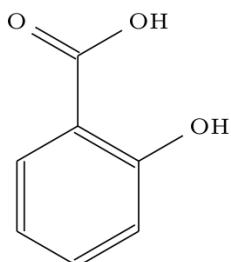


Categorical Chemistry: The architecture of a chemical computer

① Reasoning in organic synthesis

1.1 Abstractions in chemistry: The languages of technological revolution

* Substances:



Molecular formula

$C_7H_6O_3$

CHO

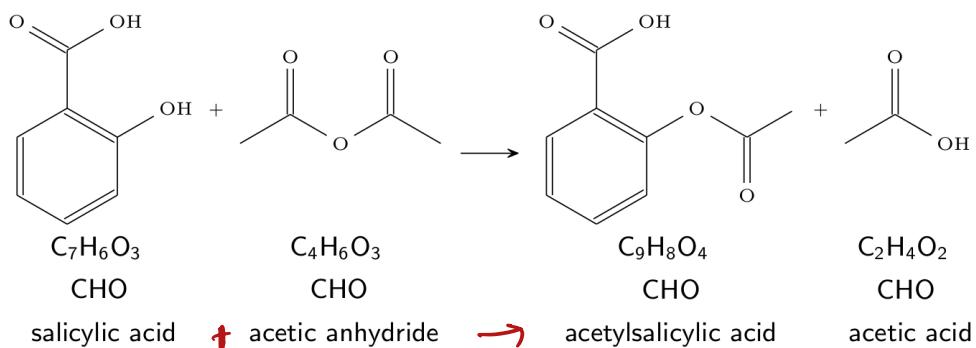
salicylic acid

Chemical formula

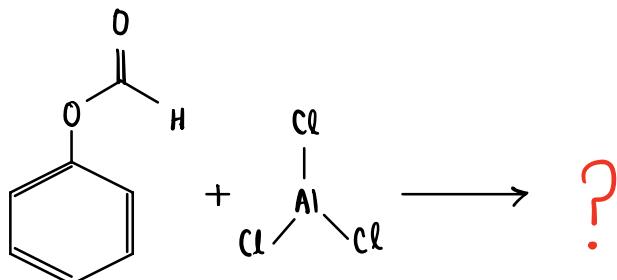
Composition

Chemical name

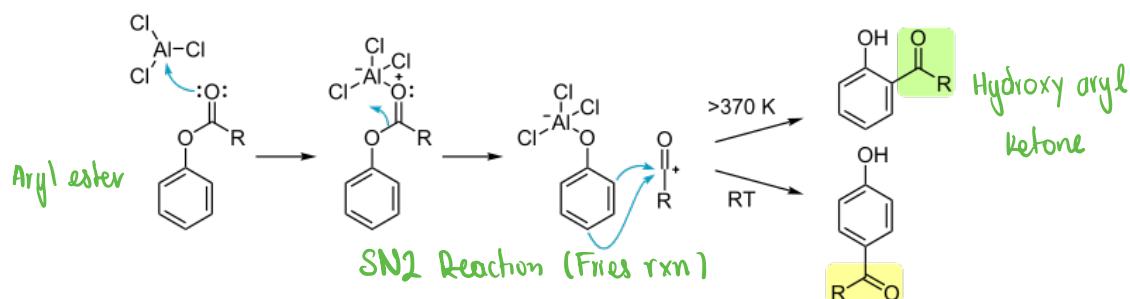
* Reactions



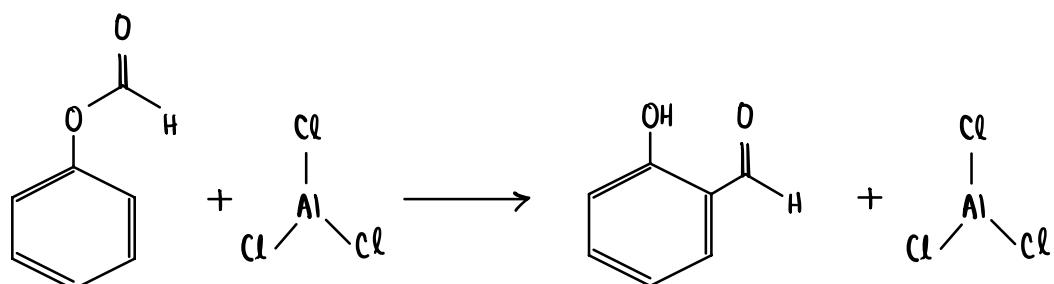
* The puzzles



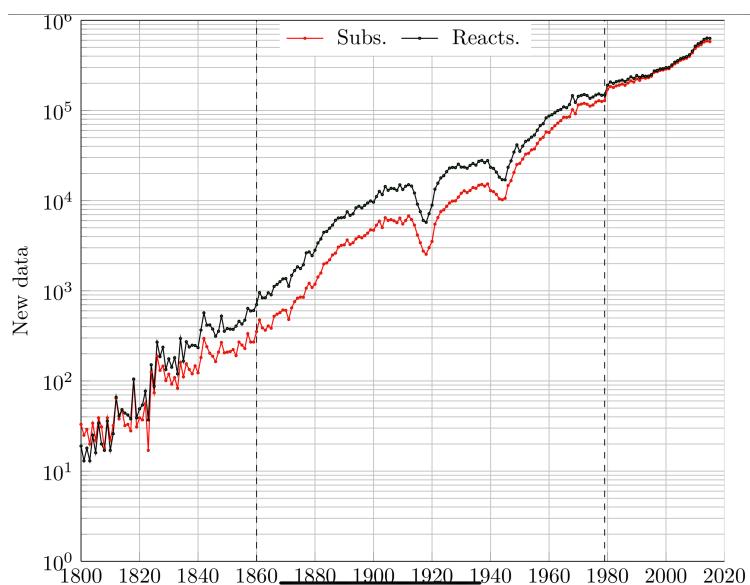
* The reasoning (chemical demonstration/proof)



* The solution



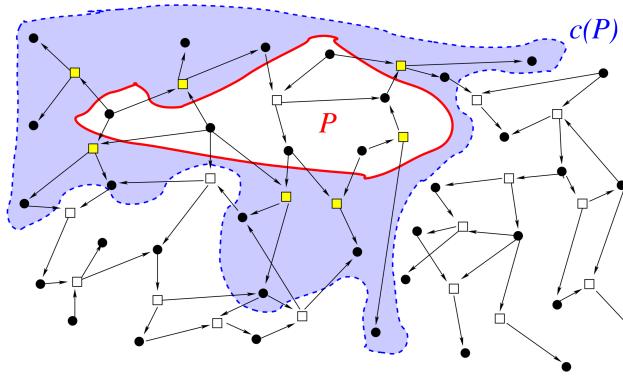
Chemists have been synthesising new substances at an exponential rate!



1.2 Goal: Just wanna write the math for the reasoning system and implement it in code

* Underlying formal spaces:

The network level

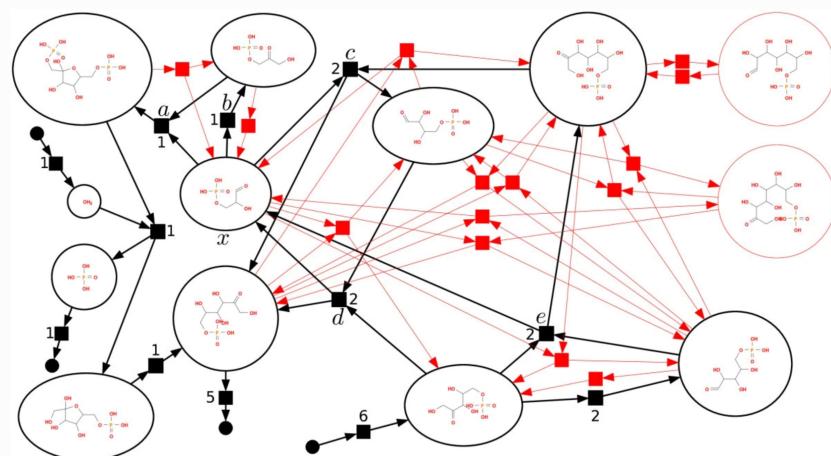


A category of directed relations where we can define: paths, reachability, connectivity, etc.

Bonus: This category should have an internal linear logic.

The molecular level

From: [Maximizing output and recognizing autocatalysis in chemical reaction networks is NP-complete](#)



Flow optimization in the pentose-phosphate reaction network. Only a small part of the chemical space is shown. We allow influx of water and ribulose-5-phosphate to generate glucose-6-phosphate as output. Phosphate is produced as waste product. An optimal solution is shown in black, using 6 ribulose-5-phosphate molecules to produce 5 glucose-6-phosphate molecules. The values of the flow $f(\cdot)$ is indicated for each hyperedge (black square), e.g., $f(a) = 1, f(b) = 1, f(c) = 2, f(d) = 2, f(e) = 2$. At each node (except the unlabelled input and output nodes) the influx and outflux is balanced. For example, at node x (glycerol-3-phosphate), we have $f(d) + f(e) = 4 = f(a) + f(b) + f(c)$.

A category where:

- Objects: Molecular formulas
- Morphisms: Reactions

The chemical similarity is affinity levels.

$A \sim A'$ if they react with similar substances to produce similar substances

Synthesis vs Analysis:

The identity of an object is given by all its interactions

Synthetic approach

Prediction of behavior \mathcal{I} $\xrightarrow{\quad}$ Prediction of internal structure \mathcal{E}

Analytical approach

The identity of an object is given by its internal structure / logic

The "Synthetic" approach in chemistry:

for each reaction $A + B_i \rightarrow C_i$ there is a reaction $A' + B'_i \rightarrow C'_i$

where,

$$A \sim A' \Leftrightarrow B_i \sim B'_i \Rightarrow C_i \sim C'_i$$



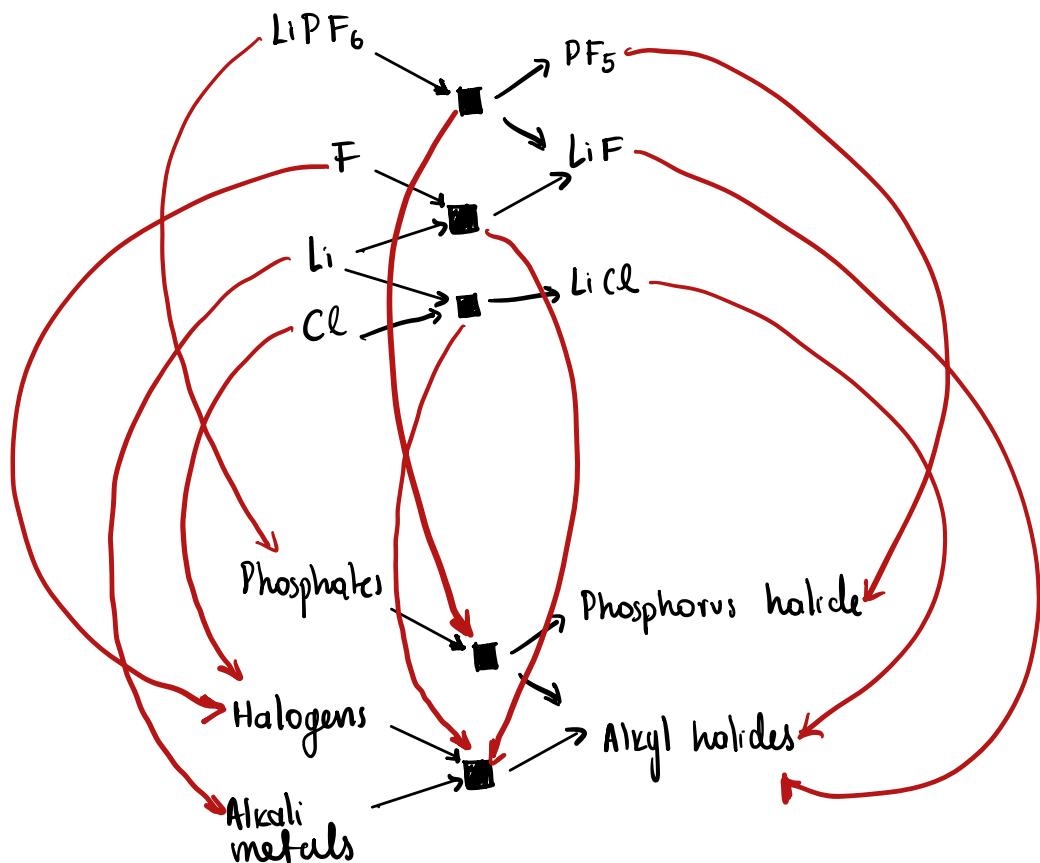
We have that $Li \sim Na$ since $F \sim Cl$, $Cl \sim Br$, $Br \sim I$

There must be a functor:

$$\mathcal{C}: N \rightarrow CN$$

where N & CN are objects of your favorite category of directed hypergraphs / Petri nets.

That sends each substance to its similarity "class" \rightarrow each reaction to its reaction class.



Note: \mathcal{C} can be defined without knowing the molecular structure.

topology solution: substances, functional groups \rightarrow Galois

UV

CONNECTIONS.

The "analytic" approach in chemistry:

The identity of a substance is given by pieces of the molecule called **functional groups**: All substances having the same **functional group** are similar.

* Data needed:

- A collection M of substances
- A collection FG of functional groups
- A relation

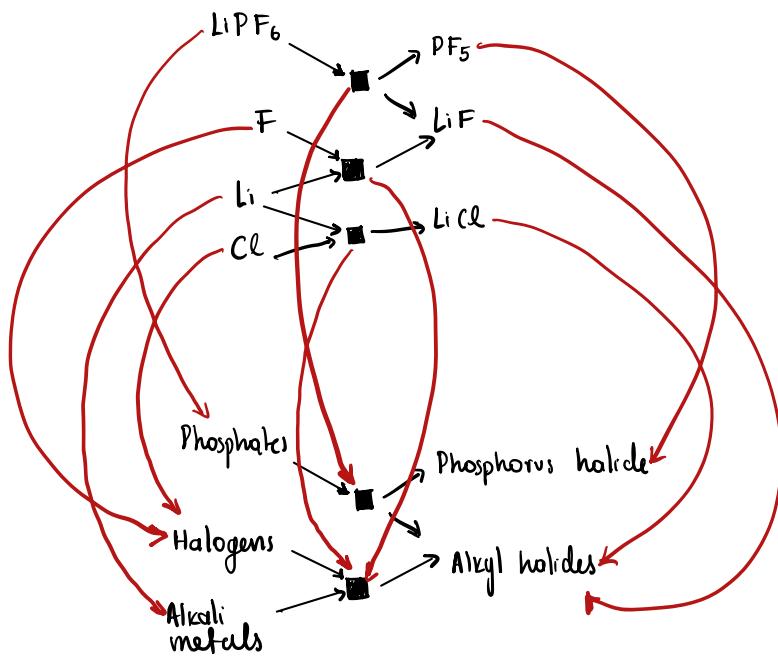
$R \subseteq M \times FG$
telling us which functional groups are present in a substance.

* Model of similarity:

$$\alpha: 2^M \rightarrow 2^{FG}$$
$$A \mapsto \alpha(A) = \{ f \in FG : \forall a \in A, aRf \}$$
$$\beta: 2^{FG} \rightarrow 2^M$$
$$B \mapsto \beta(B) = \{ a \in M : \forall f \in B, aRf \}$$
$$S: 2^M \rightarrow 2^M$$

is a closure operator whose fixed points (closed sets) form a topological structure to M that we will call **similarity**:

$$\epsilon: N \rightarrow CN$$

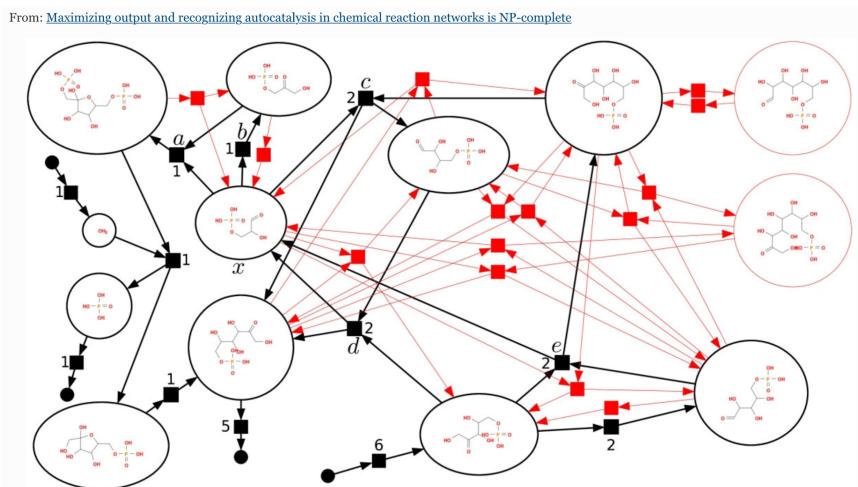


The construction above is a passage between the network level and similarity level.

Note: We introduced a chemical structure model!

This is a formal model for chemical structure used before 1860.

The molecular level:



A category where:

- Objects: Molecular structures
- Morphisms: Reactions

* Approach:

- Objects: Presheaves over a schema Order $\leftarrow B \begin{smallmatrix} \hookrightarrow \\ \curvearrowright \end{smallmatrix} A \rightarrow$ Type
- Reactions: Rewritings of these Functors

Proposition: The molecular rewriting level is the resulting:

- Double category
- Category of isomorphic classes of Rewritings.

What are functional groups at this level?