

[P52] Synthesis Route Optimisation Using the Network of Organic Chemistry

Philipp-Maximilian Jacob, Nicolai Peremezhney, Alexei Lapkin

Department of Chemical Engineering and Biotechnology, University of Cambridge, Pembroke Street, CB2 3RA Cambridge, UK

In a recent paper [1] we have proposed a new concept for the processing of biofeedstocks in the formulated consumer products industry. The paper proposes two new approaches, one based on metabolic engineering and dynamic combinatorial chemistry and the other using knowledge-based expert systems. This communication shall seek to further develop the knowledge-based expert system approach. Particular interest is paid to the Network of Organic Chemistry (NOC) first proposed in Ref. [2]. Such a network seeks to convert the Reaxys database of chemical reactions into a network, thus allowing graph theoretical analysis of the reactions and the molecules' interconnections. As a result, synthesis planning can in theory be carried out using suitably designed network traversal algorithms drastically reducing the time spent on literature research. When developing competing process models it is of vital importance to include supply chain and market considerations to obtain meaningful results. To this end it would be desirable to be able to quickly generate suitable synthesis routes and the NOC could thus be very useful. The NOC however only considers existing chemical knowledge and we feel is thus limited in the biofeedstock area, where many conversions are still to be developed. The future investigation will thus pursue two main avenues to target the discovery process for new reactions in targeted areas initially as proof-of-concept. Firstly the network will be analysed statistically using a hierarchical network algorithm developed by [3] for the purpose of edge prediction in reaction networks. This will be carried out with the aim of demonstrating the suitability of statistical tools in the prediction of new reactions within the scope of the NOC when planning a synthesis. In parallel we aim to implement a methodology for basic catalyst screening similar to that proposed in Ref. [4] in order to ultimately use this methodology in the context of the NOC to investigate the potential of catalytically improving existing routes or facilitating new ones. A vital area of ongoing investigation is the development of heuristics triggering these suggested methodologies during the network search based upon the information associated with the compounds and reactions stored in the network as well as a suitable database architecture to ensure that the data and network can be queried quickly and that all data is available in a machine-readable form to allow its automatic use.

[1] Peremezhney, N.; Jacob, P.-M.; Lapkin, A. *FCHEM*. 2 (2014) 1–7.

[2] Fialkowski, M.; Bishop, K. J. M.; Chubukov, V. A.; Campbell, C. J.; Grzybowski, B. A. *Angew. Chem. Int. Edit.* 44 (2005) 7263–7269.

[3] Clauset, A.; Moore, C.; Newman, M. E. J. *Nature*. 453 (2008) 98–101.

[4] Nørskov, J. K.; Abild-Pedersen, F.; Studt, F.; Bligaard, T. P. *Natl. Acad. Sci. USA*. 108 (2011) 937–943.