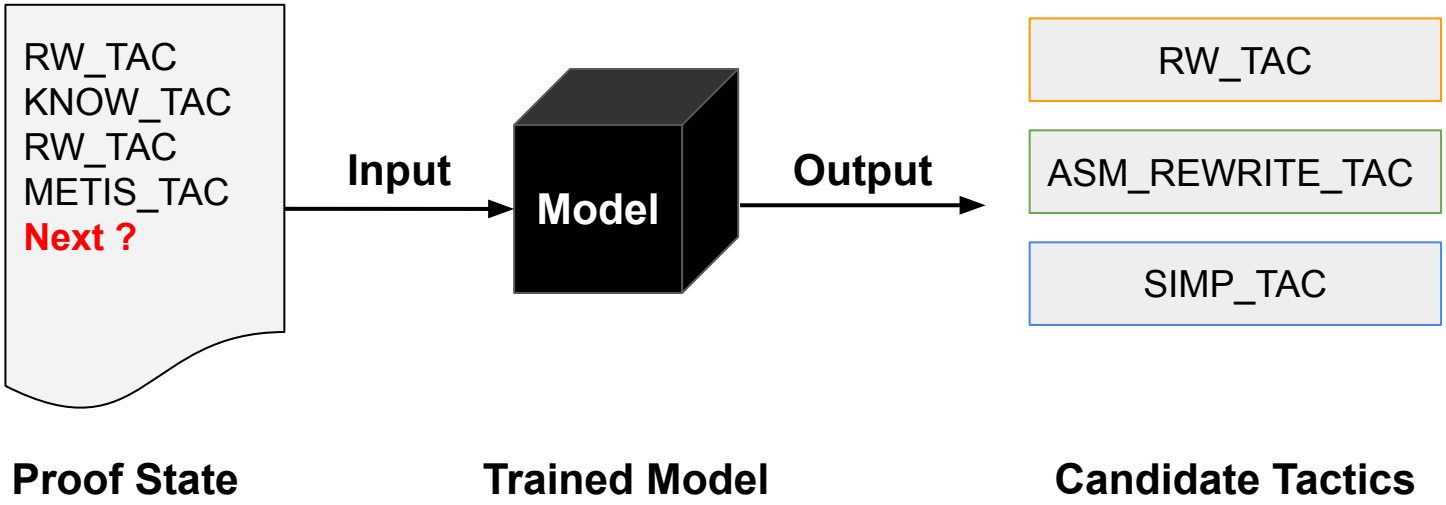
Datasets

Proof: A sequence of n tactics used to prove a theorem.

Proof States: All possible tactic sequences recorded from a proof.

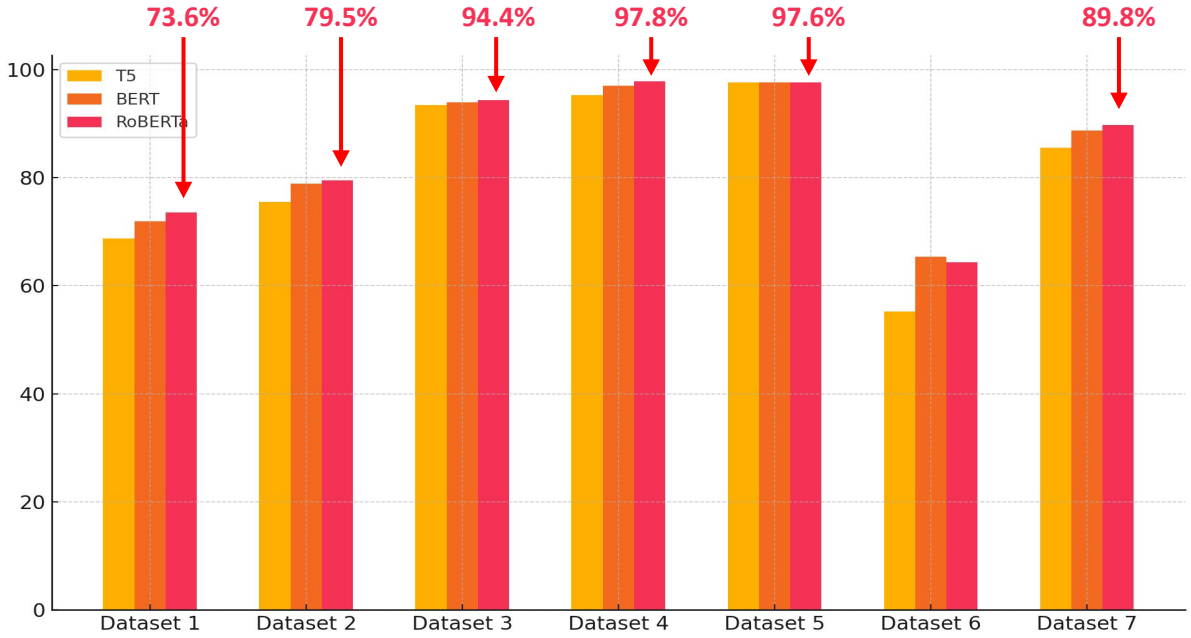
	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6	Dataset 7
Proofs	1,873	2,475	153	295	61	279	5,136
Proof States	43,167	57,602	2,973	7,371	1,784	3,259	116,156

Train Model


RW_TAC
KNOW_TAC
RW_TAC
METIS_TAC
Next ?
Input
Model
Output
RW_TAC
ASM_REWRITE_TAC
SIMP_TAC
Proof State
Trained Model
Candidate Tactics

Experimental Results

7-Correctness Rate: Probability that the correct tactic is among the top 7 recommendations.



HOL4PRS deploys RoBERTA for Top-7 recommendations

Future Work

- Integrate more contextual information with the input to improve the accuracy and relevance of proof step recommendations.

- Autonomously generate complete proofs without human intervention.

Demo