Chemistry as a graph rewriting system: generative exploration of chemical space

Jakob Lykke Andersen[1], Christoph Flamm[2], Daniel Merkle[1], Bärbel M. R. Stadler[3], and Peter F. Stadler[4]

[1]Department of Mathematics and Computer Science, University of Southern Denmark, Odense M DK-5230, Denmark, jlandersen,daniel@imada.sdu.dk
[2]Institute for Theoretical Chemistry, University of Vienna, Wien A-1090, Austria, xtof@tbi.univie.ac.at
[3]Max Planck Institute for Mathematics in the Sciences, Inselstrae 22, D-04103 Leipzig, Germany, baer@bioinf.uni-leipzig.de
[4]Bioinformatics Group, Department of Computer Science and Interdisciplinary Center for Bioinformatics, University of Leipzig, Leipzig D-04107, Germany, stadler@bioinf.uni-leipzig.de

Undirected, labeled graph have been recognized as appropriate models of molecules, and indeed underly the established notation in chemistry. Chemical reactions thus are naturally viewed as transformations of (usually not connected) graph. This simple idea is rigorously formalized by graph grammars and thus amenable to practical implementation. Grounded soundly in category theory, the framework can be extended to accomodate also geometric aspects beyond pure topology, i.e., stereochemistry. An important aspect of the mathematical framework is its capability to describe composite rules, setting the stage for computational approaches to disentangle "overall" chemical reactions into elementary steps. The rule based framework, furthermore, provides explicit atom-atom maps.

Starting molecules and reaction rules naturally span chemical spaces, in which molecular types are points and chemical reactions form directed hyperedges. These spaces may be finite or infinite, as in the case of reaction rules that support polymerization. These structure of these spaces fall outside the realm of standard topology, but retains a rich structure that naturally generalizes certain types of proximity or - equivalently - separation spaces. These structures in turn are equivalent to abstract notions of connectivity and reachability.

References

[1] J. L. Andersen, Ch. Flamm, D. Merkle and P. F. Stadler, Inferring chemical reaction patterns using graph grammar rule composition, Journal of Systems Chemistry **4** (2013), DOI 10.1186/1759-2208-4-4.

[2] J. L. Andersen, Ch. Flamm, D. Merkle and P. F. Stadler, Generic strategies for chemical space exploration, International Journal of Computational Biology and Drug Design **7** (2014), 225–258.

[3] J. L. Andersen, Ch. Flamm, D. Merkle and P. F. Stadler, Chemical Graph Transformation with Stereo-information, ICGT 2017, Lecture Notes Computer Science **10373** (2016), 54–69.

[4] J. L. Andersen, Ch. Flamm, D. Merkle and P. F. Stadler, An intermediate level of abstraction for computational systems chemistry, Philos. Trans. Roy. Soc. A **375** (2017), ID 20160354.

[5] J. L. Andersen, Ch. Flamm, D. Merkle and P. F. Stadler, Chemical transformation Motifs-modelling pathways as integer hyperflows, IEEE/ACM Transactions on Computational Biology and Bioinformatics (2018), DOI 10.1109/TCBB.2017.2781724.

[6] B. M. R. Stadler and P. F. Stadler, Reachability, connectivity, and proximity in chemical spaces, MATCH Commun. Math. Comput. Chem. (to appear).