Learning-Guided *MCTS* for Generalizable and Interpretable Synthesis Planning Using Generic Reactions

Derin Ozer[1], **Benoit Da Mota**[1], **Nicolas Gutowski**[1], **Sylvain Lamprier**[1], **Thomas Cauchy**[2]
[1] Université d'Angers, LERIA, SFR MathSTIC, F-49000 Angers, France
[2] Université d'Angers, CNRS, MOLTECH-ANJOU, SFR MATRIX, F-49000 Angers, France
{derin.ozer, benoit.damota, nicolas.gutowski, sylvain.lamprier, thomas.cauchy}@univ-angers.fr

GDR MaDICS
27 - 28 May 2025

Context

What is synthesis planning?

Finding sequences of chemical reactions to produce target molecules from available building blocks [Cor67].

Why does it matter?

— Accelerates drug and material discovery.
— Reduces cost and lab trial-and-error.
— Enables exploration of novel molecular space.

Why is it hard?

— **NP-hard :** Exponential number of possible routes.
— **Chemical constraints :** Not all reactions are valid or feasible.

Why is it still unresolved?

— **Shallow reasoning :** Most models lack grounding in chemical logic.
— **Template issues :**
  — Template-based → limited coverage, hard to extract.
  — Template-free → poor generalization, overfit to seen transformations.
— **Benchmark bias (*USPTO* [Low12]) :**
  — Over-represents frequent transformations.
  — Misses rare or foundational reactions.
  — Encourages memorization over reasoning.

Takeaway : Most systems perform well on benchmarks but fail to generate *trustworthy, generalizable, and useful* synthesis routes in real-world settings.

Our Approach

We address key limitations in current synthesis planning by introducing a method that is :

Based on generic reactions :

— We define a minimal but expressive set of generic reactions that cover foundational and innovative chemistry : the ***Broad Reaction Set (BRS)***.

Search-based :

— We use ***Monte Carlo Tree Search (MCTS)*** to explore synthesis paths through combinatorial application of *BRS*, without relying on templates or memorized routes.

Learning-guided :

— A lightweight *Q-model* predicts the distance to the target molecule and guides *MCTS* toward promising directions during search.

The Broad Reaction Set (*BRS*) [OLC+25]

We define 20 **generic reaction patterns** in SMARTS notation [Inc] that can be flexibly applied across various molecules and to multiple sites within the same molecule.

Why *BRS*?

— **Broad coverage :** Can represent all types of transformations.
— **Multi-site reactivity :** Supports combinatorial expansion.
— **Dataset construction ready :** Can generate diverse and balanced synthesis datasets.
— **Bridging paradigms :** Combines strengths of template-based and template-free approaches.

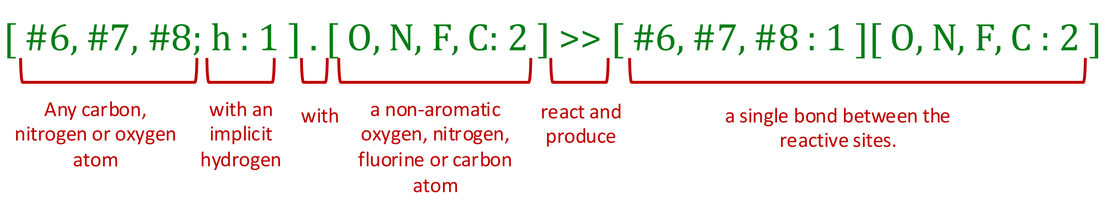

FIGURE 1 – An example reaction from *BRS*, written in the SMARTS chemical language.

Monte Carlo Tree Search (*MCTS*) for Synthesis Planning

We use **Monte Carlo Tree Search (MCTS)** to construct full synthesis routes from ***USPTO* starting molecules** to target products using only reactions from ***BRS***.

Graph structure :

— The tree alternates between molecule nodes (s) and reaction nodes (a).
— Edges represent applying a reaction or generating resulting molecules.

Search objective :

— Find a synthesis path from starting materials to the target (g) using only *BRS* reactions.
— Tests the capacity of *BRS* to reconstruct benchmark synthesis routes.

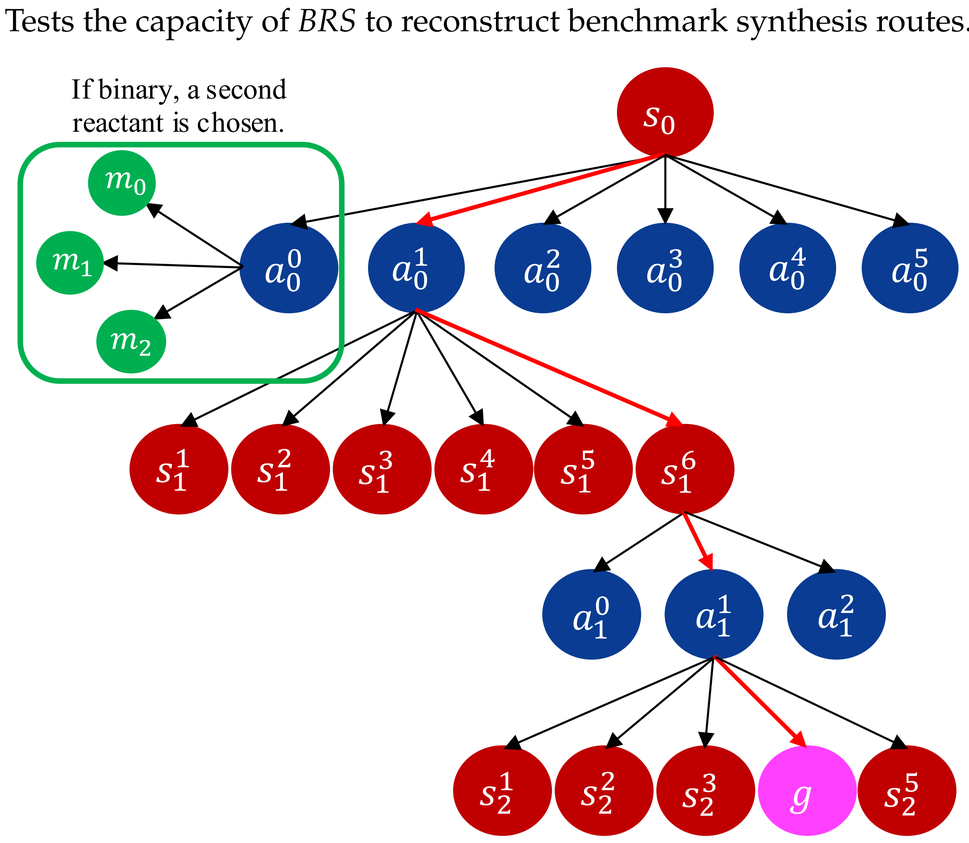

FIGURE 2 – Example of an *MCTS* search tree for synthesis planning. Molecule nodes (s_n^k) and reaction nodes (a_n^k) alternate, where n denotes the depth and k the leaf index at that depth. Binary reactions require a second reactant m. The goal is to reach the target molecule g.

Algorithm phases :

— **Selection :** Traverse the tree using a *UCT* [KS06] guided by a learned *Q-value*.
— **Expansion :**
  — At a **molecule node** : add applicable *BRS* reactions
  — At a **reaction node** : apply it to generate product molecules
— **Evaluation :** Use the *Q-network* to estimate the distance in number of reactions to the target molecule and turn it into a reward signal :

$$r(s') = \exp(-\alpha \cdot V_Q(s', g))$$

— **Backpropagation :** The reward, which tells us how promising a branch is, is passed back up the tree.

Training :

— The *Q-network* is first **pretrained** on synthetic synthesis routes created by applying *BRS* reactions to *USPTO* molecules.
— Fine-tuning online during *MCTS* episodes. The *Q-network* (a *T5* [RSR+20] encoder + regression head) is updated using visited molecule-target pairs through the successful path.
— The training objective minimizes the squared error between predicted and actual steps-to-goal :

$$\mathcal{L}_{\text{dist}} = (V_Q(s, g) - d(s, g))^2$$

Results : We divided *USPTO* into five estimated difficulty levels based on synthesis route length. Note that each level includes routes of varying lengths, and level 5 does not imply 5 reaction steps.

Difficulty	1	2	3	4	5
Success Rate	96.2%	88.7%	81.1%	–	–

TABLE 1 – Percentage of synthesis routes found at each difficulty level using BRS + MCTS.

Higher difficulty levels require deeper planning and involve more complex branching, making them harder for the *Q-network* to generalize to, especially since the pretraining examples only covered routes up to depth 20.

Références

[Cor67] Elias James Corey. General methods for the construction of complex molecules. *Pure and Applied chemistry*, 14(1) :19–38, 1967.

[KS06] Levente Kocsis and Csaba Szepesvári. Bandit based monte-carlo planning. In *European conference on machine learning*, pages 282–293. Springer, 2006.

[Low12] Daniel Mark Lowe. *Extraction of chemical structures and reactions from the literature*. PhD thesis, 2012.

[OLC+25] Derin Ozer, Sylvain Lamprier, Thomas Cauchy, Nicolas Gutowski, and Benoit Da Mota. A transformer model for predicting chemical reaction products from generic templates. *arXiv preprint arXiv :2503.05810*, 2025.

[RSR+20] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu. Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of machine learning research*, 21(140) :1–67, 2020.