Acceleration of Designing Chemical Reactions by Contextual Multi-armed Bandit Algorithm

Tamiki Komatsuzaki
Professor
Affiliation: ICReDD and Research Center of Mathematics for Social Creativity, Research Institute for Electronic Science (RIES), Hokkaido University
e-mail: tamiki@es.hokudai.ac.jp
Specialty: Mathematical and Information Science, Chemical and Biological Physics

ABSTRACT

Most typical machine learning approaches for chemical reactions is perhaps first to map the desired quantities such as reaction yield, enantioselectivity etc measured for many systems with different catalysts, substituents, and experimental conditions for a reaction to be explored to a certain feature space, and then to infer which feature(s) are most relevant to the desired quantities, which may help to design a desired reaction system with high yield and/or high enantioselectivity. However, this approach requires a huge set of reaction data *a priori*, which includes lots of undesired reactions of low performance to reveal the important features.

Reinforcement learning, utilized in the Monte Carlo tree search in alpha-GO, does not necessarily require any set of data *a priori*, and is rather aimed at providing an experimentation protocol on how to choose next experiment (that may have different catalyst, different substituents and forth) in order to discover the desired reaction system(s) with as fewer experiments as possible, with guaranteeing the accuracy of the prediction.

In this talk, we present our recent study on designing chemical reactions in which contextual multi-arm Bandits algorithm[1,2] in reinforcement learning is introduced to choose a catalyst to yield high enantioselectivity for an asymmetric hydroalkoxylation reaction. Here, as a proof-of-concept of our algorithm, we used 38 reaction data associated with enantioselectivity and more than 500 dimensional molecular features extracted by ISIDA descriptor, with the knowledge of the most desired reaction[3]. We show how one can accelerate to discover the most desired reaction before being performed all experiments.

This is a joint collaboration with List group (Dr. Tsuji) and Varnek group (Dr. Sidorov) in ICReDD.

REF.

1. K. Tabata, A. Nakamura, J. Honda, T. Komatsuzaki, A bad arm existence checking problem: How to utilize asymmetric problem structure? **Machine Learning,** 109, 327-372 (2020)
2. K. Tabata, A. Nakamura, T. Komatsuzaki, Classification Bandits: Classification Using Expected Rewards as Imperfect Discriminators, **Pacific-Asia Conference on Knowledge Discovery and Data Mining,** 57-69 (2021)
3. N. Tsuji, P. Sidorov, Y. Nagata, A. Varnek, B. List, *manuscript in preparation.*