Genetic-guided GFlowNets for Sample Efficient Molecular Optimization

Hyeonah Kim, Minsu Kim, Sanghyeok Choi, Jinkyoo Park
Korea Advanced Institute of Science and Technology (KAIST)

NEURAL INFORMATION PROCESSING SYSTEMS

Summary

TL;DR. Our method enhances the sample-efficiency in molecular optimization by **distilling a powerful genetic algorithm** into deep generative policy using **GFlowNets training**.

Key ideas

- A factorized string-based generative policy
- Exploration with graph-based genetic algorithm for molecular design
- Employ generative flow networks (GFlowNets) for off-policy training

Background

Sample-efficient Molecular Optimization

- *De novo* molecular design: generate a new molecule from scratch with the desired property
- Evaluating chemical or biological properties of a molecule is expensive (a black-box oracle function f)
- Combinatorially vast space with limited oracle budget ($\leq 10K$)

Genetic algorithms for molecule design

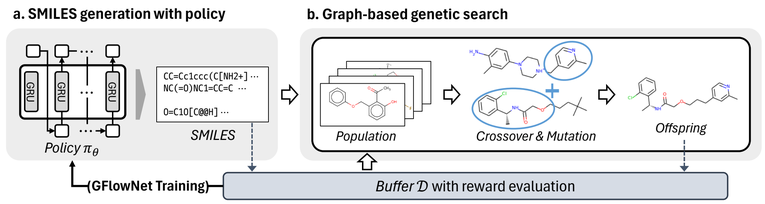
- GAs often outperform recent deep learning methods
- They iteratively evolve a population of candidates as follows:
  1. **Initialize a population**
  2. **Generate offspring**: A child is generated by *crossover* and *mutation* designed with expertise.
  3. **Select a new population** and go bact to 2.

Generative flow networks (GFlowNets)

- Sample a discrete compositional object $x \in \mathcal{X}$ proportional to its reward, i.e., $p(x) \propto R(x)$
- The forward policy P_F generates state transitions sequentially through trajectories $\tau = (s_0 \rightarrow \ldots \rightarrow s_T = x)$
- Trajectory balance loss (**Off-policy** training)

$$\mathcal{L}_{\mathsf{TB}}(\tau;\theta) = \left(\log \frac{Z_\theta \prod_{t=1}^{T} P_F(s_t|s_{t-1};\theta)}{R(x)\prod_{t=1}^{T} P_B(s_{t-1}|s_t;\theta)}\right)^2 \quad (1)$$

Genetic-guided GFlowNets



Step 1: Factorized string-based generative policy and unsupervised pre-training. The probability $\pi_\theta(\boldsymbol{x}) = \prod_{t=1}^{n} \pi_\theta(x_t|x_1, \ldots, x_{t-1})$, where $x_1, \ldots, x_n$ are characters of SMILES representation of $\boldsymbol{x}$. The policy is pretrained on existing chemical datasets $\mathcal{D}_{\text{pre}}$:

$$\mathcal{L}_{\text{pre}}(\boldsymbol{x}) = -\sum_{t=1}^{n} \log \pi_\theta(x_t|x_1, \ldots, x_{t-1}). \quad (2)$$

Step 2: Fine-tune the policy with genetic search

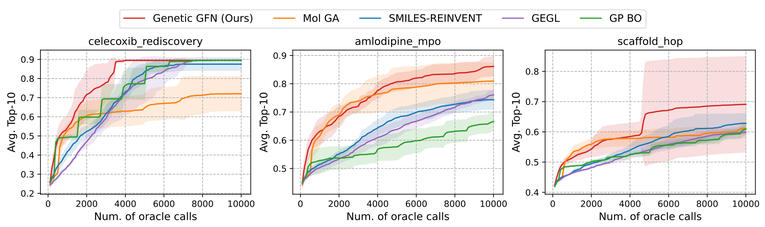
a. Generates SMILES using the policy $\pi_\theta(\boldsymbol{x})$ and evaluate
b. Select a population from the buffer based on the reward and generate offspring (graph-based crossover & mutation)
c. Train the policy with a GFlowNet loss (+ KL penalty)

$$\mathcal{L} = \mathcal{L}_{\mathsf{TB}}(\tau;\theta) + \alpha \mathsf{KL}(\pi_\theta(x)||\pi_{\text{pre}}(x)), \quad (3)$$

where $R(\boldsymbol{x}) = e^{-\beta f(\boldsymbol{x})}$.

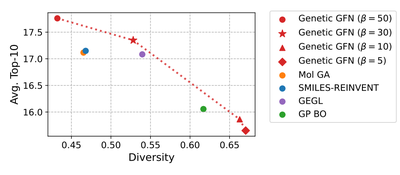
Experimental Results (1)

Main results in the Practical Molecular Optimization (PMO) benchmark



Controllability of the scores-diversity trade-off

	$\beta = 1$	$\beta = 5$	$\beta = 10$	$\beta = 30$	$\beta = 50$
AUC-10	11.083	14.597	14.735	15.815	*16.213*
Div.	0.812	0.670	0.663	0.528	*0.432*



Experimental Results (2)

Results of 23 oracles in the PMO benchmark

ID	Genetic GFN (Ours)	Mol GA [16]	SMILES REINVENT [3]	GEGL [17]
#1	**0.949 ± 0.010**	0.928 ± 0.015	0.881 ± 0.016	0.842 ± 0.019
#2	**0.761 ± 0.019**	0.740 ± 0.055	0.644 ± 0.019	0.626 ± 0.018
#3	**0.802 ± 0.029**	0.629 ± 0.062	0.717 ± 0.027	0.699 ± 0.041
#4	0.733 ± 0.109	0.656 ± 0.013	0.662 ± 0.044	0.656 ± 0.039
#5	**0.974 ± 0.006**	0.950 ± 0.004	0.957 ± 0.007	0.898 ± 0.015
#6	**0.856 ± 0.039**	0.835 ± 0.012	0.781 ± 0.013	0.769 ± 0.009
#7	0.881 ± 0.042	**0.894 ± 0.025**	0.885 ± 0.031	0.816 ± 0.027
#8	**0.969 ± 0.003**	0.926 ± 0.014	0.942 ± 0.012	0.930 ± 0.011
#9	**0.897 ± 0.007**	0.894 ± 0.005	0.838 ± 0.030	0.808 ± 0.007
#10	0.764 ± 0.069	**0.835 ± 0.040**	0.782 ± 0.029	0.580 ± 0.086
#11	**0.379 ± 0.010**	0.329 ± 0.006	0.363 ± 0.011	0.338 ± 0.016
#12	0.294 ± 0.007	0.284 ± 0.035	0.281 ± 0.002	0.274 ± 0.007
#13	0.708 ± 0.057	**0.762 ± 0.048**	0.634 ± 0.042	0.599 ± 0.035
#14	**0.860 ± 0.008**	0.853 ± 0.005	0.834 ± 0.010	0.832 ± 0.005
#15	0.595 ± 0.014	**0.610 ± 0.038**	0.535 ± 0.015	0.537 ± 0.015
#16	**0.942 ± 0.000**	0.941 ± 0.001	0.941 ± 0.000	0.941 ± 0.001
#17	0.819 ± 0.018	**0.830 ± 0.010**	0.770 ± 0.005	0.730 ± 0.011
#18	**0.615 ± 0.100**	0.568 ± 0.017	0.551 ± 0.024	0.531 ± 0.010
#19	0.634 ± 0.039	**0.677 ± 0.055**	0.470 ± 0.041	0.402 ± 0.024
#20	**0.583 ± 0.034**	0.544 ± 0.067	0.544 ± 0.026	0.515 ± 0.028
#21	**0.511 ± 0.054**	0.487 ± 0.024	0.458 ± 0.018	0.420 ± 0.031
#22	0.135 ± 0.271	0.000 ± 0.000	**0.182 ± 0.363**	0.119 ± 0.238
#23	**0.552 ± 0.033**	0.514 ± 0.033	0.533 ± 0.009	0.492 ± 0.021
Sum	**16.213**	15.686	15.185	14.354

Comparisons with GFlowNets variants

(a) Average and standard deviation of AUC scores (↑)

	SMILES		Fragment-based	
	Genetic GFN	LS-GFN [18]	GFN [6]	GFN-AL [10]
AUC Top-1	**16.530 ± 0.198**	15.514 ± 0.269	10.957 ± 0.033	11.032 ± 0.016
AUC Top-10	**16.213 ± 0.173**	15.230 ± 0.026	9.918 ± 0.027	9.928 ± 0.027
AUC Top-100	**15.516 ± 0.127**	14.619 ± 0.027	8.416 ± 0.024	8.064 ± 0.005

(b) Search distances (↑)

(GSK3β)	SMILES	Molecule
Genetic GFN	0.740	0.528
LS-GFN	0.374	0.494

(JNK3)	SMILES	Molecule
Genetic GFN	0.706	0.536
LS-GFN	0.403	0.512

Designing SARS-CoV-2 Inhibitors

Goal: maximizing the binding affinity to the target protein + QED (Quantitative Estimate of Drug-likeness) and SA (Synthetic Accessibility)

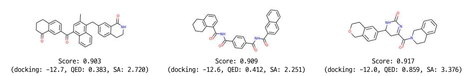


Figure 4: The final candidates for the PLPr_7JIR target with 100 steps.

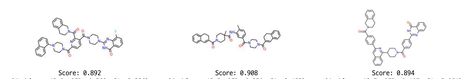


Figure 5: The final candidates for the RdRp_6YYT target with 100 steps.

Table 5: Average Top-100 scores (↑). Ours outperforms baselines with 10 times fewer steps. The **bold** denotes the best scores.

	PLPro	RdRp
JT-VAE	0.272	0.216
GFlowNet	0.326	0.280
Graph GA	0.723	0.786
REINVENT	0.717	0.799
MolRL-MGPT	0.772	0.854
Genetic GFN (100)	0.891	0.873
Genetic GFN (1000)	**0.925**	**0.902**