Monte Carlo Tree Search for Chemistry and Biology

Université d'Angers

16 Mai 2025

Tristan Cazenave

LAMSADE, Université Paris Dauphine - PSL

Tristan.Cazenave@dauphine.fr

Outline

- Monte Carlo Tree Search
- Modeling Gene Regulatory Networks
- Nested Monte Carlo Search
- Retrosynthesis
- Drug Discovery
- Nested Rollout Policy Adaptation
- RNA Design

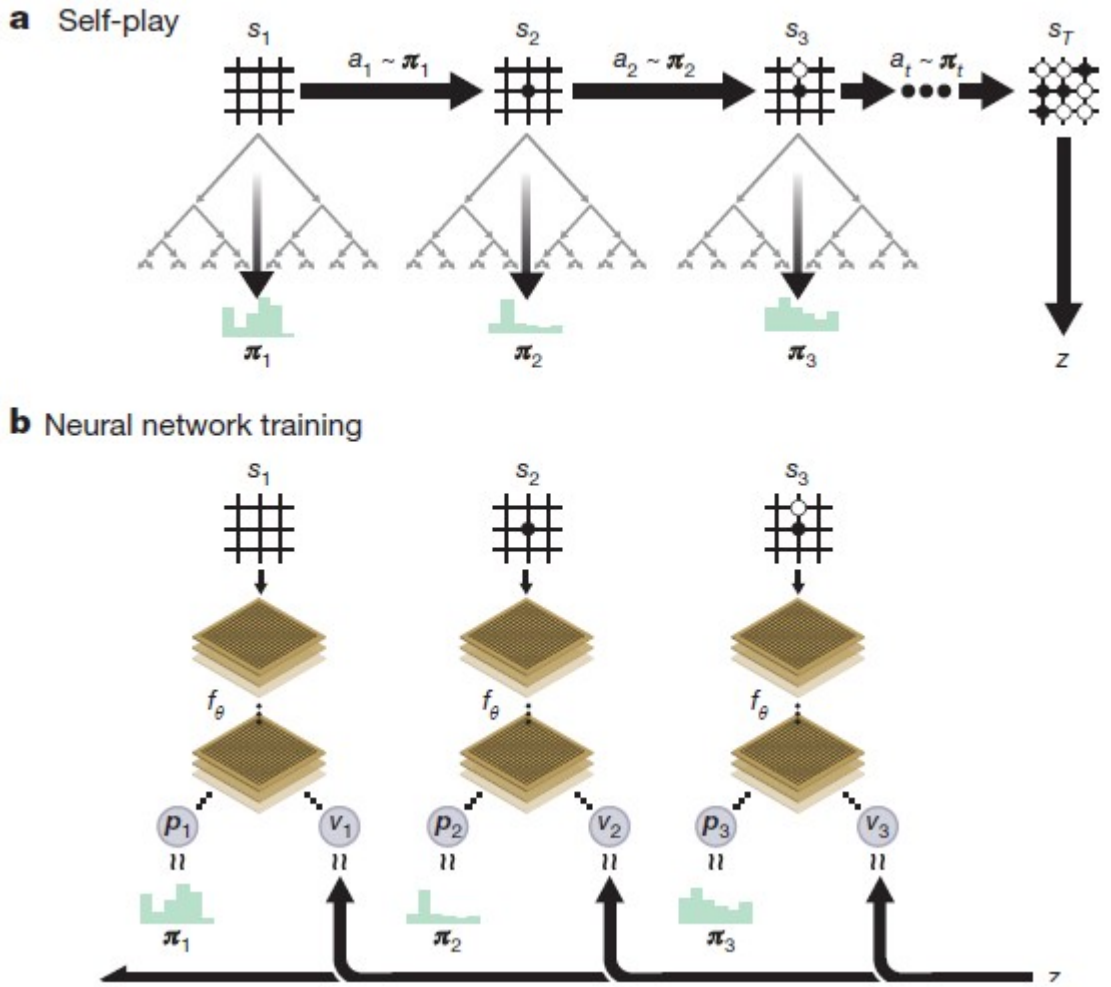
Monte Carlo Tree Search

AlphaGo

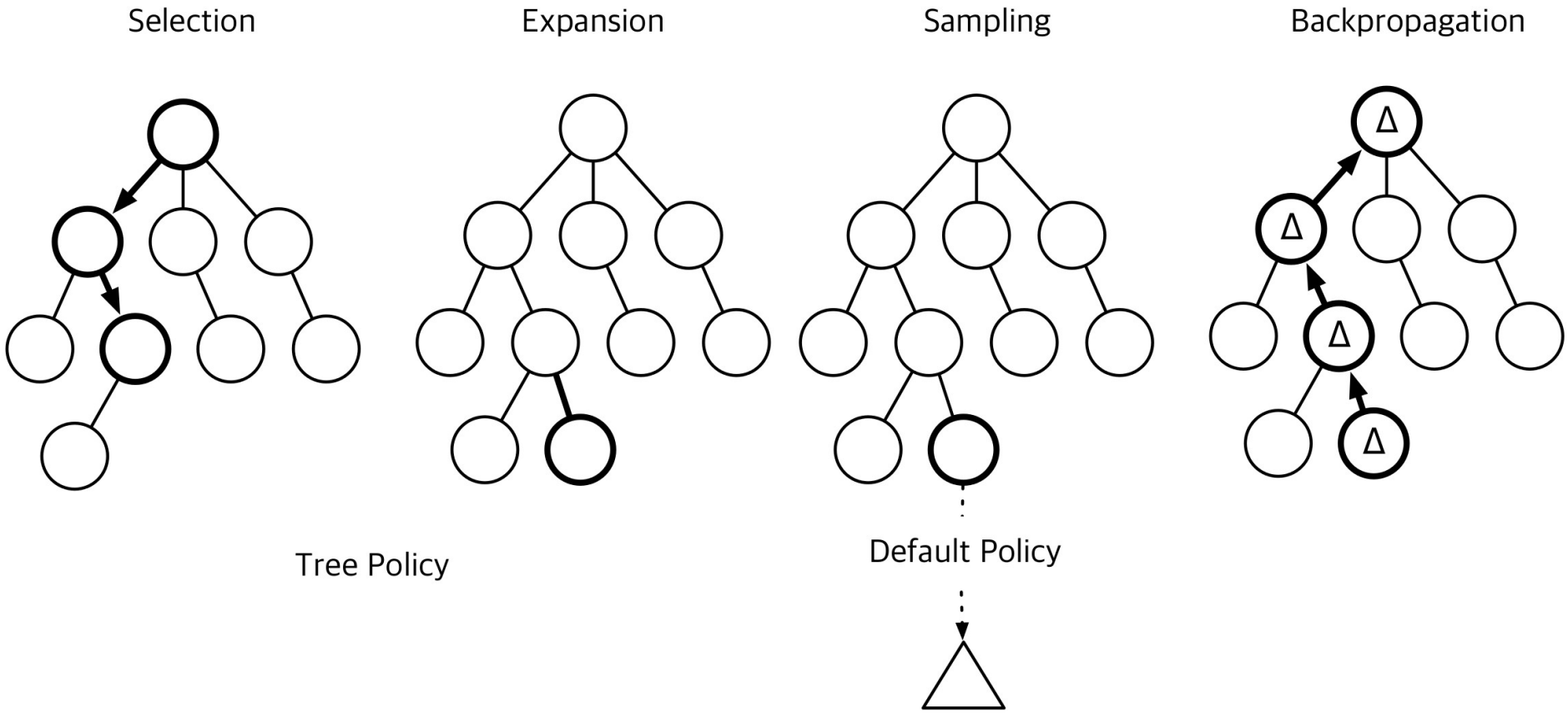
Lee Sedol is among the strongest and the most famous 9p Go player :

AlphaGo Lee won 4-1 against Lee Sedol in march 2016.

AlphaGo Zero



MCTS


Selection
Expansion
Sampling
Backpropagation
Tree Policy
Default Policy

GRAVE

- State of the art in General Game Playing (GGP)
- Best AI of the Ludii system (https://ludii.games/)
- Simple modification of RAVE
- Uses statistics both for Black and White at all nodes
- “In principle it is also possible to incorporate the AMAF values, from ancestor subtrees. However, in our experiments, combining ancestor AMAF values did not appear to confer any advantage.”

Continuous MCTS

- Infinite number of moves
- Progressive widening
- Action Decomposition (AD)
- Constraints on the possible actions (CSP)
- cRAVE and cGRAVE
- Application : Biology

Modeling Gene Regulatory Networks

Improving continuous Monte Carlo Tree Search for identifying parameters in hybrid Gene Regulatory Networks

Romain Michelucci[1], Denis Pallez[1], Tristan Cazenave[2], and Jean-Paul Comet[1]

Université Côte d'Azur, CNRS, I3S, Sophia Antipolis, France
firstname.name@univ-cotedazur.fr
LAMSADE, Université Paris Dauphine - PSL, CNRS, Paris, France
firstname.name@lamsade.dauphine.fr

Abstract. Monte-Carlo Tree Search (MCTS) is largely responsible for the improvement not only of many computer games, including Go and General Game Playing (GPP), but also of real-world continuous Markov decision process problems. MCTS initially uses the Upper Confidence bounds applied to Trees (UCT), but the Rapid Action Value Estimation (RAVE) heuristic has rapidly taken over in the discrete and continuous domains. Recently, generalized RAVE (GRAVE) outperformed such heuristics in the discrete domain. This paper is concerned with extending the GRAVE heuristic to continuous action and state spaces (cGRAVE). To enhance its performance, we suggest an action decomposition strategy to break down multidimensional actions into multiple unidimensional actions, and we propose a selective policy based on constraints that bias the playouts and select promising actions in the search tree. The approach is experimentally validated on a real-world biological problem: the goal is to identify the continuous parameters of gene regulatory networks (GRNs).

Keywords: MCTS · continuous GRAVE · constraints-based selective policy · action decomposition · chronotherapy · hybrid GRN.

1 Introduction

MCTS is a general decision-time planning algorithm that was initially designed for the improvement of computer Go [13]. The MCTS core idea is to incrementally build a search tree whose nodes represent the states of the environment and edges represent the actions taken from one state to a successor state. MCTS has proved to be effective in a wide variety of settings, including General Game Playing (GGP) [15, 23] but is not limited to games [5, 26]: it can be effective for single-agent sequential decision problems if there is an environment model simple enough for fast multistep simulation. The most popular MCTS algorithm is Upper Confidence bounds applied to Trees (UCT) [19], which addresses the exploration *versus* exploitation trade-off in each state of the tree search using the Upper Confidence Bound [1]. The Rapid Action Value Estimate [16, 17] is a

Hybrid Gene Regulatory Networks

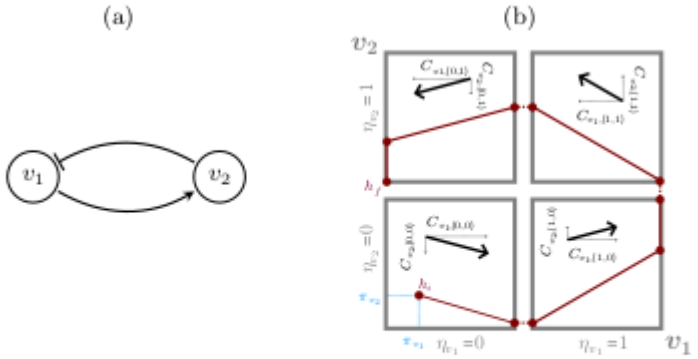


Fig. 2: Example of a hGRN depicted as a directed graph (a), and a possible hybrid state graph (b). The hGRN dynamic parameters are depicted as black arrows.

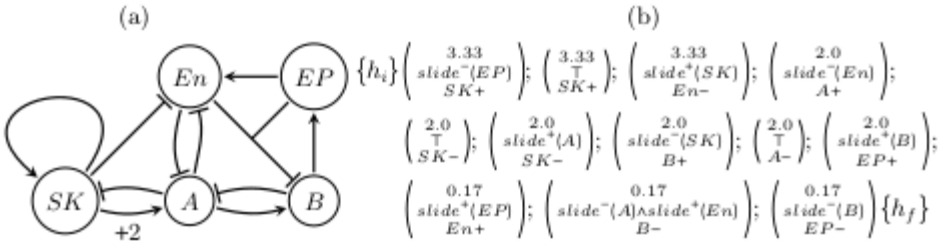


Fig. 3: Interaction graph of the 5-genes hGRN (a) and its corresponding biological knowledge (b).

cGRAVE

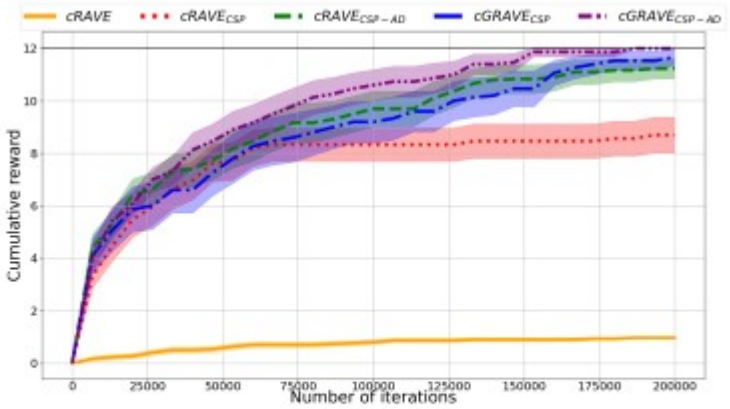
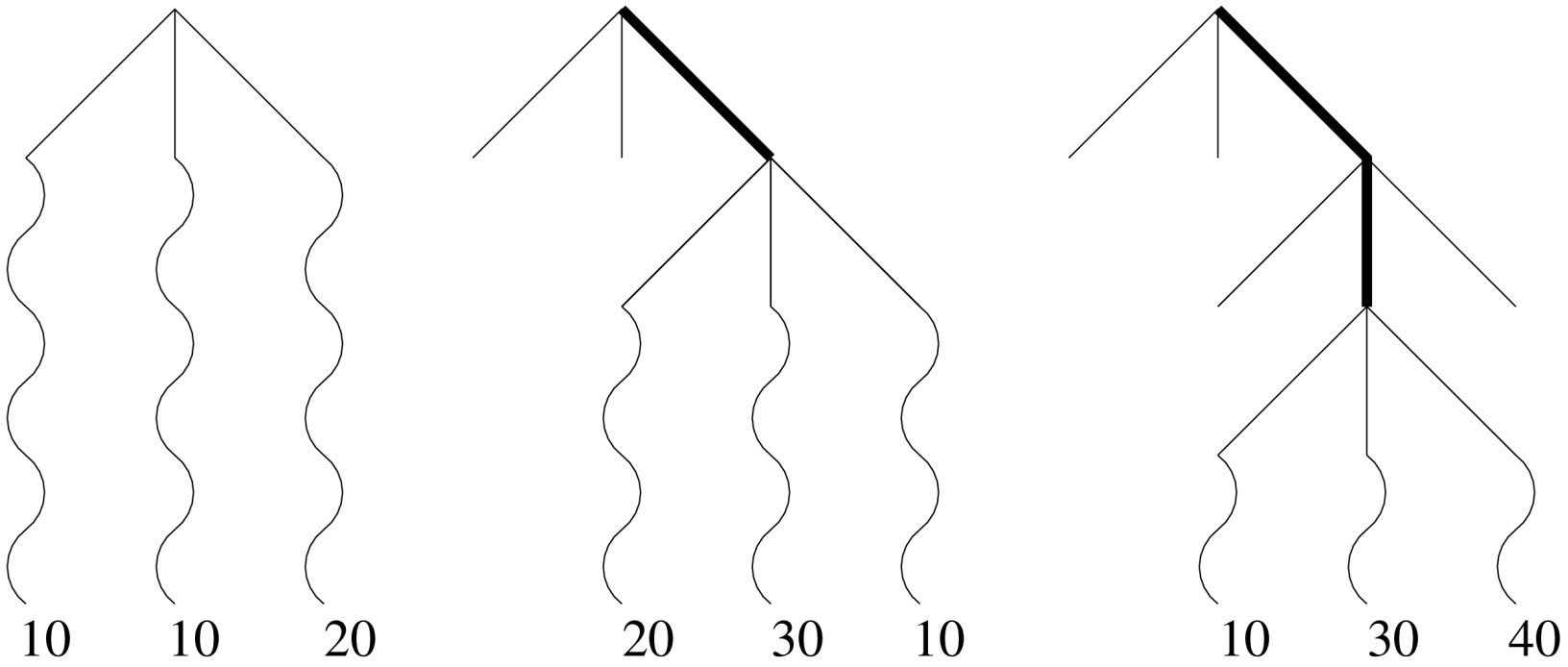


Fig. 4: Comparative performances (cumulative reward) of the different variants on the 5 genes hGRN, versus the computational budget (number of iterations). The upper the better: a reward of 12 means that a solution is found.

Nested Monte Carlo Search

Nested Monte-Carlo Search



Refutation of Spectral Graph Theory Conjectures

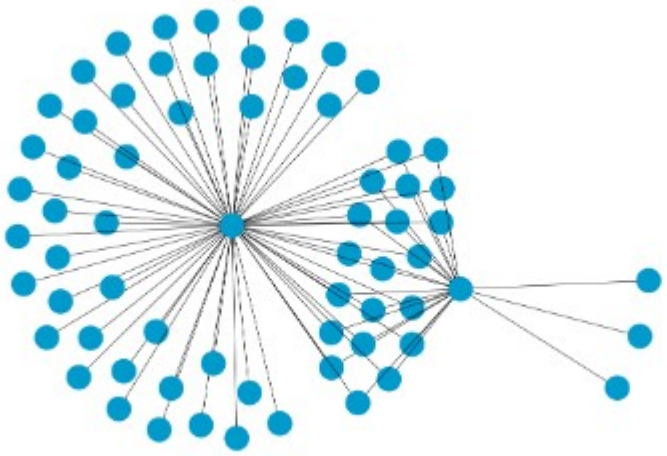


Figure 5. A counter-example of Graffiti 137 of size 67 (second largest eigenvalue $\leq$ harmonic)

- Monte Carlo Search better than Deep RL [Roucairol & Cazenave 2022]

Coalition Structure Generation

- Lazy Nested Monte Carlo Search with clever state space [Roucairol et al. 2024] :

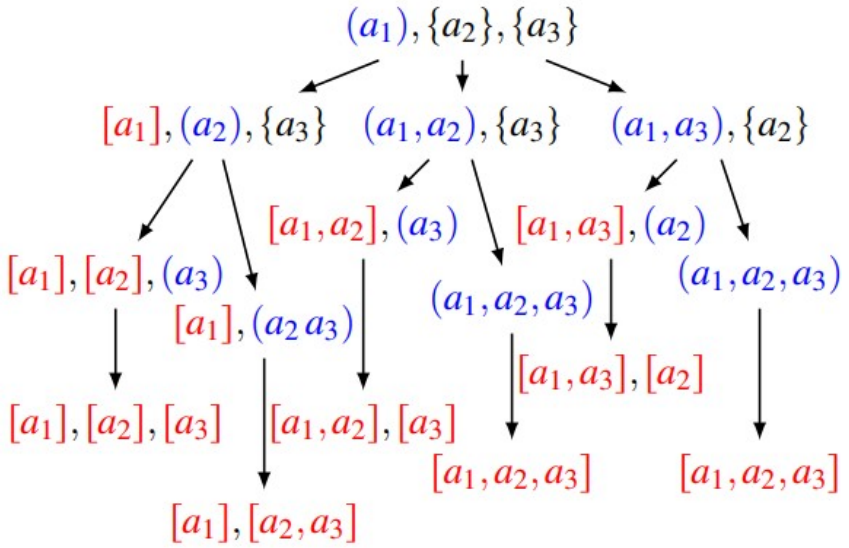


Figure 2: Model B: an example with three agents. We denote {} when the coalition is not locked and not active, () when the coalition is not locked and active, and [] when the coalition is locked.

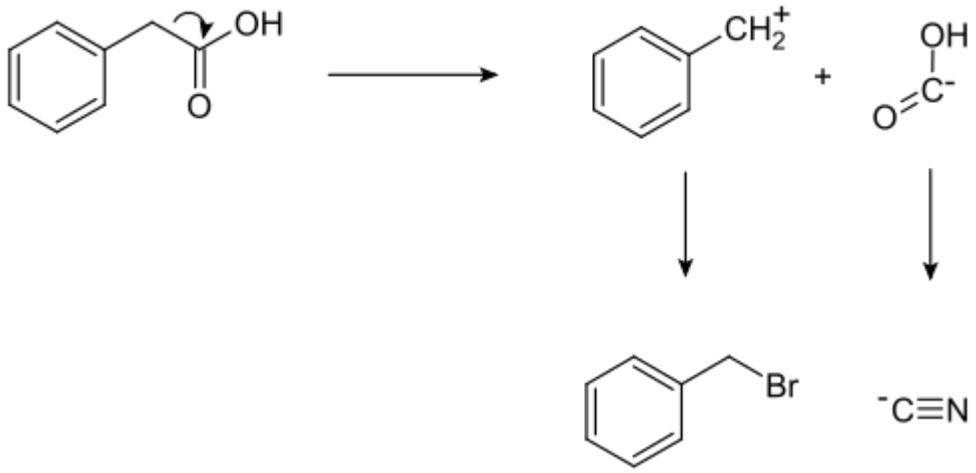
Applications

Nested Monte Carlo Search :

- Morpion Solitaire [Cazenave 2009]
- SameGame [Cazenave 2009]
- Sudoku [Cazenave 2009]
- Expression Discovery [Cazenave 2010]
- The Snake in the Box [Kinny 2012]
- Cooperative Pathfinding [Bouzy 2013]
- Software Testing [Poulding et al. 2014]
- Heuristic Model-Checking [Poulding et al. 2015]
- Pancake problem [Bouzy 2015]
- Games [Cazenave et al. 2016]
- Cryptography [Dwivedi et al. 2018]
- RNA inverse folding problem [Portela 2019]
- Perfect Rectangle Packing [Doux et al. 2022]
- Refutation of Spectral Graph Theory Conjectures [Roucairol et al. 2022]
- Coalition Structure Generation [Roucairol et al. 2024]
- Optimization of Radars [Ardon et al. 2024]
- Neural Architecture Search [Lallouet et al. 2024]
- Retrosynthesis [Roucairol et al. 2024]
- Drug-like molecule generation [Roucairol et al. 2024]
- …

Retrosynthesis

Retrosynthesis



Retrosynthesis

- Find a set of chemical reactions that enable to synthetize a given molecule.
- The state space is an AND/OR tree as in games.
- DF-PN and MCTS have been used to find retrosynthesis pathways.
- Alphachem [Segler et al. 2017].
- AiZynthFinder [Genheden et al. 2020].

Received: 26 September 2023 | Revised: 19 February 2024 | Accepted: 26 March 2024

DOI: 10.1002/minf.202300259

RESEARCH ARTICLE

molecular informatics

Comparing search algorithms on the retrosynthesis problem

Milo Roucairol | **Tristan Cazenave**

LAMSADE, Université Paris Dauphine - PSL, Paris, France

Correspondence
Milo Roucairol, LAMSADE, Université Paris Dauphine - PSL, Paris, France.
Email: milo.roucairol@dauphine.eu

Funding information
French government under the management of Agence Nationale de la Recherche, Grant/Award Number: ANR19-P3IA-0001

Abstract

In this article we try different algorithms, namely Nested Monte Carlo Search and Greedy Best First Search, on AstraZeneca's open source retrosynthetic tool : AiZynthFinder. We compare these algorithms to AiZynthFinder's base Monte Carlo Tree Search on a benchmark selected from the PubChem database and by Bayer's chemists. We show that both Nested Monte Carlo Search and Greedy Best First Search outperform AstraZeneca's Monte Carlo Tree Search, with a slight advantage for Nested Monte Carlo Search while experimenting on a playout heuristic. We also show how the search algorithms are bounded by the quality of the policy network, in order to improve our results the next step is to improve the policy network.

KEYWORDS
MCTS, Monte Carlo Tree Search, retrosynthesis, search algorithm

1 | INTRODUCTION

Retrosynthesis is a domain of organic chemistry that consists of finding a synthetic route (a sequence of reactions) for a given molecule in order to synthesize it from a given set of available precursor molecules [1]. It is an important part of organic chemistry molecule synthesis, and can be used to produce newfound drugs. What we aim for in this paper is to evaluate the strengths and weaknesses of two search algorithms by comparing them to AiZynthFinder's Monte Carlo Tree Search (MCTS) on a small benchmark consisting of curated and complex molecules, covering many reactions encountered by chemists.

The second section presents the retrosynthesis problem, the third section presents the AiZynthFinder retrosynthesis tool, the fourth section describes the search algorithms we compare, the fifth section details the benchmark used to compare the search algorithms, and the sixth section gives experimental results.

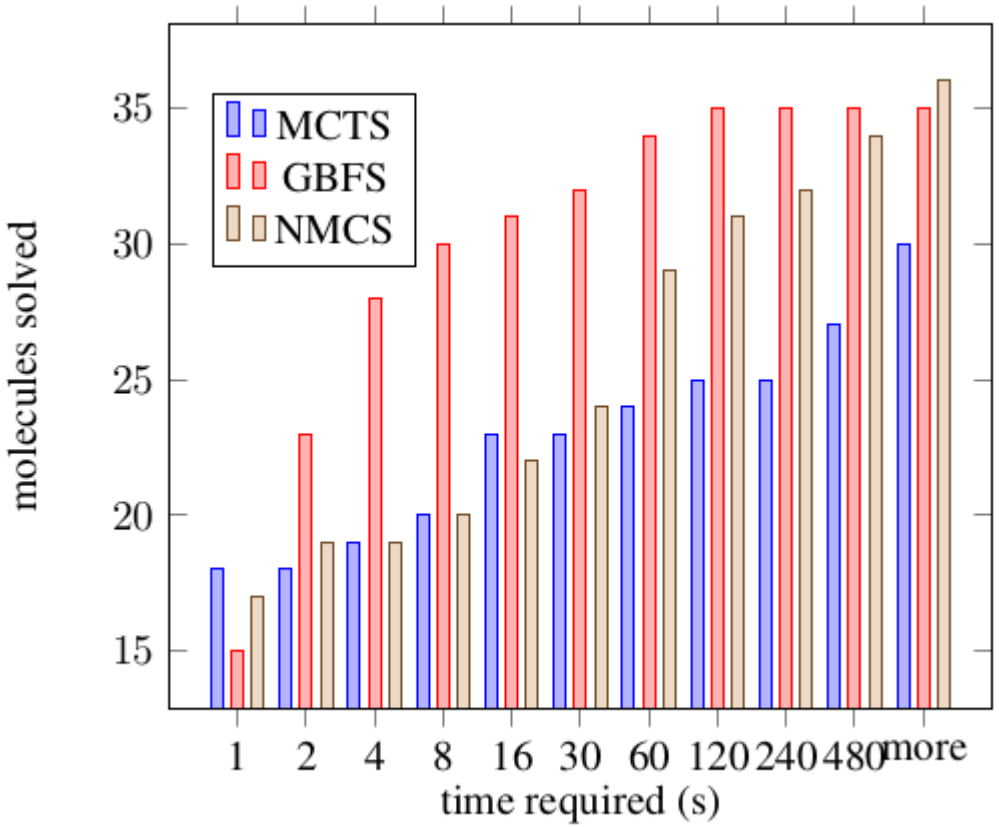
2 | THE RETROSYNTHESIS PROBLEM

Before diving into the details, let's broadly present the retrosynthesis problem.

- precursors: molecules that form one or multiple product molecules when they react together. ZINC [2] is a database of precursors that are available on the market.
- reaction template: a patent predicting the product of the reaction of one or multiple molecules. USPTO is a database of reaction template patents.
- One step retrosynthesis: an important part of retrosynthesis is selecting a few promising reaction templates before applying them as MCTS moves, this step uses a neural network.

As said before: the retrosynthetic analysis of a molecule is trying to find a sequence of reactions from a

Retrosynthesis



Drug Discovery

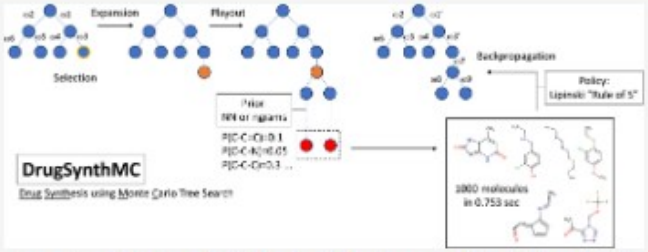
JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim
Article

DrugSynthMC: An Atom-Based Generation of Drug-like Molecules with Monte Carlo Search

Milo Roucairol, Alexios Georgiou, Tristan Cazenave,* Filippo Prischi,* and Olivier E. Pardo*

Cite This: https://doi.org/10.1021/acs.jcim.4c01451
Read Online



ABSTRACT: A growing number of deep learning (DL) methodologies have recently been developed to design novel compounds and expand the chemical space within virtual libraries. Most of these neural network approaches design molecules to specifically bind a target based on its structural information and/or knowledge of previously identified binders. Fewer attempts have been made to develop approaches for *de novo* design of virtual libraries, as synthesizability of generated molecules remains a challenge. In this work, we developed a new Monte Carlo Search (MCS) algorithm, DrugSynthMC (Drug Synthesis using Monte Carlo), in conjunction with DL and statistical-based priors to generate thousands of interpretable chemical structures and novel drug-like molecules per second. DrugSynthMC produces drug-like compounds using an atom-based search model that builds molecules as SMILES, character by character. Designed molecules follow Lipinski's "rule of 5", show a high proportion of highly water-soluble nontoxic predicted-to-be synthesizable compounds, and efficiently expand the chemical space within the libraries, without reliance on training data sets, synthesizability metrics, or enforcing during SMILES generation. Our approach can function with or without an underlying neural network and is thus easily explainable and versatile. This ease in drug-like molecule generation allows for future integration of score functions aimed at different target- or job-oriented goals. Thus, DrugSynthMC is expected to enable the functional assessment of large compound libraries covering an extensive novel chemical space, overcoming the limitations of existing drug collections. The software is available at https://github.com/RoucairolMilo/DrugSynthMC.

INTRODUCTION

Since the 1980s, *in silico* approaches have been extensively and routinely used in drug discovery and have transformed the medicinal chemistry field,[1] with expectation to do so even more in the future. The need for rapid response, highlighted by the emergence of resistant bacteria and, among others, the COVID-19 pandemic, has fueled the development of novel computational tools for drug design and screening.[2] *In silico* virtual-library screening (VS) is usually the first critical step in structure-based drug discovery, where the algorithm aims to predict the best matching binding mode of a ligand to a receptor.[3] Despite the many attempts to improve accuracy of VS methods,[4,5] the relatively limited chemical diversity of compounds in libraries reduces the ability of structure-based VS to identify hits and leads.[6,7] Indeed, it has been estimated that only a small portion (10^{8}–10^{2}) of the 10^{63} drug-like molecules predicted to be synthetically accessible has been explored.[8]

Several studies have shown that screening larger libraries that expand the accessible molecules by several order of magnitude ($\sim 10^{11}$) improves the rate of true high affinity (nM-pM) binders.[9−13] To further expand the chemical space within virtual libraries, generative models based on deep learning (DL) methodologies have been used to produce molecules

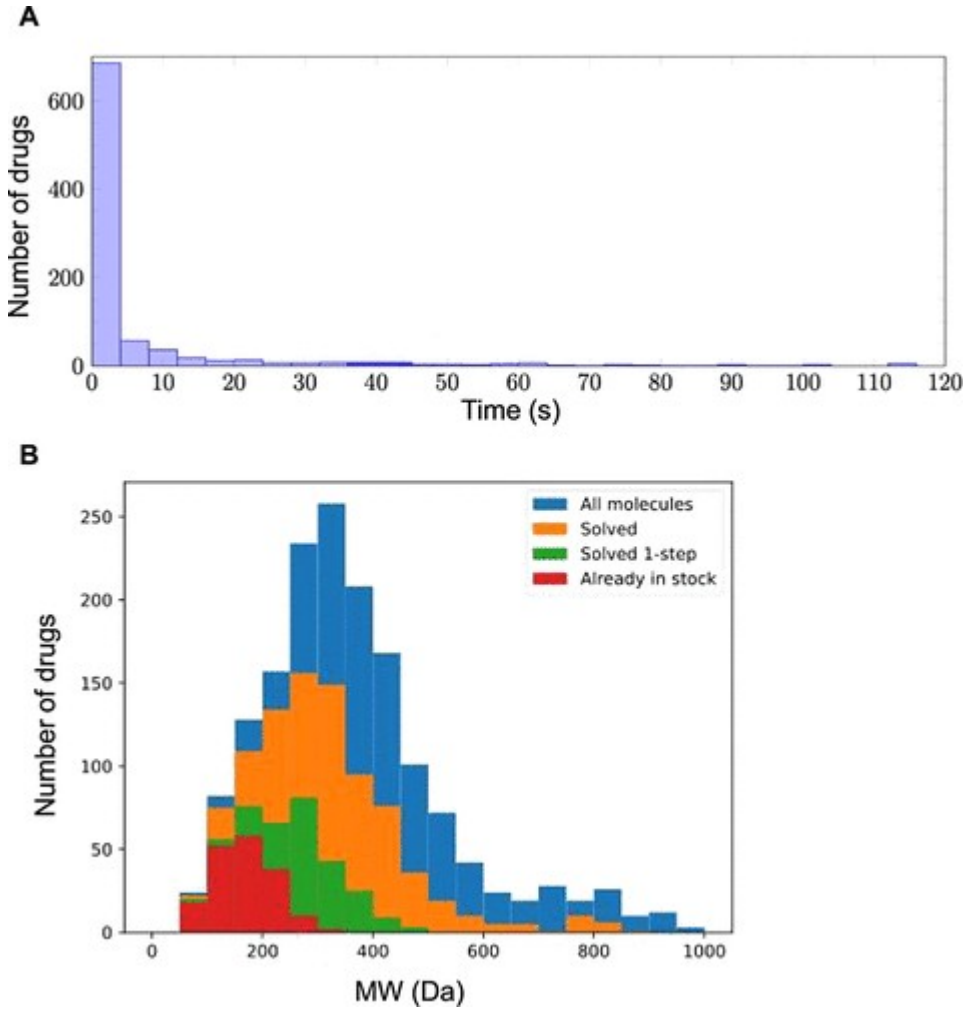
Received: August 12, 2024
Revised: August 16, 2024
Accepted: August 16, 2024

https://doi.org/10.1021/acs.jcim.4c01451
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

DrugSynthMC

- De novo design of virtual libraries
- Statistics on ngrams with the molecules of the FDA
- Lipinski rule of 5
- Synthesizability with AIZynthfinder
- Thousands of novel drug-like molecules per second
- Very small dataset used to train the ngrams (FDA)
- Future work : target oriented evaluation

DrugSynthMC

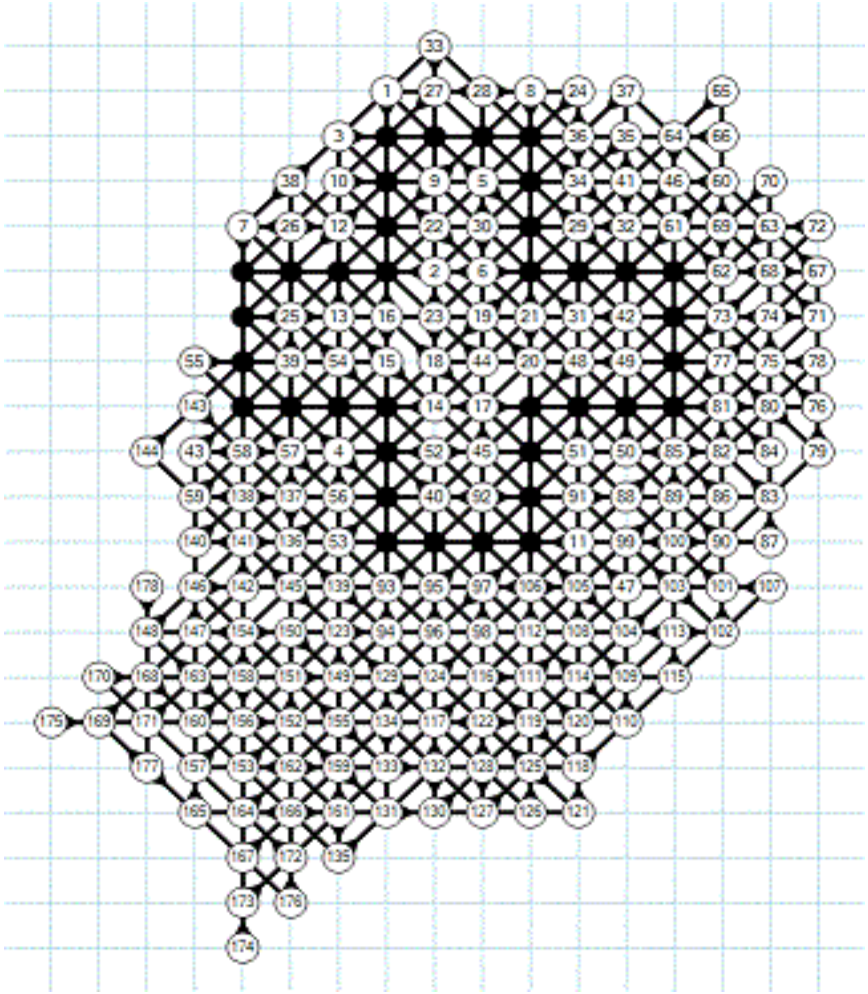


Nested Rollout Policy Adaptation

Nested Rollout Policy Adaptation

- NRPA is NMCS with policy learning.
- It uses sampling with a softmax of the move weights as a playout policy.
- There are recursive levels of best sequences as in NMCS.
- There is a policy at each level.
- The policy is reinforced on the best sequence.

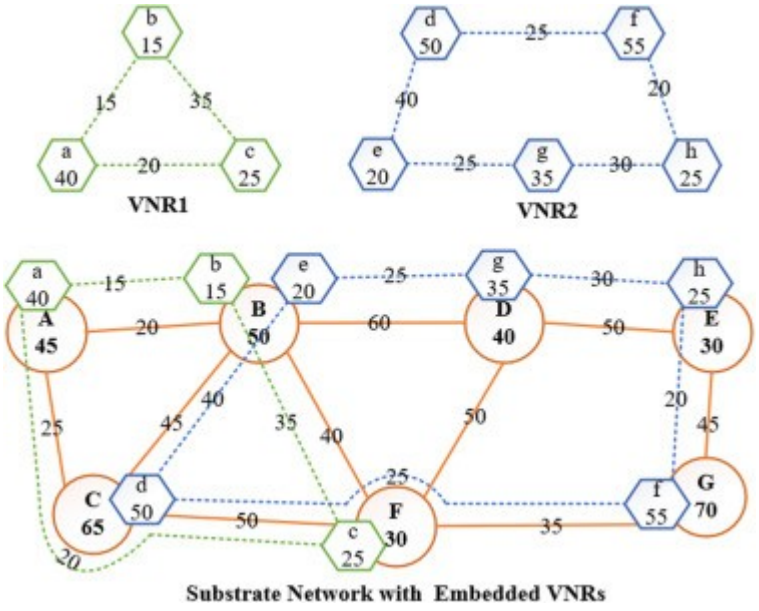
Morpion Solitaire



World record [Rosin 2011]

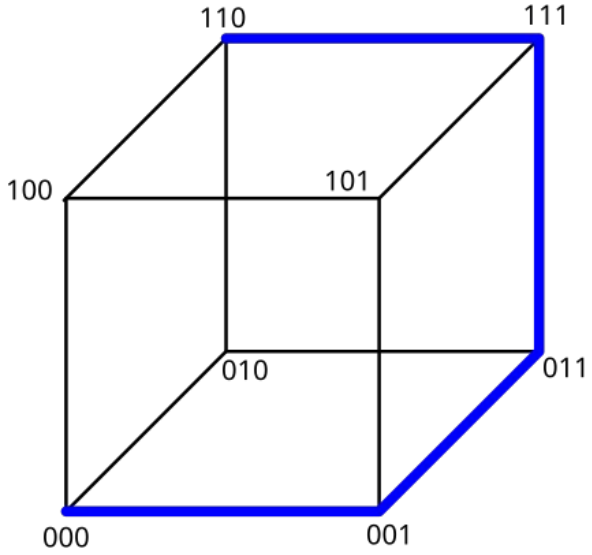
Virtual Network Embedding

- MCTS for 5G network slicing [Elkael 2023]



Snake in the Box

- Find a long path in an hypercube :



Dimension	Meta-NRPA-fe	Best Known Score
7	50	50
8	97	98
9	188	190
10	**373**	370
11	**721**	712
12	**1383**	1373
13	**2709**	2687

Table 5: Comparison of Meta-NRPA with known lower bounds on the Snake-in-the-Box

- Improved lower bounds [Dang & al. 2023]

Nested Rollout Policy Adaptation

- Morpion Solitaire [Rosin 2011]
- CrossWords [Rosin 2011]
- Traveling Salesman Problem with Time Windows [Cazenave et al. 2012]
- 3D Packing with Object Orientation [Edelkamp et al. 2014]
- Multiple Sequence Alignment [Edelkamp et al. 2015]
- SameGame [Cazenave et al. 2016]
- Vehicle Routing Problems [Edelkamp et al. 2016, Cazenave et al. 2020]
- Graph Coloring [Cazenave et al. 2020]
- RNA Design [Cazenave & Fournier 2020]
- Network Traffic Engineering [Dang & al. 2021]
- Refutation of Spectral Graph Theory Conjectures [Roucairol & Cazenave 2022]
- Slicing 5G [Elkael et al. 2023]
- Snake in the Box [Dang et al. 2023]
- Latin Square Completion and Kakuro [Cazenave 2024]
- Flexible Job Shop Scheduling [Kobrosly et al. 2025]
- …

RNA Design

RNA Design

- Molecule Design as a Search Problem
- Find the sequence of nucleotides that gives a predefined structure
- Useful for synthetic biology, medicine, and nanotechnology
- GREED-RNA: Greedy Local Search

eterna
Solve Puzzles. Invent Medicine.

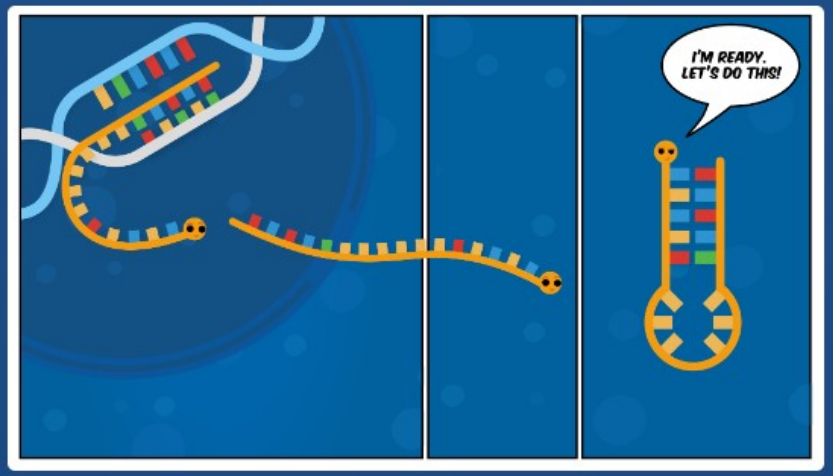
Log in Register

HOME PLAY LEARN COMMUNITY NEWS ABOUT

What is RNA?

RNAs are tiny molecules in the cells of every living thing. They copy information from DNA and use it to make things happen in the cell.

Like DNA, RNA is made up of four bases. Each RNA folds into a shape that determines its function, and the shape is defined by the pattern of the bases.



Eterna is Solved

Tristan Cazenave

LAMSADE, Université Paris Dauphine - PSL, CNRS, Paris, France

Abstract. RNA design consists of discovering a nucleotide sequence that folds into a target secondary structure. It is useful for synthetic biology, medicine, and nanotechnology. We propose Montparnasse, a Multi Objective Generalized Nested Rollout Policy Adaptation with Limited Repetition (MOGNRPALR) RNA design algorithm. It solves the Eterna benchmark.

1 Introduction

The design of molecules with specific properties is an important topic for research related to health. The RNA design problem, also named the Inverse RNA Folding problem, is a difficult combinatorial problem. This problem is important for scientific fields such as bioengineering, pharmaceutical research, biochemistry, synthetic biology, and RNA nanostructures [20].

RNA is involved in many biological functions. Synthetic RNA can be easily produced [21] and has many applications in synthetic biology, as well as in drug design with the building of riboswitches and ribozymes.

RNA design consists of finding a nucleotide sequence that folds into a desired target structure. Eterna is a standard benchmark for RNA design algorithms. Many algorithms have been applied to this problem over the years. However, none have successfully solved all the Eterna problems. This paper presents a simple algorithm that solves the Eterna benchmark.

RNA molecules are long molecules composed of four possible nucleotides. Molecules can be represented as strings composed of the four characters A (Adenine), C (Cytosine), G (Guanine), and U (Uracil). For RNA molecules of length N, the size of the state space of possible strings is exponential in N. It can be very large for long molecules. The sequence of nucleotides folds back on itself to form what is called its secondary structure. It is possible to find in polynomial time the folded structure of a given sequence. However, the opposite, which is the Inverse RNA Folding problem, is hard [2].

RNA functions are determined by its tertiary structure. The secondary structure is used to determine the tertiary structure according to the base pairing interactions. The bonds between two nucleotides are given by the six possible base pairs (CG, GC, AU, UA, UG, GU). The dot-bracket notation is used to represent the secondary structure, the opening and closing brackets represent the base pairs, and the dots represent the unbounded sites.

The paper is organized as follows: the second section is about previous attempts at designing RNA. The third section presents the algorithms used in Montparnasse. The fourth section details the experimental results.

Eterna is Solved

- Montparnasse
- Multi Objective Generalized Nested Rollout Policy Adaptation with Limited Repetitions
- Base Pair Distance (BPD), Hamming Distance, …
- Stop search at a level if the same best sequence is found a second time.
- Prior on CG, GC and A.

Eterna is Solved

(a) Gladius: problem 90

(b) Shooting Star: problem 99

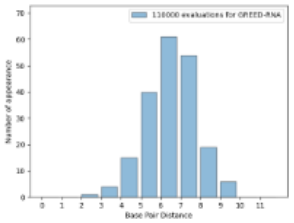
(c) Teslagon: problem 100

Eterna is Solved

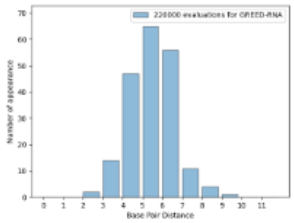
Table 1: Distributions of the BPD of the various algorithms after 270 000 evaluations for problem 99.

BPD	0	1	2	3	4	5	6	7	8
GREED-RNA	6	22	49	66	38	17	2	0	0
MOGRLS	19	46	63	39	22	7	2	2	0
PN	28	72	64	28	8	0	0	0	0
MOGNRPALR	120	78	2	0	0	0	0	0	0

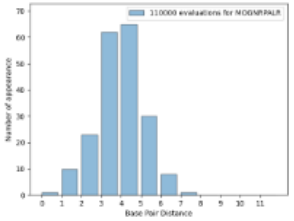
Eterna is Solved



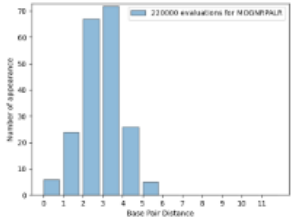
(a) Distribution of the BPD for problem 90 after 110 000 evaluations by GREED-RNA.



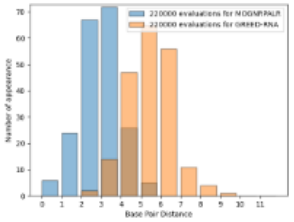
(b) Distribution of the BPD for problem 90 after 220 000 evaluations by GREED-RNA.



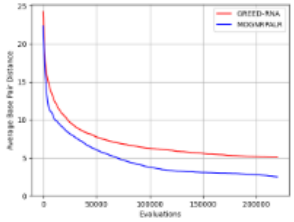
(c) Distribution of the BPD for problem 90 after 110 000 evaluations by MOGNR-PALR.



(d) Distribution of the BPD for problem 90 after 220 000 evaluations by MOGNR-PALR.



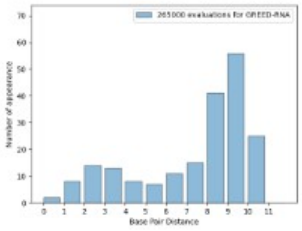
(e) Comparison of the distributions after 220 000 evaluations between GREED-RNA and MOGNRPALR.



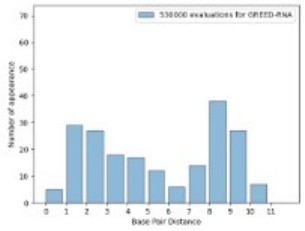
(f) Evolution of the average BPD of GREED-RNA and MOGNRPALR for problem 90.

Fig. 5: Comparison of the BPD on problem 90 for GREED-RNA and MOGNRPALR for increasing numbers of evaluations. 220 000 evaluations by one process takes one day. GREED-RNA is stuck and does not solve the problem while MOGNRPALR progresses and solves the problem 6 times out of 200 runs.

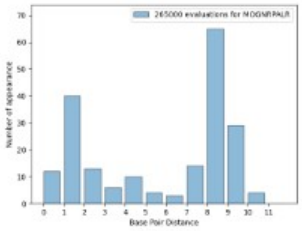
Eterna is Solved



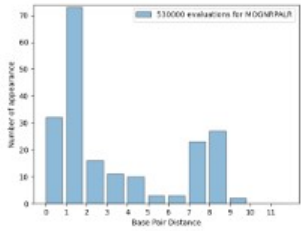
(a) Distribution of the BPD for problem 100 after 265 000 evaluations by GREED-RNA.



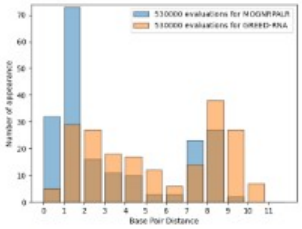
(b) Distribution of the BPD for problem 100 after 530 000 evaluations by GREED-RNA.



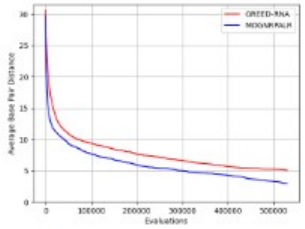
(c) Distribution of the BPD for problem 100 after 265 000 evaluations by MOGN-RPALR.



(d) Distribution of the BPD for problem 100 after 530 000 evaluations by MOGN-RPALR.



(e) Comparison of the distributions after 530 000 evaluations between GREED-RNA and MOGNRPALR.



(f) Evolution of the average BPD of GREED-RNA and MOGNRPALR for problem 100.

Fig. 6: Comparison of the BPD on problem 100 for GREED-RNA and MOGNRPALR for increasing numbers of evaluations. 530 000 evaluations by one process takes one day. GREED-RNA solves problem 100 less frequently than MOGNRPALR.

Eterna is Solved

- First time that the most difficult problems from Eterna are solved within one day.
- Eterna consists of puzzles for the secondary structure.
- Next step : 3D design.

Conclusion

- Monte Carlo Search has many applications to Chemistry and Biology:
  - Modeling Gene Regulatory Networks
  - Retrosynthesis
  - Drug Discovery
  - RNA Design