Exploiting symmetries and axiom reformulation in automated generation of formal proofs

Sana Stojanović
Faculty of Mathematics, University of Belgrade, Serbia
URL: www.matf.bg.ac.rs/ ~sana

Workshop on Formal and Automated Theorem Proving and Applications
Belgrade, Serbia, February 04, 2012.

Agenda

- Motivation
- Considered techniques
  - Automation using symmetric predicates
  - Modifying the axiomatic system
- Conclusions and future work

Motivation

- Formalization of geometry using interactive proof assistants (Isabelle, Coq), Meikle and Fleuriot, Schreck, Narboux

- Semi-automated approach, Scott, Meikle and Fleuriot

- ArgoCLP, a coherent logic based prover

  – Automatically produce formal proofs that resemble proofs that could be found in textbooks

ArgoCLP

- Coherent logic is a fragment of first-order logic, consisting of formulae of the following form:
  $A_1(\vec{x}) \wedge \ldots \wedge A_n(\vec{x}) \Rightarrow \exists \vec{y_1} B_1(\vec{x}, \vec{y_1}) \vee \ldots \vee \exists \vec{y_m} B_m(\vec{x}, \vec{y_m})$

- Simple proof procedure with forward chaining and iterative deepening

- Problems with increasing number of constants and derived facts during proof process

Improvement of ArgoCLP prover

- Reduction in number of constants and facts (dealing with symmetric predicates, reformulation of axioms)

- For theorems with proofs that rely on symmetry and axioms that introduce several witnesses, the proofs become 60% shorter (otherwise, there is no effect or there is a small negative effect)

- Automatical discovery of symmetric relations and theorems obtained by reformulation of axioms

Use of symmetric predicates

- $R(a_1, \ldots, a_i, \ldots, a_j, \ldots, a_n) \Leftrightarrow R(a_1, \ldots, a_j, \ldots, a_i, \ldots, a_n)$
- Set of constants used in the program can be enumerated and all permutations could be replaced with sorted one
- Example: $col(B, C, A)$ holds and an axiom has $col(A, C, B)$ in its premises
  - $col(B, C, A) \Rightarrow col(A, B, C)$
  - $col(A, B, C) \Rightarrow col(A, C, B)$

Symmetric predicates

- Definitions:
  - $collinear(S) \equiv (\exists p)(\forall A \in S) on\_line(A, p)$ — Isabelle
  - $collinear(A, B, C) \equiv (\exists p)(A \in p \wedge B \in p \wedge C \in p)$ — Hilbert
  - $collinear(A, B, C) \equiv (signedArea(A, B, C) = 0)$ — Meikle and Fleuriot

- Deriving symmetries from axioms along the proof search is expensive

- Automation is beneficial (Meikle and Fleuriot 2010)
  - *simp* tactic which simplifies subgoals using rewriting and decision procedures

Automated detection of symmetric predicates

- Generators of the set of permutation:
  - $R(a_1, a_2, \ldots, a_n) \Leftrightarrow R(a_2, a_1, \ldots, a_n)$
  - $R(a_1, a_2, \ldots, a_n) \Leftrightarrow R(a_2, a_3, \ldots, a_n, a_1)$
- Statements of this form could be automatically generated and then ArgoCLP prover can try to prove them
- Generated theorems will not be added to the set of axioms
- Used only for proof completion when symmetry is exploited

Automatic reformulation of axioms — example

- Application of an axiom that have more than one existential quantifier in conclusions can be replaced with an application of one or more theorems:
  - *Axiom:* There exist at least two points on a line
  - *Theorem:* For line p and a point A that lies on p, there exists a point B different from A that also lies on p

- In general, axiom
  $A_1(\vec{x}) \wedge \ldots \wedge A_n(\vec{x}) \Rightarrow \exists y_1 \exists y_2 \ldots \exists y_k B(\vec{x}, y_1, y_2, \ldots, y_k)$
  can generate following statement
  $A_1(\vec{x}) \wedge \ldots \wedge A_n(\vec{x}) \wedge B_1(\vec{x}, y_1) \Rightarrow \exists y_2 \ldots \exists y_k B_2(\vec{x}, y_1, y_2, \ldots y_k)$

- Statements of this form could be automatically generated and then ArgoCLP prover can try to prove them

Automatic reformulation of axioms

- *Axiom:* There exists 3 noncolinear points
  - Given point A there exist two points B and C so that $\neg col(A, B, C)$ holds
  - Given points A and B there exists point C such that $\neg col(A, B, C)$ holds (missing $A \neq B$)
- Theorems are provided to the prover and used instead of original axiom (when possible)
- User assistance for obtained statements that can not be proved

Conclusions

- Isabelle tactics are not changed, only set of theorems are given to the prover
- Proofs are completed to the level of axiom application
- Generic approach, applicable to any coherent theory

Future work

- Automatic attempts to discover premises that are missing in statements obtained by axiom reformulation

- Replace application of an axiom that introduces more witnesses with suitably chosen theorem

- Using the new version of the prover for proving theorems from different axiomatic systems (Hilbert, Tarski, Avigad)