Daniel Merkle

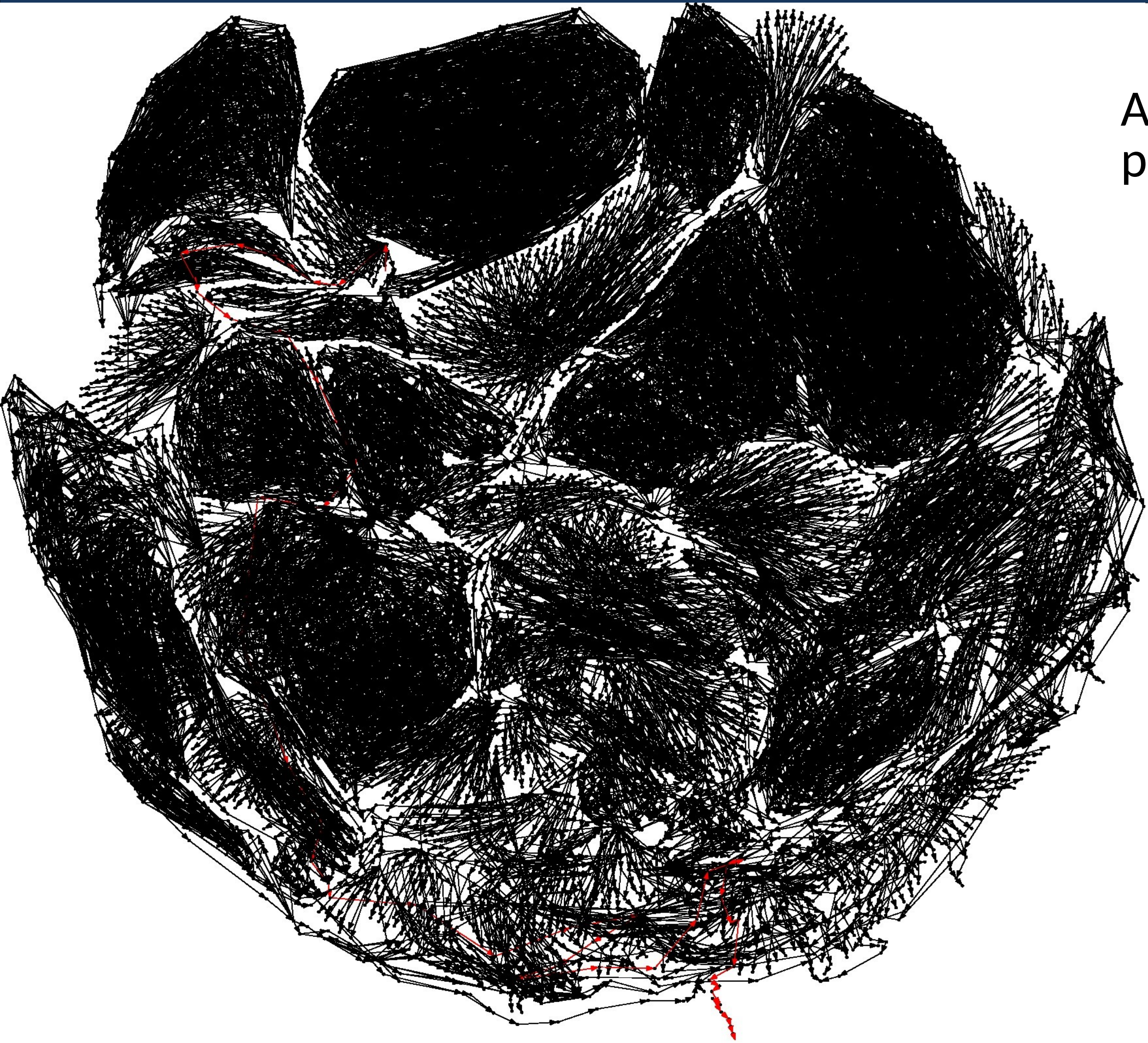
LETS START WITH A GAME

(CATALAN)

https://www.imada.sdu.dk/~daniel/DM840-2020/assignment1/assign1-2020.html

https://www.imada.sdu.dk/~daniel/DM840-2021/ (For this slide set)

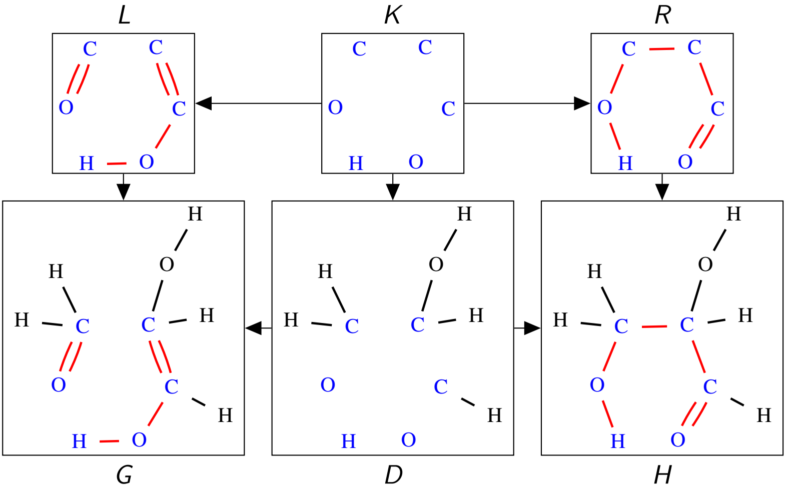
DM840 – Algorithms in Cheminformatics



A ”solution” is a shortest path in a directed graph

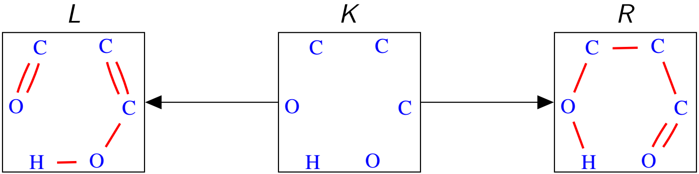
Reactions – Application of Transformation Rules

1,2-ethenediol + formaldehyde $\xrightarrow{\text{aldol addition}}$ glyceraldehyde

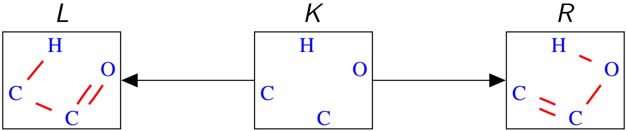


The Formose Chemistry

4 reaction patterns:

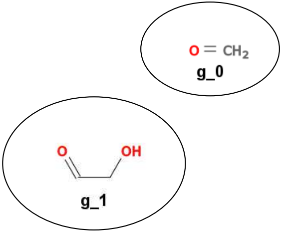


(a) Aldol addition (r_2). (Reverse aldol addition (r_3))



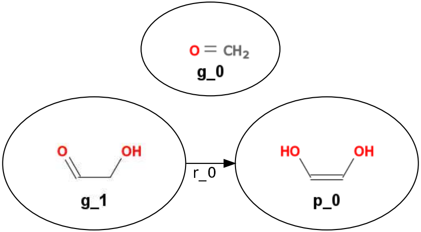
(b) Keto-to-enol (r_0). (Enol-to-keto (r_1))

Reaction Network for Formose



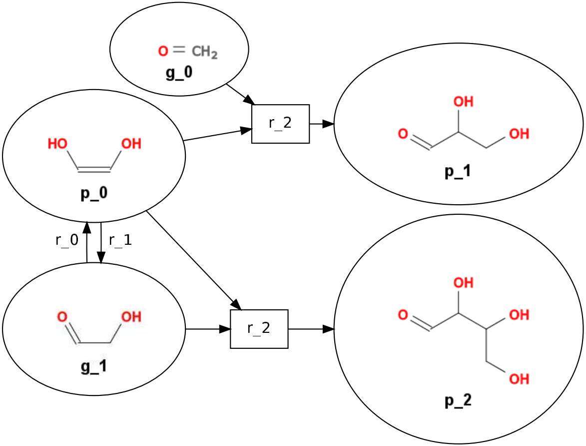
Initial State

Reaction Network for Formose



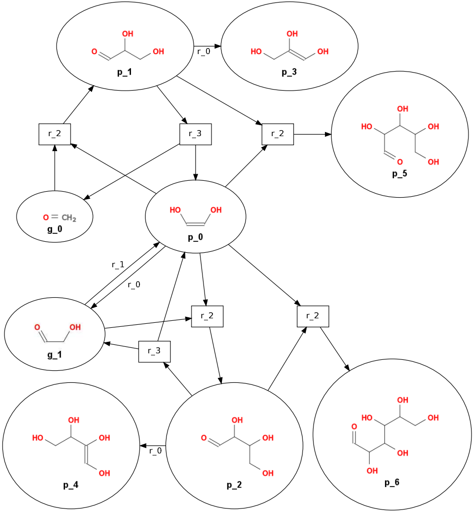
Generation 1

Reaction Network for Formose



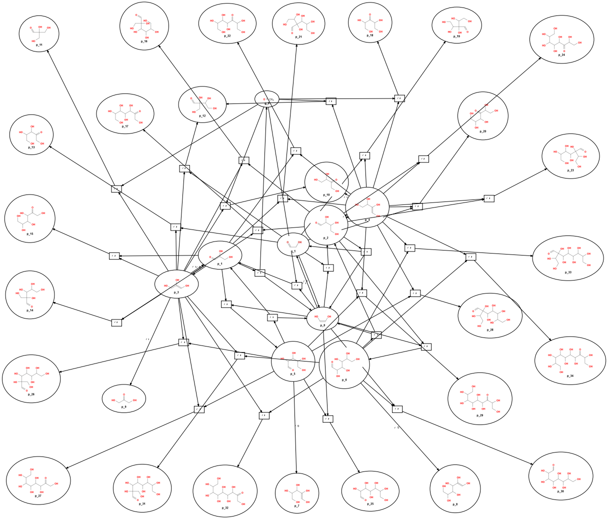
Generation 2

Reaction Network for Formose



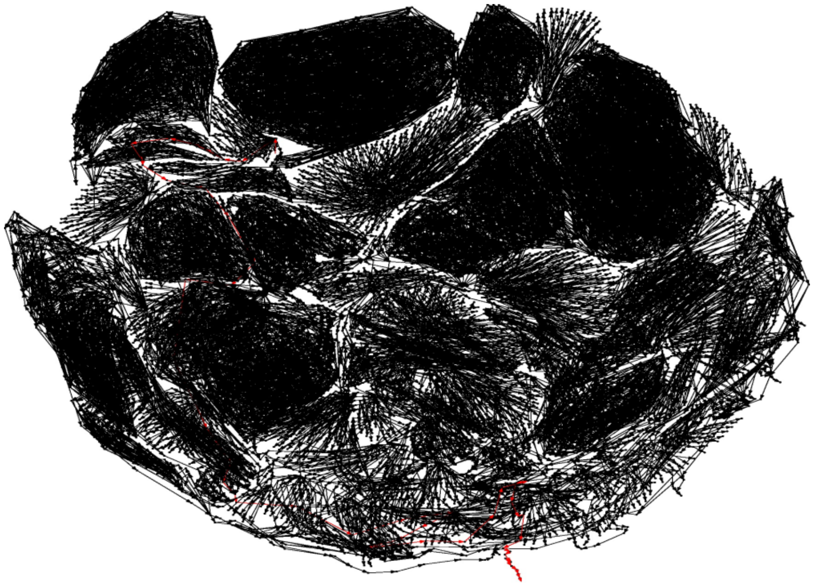
Generation 3

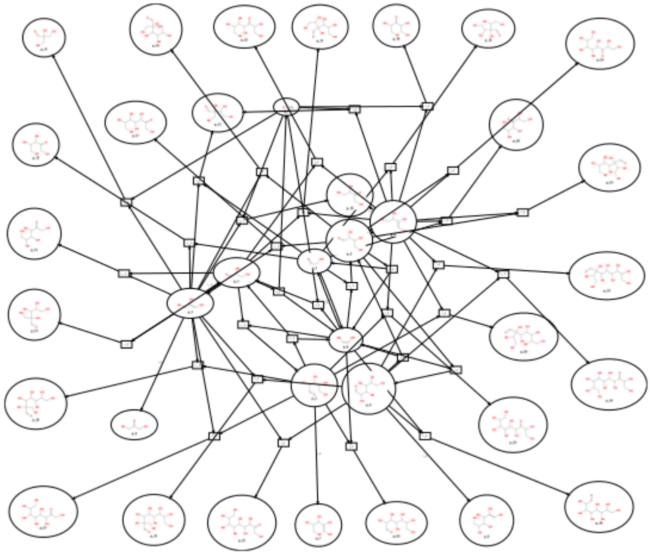
Reaction Network for Formose



Generation 4

Shortest Path(s)

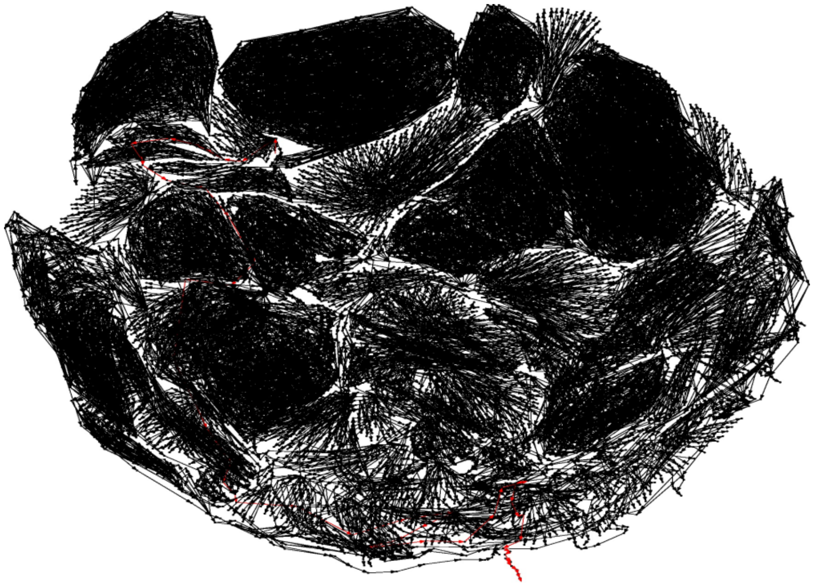




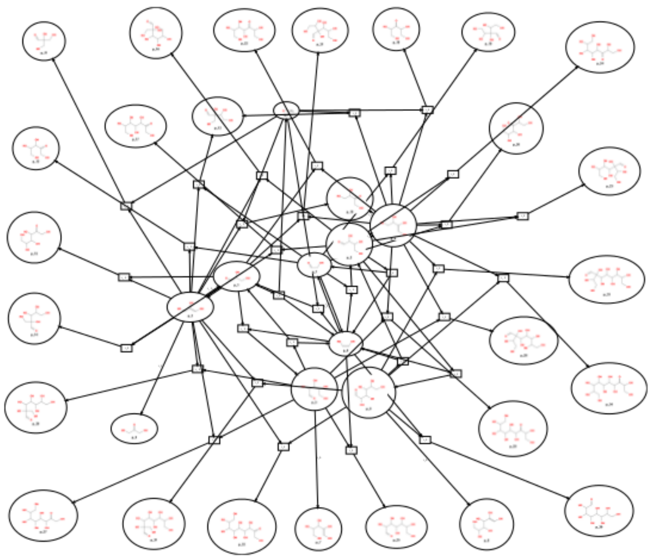
A ”solution” is a shortest path in a directed graph

http://cheminf.imada.sdu.dk/mod

Shortest Path(s)



A ”solution” is a shortest path in a directed graph



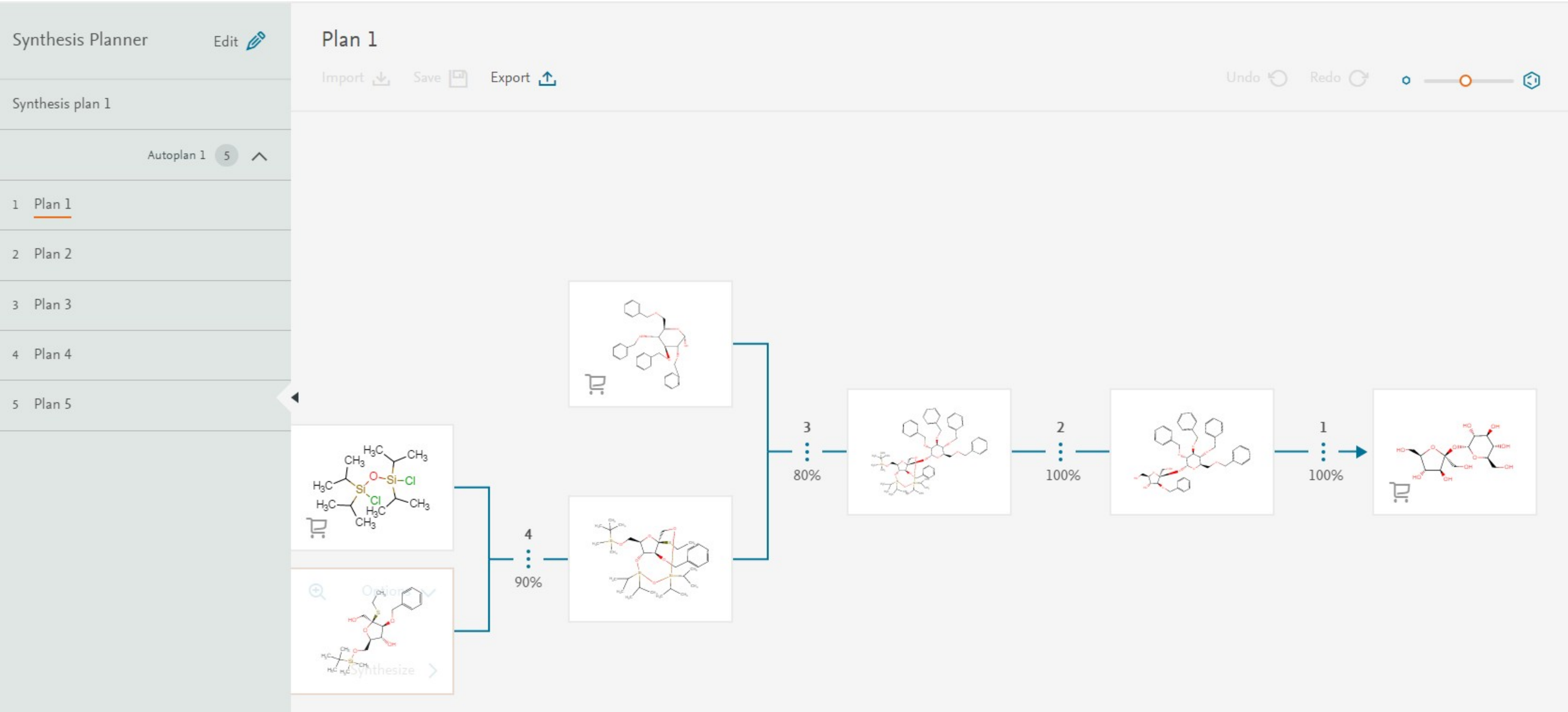
How to find the k best solutions?

In hypergraphs?

Algorithms?

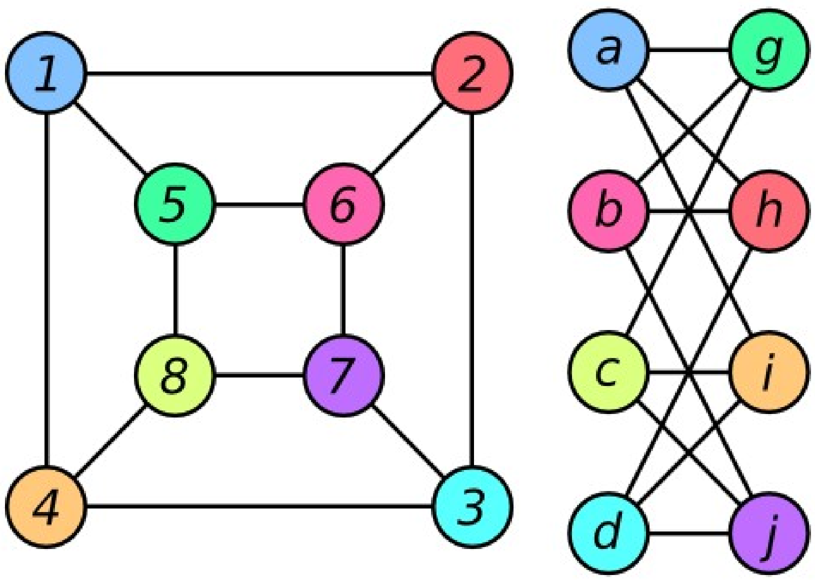
Computational Complexity?

How chemists do it



(Sub-)Graph Isomorphism / Canoncial Representaions of Graphs

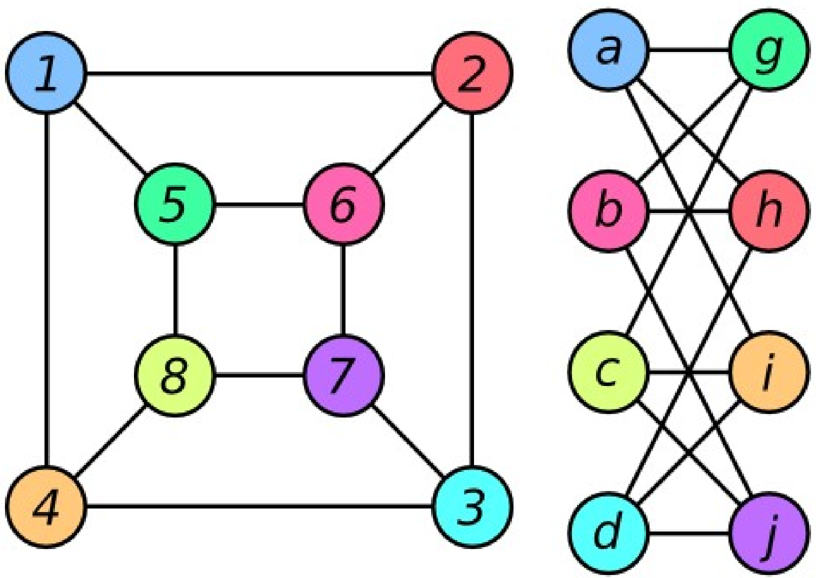
Are these 2 colored graphs isomorph?



Algorithms? Computational Complexity? State-of-the-Art?

(Sub-)Graph Isomorphism / Canoncial Representaions of Graphs

Are these 2 colored graphs isomorph?



A Quasipolynomial Time Algorithm for Graph Isomorphism: The Details

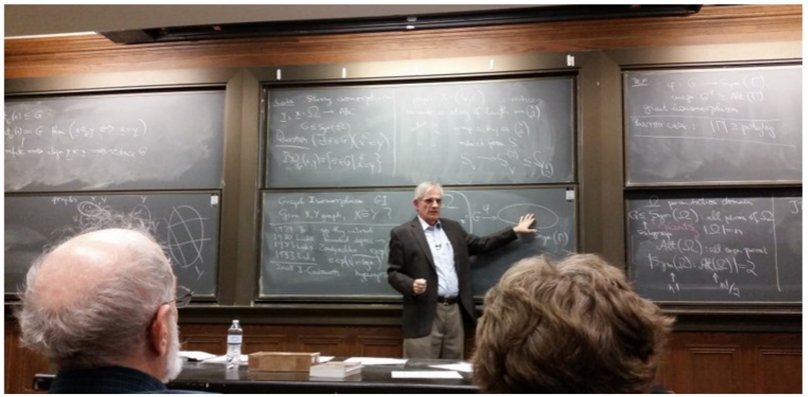
Posted on November 12, 2015 by j2kun

Update 2015-11-21: Ken Regan and Dick Lipton posted an article with some more details, and a high level overview of how the techniques fit into the larger picture of CS theory.

Update 2015-11-16: Laci has posted the talk on his website. It's an hour and a half long, and I encourage you to watch it if you have the time 🙂

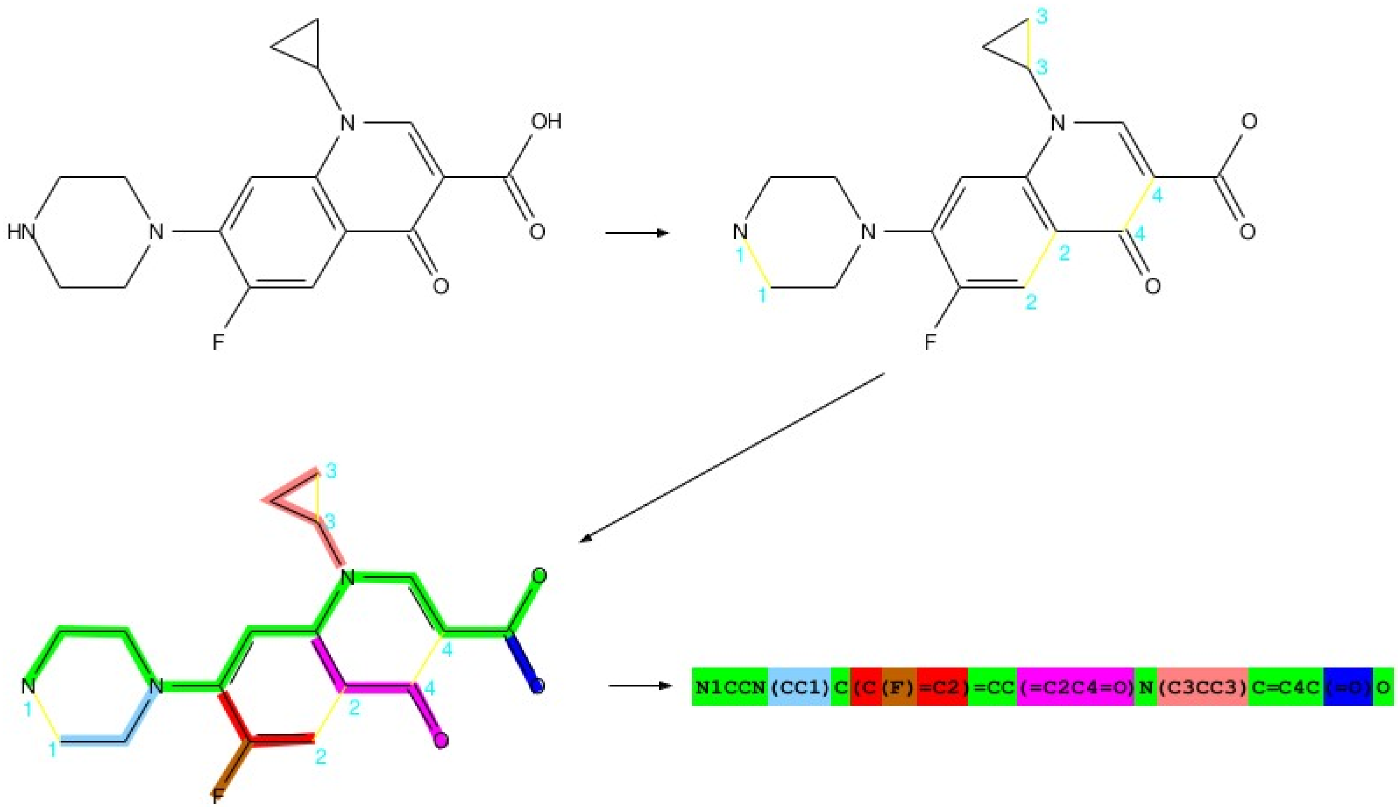
Laszlo Babai has claimed an astounding theorem, that the Graph Isomorphism problem can be solved in quasipolynomial time. On Tuesday I was at Babai's talk on this topic (he has yet to release a preprint), and I've compiled my notes here. As in Babai's talk, familiarity with basic group theory and graph theory is assumed, and if you're a casual (i.e., math-phobic) reader looking to understand what the fuss is all about, this is probably not the right post for you. This post is research level theoretical computer science. We're here for the juicy, glorious details.

Note: this blog post will receive periodic updates as my understanding of the details improve.



Algorithms? Computational Complexity? State-of-the-Art?

(Sub-)Graph Isomorphism / Canoncial Representaions of Graphs

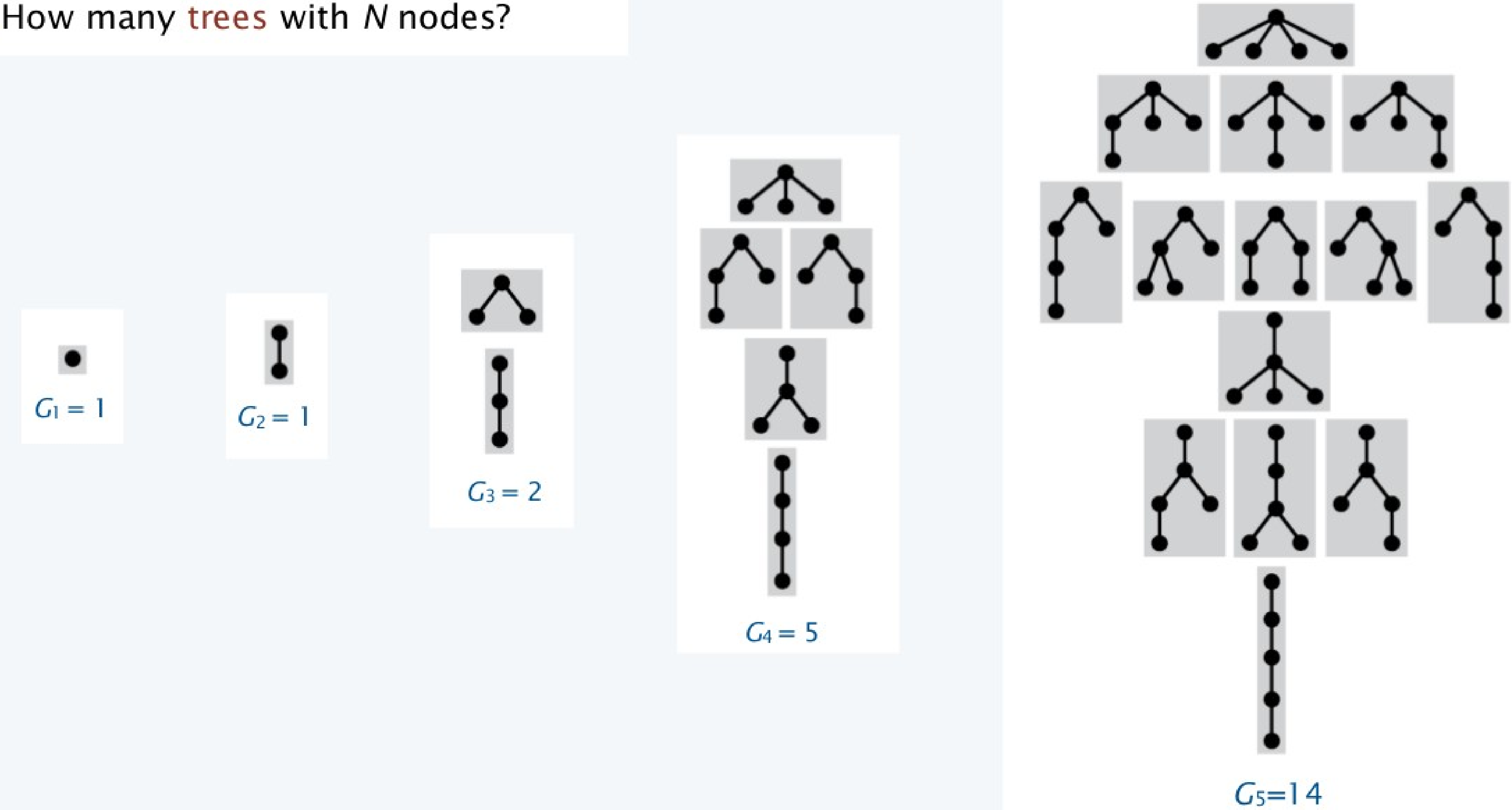


Canoncial representations of graphs (i.e. molecules):

Algorithms? Computational Complexity?
And … w.t.h. went wrong in Chemistry?

Counting

Q. How many trees with N nodes?

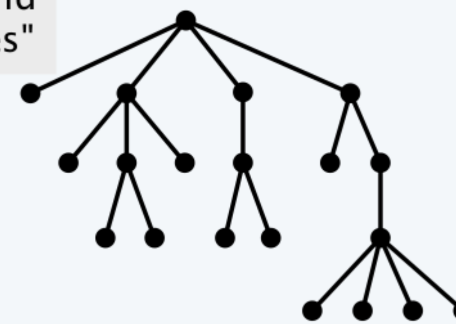


More Counting

Variations on a theme 1: Trees

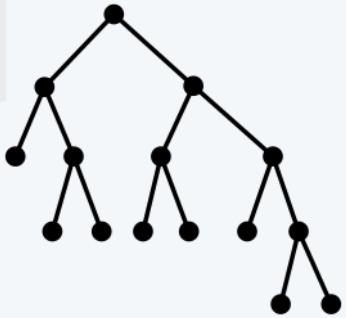
Fundamental construct

Combinatorial class	G, the class of all trees	"a tree is a node and a sequence of trees"
Construction	$G = \bullet \times SEQ(G)$	
OGF equation	$G(z) = z(1 + G(z) + G(z)^2 + G(z)^3 + \ldots) = \dfrac{z}{1 - G(z)}$	
	$G(z) - G(z)^2 = z$	



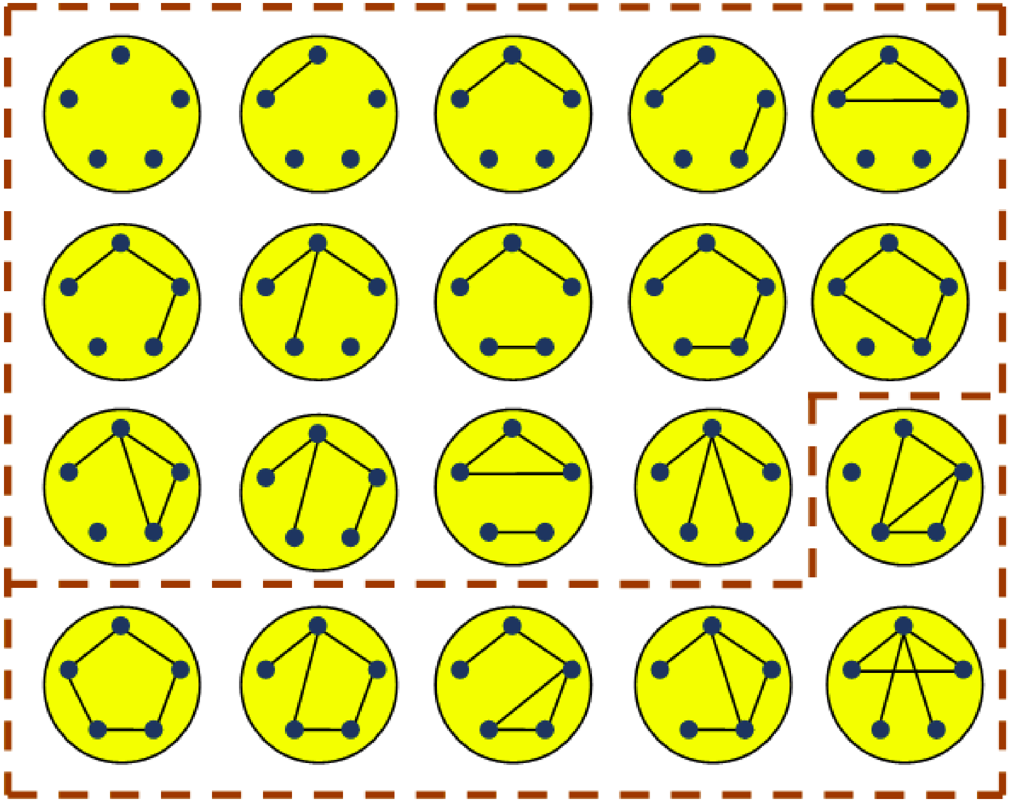
Variation on the theme: *restrict each node to* 0 *or* 2 *children*

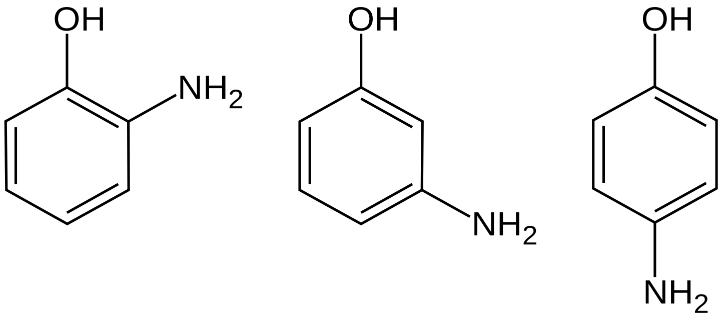
Combinatorial class	T, the class of binary trees	"a binary tree is a node and a sequence of 0 or 2 binary trees"
Construction	$T = \bullet \times SEQ_{0,2}(T)$	
OGF equation	$T(z) = z(1 + T(z)^2)$	



Applications: Sampling, Analysis of Algorithms, Chemistry

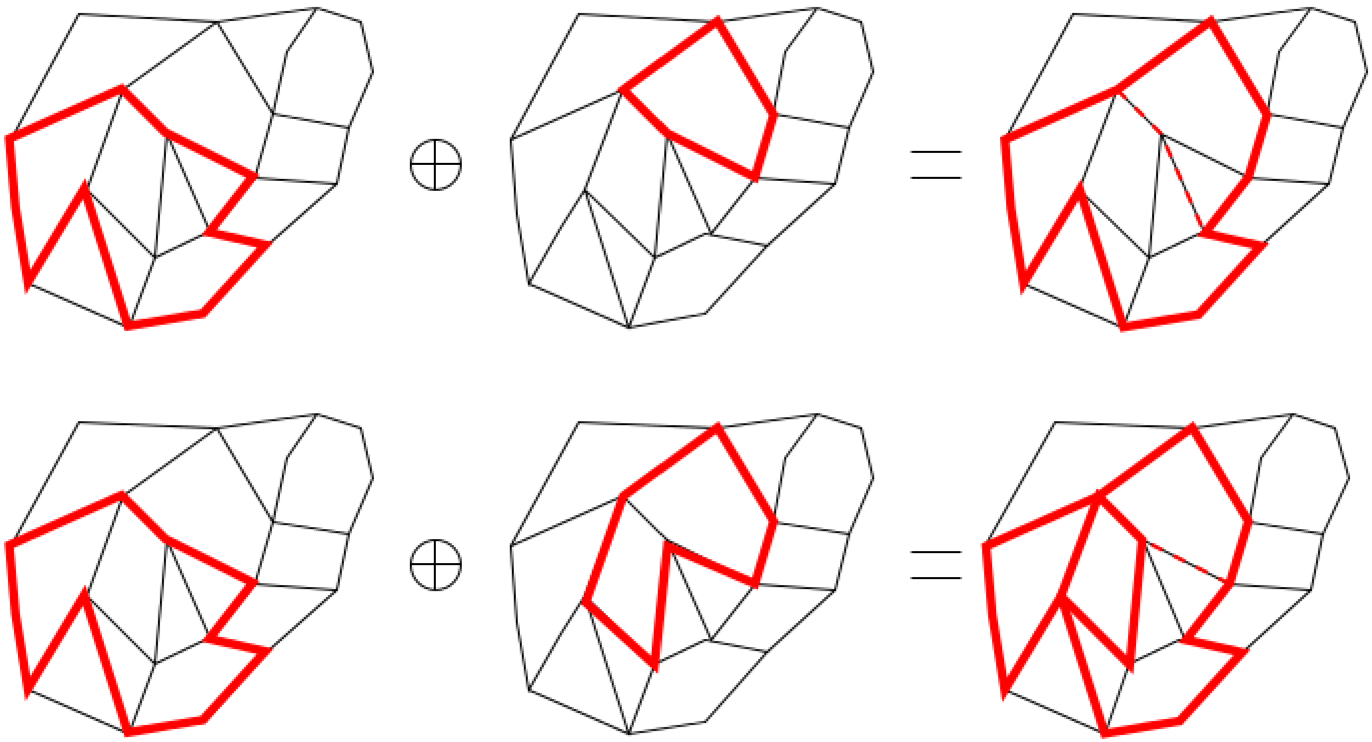
Even more counting





More Topics

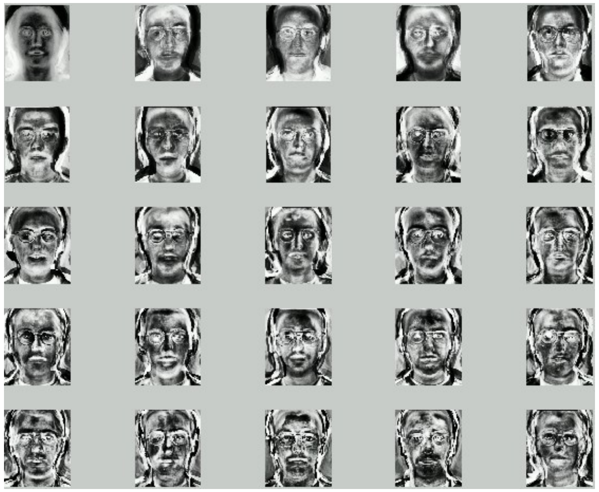
Cycle Bases (Scheduling, Distributed Computing, Chemistry)



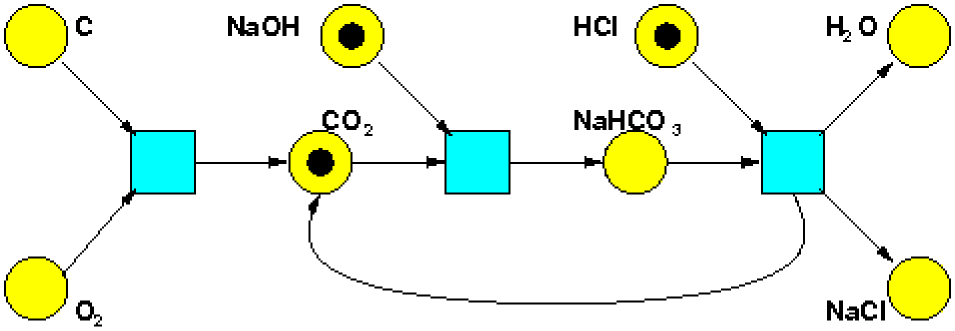
How to find a minimal cycle basis?
Algorithms?
Computaional Complexity?
...and w.t.h. went wrong in chemistry?

More Topics

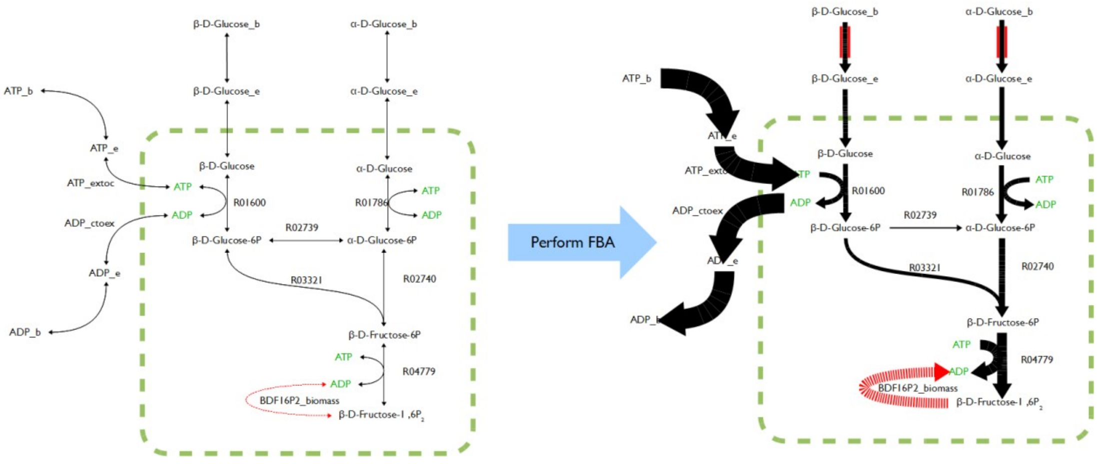
QSAR / PCA

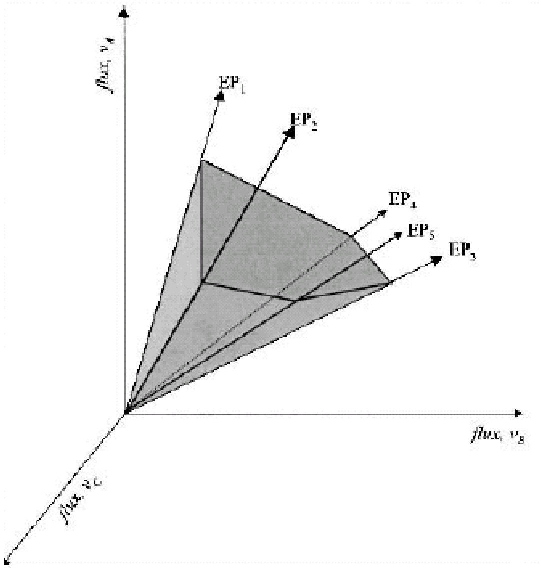


Petri Nets



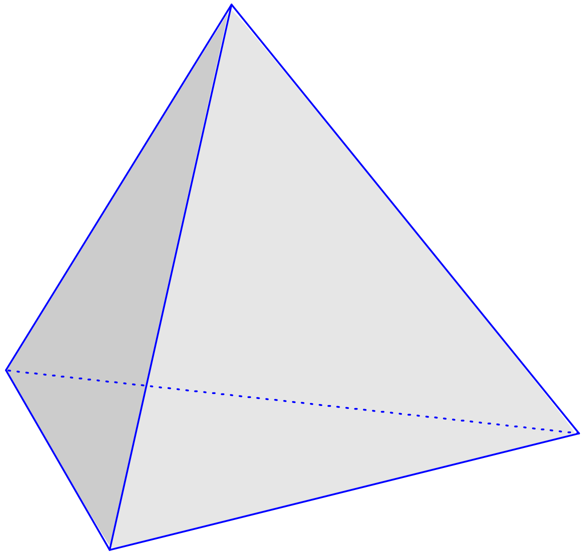
Flux Balanced Analysis





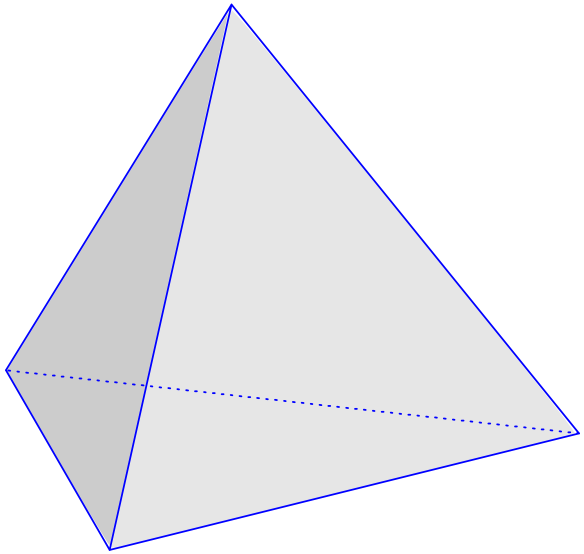
More Topics

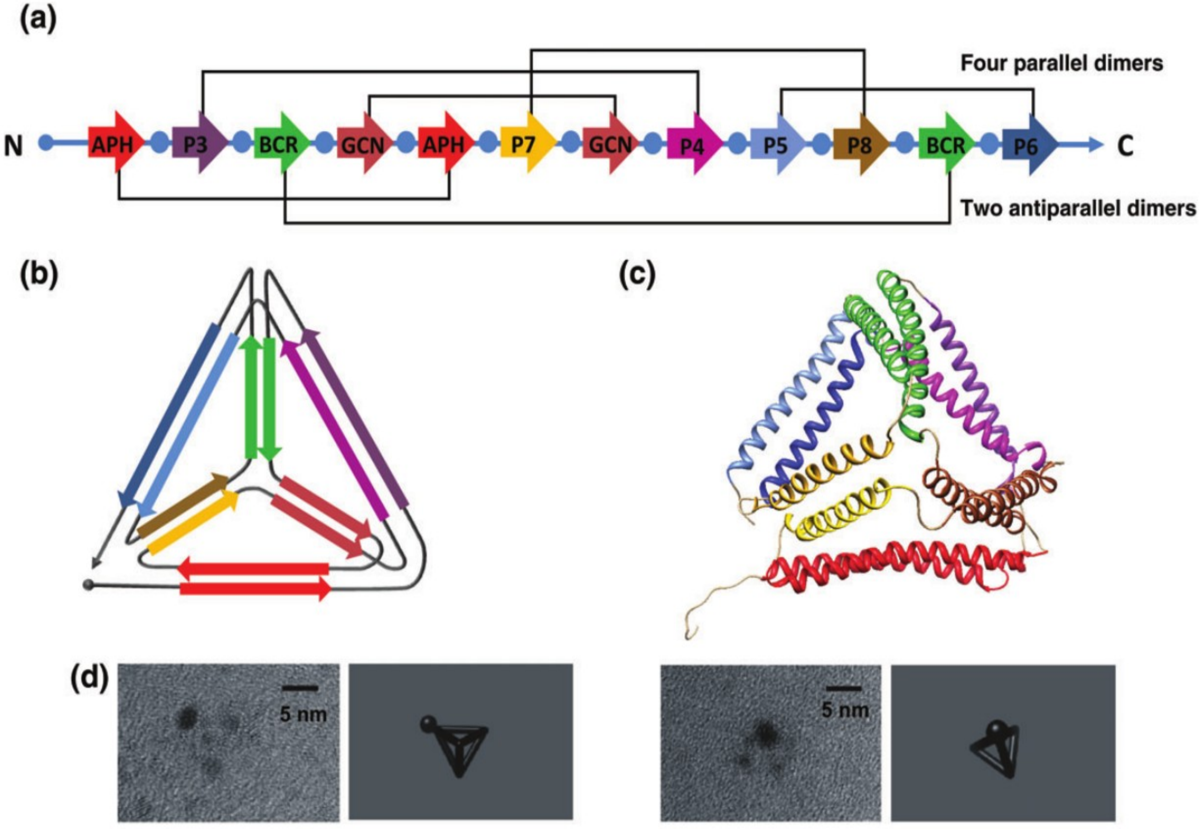
...and many more
For example: Traverse a tetrahedron, each edge twice.



More Topics

...and many more
For example: Traverse a tetrahedron, each edge twice.





Algorithms in Cheminformatics - DM840

Topics

Canonical Representations
Structure Descriptors
Graph Isomorphism
Pólya's **Counting Theory**
Artificial Chemistries
Graph Grammars
Cycle Bases
Metabolic Networks and Metabolic Pathways
Concurrency Theory
Petrinets
Applications in Industry
...

Formal Model

Algorithms

Computational Complexity

Applications (including Chemistry)

Algorithms in Cheminformatics - DM840

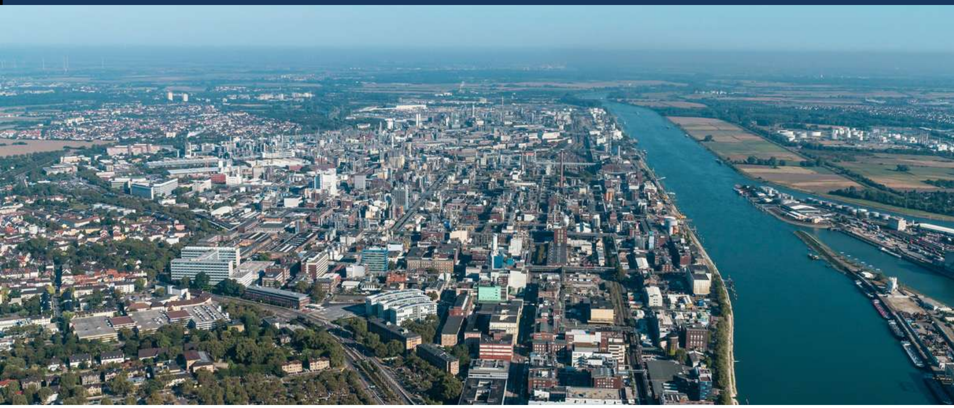
Topics

Canonical Representations
Structure Descriptors
Graph Isomorphism
Pólya's **Counting Theory**
Artificial Chemistries
Graph Grammars
Cycle Bases
Metabolic Networks and Metabolic Pathways
Concurrency Theory
Petrinets
Applications in Industry
...

- 10 ECTS

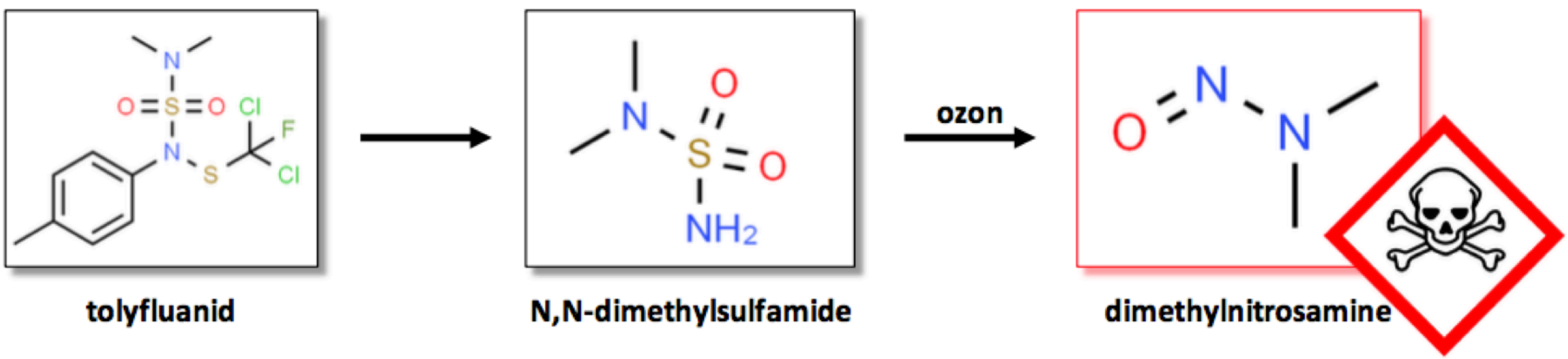
- max. 2 programming projects

no chemistry knowledge needed!

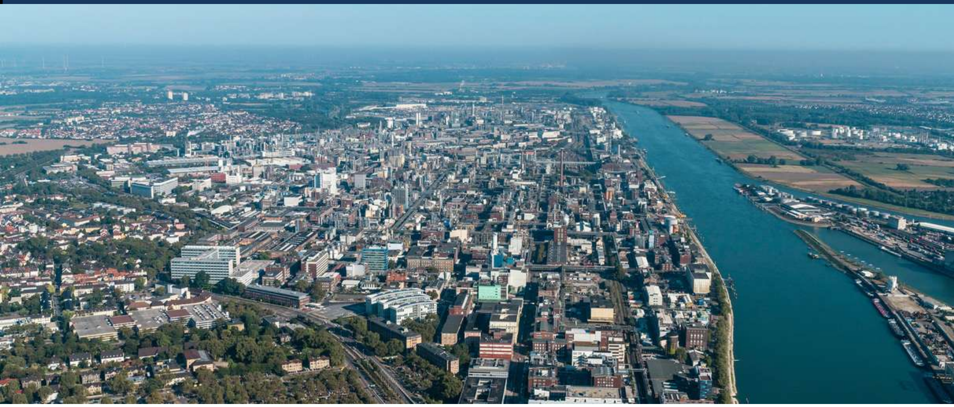


BASF
We create chemistry

HARVARD
MEDICAL SCHOOL

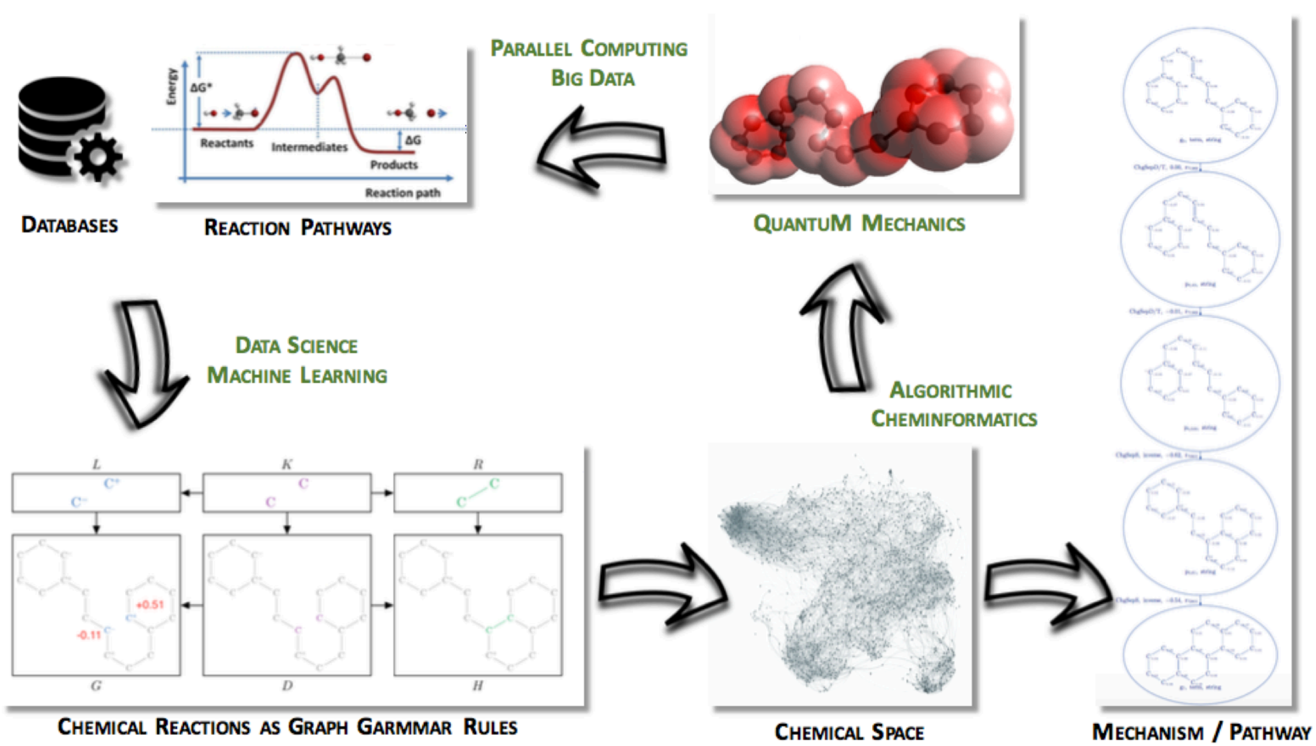


Algorithms in Cheminformatics - DM840



BASF
We create chemistry

HARVARD
MEDICAL SCHOOL



Enzyme Design

https://cheminf.imada.sdu.dk/novo-synergy/