Bond-Energy-Guided Program Synthesis for Chemical Reaction Network Discovery

Thomas Bood
T.M.Bood@student.tudelft.nl
Responsible Professor: Sebastijan Dumančić
Supervisor: Reuben Gardos Reid

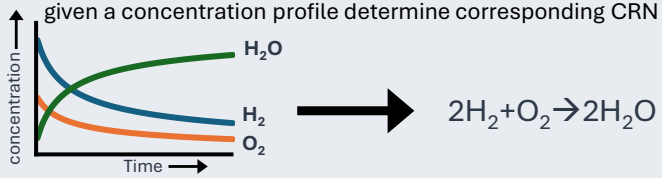
1 Background

Chemical Reaction Network (CRN)

$CH_2O_2 + C_2H_6O \rightarrow C_3H_6O_2 + H_2O$
$C_3H_6O_2 + CH_4O \rightarrow C_2H_4O_2 + C_2H_6O$

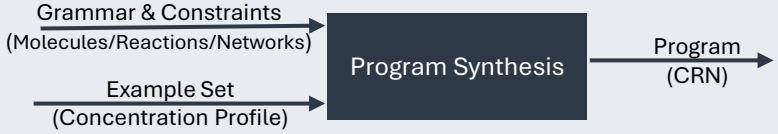
The Problem of CRN Discovery

given a concentration profile determine corresponding CRN



Program Synthesis:

The automated process of deriving programs that satisfy certain behaviour.

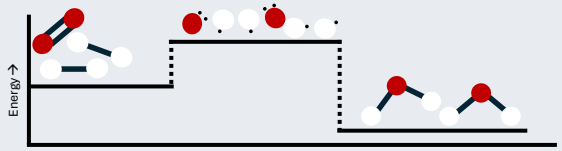


The used system makes use of a pipeline structure:

Atoms → Molecules → Reactions → Network

Bond Breaking Energy

Energy required to break chemical bonds. Lower energy reactions are more thermodynamically favourable, and thus more likely.



2 Research Question

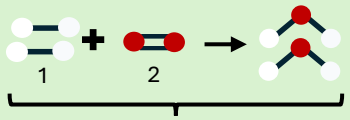
Can bond-energy-guided heuristic search in program synthesis of chemical reaction networks reduce the number of evaluated candidates and total runtime before the target network is found compared to uninformed search?

Why this matters

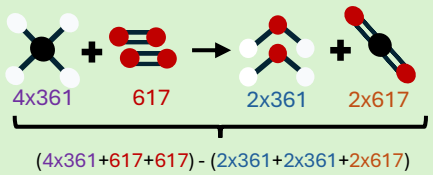
- For real-world CRN problems the search space is too large for exhaustive search.
- A reduction in candidates or runtime means more efficient search

3 Energy Based Heuristics

Max Bond Order:
Prioritise reactions that break bonds of lower order.



Delta Energy:
Take bond energies of reactants minus bond energies of products. Prioritising lower sums. Bond energies are estimated KJ/mol (single = 361; double = 617; triple = 838)



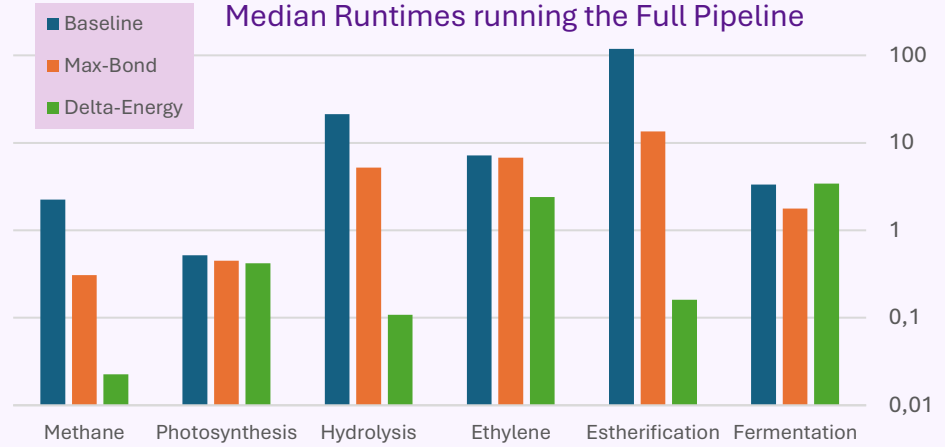
4 Results

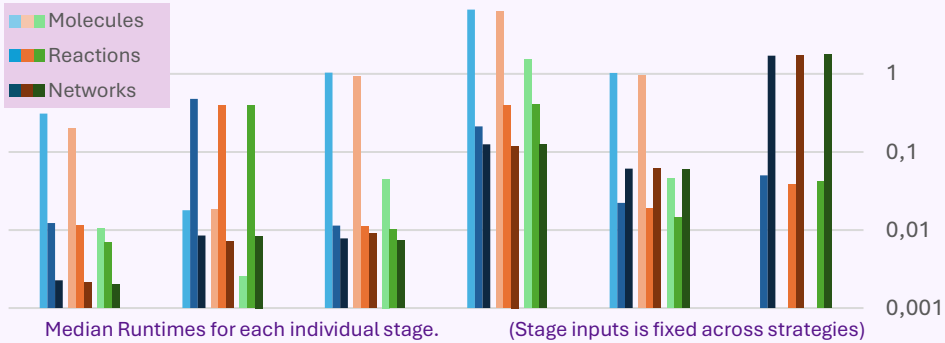
Speedups achieved by heuristic guided search over baseline

Problem	Max-Bond	Delta-Energy
Methane	7.2x	99x
Photosynthesis	1.2x	1.2x
Hydrolysis	4.1x	192x
Ethylene	1.1x	3.0x
Esterification	8.8x	744x
Fermentation	1.9x	0.97x

		Molecules	Reactions	Networks
Full	Max-bond	50	14095	8
	Delta-Energy	22	431	6
stage	Max-bond	48	86	4
	Delta-Energy	20	81	4

Candidates explored in the Hydrolysis problem
(full pipeline vs individual stages with fixed input)





Median Runtimes for each individual stage. (Stage inputs is fixed across strategies)

5 Conclusions

- **Delta Energy > Max Bond Order > BFS**
  Delta Energy generally outperforms Max Bond Order. Both heuristics overall perform better than BFS.
- **Up to 8.8x Faster Runtimes for Max Bond Order** compared to BFS
- **Up to 744x Faster Runtimes for Delta Energy** compared to BFS
- **The heuristics main problem seem to be the inability to properly differentiate similar reactions**
- **The staged pipeline structure leads to earlier stages being more important as improvements compound**

Future Work

- ❑ More Sophisticated Heuristics (ex. based on connected atoms)
- ❑ Integrate effects of stage specific heuristics.
- ❑ Scale to larger, industrial scale CRNs