A graph-convolutional NN model for the prediction of chemical reactivity

Egor Gladin

MIPT

22.04.19

Plan

Статья

Connor W. Coley, Wengong Jin, Luke Rogers, Timothy F. Jamison, Tommi S. Jaakkola, William H. Green, Regina Barzilay, Klavs F. Jensen

A graph-convolutional NN model for the prediction of chemical reactivity
Chemical Science, 2018

Постановка задачи

Задача

Предсказать продукты реакции по веществам и растворам, участвующим в реакции

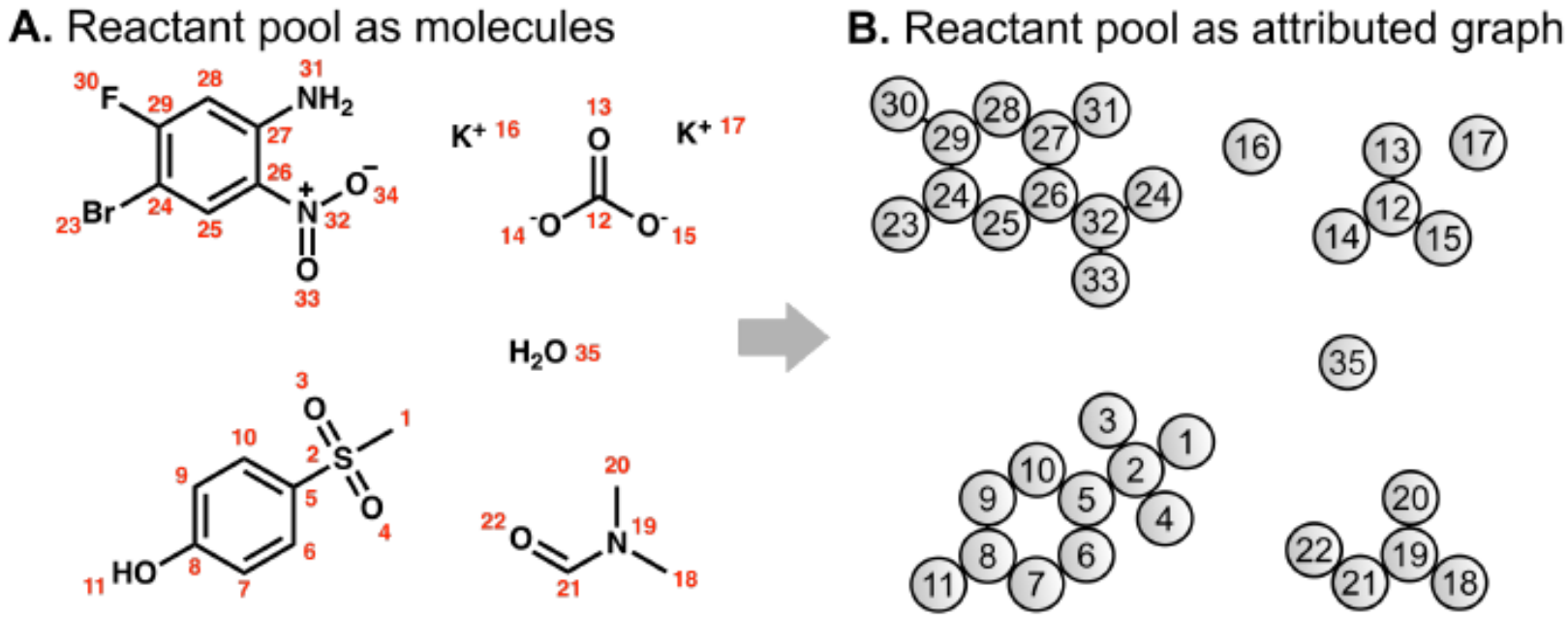


Схема подхода

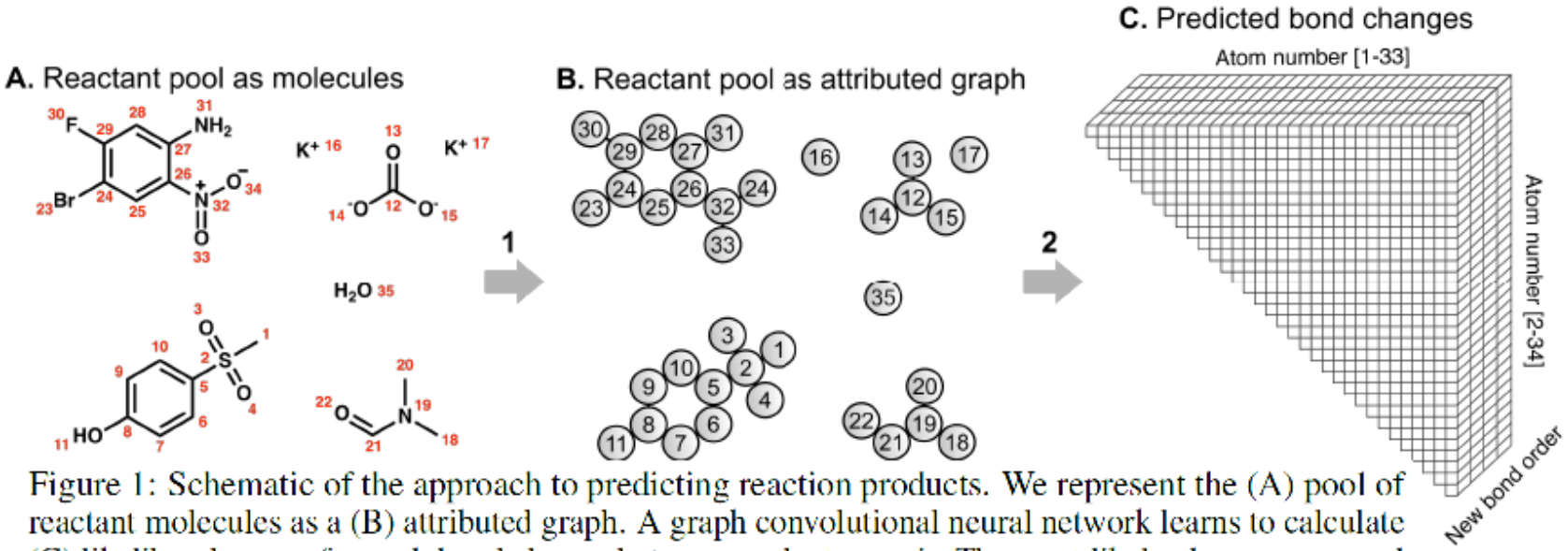


Figure 1: Schematic of the approach to predicting reaction products. We represent the (A) pool of reactant molecules as a (B) attributed graph. A graph convolutional neural network learns to calculate (C) likelihood scores for each bond change between each atom pair. The most likely changes are used to perform a focused, ranked enumeration of (D) candidate products, which are filtered by chemical valence rules. These candidates are then rescored by another graph convolutional network to yield (E) a probability distribution over predicted product species.

Схема подхода

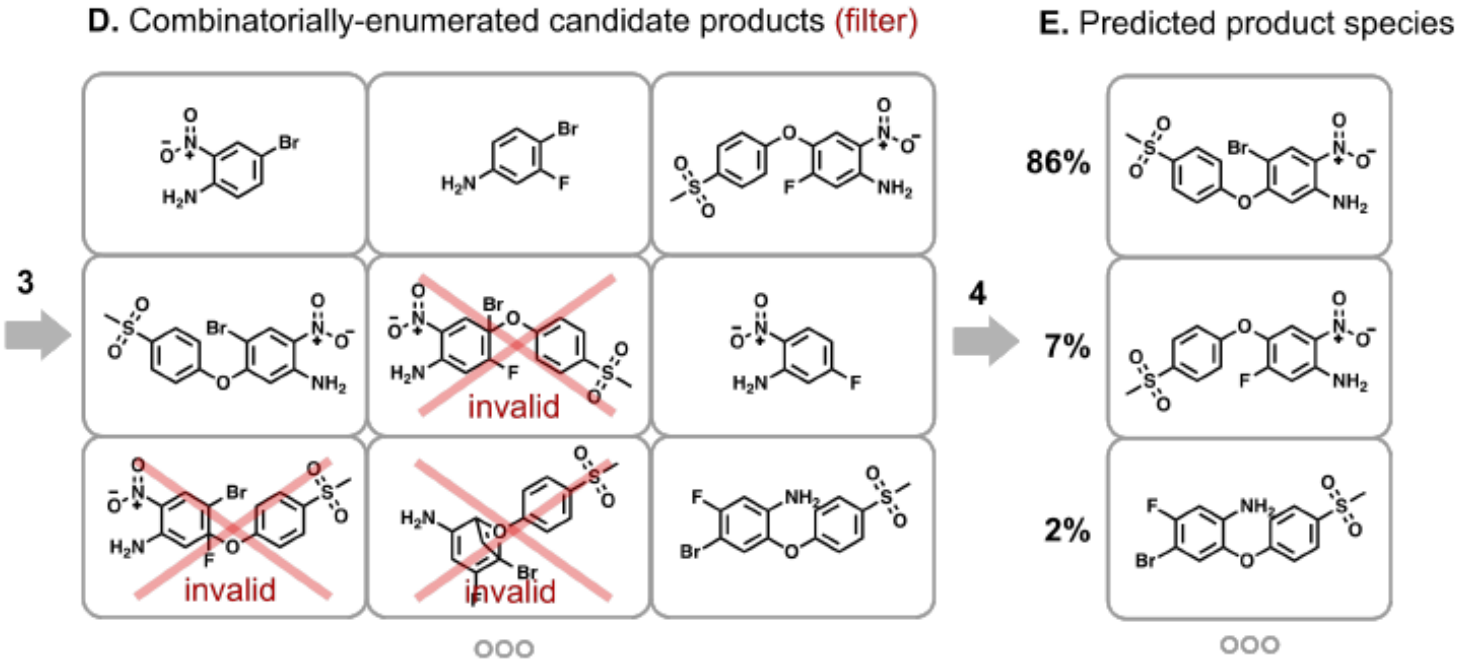
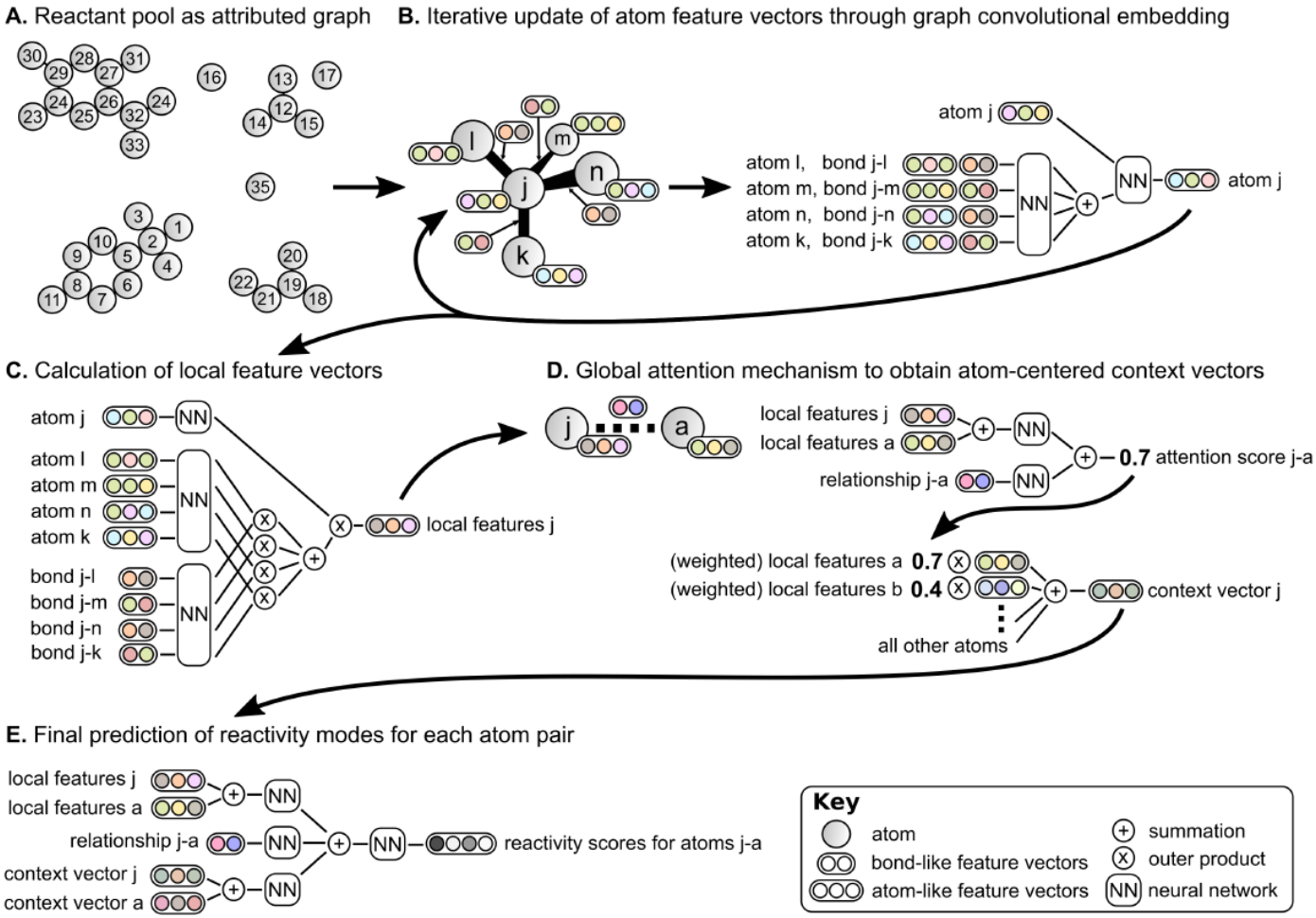


Figure 1: Schematic of the approach to predicting reaction products. We represent the (A) pool of reactant molecules as a (B) attributed graph. A graph convolutional neural network learns to calculate (C) likelihood scores for each bond change between each atom pair. The most likely changes are used to perform a focused, ranked enumeration of (D) candidate products, which are filtered by chemical valence rules. These candidates are then rescored by another graph convolutional network to yield (E) a probability distribution over predicted product species.

WLN for predicting likely changes in bond order



Local atom representation

Обозначения

- f_v^l - feature vector атома v на итерации l
- f_{uv} - стартовый feature vector связи (u, v)
- U_i, V_i, W_i - матрицы WLN (обучаемые веса)
- c_v - конечное локальное представление атома v

$$f_v^l = \tau\left(U_1 f_v^{l-1} + U_2 \sum_{u \in N(v)} \tau(V_1 f_u^{l-1} + V_2 f_{uv})\right) \quad (1 \le l \le L)$$

$$c_v = \sum_{u \in N(v)} W_1 f_u^L \odot W_2 f_{uv} \odot W_3 f_v^L$$

Global atom representation

Обозначения

- α_{vz} - attention score of atom v upon atom z
- b_{uv} - feature vector, отвечающий за отношение (u, v)
- P_a, P_b, u - обучаемые веса модели

$$\alpha_{vz} = \sigma(u^T \tau(P_a c_v + P_a c_z + P_b b_{vz}))$$

$$\tilde{c}_v = \sum_v \alpha_{vz} c_z$$

Reaction center

Центр реакции

(u, v, b) находится в центре реакции $\iff$ тип связи между атомами u и v сменился на b

Обозначения

- $s_{u,v,b}$ - вероятность того, что (u, v, b) в центре реакции
- M_a, M_b, P_a, u_b - обучаемые веса модели
- $y_{u,v,b} = 1 \iff (u, v, b)$ в центре реакции

$$s_{u,v,b} = \sigma\left(u_b^T \tau(M_a \tilde{c}_u + M_a \tilde{c}_v + P_a c_u + P_a c_v + M_b f_{uv})\right)$$

$$- \sum_{u,v,b;u \neq v} y_{u,v,b} \log s_{u,v,b} + (1 - y_{u,v,b}) \log(1 - s_{u,v,b})$$

Reaction center prediction

Обозначения

- r - реагенты
- p_i - i-й потенциальный продукт реакции
- $c_v^{(p_i)}$ - feature vector атома v в продукте p_i
- $d_v^{(p_i)} \triangleq c_v^{(p_i)} - c_v^{(r)}$, $h_v^{(p_i,0)} = d_v^{(p_i)}$
- $RC(p_i)$ - множество связей, изменившихся в p_i

$$h_v^{(p_i,l)} = \tau\left(U_1 h_v^{(p_i,l-1)} + U_2 \sum_{u \in N(v)} \tau\left(V_1 h_u^{(p_i,l-1)} + V_2 f_{uv}\right)\right) \qquad (1 \le l \le L)$$

$$g_v^{(p_i)} = \sum_{u \in N(v)} W_1 h_u^{(p_i,L)} \odot W_2 f_{uv} \odot W_3 h_v^{(p_i,L)}$$

$$s(p_i) = u^T \tau\left(M \sum_{v \in p_i} g_v^{(p_i)}\right) + \sum_{(u,v,b) \in RC(p_i)} s_{u,v,b}$$

Results

Точность на уровне экспертов
Скорость: 100 ms per example on a single consumer GPU

Method	$\|\theta\|$	Top-1 [%]	Top-2 [%]	Top-3 [%]	Top-5 [%]
WLN/WLDN [16]	3.2 M	79.6	-	87.7	89.2
Sequence-to-sequence [31]	30 M*	80.3	84.7	86.2	87.5
This work	2.6 M	**85.6**	**90.5**	**92.8**	**93.4**