HIGHER - DIMENSIONAL SUBDIAGRAM MATCHING

Amar Hadzihasanovic

TALTECH

6.7.2023

NCM - GETCO

HIGHER-DIMENSIONAL REWRITING :

A rewrite from l to r

is

an $(n+1)$-cell whose n-dim boundary is subdivided into an input and an output half

CELL COMPLEXES : HOMOTOPY TYPES

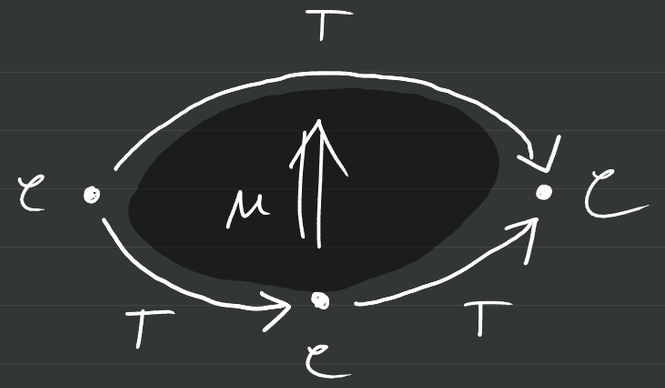
=

HIGHER REWRITE SYSTEMS : HIGHER CATEGORIES

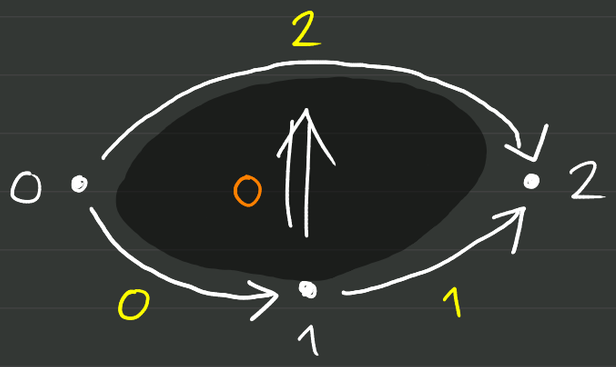
As there are many models of cell complexes (point-set, combinatorial, ...)

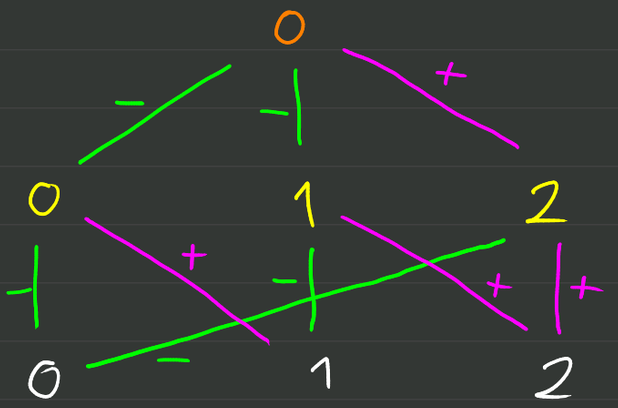
there can be many models of higher-dimensional rewrite systems.

DIAGRAM



$=$ SHAPE U





$+$ LABELLING $\ell : U \longrightarrow \mathcal{X}$

$0, 1, 2 \longmapsto e$

$0, 1, 2 \longmapsto T$ $\quad 0 \longmapsto \mu$

Def An oriented graded poset is a graded poset together with an edge-labelling of its Hasse diagram in $\{-, +\}$

For $x \in P$,

$$\Delta x := \{ y \in P \mid \begin{matrix} x \\ | \\ y \end{matrix} \text{ in } \mathcal{H}P \}$$ FACES

$$\Delta^- x := \{ y \in P \mid \begin{matrix} x \\ |- \\ y \end{matrix} \text{ in } \mathcal{H}P \}$$ INPUT FACES

$$\Delta^+ x := \{ y \in P \mid \begin{matrix} x \\ |+ \\ y \end{matrix} \text{ in } \mathcal{H}P \}$$ OUTPUT FACES

Basic terminology about o.g. posets:

① CLOSED SUBSET $U \subseteq P$:
lower set i.e. downwards closed subset

$\downarrow U$: downward closure of U

② DIMENSION $\dim x$: rank/grade of x

$$\dim U : \begin{cases} -1 & \text{if } U = \emptyset \\ \max\{\dim x \mid x \in U\} & \text{otherwise} \end{cases}$$

③ COFACES $\nabla^\alpha x$ in $P \equiv \Delta^\alpha x$ in P^{op}

④ $U_k := \{x \in U \mid \dim x = k\}$

$$\left(\alpha \in \{-, +\} \quad , \quad k \in \mathbb{N}\right)$$

Def $\Delta_k^{\alpha} U := \{ x \in U_k \mid \nabla^{-\alpha} x \cap U = \emptyset \}$

Def $(\mathrm{Max}\, U)_k := \Delta_k^{-} U \cap \Delta_k^{+} U$

i.e. elements of U_k that are maximal in U

Def

$$\partial_k^{\alpha} U := \mathrm{cl}\left(\Delta_k^{\alpha} U \cup \bigcup_{J<k} (\mathrm{Max}\, U)_J \right)$$

INPUT & OUTPUT k-BOUNDARY

WELL-FORMED SHAPES OF DIAGRAMS

≀

"REGULAR MOLECULES"

$\mathbb{R}$ inductive subclass of o.g. posets

① The point $\boxed{1}$ is in $\mathbb{R}$

② If $U, V \in \mathbb{R}$ and $\partial^+_k U \simeq \partial^-_k V$, $\boxed{U \#_k V}$, the pasting of U and V along the k-boundary, is in $\mathbb{R}$

③ If $U, V \in \mathbb{R}$, $\dim U = \dim V$, $\partial^\alpha U \simeq \partial^\alpha V$, and U, V are round, then $\boxed{U \Rightarrow V}$, the rewrite of U into V, is in $\mathbb{R}$

ISOMORPHISM CLASSES OF

REGULAR MOLECULES

+

PASTING

form a strict ω-category

In our model,

a diagram is

- a regular molecule U, with
- a labelling $\ell : U \longrightarrow \mathcal{X}$

in some set of names/variables.

(The diagram is a cell

if U has a greatest element.)

Our implementation:

rewalt

a Python library for higher-dim rewriting

Documentation: rewalt.readthedocs.io

Joint work with Diana Kessler:

Understanding algorithmic aspects of higher-dim rewriting in this model

① Data structures for topologically sound higher-dimensional diagram rewriting

ACT 2022, arXiv: 2209.09509

② Higher-dimensional subdiagram matching

LICS 2023, arXiv: 2304.09216

An abstract machine operating by n-dimensional diagram rewriting:

- input: an n-dim diagram
  $t : U \longrightarrow \mathcal{X}$
- The machine goes through a finite list $(r_i : V_i \longrightarrow \mathcal{X})_{i=1}^{m}$ of admissible rewriters, in the form of $(n+1)$-cells, and searches for matches of $\partial^- r_i = r_i|_{\partial^- V_i}$ with subdiagrams $s = t|_W$, $W \simeq \partial^- V_i$, $W \subseteq U$

- if it finds a match of $\partial^- r_i$ with $s = t/w$, it substitutes $\partial^+ r_i$ for s in T.

$W \sqsubseteq U$:

there exists a pasting decomposition of U in which W is a factor

Necessary for substitution to always produce a well-formed diagram!

If the resulting n-dimensional rewrite system is convergent (confluent + terminating),

the machine can be seen as computing a function from n-dim diagrams to n-dim diagrams.

Q: How does the derivational complexity of the rewrite system relate to the worst-case time complexity of an implementation of this abstract machine?

Depends on the computational complexity of subdiagram matching in dimension n!

SUBDIAGRAM MATCHING PROBLEM:

Given diagrams $T: U \longrightarrow \mathcal{X}$,

$S: V \longrightarrow \mathcal{X}$

s.t. $n = \dim U = \dim V$ and V is round,

Find, if any, all embeddings $V \overset{\iota}{\hookrightarrow} U$ s.t.

① $\iota(V) \subseteq U$

② $t \circ \iota = s$.

We can split into subproblems:

① Find all embeddings $V \overset{l}{\hookrightarrow} U$

(MOLECULE MATCHING PROBLEM)

② Given $V \overset{l}{\hookrightarrow} U$, decide if $l(V) \trianglelefteq U$

(REWRITABLE SUBMOLECULE PROBLEM)

③ Check if labellings match.

↖ EASY

In our ACT 2022 paper:

$O(N^2 \log N)$ algorithm for the molecule isomorphism problem

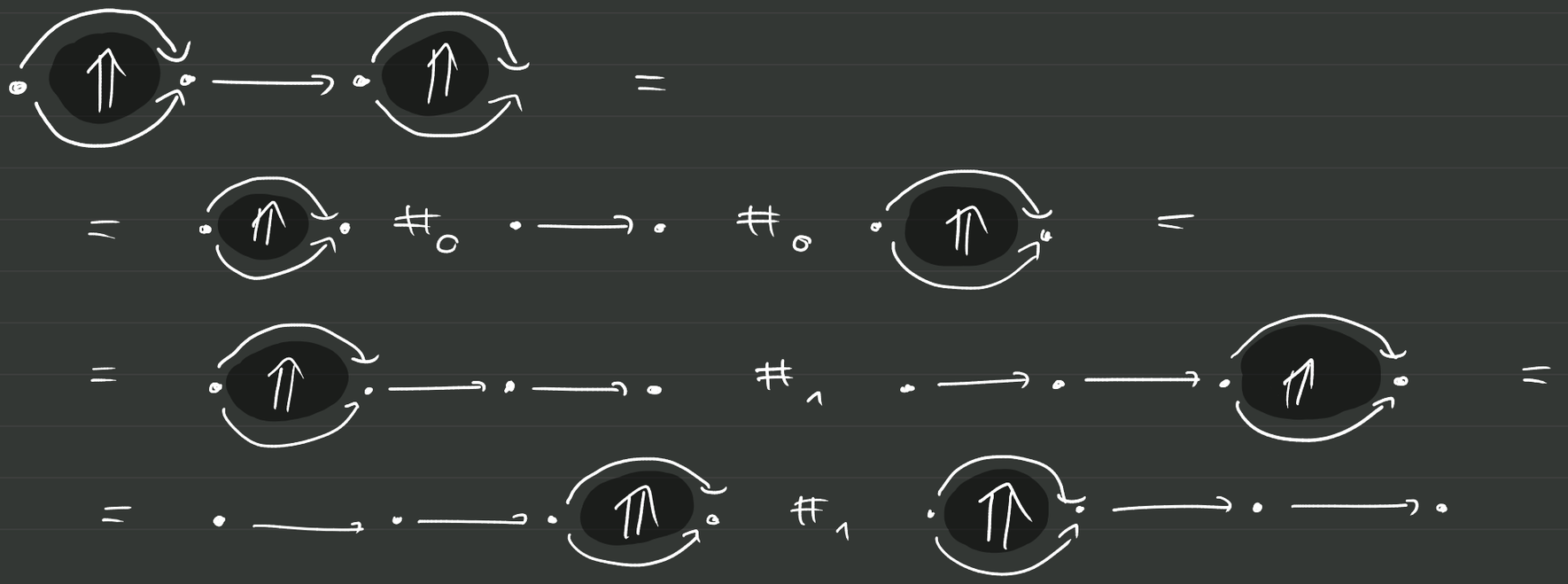
It turns out that molecule matching of V into U can be solved with $|U_h|$ calls to the molecule isomorphism algorithm for V.

OUR RESULTS ON THE REWRITABLE SUBMOLECULE PROBLEM

1. A decision algorithm valid in all dimensions, with a super-polynomial (factorial) upper bound;
2. A linear-time improvement for a class of molecules satisfying a condition we call stable frame-acyclicity, which includes all molecules up to dim 3

The difficulty:

Pasting decompositions are highly non-unique due to interchange!



Let U be a regular molecule.

For $k \geq -1$, consider the set

$$\bigcup_{j>k} (\mathrm{Max}\, U)_j \qquad (*)$$

Def A k-*layering* of U is a sequence $(U^{(i)})_{i=1}^{m}$ of submolecules of U s.t.

① $U \simeq U^{(1)} \#_k \dots \#_k U^{(m)}$

② each $U^{(i)}$ contains *exactly* one element of $(*)$

IDEA :

We can, essentially, reduce the rewritable submolecule problem to the problem of finding an $(n-1)$-layering of U in which maximal elements in the image of V are consecutive, + extra conditions

Def The layering dimension of U is the integer

$$\mathrm{lydim}(U) := \min\left\{ k \geq -1 \;\middle|\; \#\bigcup_{j>k+1} (\mathrm{Max}\, U)_j \leq 1 \right\}$$

Def The frame dimension of U is the integer

$$\mathrm{frdim}(U) :=$$

$$\dim \bigcup \left\{ \mathrm{cl}\{x\} \cap \mathrm{cl}\{y\} \;\middle|\; x, y \in \mathrm{Max}\, U, \quad x \neq y \right\}$$

Theorem U regular molecule

$$\mathrm{frdim}(U)$$
$$\geq$$
$$\min\{k \mid \text{a } k\text{-layering of } U \text{ exists}\}$$
$$\geq$$
$$\mathrm{lydim}(U)$$
$$\geq$$
$$\dim(U)$$

Def The maximal k-flow graph of U is the directed graph $\mathcal{M}_k U$ with

- vertex set $\bigcup_{j>k} (\operatorname{Max} U)_j$

- an edge $x \longrightarrow y$ iff $\Delta_k^+ x \cap \Delta_k^- y \neq \emptyset$.

Def A k-ordering of U is a topological sort of $\mathcal{M}_k U$.

Proposition

Every k-layering of U determines a unique k-ordering.

THE CONVERSE IS NOT TRUE IN GENERAL!

IDEA OF THE ALGORITHM:

1. Compute $M_{n-1} U$.
2. Find a topological sort in which vertices in V are consecutive.
3. Check if it is induced by an $(n-1)$-layering satisfying the conditions; if not, iterate.

Each step can be computed efficiently...

but we may have to try out factorially many topological sorts before we find a layering.

Def U is **frame-acyclic** if $\forall V \sqsubseteq U$, if $r := \mathrm{fr\,dim}\, V$, then $M_r V$ is acyclic.

Seems contrived, but has some remarkable consequences for which it appears to be somewhat tight:

- k-layerings are bijection with k-orderings, for all k
- U "is equivalent" to a polygraph

Def U is stably frame-acyclic if $\forall V \in U$, every substitution in V produces a frame-acyclic molecule.

When U is stably frame-acyclic, either...

ALL $(n-1)$-layerings such that vertices in V are consecutive satisfy the extra conditions, or

NONE do,

so there is no need to iterate!

Proposition

If $\dim U \leq 3$, then U is stably frame-acyclic.

There is a counterexample in dim 4!

- Is there a PTIME algorithm for subdiagram matching in dim 4?
- Or, the more interesting option: is it NP-complete?

THANK YOU FOR YOUR ATTENTION!

ioc.ee/namar