Condensed Graph of Reaction - a Swiss-army knife for chemical reactions mining

Alexandre Varnek, University of Strasbourg

Chemical reactions are difficult objects in chemoinformatics because they involve several species of two different types: reactants and products. The *Condensed Graph of Reaction* (CGR) approach opens new perspectives in reaction mining since it allows to transform several 2D molecular graphs describing a chemical reaction into a single graph. Besides conventional chemical bonds (simple, double, aromatic, etc), a CGR contains dynamical bonds corresponding to created, broken or transformed bonds, and dynamic atoms describing the atoms which change their formal charge upon chemical transformation. Thus, a chemical reactions database can be transformed into a set of "pseudo-molecules" which can be encoded by special CGR/SMILES strings or related fingerprints. In such a way, most of chemoinformatics methods developed for individual molecules – similarity searching, clustering, machine-learning modeling, etc. - can also be applied to chemical reactions.

In this presentation we demonstrate how CGR approach can efficiently be used in different reaction informatics tasks. Some recent applications of CGR to predictive modeling of thermodynamic and kinetic parameters of reactions, design of new enantioselective catalysts, and automatized processing of information in large reaction databases will be discussed. Particular attention will be paid to the integration of CGR with modern Artificial Intelligence tools leading to discovery of novel chemical transformations.