

Posudek vedoucího na diplomovou práci Petry Kuřinové
Polynomiální algoritmus pro binární PCP

Práce popisuje algoritmus řešící binární verzi Postova korespondenčního problému. Principy tohoto algoritmu jsou známy již třicet let, ale jeho složitost závisela na dvou otázkách: struktuře tzv. následníků a tzv. koncových bloků, která byla objasněna teprve nedávno. Z příslušných výsledků plyne, že algoritmus může být polynomiální. Cílem předkládané práce byl podrobný popis a implementace tohoto algoritmu. Přestože je základní myšlenka algoritmu velmi jednoduchá, naráží na několik komplikací, které měla uchazečka podrobně prostudovat a přesvědčivě ukázat, že polynomiálnost algoritmu neohrožují. Vedlejším úkolem práce bylo pokusit se nalézt instance problému zajímavé v tom smyslu, že ilustrují právě zmíněné komplikace, případně konstatovat, že nalézt takové instance je obtížné (je dokonce možné, že neexistují).

První čtyři kapitoly představují teorii v pozadí a známé výsledky. Výklad velmi těsně sleduje literaturu, ze které čerpá. Stálo by za úvahu, zda by velká část kapitoly 2.2., která reprodukuje vysoce technickou charakterizaci homomorfismů se stálou sufixovou složitostí, nemohla být vynechána ve prospěch důkazů základních nástrojů kombinatoriky na slovech, což by mohlo být užitečné nejen pro čtenáře, ale i pro autorku.

Jak je vidět z nadpisu kapitoly 5, měla by právě ona být jádrem práce. Popis algoritmu, který nabízí, je ale spíše neformálním přiblížením základních myšlenek, a není tedy příliš vzdálen od bodu, ve kterém práce začínala. Nejsem si jist, zda tato kapitola algoritmus objasní tomu, kdo ho již dobře nezná. V tomto směru jsou zřejmě užitečné uvedené příklady a srozumitelně představuje základní myšlenky také Úvod. V každém případě však pátá kapitola neposkytuje přesvědčivé důvody pro správnost algoritmu, což mělo být klíčovou přidanou hodnotou. Tato správnost závisí na teoretických poznacích, které jsou obsaženy ve třetí kapitole.

Kapitola 3 tedy tvoří přes svou krátkost samotné jádro diplomové práce a také měla reprezentovat vlastní přínos autorky z teoretického hlediska (tedy odhlédnuto od programátorské práce). Musím s lítostí konstatovat, že tato klíčová kapitola obsahuje velké nedostatky.

Méně důležitým problémem je čtenáři (mírně řečeno) nepřátelská terminologie: „dostatečný kandidát“, „maximální kandidát“ a „maximální dostatečný kandidát“.

(Podobně je tomu v kapitole 2.3 s termíny „dlouhé“, „krátké“ a „nedlouhé“ řešení, které umožňují řešení, které sice není „krátké“, ale je „nedlouhé“. Definice *dlouhého* řešení navíc obsahuje drobnou vadu, slovo τ by zřejmě mělo splňovat všechny požadavky Tvrzení 2.19, nejen rovnost (2.4).)

Závažnější je to, jak se s pojmem *maximálního dostatečného kandidáta*, který je pro správnost algoritmu rozhodující, zachází. Tento pojem je sám definován algoritmicke a nejsou zkoumány ani zformulovány vlastnosti, které by jeho funkci v algoritmu ospravedlňovaly. Definice je bez jasného důvodu uvedena dvakrát (dvě odlišné parafráze téhož algoritmického postupu), na str. 33 a na str. 44. Žádoucí vlastnosti maximálního dostatečného kandidáta jsou v zárodku přítomny nejspíš v Tvrzeních 3.3. a 3.4., které předcházejí jeho (prvnímu) popisu, ačkoli ho důkaz Tvrzení 3.4. již používá.

Tvrzení 3.3. a 3.4. na str. 32 jsou konečně tak, jak jsou formulována, neplatná a jejich důkazy jsou zmatečné. Ukážu několik problémů.

- Autorka se zřejmě sama nechala zmást svou krkolomnou terminologií. V Tvrzeních 3.3. a 3.4. nic nebrání tomu, aby slova w'_g a w'_h obsahovala dvě speciální situace, a tudíž žádné jejich prodloužení nemusí být maximálním kandidátem, který toto vylučuje. To by se ale dalo snadno napravit přidáním příslušné podmínky.
- Pokud by platily obě vysazené rovnosti důkazu Tvrzení 3.3., muselo by platit $m = n$. Příklad $m \neq n$ tedy vůbec není řešen, což je zvláštní zejména proto, že v důkazu hned následujícího Tvrzení 3.4. je podobný rozbor případů proveden.
- Jak již bylo řečeno, důkaz Tvrzení 3.4. používá dosud nedefinovaný pojem, smysl poslední věty bodu 1) tedy nedává žádný dobrý smysl.
- Oba důkazy navíc obsahují kroky, jejichž správnost není vysvětlena, je snad závislá na pomocných tvrzeních, která nikdy nebyla zformulována. Argumentace proto zůstává na příliš intuitivní úrovni, což zřejmě vysvětluje díry v důkazu.

Vzhledem k uvedeným problémům a nejasnostem nepovažuji za možné ani užitečné hodnotit podrobně implementaci algoritmu. Práce obsahuje několik ilustrativních příkladů, nikoli však příklady ve smyslu uvedeném na začátku tohoto posudku.

Je zřejmé, že práce naplňuje požadavky jen částečně. Přesto doporučuji, aby se ji autorka pokusila obhájit.

Praha 24. ledna 2013

Štěpán Holub