PhD Position at the heart of Europe, Nancy - France

Coupling artificial intelligence and molecular modeling tools for fast and reliable chemical product design

Laboratory of Reactions and Process Engineering, LRGP CNRS UMR 7274, at Nancy, France.

Keywords: **Modeling tools, machine learning, molecular modeling, Monte Carlo methods, statistical analysis of data, thermodynamic properties of molecules, deep learning**

General context

The accurate modeling of thermodynamic and kinetic properties of mixtures requires knowledge of thermochemical and thermophysical molecular properties. However, experimental data of pure molecules are not always available for numerous species of interest and, in certain cases, some properties of molecules in their pure form cannot even be measured in practice. The development of computational tools enabling the prediction of such properties is thus a pivotal requirement for the characterization of these products and for the further development of innovative solutions in different scientific areas, such as the areas of product engineering, energy, chemical reaction kinetics and retrosynthesis, process simulation, catalyst design, drug discovery, and, in general, computer-aided molecular design (CAMD).

Today, Group-Contribution (GC) methods, Quantum Chemistry (QC) and Monte Carlo (MC) simulations are used for the prediction of these properties, but their reliability is often subject to question, especially when it comes to newly developed molecules. Accordingly, great effort is being directed, over the last years, to the development of alternative predictive tools that could act either synergistically or as alternative solutions to the aforementioned approaches. These efforts are principally based on the implementation of data-driven techniques and, more specifically, of methods belonging to the greater family of machine learning (ML) or artificial intelligence (AI).

Objectives of the PhD

In this thesis, a coupling of QC, MC and ML techniques is proposed in order to build a powerful predictive tool for the determination of thermochemical (e.g., enthalpy and entropy of formation) and thermophysical (critical temperature, critical pressure, acentric factor and the specific heat capacity) properties of different classes of molecules. The automatization of the calculation of QC properties as well as the postprocessing of its outputs will also be part of the objectives of the thesis. In addition, the use of graphical methods, such as Graph Neural-Networks (GNN) will be investigated as alternative to the use of molecular descriptors, for a more robust representation of the molecules and control of the desired molecular characteristics that need to be considered as inputs to the model. Finally, the developed models will be benchmarked, whenever possible, against GC and MC results and will be further employed in applications for which the implementation of the latter is constrained or infeasible. A preliminary framework for this study has already been established in LRGP in the framework of previous PhD and Master theses.

Student Profile

We are looking for a highly-motivated PhD candidate, with a background in chemical or process engineering. The candidate should demonstrate a specific preference and talent for computer programming and simulation. More specifically:

Prerequisites

- An established background in statistical analysis and/or data-driven modeling techniques.
- Good knowledge of Python and/or Matlab, or any other similar package (e.g., Scilab).
- Excellent written and spoken English.
- Solid background in thermodynamics.
- Motivation to carry out a major part of the PhD in developing and running PC programs.

- Organizational skills and capacity to adapt and evolve his/her learning methodology.

Additional appreciated skills

- Master degree or certifications in related fields (e.g., data treatment, molecular simulations, machine learning etc.).
- Any level of French and/or willingness to improve it.

Related indicative literature

Trinh C., et al., 2022, Proceedings of the 32nd European Symposium on Computer Aided Process Engineering, ESCAPE 32, 1471. https://doi.org/10.1016/B978-0-323-95879-0.50246-0.
Yalamanchi et al., 2020, J. Phys. Chem. A 2020, 124, 6270−6276. https://dx.doi.org/10.1021/acs.jpca.0c02785

Applications should be sent, by June 11, to the following addresses (domain is: “univ-lorraine.fr”):
dimitrios.meimaroglou [at] domain
olivier.herbinet [at] domain