DISCOVERY INFORMATICS: A MODEL-DRIVEN MULTI-SCALE INTEGRATED FRAMEWORK FOR PRODUCT DESIGN AND ENGINEERING

Venkat Venkatasubramanian, James M. Caruthers, W. Nicholas Delgass
Center for Catalyst Design
School of Chemical Engineering, Purdue University
West Lafayette, IN 47907, USA

Abstract

Designing new materials and formulations with desired properties is an important and challenging problem, encompassing a wide variety of products. Traditional trial-and-error design approaches are laborious and expensive, delay time-to-market, and miss some potential solutions. Furthermore, the growing avalanche of high throughput experimentation data has created both an opportunity and a major informatics challenge for materials design and discovery. However, we cannot use experimental data alone to "find" new materials, as we must have knowledge to guide the search more effectively. A new paradigm is needed that increases the generation of potential leads significantly, broadens the search horizon, and extracts knowledge from today's successes and failures in order to accelerate the discovery of new products tomorrow. Towards this end, over the past decade we have been developing a novel model-based, multi-scale integrated methodology called *Discovery Informatics*. This framework enables the management of large complex data sets, systematic extraction and accumulation of knowledge, iterative model refinement via hypotheses testing by interaction with experiments, and efficient search for new materials with desired performance characteristics. This paradigm has two main components: a *forward model* connecting descriptors of material structure or formulation to performance and an *inverse process* to determine optimal material structure or formulation from desired performance. While a rigorous first-principles forward model is desirable, a totally fundamental approach is often not practical. Therefore, a hybrid approach is adopted, where one can guide *convergence* to an increasingly fundamental model in concert with an information-rich data flow. Such a hybrid model combines fundamental knowledge, expert rules and statistical correlations of experimental data. We have developed an enabling intelligent modeling system – Reaction Modeling Suite that aids the human expert in building kinetic forward models facilitating knowledge extraction. Once a robust forward model is available, the inverse problem is addressed using evolutionary algorithms that exploit system knowledge to efficiently search the combinatorial space for materials matching performance targets. We will present the application of this paradigm on three widely different industrial product design domains: fuel-additives design, sulfur-vulcanized elastomers formulation and catalyst design.

Keywords: Material design, Product engineering, Informatics, Artificial intelligence, Knowledge engineering.