



## РЕЦЕНЗИЯ

от проф. д-р Тинко Величков Тинчев

за дисертационния труд „Finite-State Automata, Transducers and Bimachines: Algorithmic Constructions and Implementations“  
представен от доц. д-р Стоян Милков Михов  
за придобиване на научната степен “доктор на науките”  
в професионално направление 4.6. Информатика и компютърни науки

### За представените документи

Представеният от доц. Михов дисертационен труд е на английски език в обем от 225 страници и е структуриран в 8 глави (210 страници), съдържание, увод, заключение и библиография, съдържаща 48 цитирани източника. В заключението са включени декларация за оригиналност и приноси – научни и научно-приложни. Дисертацията е базирана на съвместна с Клаус Шулц монография, публикувана от Cambridge University Press в серията Cambridge Tracts in Theoretical Computer Science през 2019 г., една глава от книга в серията Trends in Linguistics–Studies and Monographs и 11 статии. Три от тези статии са в реномирани списания с импакт фактор – две в Computational Linguistics (Q1) и една в Theoretical Computer Science (Q3); други седем са в списания и сборници с SJR фактор. Според Scopus те са цитирани 227 пъти (без автоцитатите). Всички цитирания на научни трудове на доц. Михов според Scopus са над 300 (без автоцитатите). Следователно наукометричните показатели на доц. Михов далеч надхвърлят минималните национални изисквания. Представени са автореферат на български език (58 стр.) и автореферат на английски език (57 стр.), които вярно отразяват дисертационния труд. Представени са и всички други изискуеми от ЗРАСРБ и правилниците за прилагането му. Няма да изреждам всички тези документи, а само ще отбележа един от тях – декларация от проф. Клаус Шулц, че във всяка от главите на съвместната им монография приносът на д-р Стоян Михов е по-голям от неговия, като главите, отнасящи се до езика  $S(M)$  и включените в монографията програми на  $S(M)$ , са принос изцяло на д-р Стоян Михов.

### За областта

В началото на 70-те години на миналия век се ширеше мнението, че теорията на крайните автомати е от математическа гледна точка зряла и до голяма степен завършена теория с теоремите на Клини, Рабин-Скот, Майхил-Нероуд, МакНотон, Салома и др.; на дневен ред са по-изразителни модели от типа на стекови автомати, безконтекстни граматика и пр. Сега, както добре знаем, истината се оказва друга – бурното разрастване на изчислителните системи предизвика глад за ефективни алгоритми и крайни структури за представяне и обработка на знание. В частност, развиха се редица разновидности на крайните преобразуватели, често пъти поради специфични приложения или за



техническо удобство на изложението. С това взните като че ли се наклониха на страната на алгебрата и логиката. Предложеният дисертационен труд предлага разумен общ алгебричен подход с ясна алгоритмична окраска. Той е в активно развивана област, както като математика за теоретичната компютърна наука, така и като приложения в компютърната лингвистика и изкуствения интелект.

## Преглед и оценка на съдържанието на дисертацията

Дисертацията е единно цялостно научно и научно-приложно изследване, следващо ясна и подходяща методология и съдържа всички необходими понятия за проследяването му, като, грубо казано, има две естествено свързани и взаимно обусловени части. Първите 6 глави съдържат математическата теория на разглежданите преобразуватели с краен брой състояния – крайни автомати, различни видове преобразуватели и бимашини, тук са и научните теоретични приноси на доц. Михов: дадени са редица ефективни алгоритмични конструкции, за които е доказана коректност и е изследвана сложността им. В, условно казано, втората част те са програмно имплементирани. Тази част съдържа съществени научно-приложни приноси на доц. Михов – разработен е удобен за математици целеви програмен език  $S(M)$  (глава 7), на който в глава 8 са реализирани алгоритмите от първата част, а така също и четири примера за приложения на развитата техника към конкретни практически важни нетривиални задачи. Тук веднага трябва да кажем, че за този език доц. Михов е създал компилатор, който превежда програмите в ефективен код на  $S$ , който след компилиране може да се изпълни. По такъв начин дисертацията съдържа коректен изпълним код за редица алгоритми.

Глава 1 е уводна и въвежда необходимите за по-нататъшно изложение основни понятия – думи,  $n$ -орки от думи, моноидни думи, езици, моноидни езици и класическите операции с думи и езици. Същевременно е мотивиран избраният моноиден подход с цел обобщение, естествено обхващащо класически понятия като думи,  $n$ -орки от думи, езици, автомати и  $n$ -лентови автомати.

В глава 2 се дефинират и изучават моноидни крайни автомати. Показана е затвореността на класа на моноидните езици относно хомоморфизми и моноидните операции обединение, произведение и звезда на Клини. Дефинират се понятията моноиден регулярен език и моноиден регулярен израз. Доказана е класическата теорема на Клини за моноидните регулярни езици и моноидните регулярни изрази. В тази глава се вижда ясно основната цел на избрания обобщен подход - не просто да се докаже конструктивно съществуването на обект с дадено свойство, а да се даде ефективна директна конструкция на обект, за който се доказва, че има желаното свойство. Обикновено елегантното и икономично от гледна точка на изложението доказателство е в бележка след доказателството, че конструираният обект има желаното свойство. В последния параграф на глава 2 се показва елиминиране на недостижимите състояния на моноиден краен автомат и два начина за елиминиране на  $e$ -преходи.

В глава 3 фокусът е детерминизация и минимизация, като основно се изучават класическите крайни автомати (моноидът е на думите в крайна азбука). Показани са ефективни конструкции, доказващи затвореността на регулярните езици относно теоретико множествени допълнение, сечение и разлика, а така също и обръщане на



думите. Причината за ограничаването до класическите крайни автомати е показана в параграф 7 и се корени във факта, че теоретико множественото допълнение на моноиден регулярен език не винаги е регулярен език. Един разумен вариант са понятията псевдо-детерминиран и псевдо-минимален моноиден краен автомат – съответният свободен компаньон да е детерминиран, съответно минимален. Показано е как се получават на базата на развитите в първите 5 параграфа конструкции за детерминизация и минимизация на класически крайни автомати. Струва си да се отбележи обстояйното изследване на релацията на Майхил-Нероуд в параграфи 4 и 5, която по-нататък се превръща в мощно средство за изследване на аналогични проблеми минимизация на разнообразни видове преобразуватели. Едно директно приложение на представената техника за минимизация е към въведените в параграф 6 оцветени детерминирани крайни автомати, които се използват съществено при псевдо-минимизация на бимашини в глава 6.

В глава 4 се изучават многолентовите крайни автомати като специален вид моноидни крайни автомати. Дават се конструкции, показващи затвореността на езиците, разпознавани с моноидни многолентови автомати, относно декартови произведения и проекции, а в случая на двулентови автомати и относно операцията обръщане. Описана е конструкцията, даваща релационната композиция за двулентови посимволни крайни автомати, с което се получава и конструкция за композицията на класически  $n$ -лентови автомати. Съществена част от главата е посветена на моноидните крайни преобразуватели и тяхната функционалност в случая, когато и втората лента е свободен моноид (крайните класически преобразуватели). Целта е да се получи ефективен метод за разпознаване на свойството функционалност на класическите крайни преобразуватели. Представената конструкция е базирана на нетривиална техника, използваща редица специални видове преобразуватели и така наречената функция на допустимия баланс. Полученият метод е необходим за изследването на детерминизацията на крайни преобразуватели в следващата глава, поради теорема на Schützenberger.

Глава 5 е посветена на детерминирани крайни преобразуватели и е особено интересна, поради важноста от гледна точка на приложенията на подпоследователните преобразуватели (функционални и детерминирани по входната лента) в компютърната лингвистика, например, за разпознаване на реч и обработка на текст. В параграф 2 се описва по същество конструкцията на Roche & Schabes за секвенциализация: итеративна процедура, която започвайки от класически реално-временен (по втората лента няма  $\varepsilon$ -преходи) краен преобразувател без недостижими състояния  $T$ , последователно строи крайни преобразуватели като, ако завърши след краен брой стъпки, то резултатът е класически подпоследователен преобразувател  $T'$ , еквивалентен на  $T$ . При това, ако началният преобразувател  $T$  е с ограничена вариация, то процедурата завършва работа след краен брой стъпки. Оттук веднага следва характеристика на представимите чрез класически подпоследователни преобразуватели регулярни функции между думи: те са точно онези регулярни функции, които са с ограничена вариация. В параграф 3 доц. Михов намира подходяща елегантна и далеч неочевидна форма на забелязана от Béal et al. връзка между свойството на  $T$  да е с ограничена вариация и валидността на елементарно неравенство, свързано с броя на състоянията на  $T$ , най-дългия преход по втората лента на  $T$  и допустимия напредък на състояние от квадратирания автомат, асоцииран с  $T$ . Това му позволява да вгради в алгоритъма за секвенциализация



разпознаването на въпросното свойство, с което алгоритъмът добива цялостност. В параграфи 4 и 5 се изследва минимизацията на крайните подпоследователни преобразуватели на базата на аналог на релацията на Myhill-Nerode и се дава подходяща конструкция, която е базирана на първоначално преобразуване в така наречената канонична форма. Именно в този алгоритъм за преобразуване в канонична нормална форма, с малка сложност и използващ изцяло чисто автоматни техники, е един от сериозните теоретични приноси на доц. Михов. В последния параграф на тази глава се демонстрира пренасянето на резултати за детерминизация, разпознаване на функционалност, разпознаване на ограничена вариация и минимизация, получени за класически подпоследователни преобразуватели, за моноидни подпоследователни преобразуватели над адитивния моноид на естествените числа.

Глава 6 е посветена на въведените от Schützenberger крайни преобразуватели, наричани бимашини и съдържа, според мен, най-важните теоретични приноси на доц. Михов: псевдо-минимизация и директна конструкция на композицията на бимашини. Ефективността на тези операции до голяма степен определя практическата използваемост на бимашините. Тук подхода е изцяло оригинален и е базиран на моноидни крайни преобразуватели. За всяка моноидна бимашина се конструира подходящ моноиден краен преобразувател. И по-интересната и важна конструкция е в обратната посока, на всеки класически реално-времеви функционален преобразувател (с минимални ограничения) се съпоставя моноидна бимашина с подходяща изходна функция. Така, от една страна, се показва, че регулярните функции върху думи, които на празната дума съпоставят празната дума, са точно представимите с бимашини функции. От друга страна, с тази конструкция се получава бимашина, изобщо казано, със значително по-малък брой състояния.

Научно-приложната част на дисертацията – глави 7 и 8 – е органично свързана с теоретичната част. В глава 7 е представен по издържан начин разработеният от доц. Михов декларативен функционален език  $S(M)$ , предназначен както за математици, които искат да пишат ефективни програми, но не могат да пишат на  $S$ , така и за специалисти в областта на компютърните технологии. Несъмнено езикът ще бъде полезен за студенти и докторанти, изучаващи или ползващи техниките най-общо казано на крайните преобразуватели. Проектирането, имплементирането и тестването на  $S(M)$  само по себе си е доволно нетривиално както в идейно, така и в техническо отношение огромна по-обем работа. Компиляторът на Михов дава ефективен код на  $S$ , включително поддръжка на паметта. Това е позволило в глава 8 да бъдат описани редица нетривиални алгоритми по един напълно разбираем начин. Едно допълнително удобство, предоставяно от компилатора, е възможността за компилиране до LaTeX-файл, съдържащ красиво оформен псевдокод.

В глава 8 са дадени имплементации на  $S(M)$  на основните конструкции, представени в дисертацията. Всички те са придружени от съдържателни обяснения и коментари. Приложени са и 4 интересни примера за приложни програми, написани на  $S(M)$  и използващи представените алгоритми. Всички те са смислени, полезни и не лесни за непознаващи технологиите на крайните преобразуватели. Дадени са подсказки за модификации, които окуражават читателя да ги адаптира за свои цели. За първата група алгоритми, отнасящи се за крайни автомати, пример 8.1.16 е за краен детерминиран



автомат разпознаващ валидните дати от грегорианския календар във фиксиран формат. За втората група алгоритми, отнасящи се за класически крайни преобразуватели, пример 8.2.13 е за програма реализираща пълната функционалност, която е необходима за правописна корекция – функция за тестване на срещане в даден речник и функция връщаща множество на близки думи в смисъл на разстоянието на Левенщайн. Третата група алгоритми, отнасяща се за детерминирани крайни преобразуватели, е илюстрирана с програма, пример 8.3.10, даваща функционалност, която се използва при синтез на реч. Конструира се минимален подпоследователен преобразувател, който преобразува число, записано в десетична бройна система, в неговата английска фонетизация. Четвъртата група алгоритми, отнасяща се за бимашини, е пример 8.4.4 за реализация на основни аритметични операции с неограничени естествени числа с помощта на бимашини.

### **Изградени цялостни впечатления**

Представени са всички необходими документи, предвидени от ЗРАСРБ и правилниците за прилагането му.

Наукометричните показатели на доц. д-р Стоян Милков Михов далеч надхвърлят минималните критерии, описани в Правилника за прилагане на ЗРАСРБ. Авторефератите на български език и на английски език правилно отразяват съдържанието на дисертацията. Авторската справка правилно отразява основните научни и научно приложни приноси в дисертацията.

Дисертацията е написана в изискан добре балансиран стил. Фигурите и примерите са добре обмислени и подпомагат за по-бързото и непосредствено разбиране.

Конструкциите са точно и ясно описани. Твърденията са коректно доказани.

Представената дисертация е великолепен пример за успешна и плодотворна симбиоза на задълбочено високо стойностно математическо изследване и полезни нетривиални практически приложения.

### **Заклучение**

Убедено смятам, че дисертационният труд съдържа съществени научни и научно-приложни приноси наред с всички други изискуеми от ЗРАСРБ и правилниците за неговото прилагане качества и убедено **препоръчвам на Научното жури да присъди научната степен “доктор на науките”** в професионално направление **4.6. Информатика и компютърни науки** на доцент д-р Стоян Милков Михов за представения дисертационен труд **„Finite-State Automata, Transducers and Bimachines: Algorithmic Constructions and Implementations“** (“Крайни автомати, преобразуватели и бимашини: алгоритмични конструкции и имплементации”).

30.03.2020 год.  
София

Рецензент:



(проф. д-р Тинко Тинчев)