A Toolbox for Handling Infinite Sets of Graphs

Radu IOSIF (VERIMAG, Grenoble)
mailto:Radu.Iosif@univ-grenoble-alpes.fr

1 Background

Regular and context-free languages form a central pillar of the theory of formal languages. This foundational role is underpinned by the variety of well-established formalisms that define these classes of languages. For instance, regular word languages can be represented by finite automata, semigroups, regular expressions, regular grammars and monadic-second order logic. These formalisms provide a robust and interconnected framework for understanding the structure of regular word languages.

Graphs are important for many areas of computing such as static analysis [10], databases with knowledge representation [2] and concurrency [7]. While formal languages of words and ranked trees are understood, the development of formal languages for graphs still raises several open problems. One such problem is the lack of a notion of automaton defining a class of graph languages closed under boolean operations and having a logical characterization. A seminal development in this direction is the work of Courcelle that introduced graph algebras [5, 6]. In particular, the *hyperedge replacement* algebra generalizes the notion of recognizability from words to graphs.

The tree-width of a graph is a positive integer measuring the distance between the graph and a tree. For instance, trees have tree-width one, series-parallel graphs (i.e., circuits with one input and one output that can be either cascaded or overlaid) have tree-width two, whereas $n \times n$ square grids have tree-width n, for any $n \geq 1$. The tree-width is a cornerstone of algorithmic tractability. For instance, many NP-complete graph problems such as Hamiltonicity and 3-Coloring become PTIME, when restricted to inputs whose tree-width is uniformly bounded by a constant, see, e.g., [9, Chapter 11].

The problem of defining finite description of infinite sets of graphs of bounded tree-width has gained recent momentum. For instance, regular expressions that capture the recognizable sets of graphs of tree-width at most 2 [8] and regular grammars that capture the recognizable sets of graphs of bounded *embeddable* tree-width [4], i.e., an over-approximation of tree-width that considers only decompositions whose backbones are spanning trees of the given graph. Based on the recent development of regular grammars for sets of graphs having tree-width at most 2, we give an algorithm that decides the inclusion problem between the languages of such grammars in 2EXPTIME [3].

2 Goal and Challenges

The goal of this internship is an implementation of a toolbox that handles sets of graphs described as regular grammars. Ideally, such a toolbox should support the boolean operations of union, intersection and complement and decide the problems of membership and inclusion, by building finite recognizer algebras from the syntactic description of sets, as regular grammars.

The main challenge lies in the fact that such recognizer algebras are at least exponential in the size of the grammar (even for words). We aim at investigating optimization techniques, such as the use of antichain algorithms [1] and on-the-fly generation to tackle this complexity.

The internship consists of both theoretical and implementation work.

3 How to Apply

Send your CV and a letter of intent to: mailto:Radu.Iosif@univ-grenoble-alpes.fr

References

[1] P. A. Abdulla, Y. Chen, L. Holík, R. Mayr, and T. Vojnar. When simulation meets antichains. In J. Esparza and R. Majumdar, editors, *Tools and Algorithms for the Construction and Analysis of Systems, 16th International Conference, TACAS 2010, Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2010, Paphos, Cyprus, March 20-28, 2010. Proceedings*, volume 6015 of *Lecture Notes in Computer Science*, pages 158–174. Springer, 2010.

[2] S. Abiteboul, P. Buneman, and D. Suciu. *Data on the Web: From Relations to Semistructured Data and XML*. Morgan Kaufmann, 2000.

[3] M. Bozga, R. Iosif, and F. Zuleger. Regular grammars for graph sets of tree-width ≤ 2. *CoRR*, abs/2408.01226, 2024.

[4] M. Chimes, R. Iosif, and F. Zuleger. Tree-verifiable graph grammars. In N. S. Bjørner, M. Heule, and A. Voronkov, editors, *LPAR 2024: Proceedings of 25th Conference on Logic for Programming, Artificial Intelligence and Reasoning, Port Louis, Mauritius, May 26-31, 2024*, volume 100 of *EPiC Series in Computing*, pages 165–180. EasyChair, 2024.

[5] B. Courcelle. The monadic second-order logic of graphs. i. recognizable sets of finite graphs. *Information and Computation*, 85(1):12–75, 1990.

[6] B. Courcelle and J. Engelfriet. *Graph Structure and Monadic Second-Order Logic: A Language-Theoretic Approach*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2012.

[7] P. Degano, R. D. Nicola, and J. Meseguer, editors. *Concurrency, Graphs and Models, Essays Dedicated to Ugo Montanari on the Occasion of His 65th Birthday*, volume 5065 of *Lecture Notes in Computer Science*. Springer, 2008.

[8] A. Doumane. Regular expressions for tree-width 2 graphs. In M. Bojanczyk, E. Merelli, and D. P. Woodruff, editors, *49th International Colloquium on Automata, Languages, and Programming, ICALP 2022, July 4-8, 2022, Paris, France*, volume 229 of *LIPIcs*, pages 121:1–121:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

[9] J. Flum and M. Grohe. *Parameterized Complexity Theory*. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2006.

[10] N. D. Jones and S. S. Muchnick. A flexible approach to interprocedural data flow analysis and programs with recursive data structures. In *Proceedings of the 9th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, POPL '82, page 66–74, New York, NY, USA, 1982. Association for Computing Machinery.