Really Big Data from HPC-enabled biomedical agent-based modeling: Pathways to Precision Medicine

Gary An and Chase Cockrell, Department of Surgery, University of Chicago

Background: Precision medicine should mean "the right drug for the right patient at the right time," and include a means of efficiently engineering the appropriate control strategies in order to fulfill this goal. Machine learning (ML) and evolutionary computing (EC) are powerful computational methods made possible by the enhanced large-scale collection of data and computational power provided by high performance computing (HPC). However, despite best efforts, the scale of data collected in the biomedical arena is too temporally sparse from too heterogeneous a group of sources to utilize the full capabilities of ML/EC for dynamic disease processes. There are also fundamental constraints to the extrapolation of laboratory data to the general case, and how computational models are currently evaluated and used in this context. We pose that overcoming the limitations of clinical, experimental and current computational methods requires the use of quasi-mechanistic dynamic simulations as proxy systems implemented on HPC environments to characterize the behavioral landscapes of biological systems, provide epistemic boundary conditions for investigative strategies and direct the development of novel methods aimed at achieving system characterization and control.
Methods: We utilize a HPC version of a previously validated agent-based model (ABM) of sepsis as a proxy model in order to identify epistemic boundary conditions for the study of sepsis and identify novel investigatory uses for multi-scale simulations. We perform mathematical analysis derived from the study of Random Dynamical Systems (RDS) to identify two novel means for characterizing system behavior: Probabilistic Basins of Attraction (PBoA) and Stochastic Trajectory Analysis (STA). We also employed a genetic algorithm (GA) as an example of how evolutionary computing can be used to search for putative control strategies to determine controllability for a complex dynamical system.
Results: Simulations were performed on the Edison Cray XC30 Supercomputer at the National Energy Research Scientific Computing Center and on the Beagle Cray XE6 Supercomputer at the University of Chicago. Behavior space characterization using > 75 million simulated sepsis patients defined a region of clinically plausible parameter space for the sepsis ABM. PBoAs demonstrated multi-dimensional attractor basins for the system, and STA characterized the probabilistic inflection "zones" between attractor outcomes. The GA performed on a specific parameter set and initial perturbation identified effective combinations of multi-modal and sequential interventions, as well as rescue therapies for initial non-responders.
Conclusions: HPC simulations of ABMs as abstracted proxy models can be used to establish "boundaries of futility" with respect to existing investigatory approaches and provide novel means of characterizing what constitutes biological heterogeneity, both in terms of system state (PBoA) and trajectory (STA). HPC-enabled simulation experiments also allow the use of ML/EC to determine the controllability of a complex dynamical system and provide a first approximation as to the scope of such a task, which in turn can direct concurrent clinical and experimental investigations to the appropriate levels of mechanistic and temporal resolution, a necessary step towards true Precision Medicine.