



AGH

**AKADEMIA GÓRNICZO-HUTNICZA
IM. STANISŁAWA STASZICA
W KRAKOWIE**

Studio projektowe 1

Benchmark solverów prover9 oraz SPASS

Autorzy:

Mateusz Grzeliński

Przemysław Michałek

1 Cel

Celem projektu jest zbadanie wydajności automatycznych metod dowodzenia twierdzeń Prover9 oraz SPASS.

2 Streszczenie

Badany prover zostaje uruchomiony kilka razy z wcześniej przygotowanym problemem SAT(Boolean satisfiability problem). Badany jest czas wykonywania oraz rezultat (czy SAT jest spełnialny)

2.1 Prover9

Prover9 korzysta z biblioteki [LADR](#) (załączona w źródłach, libladr.a, ladrladr.h). Kompiluje się ona do pliku wykonywalnego. Prover9 przyjmuje składnię LADR. Konwerter TPTP to LADR jest dostępny w bibliotece LADR

2.2 SPASS

[SPASS](#) nie korzysta z bibliotek, kompiluje się do pliku wykonywalnego. Jest duża szansa że można skompilować go do biblioteki. Może korzystać z własnej składni lub ze składni TPTP.

2.3 TPTP

[Link do składni BNF](#)

[Przykład pliku](#)

TPTP udostępnia wiele problemów do testowania: [link](#)

Te problemy są sklasyfikowane przez domeny (3 literowe skróty), przykładowo

LCL - Logic Calculi

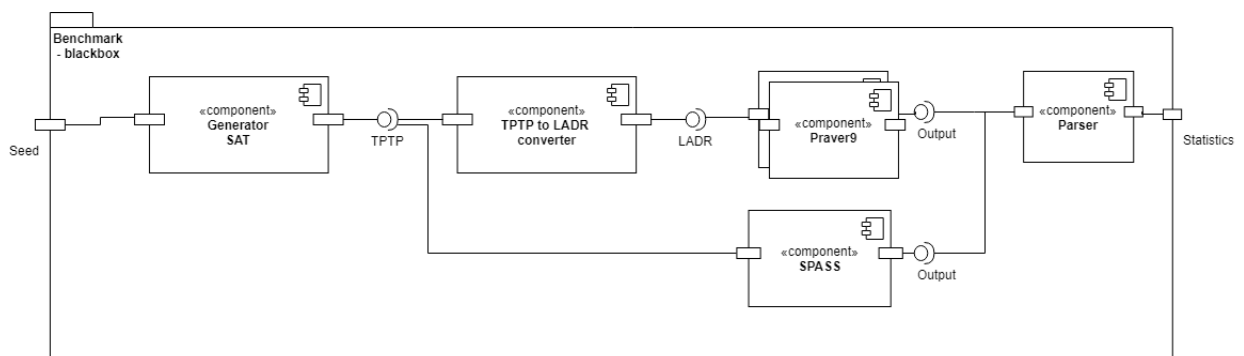
COL - Combinatory Logic

Pełny opis skrótów dostępny jest [tutaj](#)

3 Benchmark - strategia blackbox

Provery traktujemy jako czarne skrzynki, statystyki są ograniczone.

- TPTP - składnia TPTP
- LADR - format wejścia Provera9
- Output - wyjście proverów, mogą być różne
- statistics - statystyki



3.1 Algorytm

- Wygeneruj formułę SAT
- Wykonaj kod w podprocesie
- Zmierz czas wykonania procesu
- Zbadaj output: dodatkowe statystyki, wynik

4 Generator SAT

4.1 Wejście

- seed - opcjonalnie - umożliwia powtarzanie losowania
- typ generatora liczb losowych
- ilość formuł
- ilość zmiennych

4.2 Wyjście

- string w formacie TPTP (dokumenty do składni: <http://www.tptp.org/>)

4.3 Algorytm

- Wypisz metadane jako komentarz (link do źródła, parametry wejściowe)

5 TO DO

- czy statystyki czas wykonania, wynik wystarczają? -
 - zależność pomiędzy długością wejścia a ilością zmiennych
 - zajętość pamięci
- jakie parametry ma przyjmować generator SAT? Czy budujemy generator CNF, czy generator ogólny?
- czy brać pod uwagę parametry takie jak k-SAT, Horn-SAT, NAE3SAT
- jaki ma być format wyjściowy generatora?
 - propozycja: TPTP - SPASS ma natywne wsparcie dla tego formatu, a prover9 ma translator TPTP to LADR
- wypisywanie do terminala będzie zabierało czas obliczeń. Trzeba będzie zminimalizować wypisywanie
- opisać składnię obu proverów

6 Słownik

SAT - problem spełnialności – zagadnienie rachunku zdań, określające czy dla danej formuły logicznej istnieje takie podstawienie (wartościowanie) zmiennych zdaniowych, żeby formuła była prawdziwa

Prover9 – jest to zautomatyzowane narzędzie udowadniające dla logiki pierwszego rzędu stworzone przez Williama McCune’a

SPASS - SPASS Theorem Prover jest narzędziem do automatycznego dowodzenia twierdzeń, należących do rachunku predykatów pierwszego rzędu.

Logika LTL – jedna z logik temporalnych. Jest oparta na liniowej strukturze czasu

Logika temporalna – logika umożliwiająca rozważanie zależności czasowych bez wprowadzania czasu explicite

TPTP (Thousands of Problems for Theorem Provers) - is a library of test problems for automated theorem proving (ATP) systems

FOF (First-Order Formula) reduction is a method of attempting to simplify a problem by reducing it to an equivalent set of independent subproblems. A trivial example is to reduce the conjecture $A \leftrightarrow B$ to the pair of problems $A \rightarrow B$ and $B \rightarrow A$.

7 Wnioski

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.