Internship proposal

Title: Algebraic approach to register minimisation in streaming string transducers

Location: Team Modelisation and Verification, Laboratoire d'Informatique et des Systèmes, Luminy
`http://www.lis-lab.fr`

Supervisors:
Nathan Lhote and Pierre-Alain Reynier
Email: `{nathan.lhote,pierre-alain.reynier}@lis-lab.fr`

Context.

The theory of regular languages is one of the major pillars of computer science. It relies on multiple characterizations, by means of expressions, logic, automata and algebra. Algebraic characterizations come with a canonical object representing a language, allowing to decide important properties of the language, the most famous example being the definability in first-order logic (FO), a result due to [Sch65, MP71].

This rich setting has been studied a lot by the research community in order to lift it to more general ones. A natural extension of languages is that of transducers, which are automata extended with outputs, allowing to represent functions, or even relations, between finite words.

Positive results have been obtained for the class of so-called *rational functions*. These are the functions that can be obtained by a non-deterministic finite-state transducer, or, equivalently, by an unambiguous finite-state transducer. This class also admits a logical representation (order-preserving MSO transductions [FR16]), a presentation using regular expressions, and an algebraic characterization by means of bimachines [Sch61]. This last characterization has been used recently to lift the result of FO-definability to the setting of rational functions in [FGL16b, FGL16a].

Another transducer model, closer from an implementation model, has been introduced few years ago in [AČ10, AČ11]. It is based on a deterministic automaton equipped with variables allowing to store intermediate computation values. This model, known as streaming string transducers (SST), is very expressive, and different subclasses have been identified allowing to match the expressiveness of existing models. For instance, a model equivalent to rational functions has been exhibited.

A key property of finite-state automata is minimization: every regular language admits a canonical minimal automaton. Minimizing size is important for complexity issues, and canonicity of the model is also useful for deciding further properties, such as FO-definability. More generally, simplifying models is an important challenge.

Objectives.

The general objective of this internship is to investigate the use of bimachines for model simplification for the class of rational functions.

Numerous simplification problems can be addressed:

- the class of multi-sequential functions corresponds to functions that can be realized as a finite union of sequential functions (realized by a one-way deterministic transducer). A characterization of these functions exists, but is based on a pattern of the transducer [JF18]. Is it possible to decide this class directly on the bimachine of the function?
- a further question, related to multi-sequential functions, corresponds to minimize the size of the union. This has been addressed recently, again using a pattern, with links to register minimization of some class of SST in [DJRV17]. It would be interesting to have another approach directly based on bimachines.
- last, a more ambitious perspective consists in studying the register minimization problem for the class of SST corresponding to rational functions. This problem has been solved in [DRT16], again using a pattern on the transducer, characterizing whether the function can be expressed using at most k registers. Being able to express this property directly on the bimachine would be an important achievement.

References

[AČ10] Rajeev Alur and Pavol Černý. Expressiveness of streaming string transducers. In *Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, volume 8 of *LIPIcs*, pages 1–12. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2010.

[AČ11] Rajeev Alur and Pavol Černý. Streaming transducers for algorithmic verification of single-pass list-processing programs. In *Proc. of the 38th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, POPL 2011*, pages 599–610. ACM, 2011.

[DJRV17] Laure Daviaud, Ismael Jecker, Pierre-Alain Reynier, and Didier Villevalois. Degree of sequentiality of weighted automata. In *Proc. 20th International Conference on Foundations of Software Science and Computation Structures (FoSSaCS 2017)*, volume 10203 of *Lecture Notes in Computer Science*, pages 215–230. Springer, 2017.

[DRT16] Laure Daviaud, Pierre-Alain Reynier, and Jean-Marc Talbot. A generalised twinning property for minimisation of cost register automata. In Martin Grohe, Eric Koskinen, and Natarajan Shankar, editors, *Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in*

Computer Science, LICS '16, New York, NY, USA, July 5-8, 2016, pages 857–866. ACM, 2016.

[FGL16a] Emmanuel Filiot, Olivier Gauwin, and Nathan Lhote. Aperiodicity of rational functions is pspace-complete. In Akash Lal, S. Akshay, Saket Saurabh, and Sandeep Sen, editors, *36th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2016, December 13-15, 2016, Chennai, India*, volume 65 of *LIPIcs*, pages 13:1–13:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016.

[FGL16b] Emmanuel Filiot, Olivier Gauwin, and Nathan Lhote. First-order definability of rational transductions: An algebraic approach. In Martin Grohe, Eric Koskinen, and Natarajan Shankar, editors, *Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science, LICS '16, New York, NY, USA, July 5-8, 2016*, pages 387–396. ACM, 2016.

[FR16] Emmanuel Filiot and Pierre-Alain Reynier. Transducers, logic and algebra for functions of finite words. *SIGLOG News*, 3(3):4–19, 2016.

[JF18] Ismaël Jecker and Emmanuel Filiot. Multi-sequential word relations. *Int. J. Found. Comput. Sci.*, 29(2):271–296, 2018.

[MP71] Robert McNaughton and Seymour Pappert. *Counter-Free Automata*. MIT Press, 1971.

[Sch61] Marcel Paul Schützenberger. A remark on finite transducers. *Inf. Control.*, 4(2-3):185–196, 1961.

[Sch65] Marcel Paul Schützenberger. On finite monoids having only trivial subgroups. *Inf. Control.*, 8(2):190–194, 1965.