TOPOLOGICAL ANALYSIS OF THE CHEMICAL SPACE: UNDERSTANDING AQUEOUS SOLUBILITY.

JACEK BRODZKI

In this talk we will present an application of topological data analysis to understand the structure of the descriptor space of molecules produced from a standard chemical informatics software. We are interested in discovering indicators of when a chemical compound is soluble in water. We have used the mapper algorithm, a TDA method that creates low-dimensional representations of data, to create a network visualization of the solubility space. While descriptors with clear chemical implications are prominent features in this space, reflecting their importance to the chemical properties, the topological analysis has uncovered new and interesting chemical properties responsible for water solubility.

We have also considered a representation of the chemical space using persistent homology applied to molecular graphs, and we have discovered that links between this chemical space and the descriptor space are in agreement with chemical heuristics.

References

[1] M. Pirashvili, F. Belchi Guillamon, L. Steinberg, M. Niranjan, J. Frey, J. Brodzki, *Improved understanding of aqueous solubility modeling through Topological Data Analysis*, Journal of Cheminformatics, in press, 2018.

(Jacek Brodzki) Centre for Geometry, Topology and Applications, School of Mathematical Sciences, University of Southampton, Southampton SO17 1BJ, UK

E-mail address: j.brodzki@soton.ac.uk

This work was supported by the EPSRC grant EP/N014189/1.