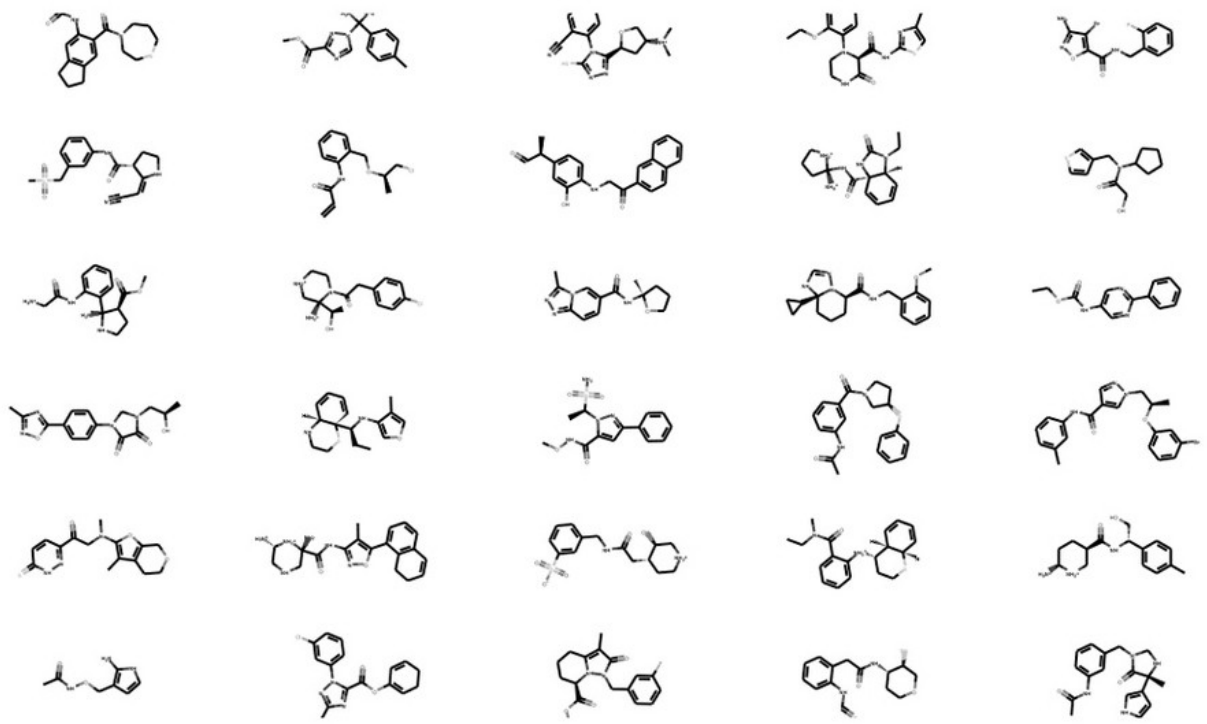
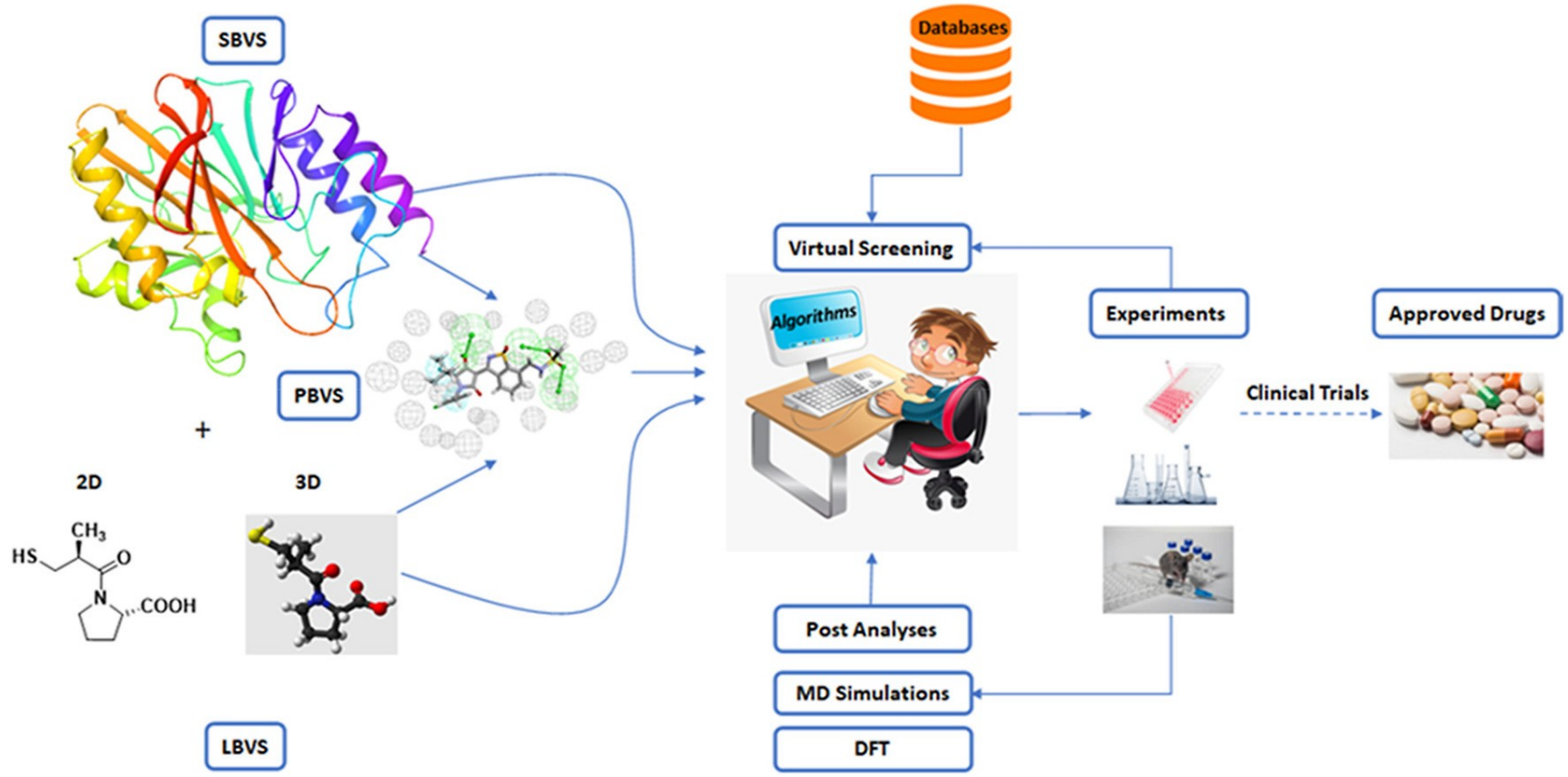
Molecular Graph Generation



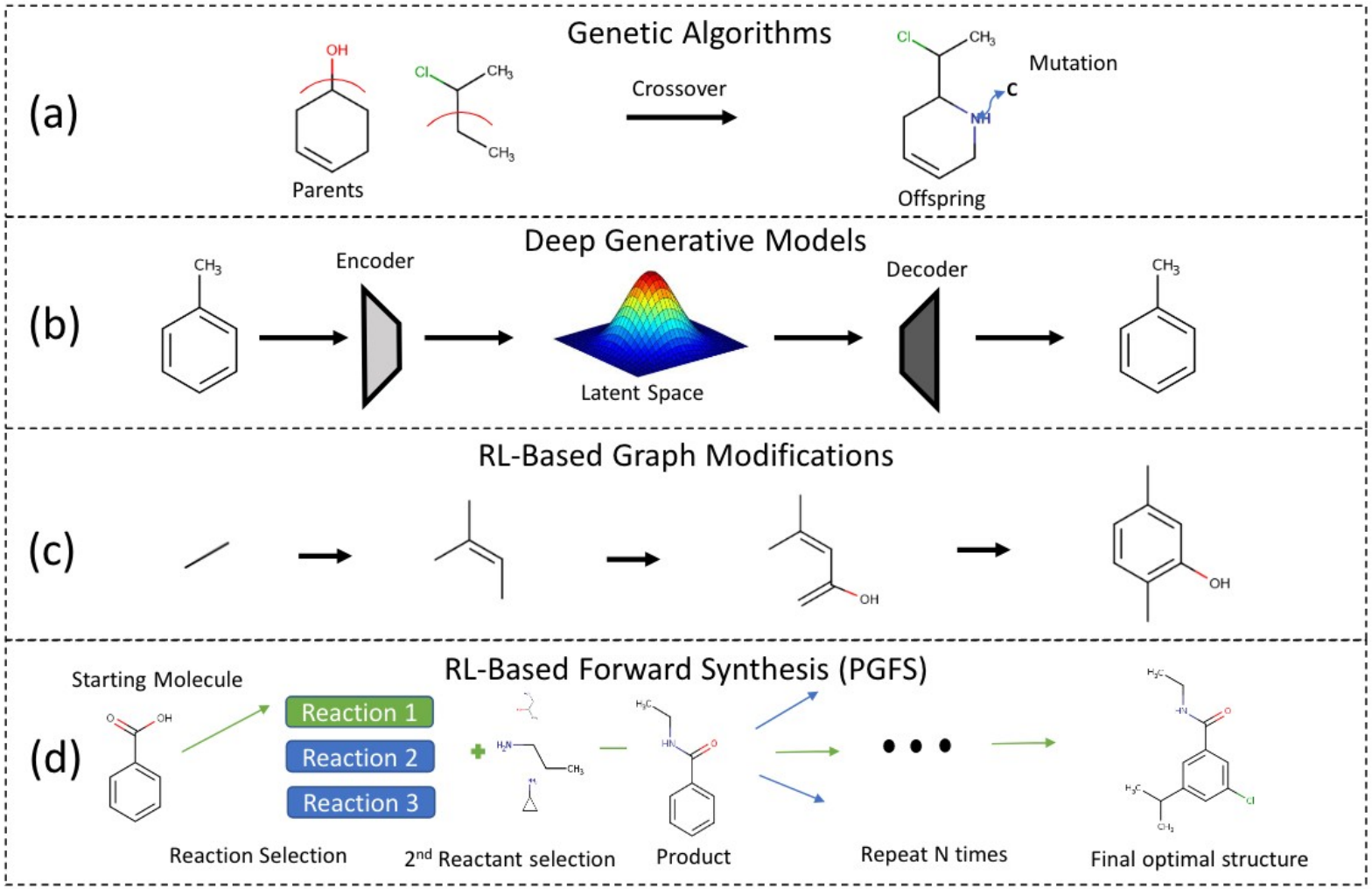
Problem:

Chemists take a significant amount of time **manually tweak** the structure of molecules,
and still not produce molecules with desired properties…



Solution:

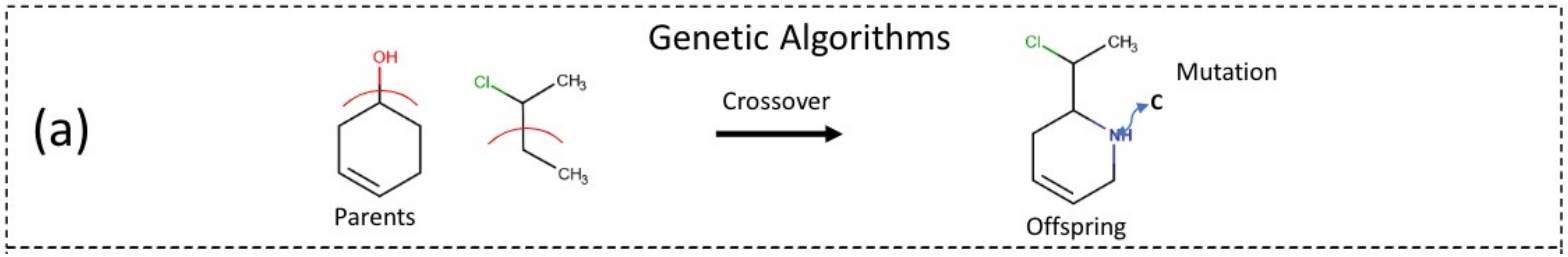
Use ML to Generate molecular graphs

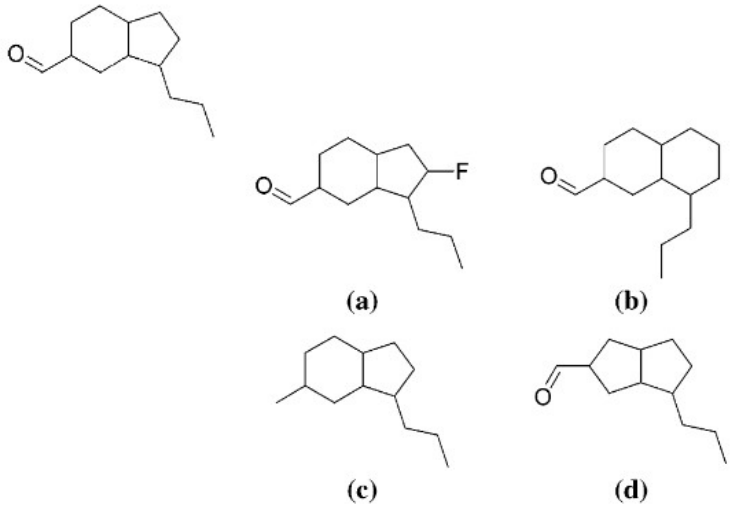


Various types of Generative Models

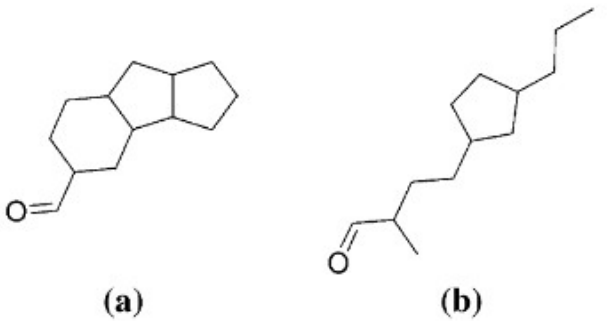
And also GFlowNets

Gottipati, S. K., Sattarov, B., Niu, S., Pathak, Y., Wei, H., Liu, S., Blackburn, S., Thomas, K., Coley, C., Tang, J., Chandar, S., & Bengio, Y. (2020). PGFS_Learning to Navigate The Synthetically Accessible Chemical Space Using Reinforcement Learning. *Proceedings of the 37th International Conference on Machine Learning*, 3668–3679. https://proceedings.mlr.press/v119/gottipati20a.html





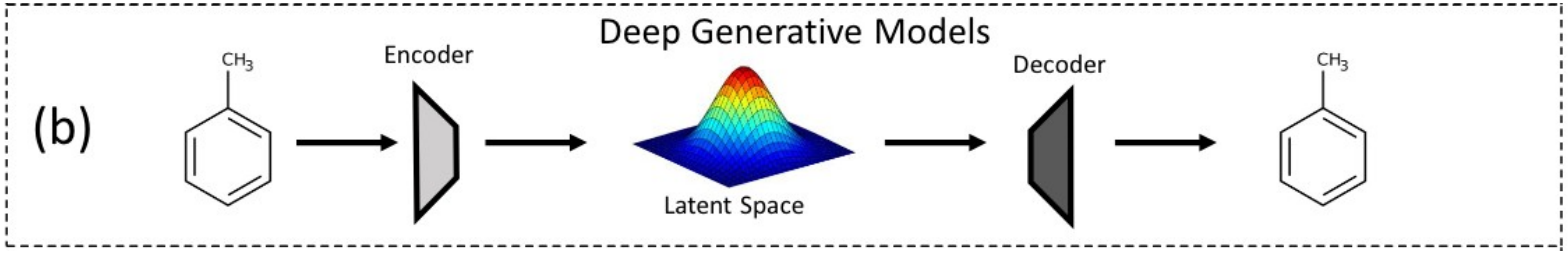
Node mutation operators (a) append, (b) insert, (c) prune, and (d) delete

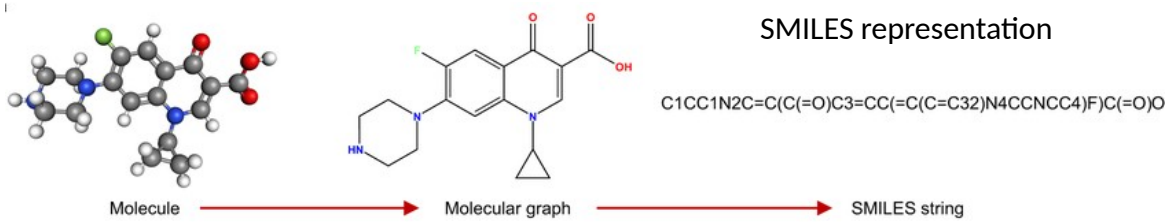


Edge mutation operators (a) add and (b) delete

Then Crossover …
Subgraph Crossover …

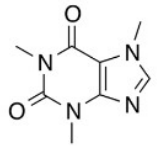
Brown, N., McKay, B., Gilardoni, F., & Gasteiger, J. (2004).
A Graph-Based Genetic Algorithm and Its Application to the Multiobjective Evolution of Median Molecules.
Journal of Chemical Information and Computer Sciences, *44*(3), 1079–1087. https://doi.org/10.1021/ci034290p





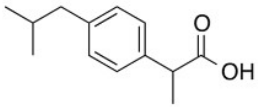
SMILES representation

C1CC1N2C=C(C(=O)C3=CC(=C(C=C32)N4CCNCC4)F)C(=O)O



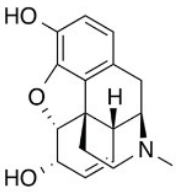
Caffeine

CN1c2ncn(C)c2C(=O)N(C)C1=O



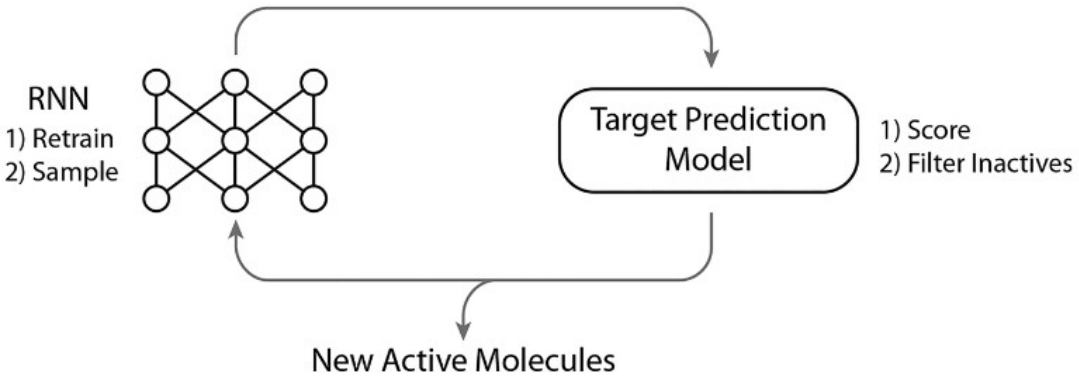
Ibuprofen

CC(C)Cc1ccc(cc1)C(C)C(O)=O



Morphine

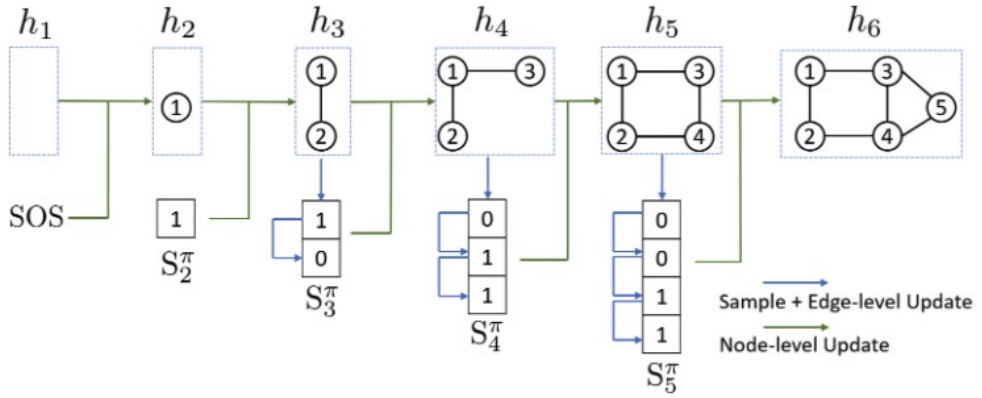
[H][C@]12C=C[C@H](O)[C@@H]3Oc4c5c(C[C@H]1N(C)CC[C@@]235)ccc4O

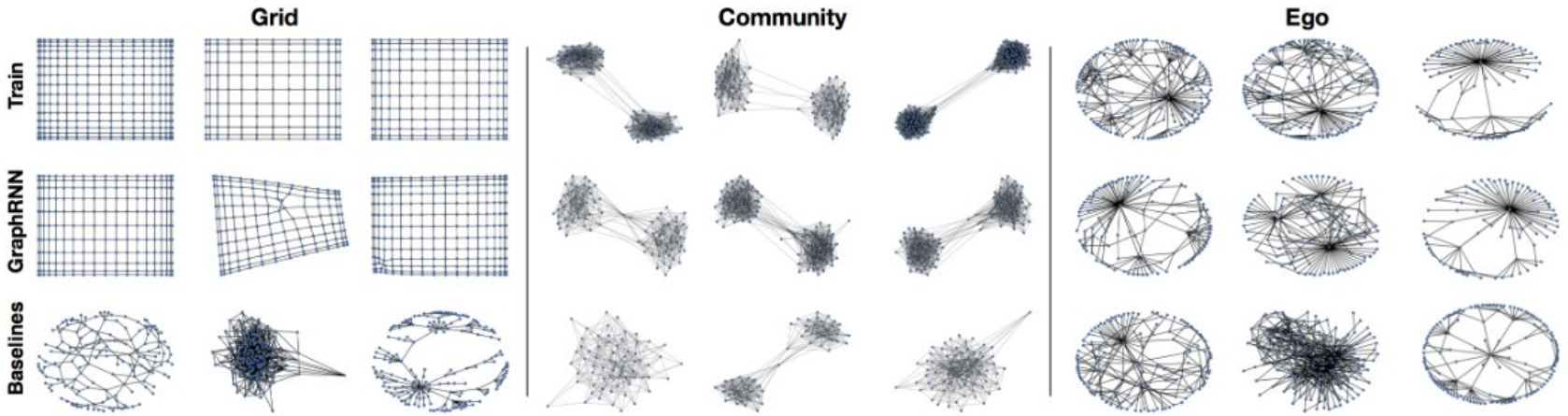


Segler, M. H. S., Kogej, T., Tyrchan, C., & Waller, M. P. (2018).
SMILES_Generating Focused Molecule Libraries for Drug Discovery with Recurrent Neural Networks.
ACS Central Science, *4*(1), 120–131. https://doi.org/10.1021/acscentsci.7b00512

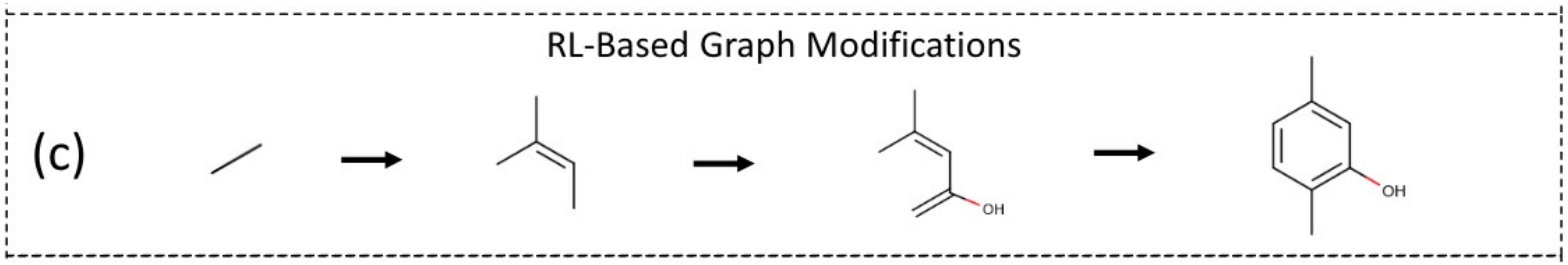
Popova, M., Shvets, M., Oliva, J., & Isayev, O. (2019).
MolecularRNN: Generating realistic molecular graphs with optimized properties (arXiv:1905.13372).
arXiv. http://arxiv.org/abs/1905.13372

GraphRNN: Deep Generative Models for Graphs

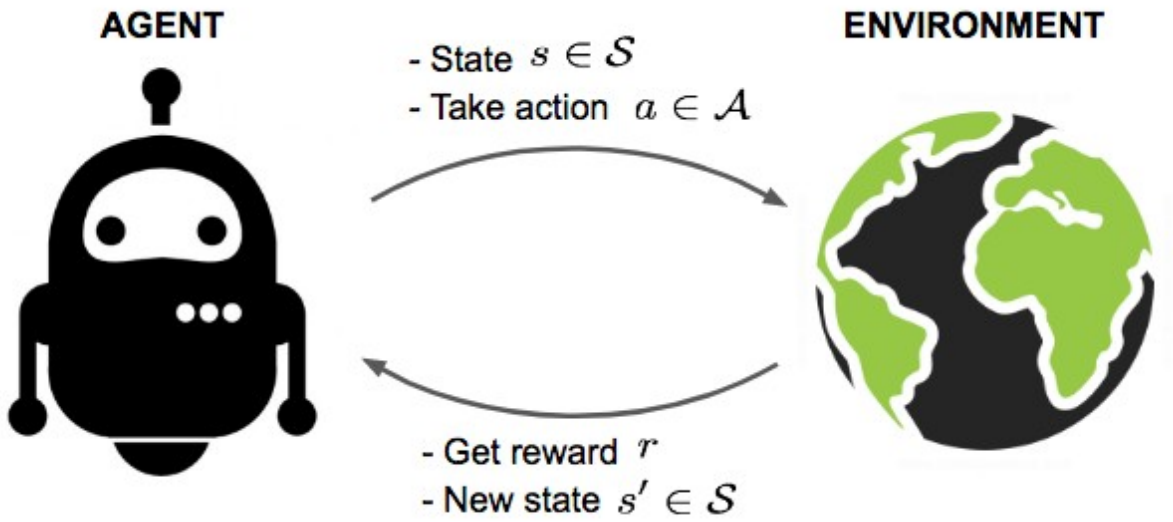




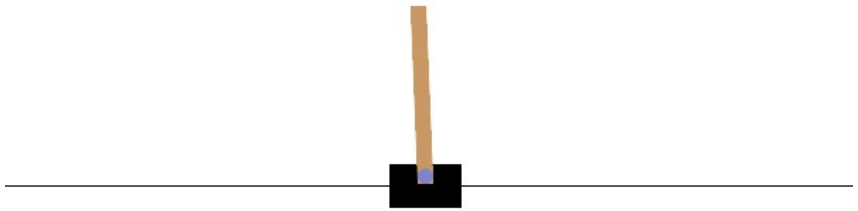
You, J., Ying, R., Ren, X., Hamilton, W., & Leskovec, J. (2018). GraphRNN: Generating Realistic Graphs with Deep Auto-regressive Models. *Proceedings of the 35th International Conference on Machine Learning*, 5708–5717. https://proceedings.mlr.press/v80/you18a.html



Brief intro of RL



e.g. Cart
Pole



Action:

- 0: Push cart to the left
- 1: Push cart to the right

State:

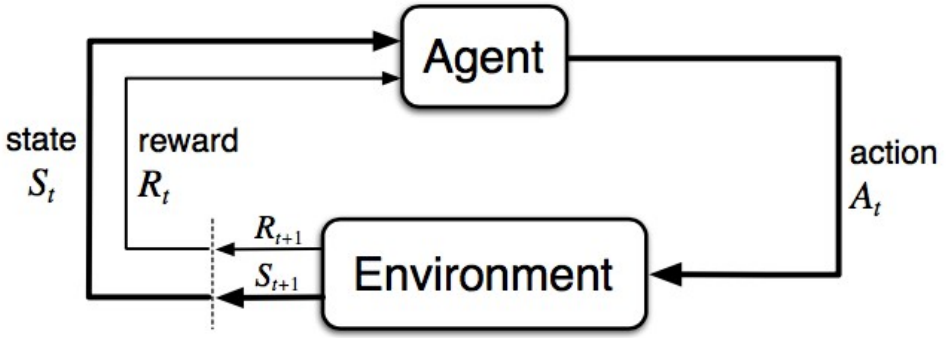
Cart Position
Cart Velocity
Pole Angle
Pole Angular Velocity

Rewards:
Since the goal is to keep the pole upright for as long as possible,
a reward of +1 for every step taken.

So What is a RL Model?

Markov Decision Processes

(mathematical formalism)



More formally, almost all the RL problems can be framed as **Markov Decision Processes** (MDP)

$$\mathbb{P}[S_{t+1}|S_t] = \mathbb{P}[S_{t+1}|S_1, \ldots, S_t]$$

one **transition** step: (s, a, s', r)

A Markov deicison process consists of five elements $M = \langle S, A, P, R, \gamma \rangle$,

- S - a set of states
- A - a set of actions
- P - transition probability function
- R - reward function
- γ - discounting factor for future rewards. (0 ~ 1, importance to the current state)

$$P^a_{ss'} = P(s'|s,a) = \mathbb{P}[S_{t+1} = s'|S_t = s, A_t = a] = \sum_{r \in \mathcal{R}} P(s', r|s, a)$$

$$R(s,a) = \mathbb{E}[R_{t+1}|S_t = s, A_t = a] = \sum_{r \in \mathcal{R}} r \sum_{s' \in \mathcal{S}} P(s', r|s, a)$$

(In an unknown environment, we do not have perfect knowledge

And What is the model Learning?

The policy function π is what we try to learn in reinforcement learning. (also, the value function V(s))

$\pi(s) = a$

We'd like to find a **optimal policy** $\pi(s)$ that provides the guideline on what is the optimal action to take in a certain state with **the goal to maximize the total rewards**.

How to update the policy during each episode?
How to evaluate the Policy, and improve the policy iteratively?
Can we Improve the policy greedily ? (go to the state w probable highest value)

trade-off between Exploitation and Exploration

No Exploration! → **Random Exploration** → **Smart Exploration**

Greedy algorithm — Є-Greedy algorithm — Upper confidence bounds (UCB), Thompson sampling

Common Approaches:

$$\pi_0 \xrightarrow{\text{evaluation}} V_{\pi_0} \xrightarrow{\text{improve}} \pi_1 \xrightarrow{\text{evaluation}} V_{\pi_1} \xrightarrow{\text{improve}} \pi_2 \xrightarrow{\text{evaluation}} \cdots \xrightarrow{\text{improve}} \pi_* \xrightarrow{\text{evaluation}} V_*$$

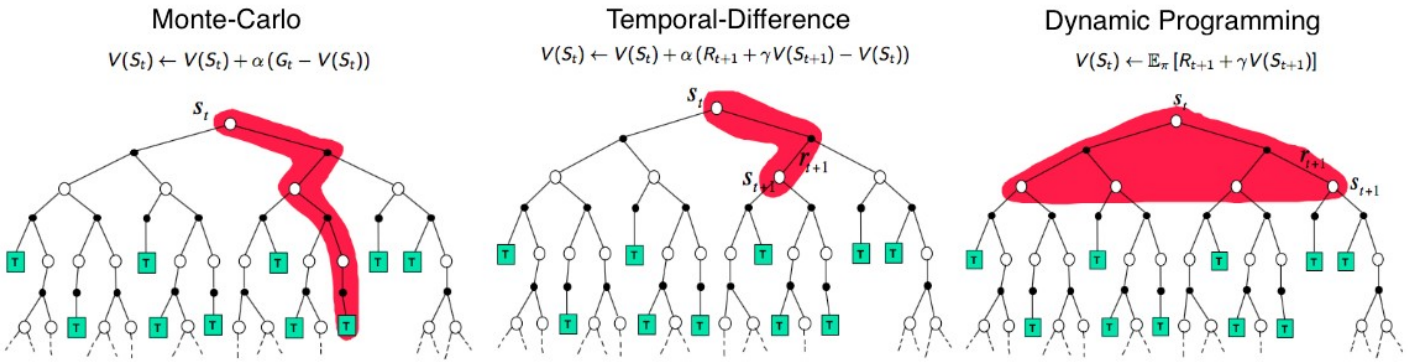
Dynamic Programming

Monte-Carlo Methods

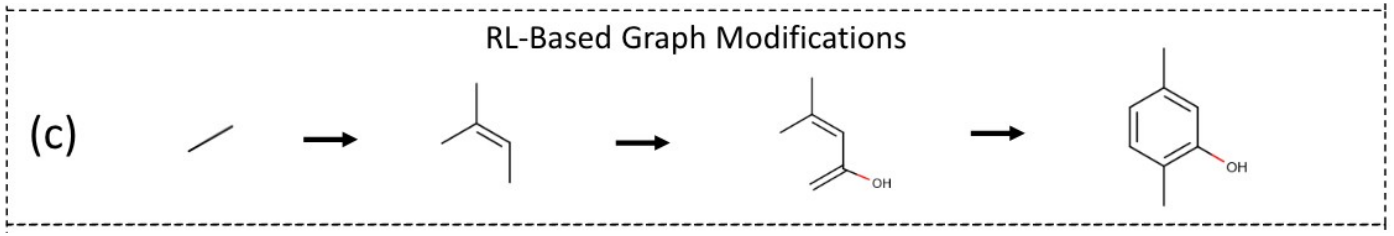
Temporal-Difference (TD Learning) (Q-Learning)

Policy Gradient (gradient ascent)

Evolution Strategies

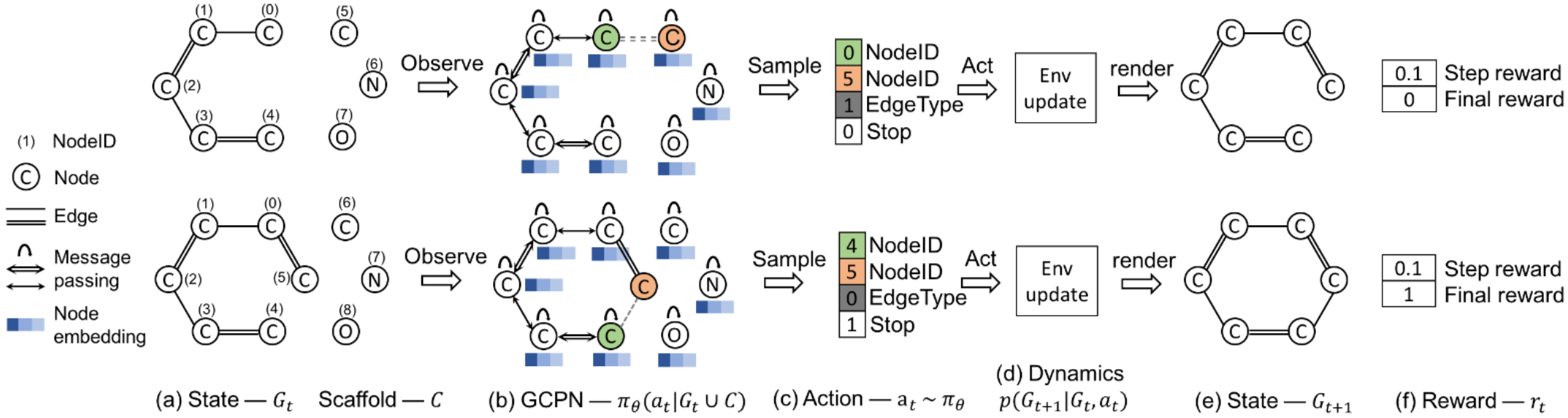


Molecular Graph Generation by RL



Compared to the previous methods,

RL allows active exploration of the molecule space beyond samples in a dataset

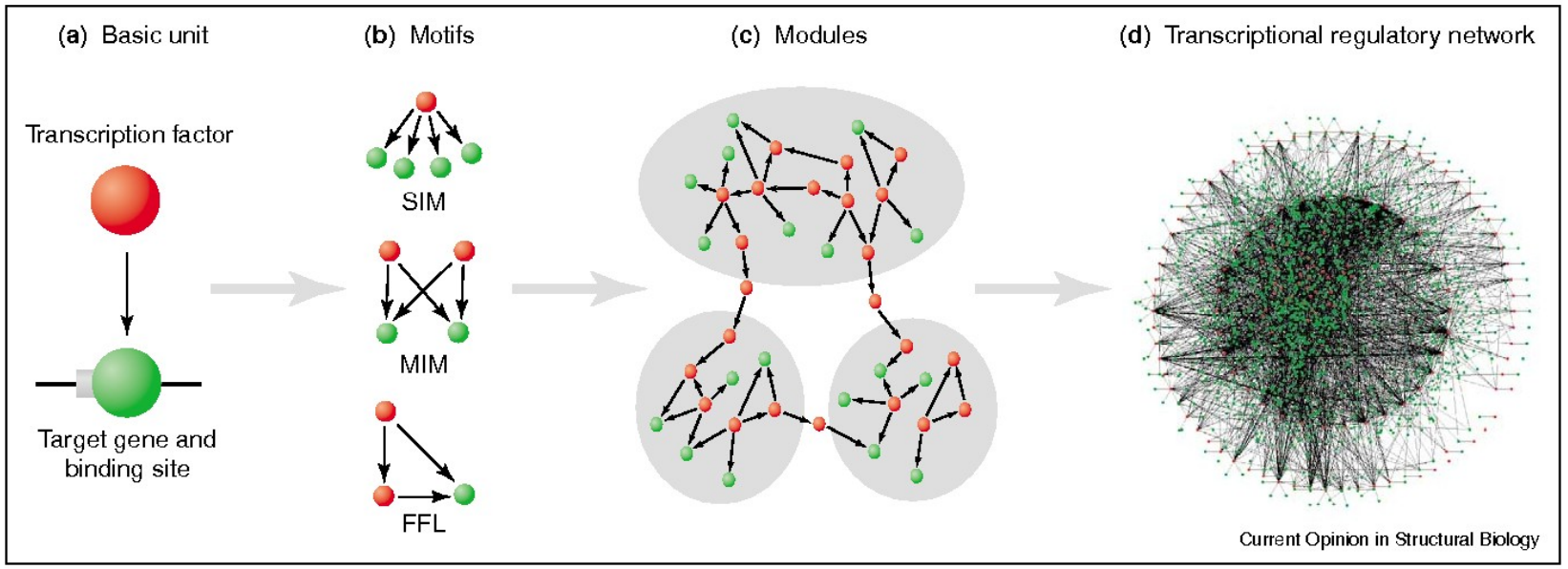


scaffold subgraphs {C1, ..., Cs} ---> link prediction ---> $G_t \cup C$

You, J., Liu, B., Ying, Z., Pande, V., & Leskovec, J. (2018).
GCPN_Graph Convolutional Policy Network for Goal-Directed Molecular Graph Generation.
Advances in Neural Information Processing Systems, *31*. https://proceedings.neurips.cc/paper_files/paper/2018/hash/d60678e8f2ba9c540798ebbde31177e8-Abstract.html

Zhou, Z., Kearnes, S., Li, L., Zare, R. N., & Riley, P. (2019).
MolDQN_Optimization of Molecules via Deep Reinforcement Learning. *Scientific Reports*, *9*(1), 10752. https://doi.org/10.1038/s41598-019-47148-x

GRN


(a) Basic unit
(b) Motifs
(c) Modules
(d) Transcriptional regulatory network
Transcription factor
Target gene and binding site
SIM
MIM
FFL
Current Opinion in Structural Biology