

# Posudek

## oponenta na disertační práci Ing. Zbyňka Křivky „Přepisující systémy s omezenými konfiguracemi“

Posuzovaná disertační práce z oblasti teorie formálních jazyků se zabývá přepisujícími systémy, inspirovanými formálními gramatikami a automaty. Zavádí nové modely přepisujících systémů a srovnává jejich vlastnosti se známými modely regulovaných formálních gramatik. Námět práce odpovídá oboru disertace a je aktuální z hlediska současného stavu výzkumu v oboru formálních jazyků.

Přínosy práce sám doktorand shrnuje na stránce 89. Jde jednak o zavádění unifikovaného přístupu ke studiu přepisujících systémů zahrnujících formální gramatiky i automaty. Doktorand se pokouší o zobecnění generativních a akceptujících systémů pod společnou koncepcí *přepisujících systémů* podle definic 2.11 a 2.12. Skutečně, základní typy generativních gramatik dle Chomského hierarchie a jim odpovídající formální automaty se dají pod pojem přepisujícího systému zahrnout. Ovšem definice přepisujícího systému selhává, jakmile jde o mírně nestandardní modely (paralelní gramatiky, regulované gramatiky, automaty s více zásobníky apod.). Kamenem úrazu je zde podmínka konečnosti relace  $R$  přepisujícího systému v kombinaci s definicí přepisující relace 2.12. Například stěžejní koncept práce, #-přepisující systém, nevyhovuje definicím 2.11 a 2.12 (je nutno zavést dodatečná omezení přepisující relace, čímž se ovšem ztrácí obecnost těchto definic).

Zavedení nového obecného konceptu je oprávněné tehdy, jestliže umožňuje užitečnou aplikaci společných matematických nástrojů na širokou třídu jeho instancí; je otázkou, zda koncept přepisujícího systému toto dovoluje. V tomto bodě nejsem zcela přesvědčen o přínosu práce.

Další přínosy práce jsou ovšem zcela jednoznačné. Jde o klasifikace jistých typů omezujících podmínek formálních modelů (konečný index,  $k$ -limitovanost u gramatik a hloubka u automatů) a jejich vzájemné vztahy. Dále pak zavedení modelu #-přepisujícího systému a studium jeho vlastností. Konečně pak studium vlastností řízených a omezených zásobníkových automatů. Společnou motivací těchto modelů je jednak dosažení generativní/akceptující síly na úrovni vyšší než bezkontextové s použitím regulovaných bezkontextových komponent. Dále pak studium jistých omezení se zřetelným vztahem k algoritmům syntaktické analýzy formálních jazyků.

Výsledky dosažené doktorandem v této oblasti byly publikovány celkem v deseti citovaných příspěvcích ve sbornících z konferencí a seminářů (převážně tuzemských) a v jednom článku v časopise IJFCS (impaktní faktor 0,5). Tento objem a kvalitu publikační činnosti považuji na úrovni disertace za dostatečný. Současně konstatuji, že výsledky zahrnuté v práci, jejich prezentace a technické provedení důkazů svědčí jednoznačně o vědecké erudici uchazeče. Dále se vyjádřím k jednotlivým aspektům práce.

Doktorand věnuje velkou péči zasazení své práce do širšího kontextu teorie formálních jazyků a prokazuje značný nadhled nad zpracovanou látkou, ze-

jména v kapitolách 2 a 3, což hodnotím jednoznačně kladně. K stěžejnímu novému konceptu práce, #-přepisujícím systémům, mám ovšem následující poznámku. Nelze si nevšimnout podobnosti #-přepisujících systémů a některých normálních forem maticových gramatik. U těchto normálních forem maticových gramatik je množina neterminálních symbolů rozdělena na dvě části, z nichž jedna pracuje prakticky shodně jako stavy u #-přepisujících systémů. Hlavní rozdíl by pak byl v tom, že maticové gramatiky přepisují jednotlivé neterminály ve větné formě, zatímco #-přepisující systémy přepisují číslované výskyty „neterminálu“ #. Této podobnosti s maticovými gramatikami doktorand nikde nevyužívá, ačkoli se lze domnívat, že by dovolila aplikovat řadu výsledků z oblasti maticových gramatik na případ #-přepisujících systémů.

Sám o sobě je koncept #-přepisujících systémů poměrně zajímavý a dovoluje podrobně studovat aspekty omezování generativních systémů pomocí indexů neterminálů. Co se týče kapitoly 5 obsahující vlastní výsledky doktoranda, tyto jsou vědecky přínosné a odborně kvalitní. Domnívám se však, že využití normálních forem regulovaných gramatik známých z literatury by mohlo některé důkazy zjednodušit.

V kapitole 6 se doktorand vyjadřuje k možným aplikacím dosažených výsledků. Jeho náměty jsou nepochybně podnětné a některé vypadají slibně. Oponent ovšem musí konstatovat, že ačkoli pracoval v oblasti regulovaných gramatických systémů přes 10 let a spolupracoval s autory stěžejních monografií v oboru, neví dosud o žádných praktických aplikacích regulovaného přepisování v oblasti počítačového hardware či software. Lze jen doufat, že (snad i s přispěním uchazeče) se tento stav v budoucnu změní. Na druhé straně regulované gramatiky již mnohokrát pomohly charakterizovat sílu jiných, více aplikačně zaměřených výpočetních modelů, v čemž je jejich jednoznačný přínos.

Po formální stránce je práce vypracována velmi přehledně až precizně, se všemi potřebnými náležitostmi. Vedle elementárních náležitostí, jako obsah, seznam literatury apod., čtenář najde přehled kapitol, popis základních matematických konvencí, přehled elementárních definic a seznam matematických symbolů, což vše činí text snadno čitelným. Rovněž po stránce typografie nemám vážnějších výhrad. Seznam literatury vykazuje odchylky od abecedního řazení (viz [22] až [37]), což poněkud ztěžuje orientaci.

Z dílčích výtěk k práci uvádím následující:

- občasné gramatické poklesky (s. 12, ř. 19, studovali se gramatiky, s. 33, ř. 12, automaty se začali) a překlapy (*trzené*, několik výskytů počínaje stranou 59)
- s. 12, ř. -13, odstavec za konceptem 2.2 má začínat „Instancí formálního modelu je například...“
- s. 12, ř. -10, zavádět pojem „správná instance“ nedává smysl; instance, která není správná, vůbec není instancí
- s. 16, proč jsou nezáporná, resp. kladná celá čísla značena  $\Psi$ , resp.  $I$  (přičemž jedno je písmeno řecké a druhé latinské abecedy), namísto obvyklého  $\mathbb{N}_0$ , resp.  $\mathbb{N}$ ? Užívání nestandardní notace činí text hůře čitelným.

- s. 22, koncept konfigurace se nezdá zcela navazovat na koncept přepisujícího systému, s nímž je svázán. Za prvé, relace  $R$  přepisujícího systému je binární a konečná, takže přepisující systém vždy pracuje nad jediným řetězcem. Naproti tomu konfigurace může být obecně  $n$ -tice komponent,  $n \geq 1$ . Je sice možno komponenty (anebo alespoň určité jejich části) během kroku přepisu propojit do jediného řetězce, aby je mohl přepisující systém modifikovat, taková konstrukce se ale nezdá být přirozená.

Za druhé, při zavedení konfigurace nejsou její komponenty nijak specifikovány (mohou to být libovolné objekty), ovšem formule na s. 22, ř. 11 implicitně předpokládá, že pro každou komponentu je definována její délka. Bylo by vhodné upřesnit charakter komponent.

- pojmosloví v konceptu 2.4. na s. 22 je kontra-intuitivní: opakem pojmu *proměnná* je pojem *pasivní symbol*, zatímco opak pojmu *aktivní proměnná*, tedy proměnná, která není aktivní v aktuální konfiguraci, není nijak pojmenován
- s. 23, nepovažuji za vhodné používání notace  $[A \rightarrow x]$  pro označování pravidel; tato notace standardně značí celou část čísla a v literatuře věnované formálním jazykům jsem se s podobným značením pravidel dosud neseťkal
- s. 23, ř. 12,  $\text{lab}(R)$  má být zřejmě  $\text{lab}(P)$
- s. 49, proč pojem *Deep Pushdown Automaton* má zkratku DTDP nepodobnou jeho názvu, kdežto zkratky dalších příbuzných modelů automatů vždy končí PDA, jak lze očekávat, například RDPDA, RCPDA?
- s. 51, první odstavec kapitoly 4.2.1 neodpovídá navazující definici 4.10, a není pravdou, že automat podle definice 4.10 nepovoluje současnou existenci pravidel  $r_1 : 2pA \rightarrow qx$  a  $r_2 : 2pC \rightarrow oy$ . Viz též příklad 4.5, kde se podobné kombinace vyskytují.
- s. 51, definice 4.10, srozumitelnější by bylo napsat, že pro každý stav  $q \in Q$ ,  $mqA \rightarrow pv \in R$  a  $m'qA' \rightarrow p'v' \in R$  implikuje  $m = m'$ , pro  $A, A' \in \Gamma - T$ ,  $v, v' \in \Gamma^+$  a  $p, p' \in Q$ . Analogicky v definici silně deterministického automatu.
- s. 58, důkaz lemmatu 5.1, co znamená  $g(p)$ , kde  $p$  je návěští pravidla z  $P$ ? Jde o množinu následných pravidel? V definici 2.21 taková notace není zavedena.
- s. 66, pro platnost důsledku 5.9 je nezbytná ještě inkluze  $\mathbf{REG} \subseteq \mathcal{L}(1\text{-RLIN}\#\mathbf{RS})$ . Jakkoli je triviální ukázat její platnost, měla by být v textu explicitně zmíněna.
- s. 69, otevřený problém či hypotéza 5.1 se zdá být snadno dokazatelný už pro systémy indexu 1 (a s jediným stavem), protože neomezená pravidla umožňují hranici  $\#$  volně migrovat po větě formě (pravidly

typu  $\#a \rightarrow a\#$  a  $a\# \rightarrow \#a$ ) a simulovat pravidla Chomského gramatiky typu 0; technický problém s kontrolou, zda větná forma již obsahuje samé terminály, se rovněž zdá být lehce řešitelný

- s. 84, definice 6.3, výraz  $\text{card}(\{l \mid l \leq i, i \in K(H, p)\})$ , kde implicitně  $l \in I$  (protože na  $l$  není uloženo žádné omezení) znamená totéž co  $\text{sup}(K(H, p))$ . Pak tedy definice říká následující: stavově-řízený- $k$ -kanonický výpočetní krok (ve stavu  $p$ ) je takový, když všechna pravidla v  $R$  se stavem  $p$  na levé straně mají index nejvýše  $k$ . Jinými slovy, koná-li systém stavově-řízené- $k$ -kanonické kroky, lze jej nahradit ekvivalentním systémem indexu nejvýše  $k$  (protože pravidla vyššího indexu a stavy s nimi asociované nelze nikdy použít). Taková definice ovšem neříká (nedefinuje) nic nového, a také jí neodpovídá Algoritmus 6.1. Něco tu zřejmě není v pořádku.

Přes uvedené dílčí výtky hodnotím práci celkově jako kvalitní, prokazující schopnosti a erudici uchazeče v daném oboru, a přinášející originální výsledky. Dle mého názoru práce jednoznačně odpovídá obecně uznávaným požadavkům na disertační práci k udělení akademického titulu.

V Opavě dne 9. září 2007

Doc. Ing. Petr Sosík, Dr.