

Review of Master's Thesis

Student: Vít Radek, Bc.
Title: Grammar-Based Translation Framework (id 20973)
Reviewer: Kolář Dušan, doc. Dr. Ing., UIFS FIT VUT

- 1. Assignment complexity** **more demanding assignment**
Zadání vidím jako obtížnější, neboť srovnání se a vymezení oproti známým nástrojům na generování analyzátorů pro syntaxi na úrovni bezkontextových gramatik považuji za velkou výzvu.
- 2. Completeness of assignment requirements** **assignment fulfilled with serious reservations**
Zadání bylo formálně naplněno, bohužel jakousi "svojskou" cestou. Student se srovnal jen s jediným nástrojem (bison), přitom ovšem v metrikách, které jsou zajímavé jen formálně (řešené konflikty, počet stavů). Student totiž nevytvořil nástroj jako je bison, nebo podobné, ale knihovnu, která navíc nepracuje nad syntaxí řízeným překladem (jako zmíněné nástroje), ale nad překladovými gramatikami. Tedy srovnání tím pádem ani není možné (například tak, že pro existující open source nástroj vygenerujeme kód novým nástrojem a srovnáme chování).
- 3. Length of technical report** **in usual extent**
Práce je v obvyklém rozmezí - na horní hranici, spíše. V práci se dá najít několik míst, která považuji za zbytná, ale nebylo by to více, jak 2-3 tištěné strany, takže to nepovažuji za chybu.
- 4. Presentation level of technical report** **60 p. (D)**
Logická struktura práce je dobrá, kapitoly na sebe navazují. Práce se však dost těžko čte. Některé věty, či pasáže je třeba číst vícekrát. Jak bylo zmíněno, zejména v úvodu je pár zbytných částí. Horší je spíše to, že student se snažil jít na minimální míru textu, aby nebyl moc dlouhý, takže občas mu někde vypadne nějaká definice nějakého symbolu, nebo překlep v definici, apod. Chybně je např. definice 2.3.9, díky překlepům algoritmus 3.12, 3.14, definice 3.4.13, a další.
- 5. Formal aspects of technical report** **60 p. (D)**
Typografická stránka je až na pár výjimek v pořádku. Práce je psána anglicky, jak jsem zmínil, je těžko čitelná. Formulace jsou těžko čitelné, student používá sice formálně správná, ale matoucí značení (např. n je z N, podsouvá, že n je přirozené číslo, nikoliv, že se jedná o non-termínál), apod. Přečíst celý text vyžaduje velkou pozornost a velkou prerekvizitní znalost. Díky překlepům a podobným chybám se nedá ani pořádně analyzovat prezentovaný důkaz, neboť hned v úvodu je chyba (co je symbol D, jak je možné sjednotit symbol a množinu, apod.).
- 6. Literature usage** **85 p. (B)**
Student zvolil literaturu vesměs vhodně, je jasné, na co se odkazuje, bohužel slovník, který používá (že zavádí věc, která je již zavedena), to kazí.
- 7. Implementation results** **50 p. (E)**
Stran realizačního výstupu mám ambivalentní názor. Student vymyslel a nakódoval nový způsob tvorby LR rozkladové tabulky tak, aby zachoval sílu LR gramatik s tím, že velikost tabulky se bude blížit LALR/IELR rozkladové tabulce. Algoritmus, jak sám tvrdí, je pomalý (o řád i více), než stávající, ale je lepší v tom, co vytvoří (na počtu stavů uvádí konkrétní čísla). Bohužel ho nezačlenil do žádného známého nástroje, ani ho neumožnil použít způsobem, jaký je v podobných nástrojích běžný, ale udělal knihovnu, která je zcela nestandardní, no a, zejména, nepracuje se syntaxí řízeným překladem, ale s překladovými gramatikami, což tuto práci činí prakticky nepoužitelnou a neporovnatelnou s ostatními nástroji, což byl bod zadání. Oceňuji nový algoritmus, zcela nechápu to, kam jej student začlenil. Na CD není ani žádné DEMO, které by bylo možné použít "na hraní". Ano, protože takové vytvořit by stálo hodně úsilí, aby se to srovnalo s tím, co lze v ostatních nástrojích.
- 8. Utilizability of results**
Podle mě nulová, tak jak to je.
- 9. Questions for defence**
Co vás vedlo k tomu vytvořit, co jste vytvořil, místo toho, abyste do známého programu (bison, ANTLR, ...) dodělal nový generátor, zoptimalizoval jeho rychlost, nabídl komunitě?
- 10. Total assessment** **60 p. satisfactory (D)**
Student napsal anglicky DP, kterou však číst není příliš jednoduché. Vytvořil nový algoritmus, který má lepší vlastnosti, než stávající. Jeho implementace má výkonnostní vadu, také to, jak byl začleněn, nad čím pracuje, degraduje jeho využití a odsouvá jej buď k zapomnění, nebo do doby, než ho někdo pořádně pochopí a udělá standardní implementaci, nebo algoritmus začlení do stávajících nástrojů.

In Brno 6. June 2019

Kolář Dušan, doc. Dr. Ing.
reviewer