Graph Transformation, Atom Tracing, and Isotope Labelling

Peter Severin Rasmussen

University of Southern Denmark (SDU)

Feb. 12, 2018
About me

- Peter Severin Rasmussen
- University of Southern Denmark (SDU)
- Computer Science student
- Master thesis
- Supervisor: Daniel Merkle
Examples from chemistry

- Isotope labelling experiments
- Mass spectrometry
- Hypothetical (prebiotic) chemistries
- Metabolic engineering
- Synthesis planning
- One-pot synthesis
My master thesis

- Isotope labelling experiments ←
- Mass spectrometry
- Hypothetical (prebiotic) chemistries
- Metabolic engineering
- Synthesis planning
- One-pot synthesis
Figure 25: Carbon trace of two glycolysis pathways. The Embden-Meyerhof-Parnas pathway (EMP) is depicted with black reaction arrows, and the Entner-Doudoroff pathway (ED) is depicted with green reaction arrows. The six carbon atoms from glucose are converted into two pyruvate molecules, highlighted in blue, in two different ways depending on whether EMP or ED was used to catabolise glucose. The trace for one pyruvate overlaps in the pathways, while the sequence of carbons is inverted in the other pyruvate.
A bit of background

The molecular model

Molecules as graphs

Reactions as graph transformations
MØD Overview

Models, methods, and concepts

- Molecules
- Reaction patterns
- Typed attributed graphs
- Graph transformation rules
- Point groups, stereochemistry
- Exploration strategies
- Rule composition
- Reaction networks
- Directed hypergraphs
- ILP
- Tree search
- Pathways
- Integer hyperflows
- Petri-nets
- Pathway realisations
- Rule composition
- Atom maps
- Atom traces

Core Graph Algorithms

- Monomorphism enum.
- Isomorphism
- Canonicalization
- Automorphism enum.

Software

- The MØD package:
  - C++ library
  - Python interface
  - Figure generation
  - GraphCanon library
  - PermGroup library
Chemical reaction patterns
Chemical reaction patterns

Chemical Reactions (Educts $\rightarrow$ Products)

http://mod.imada.sdu.dk 6/46
Chemical reaction patterns

\[
\begin{align*}
\text{Reactant} & \quad \rightarrow \quad \text{Product} \\
\end{align*}
\]
Chemical reaction patterns

Rule

Educts

Products

pattern match

Reaction

Rule

http://mod.imada.sdu.dk 8/46
Chemical reaction patterns

Rule

pattern match

Educts

http://mod.imada.sdu.dk
Chemical reaction patterns
Example: Formose

Molecules

Formaldehyde

\[
\begin{align*}
\text{H}&\langle3\rangle \\
\text{C}&\langle0\rangle \\
\text{O}&\langle1\rangle \\
\text{H}&\langle2\rangle \\
\end{align*}
\]

Glycolaldehyde

\[
\begin{align*}
\text{H}&\langle6\rangle \\
\text{C}&\langle1\rangle \\
\text{O}&\langle0\rangle \\
\text{H}&\langle5\rangle \\
\text{H}&\langle4\rangle \\
\text{C}&\langle2\rangle \\
\text{H}&\langle7\rangle \\
\text{O}&\langle3\rangle \\
\end{align*}
\]
Example: Formose

Rules

Keto-Enol Isomerization

Aldol Addition
Example: Formose
Derivation Graph / Chemical network
Example: Formose
Derivation Graph / Chemical network

Generation 1
Example: Formose

Derivation Graph / Chemical network

Generation 2
Example: Formose
Derivation Graph / Chemical network

Generation 3
Example: Formose
Derivation Graph / Chemical network
Example: Formose
Derivation Graph / Chemical network

Limited to molecules with \( \leq 5 \) carbons
Derivation Graph as Semigroup of Transformations

1. Groups
2. Permutation Groups
3. Construction
4. Semigroups
Group Theory

Group: \((G, \bullet)\)

**Closure** If \(g, h \in G\), then \(g \bullet h \in G\).

**Associativity** For all \(g, h, k \in G\), then \((g \bullet h) \bullet k = g \bullet (h \bullet k)\).

**Identity** There exists \(e \in G\) s.t. for all \(g \in G\), then

\[
e \bullet g = g = g \bullet e
\]

**Inverse** For all \(g \in G\), there exists \(g^{-1} \in G\) s.t.

\[
g^{-1} \bullet g = e = g \bullet g^{-1}
\]
Group Theory: Example

\[ G = \{0, 1, 2, 3\} \]
\[ \bullet = \_ + \_ \mod 4 \]

Identity: 0

\[ 0 + 2 = 2 \]
\[ 2 + 0 = 2 \]
Group Theory: Example

\[ G = \{0, 1, 2, 3\} \]
\[ \bullet = \_ + \_ \mod 4 \]

Inverse: \(-x\)

\[ 1 + (-1) = 1 + 3 = 4 \mod 4 = 0 \]
Group Theory: Example

\[ G = \{0, 1, 2, 3\} \]
\[ \bullet = \_ + \_ \mod 4 \]

Closure

\[ 1 + 1 = 2 \]
\[ 2 + 3 = 5 \mod 4 = 1 \]
Group Theory: Example

$G = \{0, 1, 2, 3\}$
$\bullet = _+ _ \mod 4$

Generators

$G = \langle 1 \rangle = \langle 1, 2 \rangle$
Permutation Groups

Points
\( \Omega = \{0, 1, 2, \ldots, n\} \)

Permutation
\( \sigma : \Omega \to \Omega \)

Example:
\( \sigma : 5 \mapsto 7 \)
\( 7 \mapsto 5 \)
\( 11 \mapsto 42 \)
\( 42 \mapsto 10 \)
\( 10 \mapsto 11 \) (rest unchanged)

Cyclic notation
\( \sigma = (5 \ 7)(11 \ 42 \ 10) \)
Permutation Groups

Permutation
\[ \sigma : \Omega \rightarrow \Omega \]
Ex:

\[ \sigma : 5 \mapsto 7 \]
\[ 7 \mapsto 5 \]
\[ 11 \mapsto 42 \]
\[ 42 \mapsto 10 \]
\[ 10 \mapsto 11 \]

(rest unchanged)
Permutation Groups

Permutation

\( \sigma : \Omega \rightarrow \Omega \)

Ex:

\[
\begin{align*}
\sigma : & \ 5 \mapsto 7 \\
 & \ 7 \mapsto 5 \\
 & \ 11 \mapsto 42 \\
 & \ 42 \mapsto 10 \\
 & \ 10 \mapsto 11 \\
\end{align*}
\]

(rest unchanged)

Cyclic notation

\[
\sigma = (5 \ 7)(11 \ 42 \ 10)
\]
Cool Tools from Group Theory

- Orbit
- Schreier-Sims algorithm
- ...
Cool Tools from Group Theory

- Orbit

\[ G = \langle (1\ 2)(3\ 4), (2\ 5) \rangle \]

\[ \text{Orbit}_G(\omega) = \{ g(\omega) \mid g \in G \} \]
\[ \text{Orbit}_G(1) = \{ \text{id}(1), g_1(1), g_2(1), (g_1 \circ g_2)(1) \} \]
\[ = \{ 1, 2, 5 \} \]

Can be done on pairs too.

- Schreier-Sims algorithm
Cool Tools from Group Theory

- Orbit
- Schreier-Sims algorithm
  - Membership testing in poly time
  - Element decomposition
Permutation Group of Derivation Graph
Permutation Group of Derivation Graph

Ω = \{0, 1, 2, \ldots, 22\}
Permutation Group of Derivation Graph
Permutation Group of Derivation Graph

(8 14)(11 15)(9 12)(10 13)
Permutation Group of Derivation Graph

\[ G = \langle (8 \ 14)(11 \ 15)(9 \ 12)(10 \ 13), \quad (13 \ 15)(12 \ 14), \quad (18 \ 16)(19 \ 15)(17 \ 14), \ldots \rangle \]
Permutation Group of Derivation Graph
Permutation Group of Derivation Graph
Permutation Group of Derivation Graph

Start

Goal
Permutation Group of Derivation Graph
Permutation Group of Derivation Graph

(0, 2, 3) → (13, 13, 13)

Start

Goal
Permutation Group of Derivation Graph

(0,2)  
?  
(13,12)  

Start  
Goal  

Glycolaldehyde

H₂C=O $\overset{\text{Formaldehyde}}{\rightarrow}$

0  
2  

13  
12  

Semigroups

Closure  If \( g, h \in G \), then \( g \bullet h \in G \).

Associativity  For all \( g, h, k \in G \), then \( (g \bullet h) \bullet k = g \bullet (h \bullet k) \).

Identity  There exists \( e \in G \) s.t. for all \( g \in G \), then

\[
\begin{align*}
e \bullet g &= g = g \bullet e
\end{align*}
\]

Inverse  For all \( g \in G \), there exists \( g^{-1} \in G \) s.t.

\[
\begin{align*}
g^{-1} \bullet g &= e = g \bullet g^{-1}
\end{align*}
\]
Semigroup Tools

- Orbit calculation: ✔
- Schreier-Sims: ✔
Demo
Current Work

- Further development of MØD.
- Python implementation of semigroup construction.
- Ability to trace atoms of specific markings via orbits.
Use Cases

- Glycolysis: ED & EMP Pathways
- TCA
- PPP
- Formose
- Polyketides
- ...

...
Future Work

- Assist in isotope labelling experiments (automated)
- Constraint-based membership testing
- Dynamic simulation w. ODEs
Questions?