RETROG: RETROSYNTHETIC PLANNING WITH TREE SEARCH AND GRAPH LEARNING

Stephen Obonyo
LIASD,
Paris 8 University,
Paris, France.
{sobonyo}@strathmore.edu

Nicolas Jouandeau
LIASD,
Paris 8 University,
Paris, France.
{n}@up8.edu

Dickson Owuor
SCES,
Strathmore University,
Nairobi, Kenya.
{dowuor}@strathmore.edu

ABSTRACT

Retrosynthesis Planning (RP) is one of the challenging problems in organic chemistry. It involves designing target molecules using compounds which are commercially available or easy to synthesize by following a series of backward steps. While the conventional RP methods were majorly expert-based, successful computer-aided RP methods have emerged in the recent past. This success is critical to the development of new drugs and the synthesis of target compounds in material science and agrochemicals domains. In this paper, we present an RP model called RetroG. Its design is based on tree search with a Graph Neural Network (GNN) as a value function. The model adapts successful reaction templates and product molecules to the route length. The evaluation of RetroG on the test benchmark datasets records new results while also presenting interesting future research areas.

1 INTRODUCTION

RP is a branch of Organic Chemistry which involve designing target molecules by chemically combining simpler building block molecules by following *recursive* backward reaction steps. The building block molecules have two key properties (i) commercially or, (ii) easily synthesizable. As an example, according to Figure 1, the root molecule is the target molecule. To design it, the leaves of the left sub-tree (first and second building block molecules) are combined in a forward chemical reaction process to produce the intermediate molecule (root's left child). This is similarly combined with the root's right child (third building block molecule) leading to the target molecule (root molecule).

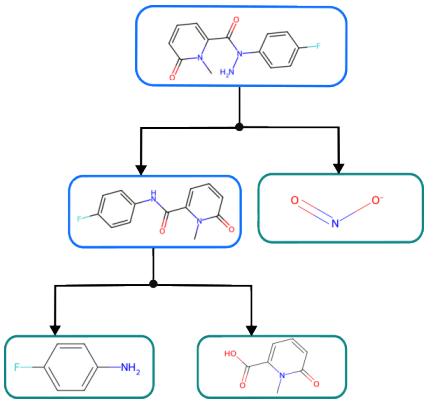

Figure 1: Retrosynthetic Planning representation as a tree. The root molecule is the target molecule. All the leaf nodes are building block molecules that are commercially available or easy to chemically synthesize in the laboratory e.g. through simple forward chemical reaction processes.

Advances and development of new and novel computer-aided RP methods can help design new drugs, agrochemicals and synthetic molecules in material science. Accordingly, many computer-aided RP methods have been proposed over the years to achieve this. Primarily, they are classified according to three major categories: (i) expert-based, (ii) template-based and, (iii) template-free.

Expert-based RP methods were the early computer-aided RP models which were designed according to bond, ring and functional group patterns. The patterns (or rules) are extracted and then subsequently inferenced to the Knowledge Base (KB) to generate the desired planning route (Pensak & Corey, 1977; Wipke et al., 1978; Bauer et al., 1988; Johnson & Marshall, 1992; Parsons et al., 1996; Chen & Baldi, 2009). Due to the dependence on expert-designed rules, expert-based RP models are hard to scale and cost-intensive as well. Additionally, several studies have reported that they are sub-optimal on complex retrosynthesis tasks. These limitations have inspired the development of template-based RP models where the chemical reaction patterns can be extracted automatically.

Template-based RP methods use similarity ranking (Coley et al., 2017b) or Machine Learning (ML) (Segler et al., 2018; Chen et al., 2020; Coley et al., 2017a) to extract the reaction templates from a chemical reaction database (Lowe, 2012) or reaction rule database (Duigou et al., 2019). Precisely, a reaction template is *chemical reaction signature*, thus, it defines the bond, ring, functional group and reaction mechanism patterns which guide a chemical reaction process. Given a product molecule, a reactant set which conforms with the reaction template can be generated automatically. The success of template-based RP is constrained by several factors: (i) *the availability of building block molecules* - all the reaction leaf nodes must be commercially available, (ii) *route length* - shorter and cost-effective routes are generally preferred, (iii) process quality - the proposed planning routes must be *chemically viable* and *sound*, (iv) *valid transformations and symmetry* - there are many valid transformations which can lead to the desired planning path including their symmetries, and (v) *combinatorial explosion* - while the tree depth is generally limited by depth (maximum ≈ 20), the branching factor is determined by valid templates which can span more than 10K transformations. These limitations have motivated the recent advances in template-free RP.

In contrast to template-based RP, template-free RP models do not require a reaction template or a building block molecule. Given a target molecule, a template-free RP can be designed as a prediction task whereby the product is the input and the set of reactants is the output. The prediction problem can be implemented by many ML algorithms such as SVM's, Logistic Regression and even by neural machine translation (NMT) algorithm. NMT is a natural language processing (NLP) method which learns a representation to map the source language to the target language using Deep Neural Networks. In the RP context, the source language is the product while the target language is the reactant set. In the recent past, NMT-based models have been reported by several studies to record unprecedented results e.g. Seq2Seq (Liu et al., 2017) and transformer-based models (Karpov et al., 2019; Zheng et al., 2019). In addition to the successful application of NMT-based algorithms to template-free RP, GNN-based (Kipf & Welling, 2016) methods have also gained a lot of attention recently (Yan et al., 2020; Shi et al., 2020; Dai et al., 2019; Shi et al., 2020). Essentially, the GNN methods model the RP problem as a graph and then learn representations and relations which allow for the molecule reconstruction from synthons according to the target. In retrosynthesis analysis or planning, a synthon is a building block fragment in a molecule which can be used to define simpler building blocks.

In the last five years, tree search and deep learning (DL) have emerged as strong template-based retrosynthetic methods (Segler et al., 2018; Chen et al., 2020; Kim et al., 2021). This combination is inspired by its successful application of the method in challenging computer game problems (Silver et al., 2016; Anthony et al., 2017), where the DL model is used as a state evaluation function as a policy. While GNN and transformer architectures have inspired the design of stronger template-free RP models, template-based RP models are still stronger. Despite this, however, template-based RP is still considered a hard problem since during the template application step, a molecule can have as many as 10K valid transformations (Szymkuć et al., 2016) with the search tree depth (route length) varying between 10 to 20 (Shibukawa et al., 2020) based on the complexity of the target molecule. Furthermore, the current template-based RP methods do not include an adaptation of the applied template or route length to the successful trajectories to guide the future search. Based on this, we present our contributions as follows:

i) **RetroG.** We present a novel template-based RP model based on tree search with a deep neural network as an evaluation function. The value function is a neural network based on GNN

and Fully Connected Layers (FC) architecture which is trained to jointly adapt the reaction templates and product molecules according to the successful planning routes.

ii) **Template-product molecule adaptation.** We present a new framework which adapts the successful planning routes to the applied reaction templates and product molecules to guide future search paths.

The rest of this paper is organized as follows; Section 2: related work, Section 4: methods including the proposed framework and datasets used in this work, Section 5: results and discussion and Section 6: conclusion and future research perspectives.

2 Related Work

This section covers existing retrosynthesis methods according to the expert-based, template-based and template-free classification scheme.

- **Expert-based retrosynthesis**. These were the first RP planning methods. Primarily, they were based on the expert-designed bond, functional group and ring matching rules. Once the matching rules are extracted from the target molecule, they are parsed and then inferenced to the KB to design the desired planning routes. Chen & Baldi (2009) proposed an expert-based model which accepts the target and reagents and then parses them using an inference engine followed by reagent and transformation rule mapping according to the KB. Earlier expert-based RP methods also imitated this design pipeline except for the KB and pattern-matching rules. Pensak & Corey (1977) proposed a model which can design molecules based on the functional, bond disconnection and structural feature strategies. The RG model proposed by Wipke et al. (1978) included a heuristic search mechanism and route evaluation with a chemistry KB and language. Other expert-based RP models include functional group proximity planner (Johnson & Marshall, 1992), theorem-prover inference engine (Johnson et al., 1989), chemistry KB with retrosynthetic synthons (Gasteiger et al., 2000), transformation rules with bond-electron (Bauer et al., 1985) and target-building block distance optimizer (Hanessian et al., 1990). Expert-based models are hard to scale due to the dependency on expert-designed KB and matching rules (Grzybowski et al., 2018; Kayala & Baldi, 2012). Consequently, in the recent past, RP template-based RP has attempted to address these challenges by proposing automatic methods to extract the chemical rules.
- **Template-based retrosynthesis**. Template-based RP methods rely on the reaction rules or chemical reaction database to extract reaction templates automatically. This is then followed by the application of the reaction template to the product molecule to generate the reactant sets. Recently, template-based RP methods have recorded unprecedented performance as a result of ML applications to predict or rank the reaction templates. 3N-MCTS model proposed by Segler et al. (2018) was composed of a Monte Carlo Tree Search (MCTS) algorithm guided by three neural networks. The first was used to select the most promising templates (expansion policy), the second to filter the plausible reactions templates (filter policy) and the third as a state value function during the rollout phase. Retro* Chen et al. (2020) proposed AND-OR RP model with an A*-like heuristic. In the model, every OR node (reaction node) requires any of its children while each AND node (product molecule) requires all its children. Designing RP as an AND-OR tree is sound and effective. Retro* design choice was adopted and extended to design *self-improving* model by Kim et al. (2021) and GNN route cost optimizer (Han et al., 2022). Earlier models such as optimization of the planning with Proof Number Search and heuristic edge initialization (Kishimoto et al., 2019) and hierarchical sampling GNN logical Network (Dai et al., 2019) also recorded very strong results. Other template-based RP methods which have registered competitive performances include template ranking and prediction with symbolic features (Segler et al., 2017), two-step DNN transformation prediction (Baylon et al., 2019), self-play with DNN evaluation (Schreck et al., 2019), template similarity scoring (Coley et al., 2017b) and chemical-aware similarity scoring (Koch et al., 2019).
- **Template-free retrosynthesis.** Template-free RP methods require a database of chemical reactions or rules to extract reaction templates. In contrast, template-free RP methods

model the planning task as a prediction; the product is an input feature while the reactant set the label. This design pipeline can allow for the application of many ML algorithms to the problem. Seq2seq model (Liu et al., 2017) learned to map the source language (product) to the target language (set of reactants). Both the input and output were coded as Simplified molecular Input Line Entry System (SMILES) (Weininger, 1988). A similar design pipeline was also adopted by Karpov et al. (2019) while using transformer architecture (Vaswani et al., 2017). While transformed-based models have recorded unprecedented success rates, LSTM-based models e.g. end-to-end chemical reaction learning with BiLSTM (Schwaller et al., 2019) have also been investigated by other studies and reported competitive and promising results. Other studies such as Kayala & Baldi (2012); Lin et al. (2020) have since extended the transformer-based models with tree search recording even stronger results. GNN-based models have also gained a lot of attention due to their ability to represent molecules better while also learning more useful graphical patterns. Generally, the GNN-based models solve the RP as follows; (i) identify the reaction centres, (ii) learn synthon relations e.g. with a relational model such as Schlichtkrull et al. (2018) and apply them, then (iii) generate and optimize the graph. Existing GNN-based methods include Graph to Graphs prediction (Shi et al., 2020), Edge Attention Network with chemical context (Xie et al., 2022), Weisfeiler-Lehman higher order molecular reconstruction (Jin et al., 2020), end-to-end learning with reactant-reagent-product (Coley et al., 2019b) and graph relation optimization (Yan et al., 2020).

3 DRUG DISCOVERY

The drug discovery process is a key research problem in Bioinformatics computation. The key fundamental objective in drug discovery is to design a candidate organic molecule that can effectively bind to a biological target such as protein or RNA and subsequently interrupt its biological functions thus inhibiting the *disease-causing* ability. Generally, the candidate molecule should have a complementary structure (or shape) and charge to the target for effective binding and efficacy (Zhou & Zhong, 2017). Designing effective candidate molecule during the drug design process can take 12 to 15 years and costs billions of money (Yu & MacKerell, 2017).

It is imperative to design candidate molecules that not only possess the optimal absorption, metabolism, distribution, and excretion properties (Lipinski et al., 1997) but also have desired pharmacokinetics - drug's course of action, pharmacogenomics - genes affect on the drug, toxicity - no long term side effects, solubility, PH stability - the stomach is acidic, chemical stability and oral bioavailability properties.

If the candidate molecule which conforms to the desired properties can be defined, then the RP method can be used to generate planning paths which can lead to the desired target. As previously argued, shorter planning can be economically viable during chemical synthesis in the real world thus a typical RP should include the ability to design planning paths with shorter routes. This characterization, however, is limited by the cost and ease of accessing the building block molecules in the real world. Figure 2 shows two planning paths which can be used to synthesize M_0 e.g through reactions R_0 and R_1. RP framework can propose many viable molecule planning routes which can be used to generate candidate target compounds in the drug discovery process.

4 DESIGNING RETROG

4.1 PRELIMINARIES

- **Molecule Space ($\mathcal{M}$).** This is a set of all the valid molecules in organic chemistry. The product of any chemical reaction, m_1 + m_2 ... m_k, is also within the chemical space e.g. $\{m_i\}_{i=0}^{k} \in \mathcal{M}$.
- **Forward Reaction ($\mathcal{F}$).** This is a forward chemical reaction composed of reactants and product molecules. Similar to the existing baseline models, reagents or catalysts are ignored in this paper.

$$m_1 + m_2 \longrightarrow m_3 \tag{1}$$

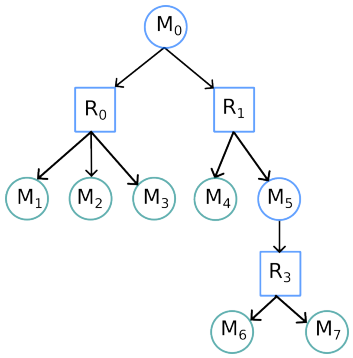

Figure 2: Two valid planning paths for target molecule M_0; it can be synthesized through the right or left sub-tree. The left sub-tree plan has a route length of 1 and 2 for the right one. All children are purchasable molecules. Shorter route lengths can be more economically sound subject to the real-world cost of constraints.

- **Backward Reaction.** This is the reverse of the forward chemical reaction. The forward chemical reaction in 1 can be rewritten in backward form as shown in Equation 2 where $\Longrightarrow$ is the retrosynthetic arrow.

$$m_3 \Longrightarrow m_1 + m_2 \tag{2}$$

- **Reaction Template** ($\mathcal{T}$). A reaction template defines the patterns or fragments for a given chemical reaction. In this research, the reaction templates were extracted from a database of chemical reactions from (Lowe, 2012) containing more than 1.8 Million chemical reaction instances.
- **One-step Retrosynthesis** ($\mathcal{O}$)**.** Given a reaction template and a product molecule, a one-step retrosynthesis applies the product molecule to the template to generate the plausible reactant set. As shown in Equation 3, m_p, a template can be applied to the product molecule leading to the reactant set. Importantly, due to a large number of valid reaction templates, only the top 50 reaction templates returned by the MLP ranking model are used. The MLP model is trained according to the Chen et al. (2020) literature and then used to predict the promising templates to apply. The application of product to template functionality is implemented in software utilities such as RDChiral Coley et al. (2019a) and RDKit Landrum (2013). Importantly, while one-step retrosynthesis has shown the ability to synthesize simple target molecules, designing complex molecules requires many one-step calls - the multi-step retrosynthesis.

$$\mathcal{O} : m_p \longrightarrow \{m_i + m_{i+1} ... m_k\} \tag{3}$$

- **Building block molecules.** Building block molecules is a set of all commercially available molecules. During the tree search, a node is considered terminal if two conditions are met: (i) the maximum depth is reached or (ii) the children of the current reaction node are in the building block molecule set. In this research, the building block molecules set was created from eMolecules database [1]. It is composed of more than 231 Million commercially available molecules.

4.2 RetroG Planning and Self-Playing

Template-based retrosynthetic planning can be formulated as Markov Decision Process (MDP) as $\langle S, A, P, R \rangle$ corresponding to state space, a valid set of actions in a given state, the transition function from one state to another and the reward obtained during the transition respectively. Accordingly, $\mathcal{M}$ can be mapped to S, reactions from $\mathcal{O}$ to valid actions, P to a deterministic value function and R to a unique scalar signal for both the terminal and non-terminal states.

[1] https://www.emolecules.com/

Tree search and learning have been applied to solve many game problems including Anthony et al. (2017); Silver et al. (2016; 2017). In these studies, the value function is modelled using a deep neural network which optionally serves as a policy. Instead of the expensive rollout simulations, the value network predicts the state value which then informs the policy. Consequently, at the end of the episode, the reward signal assigned is used as the target for all the observed states. With the reward-state tuple combinations, the value network is then trained in a supervised learning fashion to improve its experience. This is the core concept underlying self-playing.

RetroG is a tree search and learning algorithm which uses self-playing to improve the value network. The states observed during the episodes are assigned a label corresponding to the route length which acts as the reward signal. The model is composed of an AND-OR tree search involving three phases: selection, expansion and backpropagation phases. Each of these is covered in the following section and graphically shown in Figure 3.

I) **Selection.** During this phase, the node which minimizes the value function, neural network (v_θ), is selected. This is based on Equation 4 where T is the tree with nodes (t). Each reaction node requires all its children (AND nodes) while a given product molecule can be synthesized by any of the reaction nodes (OR nodes).

$$argmin_{t \in T} v_\theta(t) \tag{4}$$

II) **Expansion.** Following the selection phase, if the resulting reaction node is non-terminal i.e. all its children are not in the building block set and the maximum tree depth (route length) is not reached, then it is expanded by adding a new reaction node with the corresponding children. The expansion is achieved using $\mathcal{O}$.

III) **Backpropagation.** If the terminal state is reached and the route is found, then the negative of the total length (tree depth) backpropagated up the tree or its positive value otherwise.

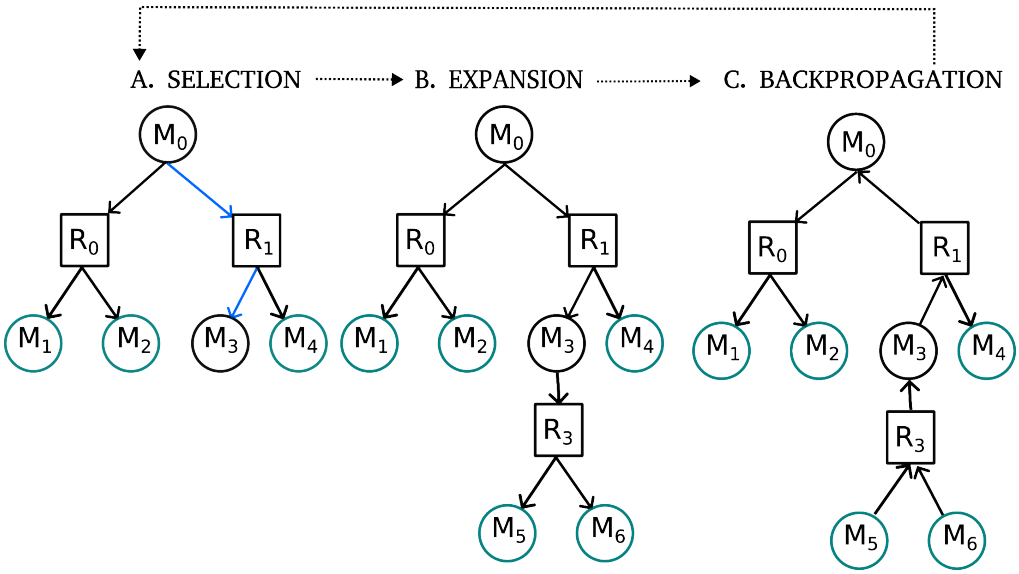

Figure 3: RetroG Retrosynthetic Framework; selection, expansion and propagation. Squared nodes are reaction nodes. All the leaves are building block molecules. M_0 can be synthesized by any of the left OR right sub-tree reactions while R_3 requires M_5 AND M_6.

4.3 Value Network

During the expansion phase of the tree search, the one-step retrosynthesis model generates new reaction nodes in the non-terminal states. This is achieved by applying the product molecule (m_p) (from the tree node) to the templates predicted by the MLP model. In the terminal state, all the

applied templates, the product molecules and route length (tree depth) are collected to train the value network in a supervised learning fashion e.g $\mathcal{D}^{value} = \{\mathcal{T}, m_p, d\}$. $\mathcal{T}$ and m_p are the input features corresponding to the template and product molecule respectively while d is the target which corresponds to the route length. The value of d is negative if a successful planning route is found and positive otherwise. The network is trained based on the Mean Absolute Error loss function shown in Equation 5 and Cosine Annealed AdamW optimizer with 0.004 as the initial learning rate. The batch size was set to 32 throughout the experiments.

$$\mathcal{L} = MAE\big(v_\theta(\mathcal{T}, m_p), d\big) \tag{5}$$

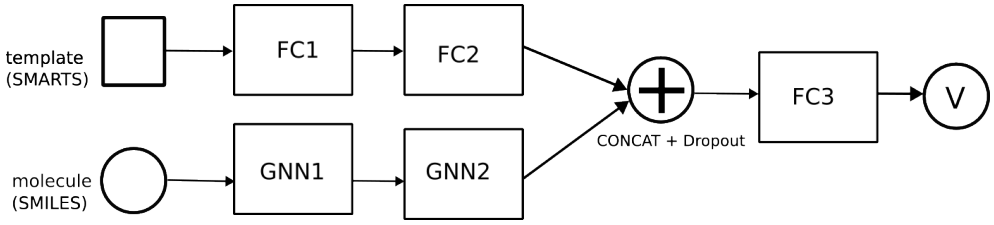

Figure 4: Value Network Architecture. It accepts the SMART-encoded template and SMILE-based product molecule and then computes the scalar value. GNN: Graph Neural Network (Kipf & Welling, 2016), FC: Fully Connected Neural Network Layer, V: predicted reward value. The input-output dimension of FC1 is 4096-512, FC2:512-512, GNN1:94-256, GNN2:256-512 and FC3:1024-1. FC and GNN layers were followed by a LayerNorm (Ba et al., 2016) except for FC3. The Dropout (Srivastava et al., 2014) probability was set to 0.5. The template feature vector was extracted by using the *CreateStructuralFingerprintForReaction* API from RDKit (Landrum, 2013).

The value network adapts successful template-product molecule trajectories to the route length in a self-playing fashion. This value function improvement method has recorded successful results in solving computer game (Anthony et al., 2017; Silver et al., 2016; 2017) and computational biology problems (Kim et al., 2021; Koch et al., 2019; Obonyo et al., 2022). The value network architecture is composed of FC and GNN layers. Recently, GNN has emerged as a powerful method for learning graphical representations (Perozzi et al., 2014; Kipf & Welling, 2016; Schlichtkrull et al., 2018). Molecules are inherently graphical hence learning product molecule representations with GNN can allow for the discovery is novel molecular patterns leading to better-performing models. In contrast, the template (encoded in SMARTS) features are learned using FC since they are not valid molecules e.g. they are one-hot encoded fragment patterns. This was achieved using RDKit (Landrum, 2013) API. Features independently learned by the GNN and FC are concatenated and then trained in an end-to-end fashion. This network design choice allows for the sharing of representations across the two paths leading to a robust model due to the multi-task learning ability (Ruder, 2017).

4.4 Dataset

The US Patent Office Dataset (USPTO) (Lowe, 2012) dataset is used in this study. It is a publicly available database containing about 1.8 Million organic chemical reactions spanning the past three decades up to the year 2016. The dataset was cleaned and split according to Chen et al. (2020) (80-10-10). The train set is used to train the MLP neural network model used to predict the possible templates used by $\mathcal{O}$. We refer the readers to Chen et al. (2020) literature on the MLP training procedure. RetroG is tested on the validation set. All the experiments were run on a single Nvidia Quadro T1200 GPU (the Turing architecture)

5 Results and Discussion

Our proposed model was compared with existing template-based methods including (i) 3N-MCTS (Segler et al., 2018), Retro-* (Chen et al., 2020) and DFPN-E (Kishimoto et al., 2019). The greedy Breadth First Search (BFS) was used as a baseline model. Based on the obtained results, RetroG

Table 1: *Success rate, route length and time.*

	success rate	**route length**	**one-step calls (time)**
BFS (Greedy)	20.64 ± 0.28	11.00 ± 0.20	400.23 ± 0.20
3N-MCTS	36.80 ± 0.39	9.95 ± 0.60	300.01 ± 0.19
Retro-*	85.09 ± 1.05	9.01 ± 0.47	250.80 ± 0.34
DFPN-E	54.69 ± 0.21	10.92 ± 0.04	350.60 ± 0.05
RetroG	80.00 ± 0.08	$\mathbf{7.95} \pm 0.38$	$\mathbf{200.51} \pm 0.89$

recorded very competitive and promising results. It recorded the shortest planning routes as well as the best planning runtime while also achieving second-best success rate scores. The runtime was computed based on the number of calls to the one-step model. This is identical to the computation method employed by Chen et al. (2020) and Kim et al. (2021) in their work. In light of these performances, we argue that through self-play, the RetroG value network learned useful representations from the successful trajectories during the tree search simulations. In addition, it also shows that the learned representations captured template-product molecules to route length patterns. Accordingly, during the selection phase of the tree search, only the template-product molecule pair is more likely to lead to the successful planning path.

In reality, shorter plans are preferred in RP, however, in some instances, longer ones can be desired. Shorter or longer design path choice is constrained by other real-world synthesis factors such as time, labour, costs, availability of building block materials and even expertise. To this end, designing plans which are not only shorter but also aligned with these constraints is imperative. Of course, this motivates the future research frontier in template-based RP.

6 Conclusions and future research

In this paper, we present a template-based retrosynthetic model called RetroG. It adapts the templates-product molecule patterns to the route length. As result, during the route path selection phase, the most promising one is selected. This is achieved by a value network trained in a self-playing fashion on both the successful and unsuccessful trajectories. The proposed model recorded the shortest route and runtime on the test set with very competitive success rate scores. While these metrics are promising and competitive, it is imperative to design a model which not only has shorter routes but also optimal success rates. To this end, as part of the future research frontier, designing new value network architecture or evaluation function with novel training mechanisms present very interesting directions.

REFERENCES

Thomas Anthony, Zheng Tian, and David Barber. Thinking fast and slow with deep learning and tree search. *Advances in Neural Information Processing Systems*, 30, 2017.

Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E Hinton. Layer normalization. *arXiv preprint arXiv:1607.06450*, 2016.

J Bauer, R Herges, E Fontain, and I Ugi. Igor and computer assisted innovation in chemistry. *Chimia*, 39(2):43–53, 1985.

Johannes Bauer, Eric Fontain, Dietmar Forstmeyer, and Ivar Ugi. Interactive generation of organic reactions by igor 2 and the pc-assisted discovery of a new reaction. *Tetrahedron Computer Methodology*, 1(2):129–132, 1988.

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.

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.

Connor W Coley, Regina Barzilay, Tommi S Jaakkola, William H Green, and Klavs F Jensen. Prediction of organic reaction outcomes using machine learning. *ACS central science*, 3(5):434–443, 2017a.

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, 2017b.

Connor W Coley, William H Green, and Klavs F Jensen. Rdchiral: 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.

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.

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.

Bartosz A Grzybowski, Sara Szymkuć, Ewa P Gajewska, Karol Molga, Piotr Dittwald, Agnieszka Wołos, and Tomasz Klucznik. Chematica: a story of computer code that started to think like a chemist. *Chem*, 4(3):390–398, 2018.

Peng Han, Peilin Zhao, Chan Lu, Junzhou Huang, Jiaxiang 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 Hanessian, Jonathan Franco, and Benoit Larouche. The psychobiological basis of heuristic synthesis planning-man, machine and the chiron approach. *Pure and applied chemistry*, 62(10): 1887–1910, 1990.

Wengong Jin, Regina Barzilay, and Tommi Jaakkola. In *Multi-objective Molecule Generation Using Interpretable Substructures*, pp. 4849–4859. PMLR, 2020.

A Peter Johnson and Chris Marshall. Starting material oriented retrosynthetic analysis in the lhasa program. 3. heuristic estimation of synthetic proximity. *Journal of chemical information and computer sciences*, 32(5):426–429, 1992.

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.

Matthew A Kayala and Pierre Baldi. Reactionpredictor: prediction of complex chemical reactions at the mechanistic level using machine learning. *Journal of chemical information and modeling*, 52(10):2526–2540, 2012.

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.

Mathilde Koch, Thomas Duigou, and Jean-Loup Faulon. Reinforcement learning for bioretrosynthesis. *ACS synthetic biology*, 9(1):157–168, 2019.

Greg Landrum. Rdkit documentation. *Release*, 1(1-79):4, 2013.

Kangjie Lin, Youjun Xu, Jianfeng Pei, and Luhua Lai. Automatic retrosynthetic route planning using template-free models. *Chemical science*, 11(12):3355–3364, 2020.

Christopher A Lipinski, Franco Lombardo, Beryl W Dominy, and Paul J Feeney. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced drug delivery reviews*, 23(1-3):3–25, 1997.

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.

Daniel Mark Lowe. *Extraction of chemical structures and reactions from the literature*. PhD thesis, University of Cambridge, 2012.

Stephen Obonyo, Nicolas Jouandeau, and Dickson Owuor. Designing RNA Sequences by Self-play. *14th International Joint Conference on Computational Intelligence*, pp. 305–312, 2022.

Philip J Parsons, Clive S Penkett, and Adrian J Shell. Tandem reactions in organic synthesis: novel strategies for natural product elaboration and the development of new synthetic methodology. *Chemical reviews*, 96(1):195–206, 1996.

David A Pensak and Elias James Corey. In *LHASA—logic and heuristics applied to synthetic analysis*. ACS Publications, 1977.

Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. Deepwalk: Online learning of social representations. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 701–710, 2014.

Sebastian Ruder. An overview of multi-task learning in deep neural networks. *arXiv preprint arXiv:1706.05098*, 2017.

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, 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.

Marwin Segler, Mike Preuß, and Mark P Waller. Towards" alphachem": Chemical synthesis planning with tree search and deep neural network policies. *arXiv preprint arXiv:1702.00020*, 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. PMLR, 2020.

Ryosuke Shibukawa, Shoichi Ishida, Kazuki Yoshizoe, Kunihiro Wasa, Kiyosei Takasu, Yasushi Okuno, Kei Terayama, and Koji Tsuda. Compret: a comprehensive recommendation framework for chemical synthesis planning with algorithmic enumeration. *Journal of cheminformatics*, 12 (1):1–14, 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, Dharshan Kumaran, Thore Graepel, et al. Mastering chess and shogi by self-play with a general reinforcement learning algorithm. *arXiv preprint arXiv:1712.01815*, 2017.

Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1):1929–1958, 2014.

Sara Szymkuć, Ewa P Gajewska, Tomasz Klucznik, Karol Molga, Piotr Dittwald, Michał Startek, Michał Bajczyk, and Bartosz A Grzybowski. Computer-assisted synthetic planning: the end of the beginning. *Angewandte Chemie International Edition*, 55(20):5904–5937, 2016.

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.

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.

W Todd Wipke, Glenn I Ouchi, and S Krishnan. Simulation and evaluation of chemical synthesis—secs: An application of artificial intelligence techniques. *Artificial Intelligence*, 11(1-2): 173–193, 1978.

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.

Wenbo Yu and Alexander D MacKerell. Computer-aided drug design methods. In *Antibiotics*, pp. 85–106. Springer, 2017.

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.

Shu-Feng Zhou and Wei-Zhu Zhong. Drug design and discovery: principles and applications. *Molecules*, 22(2):279, 2017.