KTH ROYAL INSTITUTE OF TECHNOLOGY

Reinforcement learning

Applications to chemistry and biology

What is needed for applying RL?

Data

Cheap, preferrably online

Simulations

Representable states and actions

Problem

Represents a process

Is a decision-making problem (ideally)

Possible to formulate a reward

Data and problems in biology

Data	Problem
2D/3D images	Image segmentation/classification
Time series	Controlling dynamical systems
Medication – results	Drug discovery
Disease – symptoms	Disease diagnosis
Gene modification – effects	Gene expression

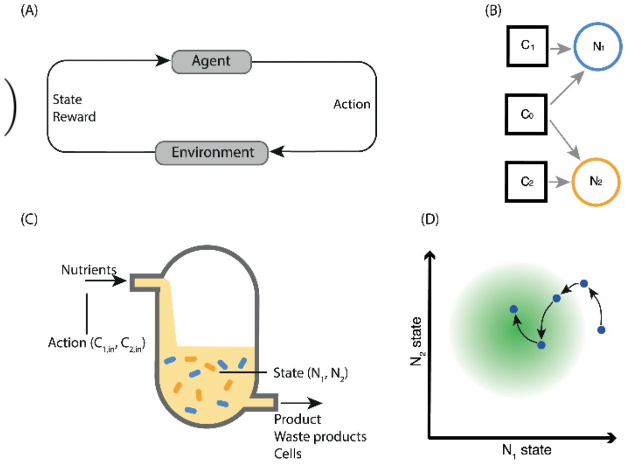
RL for (biological) control

Deep reinforcement learning for the control of microbial co-cultures in bioreactors

Treloar NJ, Fedorec AJH, Ingalls B, Barnes CP (2020) PLOS Comput Biol 16(4)

Problem: how to feed two competitive microbial cultures to keep their populations within a desired range?

$$r = \frac{1}{10}\left(1 - \frac{1}{2}\left(\frac{|N_1 - N_1^{target}|}{20\times 10^9} + \frac{|N_2 - N_2^{target}|}{30\times 10^9}\right)\right)$$



$$\frac{d}{dt}C_0(t) = q(C_{0,in} - C_0(t)) - \sum_{i=1}^{m} \frac{1}{\gamma_{0,i}} \boldsymbol{\mu}_i(t)\mathbf{N}_i(t)$$

$$\boldsymbol{\mu}_i = \boldsymbol{\mu}_{max,i} \frac{\mathbf{C}_i}{\mathbf{K}_{s,i} + \mathbf{C}_i} \frac{C_0}{\mathbf{K}_{s0,i} + C_0}$$

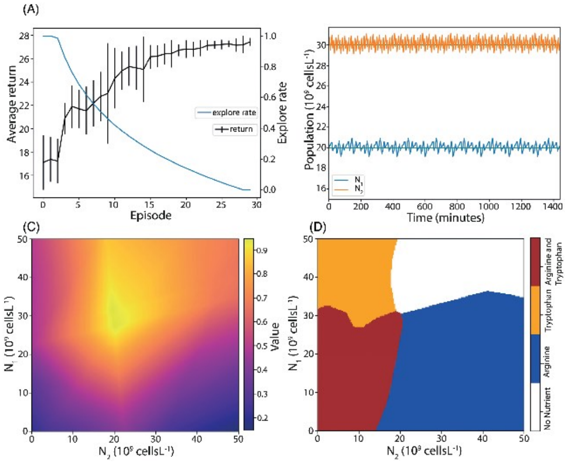
$$\frac{d}{dt}\mathbf{N}_i(t) = (\boldsymbol{\mu}_i(t) - q)\mathbf{N}_i(t)$$

RL for (biological) control

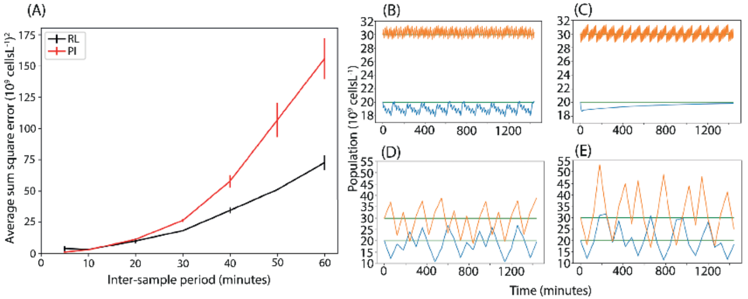
Deep reinforcement learning for the control of microbial co-cultures in bioreactors
Treloar NJ, Fedorec AJH, Ingalls B, Barnes CP (2020) PLOS Comput Biol 16(4)

Problem: how to feed two competitive microbial cultures to keep their populations within a desired range?

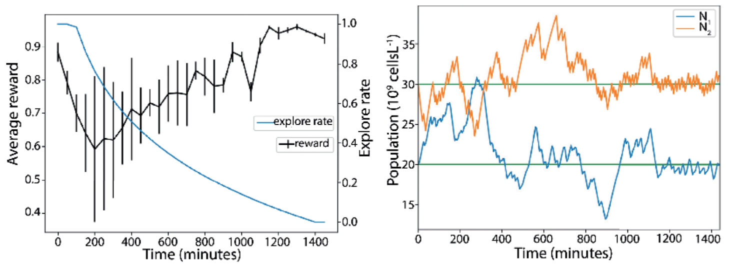
Trained on the simulated data



RL vs PID (on the simulated data)



24 h online data collection



 Treloar et al. PLOS Comput. Biol. 16(4), 2020. Images are reproduced according to CC BY 4.0

RL for treatment strategy

Improving sepsis treatment strategies by combining deep and kernel-based reinforcement learning
Peng X, Ding Y, Wihl D, Gottesman O, Komorowski M, Lehman LH, Ross A, Faisal A, Doshi-Velez F (2018) AMIA Annual Symposium

Problem: how to improve sepsis treatment to decrease mortality?

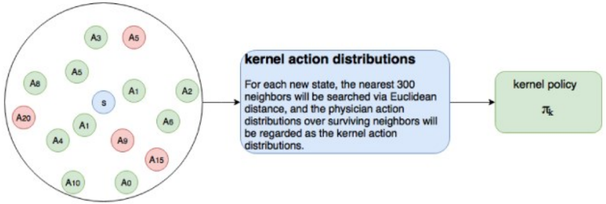
Inputs: age, Weight_kg, GCS, HR, SysBP, MeanBP, DiaBP, RR, Temp_C, FiO2_1, Potassium, Sodium, Chloride, Glucose, Magnesium, Calcium, Hb, WBC_count, Platelets_count, PTT, PT, Arterial_pH, paO2, paCO2, Arterial_BE, HCO3, Arterial_lactate, SOFA, SIRS, Shock_Index, PaO2_FiO2, cumulated_balance_tev, Elixhauser, Albumin, CO2_mEqL, Ionised_Ca, max_dose_vaso, SpO2, BUN, Creatinine, SGOT, SGPT, Total_bili, INR, input_total_tev, input_4hourly-tev, output_total, output_4hourly

Output: administration levels of 2 drugs (0-4 each)

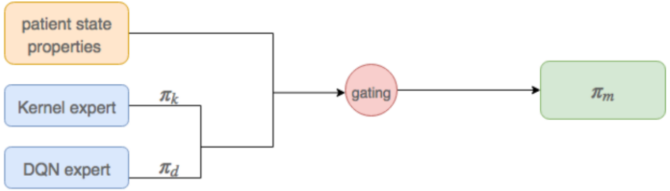
Reward: $$r(o,a,o') = -\log \frac{f(o')}{1-f(o')} f(o') + \log \frac{f(o)}{1-f(o)}$$

$f(o)$ – probability of mortality (pre-trained NN)

Policy: Kernel vs Deep



DQN policy has no guarantees!

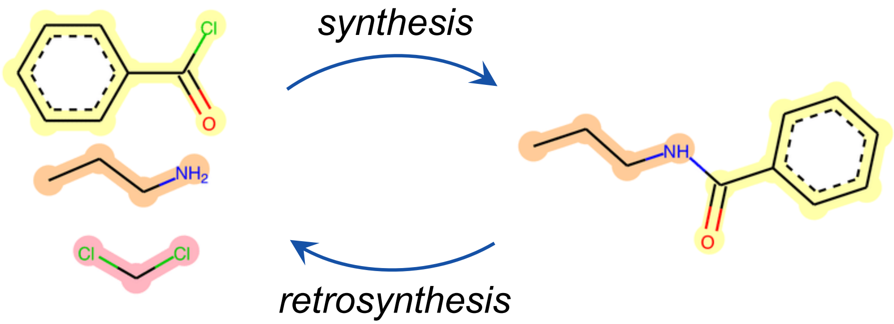


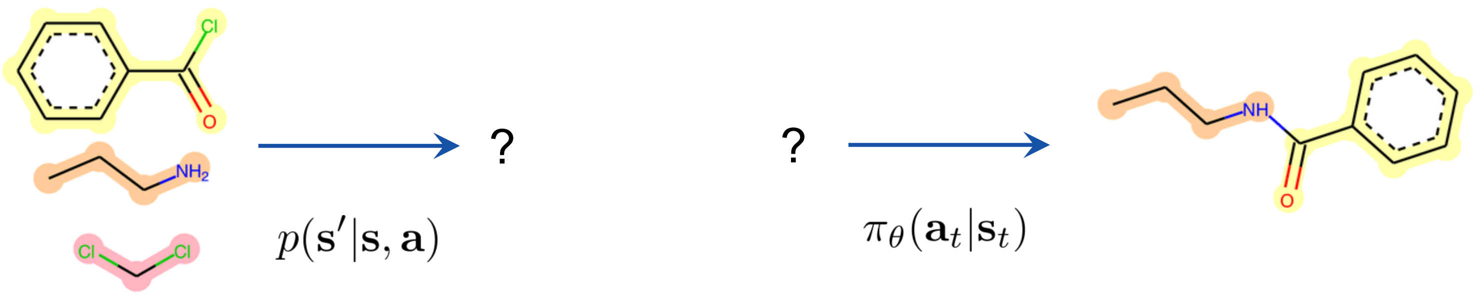
Peng et al. AMIA Annual Symp., 2018. Images are reproduced with permission.

Data and problems in chemistry

Data	Problem
Chemical reactions	Predict reaction outcome
Reaction kinetics	Control reaction process
Molecule – properties	Predict molecular property
Interaction graphs	Predict interactions
Molecule datasets	(Conditionally) generate molecules

Retrosynthesis



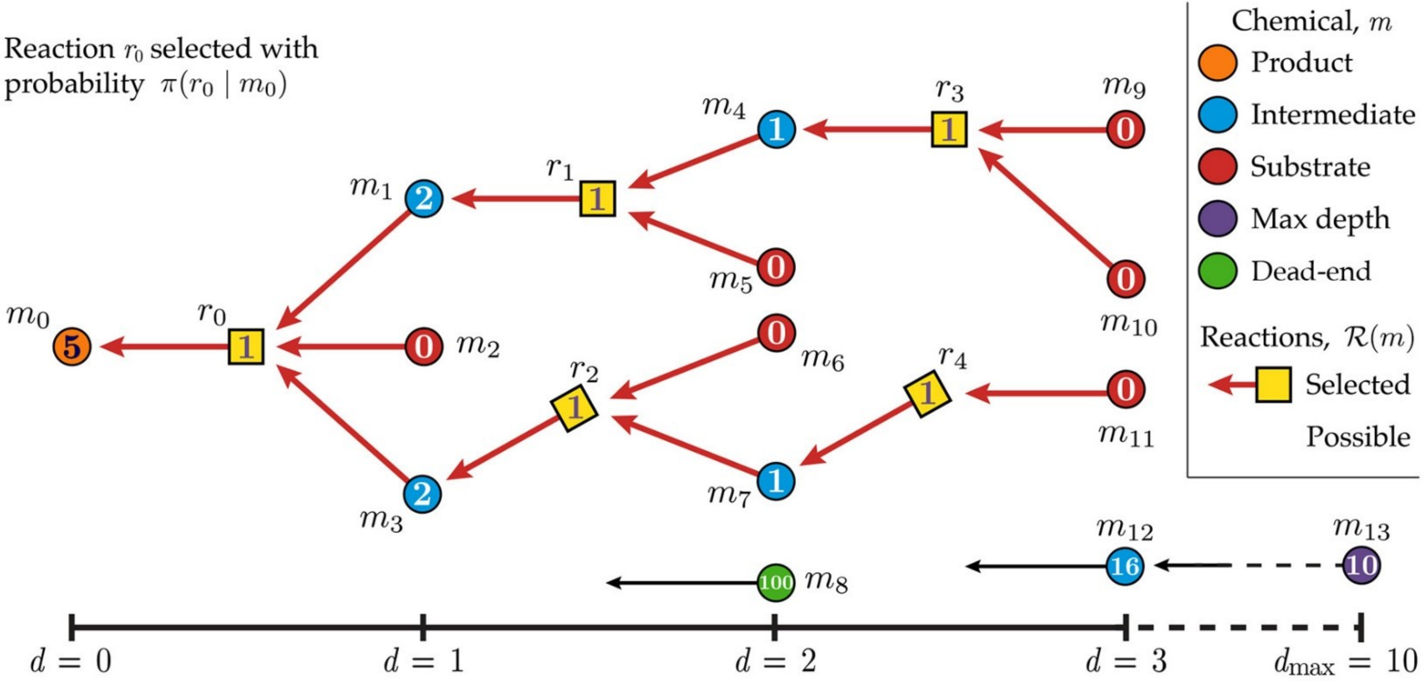


$p(\mathbf{s}'|\mathbf{s}, \mathbf{a})$

$\pi_\theta(\mathbf{a}_t|\mathbf{s}_t)$

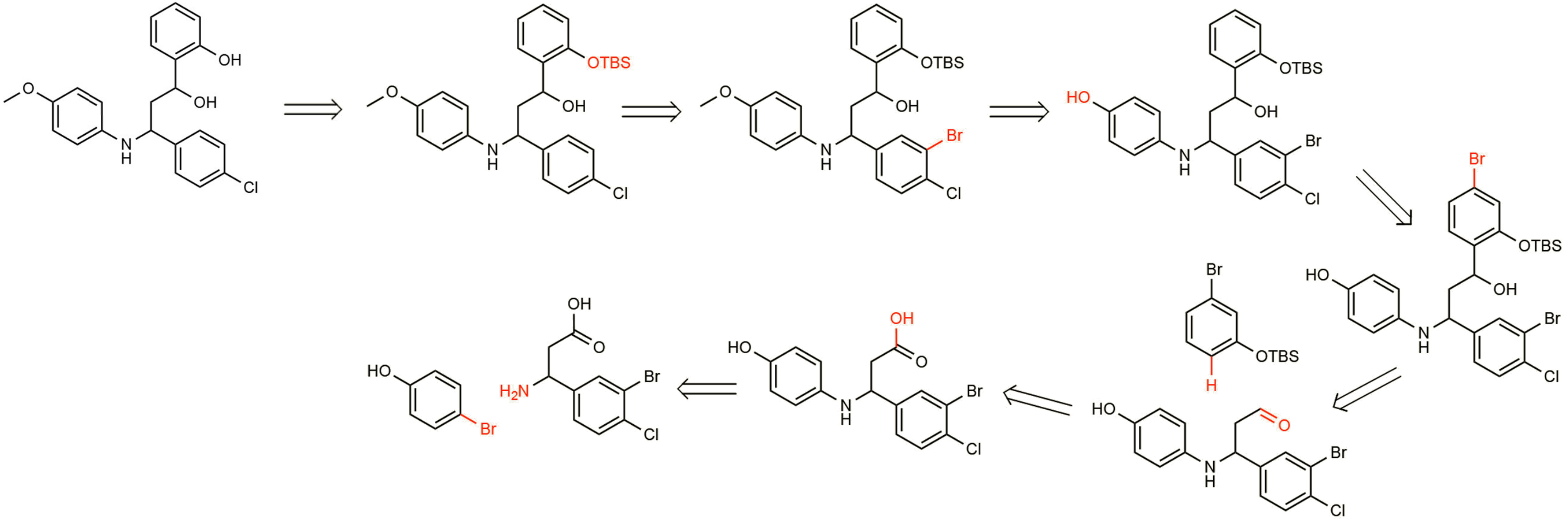
Retrosynthesis

Learning Retrosynthetic Planning through Simulated Experience
Schreck JS, Coley CW, Bishop KJM (2019) ACS Cent. Sci. 5(6)



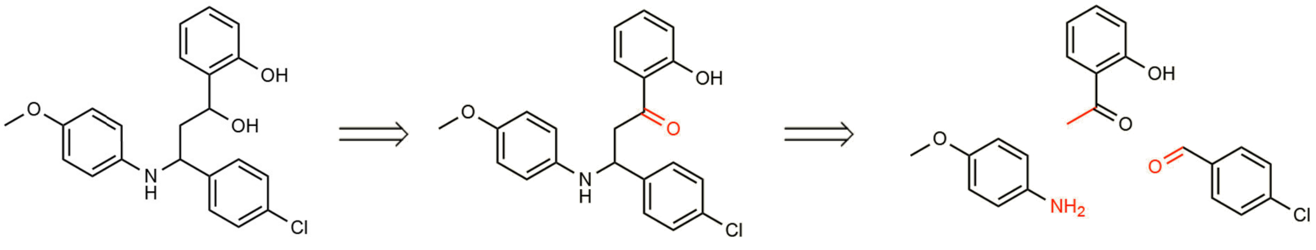
Schreck et al. ACS Cent. Sci., 2019, 5(6). Image is reproduced according to ACS AuthorChoice.

Retrosynthesis – initial policy

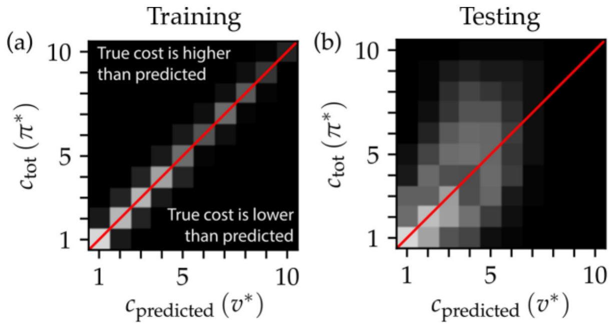


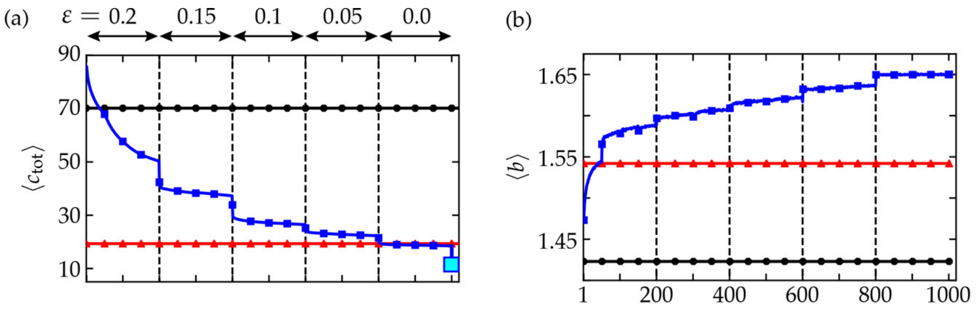
 Schreck et al. ACS Cent. Sci., 2019, 5(6). Image is reproduced according to ACS AuthorChoice.

Retrosynthesis – optimal policy



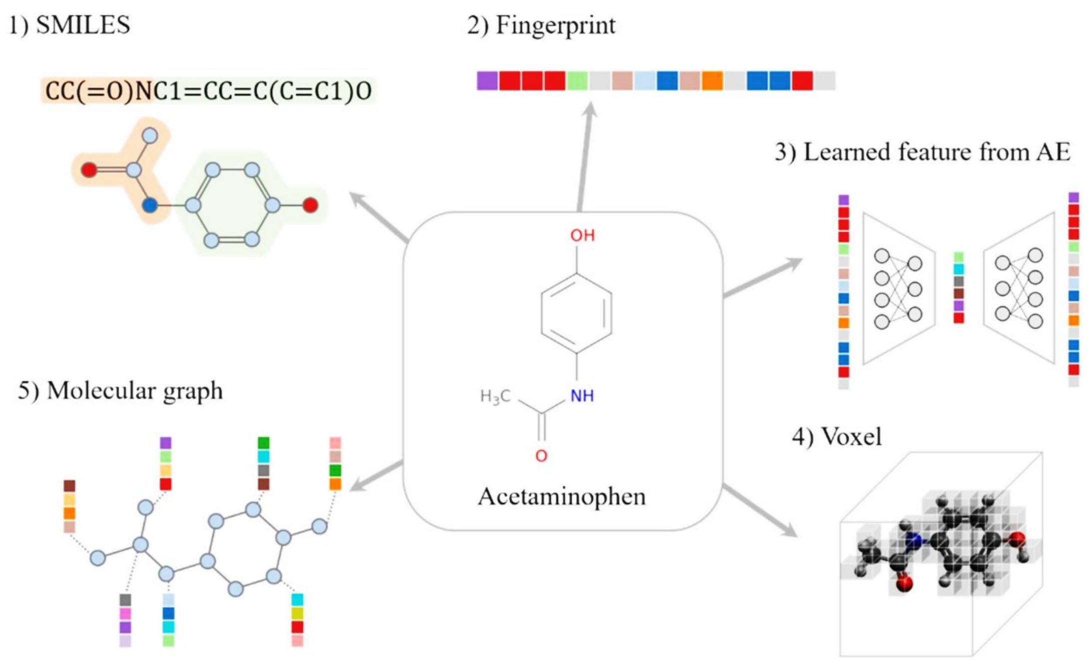
Benefit of using a meaningful reward function:





Schreck et al. ACS Cent. Sci., 2019, 5(6). Images are reproduced according to ACS AuthorChoice.

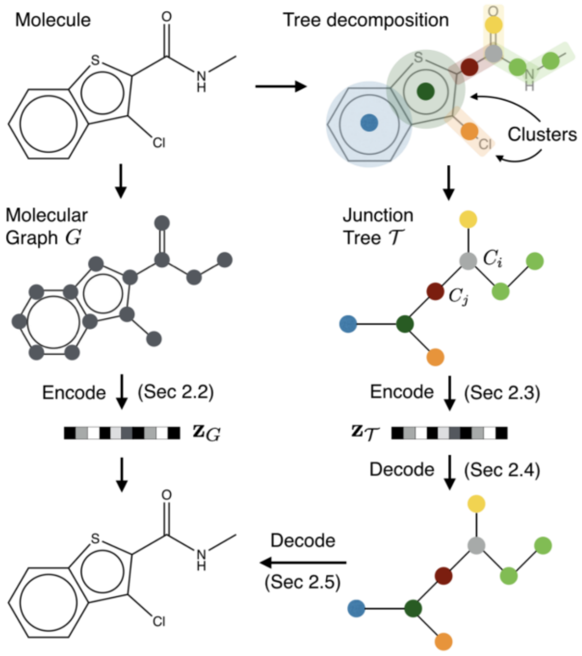
Molecular representations



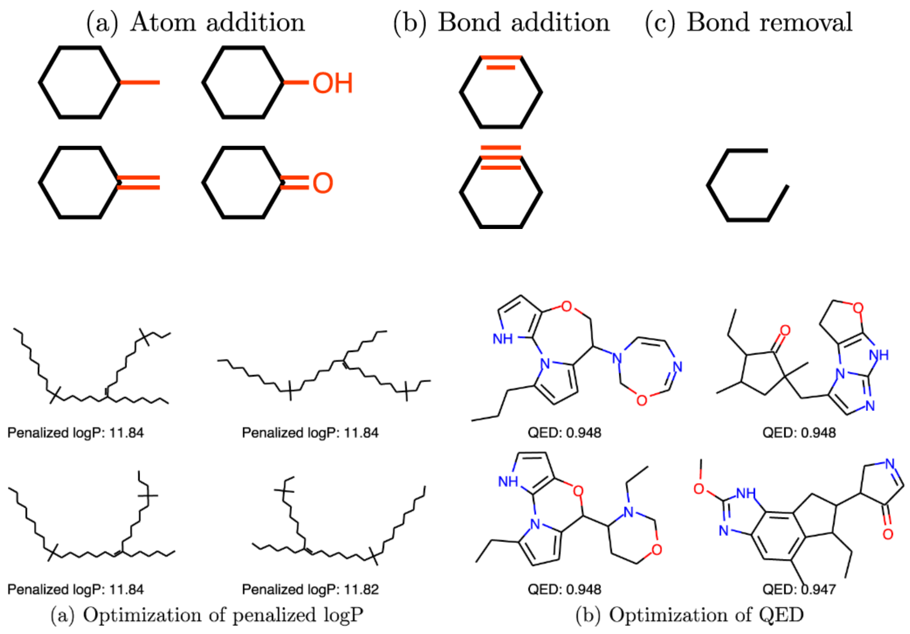
Kim et al. Int. J. Mol. Sci., 2021, 22(18). Image is reproduced according to CC BY 4.0

Molecule generation

"Direct" latent generative model



Stepwise modification model – formulated as RL problem



Jin et al. PMLR 80, 2018.; Zhou et al. Sci. Rep., 2019, 9:10752.

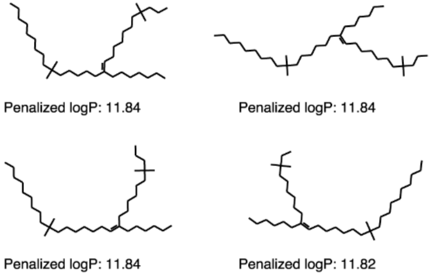
Molecule generation

Optimization of molecules via deep reinforcement learning

Zhou Z, Kearnes S, Li L, Zare RN, Riley P (2019) Sci. Rep. 9:10752

Objective: modify given molecule maximizing one or several of (QED, logP) while retaining similarity

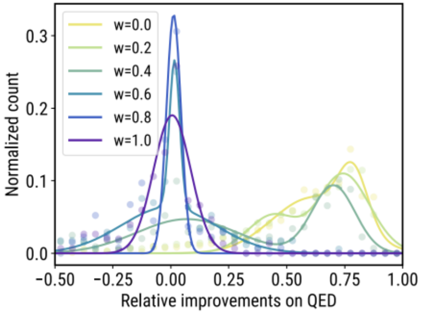
Constrained optimization



$$\mathcal{R}(s) = \begin{cases} \text{logP}(m) - \lambda \times (\delta - \text{SIM}(m, m_0)) & \text{if SIM}(m, m_0) < \delta \\ \text{logP}(m) & \text{otherwise} \end{cases}$$

Multi-objective optimization

$$\mathcal{R}(s) = w \times \text{SIM}(s) + (1 - w) \times \text{QED}(s)$$



Example