Higher-dimensional subdiagram matching

Amar Hadzihasanovic
[1] Tallinn University of Technology
[2] Quantinuum, 17 Beaumont Street, Oxford OX1 2NA, UK
amar@ioc.ee

Diana Kessler
Tallinn University of Technology
diana-maria.kessler@taltech.ee

Joint work with Amar Hadzihasanovic

Abstract of arXiv:2304.09216. Accepted for LICS 2023.

Higher-dimensional rewriting [1, 2] is founded on a duality of rewrite systems and cell complexes, connecting computational mathematics to higher categories and homotopy theory: the two sides of a rewrite rule are two halves of the boundary of an $(n+1)$-cell, which are (pasting) diagrams of n-cells. We continue the work started in [4] studying higher-dimensional rewriting as a *mechanism of computation*, and try to answer the question: given a machine that operates by higher-dimensional rewriting, is the obvious cost model that attributes constant cost to each rewrite step a "reasonable" cost model?

The basic computational step of any such machine may be described as follows. The machine has a list $(r_i)_{i=1}^m$ of rewrite rules, which are $(n+1)$-cells, and whose input boundaries $(\partial^- r_i)_{i=1}^m$ and output boundaries $(\partial^+ r_i)_{i=1}^m$ are n-dimensional diagrams. Given an n-dimensional diagram t as input, the machine tries to match one of the input boundaries with a rewritable *subdiagram* of t. If it finds a match with $\partial^- r_i$ for some $i \in \{1,\dots,m\}$, it substitutes $\partial^+ r_i$ for the match in t; otherwise it stops.

Evidently, our question is answered in the affirmative for such a machine if and only if the *subdiagram matching problem* admits a (preferably low-degree) polynomial-time algorithm with respect to a reasonable size measure for diagrams. Since a cognate problem such as subgraph matching is notoriously NP-complete, it is not at all obvious that this should be true.

We study the higher-dimensional subdiagram matching problem in the "topologically sound" *diagrammatic sets* model [3]. Our main contribution is an algorithm for subdiagram matching in arbitrary dimension.

In the theory of diagrammatic sets, we represent the shape of a diagram with its *face poset* which records whether an $(n-1)$-dimensional cell is in the input or output half of the boundary of an n-dimensional cell. We call this structure an *oriented graded poset* and represent it as a Hasse diagram with oriented edges. The well-formed shapes of diagrams form an inductive subclass of oriented graded posets called *regular molecules*. A diagram, in our setting, is then a labelling $t\colon U \to \mathbb{V}$ of a regular molecule U into a set of variables.

Given diagrams $s\colon V \to \mathbb{V}$ and $t\colon U \to \mathbb{V}$, the subdiagram matching problem can be split into three subproblems:

1. find, if any, the inclusions of the shape of s into the shape of t;
2. decide if an inclusion is a *subdiagram* inclusion;
3. check that the labelling is preserved.

The subdiagram inclusions are those corresponding to portions of a diagram that can be rewritten producing another well-formed diagram; these are a proper subclass of all inclusions. The third problem is easy, so we focus on the first two.

Submitted to: ACT 2023

For a fixed regular molecule, U, we let U_n be the subset of n-dimensional elements in U, $|U|$ be the number of elements in U, $|U_n|$ the number of n-dimensional elements in U, $|\mathscr{E}_n U|$ the number of edges between n and $(n-1)$-dimensional elements in the Hasse diagram of U, and $|\mathscr{E}_\vee U|$ the maximum of the $|\mathscr{E}_n U|$ if non-zero, 1 otherwise.

Theorem 1 — *The problem of finding all inclusions in dimension n can be solved in time*

$$O(|U_n|\,|V_n|\,|V|\,|\mathscr{E}_\vee V|\log|\mathscr{E}_\vee V|).$$

The second problem turns out to be highly non-trivial. Our best general solution has the following worst-case time complexity upper bound.

Theorem 2 — *The problem of deciding if an inclusion is a subdiagram inclusion in dimension n can be solved in time*

$$O\left(\prod_{k\leq n}|U_k|!\,|U_k|\right).$$

The algorithm relies on new combinatorial results on *layerings* of diagrams, which are decompositions into layers containing each a single cell of the highest dimension. One can associate to each n-dimensional diagram a directed acyclic graph whose vertex set is U_n, such that each layering induces a distinct topological sort, which we call an *ordering*. The factorial factors come from the fact that our algorithm searches for a valid layering by trying out different orderings, which are as many as topological sorts of the graph, which are factorial in the number of vertices in the worst case of a discrete graph.

If we restrict our attention to a special class of diagrams that we call *stably-frame acyclic*, then there is a perfect correspondence between layerings and orderings, and we can avoid this iteration.

Theorem 3 — *If a diagram is guaranteed to be stably frame-acyclic, the problem of deciding if an inclusion is a subdiagram inclusion can be solved in linear time in the size of its Hasse diagram.*

Fortunately, the following holds.

Theorem 4 — *Every diagram of dimension ≤ 3 is stably frame-acyclic.*

We conclude that subdiagram matching is feasible up to dimension 3.

In the 4-dimensional case, not every shape of diagram is stably frame-acyclic and the main problem to overcome is the expensive iteration on orderings. The problem seems to arise from the possibility of rewrites causing non-local obstructions which may prevent other "disjoint" rewrites to be applied. This behaviour may hint to a potential gap in complexity between the 3-dimensional and 4-dimensional cases. We leave the existence of a polynomial time algorithm for subdiagram matching in dimension 4 or higher as an open problem and we hope that a deeper understanding of such cases will either lead to an improved algorithm or a proof of NP-completeness.

References

[1] A. Burroni (1993): *Higher-dimensional word problems with applications to equational logic*. *Theoretical Computer Science* 115(1), pp. 43–62, doi:10.1016/0304-3975(93)90054-w. Available at `https://doi.org/10.1016%2F0304-3975%2893%2990054-w`.

[2] Y. Guiraud (2019): Rewriting methods in higher algebra. Thèse d'habilitation à diriger des recherches, Université Paris 7.

[3] A. Hadzihasanovic (2020): *Diagrammatic sets and rewriting in weak higher categories*. *arXiv preprint arXiv:2007.14505*.

[4] A. Hadzihasanovic & D. Kessler (2022): *Data structures for topologically sound higher-dimensional diagram rewriting*. In: *Proceedings Fifth International Conference on Applied Category Theory (ACT2022)*.