Posudek oponenta bakalářské práce

Název: Probabilistické algoritmy pro prvočíselnost
Autor: Natálie Tejkalová

Shrnutí obsahu práce

Práce podává popis dvou klasických probabilistických algoritmů na testování prvočíselnosti – Rabin-Millerova a Solovay-Strassenova. Zvláštní pozornost přitom autorka u obou algoritmů klade na odvození horních mezí pro pravděpodobnost chyby v jedné iteraci algoritmu a pokouší se podat ucelený výklad k tomuto účelu nezbytné teorie. Je též prezentováno několik „naivních" probabilistických testů, u nichž je vždy naznačen důvod jejich praktické nepoužitelnosti pro testování velkých čísel. V druhé části práce se autorka zabývá formalizací pojmu probabilistického algoritmu a předvádí několik tříd výpočetní složitosti odpovídajících různým možným formalizacím. Je též pojednáno o problému generování náhodných čísel v praxi.

Celkové hodnocení práce

Téma práce. Obtížnost tématu je odpovídající bakalářské práci. Zadání práce – nastudovat Rabin-Millerův a Solovay-Strassenův algoritmus a podat výklad matematické teorie potřebné pro pochopení jejich fungování – autorka splnila (s výhradami uvedenými níže).

Vlastní příspěvek. Práce je koncipována jako rešeršní, bez aspirace na dosažení nových výsledků. Prezentované definice a důkazy jsou klasické a většinou dobře známé. Vlastní autorčin příspěvek spočívá zejména ve vymezení a výkladu poznatků potřebných pro důkazy korektnosti zkoumaných algoritmů. Tento výklad je ucelený a dostatečně obsáhlý avšak ne vždy kvalitně a jednotně podaný (podrobněji viz Matematická úroveň níže).

Matematická úroveň. Celková matematická úroveň práce je podprůměrná – zejména kvůli celkově velmi slabé kapitole 5 a velkému množství chyb v jinak vcelku dobrém výkladu kapitol 2, 3 a 4.

Kapitola 5 věnovaná obecnějšímu zkoumání pojmu probabilistického algoritmu je nejslabší částí práce. Zde podaný výklad formalizace pojmu probabilistického algoritmu pomocí Turingových strojů je velmi vágní, jdoucí po povrchu a obsahující mnoho nepřesností. Kapitola obsahuje i několik vyložených chyb a smysl postrádajících formulací ukazujících na nepochopení (či snad jen neschopnost reformulace) z literatury načtených pasáží (např. definice pravděpodobnostního algoritmu na stranách 33-34 je v uvedeném zdroji [8] správně, v práci je však podána zcela chybně).

Kapitoly 2, 3 a 4 podávají ucelený výklad Rabin-Millerova a Solovay-Strassenova algoritmu. Tento výklad je úplný – shrnuje nezbytnou matematickou teorii a podává důkaz korektnosti obou algoritmů – a (odmyslíme-li značné množství chyb – viz níže) rigorózně podaný. Důkazy jsou dostatečně podrobné a většinou neobsahují vážné logické chyby či mezery. Definice a znění tvrzení jsou v principu formulovány správně (což lze ovšem u kompilační práce pokládat za samozřejmost). Autorka tak v této části prokázala porozumění potřebné matematické teorii.

Kvalitu této části práce však sráží dolů velké množství nejrůznějších chyb a nedostatků (neuvádím zdaleka kompletní výčet, jen typické chyby a příklady jejich výskytu):

- **Velmi časté opomenuté předpoklady v tvrzeních:**
  - Věta 6 neplatí (ani nedává smysl) pro $p = 2$.

- V lemmatu 6 je opomenut předpoklad $N > 0$.
- V definici Jacobiho symbolu je zapomenuto na případ $N = 1$.
- Další podobné chyby skrze celý text.

- **Logicky nepřesné argumenty:**
  - V důkazu věty 5 se na třetím řádku tvrdí, že (schematicky vyjádřeno) $\varphi \Rightarrow \psi$ „a také opačně“ (sic!) $\neg\psi \Rightarrow \neg\varphi$.
  - V důkazu lemmatu 1 a) či předposlední větě důkazu lemmatu 7 dokazované samozřejmě platí ale ne díky tam uvedenému zdůvodnění.
- **Nepřesnosti a chyby v důkazech:**
  - V důkazu věty 5 ve druhém odstavci není $Ab_0 \subseteq B$ – to platí až modulo m. Dále se však v důkazu pracuje jako kdyby inkluze platila doslovně.
  - V důkazu věty 9 v případě, kdy n má právě dva prvočíselné dělitele, se n vyjadřuje jako $n = pq$ (zapomenuté mocniny).
  - V důkazu lemmatu 10 při počítání počtů řešení jednotlivých kongruencí a jejich soustav není uváděno na jakých intervalech se počet řešení počítá. Celá pasáž důkazu je tak značně nejasná.
- **Zmatené a nekonzistentní značení:**
  - Na mnoha místech je nekonzistence v označení téže veličiny ve znění tvrzení a jeho důkazu nebo dokonce i v rámci jednoho znění tvrzení či důkazu:
    - V definici Jacobiho symbolu je v několika výskytech M zaměněno za b.
    - Ve znění lemmatu 6 se vyskytující N a M_1, M_2 jsou v důkazu bez upozornění značeny po řadě jako n a a, b.
    - Ve znění i důkazu věty 7 je několikrát náhodně zaměněno p a n.
    - Další (méně křiklavé) obdobné případy skrze celý text.
  - Na straně 18 je nekonzistence v pořadí argumentů M, N funkce Solovay-Strassen mezi její definicí v algoritmu 2 a použitím v následujícím příkladu.
- **Nejednotnost a neuspořádanost terminologie:**
  - Řád grupy je v důkazu tvrzení 2 označován jako „stupeň“, ve zbytku práce jako „řád“.
  - Některé definované pojmy nejsou v dalším výkladu vůbec použity, přestože by to mnohdy bylo vhodné (příkladem je pojem *modulární reprezentace* definovaný na str. 6).
  - Některé již pojmenované fenomény jsou v následujících pasážích práce pojmenovány znovu jiným názvem a definovány zjevně ekvivalentní definicí, načež jsou oba názvy v dalším používány střídavě (např. generátor grupy $\Phi(n)$ a primitivní kořen modulo n).
- **Zbytečně duplicitní informace:** Ze vzorců pro výpočet Eulerovy funkce na straně 4 jsou první dva uvedené triviálními důsledky třetího.
- **Špatně zapsané vzorce:**
  - Na mnoha místech chybějí závorky v místech, kde jsou nezbytné:
    - důkaz věty 5, odstavec 2, řádek 3,
    - důkaz věty 6, případ A),
    - několik dalších obdobných situací.
  - Chybí části vzorců a formulí:
    - poslední věta důkazu lemmatu 6,
    - dolní mez ve znění věty 10,

    - několik dalších obdobných případů.
  - Relace dělitelnosti či nedělitelnosti je několikrát zapsána v opačném pořadí
    - v příkladě fungování Solovay-Strassenova algoritmu na straně 18,
    - na konci prvního odstavce důkazu věty 8.
- **Různorodé drobné matematické chyby** (uvádím jen ty, které se několikrát opakují):
  - záměna či posunutí indexů o ± 1,
  - záměna $\leq$ a $<$,
  - záměna $+1$ a -1,
  - záměny značení (n místo N apod.)
- **Množství překlepů všeho druhu**

Práce se zdroji. Práce obsahuje poměrně rozsáhlou bibliografii. Ve většině případů je uvedeno jakého zdroje se autorka při svém výkladu přidržuje (snad s výjimkou výkladu obecné algebraické teorie v úvodních kapitolách, kterou však lze označit za matematický folklor nevyžadující citování). Oponent si není vědom žádných zkopírovaných či doslovně otrocky přeložených ani zásadních rozsáhlejších převzatých a přitom necitovaných pasáží.

Formální úprava. Formální úpravu práce lze hodnotit vzhledem k pravděpodobné nezkušenosti autorky se systémem LaTeX jako dostačující. Vytknout lze jen časté nekonrolované rozdělení vzorců okrajem řádku a poněkud nepřehledný způsob číslování tvrzení (pro každý typ tvrzení – lemma, tvrzení, věta – samostatná číselná řada). Na několika místech je též krátký matematický text (typicky proměnná) vysázen mimo matematický mód a tudíž graficky odlišně.

Použité jazykové prostředky odpovídají charakteru práce.

Závěr

Autorka splnila zadání práce co do obsahu avšak kvalita textu je podprůměrná. Množství chyb i v pasážích, kde bylo možné se blízce držet kvalitní existující literatury, vyvolává dojem, že práci nebylo věnováno dostatečné úsilí či čas. S výjimkou kapitoly 5 by přitom práce po opravení uvedených nedostatků mohla být poměrně kvalitní a poskytnout zajímavý a ucelený výklad o Rabin-Millerově a Solovay-Strassenově algoritmu.

Práci považuji za podprůměrnou avšak dostačující. Doporučuji ji uznat jako bakalářskou práci a ohodnotit ji známkou dobře (3).

Petr Glivický
Katedra teoretické informatiky a matematické logiky, MFF UK v Praze
11.9.2013