Master's Thesis: ***One ring break model to rule them all?***

Bob Wilhelmus van Schendel

Academic Supervisor: **Dr. Mike Preuss**

Industrial supervisor: **Dr. Samuel Genheden**

A thesis presented for the degree of
Master of Science

Universiteit
Leiden
The Netherlands

Leiden Institute of Advanced Computer Science
Leiden Universiteit
The Netherlands
2023-07-04

Contents

Abstract

AI and ML are crucial in the continued development of Drug Discovery. Retrosynthesis plays a central role in this field and is a good example of how to leverage AI and ML methods and techniques. We do this by focusing on a specific technique for computer-assisted synthesis planning, currently in use by the software AiZynthFinder. This technique consists of Monte Carlo Tree Search combined with a single-step expansion policy for producing synthesis routes. Despite its strengths and although it is working to produce valid synthesis pathways, it has a shortcoming. This shortcoming manifests itself in the underutilization of ring breaking templates in the produced synthesis pathways. We identify the concrete problem and define a Machine Learning task focused on learning when to use ring breaking reactions. We solve this task by applying Graph Neural Networks and developing a new expansion policy to combine multiple models. Then we apply them to AiZynthFinder to improve the performance metrics of produced routes. Due to a possible number of limitations, this new expansion policy does not outperform the current in-use policy. We identify possible inherent limitations in the applied technique and developments that could shed further light on the appropriate application of ring breaking templates and to further improve this technique of AI-assisted synthesis planning.

1 Introduction

Chemical synthesis: The execution of specific *chemical reactions* to convert one or more *reagents* into one or more *products*.

The techniques around how to synthesize a chemical compound, also called a synthesis pathway, have come a long way since their inception. Where, initially, finding such a pathway was largely based on manual trial-and-error by chemists, now, software is able to perform this function largely automatically. This process of finding a synthesis pathway from a target compound is called retrosynthesis, using software to execute this is a part of the area called Computer-assisted Synthesis Planning (**CASP**).

Retrosynthesis is a strategy for deconstructing complex molecules into simpler molecules in order to design optimal synthesis routes for these molecules.[1] By working in reverse, it involves breaking down the molecular structure of a target compound into simpler starting materials and applying chemical reactions to synthesize the molecule. The goal of CASP is to use retrosynthesis to simplify the complex target molecule into more accessible precursors while identifying the optimal synthetic pathway.

The methods underlying retrosynthesis were developed by Elias James Corey in the early 1960s, which revolutionized the field of organic chemistry.[2] Corey's contribution to the field was recognized with the Nobel Prize in Chemistry in 1990. Before the advent of retrosynthesis, chemists relied on manual trial and error to determine synthetic routes for target molecules, which was time-consuming and often resulted in suboptimal yields.[3] The approach is often used in drug discovery, where synthesizing novel compounds with specific therapeutic properties is crucial. This has significantly accelerated the synthesis process, contributing to faster drug discovery.

Today, retrosynthesis serves as the cornerstone of drug development. It seamlessly integrates with the systematic approach of the Design-Make-Test-Analyze (DMTA) cycle, a central optimization protocol in drug development,[4] visualized in Figure 1. This process begins with the design or generation of potential drug candidates in the Design phase. Using computational tools and AI algorithms to identify compounds with desired properties. CASP comes in at the second step, Make, where the goal is to find a synthesis pathway and execute it. This not only involves finding a technically possible route, but also takes into account resource availability and cost of execution.

After synthesis, the third step, Test, involves testing the compounds for biological activity or toxicology using in silico assays. The final step is suggesting future steps or modifications by analyzing the integrated results of all the previous steps in the analyze phase. By using the DMTA cycle, researchers can systematically optimize drug candidates, identify drug candidates more efficiently, reduce the cost and time required for development, and ultimately develop better drugs more effectively.

Artificial Intelligence (**AI**) and Machine Learning (**ML**) are shaping the next phase of developments in Drug discovery and CASP.[5,6] Already Generative AI has shown the ability to automate and shorten parts of the Design phase.[7] AI-assisted Synthesis Planning helps chemists to choose the most efficient and cost-effective synthetic routes, accelerating the Make phase.[8] The final Analyze phase is being accelerated by applications of few-shot learning[9,10] and data imputation techniques.[11] As the relatively easy challenges of a field are solved, more sophisticated techniques are required for further improvement, this is also the case for the applications of AI and ML themselves. In this thesis we will focus on improving a methodology in the field of retrosynthesis and understanding the limitations holding back further improvement.

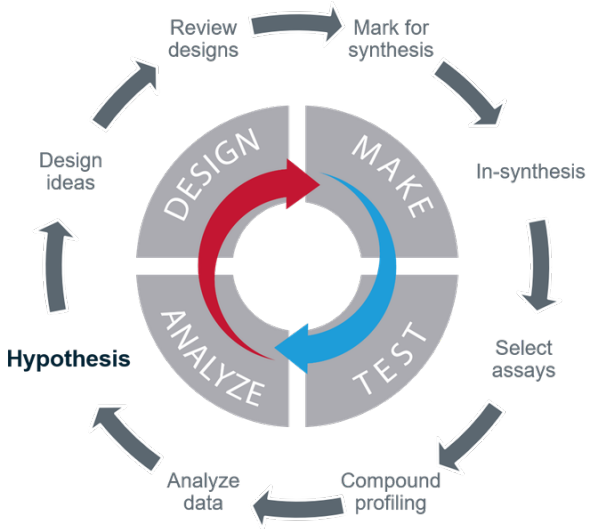

Figure 1: A visualization of the Design-Make-Test-Analyze cycle.

A more specific example of how these developments impact Drug discovery is how retrosynthesis models are utilized in the aforementioned Design-Make-Test-Analyze (DMTA) cycle. Consider REINVENT, a production-ready tool for de novo design of molecules, focusing on the domain of drug discovery.[12] By harnessing the power of deep learning and artificial intelligence, REINVENT enables the generation of chemical compounds that address the exploration or exploitation challenges encountered during the drug development process. This approach aids researchers in the idea generation phase by identifying the most promising compounds aligned with desired properties, shortening DMTA cycles and expediting the entire process.[6]

Another example of a methodology which has itself significantly impacted drug discovery and is being further improved by the application of AI and ML is High Throughput Screening (HTS). It achieved this by providing a rapid and systematic approach to screen large chemical libraries to identify compounds exhibiting desired biological activity.[13] HTS has replaced the traditional time-consuming and expensive manual screening process by automating the process and enabling the screening of thousands of compounds at once. This approach has led to the identification of potential drug leads for further development and optimization. This method is being improved by using Machine Learning (ML), for example by selecting the most promising compounds in a HTS batch for further screening without any manual and labor-intensive analyzing being necessary. It allows more efficient screening of large numbers of compounds, while finding active compounds more easily. Applying AI is a logical approach and further optimizes the efficiency of HTS and thus drug development.[14]

Due to the virtually infinite search space, integrating AI, ML and other technologies is crucial to the future of drug discovery. It allows optimization of the design, synthesis, testing, and analysis of new compounds for potential therapeutic applications. This integration is currently also underway in the field of CASP.[15]

One of the challenges of CASP is the enormous amount of applicable reactions for given molecules. The possible applicable reactions easily exceed 1 million. This makes the search space too large to search exhaustively thus necessitating more sophisticated approaches. ML is used to learn the best transformations for breaking down molecules and AI for guiding and automating the process. This development has resulted in multiple software programs in-

cluding AiZynthFinder, Spaya.ai and Chemistry42 currently usable by users in the lab.[16,17]

These programs are effective but have limitations and face challenges in their application, especially in the human-likeness of the produced routes. In this thesis we will look at a shortcoming of Monte Carlo Tree Search combined with an expansion policy in retrosynthesis. We will analyze the problem and use several methods to approach it. Then we show that ML models, specifically Graph Neural Networks (GNN), can be used to model this specific structural task. We will provide a solution and show how to integrate it into AiZynthFinder, in order to improve the methods employed in this software. As we will focus on the techniques, this solution also applies to the method used in any implementation.

2 Background

This thesis is on chemoinformatics, which lies on the intersection of Computer Science and Chemistry. Many concepts are mentioned which may need clarification for researchers from either field, which is done in this section. It is separated into two parts, Chemistry and Computer Science, though they often blend into one another.

Chemistry

2.1 Reactions

Chemical reactions play a fundamental role in the transformation of chemical substances from one form to another. In its simplest definition, a chemical reaction is a process in which reactants are transformed into product(s). Solvents, catalysts, side-products or reaction conditions which can influence the rate or direction of a reaction, are not taken into account in this thesis and are ignored.

To represent a chemical reaction, a diagram is typically employed in which the reactants are listed on the left-hand side and the product(s) on the right-hand side. An example is provided in Figure 2. This arrangement provides a visual indication of the reactants and the product(s), with arrows denoting the direction of the reaction. Chemical reactions can take on a variety of forms, from simple substitutions to complex, multi-step processes.

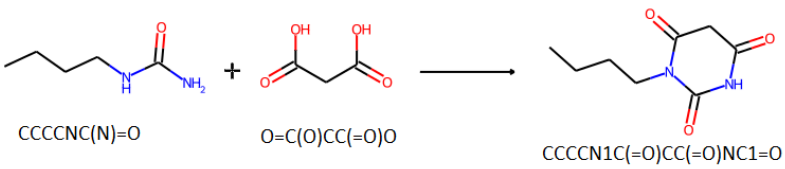

Figure 2: A chemical reaction with two reactants (on the left side of the arrow) and one product (the molecule on the right side) containing a ring structure. If this reaction were to take place both the reactants molecules would be combined into one product molecule. The corresponding SMILES strings of the molecules have been added underneath.

2.2 Templates

Chemical reactions are typically described in terms of fixed, complete reactants and products, but this approach can be limiting when dealing with complex molecules. This is because the reaction could actually be applied to some part of the molecule while the entire molecule does not fit the exact structure specified in its formula. However, we can isolate

the reaction center from its description, only looking at the atoms and bonds changing in the reaction and those occurring one bond away. By doing this we capture the generalized form of this reaction and refer to this as the reaction template.[18,19] This allows us to apply reactions in a more general way and helps us to find novel synthesis routes. An example of two templates used in AiZynthFinder are shown in Figure 3. This approach has been shown to be effective in a variety of fields, including pharmaceuticals.[20,21]

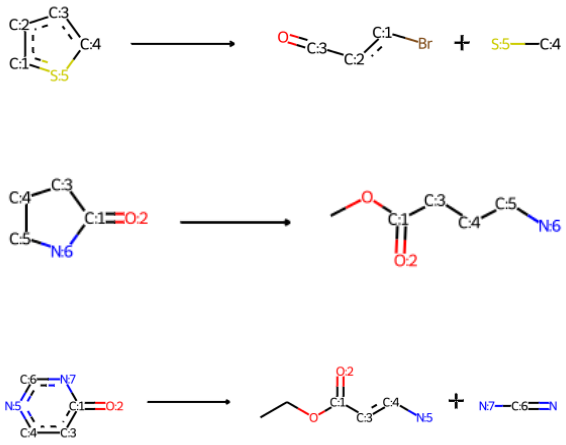

Figure 3: Three examples of reaction templates rendered with RDkit.[22] These templates are all The upper template is *Benzothiophene synthesis*, the middle template is Carboxylic ester + amine reaction and the lower template shows a Pyrimidin-4-one synthesis.

2.3 Ring structures

In the context of chemical structure, a ring can be defined as a specific combination of distinct bonds and atoms that form a loop within a molecule, as per the IUPAC definition.[23] The product of the reaction in Figure 2 shows a ring structure surrounded by double-bonded oxygen atoms. While the terms ring and *ring system* are often used interchangeably in the literature, for the purposes of this thesis, the term "ring" will be employed. Rings have long been recognized as a key feature in drug discovery, as they can significantly influence the chemical properties and activity of a molecule. Consequently, rings have been employed as scaffolds in the generation of drug-like molecules in numerous studies.[24–26] This thesis will explore the role of rings in the context of retrosynthesis.

2.4 Synthesis pathway

In the realm of synthetic organic chemistry, a crucial step in the development of new drugs and materials is the synthesis pathway, which is a sequence of chemical reactions that, when executed, leads to the desired end product. Monte Carlo Tree Search (**MCTS**), further defined in Section 2.8.1 is a powerful algorithmic technique that can be used to identify and plan such synthesis pathways. By following the tree from the terminal node to the root node, MCTS hopefully yields a synthesis pathway that is both efficient and high-yielding. One commonly used representation of retrosynthesis is the synthesis tree, which starts with the target molecule at the root node and includes the precursors at each stage of the retrosynthesis, see Figure 4. The edges of the tree correspond to the application of

specific chemical reactions or templates to molecules or precursors at a given node, leading to the retro output of that template and the next node in the tree. When a node contains only precursors that can be directly procured, it is considered a terminal node, and a synthesis pathway can be constructed simply by collecting the precursors in that node and applying all the chemical reactions from the terminal node to the root node.

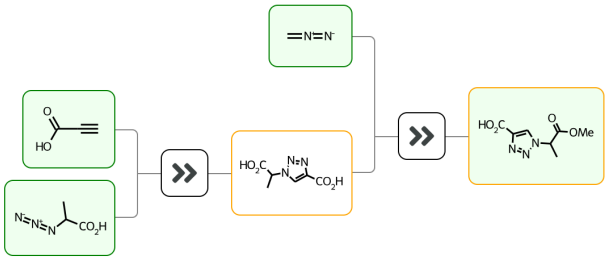

Figure 4: A produced synthesis pathway with the target compound on the right and the set of precursors being the 3 leaves on the tree. This synthesis pathway uses 2 reactions, represented by the double arrows, with an intermediate molecule between the arrows.

2.5 Retrosynthesis

The process of retrosynthesis is a critical aspect of synthetic organic chemistry, enabling the deconstruction of a target molecule into a series of simpler precursor structures that can be synthesized through a series of chemical reactions. By working backwards from the target molecule to identify the appropriate bond disconnections we can approach the problem in a structured way. This can be seen as a search problem, where we intend to find a candidate set of precursor structures through transformation steps of reversed chemical reactions. This structure also makes it possible to use algorithms to guide our search in this space, to make trade-offs between exploration and exploitation and perform optimization.

Once the precursor structures have been identified, chemists can plan the necessary synthetic steps to create them. These synthetic steps may include the use of protecting groups, which allow specific chemical reactions to take place without altering other parts of the molecule, as well as the application of various reaction conditions to optimize yields and selectivity. This space of sets of precursor structures is large, making algorithmic optimization a requirement.

Computer Science

2.6 Molecular representation

Molecular representation is an important aspect of machine learning applications in the field of drug discovery. The structures of molecules needs to be converted into representations that can be used as inputs for machine learning models. Typically, molecules are created in 2D and their 3D conformation is ignored. There are many ways of representing molecules due to their complex structural and chemical characteristics. Among them are so called fingerprints, which function as fixed-length molecular descriptors, and graph-based representations.[27,28] The choice of representation is essential for the performance of ML algorithms.[29] Each representation method has its merits and drawbacks, we shall describe them and some of the concepts related to them.

2.6.1 Molecular fingerprints

There are many of expert-designed and learned descriptors at the molecule level available. These descriptors encode chemical or structural information as a fixed-length sequence. There are several types of molecular fingerprints, including circular fingerprints, linear fingerprints, and extended-connectivity fingerprints, among others.[27]

Essentially, molecular fingerprints are simply long sequences of bits (binary digits) that hold information on whether a molecule holds a certain structural fragment or other specific property. While this structure has many advantages, such as computational efficiency and flexibility, one disadvantage is that it discards much of the structural information during creation, also called featurization. Even though some fingerprints encode structural elements or fragments, they do not capture the global structure of these fragments or elements. This can limit the ability of machine learning models to capture more complex structural relationships of molecules.[28]

2.6.2 SMILES

A widely used line notation representation of molecules is the Simplified molecular input line entry system (SMILES).[30] SMILES strings use a combination of letters, numbers, and symbols to represent the atoms and bonds in a molecule. See figure 2 shows the SMILES strings of the molecules pictures below their structural formula. The advantages of this format are compactness, readability and ease of editing. One of the disadvantages is that this structure does not (directly) encode all information for a molecule, such as 3D conformation.

2.6.3 Graph representation

The representation of molecules as molecular graphs offers a powerful tool for mapping the structural information of a molecule to a set of nodes and edges. Typically, atoms and bonds are mapped to the sets of nodes and edges, respectively, which can then be encoded as an adjacency matrix. This representation can be further expanded through the addition of node or edge features, characterizing information such as specific bond types. While a graph representation holds all the structural information in the form of matrices, a more natural of thinking about this is by using visualizations such as in Figure 2.

The main advantages of the molecular graph representation are twofold. First, this approach enables the encoding of all available structural information, including global structural information. Second, the graph representation provides a highly interpretable format for domain experts, allowing for efficient and accurate analysis of the underlying structure and properties of the molecule, as well as its subgraphs.[28]

2.7 Tree

In Computer science terms a tree is a structure for encoding information where each element has references to (usually two) related elements. These elements are called nodes and the connections are called edges. The tree lends its use to the fact that searching in a tree leads to elimination of many options with each step, as a node is only connected to a small subset of related nodes. Imagine wanting to find a synthesis route for a compound with a tree structure. The completed compound would be located in the root node (with no incoming edges as it is complete). Each edge leading from the root node would represent a template applied to the compound and the nodes they point to would hold the resulting molecule(s). If we would follow the nodes down and remember all the steps, we would end up with a node holding only fragments of the original molecule easy to procure. This way we would solve a complex problem with a simple step-wise process. The synthesis pathway in Figure

$$\frac{w_i}{n_i} + c \cdot \sqrt{\frac{\ln(N_i)}{n_i}}$$

Figure 6: The formula for calculating the Upper Confidence Bound. w_i stands for the number of positive results for the node considered after the ith move. n_i stands for the number of simulations for the node considered after the ith move. N_i stands for the total number of simulations after the ith move run by the parent node of the considered node. c is the exploration parameter.

4 uses a tree, albeit with a slightly distinct structure. This structure takes the molecules as nodes and the applied templates as the edges. The edges used here lead from 1 node to many nodes, indicating the resulting retro outputs of an applied template.

2.8 Tree Search

Tree search is a powerful computational tool for identifying viable pathways for synthesizing target molecules. In retrosynthesis, a tree search algorithm can explore the vast space of possible reactions that could be used to synthesize a target molecule. As mentioned in the previous subsection, nodes represent a collection of precursors (of the target molecule) and edges represent a reaction template applied to a molecule in a node. By searching a generated tree, the algorithm can identify the most efficient and feasible pathway for synthesizing the target molecule. This way, various factors such as the availability of starting materials, reaction yields, reaction condition compatibility, could be automatically taken into account. Tree search is a valuable tool in retrosynthesis, allowing chemists to quickly explore a large number of potential pathways and select the most promising ones for further investigation. In this thesis we will focus on MCTS but other interesting and relevant algorithms include Depth-First Proof Number[31] and A*[32] are also search algorithms on Tree structures.

2.8.1 Monte Carlo Tree Search

A specific powerful algorithm that can be used in tree search is MCTS.[18] This search algorithm incrementally searches and expands a tree exploring the search space. It combines this with random sampling to efficiently select candidates for expansion.

As mentioned, MCTS incrementally builds a tree exploring the problem space. Each iteration consists of four steps:

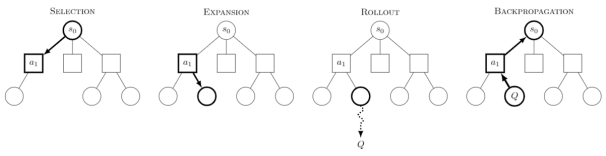

Figure 5: A diagram showing the 4 steps of Monte Carlo Tree Search and their effect on the tree being explored.

Selection: Starting from the root node, the algorithm explores the tree by selecting child nodes based on certain criteria. One common criterion is the UCB, explained in Figure 6 , which automatically balances exploration and exploitation.

Expansion: Once a leaf node is selected, the algorithm expands the tree by adding one or more child nodes corresponding to unexplored actions from that state. These actions

represent moves, transformations or decisions that can be taken from the state of that particular node.

Simulation or Rollout: The algorithm performs a simulated or random playout from the newly added child nodes. It continues playing the game, or performs some other heuristic-based simulation from that state until a terminal state is reached.

Backpropagation: After the playout is complete, the results are backpropagated up the tree to update the statistics of the visited nodes and improve further iterations of the algorithm.

These steps are repeated for a predefined number of iterations, until a computational budget is exhausted or until a target has been reached. MCTS can enable chemists to automatically explore a huge number of possible synthesis pathways and identify the most feasible and efficient ones. It is particularly effective in situations where the state space is large and difficult to exhaustively explore, such as in games or the process of Retrosynthesis.

2.9 Retrosynthetic reactions

In the context of retrosynthesis, the conventional terms of reactant and products used in chemical reactions are reversed, with retro input being used instead of products, and retro output instead of reactants. This reversal of terms reflects the thinking process involved in retrosynthesis. Retro output represents the precursor structures, which are the starting materials for that step of the synthetic pathway, while retro input represents the target molecule, which is the final product of the synthetic pathway. This is visualized in Figure 2.

$$\textbf{Original} : \textit{Reactants} \rightarrow \text{Products} \tag{1}$$

$$\textbf{Retrosynthetic} : \textit{Retro output} \leftarrow \text{Retro input} \tag{2}$$

2.10 AI-assisted synthesis planning

As this thesis is focused on MCTS and expansion policies, we will explain their main moving parts. These methods are applied in AI-assisted Synthesis Planning in general, but also in AiZynthFinder in particular. So, even though we define our problems and solutions in the context of AiZynthFinder, the results and implications actually apply to these methods in general and they can be viewed as agnostic to the specific application.

2.11 Expansion policy

The described underutilization has to do with the expansion policy of AiZynthFinder. As this model expands nodes it is directly related to which reactions get rated highly and thus are likely to be used in the overall synthesis pathway. This model is trained by extracting reaction templates from the reaction databases Reaxys,[33] Pistachio[34] and internal AstraZeneca Electronic Lab Notebooks. Due to the biases in these reaction databases we get a model that shows the same tendencies. In this case we have few reactions for making or breaking ring structures in the datasets, even though they are ubiquitous in modern day pharmacology.[18] This kind of bias is not only present when it comes to ring breaking reactions but is more generally a problem in data. This leads to problem in CASP in general,[35] indicating the importance of finding a solution.

2.12 Single-step retrosynthesis

The overall goal of retrosynthesis is to systematically break down a target compound into simpler precursor molecules. This strategic planning tool is used a lot in organic chemistry

and can be generally divided into two main parts: Single-step retrosynthesis and multi-step retrosynthesis.

Single-step retrosynthesis focuses on effectively finding precursors for a molecule. Single-step retrosynthesis is generally approached as a supervised learning problem, with template data or sequences of SMILES being used as the training data. The former is called template-based[36] and is the category that AiZynthFinder as well as this thesis belong to. The idea is to use chemical reaction data to extract reaction templates, which can be used for training. The latter approach is called template-free and it is a promising direction of development in CASP.[37–39] Template-free approaches are, however, not within scope of this thesis and we will restrict ourselves to template-based methods.

2.13 Multi-step retrosynthesis

Multi-step retrosynthesis is centered on how to discover viable and efficient synthesis pathways from target compound to starting materials.[40] It involves deconstructing molecules in successive sets of precursors. This process is facilitated by a, generally fixed, single-step retrosynthesis model to suggest expansions to apply to the molecule in question.[41] The focus here is on optimization and the task is often approached as a search problem; "How do we efficiently leverage the single-step model to find a viable and efficient route?" Some of the pioneering contributions to this field are template-based and in the area of NN-guided MCTS.[18]

This is also the area this thesis intends to improve. Combining expansion models or, equivalently, single-step retrosynthesis models, is a promising area of research able to improve performance of single-step and multi-step retrosynthesis as a whole.

2.14 AiZynthFinder

Our software of choice is AiZynthFinder.[42] This is a software for multi-step retrosynthesis currently in use at the pharmaceutical company AstraZeneca. It uses templates to apply deconstruction steps, this is also referred to as being template-based. AiZynthFinder combines MCTS, for the overall route planning, with an expansion policy at each individual step to limit the growth in breadth of the search tree to a manageable degree. AiZynthFinder approaches retrosynthesis as a tree search problem, where we begin from the target compound and where the goal is to find sets of precursors that are commercially available. The nodes hold a set of precursors each and every iteration a number of transformations are applied to one of the molecules in a leaf node, expanding it through edges into new nodes. Note that the expansion policy is functionally the same as a single-step retrosynthesis model. We continue using the term expansion policy as we consider the methodology in the context of MCTS, not just AiZynthFinder.

The mentioned expansion policy is a model that predicts the most promising candidate templates to apply to a given molecule. The expansion policy is not fixed in AiZynthFinder and there are two available models for use as a policy:

- **A general model**, trained on all the available templates.
- A specialized **ring breaking model**, trained only on ring breaking templates.[43]

Figure 7 shows a snapshot of the execution of AiZynthFinder visualized as a diagram. The most important thing to keep in mind is that AiZynthFinder uses a tree for searching through the problem space in which the nodes hold sets of precursors (to the nodes preceding it) and edges represent specific templates being applied to a given molecule in a preceding node. As more nodes get expanded into new nodes with precursors, the depth of this tree

increases and eventually we reach a set of precursors that is completely procurable. This would not necessarily end our search as diversity of synthesis pathway is valued highly so AiZynthFinder will exploit the found route to produce more than one synthesis pathway until the iteration or time limit is reached.

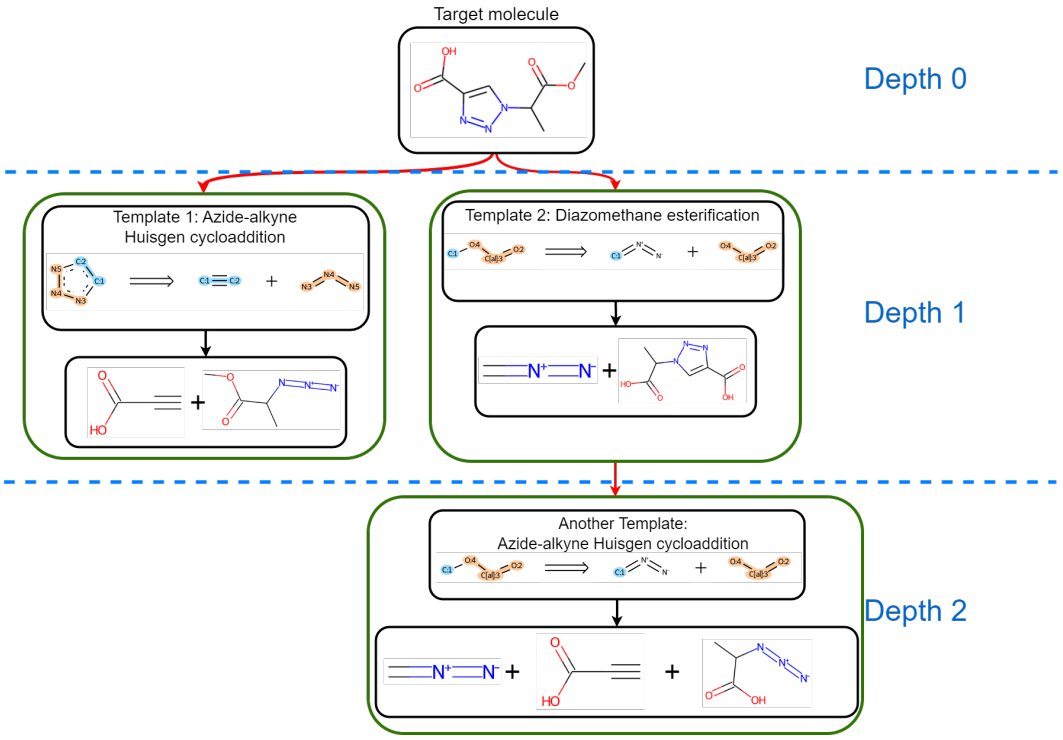

Figure 7: A visualization of AiZynthFinder during execution. At the picture moment, there are two branches of the root node explored. The left branch leads to the Azide-alkyne Huisgen cycloaddition which directly deconstructs the target molecule into a set of two procurable molecules. The right branch first applies the Diazomethane esterification and subsequently applies a different Azide-alkyne Huisgen cycloaddition to arrive at a set of three procurable molecules. How these two separate routes are valued depends on the scoring of the templates and procurable molecules, this was too extensive to be added in this image. Nonetheless, it shows a snapshot oh how AiZynthFinder functions.

3 Problem statement

AiZynthFinder is in use today and can provide chemists with reasonable synthesis pathways for novel molecules. Despite this, the generated synthesis pathways sometimes suffer from inaccuracies and there are limitations that can be addressed. One of these limitations pertains to a characteristic of the synthesis paths created by the program. There is a distinct lack of reactions that create ring structures in the produced synthesis pathways.[44] This means that there are too few ring breaking templates being applied during the retrosynthetic process. We will focus on this issue as it is one of the most problematic limitations, as indicated by chemists. The overarching goal to solving or ameliorating this underutilization is that it may increase the number of compounds that we can find synthesis routes for,

or the quality of these produced synthesis routes. This would increase the effectiveness of retrosynthesis as a whole and AiZynthFinder in particular.

Being able to use the different models that we have would be able to improve the effectiveness of the expansion policy by leveraging the strengths of each model. Finding methods to combine multiple expansion models will be a big step in the right direction, as it will allow greater specialization and efficiency of the template selection process.

A policy model predicts a number of promising templates and associated probabilities. These probabilities are also referred to as priors and used for updating the MCTS values. A big hurdle in how to combine these models is that they are trained separately and consequently, after training, draw from separate probability distributions. This makes combining the probabilities from separate models mathematically unreliable. Despite this, combining the predictions from these different distributions with some method is an avenue worth exploring.

3.1 Problem definition

The core problem of this thesis is the underutilization of ring breaking templates in produced synthesis pathways of AiZynthFinder. This problem is directly related to the constrained potential of the expansion policy in the context of its use with MCTS. This is due to bias in the data that are used for training the models in question and this limitation cannot be directly addressed so this thesis intends to take another approach. Our proposed solution is to experiment with combinations of the available expansion policy models, to leverage the strengths of multiple models in a way that improves the overall template prediction. This could then lead to improvement in the overall performance of AiZynthFinder and other programs implementing MCTS and Expansion policies. This leads to the main problem definition, which will then be divided into subproblems in Sections 4 and 5.

> *Main problem: Can we find a method to combine multiple expansion policies to improve the performance of AiZynthFinder*?*

*The performance of AiZynthFinder is defined as the quality of the produced synthesis pathways or the diversity of these pathways. This is further explained and defined in Section 5.

4 Expansion policy experiments

There are a number of ways to combine the mentioned policy models, some more ad-hoc than others. Let us introduce some notation to make the explanation easier to follow.

Name the probability distributions for the General model G and for the ring breaking (**RB**) model R. For an arbitrary prediction we name the produced templates t_G^i and the priors p_G^i for the general model and t_R^i and p_R^i for the templates and priors of R, here superscript i is the index of the respective template and prior, beginning from 1. We also assume that the priors and templates are ordered from best to worst, so t_G^1 is the best template from the General model for some given molecule and p_G^1 the prior of that template.

At the beginning of the thesis two options were explored. The first was a policy only using the General model for every molecule, this was named the **Standard policy** and taken as our baseline.

The second option was to train a classification model that could predict whether to use the General or RB model. If we name our classification model C and a given molecule m then our model would predict that we should apply the RB model with certainty $C(m)$ with the certainty of the inverse being $1 - C(m)$. This classification model would then be

implemented in a policy and called **Classification policy**. There is an additional property available for optimization, it is the strictness with which we wish to apply the General or RB model. We could use a threshold of $\frac{1}{2}$ This policy and its experiments are in Section 4.1

Two other policies were explored later in the thesis, both ways of combining the predicted templates without an auxiliary model like in the classification policy. The **Concatenated** policy is currently used in the production version of AiZynthFinder. This uses both models for prediction at every step and takes the top N of both predicted lists to expand into new nodes.

The last method explored was the **Naive combine** policy. This policy uses both models for prediction at every step and then performs the softmax function on the *individual* lists. The lists are then concatenated, ordered again and the top N are then expanded into new nodes. The softmax function normalizes all elements in the lists so that the absolute scoring from the policy models are overriden and ignored. This gives each model the same amount of influence over which templates get selected and could produce more varied template selection.

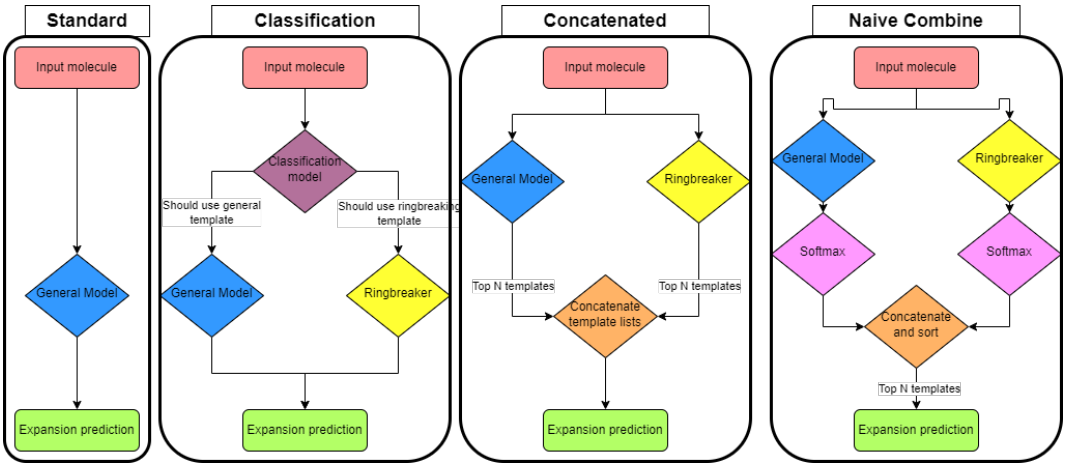

Figure 8: A diagram indicating the flow of data through the individual policies explained in Section 4.

Developing the classification model is a task in and of itself and thus requires a proper problem statement:

> *Classification problem: Can we develop a model to accurately identify, for a given molecule, whether a ring breaking reaction is advantageous or not?*

4.1 Classification

As the classification policy needs a separate model, this section will go into the process of that. At every iteration this model would take a novel molecule and predict whether a ring breaking template was advantageous to apply. The molecules at every step are novel and thus we need a method that can analyze a general molecular structure and give a reliable prediction.

The classification model will serve as a proxy for a chemist working in CASP and needs to decide whether to use a ring breaking reaction or not.

4.2 Classification data

We have a dataset of synthesis pathways created by human chemists available. We used the same method published in the PaRoutes paper[45] to gather routes extracted from literature.[33] Of this set of extracted routes we used broader limitations, such as not filtering routes that had overlapping steps and including routes that had only one step, to end up with as many routes as possible. These pathways contained information on the products, reactants and type of every reaction in the pathway, including whether these reactions broke a ring structure, i.e. were ring breaking. Whether a reaction was ring breaking was determined by using the RDKit software.[22] This dataset contains 465.884 synthesis pathways with 6.932.478 individual reactions. Of these reactions $\approx$ 18% or 1.247.846 are ring breaking. For balancing of the dataset we applied oversampling on the available ring breaking reactions to get to a dataset containing 11.3 million reactions. This dataset was randomly split 80/10/10 for training/validation/testing.

4.3 Classification model

Whether to use a specific reaction in the process of deconstruction is an inherently structural question where the global structure it very important. Thus we need to take this into account with the model that we use. To give an example, imagine a small fragment present on the outside of a ring structure needing to be removed before the ring structure can be broken efficiently. This fragment could have a negligible effect on the molecule-level chemistry but requires a non-ring breaking reaction to be removed. This non-ring breaking reaction would have to be applied before the ring breaking reaction, changing the optimal reaction for the initial molecule. For this reason, we chose to approach the problem with models that more naturally take the global structure of the molecules into account. The aim was to use graph neural networks (GNNs) to be able to leverage all information that we had of the molecules in question, every bond (type) and every atom (type). A number of models were used, both Graph Convolutional and Message-passing (GCN and MPNN respectively). The models that were tested and applied to the problem were:

- **TransformerConv**:[46] Officially named: Unified Message Passing model for Semi-supervised Classification. This is an advanced machine learning model specifically designed for graph representation learning. It applies the principles of self-attention and transformer architecture to graph data. Compared to traditional graph convolutional methods, TransformerConv brings several improvements. It uses self-attention mechanisms to capture global dependencies in the graph, enabling the model to effectively reason about relationships between distant nodes. Additionally, it combines feature and label propagation in this attention transformer network, leading to improved performance in tasks such as node classification and graph classification.

- **GATv2Conv:**[47] This is an advanced machine learning model designed for graph representation learning tasks. By incorporating attention mechanisms, GATv2Conv assigns varying importance weights to neighboring nodes, enabling it to effectively process information from the graph. It improves upon its predecessor: GATConv by the added dynamic attention mechanism, which increases the expressiveness of the mode. In summary, GATv2Conv is a significant improvement over GATConv, providing better modeling capabilities and delivering state-of-the-art results in graph representation learning.

- **Chemprop:**[48] This is a machine learning model tailored specifically to chemical property prediction. Its distinguishing feature lies in its use of directed message passing neural networks (D-MPNNs). By representing molecular structures as graphs,

$$\text{BCEloss} = -y \cdot \log(p) - (1 - y) \cdot \log(1 - p)$$

Figure 9: The formula for the loss function Binary Cross Entropy. Here $p = C(m)$ explained in Section 4 and y is the true label of the molecule.

Value	Chemprop	TransformerConv	GATv2Conv
Embedding size	300	100	200
Dropout	0.1	0.1	0.2
Number of convolutional layers	10	9	10
Number of dense layers	8	2	5
Number of heads	N/A	5	6
Learning rate	$3 \cdot 10^{-4}$	$6.7 \cdot 10^{-5}$	$7.4 \cdot 10^{-5}$
Weight decay*	0	0	0
$\beta 0$*	0.914	0.889	0.887
$\beta 1$*	0.992	0.969	0.958

Table 1: The optimal values of the hyperparameters tested for the 3 classification models.

Chemprop enables the effective encoding of local and global interactions. The D-MPNNs play a crucial role in this process, facilitating the passing of messages through the graph while considering the directionality of connections.

4.4 Training and experiments

For the classification model, we trained and tested the three aforementioned models. Before training we took a random subset of the available reaction data of 50k reactions for hyperparameter optimization. Using Lightning[49] we executed Bayesian Optimization with HyperBand (**BOHB**)[50] for 7 days of computation. This method was used due to the unknown influence of each of the parameters on the prediction performance. Our knowledge of how parameters were related to and affected each other was limited, so BOHB was used as it combines the benefits of Bayesian Optimization and Hyperband to make the optimization process more efficient and adaptive to local minima while limiting the number of evaluations. The hyperparameters optimized as well as the optimal values are shown in Table 4.4. All the models are implemented with a certain number of dense layers with dropout after the convolutional or message-passing layers. This is to increase the generalizability of the networks. As an optimizer AdamW was used for every model, this is a PyTorch implementation of the Decoupled Weight Decay Regularization algorithm.[51] During training and hyperparameter optimization our loss function is Binary Cross Entropy, the formula for its calculation is shown in Figure 9.

Some elaboration on which values mean what: Embedding size is the dimension of the hidden layers of the models. This functions the same way in all the models but in Chemprop it's called the Embedding dimension. Dropout is the fraction of dropout in the dense layers at the end of the networks. Number of convolutional layers is number of convolutional or message-passing layers following each other before the dense layers of the networks. Number of dense layers pertains to the number of dense layers at the end of the network. Number of heads is only applicable to TransformerConv and GATv2Conv and represents the number of attention heads producing differently weighted parts of the input. The values 'Learning rate', 'Weight decay', '$\beta 0$' and '$\beta 1$' relate directly to the values used in the AdamW optimizer.

Model	Test accuracy	Test AUC	Test F1
Chemprop	0.88	0.95	0.88
GATv2Conv	0.81	0.92	0.79
TransformerConv	0.83	0.93	0.81

Table 2: The produced accuracy, Area under curve and F1 scores of the models trained on the full training and validation dataset and tested on the full test set.

4.5 Results

After hyperparameter optimization the best sets of parameters were used for training, validating and testing the models on the respective sets of reaction data. From the results in table 4.5 we can conclude that the Chemprop model performed the best in general and it was chosen as our classification model going forward.

All the models used performed quite well, achieving a high accuracy, AUC and F1 on the test set. In our case, it is more important to use the AUC and F1 scores, as the penalty of false positives is higher than that of false negatives. To clarify, false positives are when C predicts a RB template should be used but it shouldn't, and false negatives are when C predicts a non-RB template but one should be used. Due to the RB model strictly only using RB templates and the General model not being limited to non-RB templates, the false positive situation is clearly worse. In the false negative case, the General model should still be able to predict a RB template, despite it generally undervaluing the importance of these templates. For classification this was not problematic as we focused only on producing the best model for the task but it becomes relevant when looking at the multi-step retrosynthesis process in Section 5. The F1 is a more important indicator of the value of a These high scores show that this model structure lends itself well to problems such as ours. Given this, we can reasonably conclude that we have sufficiently answered the *classification problem* defined in Section 4. We will use this model as part of an Expansion Policy and continue to the next step of our main problem in Section 5

5 Multi-step Synthesis experiments

Given that an accurate model for the classification policy is available, all four policies can be tested with AiZynthFinder. AiZynthFinder will be used to find synthesis routes for molecules that are realistic to use as target molecules in real-life. This would produce four collections of synthesis routes that can be compared based on quantitative metrics. These metrics can be used to show performance differences in AiZynthFinder. These differences would then be able to answer, more generally, the main problem defined in Section 3.1.

As mentioned, the Standard policy will be used as the baseline method, so the pathways produced by this policy and the performance characteristics of this will be considered baseline as well. The Concatenated and Naive combine policies will be run as is.

For the classification policy, a strictness parameter will be applied and tested rather than just applying the RB model when $C(m) \geq \frac{1}{2}$. This parameter s will be used to allow higher strictness for when to apply the RB model, that will hopefully counteract the negative effect that a high false positive error has on the overall performance. Different values of s must be tested to find the value that properly balances the effects of high false positive rate and actual application of the RB model.

5.1 Data

For generating a body of synthesis pathways representative of real-life queries for software such as AiZynthFinder, realistic target compounds are required. For this an internal AstraZeneca dataset containing 28k molecules, further referred to as the AZ target set, and a public dataset of 10k molecules which is a of ChEMBL molecules. The 28k molecules were internal and cannot be shown in this thesis. However, they are molecules that have actually been used for retrosynthesis planning in the past so they fit our purpose very well. The 10k subset is taken from the AiZynthTrain paper[52] and is a representative sample of relevant ChEMBL[53] molecules. This means that some filtering has been applied on the ChEMBL dataset such as filtering all molecules exceeding 700 daltons.

These molecules will be used side-by-side to generate two bodies of synthesis routes. When AiZynthFinder executes on these molecules, it will intend to find as many valid synthesis routes per molecule as possible within given computational limits. If routes are found for a molecule they are saved together with performance characteristics per molecule.

The target compounds is not the only important dataset during execution. AiZynthFinder works with a number of data sources that affect its ability to find synthesis routes for molecules. The set of templates used is perhaps most critical. There is a set of templates used for the General model and one for the RB model, as explained before the RB set of templates only contains RB templates. The source of these templates is the same, they are extracted from reaction data in the database from Reaxys[33] and the United States Patent and Trademark Office[54] (**USPTO**). Another important dataset is of the molecules that are part of the stock and are thus procurable. This is used for assessing when a set of starting materials for a synthesis route has been found. In this case a dataset from eMolecules has been used for all policies.[55]

5.2 Performance metrics

After execution, a number of routes (≥ 0) is produced for each molecule along with a number of performance characteristics. These will be used for performance assessment of the policy in general.

The following metrics were used:

- **Solution percentage**: If any synthesis pathway can be found for a molecule for which all the starting materials are in stock, this means we have succesfully found a synthesis pathway. In this case the molecule is referred to as *solved*. The percentage of molecules for which > 0 routes have been found is an accurate way of gauging the performance of multi-step retrosynthesis as, eventually, this is what we are mostly interested in.

- **Route length**: When a synthesis route in particular is examined, the length of the route is a (rough) indicator of its quality. With increasing synthesis pathway length comes an increase in reactions and thus dropping yield. This is something we wish to avoid and so a shorter synthesis route length is a positive marker for a specific synthesis pathway.

- **First solution iteration**: If AiZynthFinder finds any synthesis route for a molecule it almost always finds more than one. The number of iterations it took for the program to find the first valid synthesis route is an indicator of the ease of finding a route. Any routes found after this first route have no influence on this value.

- **Number of routes**: The total number of valid synthesis routes found for a given molecule after the search ends.

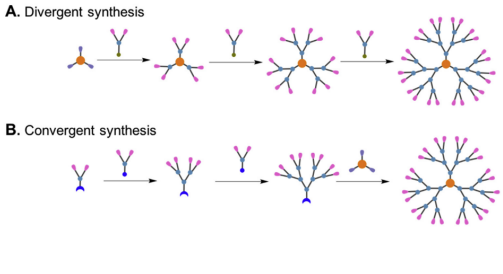

Figure 10: An example of a convergent and divergent synthesis of a dendrite. The convergent synthesis assembles parts of the target molecule in separate reactions and then puts those larger parts together in a later reaction. The divergent synthesis starts with the core of the target molecule and then adds parts step-by-step onto it, slowly increasing the size of the dendrite.

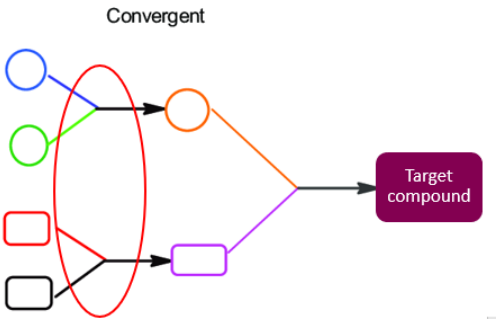

Figure 11: The way the Convergence rate is calculated is by counting reactions on the same level of the tree. For every layer we add the amount of reactions (≥ 1) and take this number as the Convergence value of a specific synthesis pathway. In this figure, the red oval indicates where in the tree these reactions would be.

- **Convergence rate**: This refers to the degree to which the synthesis pathway is Convergent.[56] A convergent synthesis is a synthesis pathway wherein multiple different parts of a molecule get assembled before being combined into the final molecule. This is counted by the total number of reactions (so > 1) in the tree that are on the same depth. The higher the number, the more efficient and thus better quality a route is. Non-convergent synthesis pathways are sometimes called Divergent but we will refer to them as regular synthesis pathways. An example of a convergent synthesis pathway and how the Convergence value is calculated per Pathway is visualized in Figures 10 and 11.

For most of these metrics the interpretation is important. The solution percentage is straightforward in its interpretation: more molecules being solved is better. However, it needs to be taken into account that models do not always solve the same molecules, even though one may have a higher solution percentage than the other. This may, for example, hint at one model being more suited for certain types of molecules. For the route count, route length, first solution iteration and convergence rate we have a similar situation, although these metrics are more dependent on the types of routes being produced. In principle these metrics have one direction that is 'better'; a higher route count, a lower route length, a faster first solution and more convergence. All these principles have pitfalls. More routes are generally better but this says nothing about the diversity or such routes, 5 very different routes can be much more useful than 20 very similar routes. Shorter routes generally have less yield, but if a route is found with one reaction, low yield and extremely expensive ingredients, it might be better to have a longer but simpler and cheaper route. A faster first solution does not say much if the first solutions found for a batch of molecules are consistently terrible across all metrics and even with convergence rate we do not always want more convergence. In fact, for some specific synthetic routes it is better to use a combination of convergence and non-convergence.[57] All these considerations are important to keep in mind when analyzing produced synthesis routes before conclusions can be drawn about the quality of produced synthesis routes.

5.3 Classification strictness

The strictness parameter s of the Classification policy must be optimized. As explained in Section 4.5, finding the optimal value for s will lead to the best performance for the Classification policy overall. For testing, a 1000 molecule subset of the ChEMBL set mentioned in Section 5.1 will be used. To gauge the performance of a specific strictness value the Solution percentage will be used as it is, in principle, the most important performance indicator in AiZynthFinder. The other properties of synthesis pathways, such as route length or first solution iteration only apply to molecules that have been solved and if less molecules are solved but the those values still rise, it might give a deceptive view of improvement. After noticing the lower accuracy than expected for $s = \frac{1}{2}$, higher values were experimented with. These indicated that only by being extremely strict with applying the ringbreaker model, did it seem to offer an actual improvement. This is reflected in Figure 12, where there is a slight but pronounced increase of the line indicating percentage of compounds solved as it approaches 100% strictness. Keep in mind that 100% strictness actually means that only the General model is applied. After all, the model is never 100% sure of its prediction for when to apply the RB model.

The best results are evidently in the region of 94% - 100% so an, albeit very strict, percentage of $s = 97\%$ was identified as most reliably being the optimal strictness value. Having established an optimally performing value for s, experiments on larger datasets and analysis thereof could be performed. Further experiments with other percentages were not performed due to simple lack of time for the project in general.

Having such a high value for s meant that the negative effects of false positive cases did not impair the solution percentage of AiZynthFinder for this policy. However, it also made the positive effect of applying the RB model much less pronounced, simply due to having it be applied so little.

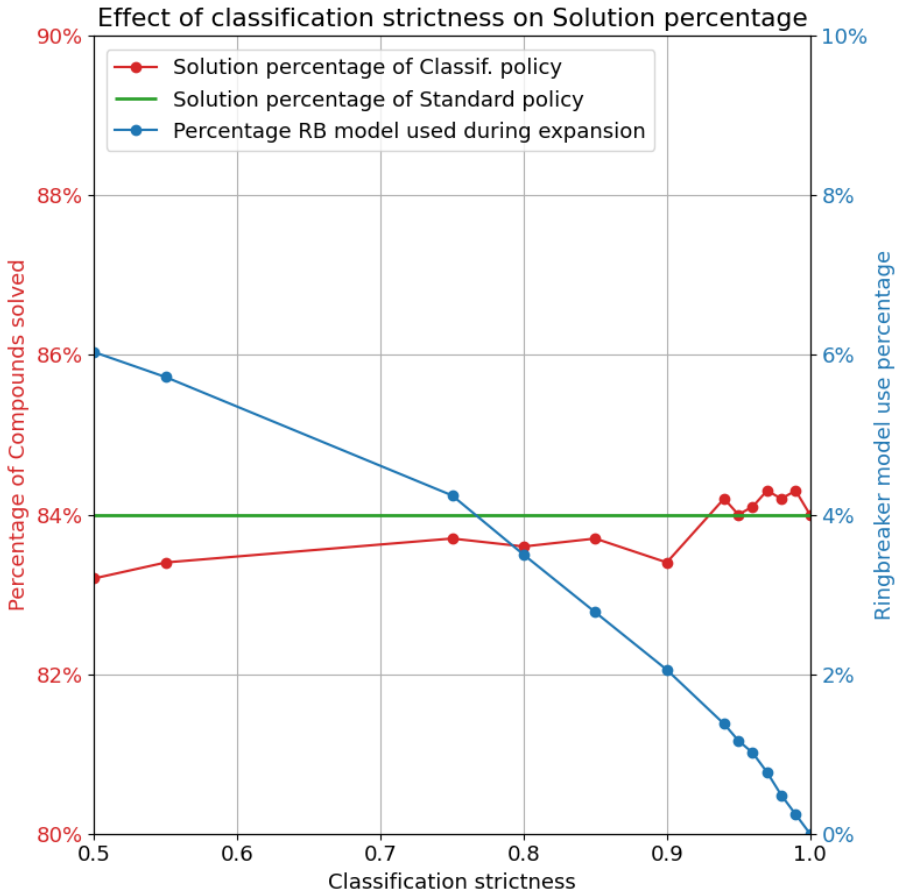

Figure 12: The strictness of the ring breaker policy as it relates to the solution percentage. This is executed on a subset of the ChEMBL dataset, containing 1000 target compounds. Note that the Classification policy with 1.0 strictness functions the same as the Standard policy, the green line indicates this for easy comparison. Additionally, the blue data points indicate the relative use of the RB model and show the inverse relation between the strictness and application of the RB model.

5.4 Multi-step Synthesis experiments

ChEMBL experiments

As mentioned before, we take the main indicator of performance for Multi-step retrosynthesis to be the solution percentage. While in some cases we value the diversity or the efficiency of routes or other properties more, generally the ability to find valid routes at all is the most important.

The four policies were executed on the ChEMBL dataset via AiZynthFinder and the performance metrics were aggregated. Figure 13 shows the resulting solution percentages of the four policies. To put these differences into perspective: The difference between the Classification policy and the Naive combining policy is 0.9%, amounting to 90 total molecules. So 90 more molecules were solved by our best policy than the worst, the Naive combining. Keep in mind that our baseline is the Standard policy, solving a mere 0.2% less molecules than our Classification, amounting to 20 molecules. This result is not as significant as we would have hoped, but it is a difference nonetheless.

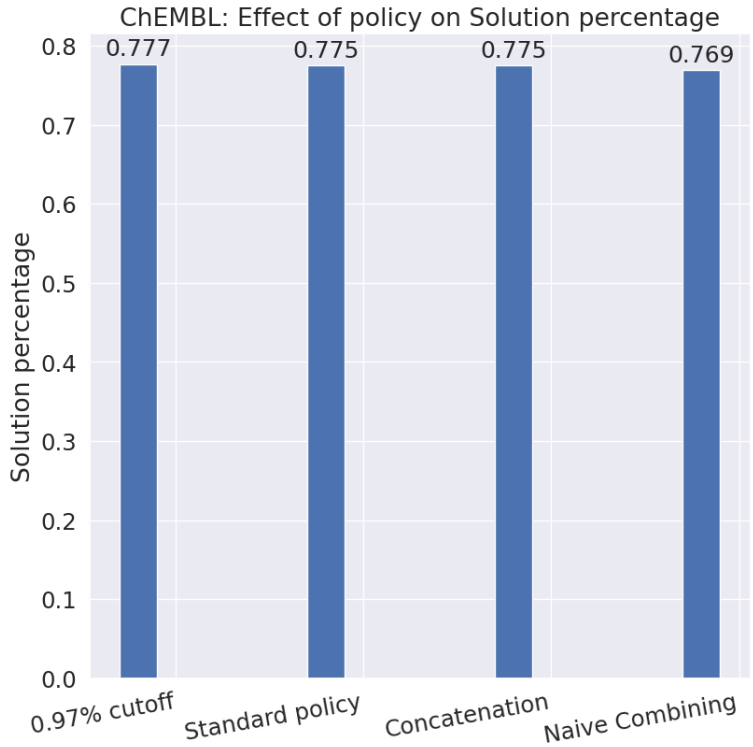

Figure 13: The ratio of solved molecules of the ChEMBL dataset (10k) plotted against the available policies.

The other performance indicators are secondary in our assessment of performance, but important nonetheless. Additionally, if there is no difference or improvement in the general solution rate of molecules then there might still be some difference or improvement in the route count per molecule, the mean route length, etc. Figure 15 shows the most important secondary performance metrics: route length, route count, first solution iteration and convergence rate. Immediately it is clear that for the convergence rate there is barely any difference, for all intents and purposes the quantitative measure of how much convergence

there is in the entire produced body of routes is the same, regardless of policy. For the route length, there is likewise a quite small difference, with only the Naive combining policy achieving generally lower scores than the rest. It is interesting that the Naive combining policy scores worse on every metric, though the difference can be small.

When the other three policies are examined in Figure 15, the differences are very small. The difference between the Classification and standard policy on the mean first solution iteration is 0.04. To put this into perspective: imagine having 9925 compounds for which the first synthesis route was found within six iterations and for 75 these were found within two iterations, this difference would amount to 100 of these compounds finding the first synthesis route four iterations sooner. While this is a significant reduction in length for these compounds, this would happen to $\approx 1\%$ of routes making it not a practically significant improvement. The same holds for the difference in the mean route count, amounting to roughly 600 compounds (of 10k) having one less total synthesis route found. The difference in performance we are most interested in is between the Classification policy and the Standard policy, due to the Classification mechanism combining our models in a non-trivial way, contrary to the Concatenation policy. The main focus of this thesis was on the Classification policy, but judging from these performance metrics, it does not seem to make a significant difference.

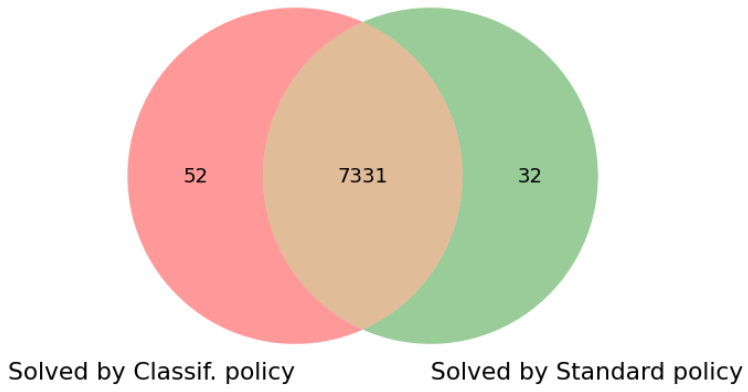

Figure 14: A Venn diagram illustrating the numerical overlap between the molecules of the ChEMBL dataset (10k) solved by the Classification policy and those solved by Standard policy. The molecules unsolved by both are now included.

Figure 14 shows the overlap between the classification policy and the Standard policy in the ChEMBL dataset. It is clear how much overlap there is in general, only $\approx 1\%$ of the total molecules are solved by one policy but not the other. If there would have been perfect overlap by one policy, so one set of solved molecules would be a subset of the other, the difference in solved molecules is still minimal. The presence of imbalance does show that the application of the Classification policy is not as effective as hoped. Given that the RB model solves single-step retrosynthesis molecules that the general model cannot,[44] it is to be expected that the multi-step performance, optimally utilizing the ring breaker model would solve strictly more molecules. This is evidently not the case, it only solves some additional molecules but at the same time misses some others.

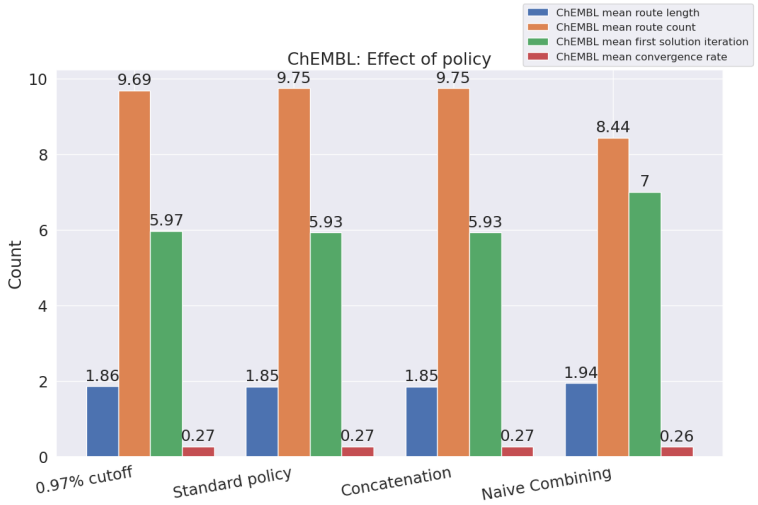

Figure 15: The mean route length, mean route count, first solution iteration and convergence rate of the ChEMBL dataset (10k) plotted against the available policies.

AZ target experiments

Due to time constraints of the project, only two of the four policies were executed on the AZ target set. Nonetheless, these were the most relevant ones, the baseline Standard policy and the Classification policy, which is the focus of the thesis. The solution percentage of the Classification is 0.2% higher than the percentage of the Standard policy. For 28k molecules this amounts to the Classification policy solving roughly 56 total molecules more than the Standard policy solves. The percentage difference for these two policies is the same for the AZ target as for the ChEMBL dataset. Both percentages are lower for the AZ target dataset, indicating that the molecules in this dataset are slightly more difficult to solve in general. This is in line with expectations, considering the source of this dataset being chemists working on state-of-the-art developments in the pharmaceutical field.

Looking at the other performance indicators shows a similar pattern as with ChEMBL. The Convergence rate and mean route length show no significant difference. The mean route count does show a small difference of 0.03, for 28k molecules this amounts to $\approx$ 1700 solved compounds having 1 route less for the Classification policy. The same goes for the first solution iteration, the difference of 0.05 means that 1400 compounds would require only 1 iteration less to find their first solution synthesis route for the Standard policy. The differences in performance metrics between the policies seem to be the same as for the ChEMBL dataset which indicates that the performance differences between the Classification and the Standard policy are consistent. This is also the case for the solution percentage. This is a positive sign, since, if the Classification model can be applied more effectively, this could lead to a general performance increase, the overall goal of developing these policies.

Concluding, even with the differences being small, it is to be expected that the molecules that are not solved by the Standard policy but are solved by the Classification Policy are generally difficult to solve, so the synthesis routes for these would be longer and perhaps have few routes per molecule. Then again, we should be careful to draw conclusions based on such small differences of solved molecules. Due to the mentioned time constraints, there was no possibility to analyze the synthesis routes of the molecules extensively. If there would have been opportunities for this points could have been identified where one or the

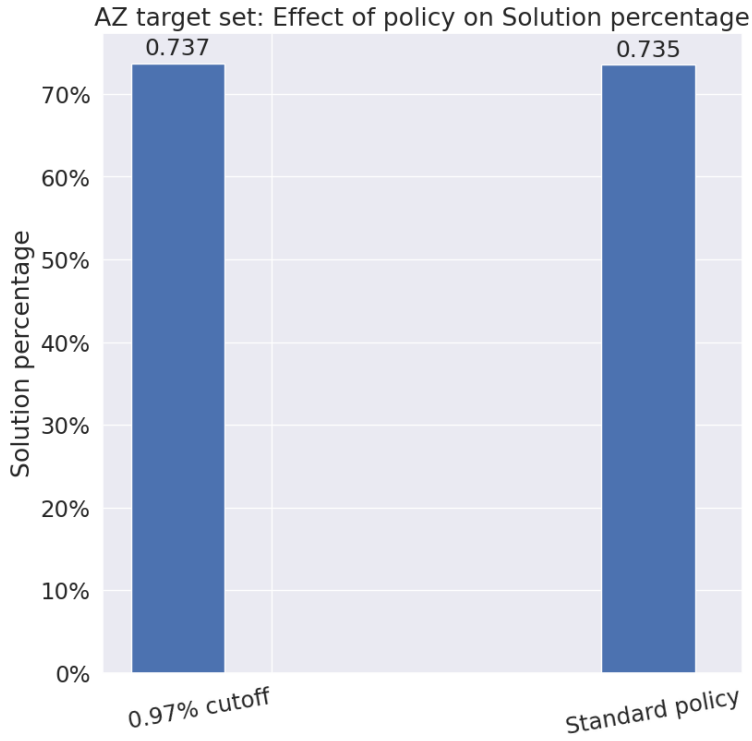

Figure 16: The ratio of solved molecules of the AZ target dataset (28k) plotted against the available policies.

other policy failed. For example by trying to identify points where only the Standard policy found valid synthesis routes, are there points where a ring breaking template would have been beneficial? If so, what prediction do we get when we run the Classification model and the RB model on those points specifically? This could have identified classes of molecules, or specific ring systems that the classification policy or the RB model fail to give a good prediction for. Comparing the synthesis routes that both policies solved is another avenue worth exploring. This could identify specific points in the synthesis route where policies gave a different prediction. Does the Standard policy still predict a ring breaker template? If not, does it predict a ring breaker further down the synthesis route? If the routes for a given molecule have very different lengths, why is this? There has been a case of a very inefficient route being produced by the application of a ring breaker template early in the retrosynthetic process, leading to a cascade of reactions, making the route so long that MCTS did not efficiently explore options from the beginning. By not using the ring breaker model entire fragments containing rings could be disconnected as starting materials, leading to an efficient synthesis pathway. Cases such as this may be common, and in-depth analysis of the produced routes is essential to properly understanding the performance of the proposed expansion policies in general.

Just like in Figure 14, in Figure 17 we have the Venn diagram visualizing the overlap of solved molecules between the Classification and Standard policies. This diagram shows the differences on the AZ target dataset. As with the performance metrics, the similarity between the ChEMBL dataset and this one is clear. There is almost exactly the same distribution between molecules only solved by the Classification policy and those only solved by the Standard policy. The pattern is the same and this motivates that we are not seeing

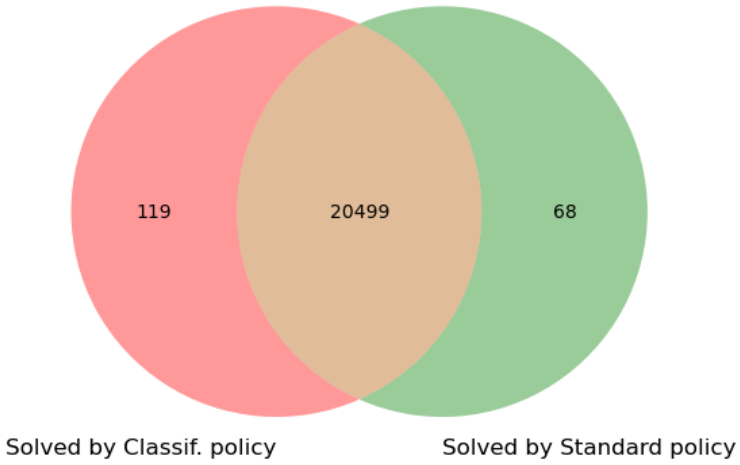

Figure 17: A Venn diagram illustrating the numerical overlap between the molecules of the AZ target dataset (28k) solved by the Classification policy and those solved by Standard policy. The molecules unsolved by both are now included.

noise. That for such a dataset $\approx 0.7\%$ of molecules will be solved by one but not by the other policy. The implications of the distribution are the same, there seems to be a big limitation in how the RB model is applied currently, but it does in principle show the possibility of improving predictions.

Figure 18: The mean route length, mean route count, first solution iteration and convergence rate of the AZ target dataset (28k) plotted against the available policies.

5.5 Conclusion

Experiments were performed on a public dataset and a private (in-house) AstraZeneca dataset. These datasets contained varied and realistic target molecules for retrosynthesis. These were chosen to mimic the use of AiZynthFinder by end-users. For these experiments AiZynthFinder was executed using the 4 different policies on the mentioned datasets, producing 4 sets of results for every dataset. The results for all the policies showed that they worked, in principle. They produced viable synthesis pathways for the vast majority of targets. When comparing the performance metrics between the policies, however, the results were less fortunate. The differences showed a non-significant difference in performance between the policies we were most interested in: the classification policy and the Standard policy. This means that the performance metrics show that the policies produced roughly the same quality synthesis routes for most of the molecules. For how we define this measure of quality, refer to Section 5.2. The absence of significant performance improvement raises the question of how to approach future research and which paths to explore more extensively.

6 Discussion

ML and AI are already indispensable in drug discovery. The practical limitations and bottle-necks of drug discovery are currently being overcome by methods from these fields. The development of these methods, both general and task-specific is therefore critically important. In this thesis we attempted to take the next step in the development of retrosynthesis models: combining single-step models. This step is necessary to further improve and combine the current developments towards better CASP. We aimed at a very distinct problem: the underrepresentation of ring breaking templates in synthesis routes generated with MCTS. This would serve as a foray into the area of combinations of expansion models or single-step retrosynthesis models that improved the performance of MCTS implementing these models. In our case this was done with the software AiZynthFinder.

The problem of underrepresentation of ring breaking templates was tackled by applying state-of-the-art graph neural networks and the already available ring breaker model.[44] This solution worked quite well and was able to confidently solve our task and generalize to unseen data. A policy combining these things was implemented and tested alongside two other policy combinations. After using a rigorous set of evaluation criteria and analyzing the results, the differences between the combinations and the baseline policy were not significant enough to draw strong conclusions about improvement of our policy combinations. Our policy seemed to be held back by an inability to effectively combine the strengths of the models and compensate for the weaknesses. We know that the RB model is very effective on specific molecules,[44] but how we identify these few molecules in particular is a challenge our policy could not solve adequately. However, the data did strongly suggest that our method works in principle. Especially the classification model seems to be effective. Perhaps a different way of applying the results from the classification model is the main thing holding increased performance back. Possibly more sophisticated methods of a multi-model expansion policy, for example a Mixture of Experts model or a gating neural network, are able to effectively divide the problem into different segments which could potentially bring about the next generation of single-step retrosynthesis models.

The insignificant results of the multi-step synthesis experiments forces us to look at several additional factors that could have contributed to this result, not all of which are within the scope of our thesis. First and foremost, the evaluation criteria of Monte Carlo Tree Search used in the architecture of AiZynthFinder could be limiting the proper applications of better expansion policies. AiZynthFinder uses information from routes that have been

(partially) solved and the backpropagation of these solutions to evaluate the expansion of new nodes. This scoring works very well for the current application of AiZynthFinder, but perhaps more sophisticated ways of combining multiple metrics would have to be applied to improve the areas currently lacking in accuracy, e.g. the utilization of ring breaking reactions. Other criteria or ways of estimating the value of a node without knowing of subsequent nodes could add great value to this process. This also concerns how these policy models are applied and how they are used in the context of the entire retrosynthetic process. If this is a limiting factor then our approach to dynamically select the expansion policy is inherently limited and developing sophisticated new single-step models would not necessarily lead to a significant improvement. This, much broader, direction of research was impossible to explore in addition to the work performed. Another, more directly relevant question is whether the accuracy of our classification model, while high on its own, is below some threshold that would lead to significant improvement in AiZynthFinder performance. The fact that the strictness of the policy model needed to be very high (97%) to achieve good results hints that this could be a factor. After all, if we properly utilized the strengths of both models, this kind of compensation would not have been necessary. Despite extensive experimentation, the accuracy of the model could not be improved further. Perhaps other methods or additions to our model could improve the results, unfortunately these are outside the scope of this thesis.

This lack of significant results for the second part of our problem does not limit the achievements in and of itself, we have succesfully explored a very concrete problem of AiZynthFinder for chemists using it at the moment. We have shown that applying a GNN to a problem like ours can reliable and generally model its structure and thus be a step in the further development of this technique. In the case of AiZynthFinder, the current avenue, improving the expansion policy to improve the overall performance is not without reward. This thesis showcases that this area can be explored systematically, decomposed into concrete problems and provide valuable insights into how to approach these problems in the future. This points in the direction of more promising research, such as more sophisticated combination of policy models. But less so in expanding the classification model as is. It may serve as an example of how to continue research in the field of retrosynthesis.

Future developments may explore the current MCTS implementation of AiZynthFinder or focus on entirely different approaches of retrosynthesis prediction. Promising questions in this area are focused around combinations as an alternative to the fixed policy model. Parameter changes or changes to the node scoring in MCTS are also worth exploring. Another, perhaps more complex, avenue is testing new policy model combinations together with added parameters in the scoring of nodes or transformations.

Concluding, it can be said that retrosynthesis and its parts are an exciting part of current research and there is a wealth of research questions just waiting to be explored. It seems that we are at the precipice of making truly big developments to the retrosynthesis space and produce methods that will be able to reach human-like performance on retrosythesis in general.

Acknowledgements

Firstly, I would like to acknowledge the advice, contributions and enthusiasm of my Industrial supervisor, Samuel Genheden. Throughout the process of writing this thesis he was consistently optimistic, motivating and funny (a valuable skill), despite the result of the thesis. I have learned much from my interactions with him about science, the industry and Sweden and hope I will be able to interact with him in the future, if only for the sci-fi novel recommendations.

I would also like to express my gratitude to Mike Preuss, my academic supervisor for his invaluable input during the coordination and writing phase of my internship in Sweden. From him I have learned quite a bit, not just about thesis-writing, but about research and science in general.

Thirdly, I would like to thank Alan Kai Hassen as my third supervisor. His advice and thoughts on Science, careers and everything else was very useful and very welcome.

Lastly, I express my gratitude to all the new friends I've made, especially Michael and Nadia and my appreciation to the entire research team at AstraZeneca, they are a wonderful bunch who are great at what they do and somehow still have time to explain things to a novice student such as myself. The environment at AstraZeneca is very motivating and constructive and inspires me to continue learning. I hope I get to spend more time in it in the future.

References

[1] K. C. Nicolaou and Jason S. Chen. *Classics in Total Synthesis III*. John Wiley & Sons, Mar 2011.

[2] E.J. Corey and W. Todd Wipke. Computer-assisted design of complex organic syntheses: Pathways for molecular synthesis can be devised with a computer and equipment for graphical communication. *Science*, 166(3902):178–192, 1969.

[3] Elias James Corey. The logic of chemical synthesis: multistep synthesis of complex carbogenic molecules (nobel lecture). *Angewandte Chemie International Edition in English*, 30(5):455–465, 1991.

[4] Steven S. Wesolowski and Dean G. Brown. The strategies wesolowski brown 2016 politics of successful design, make, test, and analyze (dmta) cycles in lead generation. *Lead Generation*, pages 487–512, 2016.

[5] Suresh Dara, Swetha Dhamercherla, Surender Singh Jadav, CH Madhu Babu, and Mohamed Jawed Ahsan. Machine learning in drug discovery: a review. *Artificial Intelligence Review*, 55(3):1947–1999, 2022.

[6] R.S.K. Vijayan, Jan Kihlberg, Jason B. Cross, and Vasanthanathan Poongavanam. Enhancing preclinical drug discovery with artificial intelligence. *Drug discovery today*, 27(4):967–984, 2022.

[7] Yuemin Bian and Xiang-Qun Xie. Generative chemistry: drug discovery with deep learning generative models. *Journal of Molecular Modeling*, 27:1–18, 2021.

[8] Connor W. Coley, William H. Green, and Klavs F. Jensen. Machine learning in computer-aided synthesis planning. *Accounts of chemical research*, 51(5):1281–1289, 2018.

[9] Daniel Vella and Jean-Paul Ebejer. Few-shot learning for low-data drug discovery. *Journal of Chemical Information and Modeling*, 63(1):27–42, 2022.

[10] Han Altae-Tran, Bharath Ramsundar, Aneesh S. Pappu, and Vijay Pande. Low data drug discovery with one-shot learning. *ACS central science*, 3(4):283–293, 2017.

[11] Thomas M. Whitehead, Benedict W.J. Irwin, P. Hunt, Matthew D. Segall, and Gareth John Conduit. Imputation of assay bioactivity data using deep learning. *Journal of chemical information and modeling*, 59(3):1197–1204, 2019.

[12] Marcus Olivecrona, Thomas Blaschke, Ola Engkvist, and Hongming Chen. Molecular de-novo design through deep reinforcement learning. *Journal of cheminformatics*, 9(1):1–14, 2017.

[13] Ricardo Macarron, Martyn N. Banks, Dejan Bojanic, David J. Burns, Dragan A. Cirovic, Tina Garyantes, Darren V.S. Green, Robert P. Hertzberg, William P. Janzen, Jeff W. Paslay, et al. Impact of high-throughput screening in biomedical research. *Nature reviews Drug discovery*, 10(3):188–195, 2011.

[14] Antonio Lavecchia. Machine-learning approaches in drug discovery: methods and applications. *Drug discovery today*, 20(3):318–331, 2015.

[15] Gisbert Schneider and Uli Fechner. Computer-based de novo design of drug-like molecules. *Nature Reviews Drug Discovery*, 4(8):649–663, 2005.

[16] Samuel Genheden, Amol Thakkar, Veronika Chadimová, Jean-Louis Reymond, Ola Engkvist, and Esben Bjerrum. Aizynthfinder: a fast, robust and flexible open-source software for retrosynthetic planning. *Journal of cheminformatics*, 12(1):70, 2020.

[17] Maud Parrot, Hamza Tajmouati, Vinicius Barros Ribeiro da Silva, Brian Atwood, Robin Fourcade, Yann Gaston-Mathé, Nicolas Do Huu, and Quentin Perron. Integrating synthetic accessibility with ai-based generative drug design. 2023.

[18] Marwin H.S. Segler, Mike Preuss, and Mark P. Waller. Planning chemical syntheses with deep neural networks and symbolic ai. *Nature*, 555(7698):604–610, 2018.

[19] Philipp Seidl, Philipp Renz, Natalia Dyubankova, Paulo Neves, Jonas Verhoeven, Jorg K. Wegner, Marwin Segler, Sepp Hochreiter, and Gunter Klambauer. Improving few-and zero-shot reaction template prediction using modern hopfield networks. *Journal of chemical information and modeling*, 62(9):2111–2120, 2022.

[20] 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, 2017.

[21] 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, 2019.

[22] RDKit: Open-source cheminformatics; http://www.rdkit.org.

[23] International Union of Pure and Applied Chemistry. Iupac compendium of chemical terminology – the gold book, 2009.

[24] Roman Kombarov, Andrea Altieri, Dmitry Genis, Mikhail Kirpichenok, Valeriy Kochubey, Natalia Rakitina, and Zoya Titarenko. Biocores: identification of a drug/natural product-based privileged structural motif for small-molecule lead discovery. *Molecular Diversity*, 14(1):193–200, May 2009.

[25] Jonghoon Kim, Heejun Kim, and Seung Bum Park. Privileged structures: Efficient chemical "navigators" toward unexplored biologically relevant chemical spaces. *Journal of the American Chemical Society*, 136(42):14629–14638, Oct 2014.

[26] Peter Ertl. Database of bioactive ring systems with calculated properties and its use in bioisosteric design and scaffold hopping. *Bioorganic and Medicinal Chemistry*, 20(18):5436–5442, Sep 2012.

[27] Roberto Todeschini and Viviana Consonni. Molecular descriptors for chemoinformatics. *Methods and Principles in Medicinal Chemistry*, Jul 2009.

[28] Laurianne David, Amol Thakkar, Rocío Mercado, and Ola Engkvist. Molecular representations in ai-driven drug discovery: a review and practical guide. *Journal of Cheminformatics*, 12(1), Sep 2020.

[29] Yoshua Bengio, Aaron Courville, and Pascal Vincent. Representation learning: A review and new perspectives. arxiv 2012. *arXiv preprint arXiv:1206.5538*, 2012.

[30] 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.

[31] Christopher Franz, Georg Mogk, Thomas Mrziglod, and Kevin Schewior. Completeness and diversity in depth-first proof-number search with applications to retrosynthesis, 2022.

[32] Binghong Chen, Chengtao Li, Hanjun Dai, and Le Song. Retro*: learning retrosynthetic planning with neural guided a* search. In *International Conference on Machine Learning*, pages 1608–1616. 2020.

[33] Reaxys, Copyright 2019 Elsevier Limited except certain content provided by third parties. Reaxys is a trademark of Elsevier Limited.

[34] J. Mayfield, D. Lowe and R. Sayle, Pistachio. 2.0 edn, Software, 2018.

[35] Xiwen Jia, Allyson Lynch, Yuheng Huang, Matthew Danielson, Immaculate Lang'at, Alexander Milder, Aaron E. Ruby, Hao Wang, Sorelle A. Friedler, Alexander J. Norquist, et al. Anthropogenic biases in chemical reaction data hinder exploratory inorganic synthesis. *Nature*, 573(7773):251–255, 2019.

[36] Amol Thakkar, Simon Johansson, Kjell Jorner, David Buttar, Jean-Louis Reymond, and Ola Engkvist. Artificial intelligence and automation in computer aided synthesis planning. *Reaction chemistry & engineering*, 6(1):27–51, 2021.

[37] 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.

[38] 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.

[39] 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*, pages 8818–8827. PMLR, 2020.

[40] Jingxin Dong, Mingyi Zhao, Yuansheng Liu, Yansen Su, and Xiangxiang Zeng. Deep learning in retrosynthesis planning: datasets, models and tools. *Briefings in Bioinformatics*, 23(1):bbab391, 2022.

[41] Alan Kai Hassen, Paula Torren-Peraire, Samuel Genheden, Jonas Verhoeven, Mike Preuss, and Igor V Tetko. Mind the retrosynthesis gap: Bridging the divide between single-step and multi-step retrosynthesis prediction, Oct 2022.

[42] Samuel Genheden, Amol Thakkar, Veronika Chadimová, Jean-Louis Reymond, Ola Engkvist, and Esben Bjerrum. Aizynthfinder: a fast, robust and flexible open-source software for retrosynthetic planning. *Journal of cheminformatics*, 12(1):70–70, 2020.

[43] Amol Thakkar, Nidhal Selmi, Jean-Louis Reymond, Ola Engkvist, and Esben Jannik Bjerrum. "Ring Breaker": Neural Network Driven Synthesis Prediction of the Ring System Chemical Space. *Journal of Medicinal Chemistry*, 63(16):8791–8808, August 2020.

[44] Amol Thakkar, Nidhal Selmi, Jean-Louis Reymond, Ola Engkvist, and Esben Jannik Bjerrum. "ring breaker": neural network driven synthesis prediction of the ring system chemical space. *Journal of medicinal chemistry*, 63(16):8791–8808, 2020.

[45] Samuel Genheden and Esben Bjerrum. PaRoutes: towards a framework for benchmarking retrosynthesis route predictions. *Digital Discovery*, 1(4):527–539, August 2022.

[46] Yunsheng Shi, Zhengjie Huang, Shikun Feng, Hui Zhong, Wenjin Wang, and Yu Sun. Masked label prediction: Unified message passing model for semi-supervised classification. *arXiv preprint arXiv:2009.03509*, 2020.

[47] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.

[48] Kevin Yang, Kyle Swanson, Wengong Jin, Connor Coley, Philipp Eiden, Hua Gao, Angel Guzman-Perez, Timothy Hopper, Brian Kelley, Miriam Mathea, et al. Analyzing learned molecular representations for property prediction. *Journal of chemical information and modeling*, 59(8):3370–3388, 2019.

[49] William Falcon and The PyTorch Lightning team. PyTorch Lightning, March 2019.

[50] Stefan Falkner, Aaron Klein, and Frank Hutter. Bohb: Robust and efficient hyperparameter optimization at scale. pages 1437–1446, 2018.

[51] Ilya Loshchilov and Frank Hutter. Decoupled Weight Decay Regularization. *arXiv*, November 2017.

[52] Samuel Genheden, Per-Ola Norrby, and Ola Engkvist. Aizynthtrain: robust, reproducible, and extensible pipelines for training synthesis prediction models. *Journal of Chemical Information and Modeling*, 63(7):1841–1846, 2023.

[53] David Mendez, Anna Gaulton, A. Patrícia Bento, Jon Chambers, Marleen De Veij, Eloy Félix, María Paula Magariños, Juan F. Mosquera, Prudence Mutowo, Michal Nowotka, María Gordillo-Marañón, Fiona Hunter, Laura Junco, Grace Mugumbate, Milagros Rodriguez-Lopez, Francis Atkinson, Nicolas Bosc, Chris J. Radoux, Aldo Segura-Cabrera, Anne Hersey, and Andrew R. Leach. ChEMBL: towards direct deposition of bioassay data. *Nucleic Acids Research*, 47(D1):930–940, January 2019.

[54] Lowe, D. Chemical Reactions from US Patents (1976 - Sep 2016). "`https://figshare.com/articles/Chemical_reactions_from_US_patents_1976-Sep2016_/5104873`" (accessed Apr 31, 2018).

[55] eMolecules Inc. eMolecules database download. www.emolecules.com/ [Accessed 1 January 2023].

[56] Michael Smith. *Organic Synthesis, 3rd Ed.* Academic Press, Jan 2010.

[57] Marie V. Walter and Michael Malkoch. Simplifying the synthesis of dendrimers: accelerated approaches. *Chemical Society Reviews*, 41(13):4593–4609, 2012.