[P6] Efficient Molecular Structure Enumeration Algorithm for Inverse QSPR/QSAR

Tomoyuki Miyao, Hiromasa Kaneko, Funatsu Kimito

Funastu laboratory, Department of Chemical System Engineering, Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, 113-8656, Tokyo, Japan.

One of the goals of molecular design is to design novel molecules which have desired properties or activities. To achieve this goal, we recently developed fragments combination-typed structure generator for inverse QSPR/QSAR analysis. The fragments are collected from the samples in the training data set, which is used to construct QSPR/QSAR models. These fragments are ring system [1] based. By combining the ring system based fragments and atom typed fragments with hydrogen atoms in tree-like way, we can enumerate structures. Concepts of applicability domain (AD), which is data domain where target QSPR/QSAR models have predictive ability, is considered automatically. QSPR/QSAR models are constructed in the form of probabilistic models and thus the result of inverse QSPR/QSAR analysis is also in the form of conditional probabilistic distribution given the desirable values of objective variables. The specific type of descriptors is used in order to enhance computational efficiency. Only descriptors whose values increase when a fragment is attached to a growing structures are adopted. We named these descriptors Monotonous Changing Descriptors (MCDs) [2]. In order to avoid making duplicated structures even when we use combining fragments, we use Mckay's canonical construction path method [3].

Although there are some restrictions with our system, i.e. using ring system based fragments, the region in chemical space being restricted with AD, using MCDs, and generating tree-like structures, we can enumerate 2D based chemical structures focusing only on the structures having desired values of properties and/or activities. Also, we calculate the values of MCDs efficiently during generating process and explain the methods in our poster in addition to overall concept of the system for inverse QSPR/QSAR.

Bibliography:

[1] G. W. Bemis, M. A. Murcko, J. Med. Chem. 39(1996), 2887-2893
[2] T. Miyao, M. Arakawa, K. Funatsu, Mol. Inf. 29(2010), 111-125
[3] B. D. McKay, J. Algorithms 26(1998), 306-324