

On AI-centered Retrosynthetic Planning: a Survey

Anonymous authors

Paper under double-blind review

Abstract

Retrosynthetic planning is one of the most challenging problems in organic chemistry up to date. It involves the process of designing a target molecule by recursively decomposing it into simpler molecules through a series of backward chemical reaction steps. Typically, these simpler molecules are either commercially available or easy to synthesize. The development of new computer-aided retrosynthetic methods can advance the development of target molecules in drug design, material science and agrochemicals. In this paper, we present the retrosynthetic planning problem and the current AI-based methods. We conduct an in-depth review of the application of AI-based methods to retrosynthetic planning while also presenting the current challenges and potential future research directions.

1 Introduction

Retrosynthetic Planning (RP) is a branch of organic chemistry that was first formalized over four decades ago following the success of early backward chemical synthesis. It entails designing a chemical pathway for synthesizing a target molecule through a series of recursive backward reactions (Corey, 1967). Given a target molecule, the retrosynthetic plan decomposes it into simpler building molecules that are either commercially available or easily synthesizable. Subsequently, the building blocks can be chained up through a series of forward chemical reaction steps to synthesize the target molecule.

Retrosynthesis also applies to biochemistry in particular metabolic engineering (Lin et al., 2019). Bio-retrosynthesis differs from retrosynthetic planning in two aspects: (i) the reactions must be catalyzed by enzymes and, (ii) the building molecules are molecules that are naturally produced by the organism that is being used.

The success of RP can have potential benefits in the development of new drugs, vaccines, agrochemicals, materials science, target compounds etc. Development of new drugs for example is a long and expensive process and takes an average of 10-15 years and billions of money (DiMasi et al., 2016). In addition, the ability to design target compounds can accelerate the development of new chemical compounds which can help solve some of the most pressing challenges in the world today such as climate change and food security. Accordingly, the advances in computing power (Moore, 1998; Dean, 2020) and novel algorithmic designs e.g. Deep Learning (DL) (LeCun et al., 1998; 2015; Krizhevsky et al., 2012) have inspired a strong interest in the development of new computer-aided RP methods in the last few years.

Conventionally, the RP processes are heavily dependent on the expertise of organic chemists. In the last 5 years, however, computer-aided RP methods have gained more attention due to their ability to design better and faster planning routes (Segler et al., 2018; Chen et al., 2020; Han et al., 2022). Consequently, this has inspired the development of several molecular and chemical reaction databases e.g. Reaxys¹, ChEMBL, (Gaulton et al., 2017), eMolecules², USPTO (Lowe, 2012) - to support such methods.

RP can be formulated as a tree. As an example, in Figure 1, M_0 is the target molecule. It can be designed by either the left or the right sub-tree. The squared internal nodes represent chemical reactions while the circled correspond to the molecules. Molecule M_0 can be synthesized by reaction R_0 or R_1 . Reaction R_0 requires molecule M_1, M_2 and M_3 while reaction R_1 requires molecule M_4 and M_5 . All the leaf nodes (encircled

¹<https://www.reaxys.com>

²<https://www.emolecules.com/>

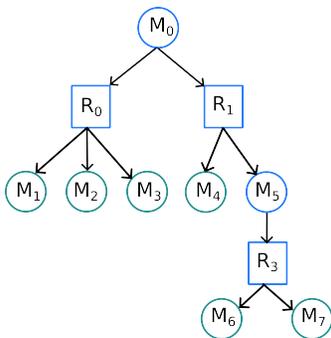


Figure 1: Retrosynthetic AND-OR tree plan. Squares represent reaction nodes and molecules represented by the circles. Each reaction node in the tree requires all of its children (AND node) while M_0 can be designed by the left or right sub-tree (OR nodes). All the leaf nodes are building block molecules.

green) are building block molecules e.g. commercially available or easy to synthesize. Molecule M_5 is not a building block thus it’s decomposed into M_6 and M_7 through reaction R_3 .

Based on the given AND-OR tree we can extract two retrosynthetic plans $P1$ and $P2$ for designing M_0 shown in Equations 1 and 2 where the symbol \implies denotes the retrosynthetic symbol e.g. the backward chemical reaction. The route lengths of the plans $P1$ and $P1$ are 1 and 2 respectively. The route length is one of the commonly used metrics used to evaluate an effective retrosynthetic plan.

$$P1 : M_0 \implies M_1 + M_2 + M_3 \quad (1)$$

$$P2 : M_0 \implies \{M_6 + M_7\} + M_4 \quad (2)$$

In RP, the proposed routes must be effective, chemically viable and feasible. While shorter routes are generally preferred, this is however contingent on the availability of the building block molecules and the real-world costs such as human capital and labor. A transformation such as $M_0 \implies M_1 + M_2 + M_3$ is not unique. As a result, other transformations can be composed to design the target e.g $M_0 \implies M_a + M_b + \dots + M_c$ given that $\{M_1, M_2, M_3\}$ and $\{M_a, M_b, M_c\} \in S$, where S is the chemical space. While the planning length typically ranges between 15-20, the number of transformations at each level in the tree for simpler molecules ranges between 80-40000 (Segler et al., 2018) with complex molecules exceeding these ranges.

Currently, there are four key AI-based approaches to RP (i) expert-based systems, (ii) template-based (iii) template-free and semi-template. In **Expert-based**, RP is designed according to the expert-engineered rules where starting material is parsed with an inference engine and then matched against the Knowledge Base (KB) of chemical reactions (Pensak & Corey, 1977; Chen & Baldi, 2009). **Template-based** RP methods are based on the chemical reaction templates. Here, the reaction rules are not predefined as expert-based systems but rather extracted automatically from a database of chemical reactions. This is achieved by first generating the signature of the chemical reactions (reaction templates) associated with a given target molecule and then applying them to generate a list of potential reactants (Chen et al., 2020; Han et al., 2022; Kim et al., 2021; Segler et al., 2018). **Template-free** alleviates the template-based dependency by modeling RP as a prediction task. The product molecule is used as the input feature while the reactants are the labels. A learning algorithm is then used to learn the mapping between the input features and the labels (Liu et al., 2017; Karpov et al., 2019; Schwaller et al., 2019). **Semi-template-based** RP model RP as a two-step process involving (i) Product to Sython (P2S): the product molecule is broken into synthons(molecular segments) conforming with the reaction centers, and (ii) Sython to Reactant (S2R): the reactants for each synthon are then generated separately. Generally, the first step can be modeled as link prediction while the second uses several methods such as Seq2Seq, generative modeling with graphs, leaving group prediction, etc. (Somnath et al., 2021; Wang et al., 2020a; Shi et al., 2020). These methods are increasingly appealing

to the research community due to their ability to model diverse chemical relationships while also being interpretable.

Template-based methods are interpretable since they can be mapped to known chemical reactions in the database, however, they are limited by the chemical template database. Template-free methods have leveraged the advances in NLP e.g. attention (Bahdanau et al., 2014) and transformers (Vaswani et al., 2017) contributing to competitive performance in RP. While they are not as interpretable as template-based RP methods, they require no chemical template database. Template-free methods also require a database of chemical reactions in the form of product-reactant mappings. In addition, they are prone to generating invalid molecules due to their generative nature. Semi-template-based methods are interpretable and can model diverse chemical relationships while also being interpretable. While this class of RP design is increasingly appealing to the research community it inherits the limitations of molecular graph modeling as well as the sequence to sequence modeling as the second step is generative. Figure 2 summarizes the current AI-based RP methods.

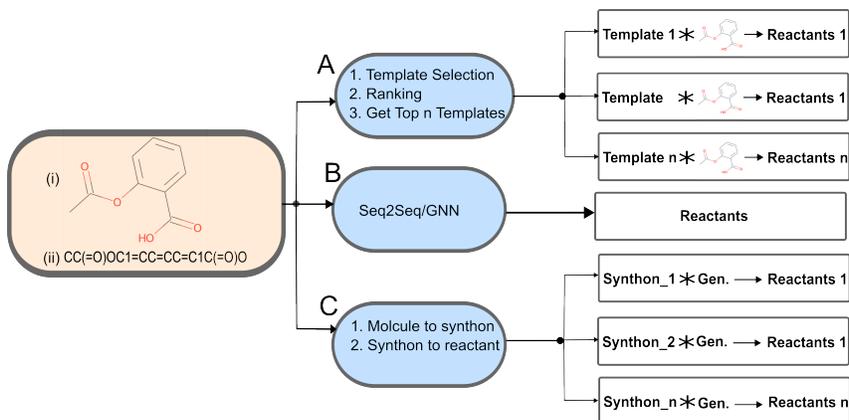


Figure 2: RP planning methods summary. A: template-based - generates a list of potential reactants by applying the predicted template, B: template-free - maps the product molecule to the reactants using Generative Models (Gen.) e.g. Seq2Seq, C: semi-template-based - breaks the product molecule into synthons and then generates the reactants for each synthon. The input can be Molecular Graph (i) or SMILES (ii)- the two are equivalent.

2 Background

2.1 Retrosynthetic Planning Modeling

RP can be represented formally as a Markov Decision Process (MDP) $\mathcal{R} = \{S, A, P, R, \sigma\}$. The search space component $s \in S$ is a set of all chemicals in the organic chemistry with the terminal states corresponding to the building block molecules or molecules that cannot be validly decomposed further. An action $a \in A$ is a chemical reaction which transform the product molecule $s \in S$ into the reactant set $\{s'_1, s'_2, \dots, s'_n\} \in S$. The transition probability $P : S \times A \rightarrow 2^S$ is a one-step chemical reaction (one-step retrosynthesis) that accepts the product molecule s and the action a and then returns the reactant set. The reward function $R : S \times A \rightarrow \mathbb{R}$ is the cost single-step chemical reaction, while σ includes a set of reaction parameters e.g. reagents that control the single-step chemical reaction. The mapping $P : S \times A \rightarrow 2^S$ is one-to-many mapping and in multi-step RP where there are several one-step retrosyntheses, the target product molecule $M_0 \in S$ is recursively decomposed into reactant molecules until the terminal state e.g. a given depth is reached or all building block molecules are found. Given that product-reactant pairs are not unique, the transition probability P is non-deterministic and thus can result in exponential branching which is a major computational challenge in RP.

2.2 Definitions

SMILES. Simplified Molecular Input Line Entry System (SMILES) (Weininger, 1988) is a chemical coding standard that is used to represent a molecule as a string of characters. Each atom, bond, ring and other molecular properties are represented by a unique character. For example, SMILES for water (H_2O) is **O** while that of ethanol (C_2H_6O) is **CCO**. SMILES represent atoms by their symbols e.g. **C** for Carbon. Hydrogen atoms are not explicitly represented as they are inferred from the valency of the atoms. The standard also includes mechanisms of representing other molecular features such as bond types - single (-) or (), double (=) and triple (#), aromaticity, chirality and ring structures.

SMARTS and Reaction Template. SMILES arbitrary target specification (SMARTS) is also a chemical coding standard that is used to represent a subgraph pattern of a molecule. It is an extension of SMILES and every SMILES is a valid SMARTS. SMILES are mostly used to encode general molecular subgraph patterns which are key in chemical reactions, as such, they can be used to define the template of a chemical reaction. A reaction template is a subgraph pattern that is used to define a chemical reaction. A chemical reaction must satisfy the reaction template to be valid. SMARTS encoding follows the *reactant_pattern* \gg *product_pattern* format where the *reactant_pattern* is the left-hand side of the reaction and the *product_pattern* the right-hand side. As a concrete example, **[#6:1][O:2] \gg [#6:1]=[O:2]** is a SMART-based reaction template which takes any Carbon (atomic number 6 and the first atom in the reaction list) which is singly bonded to the Oxygen (second atom) and gives a product having double bonds (=). There are many reactions which can be derived from this template. Simply put, a reaction template defines subgraph patterns for a chemical reaction, thus, **a chemical reaction is an instance of a reaction template**. In template-based RP, reaction templates are defined as SMARTS.

Molecular Fingerprints. Molecular fingerprints are fixed-length binary bits that are used to represent a molecule. They are used to encode molecular features such as the presence of a particular atom, bond type, ring structure, chirality, substructure pattern, functional groups, unique ions, leaving groups, etc. Common fingerprint encoding methods include Extended-Connectivity Fingerprints (ECFP) (Rogers & Hahn, 2010) - topological fingerprint method encoding structural features - and Morgan Fingerprints (Rogers & Hahn, 2010) - a circular fingerprint method encoding key molecular properties by encoding the subgraph patterns of the molecule. Other fingerprint encoding methods are based on count vectors, hashing, etc.

Molecular Graph. A molecular graph is a graph representation of a molecule. It is composed of *atoms* and *bonds* which are represented by the nodes and edges respectively. The atoms are connected by the bonds. Formally, a molecular graph is a tuple $G = (V, E)$ where V is the set of *atoms* and E is the set of *bonds*. Each atom $v \in V$ is a tuple $v = (a, x)$ where a is the atom type and x is the atom feature vector. Each bond $e \in E$ is a tuple $e = (v_i, v_j, b)$ where v_i and v_j are the atoms connected by the bond b . GNN (Kipf & Welling, 2016) is a popular DL method which can be used to encode molecular graphs. GNNs are composed of several layers and each layer is composed of a message passing and a readout function. The message-passing function updates the node features while the readout function aggregates the node features to give the graph features. The graph features can then be used to predict the target molecule or other molecular properties.

One-step and Multi-step Retrosynthesis. A one-step retrosynthesis involves a single chemical reaction where given a target molecule, the objective is to design a valid set of reactants that can lead to the target molecule e.g. in Figure 1, M_0 is the target molecule while $\{M_1, M_2, M_3\}$ is the reactant set which obtained by through one-step retrosynthesis via reaction R_0 . Simpler molecules can be designed through a single-step retrosynthesis, however, complex molecules require a series of one-step retrosyntheses: multi-step retrosynthesis i.e. in Figure 1, M_0 can be designed through the left or right sub-tree. The right sub-tree is composed of two one-step retrosyntheses (multi-step) as there are two reactions - R_1 and R_3 - which are required to design M_0 . Most molecules require a series of one-step retrosyntheses to be designed.

Synthon. Given an atom-mapped product molecule and reactants, the changes in bonds and atoms can identified automatically between them. This is achieved by breaking the product molecule into segments called synthons. A synthon is a subgraph of the product molecule which is connected by the reaction center. In semi-template RP a molecule is first broken into synthons according to the identified reaction centers. Then, reactants for each synthon are generated separately and combined to give the target molecule (Shi et al., 2020).

3 Existing AI-based RP Methods

In this section, we present a general overview of existing AI-based methods while also presenting their taxonomy in Figure 3.

- (i) **Expert-based:** These are the earliest RP methods designed according to the expert-engineered rules. RP research work under this category includes early models such as SYNLIMA (Johnson et al., 1989), WODCA (Gasteiger et al., 2000) and LHASA (Pensak & Corey, 1977). Generally, these methods code chemical reaction concepts such as leaving groups, functional groups, reaction centers etc. in a Knowledge Base (KB) and then attempting to match against the starting material provided by the user as a query. The models marked the shift from the manual to the computer-aided RP. Despite this, however, they are limited by scalability, generalization and the KB coding. This is despite the recent resurgence of modern rule-based methods such as Chen & Baldi (2009). The limitations and increase in computing power have inspired the design of new AI-based RP method: template-based, template-free and semi-template-based.
- (ii) **Template-based:** Here, the reaction rules are not predefined as expert-based systems but rather extracted automatically from a database of chemical reactions. This is achieved by first generating the reaction template associated with a given target molecule and then apply them to generate a list of potential reactants. The template selection can be achieved by a learning model such as MLP neural network (Chen et al., 2020; Han et al., 2022) or by a template similarity measure (Coley et al., 2017). Currently, several template-based retrosynthetic models have been proposed leading to strong performance on several benchmark datasets. These include Retro* (Chen et al., 2020) an A*-like inspired tree search method with neural evaluation to select the most promising RP path. In Retro* the search tree is modelled as an AND-OR tree whereby each reaction node (OR node) can lead to a target molecule that requires all its children (AND nodes). Similar to Retro*, Kim et al. (2021) also uses an AND-OR tree to model the RP. However, in contrast to Retro*, their model learns to improve the policy by self-play - a concept from Reinforcement Learning (RL) where the model learns to improve its performance by playing against itself (Silver et al., 2017). Prior to the two RP models, (Segler et al., 2018) proposed a Monte Carlo Tree Search (MCTS) based model called 3N-MCTS which is a multi-step RP model that uses 3 deep neural networks (DNN) to model the expansion policy, filter policy and value function. The expansion policy is used to select the most promising templates and the filter policy is to select the most *realistic* reactions following after template application. The value network is used as an evaluation function during the tree search rollout phase. While 3N-MCTS recorded strong RP results it was limited by the tree representation scheme. Each node in the tree was modeled to include more than one molecule which proved more computationally expensive and limited their molecular design choice. This limitation was addressed by Kishimoto et al. (2019) which first showed that the AND-OR tree is an apt molecular representation scheme for RP. This design choice was adopted by (Chen et al., 2020; Kim et al., 2021; Han et al., 2022); Han et al. (2022) proposed a GNN-based model with route cost optimization. GNN-based models such as Han et al. (2022) have recorded strong performance due to their ability to model molecules as graphs which are more chemically aware compared to fingerprint methods e.g. Chen et al. (2020); Kim et al. (2021). Other notable template-based RP include template similarity ranking (Coley et al., 2017), template scoring and ranking based on the symbolic features (Segler & Waller, 2017), two-step reaction with reaction group prediction and molecule generation with DNN (Baylon et al., 2019), tree search with DL evaluation (Schreck et al., 2019), MCTS with self-playing policy improvement (Koch et al., 2019) and optimizing one-step RP with RL (Liu et al., 2023a). Overall, strong template-based RP methods are based on tree search with DL evaluation algorithmic design. While template-based methods are interpretable, they can only be mapped to known chemical reactions in the database thus limiting their generalization. This limitation has inspired the design of template-free methods.
- (iii) **Template-free.** Here, the RP problem is modeled as a prediction task. The product molecule is used as the input feature while the reactants are the labels. A learning algorithm is then used to learn the mapping between the input features and the labels. Based on this, the RP problem can be designed as a classification task where the labels are the reactants. Alternatively, it can be designed

as a machine translation (a sequence-to-sequence) task where the product is the source language while the reactants are the target language. Many strong template-free RP methods are based on sequence-to-sequence models (Sutskever et al., 2014) e.g. LSTM Seq2Seq (Schwaller et al., 2019), LSTM Seq2Seq with attention (Liu et al., 2017) and Seq2Seq with the transformer architecture (Karpov et al., 2019). While the transformer-based models have recorded state-of-the-art performance in the last five years due to the attention mechanisms (Bahdanau et al., 2014) to model inter-dependencies between the input tokens, they are limited by chemical awareness as SMILES input tokens are treated as characters and not chemical entities. This limitation has led to increased adoption of molecular graph-based RP modeling including graph translation model with edge attention (Xie et al., 2022), higher-order molecular graphical reconstruction (Jin et al., 2020) and end-to-end learning with reaction metadata (Coley et al., 2019b; Xie et al., 2022). Similar to template-based methods, GNN template-free methods have recorded better performance and are more interpretable compared to the sequence-to-sequence models. Despite this, template-free models are constrained by their generative nature and thus prone to generating invalid molecules: whether seq2seq or GNN-based.

- (iv) **Semi-template.** Template-based methods are limited by the chemical template database. Template-free methods are limited by awareness and are prone to generating invalid molecules due to their generative nature. Semi-template RP attempts to some of these challenges by modeling RP as a two-stage process: (i) Product to Synthon (P2S) and (ii) Synthon to Reactant mapping. In the first stage, the product molecule is first broken into synthons according to identified reaction centers. This phase can be formulated as a link prediction problem (Daud et al., 2020; Rossi et al., 2021). In the second stage, the reactants for each synthon are then predicted separately and then combined to give the reactant set. This phase can be solved by deep generative graphs, Seq2Seq or template-based methods. As an example, early semi-template-based work was presented by Shi et al. (2020) which proposed a product-synthon, synthon-reactant model. Product-synthon step was modeled as a link prediction between bonds and atoms using Relational GNNs (Schlichtkrull et al., 2018) while the synthon-reactant step was modeled as a graph generation problem constrained as MDP. In Yan et al. (2020) the authors followed a similar design philosophy, however, the second stage was modeled using a transformer (Vaswani et al., 2017) based on SMILES. The last two laid the design pipeline for follow-up work: (i) Somnath et al. (2021); a product for reactant center prediction in the first stage is followed by leaving group prediction corresponding to each synthon segment, and (ii) Wang et al. (2020a); modeled both the first and second stages as machine translation while introducing novel methods for selecting potential reactants in the second phase.

4 AI-methods Overview.

A summary of template-based, template-free and semi-template methods is presented in Table 1. Key AI-based methods applied include (i) Tree Search, (ii) Proof Number Search (PNS), (iii) Machine translation (iv) Graph Neural Networks and (v) Reinforcement Learning.

Tree search. Tree search algorithms are effective in solving combinatorial optimization problems. A*-like methods (Liu et al., 2023a; Chen et al., 2020; Kim et al., 2021; Han et al., 2022) and Monte Carlo Tree Search (MCTS) (Koch et al., 2019; Segler et al., 2018) dominate template-based RP with state-of-the-art performance. A*-like methods are inspired by the A* search which is characteristically *optimal, complete and effective* including a cost function heuristic which helps to find paths leading to the goal faster.

MCTS incrementally builds a search tree while effectively balancing the exploitation and exploration (Kocsis & Szepesvári, 2006) in the search space. The algorithm is composed of four steps: (i) selection, (ii) expansion, (iii) simulation and (iv) backpropagation. In RP during the selection phase, the most promising reaction template is selected. The expansion phase involves applying the selected template to the product molecule to generate the reactant set. In the simulation phase, chemical templates are sampled from the policy function and then applied until the terminal state or given depth is reached. The backpropagation phase involves updating the Q values and the number of visits for each node in the tree with the reward e.g. +1 if the path leading to the terminal state includes valid building block molecules or -1 otherwise. Silver et al.

Table 1: Summary of the current AI-based RP methods.

| | <i>Novelty</i> | <i>Multi-step</i> | <i>Fingerprints</i> | <i>SMILES/SMARTS</i> | <i>USTPO</i> | <i>Accuracy</i> | <i>Software</i> |
|-------------------------|--------------------------------------|-------------------|---------------------|----------------------|--------------|-----------------|--------------------|
| TEMPLATE-BASED | | | | | | | |
| Liu et al. (2023a) | One-step optimization with RL | x | x | x | x | 86-99 | Torch, RDChiral |
| Lee et al. (2023) | Rule generalization DL memory | x | - | - | - | 23-93 | RXNMapper |
| Han et al. (2022) | Reaction cost GNN optimization | x | x | x | x | 80-91 | Torch, RDChiral |
| Yu et al. (2022) | Goal driven Actor Criticv RL | x | - | - | x | 58-88 | - |
| Balzerani et al. (2022) | MCTS with DL & heuristics | x | - | - | - | 77-83 | RetroPath |
| Kim et al. (2021) | A*-like with DNN and self-play | x | x | x | x | 57-96 | Torch, RDChiral |
| Chen et al. (2020) | A*-like tree search with DNN | x | x | x | x | 79-86 | Torch, RDChiral |
| Wang et al. (2020b) | Self-play policy MCTS | x | - | - | - | 55-76 | ASKCOS |
| Kishimoto et al. (2019) | PNS with heuristic | x | x | x | x | 25-81 | TF1, RDKit |
| Koch et al. (2019) | MCTS with similarity ranking | x | - | - | - | 77-83 | RDKit, RetroPath |
| Baylon et al. (2019) | Template prediction with DHN | x | x | x | x | 58-82 | - |
| Schreck et al. (2019) | Self-play with DL evaluation | x | x | - | - | 64-73 | TF1, RDChiral |
| Dai et al. (2019) | GNN with sampling | x | x | x | x | 64-93 | Torch, RDKit |
| Delépine et al. (2018) | BFS template similarity ranking | x | x | x | - | 63-81 | RDKit |
| Segler et al. (2018) | MCTS with DNN Policy Rollout | x | x | x | x | 85-95 | - |
| Segler & Waller (2017) | Template ranking with rules | x | - | x | - | 62-99 | Keras, RDChiral |
| Coley et al. (2017) | Template similarity ranking | x | x | x | x | 52-92 | RDKit, RDChiral |
| SEMI-TEMPLATE | | | | | | | |
| Chen et al. (2023) | Sequential optimal S2R completion | - | - | x | x | 63-91 | Sklearn, Torch |
| Somnath et al. (2021) | P2S and leaving group predictions | - | - | x | x | 63-88 | - |
| Shi et al. (2020) | P2S & S2R w/ MG transformer | - | - | - | x | 61-88 | Torch |
| Yan et al. (2020) | P2S & S2R w/ transformer | - | - | x | x | 62-80 | Torch, NMT |
| Wang et al. (2020a) | Both P2S & S2R as transformers | - | - | x | x | 64-86 | Torch, RDKit |
| TEMPLATE-FREE | | | | | | | |
| Liu et al. (2023b) | End-to-End In-context Learning | - | - | x | x | 60-94 | Torch, RDKit |
| Wan et al. (2022) | Transformer molecular attention | - | x | x | x | 57-99 | RDChiral |
| Irwin et al. (2022) | Transformer-based RP model | - | x | x | x | 53-95 | Deepspeed, Torch |
| Lin et al. (2022) | RP with Energy-based model | - | x | x | x | 52-94 | Torch, RDChiral |
| Tu & Coley (2022) | Local & global MG feature encoding | - | x | x | x | 52-85 | NMT, Torch |
| Xie et al. (2022) | GNN with edge attention | x | - | x | x | 50-72 | Torch, DGL |
| Sacha et al. (2021) | Graph Se2Seq w/ MG attention | - | - | x | x | 48-91 | Torch, RDKit |
| Sun et al. (2020) | Energy-based Seq2Seq with GNN | - | - | x | x | 53-85 | - |
| Schwaller et al. (2019) | Reactant-product attention | - | - | x | x | 76-92 | IBM RXN, NMT |
| Lin et al. (2020) | MCTS with transformer evaluation | x | - | x | x | 54-81 | TF1, RDKit |
| Coley et al. (2019b) | End-to-end Weisfeiler-Lehman GNN | - | - | - | x | 85-93 | TF1, RDKit |
| Zheng et al. (2019) | Self-correcting DNN with transformer | - | - | x | x | 74-81 | - |
| Karpov et al. (2019) | Template-free RP using transformers | - | - | x | x | 40-69 | TF1, NMT |
| Schwaller et al. (2018) | End-to-end seq2seq BiLSTM | - | - | x | x | 80-87 | RDKit |
| Bradshaw et al. (2018) | Electron Gated GNN | - | - | x | x | 70-94 | RDKit |
| Liu et al. (2017) | Product-reactant Seq2Seq | - | - | x | x | 37-70 | TF1 |
| Wei et al. (2016) | Fingerprint fusion with DNN | - | x | x | - | 85-86 | Sklearn, Hyperplot |

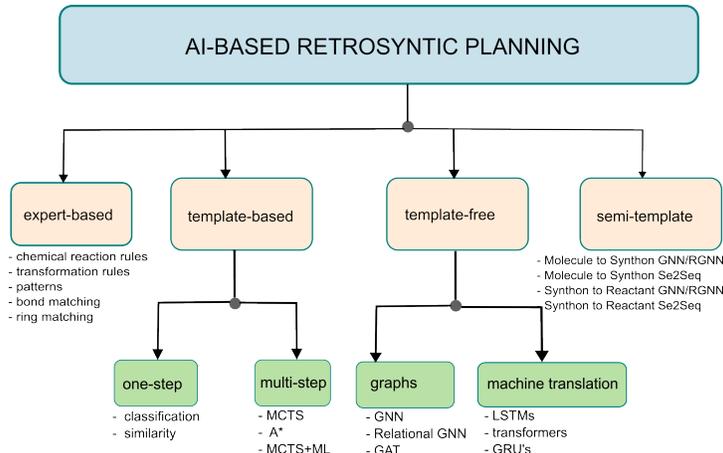


Figure 3: Taxonomy of current of AI-based retrosynthesis.

(2016) presented an effective mechanism of combining MCTS with DL as such the rollout phase of MCTS is increasingly being replaced with a prediction model. As reported, this design choice is influenced by the sub-optimality of the random policy. Similar designs have been employed by RP methods such as Koch et al. (2019). Current MCTS includes selection, expansion and backpropagation with DL as rollout policy. The pseudocode for this pipeline is presented in Algorithm 1.

Algorithm 1: MCTS with DL evaluation.

```

1 root ← tree(state);
2 while time do
3   node ← select_node(root);
4   node ← expand(node);
5   v ← DL_eval(node);
6   back_propagate(node, v);
7 return best_child(root);
  
```

Apart from MCTS and A*-like methods, other tree search algorithms that have been applied to RP include Breadth-First Search (BFS) and Depth First Search (DFS). BFS is a tree traversal algorithm that searches for all the nodes *level-by-level*. It does not require any heuristic or value function to determine the node visit but rather *uniformly*. In contrast, DFS visits all the nodes by going as deep as possible and then backtracks if the search value is not found. DFS and BFS methods have recorded very promising results in Delépine et al. (2018) while also playing a key role as baseline models Han et al. (2022); Kim et al. (2021); Chen et al. (2020); Segler et al. (2018).

Proof Number Search (PNS). PNS can be used to model the game state in a perfect information game using an AND-OR tree structure. During the playouts, the maximizing player is mapped to the AND tree nodes and the minimizing player to the OR nodes. For each node in the game tree, the proof and disproof numbers are recorded and subsequently updated. During the expansion, nodes with the least proof value and nodes with the least disproof values are selected for OR and AND nodes respectively. RP can be represented as an AND-OR tree such that given the target molecule, it can be synthesized by either the reaction node (OR) and each reaction node requires all its children (AND). Many sequential and parallel variants exist including PPNS, PN2 and DFPNS. Modeling RP with an AND-OR tree is sound and effective, thus, recent methods are increasingly employing this modeling approach (Kishimoto et al., 2019; Kim et al., 2021; Chen et al., 2020) recording state-of-the-art results. These implementations are reinforced with DL evaluation functions or policies.

Machine translation. These algorithms are based on the encoder-decoder architecture where the input sequence is encoded into a fixed-length vector by the encoder and then decoded into the output sequence by the decoder (Sutskever et al., 2014) (Seq2Seq). The Seq2Seq models were proposed to solve the machine translation problem in NLP. However, through transfer learning, they have been increasingly adopted to solve other hard NLP problems. Further, following the introduction of the attention mechanism (Bahdanau et al., 2014; Vaswani et al., 2017), Seq2Seq models have gained more traction due to their ability to model long-range dependencies leading to better performance across many NLP tasks. Seq2Seq is a primary algorithmic design in template-free RP. Existing RP methods are based on three key approaches including Seq2Seq LSTM (Schwaller et al., 2019), Seq2Seq with attention (Lin et al., 2020) and Seq2Seq with transformer (Karpov et al., 2019).

Graph Neural Networks (GNNs). This is a class of neural networks suited to learn and generalize on graph-structured data (Kipf & Welling, 2016). Molecules are inherently graphical and as such GNNs are increasingly being used to model them. While the input feature in RP can be fingerprints, SMILES or molecular graphs, the latter is more chemically aware and as such is gaining more attention in all categories of RP methods e.g. template-based (Han et al., 2022), template-free (Xie et al., 2022) and semi-template (Shi et al., 2020).

Reinforcement Learning (RL). RL algorithms learn by interacting with the environment. The agent learns policy or value function which maximizes the reward by taking actions in the environment (Sutton & Barto, 2018). While not very common in RP, possibly due to convergence and reward design issues, some studies have explored this approach. Liu et al. (2023a) proposed an RL-based algorithm composed of two policy network networks, a synthesizability value network and a cost value network. The policy networks are used in the planning phase while the value networks in the update phase to evaluate the most promising targets. Koch et al. (2019) on the other hand proposed an RL-based algorithm similar to AlphaZero (Silver et al., 2017) which combines MCTS with DL while improving the policy through self-play. A closely related work by Yu et al. (2022) employed a goal-driven actor-critic RL algorithm to solve the RP by learning the policy and value function.

Table 1 shows a summary of the current template-based and template-free methods. The abbreviations in Table are as follows; The Novelty summarizes the key idea in the paper, Multi-step: whether the method is multi-step or one-step, Fingerprint: whether the method uses fingerprints as input features, SMILES/SMARTS: whether the method uses SMILES or SMARTS as input features, USPTO: whether the method uses the USPTO dataset or its variant, Accuracy: this corresponds to the performance ranges of the method according to Top @K accuracy where the value of K ranges across different methods, Software: the software used to implement the method. The abbreviations used in the table are as follows: TF1: TensorFlow V1, GNN: Graph Neural Network, DNN: Deep Neural Network, USPTO: US Patent Office Dataset (Lowe, 2012), RDChiral (Coley et al., 2019a), DHN:Deep Highway Network Srivastava et al. (2015), ZINC (Irwin et al., 2012), MG: Molecular Graph, P2S: Product to Synthon, S2R: Synthon to Reactant, RXNMapper (Schwaller et al., 2021), AGREDA (Blasco et al., 2021), RetroPath: RetroPath RL (Delépine et al., 2018), NMT: OpenNMT (Klein et al., 2018), deepspeed: (Aminabadi et al., 2022),

On the methodology, template-free methods are limited by their generative nature however the ability to include chemical awareness and modeling global and local molecular-based attention such as in Wan et al. (2022) have shown that their performance can be improved. Semi-template methods are increasingly gaining attention due to their diverse ability to include chemical awareness in the design process. Since their inception by Shi et al. (2020), they have increasingly recorded better performance up to date. Somnath et al. (2021) and (Shi et al., 2020) even reported instances where such methods recorded better performance than template-based. Generally, template-based are more stronger and interpretable. At the feature level, while fingerprint encoding has dominated template-based methods, other molecular representations such as InChI (Heller et al., 2015) have also been used - though not very common as only a single work has reported it (Koch et al., 2019). Top@K accuracy is the primary evaluation metric used by the majority of current RP methods. It involves computing the percentage of the target molecules that are included in the top K predictions. About the datasets, the majority of the template-based methods used the USPTO (Lowe, 2012) database. Template-free methods, on the other hand, used a subset of USPTO (USPTO-5OK) or custom datasets. This is similar to semi-template methods. The USPTO is composed of more than 1.8M chemical

reactions from 1976 to 2016. Other databases have also been used by some methods including DrugBank (Wishart et al., 2008), LASER (Wishart et al., 2008), Retrorules (Duigou et al., 2019) and literary texts (Delépine et al., 2018). Major building block molecule databases include ZINC (Irwin et al., 2012), eMolecules³, and Reaxys⁴,

5 Open challenges

- (i) **Metrics.** There are no widely agreed metrics for evaluating the performance of AI-based RP methods. While shorter routes are generally preferred, other factors such as the ease of synthesis from the proposed plans or the availability of reagents to catalyze the reactions are also key. Based on this constraint, some longer planning routes can be preferred over the shorter ones. Also, in template-free retrosynthesis, different scoring metrics have been proposed by different methods e.g. Pearson coefficient and top@K accuracy. These evaluation variations currently, present a challenge while conducting an in-depth comparative analysis.
- (ii) **End-to-end design with metadata.** Reagents are chemical compounds which catalyze chemical reactions. The majority of the existing template-based and template-free solvers do not include the reagents in their design pipeline. Given that the reagents contain rich metadata information, their inclusion in the the design process has two key benefits (i) soundness: every chemical reaction is composed of reactants, *reagents* and products, (ii) robust DL-based evaluation models: DL-based evaluation networks can be trained on the richer dataset.
- (iii) **Databases.** USPTO is the major publicly available reaction database for both template-based and template-free methods. Currently, over 95% of the existing methods are trained and tested on this dataset. The creation of new databases that can allow for true *out of sample* validation and testing is thus imperative. Importantly, such new datasets can also include a uniform evaluation metric for both the template-based and template-free methods. In consequence, this is likely to help address metric variation inconsistencies as previously discussed. Additionally, other researchers have proposed retrosynthetic solvers based on custom datasets e.g. extraction from the literary text. While this approach is a very promising research frontier, it is likely to compound the comparative analysis issues.
- (iv) **Sparsity.** Molecular fingerprints encode the presence of subgraph features in a molecule as a sparse row vector. Such, among others may include the presence of functional groups, rings or bond types. The existing template-based and template-free methods are dominated sparse molecular fingerprints method - ECFPF and Morgan Fingerprints- as such they inherit the sparsity issues such as the inability to capture *richer* information and increased computational complexity for longer fingerprints.
- (v) **Natural product synthesis.** Synthesis of naturally occurring products though existing RP methods are still a challenge. This is attributed to the fact that the majority of the existing RP methods are limited by the chemical reaction databases. This limitation has been reported in a number of RP studies including Lin et al. (2020) and Segler et al. (2018). In contrast, in bio-retrosynthesis, several successful natural product syntheses have been reported (Lin et al., 2019; Zheng et al., 2022).

6 Future Research Directions

In this section, we present a summary of the key future research directions. Firstly, the **template adaptation**. The current template-based methods use DL greedily selects the most promising templates which are consequently applied to a target molecule to generate the reactant set. While this approach has recorded strong and very promising results, designing RP methods that not only select the templates to apply but also adapt the *template to the molecule* or *template to the route length* can lead to better planners. These adaptations can allow the RP solver to learn reusable templates that can be applied to future similar states in the design process.

³<https://www.emolecules.com/>

⁴<https://www.elsevier.com/solutions/reaxys>

Secondly, representation of **Molecules as Graphs** is a very promising research frontier as molecules are inherently graphical. In addition, graph-based representations such as GNNs are more chemically aware and thus can capture richer features compared to molecular fingerprints. In addition, the design choice can also help address the sparsity issues associated with molecular fingerprints. Thirdly, **Stronger RL Policies**, since Silver et al. (2017) the random simulation phase has been replaced with a prediction model as a result of the sub-optimality, in consequence of the sub-optimality. While some existing template-based RP methods have increasingly adopted these new design choices there is a need to design models with stronger policies. Fourthly, **On generative Modelling**, the Seq2Seq has inspired the development of strong template-free methods such as (Lin et al., 2020). While the transformer Seq2Seq model is currently phasing out the LSTMs, the current limitation is still chemical awareness. The transformer can model the long-range dependencies in the input sequence, however, it is still limited by the SMILES-based character-level attention. While the model proposed by Wan et al. (2022) is a step in the right direction towards modeling Seq2Seq with molecular attention, it is still limited by expressiveness.

Fifthly, **Active and transfer learning** is a key design choice. Molecules share a lot of chemical similarities. The classical reaction USPTO database, for instance, contains over 1.8M reaction samples. Many of these samples have similar rings, bonds or functional group patterns which once learned, can be *transferred* or *generalized* to help learn faster and reduce the search space. These properties can be achieved by transfer learning or active learning methods, or by a combination of the two (active transfer learning).

Lastly, **End-to-end RP, parallelization and multi-task learning** are performant design approaches in many machine learning problems as such can also be applied to RP. End-to-end learning with products, reagents, reactants and catalysts in the design pipeline is key to designing robust RP methods. The majority of the existing methods, however, do not explicitly model the reagents and catalysts RP components. The majority of the existing methods are not parallelized by design. Designing parallel RP methods has two main advantages (i) improved performance: the workload is distributed and (ii) richer representations: each parallel instance can learn local representations which can be *shared* globally. The current RP methods are embarrassingly parallel, however, *less than 1%* are parallelized by design. Multi-task learning is ubiquitous in modeling computational biology problems. They are generally characterized by *the ability to optimize more than one objective*. In template-based RP for instance a template filter network is trained separately from the template selection policy yet the two can be learned by optimizing a joint objective function. Similarly, template-molecule and template-route adaptation can also be learned in an end-to-end fashion instead of separating the two. Such models are easier to train and allow for the *sharing* of the learned representations leading to more robust models.

7 Conclusion

In this work, we presented a comprehensive survey of the current AI-based retrosynthesis methods: expert systems, template-based, template-free and semi-template. We presented a taxonomy of the current method according to their design philosophy. We also presented a summary of the key AI-based methods applied in retrosynthesis. We then presented the current open challenges and future research directions. We hope that this work will help researchers understand the current state of the art and also help in the design of stronger retrosynthetic planners.

Ethical Statement

We have no ethical issues to report about this research work.

References

- Reza Yazdani Aminabadi, Samyam Rajbhandari, Ammar Ahmad Awan, Cheng Li, Du Li, Elton Zheng, Olatunji Ruwase, Shaden Smith, Minjia Zhang, Jeff Rasley, et al. Deepspeed-inference: enabling efficient inference of transformer models at unprecedented scale. In *SC22: International Conference for High Performance Computing, Networking, Storage and Analysis*, pp. 1–15. IEEE, 2022.
- Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*, 2014.
- Francesco Balzerani, Daniel Hinojosa-Nogueira, Xabier Cendoya, Telmo Blasco, Sergio Pérez-Burillo, Iñigo Apaolaza, M Pilar Francino, and José Ángel Rufián-Henares. Prediction of degradation pathways of phenolic compounds in the human gut microbiota through enzyme promiscuity methods. *NPJ systems biology and applications*, 8(1):1–9, 2022.
- Javier L Baylon, Nicholas A Cilfone, Jeffrey R Gulcher, and Thomas W Chittenden. Enhancing retrosynthetic reaction prediction with deep learning using multiscale reaction classification. *Journal of chemical information and modeling*, 59(2):673–688, 2019.
- Telmo Blasco, Sergio Pérez-Burillo, Francesco Balzerani, Daniel Hinojosa-Nogueira, Alberto Lerma-Aguilera, Silvia Pastoriza, Xabier Cendoya, Ángel Rubio, María José Gosalbes, Nuria Jiménez-Hernández, et al. An extended reconstruction of human gut microbiota metabolism of dietary compounds. *Nature communications*, 12(1):4728, 2021.
- John Bradshaw, Matt J Kusner, Brooks Paige, Marwin HS Segler, and José Miguel Hernández-Lobato. A generative model for electron paths. *arXiv preprint arXiv:1805.10970*, 2018.
- Binghong Chen, Chengtao Li, Hanjun Dai, and Le Song. Retro*: learning retrosynthetic planning with neural guided a* search. In *International Conference on Machine Learning*, pp. 1608–1616. PMLR, 2020.
- Jonathan H Chen and Pierre Baldi. No electron left behind: a rule-based expert system to predict chemical reactions and reaction mechanisms. *Journal of chemical information and modeling*, 49(9):2034–2043, 2009.
- Ziqi Chen, Oluwatosin R Ayinde, James R Fuchs, Huan Sun, and Xia Ning. G2retro as a two-step graph generative models for retrosynthesis prediction. *Communications Chemistry*, 6(1):102, 2023.
- Connor W Coley, Luke Rogers, William H Green, and Klavs F Jensen. Computer-assisted retrosynthesis based on molecular similarity. *ACS central science*, 3(12):1237–1245, 2017.
- Connor W Coley, William H Green, and Klavs F Jensen. Rdcchiral: An rdkit wrapper for handling stereochemistry in retrosynthetic template extraction and application. *Journal of chemical information and modeling*, 59(6):2529–2537, 2019a.
- Connor W Coley, Wengong Jin, Luke Rogers, Timothy F Jamison, Tommi S Jaakkola, William H Green, Regina Barzilay, and Klavs F Jensen. A graph-convolutional neural network model for the prediction of chemical reactivity. *Chemical science*, 10(2):370–377, 2019b.
- Elias James Corey. General methods for the construction of complex molecules. *Pure and Applied chemistry*, 14(1):19–38, 1967.
- Hanjun Dai, Chengtao Li, Connor Coley, Bo Dai, and Le Song. Retrosynthesis prediction with conditional graph logic network. *Advances in Neural Information Processing Systems*, 32, 2019.
- Nur Nasuha Daud, Siti Hafizah Ab Hamid, Muntadher Saadoon, Firdaus Sahran, and Nor Badrul Anuar. Applications of link prediction in social networks: A review. *Journal of Network and Computer Applications*, 166:102716, 2020.
- Jeffrey Dean. 1.1 the deep learning revolution and its implications for computer architecture and chip design. In *2020 IEEE International Solid-State Circuits Conference-(ISSCC)*, pp. 8–14. IEEE, 2020.

- Baudoin Delépine, Thomas Duigou, Pablo Carbonell, and Jean-Loup Faulon. Retropath2. 0: A retrosynthesis workflow for metabolic engineers. *Metabolic engineering*, 45:158–170, 2018.
- Joseph A DiMasi, Henry G Grabowski, and Ronald W Hansen. Innovation in the pharmaceutical industry: new estimates of r&d costs. *Journal of health economics*, 47:20–33, 2016.
- Thomas Duigou, Melchior Du Lac, Pablo Carbonell, and Jean-Loup Faulon. Retrorules: a database of reaction rules for engineering biology. *Nucleic acids research*, 47(D1):D1229–D1235, 2019.
- Johann Gasteiger, Matthias Pförtner, Markus Sitzmann, Robert Höllering, Oliver Sacher, Thomas Kostka, and Norbert Karg. Computer-assisted synthesis and reaction planning in combinatorial chemistry. *Perspectives in Drug Discovery and Design*, 20:245–264, 2000.
- Anna Gaulton, Anne Hersey, Michał Nowotka, A Patricia Bento, Jon Chambers, David Mendez, Prudence Mutowo, Francis Atkinson, Louisa J Bellis, Elena Cibrián-Uhalte, et al. The chembl database in 2017. *Nucleic acids research*, 45(D1):D945–D954, 2017.
- Peng Han, Peilin Zhao, Chan Lu, Junzhou Huang, Jiayang Wu, Shuo Shang, Bin Yao, and Xiangliang Zhang. Gnn-retro: Retrosynthetic planning with graph neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pp. 4014–4021, 2022.
- Stephen R Heller, Alan McNaught, Igor Pletnev, Stephen Stein, and Dmitrii Tchekhovskoi. Inchi, the iupac international chemical identifier. *Journal of cheminformatics*, 7(1):1–34, 2015.
- John J Irwin, Teague Sterling, Michael M Mysinger, Erin S Bolstad, and Ryan G Coleman. Zinc: a free tool to discover chemistry for biology. *Journal of chemical information and modeling*, 52(7):1757–1768, 2012.
- Ross Irwin, Spyridon Dimitriadis, Jiazhen He, and Esben Jannik Bjerrum. Chemformer: a pre-trained transformer for computational chemistry. *Machine Learning: Science and Technology*, 3(1):015022, 2022.
- Wengong Jin, Regina Barzilay, and Tommi Jaakkola. Multi-objective molecule generation using interpretable substructures. In *International conference on machine learning*, pp. 4849–4859. PMLR, 2020.
- Peter Y Johnson, Ilene Burnstein, John Crary, Martha Evans, and Tunghwa Wang. Designing an expert system for organic synthesis: the need for strategic planning. ACS Publications, 1989.
- Pavel Karpov, Guillaume Godin, and Igor V Tetko. A transformer model for retrosynthesis. In *International Conference on Artificial Neural Networks*, pp. 817–830. Springer, 2019.
- Junsu Kim, Sungsoo Ahn, Hankook Lee, and Jinwoo Shin. Self-improved retrosynthetic planning. In *International Conference on Machine Learning*, pp. 5486–5495. PMLR, 2021.
- Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.
- Akihiro Kishimoto, Beat Buesser, Bei Chen, and Adi Botea. Depth-first proof-number search with heuristic edge cost and application to chemical synthesis planning. *Advances in Neural Information Processing Systems*, 32, 2019.
- Guillaume Klein, Yoon Kim, Yuntian Deng, Vincent Nguyen, Jean Senellart, and Alexander M. Rush. Opennmt: Neural machine translation toolkit, 2018.
- Mathilde Koch, Thomas Duigou, and Jean-Loup Faulon. Reinforcement learning for bioretrosynthesis. *ACS synthetic biology*, 9(1):157–168, 2019.
- Levente Kocsis and Csaba Szepesvári. Bandit based monte-carlo planning. In *European conference on machine learning*, pp. 282–293. Springer, 2006.
- Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 2012.

- Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436–444, 2015.
- Seul Lee, Taein Kim, Min-Soo Choi, Yejin Kwak, Jeongbin Park, Sung Ju Hwang, and Sang-Gyu Kim. Read-retro: Natural product biosynthesis planning with retrieval-augmented dual-view retrosynthesis. *bioRxiv*, pp. 2023–03, 2023.
- Geng-Min Lin, Robert Warden-Rothman, and Christopher A Voigt. Retrosynthetic design of metabolic pathways to chemicals not found in nature. *Current Opinion in Systems Biology*, 14:82–107, 2019.
- Kangjie Lin, Youjun Xu, Jianfeng Pei, and Luhua Lai. Automatic retrosynthetic route planning using template-free models. *Chemical science*, 11(12):3355–3364, 2020.
- Min Htoo Lin, Zhengkai Tu, and Connor W Coley. Improving the performance of models for one-step retrosynthesis through re-ranking. *Journal of cheminformatics*, 14(1):1–13, 2022.
- Bowen Liu, Bharath Ramsundar, Prasad Kawthekar, Jade Shi, Joseph Gomes, Quang Luu Nguyen, Stephen Ho, Jack Sloane, Paul Wender, and Vijay Pande. Retrosynthetic reaction prediction using neural sequence-to-sequence models. *ACS central science*, 3(10):1103–1113, 2017.
- Guoqing Liu, Di Xue, Shufang Xie, Yingce Xia, Austin Tripp, Krzysztof Maziarz, Marwin Segler, Tao Qin, Zongzhang Zhang, and Tie-Yan Liu. Retrosynthetic planning with dual value networks. *arXiv preprint arXiv:2301.13755*, 2023a.
- Songtao Liu, Zhengkai Tu, Minkai Xu, Zuobai Zhang, Lu Lin, Rex Ying, Jian Tang, Peilin Zhao, and Dinghao Wu. Fusionretro: molecule representation fusion via in-context learning for retrosynthetic planning. In *International Conference on Machine Learning*, pp. 22028–22041. PMLR, 2023b.
- Daniel Mark Lowe. *Extraction of chemical structures and reactions from the literature*. PhD thesis, University of Cambridge, 2012.
- Gordon E Moore. Cramming more components onto integrated circuits. *Proceedings of the IEEE*, 86(1):82–85, 1998.
- David A Pensak and Elias James Corey. Lhasa—logic and heuristics applied to synthetic analysis. ACS Publications, 1977.
- David Rogers and Mathew Hahn. Extended-connectivity fingerprints. *Journal of chemical information and modeling*, 50(5):742–754, 2010.
- Andrea Rossi, Denilson Barbosa, Donatella Firmani, Antonio Matinata, and Paolo Merialdo. Knowledge graph embedding for link prediction: A comparative analysis. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 15(2):1–49, 2021.
- Mikołaj Sacha, Mikołaj Błaz, Piotr Byrski, Paweł Dabrowski-Tumanski, Mikołaj Chrominski, Rafał Loska, Paweł Włodarczyk-Pruszyński, and Stanisław Jastrzebski. Molecule edit graph attention network: modeling chemical reactions as sequences of graph edits. *Journal of Chemical Information and Modeling*, 61(7):3273–3284, 2021.
- Michael Schlichtkrull, Thomas N Kipf, Peter Bloem, Rianne van den Berg, Ivan Titov, and Max Welling. Modeling relational data with graph convolutional networks. In *European semantic web conference*, pp. 593–607. Springer, 2018.
- John S Schreck, Connor W Coley, and Kyle JM Bishop. Learning retrosynthetic planning through simulated experience. *ACS central science*, 5(6):970–981, 2019.
- Philippe Schwaller, Theophile Gaudin, David Lanyi, Costas Bekas, and Teodoro Laino. “found in translation”: predicting outcomes of complex organic chemistry reactions using neural sequence-to-sequence models. *Chemical science*, 9(28):6091–6098, 2018.

- Philippe Schwaller, Teodoro Laino, Théophile Gaudin, Peter Bolgar, Christopher A Hunter, Costas Bekas, and Alpha A Lee. Molecular transformer: a model for uncertainty-calibrated chemical reaction prediction. *ACS central science*, 5(9):1572–1583, 2019.
- Philippe Schwaller, Benjamin Hoover, Jean-Louis Reymond, Hendrik Strobelt, and Teodoro Laino. Extraction of organic chemistry grammar from unsupervised learning of chemical reactions. *Science Advances*, 7(15):eabe4166, 2021.
- Marwin HS Segler and Mark P Waller. Neural-symbolic machine learning for retrosynthesis and reaction prediction. *Chemistry—A European Journal*, 23(25):5966–5971, 2017.
- Marwin HS Segler, Mike Preuss, and Mark P Waller. Planning chemical syntheses with deep neural networks and symbolic ai. *Nature*, 555(7698):604–610, 2018.
- Chence Shi, Minkai Xu, Hongyu Guo, Ming Zhang, and Jian Tang. A graph to graphs framework for retrosynthesis prediction. In *International conference on machine learning*, pp. 8818–8827, 2020.
- David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484–489, 2016.
- David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dhharshan Kumaran, Thore Graepel, et al. Mastering chess and shogi by self-play with a general reinforcement learning algorithm. *arXiv preprint arXiv:1712.01815*, 2017.
- Vignesh Ram Somnath, Charlotte Bunne, Connor Coley, Andreas Krause, and Regina Barzilay. Learning graph models for retrosynthesis prediction. *Advances in Neural Information Processing Systems*, 34:9405–9415, 2021.
- Rupesh Kumar Srivastava, Klaus Greff, and Jürgen Schmidhuber. Highway networks. *arXiv preprint arXiv:1505.00387*, 2015.
- Ruoxi Sun, Hanjun Dai, Li Li, Steven Kearnes, and Bo Dai. Energy-based view of retrosynthesis. *arXiv preprint arXiv:2007.13437*, 2020.
- Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to sequence learning with neural networks. *Advances in neural information processing systems*, 27, 2014.
- Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.
- Zheng kai Tu and Connor W Coley. Permutation invariant graph-to-sequence model for template-free retrosynthesis and reaction prediction. *Journal of chemical information and modeling*, 62(15):3503–3513, 2022.
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.
- Yue Wan, Chang-Yu Hsieh, Ben Liao, and Shengyu Zhang. Retroformer: Pushing the limits of end-to-end retrosynthesis transformer. In *International Conference on Machine Learning*, pp. 22475–22490. PMLR, 2022.
- Xiaorui Wang, Jiezhong Qiu, Yuquan Li, Guangyong Chen, Huanxiang Liu, Ben Liao, Chang-Yu Hsieh, and Xiaojun Yao. Retroprime: A chemistry-inspired and transformer-based method for retrosynthesis predictions. 2020a.
- Xiaoxue Wang, Yujie Qian, Hanyu Gao, Connor W Coley, Yiming Mo, Regina Barzilay, and Klavs F Jensen. Towards efficient discovery of green synthetic pathways with monte carlo tree search and reinforcement learning. *Chemical science*, 11(40):10959–10972, 2020b.

- Jennifer N Wei, David Duvenaud, and Alán Aspuru-Guzik. Neural networks for the prediction of organic chemistry reactions. *ACS central science*, 2(10):725–732, 2016.
- David Weininger. Smiles, a chemical language and information system. 1. introduction to methodology and encoding rules. *Journal of chemical information and computer sciences*, 28(1):31–36, 1988.
- David S Wishart, Craig Knox, An Chi Guo, Dean Cheng, Savita Shrivastava, Dan Tzur, Bijaya Gautam, and Murtaza Hassanali. Drugbank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic acids research*, 36(suppl_1):D901–D906, 2008.
- Shufang Xie, Rui Yan, Peng Han, Yingce Xia, Lijun Wu, Chenjuan Guo, Bin Yang, and Tao Qin. Retrograph: Retrosynthetic planning with graph search. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 2120–2129, 2022.
- Chaochao Yan, Qianggang Ding, Peilin Zhao, Shuangjia Zheng, Jinyu Yang, Yang Yu, and Junzhou Huang. Retroxpert: Decompose retrosynthesis prediction like a chemist. *Advances in Neural Information Processing Systems*, 33:11248–11258, 2020.
- Yemin Yu, Ying Wei, Kun Kuang, Zhengxing Huang, Huaxiu Yao, and Fei Wu. Grasp: Navigating retrosynthetic planning with goal-driven policy. *Advances in Neural Information Processing Systems*, 35: 10257–10268, 2022.
- Shuangjia Zheng, Jiahua Rao, Zhongyue Zhang, Jun Xu, and Yuedong Yang. Predicting retrosynthetic reactions using self-corrected transformer neural networks. *Journal of chemical information and modeling*, 60(1):47–55, 2019.
- Shuangjia Zheng, Tao Zeng, Chengtao Li, Binghong Chen, Connor W Coley, Yuedong Yang, and Ruibo Wu. Deep learning driven biosynthetic pathways navigation for natural products with bionavi-np. *Nature Communications*, 13(1):1–9, 2022.