Matematický ústav
Slovenská akadémia vied

Alexander Szabari

Autoreferát dizertačnej práce

Popisná zložitosť regulárnych jazykov

na získanie vedecko-akademickej hodnosti
Philosophiae Doctor
v odbore doktorandského štúdia 11-80-9 Teoretická informatika
Košice, 2010

Dizertačná práca bola vypracovaná v externej forme doktorandského štúdia na Matematickom ústave Slovenskej akadémie vied.

Predkladateľ: Alexander Szabari
Matematický ústav SAV, dislokované pracovisko Košice
Košice

Školiteľ: RNDr. Galina Jirásková, CSc.
Matematický ústav SAV, dislokované pracovisko Košice
Košice

Oponenti: Prof. RNDr. Viliam Geffert, DrSc.
Ústav informatiky
Prírodovedecká fakulta
UPJŠ v Košiciach

Doc. RNDr. Libor Polák, CSc.
Ústav matematiky a statistiky
Přírodovědecká fakulta
Masarykova univerzita, Brno

RNDr. Tomáš Masopust, Ph.D.
Matematický ústav
Akademie věd České republiky
pracovisko Brno

Autoreferát bol rozoslaný dňa 30.9.2010.

Obhajoba dizertačnej práce sa koná dňa 12.11.2010 o 13.00 hod. v miestnosti M-213 na FMFI UK Bratislava pred komisiou pre obhajobu dizertačnej práce v odbore doktoranského štúdia, vymenovanou dňa 29.9.2010 predsedom spoločnej odborovej komisie vo vednom odbore 11-80-9 Teoretická informatika.

Predseda spoločnej odborovej komisie:
Prof. RNDr. Branislav Rovan, PhD.
Katedra informatiky
Fakulta matematiky, fyziky a informatiky
Univerzity Komenského
Mlynská dolina, 842 48 Bratislava

1 Úvod

Regulárne jazyky sú najjednoduchšie jazyky v Chomského hierarchii. Boli intenzívne skúmané po niekoľko desaťročí. Napriek ich jednoduchosti, zostávajú niektoré dôležité problémy nevyriešené. Spomeňme napríklad otázku, koľko stavov je potrebných a nevyhnutných v najhoršom prípade pre dvojsmerné konečnostavové automaty na simuláciu nedeterministických stavových automatov. Táto otázka úzko súvisí s veľmi dobre známym otvoreným problémom, či medzi triedami jazykov DLOGSPACE a NLOSPACE platí rovnosť [3, 31, 41].

V posledných rokoch bol zvýšený záujem o výskum v teórii automatov. Diskusie na túto tému možno nájsť v [20, 46]. Mnoho aspektov v tejto oblasti sa teraz skúma do podrobnosti. Jedným z takýchto aspektov je popisná zložitosť, ktorá študuje cenu popisu jazykov v rôznych formálnych systémoch.

Stavová zložitosť regulárneho jazyka je najmenší počet stavov v akomkoľvek deterministickom konečnostavovom automate (DFA), pre daný jazyk. Nedeterministická stavová zložitosť regulárneho jazyka je definovaná ako najmenší počet stavov v akomkoľvek nedeterministickom konečnostavovom automate (NFA) potrebný na rozpoznanie daného jazyka. Stavová zložitosť (nedeterministická stavová zložitosť) operácie na regulárnych jazykoch reprezentovaných deterministickými (resp. nedeterministickými) konečnostavovými automatmi je počet stavov, ktoré sú nevyhnutné a potrebné v najhoršom prípade pre deterministický (nedeterministický) automat pre rozpoznanie jazyka, ktorý dostaneme operáciou na týchto jazykoch.

2 Prehľad o súčasnom stave problematiky, ktorá je predmetom dizertačnej práce

Niektoré dávnejšie výsledky stavovej zložitosti možno nájsť v [33, 34, 36]. Stavová zložitosť niektorých operácií, ako sú zjednotenie, prienik, zreťazenie a uzáver jazykov reprezentovaných neúplnými deterministickými konečnostavovými automatmi boli skúmané Maslovom [35]. Podobné výsledky pre úplné deterministické automaty získali Yu, Zhuang a Salomaa [44]. Túto prvú systematickú štúdiu o stavovej zložitosti operácií nad regulárnymi jazykmi nasledovalo niekoľko článkov, ktoré skúmali stavovú zložitosť operácií nad konečnými a unárnymi jazykmi [6, 39].

Nedeterministická stavová zložitosť operácií nad regulárnymi jazykmi bola študovaná Holzerom a Kutribom v [18]. Domaratzki [9] skúmal proporcionálne odoberanie, Campeanu *et al.* [7] študovali stavovú zložitosť shuffle operácie, a Salomaa *et al.* [42] sa zaoberali stavovou zložitosťou zrkadlového obrazu jazykov. Ďalšie výsledky na túto tému možno nájsť v [11, 15].

V našej práci využijeme výsledky týkajúce sa determinizácie nedeterministických automatov [40], stavovej zložitosti zreťazenia, zjednotenia a prieniku [35, 44], nedeterministickej zložitosti zjednotenia a prieniku [18], ako aj nedeterministickej zložitosti doplnku [5, 29].

V [22] Iwama, Kambayashi a Takaki formulovali otázku, či existuje minimálny n-stavový nedeterministický automat, ktorého ekvivalentný minimálny deterministický automat má α stavov pre všetky prirodzené čísla n a α také, že $n \leqslant \alpha \leqslant 2^n$. Túto otázku rozoberali tiež Iwama, Matsuura a Paterson v [23]. V týchto dvoch článkoch je dokázané, že ak $\alpha = 2^n - 2^k$ alebo $\alpha = 2^n - 2^k - 1$, kde $0 \leqslant k \leqslant n/2 - 2$, alebo ak $\alpha = 2^n - k$, kde $2 \leqslant k \leqslant 2n - 2$ a platí istá podmienka nesúdeliteľnosti, potom zodpovedajúce binárne n-stavové nedeterministické automaty, vyžadujúce α stavov v deterministickom konečnostavovom automate, existujú. V [26] boli príslušné nedeterministické automaty popísané pre všetky hodnoty n a α, avšak veľkosť vstupnej abecedy pre tieto automaty rastie exponenciálne v závislosti od n. Neskôr, v [12] bola veľkosť vstupnej abecedy pre nedeterministické automaty znížená na $n + 2$. Možné diery v hierarchii stavovej zložitosti sa v literatúre nazývajú magické čísla. Nedávno Geffert [13] dokázal, že v prípade unárnej abecedy existuje veľa takýchto magických čísel.

3 Ciele dizertačnej práce

Hlavné ciele tejto práce sú nasledujúce:

- študovať stavovú zložitosť zreťazenia dvoch jazykov, ktoré sú popísané deterministickými konečnostavovými automatmi;
- porovnať nedeterministickú stavovú zložitosť regulárneho jazyka a jeho doplnku;
- skúmať magické čísla pre determinizáciu nedeterministického automatu s konštantnou vstupnou abecedou;
- študovať magické čísla pre zjednotenie a prienik v deterministickom aj nedeterministickom prípade.

4 Metódy použité v dizertačnej práci

V tejto dizertačnej práci sú použité všeobecné metódy vedy ako indukcia, porovnávanie, dedukcia a sumarizácia. Zvyčajne sa zaoberáme odhadom hornej a dolnej hranice stavovej zložitosti. Cieľom je ukázať, že sa obidve hranice zhodujú.

K získaniu hornej hranice používame konštruktívne metódy: popíšeme konštrukciu vhodného zariadenia - deterministického alebo nedeterministického automatu, veľkosť ktorého nie je väčšia než je horná hranica.

Získanie dolnej hranice je zvyčajne oveľa zložitejší problém. Metódy, ktoré sa používajú, závisia na popise regulárnych jazykov.

Na preukázanie minimality deterministického automatu nám stačí ukázať dosiahnuteľnosť a neekvivalenciu jeho stavov. Dosiahnuteľnosť dokazujeme metódou matematickej indukcie a neekvivalenciu nájdením reťazcov, ktoré rozlišujú stavy.

Na preukázanie minimality nedeterministického automatu pre nejaký regulárny jazyk používame špeciálnu metódu tzv. klamúcej množiny: Popisujeme množinu dvojíc reťazcov tak, aby zreťazenie reťazcov v každom páre bolo v danom jazyku, zatiaľ čo takéto zreťazenie pre dva rôzne páry nie je v tomto jazyku. Veľkosť takejto klamúcej množiny potom poskytuje dolnú hranicu pre počet stavov v akomkoľvek nedeterministickom automate pre daný jazyk.

Ak chceme získať čo najlepšie dolné hranice, musíme nájsť vhodný príklad, zložitosť ktorého je čo najvyššia. Na začiatku to vedie k nejakým experimentom. Zostrojíme nie príliš veľké príklady pomocou metódy pokusov a omylov a overíme požadované vlastnosti prípadným využitím špeciálneho softvéru.

5 Hlavné výsledky dizertačnej práce

V kapitole 3 študujeme stavovú zložitosť zreťazenia dvoch regulárnych jazykov reprezentovaných m-stavovým a n-stavovým deterministickým konečnostavovým automatom, pričom prvý automat má k koncových stavov. Bolo známe [44], že horná hranica pre túto zložitosť je $m2^n - k2^{n-1}$. Zároveň v tomto článku autori ukázali, že táto hranica je dosiahnuteľná pre trojpísmenkovú abecedu a $k = 1$. V práci dokazujeme dosiahnuteľnosť tejto hranice na dvojpísmenkovej abecede a pre všetky k medzi 1 a $m - 1$.

Veta 5.1 *Pre všetky prirodzené čísla m, n, k, pre ktoré platí $m \geq 2, n \geq 2$ a $0 < k < m$, existuje binárny m-stavový DFA A s k koncovými stavmi, a binárny n-stavový DFA B taký, že každý DFA akceptujúci jazyk $L(A)L(B)$ potrebuje aspoň $m2^n - k2^{n-1}$ stavov.*

Kapitola 4.1 sa zaoberá problémom magických čísel pre determinizáciu nedeterministických automatov. Číslo α medzi n a 2^n sa nazýva magické ak neexistuje minimálny n-stavový nedeterministický automat, ktorého ekvivalentný minimálny deterministický automat má α stavov. Geffert [12, 13] ukázal, že na unárnej abecede existuje mnoho magických čísel, zatiaľ čo pre lineárnu abecedu veľkosti $n + 2$ neexistuje žiadne magické číslo. Najdôležitejším výsledkom našej práce je zníženie tejto lineárnej abecedy na konštanú štvorpísmenkovú abecedu.

Veta 5.2 *Pre všetky kladné prirodzené čísla n a α také, že $n \leqslant \alpha \leqslant 2^n$, existuje minimálny n-stavový nedeterministický konečnostavový automat so štvorpísmenkovou vstupnou abecedou, ktorého ekvivalentný minimálny deterministický konečnostavový automat má α stavov.*

Podobne v kapitole 4.2 hľadáme magické čísla pre zjednotenie a prienik dvoch regulárnych jazykov reprezentovaných deterministickými alebo nedeterministickými automatmi. Horná hranica pre tieto dve operácie je mn [35, 44]. Na druhej strane, prienik dvoch disjunkných jazykov je prázdna množina akceptovaná jednostavovým automatom. Prvý výsledok v tejto časti sa týka prieniku jazykov reprezentovaných deterministickými automatmi a ukazuje neexistenciu magických čísel v prípade, že automaty majú aspoň dva stavy.

Veta 5.3 *Pre všetky kladné prirodzené čísla* m, n *a* α *také, že* $m \geqslant 2$, $n \geqslant 2$ *a* $1 \leqslant \alpha \leqslant mn$, *existuje minimálny* m*-stavový binárny DFA* A *a minimálny* n*-stavový binárny DFA* B *také, že minimálny DFA pre jazyk* $L(A) \cap L(B)$ *má* α *stavov.*

Podobný výsledok pre zjednotenie dostaneme využitím De Morganovho pravidla a faktu, že jazyk a jeho doplnok majú rovnakú deterministickú stavovú zložitosť.

Veta 5.4 *Pre všetky kladné prirodzené čísla* m, n *a* α *také, že* $m \geqslant 2$, $n \geqslant 2$ *a* $1 \leqslant \alpha \leqslant mn$, *existuje minimálny* m*-stavový binárny DFA* A *a minimálny* n*-stavový binárny DFA* B *také, že minimálny DFA pre jazyk* $L(A) \cup L(B)$ *má* α *stavov.*

V nedeterministickom prípade je horná hranica pre zjednotenie $m + n + 1$ [18]. Ďalší výsledok ukazuje, že ani v tomto prípade magické čísla neexistujú už na binárnej abecede, s výnimkou ak $m = 1$ a $n = 1$, kedy 2 je magické číslo.

Veta 5.5 *Pre všetky kladné prirodzené čísla* m, n, α *také, že* $m \geqslant 2$ *alebo* $n \geqslant 2$, *a* $1 \leqslant \alpha \leqslant m+n+1$, *existuje minimálny* m*-stavový binárny NFA* A *a minimálny* n*-stavový binárny NFA* B *také, že každý minimálny NFA pre jazyk* $L(A) \cup L(B)$ *má* α *stavov.*

V prípade prieniku dvoch jazykov reprezentovaných nedeterministickými automatmi je horná hranica mn [18]. Na dôkaz neexistencie magických čísel používa nasledujúca veta trojpísmenkovú abecedu. Problém ostáva otvorený pre dvojpísmenkovú abecedu.

Veta 5.6 *Pre všetky kladné prirodzené čísla* m, n, α *také, že* $1 \leqslant \alpha \leqslant mn$, *existuje minimálny* m*-stavový NFA* A *a minimálny* n*-stavový NFA* B *také, že každý minimálny NFA pre jazyk* $L(A) \cap L(B)$ *má* α *stavov.*

Kapitola 4.3 skúma magické čísla pre operáciu doplnku regulárneho jazyka reprezentovaného n-stavovým nedeterministickým automatom; v deterministickom prípade je tento problém triviálny, pretože deterministická stavová zložitosť jazyka a jeho doplnku je taká istá. Determinizáciou nedeterministického automatu dostaneme ekvivalentný deterministický automat, ktorý má najviac 2^n stavov. Výmenou koncových a nekoncových stavov dostaneme rovnako veľký, a teda najviac 2^n-stavový, dokonca deterministický, automat pre doplnok pôvodného jazyka. Preto je horná hranica v tomto prípade 2^n. A keďže doplnok doplnku je pôvodný jazyk, dolná hranica je $\log n$. Náš posledný výsledok dokazuje neexistenciu magických čísel pre nedeterministický doplnok na $2n$-písmenkovej abecede.

Veta 5.7 *Pre všetky kladné prirodzené čísla* n *a* α *také, že* $\log n \leq \alpha \leq 2^n$, *existuje* n*-stavový minimálny NFA* M *s* $2n$*-písmenkovou vstupnou abecedou taký, že každý minimálny NFA pre doplnok jazyka* $L(M)$ *má presne* α *stavov.*

6 Zoznam prác uchádzača, ktoré majú vzťah k skúmanej problematike

1. Jirásek, J., Jirásková, G., Szabari, A.: State complexity of concatenation and complementation. *International Journal of Foundations of Computer Science* **16** (2005) 511–529. (ADCA)

   Preliminary version in: *Proceedings of the 9th International Conference on Implementation and Application of Automata (CIAA 2004)*, Domaratzki, M., Okhotin, A., Salomaa, K., Yu, S. (eds), Lecture Notes in Computer Science, vol. 3317, Springer, Heidelberg, 2005, pp. 178–189. (AEC)

2. Hricko M., Jirásková G., Szabari A.: Union and intersection of regular languages and descriptional complexity. In: *Proceedings of the 7th International Workshop on Descriptional Complexity of Formal Systems (DCFS 2005)*, Mereghetti, C., Palano, B., Pighizzini, G., Wotschke, D. (eds), University of Milano, Milano, Italy, 2005, pp. 170–181. (AEC)

3. Jirásková G., Szabari A.: On the nondeterministic state complexity of complements of regular languages. In: *Proceedings of the 6th Workshop on Information Technologies - Applications and Theory (ITAT 2006)*, P. Vojtáš (ed), PONT, Seňa, Slovakia, 2006, pp. 63–68. (AED)

4. Jirásek J., Jirásková G., Szabari A.: Deterministic blow-ups of minimal nondeterministic finite automata over a fixed alphabet. *International Journal of Foundations of Computer Science* **19** (2008) 617–632. (ADCA)

   Preliminary version in: *Proceedings of the 11th International Conference on Developments in Language Theory (DLT 2007)*, Harju, T., Karhumäki, J., Lepistö, A. (eds), Lecture Notes in Computer Science, vol. 4588, Springer, Heidelberg, 2007, pp. 254–265. (AEC)

7 Doteraz zistené ohlasy na práce uchádzača

Sci citácie (8):

[1] citované v:

1. Han, Y.-S., Salomaa, K.: State complexity of union and intersection of finite languages. *International Journal of Foundations of Computer Science* **19** (2008) 581–595.
2. Gao, Y., Salomaa, A., Salomaa, K., Yu, S.: The state complexity of two combined operations: Star of catenation and star of reversal. *Fundamenta Informaticae* **83** (2008) 75–89.
3. Salomaa, K., Yu, S.: On the state complexity of combined operations and their estimation. *International Journal of Foundations of Computer Science* **18** (2007) 683–698.
4. Salomaa, A., Salomaa, K., Yu, S.: State complexity of combined operations. *Theoretical Computer Science* **383** (2007) 140–152.
5. Geffert, V., Mereghetti, C., Pighizzini, G.: Complementing two-way finite automata. *Information and Computation* **205** (2007) 1173–1187.
6. Rampersad, N.: The state complexity of L2 and Lk. *Information Processing Letters* **98** (2006) 231–234.
7. Yu, S.: On the state complexity of combined operations. In: *Proceedings of the 11th International Conference on Implementation and Application of Automata (CIAA 2006)*, Ibarra, O. H., Yen, H. C. (eds), Lecture Notes in Computer Science, vol. 4094, Springer, Heidelberg, 2006, pp. 11–22.

[2] citované v:

1. Han, Y.-S., Salomaa, K.: State complexity of union and intersection of finite languages. *International Journal of Foundations of Computer Science* **19** (2008) 581–595.

Citácie podľa iných indexov a databáz (5):

[1] citované v:

1. Esik, Z. Gao, Y., Liu, G., Yu, S.: Estimation of State Complexity of Combined Operations. In: *Proceedings of the 10th International Workshop on Descriptional Complexity of Formal Systems (DCFS 2008)*, to appear.
2. Gao, Y., Salomaa, K., Yu, S.: State complexity of catenation and reversal combined with star. In: *Proceedings of the 8th International Workshop on Descriptional Complexity of Formal Systems (DCFS 2006)*, Leung, H., Pighizzini, G. (eds), New Mexico State University, Las Cruces, 2006, pp. 153–164.
3. Kari, L., Konstantinidis, S., Sosík, P., Thierrin, G.: On hairpin-free words and languages. In: *Proceedings of the 9th International Conference on Developments in Language Theory (DLT 2005),* De Felice C., Restivo, A. (eds), Lecture Notes in Computer Science, vol. 3572, Springer, Heidelberg, 2005, pp. 296–307.
4. Yu, S.: State complexity: Recent results and open problems. *Fundamenta Informaticae* **64** (2005) 471–480.

[4] citované v:

1. Matsuura, A., Saito, Y.: Equivalent Transformation of Minimal Finite Automata over a Two-Letter Alphabet. In: *Proceedings of the 10th International Workshop on Descriptional Complexity of Formal Systems (DCFS 2008)*, Câmpeanu, C., Pighizzini, G. (eds), University of Prince Edward Island, Charlottetown, Canada, 2008, pp. 224–232.

Literatúra

[1] H. N. Adorna, 3-Party message complexity is better than 2-party ones for proving lower bounds on the size of minimal nondeterministic finite automata. *J. Autom. Lang. Comb.* **7** (2002) 419-432.

[2] A. V. Aho, J. D. Ullman, and M. Yannakakis, On notions of information transfer in VLSI circuits. In: *Proc. 15th Annual ACM Symp. on Theory of Computing (STOC),* 1983, pp. 133–139.

[3] P. Berman, A. Lingas, *On the complexity of regular languages in terms of finite automata,* Technical Report 304, Polish Academy of Sciences, 1977.

[4] J.C. Birget, Intersection and union of regular languages and state complexity, *Inform. Process. Lett.* **43** (1992) 185–190.

[5] J.C. Birget, Partial orders on words, minimal elements of regular languages, and state complexity, *Theoret. Comput. Sci.* **119** (1993) 267–291. ERRATUM: Partial orders on words, minimal elements of regular languages, and state complexity, 2002. Available at `http://clam.rutgers.edu/∼birget/papers.html`.

[6] C. Câmpeanu, K. Culik II, K. Salomaa, S. Yu, State complexity of basic operations on finite languages, in: O. Boldt, H. Jürgensen (Eds.), *Proc. 4th International Workshop on Implementing Automata (WIA'99),* LNCS 2214, Springer-Verlag, Heidelberg, 2001, pp. 60–70.

[7] C. Câmpeanu, K. Salomaa, S. Yu, Tight lower bound for the state complexity of shuffle of regular languages, *J. Autom. Lang. Comb.* **7** (2002) 303–310.

[8] M. Chrobak, Finite automata and unary languages. *Theoret. Comput. Sci.* **47** (1986) 149–158. ERRATUM: *Theoret. Comput. Sci.* **302** (2003) 497–498.

[9] M. Domaratzki, State complexity and proportional removals, *J. Autom. Lang. Comb.* **7** (2002) 455–468.

[10] K. Ellul, *Descriptional complexity measures of regular languages,* Master's thesis, University of Waterloo, 2002.

[11] K. Ellul, B. Krawetz, J. Shallit, and M. W. Wang, Regular expressions: new results and open problems, *J. Autom. Lang. Comb.* 10 (2005) 407–437.

[12] V. Geffert, (Non)determinism and the size of one-way finite automata. In: C. Mereghetti, B. Palano, G. Pighizzini, D. Wotschke (Eds.), *Proc. 7th Workshop on Descriptional Complexity of Formal Systems (DCFS 2005),* University of Milano, Italy, 2005, pp. 23–37.

[13] V. Geffert, Magic numbers in the state hierarchy of finite automata. In: R. Královič, P. Urzyczyn (Eds.) *Proc. 31st International Symposium on Mathematical Foundations of Computer Science (MFCS 2006),* Lecture Notes in Comput. Sci., 4162, Springer, Berlin, 2006, pp. 412–423.

[14] I. Glaister, J. Shallit, A lower bound technique for the size of nondeterministic finite automata, *Inform. Process. Lett.* **59** (1996) 75–77.

[15] M. Holzer, K. Salomaa, and S. Yu, On the state complexity of k-entry deterministic finite automata, *J. Autom. Lang. Comb.* **6** (2001) 453–466.

[16] M. Holzer, M. Kutrib, State complexity of basic operations on nondeterministic finite automata, in: J.M. Champarnaud, D. Maurel (Eds.), *Implementation and Application of Automata (CIAA 2002),* LNCS 2608, Springer-Verlag, Heidelberg, 2003, pp. 148–157.

[17] M. Holzer, M. Kutrib, Unary language operations and their nondeterministic state complexity, in: M. Ito, M. Toyama (Eds.), *Developments in Language Theory (DLT 2002),* LNCS 2450, Springer-Verlag, Heidelberg, 2003, pp. 162–172.

[18] M. Holzer, M. Kutrib, Nondeterministic descriptional complexity of regular languages, *Internat. J. Found. Comput. Sci.* **14** (2003) 1087-1102.

[19] J. Hromkovič, *Communication Complexity and Parallel Computing,* Springer-Verlag, Berlin, Heidelberg, 1997.

[20] J. Hromkovič, Descriptional complexity of finite automata: concepts and open problems, *J. Autom. Lang. Comb.* **7** (2002) 519–531.

[21] J. Hromkovič, S. Seibert, J. Karhumäki, H. Klauck, and G. Schnitger, Communication complexity method for measuring nondeterminism in finite automata. *Inform. and Comput.* **172** (2002) 202–217.

[22] K. Iwama, Y. Kambayashi and K. Takaki, Tight bounds on the number of states of DFAs that are equivalent to n-state NFAs. *Theoret. Comput. Sci.* **237** (2000) 485–494.

[23] K. Iwama, A. Matsuura and M. Paterson, A family of NFAs which need $2^n - \alpha$ deterministic states, *Theoret. Comput. Sci.* **301** (2003) 451–462.

[24] J. Jirásek, G. Jirásková, A. Szabari, State complexity of concatenation and complementation, *Internat. J. Found. Comput. Sci.* **16** (2005) 511–529.

[25] G. Jirásková, Note on minimal finite automata, in: *Proc. MFCS 2001,* Lecture Notes in Comput. Sci., 2136, Springer, Berlin, 2001, pp. 421–431.

[26] G. Jirásková, Note on minimal finite automata. In: J. Sgall, A. Pultr, P. Kolman (Eds.), *Proc. 26th International Symposium on Mathematical Foundations of Computer Science (MFSC 2001),* Lecture Notes in Comput. Sci., 2136, Springer, Berlin, 2001, pp. 421–431.

[27] G. Jirásková, Note on minimal automata and uniform communication protocols, in: C. Martin-Vide and V. Mitrana (Eds.), *Grammars and Automata for String Processing: From Mathematics and Computer Science to Biology, and Back,* Taylor and Francis, London, 2003, pp. 163–170.

[28] G. Jirásková, State complexity of some operations on regular languages, in: E. Csuhaj-Varjú, C. Kintala, D. Wotschke, Gy. Vaszil (Eds.), *Proc. 5th Workshop Descriptional Complexity of Formal Systems,* MTA SZTAKI, Budapest, 2003, pp. 114–125.

[29] G. Jirásková, State complexity of some operations on binary regular languages, *Theoret. Comput. Sci.* **330** (2005) 287–298.

[30] G. Jirásková: Magic numbers and ternary alphabet. In: Proceedings of the 13th International Conference on Developments in Language Theory (DLT 2009, Stuttgart, Germany, June 30 - July 3), Diekert, V., Nowotka, D. (eds.), Lecture Notes in Computer Science 5583, Springer, Heidelberg, 2009, pp. 300-311.

[31] A. Kapoutsis, Ch.A.: Size Complexity of Two-Way Finite Automata. Developments in Language Theory, 47-66, 2009.

[32] E. Leiss, Succinct representation of regular languages by boolean automata, *Theoret. Comput. Sci.* **13** (1981) 323–330.

[33] O. B. Lupanov, A comparison of two types of finite automata, *Problemy Kibernetiki,* **9** (1963) 321-326 (in Russian).

[34] Yu. I. Lyubich, Estimates for optimal determinization of nondeterministic autonomous automata, *Sibirskii Matematichskii Zhurnal,* **5** (1964) 337-355 (in Russian).

[35] A. N. Maslov, Estimates of the number of states of finite automata, *Dokl. Akad. Nauk SSSR,* **194** (1970) 1266–1268 (in Russian). English translation: *Soviet Math. Dokl.* **11** (1970) 1373–1375.

[36] F. R. Moore, On the bounds for state-set size in the proofs of equivalence between deterministic, nondeterministic, and two-way finite automata, *IEEE Trans. Comput.* **20** (1971) 1211–1214.

[37] A.R. Meyer and M.J. Fischer, Economy of description by automata, grammars and formal systems, in: *Proc. 12th Annual Symposium on Switching and Automata Theory,* 1971, pp. 188–191.

[38] G. Pighizzini, Unary language concatenation and its state complexity, in: S. Yu, A. Pun (Eds.), *Implementation and Application of Automata: 5th International Conference, CIAA 2000,* LNCS 2088, Springer-Verlag, 2001, pp. 252–262.

[39] G. Pighizzini, J. Shallit, Unary language operations, state complexity and Jacobsthal's function, *Internat. J. Found. Comput. Sci.* **13** (2002) 145–159.

[40] M. Rabin, D. Scott, Finite automata and their decision problems, *IBM Res. Develop.* **3** (1959) 114–129.

[41] W.J. Sakoda, M. Sipser, Nondeterminism and the size of two-way finite automata, in: *Proc. 10th Annual ACM Symposium on Theory of Computing,* 1978, pp. 275–286.

[42] A. Salomaa, D. Wood, and S. Yu, On the state complexity of reversals of regular languages, *Theoret. Comput. Sci.* **320** (2004) 315–329.

[43] M. Sipser, *Introduction to the theory of computation,* PWS Publishing Company, Boston, 1997.

[44] S. Yu, Q. Zhuang, K. Salomaa, The state complexity of some basic operations on regular languages, *Theoret. Comput. Sci.* **125** (1994) 315–328.

[45] S. Yu, Chapter 2: Regular languages, in: G. Rozenberg, A. Salomaa, (Eds.), *Handbook of Formal Languages - Vol. I,* Springer-Verlag, Berlin, New York, pp. 41–110.

[46] S. Yu, A renaissance of automata theory? *Bull. Eur. Assoc. Theor. Comput. Sci. EATCS* **72** (2000) 270–272.

[47] S. Yu, State complexity of finite and infinite regular languages, *Bull. Eur. Assoc. Theor. Comput. Sci. EATCS* **76** (2000) 270–272.

[48] S. Yu, State complexity of regular languages, *J. Autom. Lang. Comb.* **6** (2001) 221–234.

[49] Zijl, L.: Magic numbers for symmetric difference NFAs. Internat. J. Found. Comput. Sci. 16, 1027–1038 (2005)

Summary

This thesis presents several results on the descriptional complexity of regular languages. We study the deterministic and nondeterministic state complexity of languages that are defined as the number of states in the minimal deterministic or a minimal nondeterministic finite automaton for the given language. By the state complexity of an operation on regular languages we mean the number of states that are sufficient and necessary in the worst case to accept the language resulting from the operation, taken as a function of the complexities of operands.

Our first result shows that the upper bounds on the state complexity of concatenation, that depend on the number of final states in the first automaton, are tight for an arbitrary number of the final states.

The second part of the thesis is devoted to the magic numbers problem. Here we are interested not only in the complexity in the worst case, but we also study all values that can be obtained as the complexity of some operations; the values that cannot be reached in this way are called magic numbers. In particular, we examine magic numbers for the NFA to DFA conversion, union and intersection, and complementation. In all three cases, we show that no magic numbers exist.