Abstract:
In many scientific domains, researchers are turning to large-scale behavioral simulations to better understand important real-world phenomena. These phenomena emerge as the result of a myriad interactions among large numbers of interdependent agents in a complex system, such as a transportation network or an ecological system. While there has been a great deal of work on simulation tools from the high-performance computing community, behavioral simulations remain challenging to program and automatically scale in parallel environments. In this talk, I will show how database techniques can solve this dilemma, by offering simulation developers a programmable environment that automatically provides for scalability. I will present the design of BRACE, the Big Red Agent-based Computation Engine. BRACE at the same time offers a high-level scripting language for simulation developers and scales scripts written in this language by modeling them as iterated spatial joins. I will discuss recent and ongoing work on techniques to execute these iterated joins efficiently in the cloud, with the goal of turning BRACE into an important platform for new mid-range scientific computing workloads.