Manifold Learning and Its Applications: Papers from the AAAI Fall Symposium (FS-10-06)

Treating Epilepsy by Reinforcement Learning Via Manifold-Based Simulation

Keith Bush and **Joelle Pineau**
McGill University
Montreal, QC, Canada
{kbush, jpineau}@cs.mcgill.ca

Introduction

The ability to take intelligent actions in real-world domains is a goal of great interest in the machine learning community. Unfortunately, the real-world is filled with systems that can be partially observed but cannot, as yet, be described by first principle models. Moreover, the traditional paradigm of direct interaction with the environment used in reinforcement learning is often prohibitively expensive in practice.

An alternative approach simultaneously solves both of these problems by using simulated interaction with the environment rather than real-world experience. The simulation in this approach is a computational model of a dynamical system. The barrier to linking intelligent control with real-world domains is, therefore, one of identifying high-quality state-spaces and transition functions from observations.

From a dynamical systems perspective, this barrier is analogous to the problem of finding high-quality manifold embeddings and a rich literature of theory and practice exists to address it. The contribution of this work is two-fold. First, we describe an approach for learning optimal control strategies directly from observations using manifold embeddings as the intermediate state representation. Second, we demonstrate how control strategies constructed in this way can answer important scientific questions. As a concrete example, we use our approach to guide experimental decisions in neurostimulation treatments of epilepsy.

Neurostimulation Treatment of Epilepsy

Epilepsy afflicts approximately 0.5–1% of the world's population (Kotsopoulos et al. 2002). Of those sufferers, approximately 30% do not respond to available anti-convulsant drugs or are not candidates for surgical resection. Development of new treatments, therefore, is a priority for epilepsy research.

Neurostimulation shows promise as an epilepsy treatment. *In vitro* studies have demonstrated that fixed-frequency external electrical stimulation applied to substructures within the hippocampus can effectively suppress seizures (Durand and Bikson 2001). However, results of *in vitro* experiments suggest that the suppression efficacy of fixed-frequency stimulation varies across epileptic neural systems. It has also been shown that stimulation can negatively impact neural tissue.

Neurostimulation experiments are constrained in three ways that suggest that optimal treatment strategies should be learned in simulation using manifold-based representations: 1) the complex dynamics of neural systems are typically observable only through low-dimensional time-series corrupted by noise, whereas optimal control requires complete state, 2) the lifespan of neural tissue *in vitro* is on the order of a few hours, which limits the use of on-line system identification, and 3) the best trade-off between suppression efficacy and minimization of stimulation side-effects is difficult to determine *a priori*. Using simulated learning experiments, we can describe the relationship between cost function parameters and experimental observations: seizure suppression efficacy and effective stimulation frequency of the adaptive policy. This knowledge allows epileptologists to select cost function parameters that maximize the chance of observing statistically significant adaptive treatment performance compared to fixed-frequency treatments *in vitro*.

Simulation Construction

We summarize our approach for constructing a manifold-based simulation of neurodynamics in Figure 1(a–d). Our biological model of epilepsy is the hippocampal-EC slice perfused with 4-aminopyridine. Field potential recordings of epileptiform activity are recorded from the entorhinal cortex (see Fig. 1(a) EC). External stimulation is applied to the subiculum (see Fig. 1(a) Sub).

We collected a dataset of field potential recordings from seven slices under fixed-frequency stimulation of 0.2, 0.5, 1.0, and 2.0 Hz as well as unstimulated (36,581 seconds including 83 seizures). Examples of the dataset are shown in Figure 1(b) for 0.2 Hz and 1.0 Hz fixed-frequency stimulation.

Using spectral subspace identification (Galka 2000), we extracted manifold embedding parameters of the dataset (see Fig. 1(c)) and then projected the dataset onto this manifold (see Fig. 1(d)). This manifold is the state-space of our simulation. We defined the transition function to be the local time-derivative of the element of the dataset that is nearest to the current simulation state (Parlitz and Merkwirth 2000). Actions in the model are simulated by conditioning the selection of the nearest neighbor derivative based on a desired

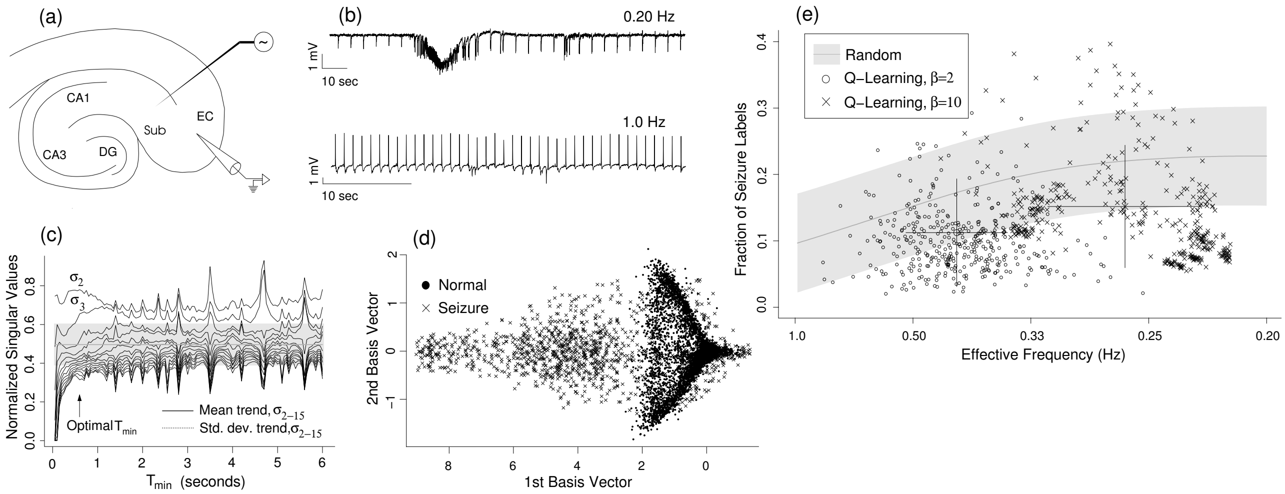

Figure 1: (a) Schematic of the hippocampal-EC slice with placement of stimulation and recording electrodes, (b) examples of field potential recordings observed under fixed frequency stimulation, (c) subspace identification spectrum, (d) neurostimulation manifold, and (e) efficacy and effective frequency comparison between random fixed-frequency and adaptive neurostimulation.

action, either *on* or *off*. Each point in the dataset was labeled with an action *on* or *off* during data acquisition.

Simulated Adaptive Control

We trained a neurostimulation agent using reinforcement learning (RL), a strategy that optimally solves multi-step decision tasks (Sutton and Barto 1998). We defined the control actions to be either 1.0 Hz or 0.2 Hz fixed-frequency stimulation. We approximated the objective function by discretizing the manifold state-space. We defined the cost function, r, to penalize both stimulation frequency and the number of seizures observed: $r = -\beta(r_{stim}) - \rho(r_{seiz})$, where $r_{stim} = 1$ if a is 1.0 Hz, otherwise $r_{stim} = 0$ and $r_{seiz} = 1$ if the current state is a seizure state, otherwise $r_{seiz} = 0$. Parameters β and ρ are tunable. We identified treatment strategies using two different cost function parameter configurations, $\beta = 2$ and $\beta = 10$ ($\rho = 1$). The parameters explore how stimulation and seizure costs influence suppression performance and effective frequency of the learned treatment.

As a reference, we also applied random control to the simulation. The random controller was implemented by uniformly sampling the actions (either the 0.2 or 1.0 Hz fixed-frequency) according to a ratio determined by the desired effective frequency. The ratio was varied to cover the entire frequency spectrum 0.2–1.0 Hz.

The seizure suppression effects of random control are summarized in Figure 1(e) as a distribution of seizure fractions over effective stimulation frequency, plotted as a bold gray line centered within a shaded region (mean$\pm$std. dev.). Individual trials of the learned control strategies for both $\beta = 2$ and $\beta = 10$ are plotted as open circles and $\times$ symbols, respectively. Distributions (mean$\pm$std. dev.) over the results of these two parameter configurations are plotted as cross-hairs. Both adaptive control policies produce seizure suppression outcomes that are significantly better ($p < 0.05$) compared to random at the respective effective frequencies.

Discussion

The primary contributions of this work are: 1) to describe how controllable simulations of complex, partially observable, poorly understood dynamic domains, such as neural systems, can be constructed, and 2) to demonstrate the utility of simulated experience in guiding parameterization decisions of real-world adaptive control experiments. Considering the expense, and in many cases the intractability, of real-world experience, we propose simulating adaptive control experiments to identify cost function parameterizations that will most likely produce significant results in real-world adaptive neurostimulation experiments. In this application manifolds play the critical role of state representation, providing RL algorithms access to real-world domains. We have also applied this approach to real-world control problems (Bush and Pineau 2009).

References

Bush, K., and Pineau, J. 2009. Manifold embeddings for model-based reinforcement learning under partial observability. In Bengio, Y.; Schuurmans, D.; Lafferty, J.; Williams, C. K. I.; and Culotta, A., eds., *Advances in Neural Information Processing Systems 22*, 189–197.

Durand, D. M., and Bikson, M. 2001. Suppression and control of epileptiform activity by electrical stimulation: A review. *Proceedings of the IEEE* 89(7):1065–1082.

Galka, A. 2000. *Topics in Nonlinear Time Series Analysis: with implications for EEG Analysis*. World Scientific.

Kotsopoulos, I.; van Merode, T.; Kessels, F.; de Krom, M.; and Knottnerus, J. 2002. Systematic review and meta-analysis of incidence studies of epilepsy and unprovoked seizures. *Epilepsia* 43(11):1402–1409.

Parlitz, U., and Merkwirth, C. 2000. Prediction of spatiotemporal time series based on reconstructed local states. *Physical Review Letters* 84(9):1890–1893.

Sutton, R., and Barto, A. 1998. *Reinforcement learning: An introduction*. Cambridge, MA: The MIT Press.