

Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy v Praze

Autor práce	Bc. Robert Brunetto		
Název práce	External sources of axioms in lean theorem proving (Externí zdroje axiomů pro štíhlé dokazování)		
Rok odevzdání	2012		
Studijní program	Informatika	Studijní obor	Teoretická informatika
Autor posudku	RNDr. Martin Suda	Role	vedoucí
Pracoviště	Katedra teoretické informatiky a matematické logiky		

Práce se zabývá rozšířením automatického dokazovače vět LeanCoP o možnost přistupovat k externím zdrojům axiomů. Navazuje na předchozí výzkum¹, kde byla tato problematika studována z pohledu saturačního dokazovače SPASS. LeanCoP vychází z úplně jiné filosofie dokazování, konkrétně je založen na connection kalkulu, a proto bylo potřeba vyvinout novou techniku pro zajištění podobného propojení. Ta byla následně naimplementována a oba přístupy byly porovnány jak na původních příkladech, tak i na příkladech nově vytvořených. Z experimentů vyplynulo, že nový systém je při komunikaci s externími zdroji efektivnější. Navíc byly identifikovány možnosti pro případná další vylepšení.

Úvodní kapitola nejprve zešíroka představuje automatické dokazování, pak se již věnuje problematice externích zdrojů axiomů. Základní motivací jsou problémy v rozsáhlých teoriích, u kterých by načtení celého vstupu před započítím výpočtu buď vůbec nebylo technicky zvládnutelné, nebo by okamžitě způsobilo „zahlcení“. Místo toho je možné část vstupní teorie oddělit do externího zdroje, dokazovači předat na vstup pouze informaci o jeho existenci a tvaru axiomů v něm uložených a nechat na něm, aby si o konkrétní axiomy žádal až za běhu a to pouze o takové, které by mohly pomoci při dokázání právě zkoumané domněnky.

Přípravná kapitola 1 připomíná základní pojmy z oblasti automatického dokazování, dále pak nastiňuje fungování dokazovače SPASS a jeho rozšíření podporující externí zdroje axiomů SPASS-XDB (1.1). Následuje neformální popis fungování connection kalkulu (1.2) a jeho prologovské implementace v dokazovači LeanCoP (1.3). Na závěr je vysvětleno, jakým způsobem se externí zdroje axiomů deklarují a jak probíhá komunikace s nimi (1.4). Sekce (1.2) se zabývá poměrně složitou problematikou a diplomant se v ní dopouští několika nepřesností², přesto z ní čtenář získá dost detailů pro porozumění následujícího textu a ve zbytku je odká-

¹ M. Suda, G. Sutcliffe, P. Wischnewski, M. Lamotte-Schubert, G. de Melo: External Sources of Axioms in Automated Theorem Proving, in Proceedings of the 32nd Annual Conference on Artificial Intelligence (eds. B. Mertsching), LNAI 5803, pp. 281-288. Springer, Heidelberg (2009).

² „Z toho plyne, že by musel být v každé klauzuli z M alespoň jeden literál, který by byl v S splněný.“ – zde nelze přímo mluvit o konkrétním splněném literálu, neboť klauzule obecně obsahuje proměnné, na kterých může splněný literál záviset. „Dojít ke sporu s existencí struktury S takové, že v M by byla nějaká klauzule nesplněná,...“ – správně by bylo „že všechny klauzule z M by byly splněny v S .“

zán na původní literaturu. V této kapitole by nejspíš měl být též vysvětlen pojem „ground atom/axiom“, který není standardní a vyskytuje se na několika místech práce.

Kapitola 2 popisuje systém LeanCoP-XDB, rozšíření dokazovače LeanCoP o přístup k externím zdrojům axiomů. Klíčovým se při jeho návrhu stal pojem *schematického důkazu* (2.1). Jedná se o důkaz spekulativně vygenerovaný LeanCoPem za předpokladu, že by externí zdroje axiomů obsahovaly určité vhodné axiomy konkrétního tvaru. Schematický důkaz se stane skutečným důkazem v okamžiku, kdy se po komunikaci s externími zdroji tento předpoklad potvrdí – schematický důkaz je *doplněn* na důkaz. (Bohužel ihned po zavedení tohoto pojmu používá diplomant pro stejnou věc i pojem *schematická odpověď*, aniž by na to čtenáře upozornil a to i dále ve zbytku textu.) V kapitole jsou poměrně obsáhle rozebrány způsoby, jak by doplňování schematických důkazů mohlo v praxi probíhat (2.2) – ukazuje se, že úloha souvisí s databázovou operací joinu tabulek – pro implementaci byl pak ale zvolen nejjednodušší korektní přístup. Další sekce se pak soustředí na návrh zbytku systému. Základní architektura byla navržena tak, aby program uměl zpracovávat více schematických důkazů současně a při tom komunikoval s více než jedním externím zdrojem. Implementace jednotlivých komponent pak ale volí jednodušší, sekvenční přístup. Jako pro čtenáře nešťastnou hodnotím strukturu této části práce, kdy jsou jednotlivé „stupně volnosti návrhu“ nejprve formou otázek prezentovány v sekci (2.3) a teprve v sekci (2.5) je na tyto otázky v uvedeném pořadí odpovídáno.

Kapitola 3 seznamuje čtenáře s jednotlivými externími zdroji axiomů použitými v experimentech a to jak se zdroji převzatými, tak i s těmi, které vznikly v rámci práce. Nejrozsáhlejší je pak kapitola 4, která popisuje vlastní experimenty. Jejich hlavním cílem je porovnání programu SPASS-XDB a nově vzniklého LeanCoP-XDB. Údaje o jednotlivých bězích programů jsou prezentovány ve formě tabulky, která je sice přehledná, ale obsahuje i nedokumentované údaje³, které by mohly čtenáře mást. Dále jsou jednotlivé experimenty doplňovány komentáři a slovním rozbořem. Z experimentů vyplývá jasná převaha systému LeanCoP-XDB nad SPASS-XDB a to nejen v naměřených časech, ale i vzhledem ke tvaru otázek, které posílá. Ty jsou totiž v případě systému LeanCoP-XDB typicky specifitější a tedy lépe „zacílené“ na potenciální odpovědi v externím zdroji. Vedlejších produktů uvedených experimentů jsou i zajímavé poznatky o chování LeanCoP-XDB v praxi, které by mohly v několika případech pomoci při dalším vylepšování programu. Tato užitečná pozorování jsou pak shrnuta v kapitole 5.

Součástí práce je na CD přiložený program LeanCoP-XDB a použitá testovací data. V textových přílohách diplomant podrobně popisuje, jak tento program instalovat a spouštět na příkladech z kapitoly 4. Dává též návod, jak vytvářet a zapisovat příklady vlastní. Součástí příloh je i popis protokolu, kterým systémy s externími zdroji komunikují, a je tedy i vysvětleno, jak by bylo možno vytvářet zdroje vlastní.

Práci jako celek hodnotím v zásadě kladně. Zejména úsilí, které bylo věnováno přípravě testovacích příkladů a s nimi související infrastruktury (nové externí zdroje, podpůrné utility,

³ Konkrétně se jedná o hodnoty na souřadnicích `term translator` a `leancoP core vs queries`, `started` a `finished`.

atd.) je jistě nadprůměrné. Bohužel její kvalitu snižují časté pravopisné chyby (převážně psaní čárek mezi větami a shoda podmětu s přísudkem), typografickou stránku kazí výskyty jednopísmenných předložek na konci řádky, či citace mimo věty (zejména v úvodu). Místy je srozumitelnost textu snížena jistou neobratností při vyjadřování. Překlepy se sice vyskytují minimálně, na několika místech je však ve větě omylem zdvojen výskyt slova. Domnívám se ale, že klady převažují nad zápory a diplomovou práci doporučuji k obhájení.

Martin Suda