Warunek jednorazowości dla automatów i transduktorów rejestrowych nad nieskończonymi alfabetami

Streszczenie rozprawy doktorskiej

Autor: Rafał Stefański

Promotor: prof. dr hab. Mikołaj Bojańczyk

Czerwiec 2022

Rozprawa ma na celu zbadanie skutków wprowadzenia warunku *jednorazowości* do definicji automatów i transduktorów rejestrowych nad nieskończonymi alfabetami. Jej celem jest również uporządkowanie i rozszerzenie wyników opublikowanych w [5].

Automaty rejestrowe zostały po raz pierwszy zdefiniowane i zbadane w [7], jako model który posiada odpowiednio dużo siły wyrazu by móc analizować słowa nad nieskończonymi alfabetami, lecz mimo to daje się analizować metodami teorii automatów (m.in. problem ich pustości jest rozstrzygalny). Już w [7] okazało się, że siła wyrazu automatów rejestrowych jest bardzo niestabilna – deterministyczne jednokierunkowe, deterministyczne dwukierunkowe oraz niedeterministyczne jednokierunkowe warianty automatów rejestrowych definiują parami nierównoważne klasy języków (w przypadku automatów skończonych wszystkie powyższe warianty definiują klasę języków regularnych). Dalsze badania (np. [11]) potwierdziły tę niestabilność – w zasadzie wszystkie warianty automatów rejestrowych są parami nierównoważne, choć ich skończone warianty są równoważne językom regularnym.

Niestabilność ta nie zmienia faktu, że automaty rejestrowe (w wielu swoich wariantach) są modelem o dużym znaczeniu zarówno praktycznym jak i teoretycznym, w wyniku czego zostały dobrze przebadane. Jeden z kierunków tych badań to użycie *zbiorów z atomami* (ang. *nominal sets* lub *sets with atoms*) do ich opisu i analizy. Podejście to zostało zapoczątkowane w pracach [2] oraz [9]. Praca [2] definiuje klasę języków rozpoznawanych przez *monoidy skończenie orbitowe* (o sile wyrazu słabszej od deterministycznych automatów rejestrowych) oraz bada jej związki z logikami $\mathrm{MSO}^{\sim}$ oraz $\mathrm{FO}^{\sim}$. Badania te były kontynuowane w [6], która definiuje *jednoznacznie strzeżony (ang. rigidly-guarded)* fragment logiki $\mathrm{MSO}^\sim$ i pokazuje, że pod względem siły wyrazu jest on równoważny monoidom skończenie orbitowym (co w kontekście automatów rejestrowych jest dosyć nietypowym wynikiem). Prace [13] (moja praca magisterska) oraz [5] kontynuują badania nad klasą języków rozpoznawaną przez monoidy skończenie orbitowe. Definiują one *warunek jednorazowości (ang. single-use restriction)* dla automatów rejestrowych (który wymaga żeby każdy odczyt wartości rejestru skutkował jej wymazaniem) i pokazują że *automaty z jednorazowymi rejestrami* definiują tę samą klasę języków co monoidy skończenie-orbitowe. Oznacza to, że ta sama klasa języków ma trzy istotnie różne definicje:

- algebraiczną – poprzez monoidy skończenie orbitowe;
- logiczną – poprzez jednoznacznie strzeżony fragment $\mathrm{MSO}^\sim$;
- maszynową – poprzez deterministyczne automaty z jednorazowymi rejestrami.

(Ostatnio trwają też prace na temat definicji topologicznej [14]). Jest to ślina przesłanka za tym, że klasa ta może mieć duże znaczenie teoretyczne, co uzasadnia potrzebę jej dokładniejszego zbadania.

Struktura rozprawy

Rozprawa składa się z czterech rozdziałów. Pierwsze dwa zajmują się automatami, drugie dwa transduktorami.

1. Alfabety nieskończone

Pierwszy rozdział prezentuje istotne dla rozprawy fragmenty dotychczasowego stanu wiedzy na temat języków nad nieskończonymi alfabetami. W szczególności zawiera on definicje automatów rejestrowych (w wielu wariantach) i przybliża metodologię zbiorów z atomami. Oparty jest głownie na książkach [3] oraz [12].

2. Warunek jednorazowości

Drugi rozdział prezentuje definicję warunku jednorazowości dla automatów rejestrowych. Następnie wprowadza klasę *jednorazowych funkcji* (jako uogólnienie warunku jednorazowości) i bada ich własności. Na końcu dowodzi, że pod względem siły wyrazu jednorazowe automaty rejestrowe nie są silniejsze od monoidów skończenie orbitowych (delegując dowód przeciwnej inkluzji do następnego rozdziału). Rozdział oparty jest głównie na [5] oraz [13], ale prezentuje też nowe wyniki.

3. Jednorazowe maszyny Mealy'a oraz ich dekompozycje w stylu Krohna-Rhodesa

Trzeci rozdział zaczyna się od definicji *maszyn Mealy'a z jednorazowymi rejestrami* (uogólnienie definicji z [10]). Następnie prezentuje on treść oryginalnego twierdzenia Krohna-Rhodesa (udowodnionego w [8]), uogólnia ją do nieskończonych alfabetów i prezentuje dowód tej uogólnionej wersji. Przy okazji wprowadza pojęcie *lokalnej monoidalnej transdukcji* oraz dowodzi brakującej inkluzji z drugiego rozdziału. Ten rozdział opiera jest w dużej mierze na [5], ale prezentuje też nowe wyniki.

4. Dwukierunkowe transduktory z atomami

Czwarty rozdział definiuje następujące modele transduktorów:

- dwukierunkowe transduktory z jednorazowymi rejestrami;
- jednorazowe i niekopiujące transduktory SST (model opisany w [1] z dodatkowo zastosowanym warunkiem jednorazowości);
- regularne funkcje na listach nad nieskończonymi alfabetami (rozszerzenie modelu z [4]);
- dwukierunkowe dekompozycje Krohna-Rhodesa nad nieskończonymi alfabetami (rozszerzenie modelu z [8]).

A następnie dowodzi, że wszystkie te modele są parami równoważne. Rozdział opiera się na [5].

Bibliografia

[1] Rajeev Alur and Pavol Černý. "Streaming transducers for algorithmic verification of single-pass list-processing programs". In: *Proceedings of the 38th annual ACM SIGPLAN-SIGACT symposium on Principles of programming languages*. 2011, pp. 599–610.

[2] Mikołaj Bojańczyk. "Nominal monoids". In: *Theory of Computing Systems* 53.2 (2013), pp. 194–222.

[3] Mikołaj Bojańczyk. *Slightly infinite sets*. 2019. URL: `https://www.mimuw.edu.pl/~bojan/paper/atom-book`.

[4] Mikołaj Bojańczyk, Laure Daviaud, and Shankara Narayanan Krishna. "Regular and first-order list functions". In: *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science*. 2018, pp. 125–134.

[5] Mikołaj Bojańczyk and Rafał Stefański. *Single use register automata for data words*. 2020. arXiv: `1907.10504 [cs.FL]`.

[6] Thomas Colcombet, Clemens Ley, and Gabriele Puppis. “Logics with rigidly guarded data tests”. In: *Log. Methods Comput. Sci.* 11.3 (2015). DOI: 10.2168/LMCS-11(3:10)2015. URL: https://doi.org/10.2168/LMCS-11(3:10)2015.

[7] Michael Kaminski and Nissim Francez. “Finite-memory automata”. In: *Theoretical Computer Science* 134.2 (1994), pp. 329–363. ISSN: 0304-3975. DOI: https://doi.org/10.1016/0304-3975(94)90242-9. URL: https://www.sciencedirect.com/science/article/pii/0304397594902429.

[8] Kenneth Krohn and John Rhodes. “Algebraic theory of machines. I. Prime decomposition theorem for finite semigroups and machines”. In: *Transactions of the American Mathematical Society* 116 (1965), pp. 450–464.

[9] Sławomir Lasota, Bartek Klin, and Mikołaj Bojańczyk. “Automata theory in nominal sets”. In: *Logical Methods in Computer Science* 10 (2014).

[10] George H Mealy. “A method for synthesizing sequential circuits”. In: *The Bell System Technical Journal* 34.5 (1955), pp. 1045–1079.

[11] Frank Neven, Thomas Schwentick, and Victor Vianu. “Finite state machines for strings over infinite alphabets”. In: *ACM Transactions on Computational Logic (TOCL)* 5.3 (2004), pp. 403–435.

[12] Andrew M. Pitts. *Nominal Sets: Names and Symmetry in Computer Science*. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 2013. DOI: 10.1017/CBO9781139084673.

[13] Rafał Stefański. “An automaton model for orbit-finite monoids”. MA thesis. University of Warsaw - Faculty of Mathematics, Informatics and Mechanics, 2018, p. 29.

[14] Henning Urbat. *Nominal Topology for Data Languages*. A Short Contribution at 16th IFIP WG 1.3 International Workshop on Coalgebraic Methods in Computer Science. Apr. 2022. URL: https://www.coalg.org/cmcs22/programme/.