TacticToe
Learning to Reason with HOL4 Tactics

Thibault Gauthier, Cezary Kaliszyk, Josef Urban

May 8, 2017
Reasoning with inference rules
Reasoning with inference rules
Reasoning with inference rules
Reasoning with tactics
Reasoning with tactics
Reasoning with tactics
Common tactics

- REWRITE_TAC
- INDUCT_TAC
- METIS_TAC
Composing tactics

THENL tactical composes the effect of tactics.
Composing tactics

THENL tactical composes the effect of tactics.
Composing tactics

THENL tactical composes the effect of tactics.
**Composing tactics**

THENL tactical composes the effect of tactics.
Composing tactics

THENL tactical composes the effect of tactics.
Tactic selection

Was the tactic successful before on similar goals?
Before: Recording tactics

- **Globalizing:**
  - Local values
    
    ```plaintext
    let val x = 5 in NTAC x INDUCT_TAC end
    ```
  - Modules
    
    `Ho_rewrite, Rewrite`

- **Wrapping:**

  `R INDUCT_TAC THENL [R REWRITE_TAC, R METIS_TAC]`

- **Database:**

  ```plaintext
  INDUCT_TAC     x + x >= x
  INDUCT_TAC     x * x >= x
  REWRITE_TAC    1 + 1 = 2
  ```
Similarity

Features: constants, subterms, names of variables, ...

INDUCT_TAC  [+ ,>= ,... ]
INDUCT_TAC  [* ,>= ,... ]
REWRITE_TAC  [1 ,2 ,+, = ,... ]

Best predicted tactic for 2 >= 1?
How to search for a proof?

Depth first search:
- Start with the conjecture
- Apply best predicted tactic
- Repeat on the new goals

A*-search:
- Cost: length of the proofs needed to create the goal
- Heuristic: evaluation of the length of the remaining proof
HOL(y)Hammer

Current Goal

HOL4  HOL(y)Hammer

ITP Proof  ATP Proof

Z3, Vampire, E-prover

TPTP

Proof Assistant  Hammer  ATPs
General results

<table>
<thead>
<tr>
<th>ID</th>
<th>7902 theorems</th>
</tr>
</thead>
<tbody>
<tr>
<td>TacticToe</td>
<td>29.73</td>
</tr>
<tr>
<td>TacticToe*</td>
<td>39.42</td>
</tr>
<tr>
<td>HolyHammer</td>
<td>32.35</td>
</tr>
</tbody>
</table>


## Results by theories

<table>
<thead>
<tr>
<th>Theory</th>
<th>arith</th>
<th>real</th>
<th>compl</th>
<th>meas</th>
</tr>
</thead>
<tbody>
<tr>
<td>TacticToe</td>
<td>37.3</td>
<td>19.7</td>
<td>42.6</td>
<td>19.6</td>
</tr>
<tr>
<td>TacticToe*</td>
<td>60.1</td>
<td>46.1</td>
<td>63.7</td>
<td>22.1</td>
</tr>
<tr>
<td>HolyHammer</td>
<td>51.9</td>
<td>66.8</td>
<td>72.3</td>
<td>13.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Theory</th>
<th>proba</th>
<th>list</th>
<th>sort</th>
<th>f_map</th>
</tr>
</thead>
<tbody>
<tr>
<td>TacticToe</td>
<td>25.3</td>
<td>48.1</td>
<td>32.7</td>
<td>53.4</td>
</tr>
<tr>
<td>TacticToe*</td>
<td>25.3</td>
<td>51.9</td>
<td>34.7</td>
<td>55.5</td>
</tr>
<tr>
<td>HolyHammer</td>
<td>25.3</td>
<td>23.3</td>
<td>16.4</td>
<td>18.1</td>
</tr>
</tbody>
</table>
Example in gcdTheory: GCD_ADD_L

\[ \forall a\ b. \ gcd(a + b)\ a = gcd\ a\ b \]

Human proof: PROVE_TAC [GCD_SYM,GCD_ADD_R]

TacticToe proof:
ARW_TAC
THEN MATCH_MP_TAC (SPECL [a, a + b] IS_GCD_UNIQUE)
THEN ARW [...] IS_GCD_MINUS_R
THEN PROVE_TAC [GCD_IS_GCD, IS_GCD_UNIQUE, IS_GCD_SYM]

HolyHammer proof: METIS_TAC [GCD_SYM,GCD_ADD_R]
Example in listTheory: DROP NIL

\[\forall ls\ n. (DROP\ n\ ls = [\ ]) \iff n \geq LENGTH\ ls\]

Human proof: Induct THEN SRW_TAC [ ] [ ] THEN DECIDE_TAC

TacticToe proof:
INDUCT_THEN list_INDUCT_ASSUME_TAC
THENL [SRW_TAC [] [], SRW_TAC [ARITH_ss] []]
Conclusion

TacticToe combines previous human proofs to solve new goals.

- Induction principle
- Simplification sets
- User-defined domain specific automation

The proofs produced are efficient HOL4 proof scripts.
Future works

- More **features** for goals:
  - Tactic arguments relation to the goal
  - Time to solve, number of tactics necessary
- Extending the **policy**: tactic argument selection
- Better **evaluation** of the difficulty of the goal