

Finding Reaction Mechanism Pathways with Deep Reinforcement Learning and Heuristic Search

A chemical reaction, as the overall transformation of reactant molecules into product molecules, can be broken down into smaller reaction steps called elementary steps. An elementary step reaction involves only a single transition state, such as making bonds, breaking bonds, or donating electrons, that transforms the chemical reactants into chemical products. While Artificial Intelligence (AI) has been used to predict the outcomes of chemical reactions, most of these reaction predictors are designed to predict the major outcome of overall transformations, skipping the chemical reactions at a mechanistic level, making it difficult to identify intermediates and byproducts of the reaction. Information on the reaction mechanisms allows practitioners to identify intermediate molecules, predict the results of similar reactions under various conditions, and validate the feasibility of that reaction. Existing models for elementary steps, like OrbChain [1], can predict reaction mechanisms using a beam search algorithm, an incomplete search algorithm that does not explicitly consider path cost. To address this, we learn a heuristic function and perform a heuristic search that is complete and explicitly takes path cost into account.

Approach, Results, and Future Work: We pose finding a sequence of reaction mechanisms as a pathfinding problem, using the United States Patent and Trademark Office (USPTO) [2] dataset for training. We create an offline dataset of precomputed start states(reactants), goal states(products), and expanded start states. It takes 4-5 days with 48 CPUs to generate 2 million datasets. We use DeepCubeA [3], with Hindsight Experience Replay [4] to learn a heuristic function that generalizes over start and goal states to guide A* search to predict the sequence of mechanistic reactions of an overall chemical transformation, from reactants to products. DeepCubeA is trained to solve puzzles for predefined goals. To handle this, we build on hindsight experience replay to train a heuristic function that generalizes over any start and goal pair. We then use the learned heuristic function with the A* search algorithm to find paths from start states to goal states. We generate a test dataset with start/goal pairs to test our trained heuristic by taking 0-6 random steps from a given start state. We also test with Tanimoto similarity [5] to guide search. DeepCubeA outperforms uniform cost search and the Tanimoto similarity. We also explore learnable representations with rxnfp [6] instead of static representations ECFPS [7], and observe that DeepCubeA heuristic with learnable representation outperforms static representations like ECFPS. The USPTO dataset has only one product called a major product, which does not contain byproducts of chemical reactions. The product can contain a set of goal states instead of a single goal. Future work includes training a heuristic function that generalizes over states and a set of goal states.

References:

1. Tavakoli, Mohammadamin, et al. "Ai for interpretable chemistry: Predicting radical mechanistic pathways via contrastive learning." *Advances in Neural Information Processing Systems* 36 (2023): 4080-4096.
2. Lowe, Daniel. "Chemical reactions from US patents (1976-Sep2016)." (*No Title*) (2017).
3. Agostinelli, Forest, et al. "Solving the Rubik's cube with deep reinforcement learning and search." *Nature Machine Intelligence* 1.8 (2019): 356-363.
4. Andrychowicz, Marcin, et al. "Hindsight experience replay." *Advances in neural information processing systems* 30 (2017).
5. Flower, Darren R. "On the properties of bit string-based measures of chemical similarity." *Journal of chemical information and computer sciences* 38.3 (1998): 379-386.
6. Schwaller, Philippe, et al. "Mapping the space of chemical reactions using attention-based neural networks." *Nature machine intelligence* 3.2 (2021): 144-152.
7. Rogers, David, and Mathew Hahn. "Extended-connectivity fingerprints." *Journal of chemical information and modeling* 50.5 (2010): 742-754.

Step/s	Solver	Path Cost	% Solved	Nodes	Secs	Nodes/Sec
Steps=0	DeepCubeA (rxnfp)	0.00	100.00%	3.09E+2	4.07	76.03
	DeepCubeA (ecfp)	0.00	100.00%	3.09E+2	3.87	79.97
	Uniform Cost Search	0.00	100.00%	3.09E+2	4.61	67.13
	Tanimoto Similarity	0.00	100.00%	3.09E+2	3.71	83.42
Steps=1	DeepCubeA (rxnfp)	1.00	100.00%	6.82E+2	10.07	67.81
	DeepCubeA (ecfp)	1.00	100.00%	7.49E+2	9.70	77.26
	Uniform Cost Search	1.00	100.00%	4.26E+4	553.33	76.95
	Tanimoto Similarity	1.00	100.00%	3.13E+4	429.29	72.97
Steps=2	DeepCubeA (rxnfp)	1.93	100.00%	1.14E+3	16.81	68.05
	DeepCubeA (ecfp)	2.07	100.00%	1.63E+4	267.16	60.87
	Uniform Cost Search	1.67	20.00%	1.32E+5	1497.77	87.96
	Tanimoto Similarity	1.75	26.67%	1.10E+5	1229.10	89.13
Steps=3	DeepCubeA (rxnfp)	2.79	93.33%	9.29E+3	127.02	73.18
	DeepCubeA (ecfp)	2.77	86.67%	4.14E+4	578.88	71.54
	Uniform Cost Search	-	0.00%	-	-	-
	Tanimoto Similarity	-	0.00%	-	-	-
Steps=4	DeepCubeA (rxnfp)	3.60	100.00%	2.63E+4	350.06	75.33
	DeepCubeA (ecfp)	3.33	60.00%	6.36E+4	821.64	77.36
	Uniform Cost Search	3.00	6.67%	1.43E+5	1962.28	73.01
	Tanimoto Similarity	3.00	6.67%	2.47E+4	272.15	90.64
Steps=5	DeepCubeA (rxnfp)	3.92	86.67%	6.37E+4	744.31	85.60
	DeepCubeA (ecfp)	3.40	33.33%	8.40E+4	968.49	86.69
	Uniform Cost Search	-	0.00%	-	-	-
	Tanimoto Similarity	-	0.00%	-	-	-
Steps=6	DeepCubeA (rxnfp)	3.55	73.33%	4.70E+4	678.27	69.29
	DeepCubeA (ecfp)	3.20	33.33%	6.14E+4	933.86	65.73
	Uniform Cost Search	-	0.00%	-	-	-
	Tanimoto Similarity	-	0.00%	-	-	-