Local Optimization for Simulation of Natural Motion

Tom Erez
Washington University in St. Louis
etom@cse.wustl.edu

Reinforcement Learning (Sutton and Barto 1998) is a theoretical framework for optimizing the behavior of artificial agents. The notion that behavior in the natural world is in some sense optimal is explored by areas such as biomechanics and physical anthropology. These fields propose a variety of candidate optimality criteria (e.g.,) as possible formulations of the principles underlying natural motion.

Recent developments in computational biomechanics allow us to create articulated models of living creatures with a significant degree of biological realism (Delp et al. 2007). I aim to bring these elements together in my research by using Reinforcement Learning to generate optimized behavior in biomechanical simulations. Such a generative approach will allow us to examine critically postulated optimality criteria and investigate hypotheses that cannot be easily studied in the real world.

Background — Reinforcement Learning

The Reinforcement Learning (RL) agent interacts with a dynamical system whose states capture all the relevant information about the current configuration of the agent and its environment. By specifying a sequence of actions, the agent alters the state transitions of this dynamical system. The optimality criterion is formalized by a reward function defined over state-action pairs, and the agent's goal is to maximize the cumulative reward.

The target of optimization in RL is the policy, mapping from states to actions — the optimal policy chooses the action that would yield maximum future rewards. Another important construct is the value function, which determines the desirability of different states by measuring what is the maximum cumulative reward that can be expected from a given state. Given the optimal value function, the optimal policy is reduced to a greedy choice.

The basic algorithms of RL in discrete state and action spaces use tabulation to approximate the value function. Such algorithms do not scale to the continuous case — if we tried to discretize a continuous state space and use discrete RL methods, we would soon hit a wall, as the number of discretized states grows exponentially with the dimensionality of the continuous state space. Therefore, continu-
task is considered, with a smoother fitness landscape. The locally-optimal solution is then used as an initial guess for a slightly harder variant. This iterative process tackles the full complexity of the domain gradually, and allows the locally optimal optimization to discover the more relevant parts of state-space.

3. Studies of the Hamilton-Jacobi-Bellman equation in optimal control have shown that the value function of a non-linear domain can be discontinuous, even if the domain itself is perfectly continuous (Crandall and Lions 1983). This is a serious hurdle for many value-function approximation algorithms. However, the same theory also suggests that in the presence of domain stochasticity, these discontinuities dissipate and smooth out (Tassa and Erez 2007). I argue that this is a crucially-beneficial feature, which can be harnessed to ensure the differentiability of dynamics with discontinuities, such as collisions (Erez and Smart).

**Proposed Work**

The most important component of my future research is the advancement and development of algorithms for model-based local optimization. This direction is not perfectly aligned with current trends in RL (that focus on discovering the dynamical model, and on partially-observable systems), but I believe that it has the potential to open the door to a whole host of high-quality studies, with contribution to both RL and potential application domains.

Ultimately, the testbed for such local optimization techniques will be the generation of simulated natural motion from first principles. In order to advance my research towards natural motion applications, I will use OpenSim (Delp et al. 2007), a general-purpose tool for biomechanical modeling and simulation. As an active member of the development team, I recently built an interface between OpenSim and MATLAB, and I can now harness existing RL algorithms to work with OpenSim models. By applying local optimization to OpenSim models, I aim to take a generative approach to the study of the various postulated optimality criteria.

There have been several studies of biomechanics and anthropology that studied the simulation of natural motion from first principles (Chowdhary and Challis 1999; Pandy and Anderson 2000; Nagano et al. 2005; Sellers and Manning 2007), employing a variety of optimization techniques, such as simulated annealing, genetic programming, and policy gradient. However, no previous study within these disciplines tackled the generation of feedback policies, or harnessed the computational efficiency afforded by local approximation of the value function. The goal of my work is to develop techniques that could be adopted as standard tools of investigation by the scientific communities of computational biomechanics and physical anthropology.

**References**


