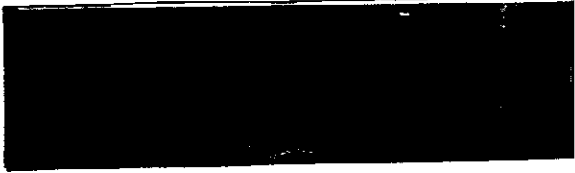


Z práce jsem poměrně rozpačitý. Za projektem stojí veliké množství práce, výsledek tomu však příliš neodpovídá. Na jednu stranu studentka zvládla nemalé množství nových věcí (jazyk Perl a práce s textem v tomto jazyce, dokazovač Prover9 a práce s jeho výstupy, principy automatického dokazování), na druhou stranu vstupy vedoucího práce byly enormní a výsledek jen částečně odpovídá původní představě. Navíc si nejsem jist hloubkou porozumění zadanému tématu.

Závěr: I přes uvedené nedostatky doporučuji uznat předloženou práci jako bakalářskou a ohodnotit stupněm **dobře**. Doporučuji komisi, aby využila doplňujících otázek u obhajoby k zjištění, nakolik studentka skutečně ví, o čem mluví.

V Praze, 5.9.2010
David Stanovský



Renáta Ševčíková:

Prezentace automaticky generovaných důkazů

posudek vedoucího práce

Tématem práce je automatické zpracování důkazů produkovaných dokazovačem Prover9 do podoby, která by byla čitelnější pro člověka. Vzhledem k okolnostem popsaným níže jsme se nakonec omezili na rovnicové důkazy a jejich transformaci dvěma způsoby: 1) primitivní strukturování do lemmat daných jednotlivými paramodulačními kroky, 2) převod důkazu do formy jedné dlouhé rovnosti, ve smyslu definice důkazu v rovnicové logice (univerzální algebře).

(Původním plánem a přirozeným rozšířením práce bylo jednak chytřejší strukturování důkazu do větších celků, tj. chytřejší definice pojmu „lemma“, a dále implementace dalších inferenčních pravidel tak, aby bylo možné zpracovat jakýkoliv důkaz vyprodukovaný dokazovačem Prover9.)

Vlastní práce se skládala ze **tří částí**: implementace výše popsaných úloh v jazyce Perl, jejich testování na problémech z databáze TPTP a nakonec vlastní text práce, který obsahuje popis problému a výsledky testování.

Zdaleka nejvíce úsilí bylo věnováno **implementaci**. Zdá se, že finální verze programu funguje správně na všech čistě rovnicových důkazech. Program je velmi neefektivní, zpracování důkazu trvá často déle než jeho nalezení dokazovačem Prover9. Cca 20% úloh, které Prover9 dokázal během 3 minut, nebyl program schopen zpracovat do hodiny. Průběh vytváření programu vyžadoval značné vstupy z mé strany, jak co se týče algoritmu, tak především testování, když jsem postupně obdržel snad 20 částečně nefunkčních verzí. Nicméně výsledný program pro mě bude v budoucnu užitečný a je jednoznačně přínosem práce.

Vzhledem k tomu, že program byl dokončen na poslední chvíli, vlastní **testování** bylo poměrně omezené. Uvedené údaje pro mě budou užitečné při budoucí práci na nastavení dokazovače i na lepším zpracování důkazů. Bohužel, grafy na straně 36 nemají dobrou vypovídací hodnotu, protože nereflektují vícenásobné hodnoty poměrů, navíc obrázek 4.2 je evidentně špatně.

Úroveň **textu** je nízká, po stylistické i obsahové stránce.

Text obsahuje řadu nepřesností (např. na str. 8 se pojem okamžitého důsledku definuje pro rovnosti a poté používá pro termy, dvě definice důkazu na str. 7 a 8 jsou navzájem nekompatibilní, bod 3. v algoritmu na str. 12 je zavádějící, protože zde jde především o provedení simplifikačních pravidel atd.).

Na několika místech jsou vysloveně chybné údaje (např. na str. 8 není pravda, že by bylo možné interpretovat množinu E jako přepisovací pravidla, obecně je třeba je používat obousměrně, na str. 13 je chybně uvedeno, že Waldmeister vrací konvergentní přepisující systém, což je pravda velmi zřídka, v sekci 3.1 je chybně přeformulovaná úvodní úloha atd.).

Text je chaotický (zkuste pochopit např. popisy grafů na str. 35, značení formulí míchá různé konvence atd.). Míchá se striktně formální text se zcela neformálním, stěžejní definice jsou často neformální („důkaz je posloupnost formulí spojených inferenčními pravidly“ na str. 7, specifikace pojmu „lemma“ nebo „graf důkazu“ v popisu algoritmu atd.), naopak některé formální definice jsou uvedeny zbytečně.

Patrně hlavním problémem je, že z textu lze špatně pochopit, co a jak vlastně program dělá. Kapitola 3 obsahuje prakticky nečitelné části pseudokódu, které nejsou dostatečně doplněny neformálním výkladem principu algoritmu. Specifikace výstupu je neformální.