Chapter 1

Introduction

In a number of problems, including Brain-Computer Interfaces (BCI), deep brain stimulation (DBS), sensory prosthetics, and spinal cord injury (SCI) therapy, complex electronic and computational systems interact with the human central nervous system. An important open question is how to control these engineered systems or agents to produce results which are in some sense optimal, e.g., efficiently decode user intent in BCI or mitigate both tremors and bradykinesia in DBS. This dissertation is concerned with electrical stimulus applied to the spinal cord via multi-electrode arrays. The purpose of this stimulation is to promote the function and rehabilitation of the remaining neural circuitry below the injury, with the goal of enabling the performance of complex motor behaviors, e.g., stepping and standing. To enable the careful tuning of these stimuli for each patient, the electrode arrays which deliver these stimuli have become increasingly more sophisticated, with a corresponding increase in the number of free parameters over which the stimulus must be optimized. Due to the exponential explosion of the sets of possible stimuli, a more rigorous, algorithmic method of selecting stimuli is necessary, particularly in light of the expense and relative inaccessibility of expert hand-tuning. The present work proposes to use a family of recent (GP-UCB, Srinivas et al., 2010) and novel active learning algorithms (GP-BUCB, Desautels et al., 2012, and GP-AUCB) for this purpose. This dissertation develops the GP-BUCB and GP-AUCB algorithms, and bounds their regret (i.e., their sub-optimality over the therapeutic period, as compared with the optimal fixed policy). It compares these novel algorithms with GP-UCB and several competing algorithms in simulation and shows that GP-BUCB and GP-AUCB are competitive with the state of the art. A variant of GP-BUCB was implemented in a closed-loop animal experiment, controlling which epidural stimulating electrodes are used in the spinal cord of an SCI rat; the results obtained are compared with concurrent stimulus tuning carried out by human experimenter. These experiments show that this algorithm is at least as effective as the human experimenter, suggesting that this algorithm can be applied to the more challenging problems of enabling and optimizing complex, sensory-dependent behaviors, such as stepping and standing in SCI patients.

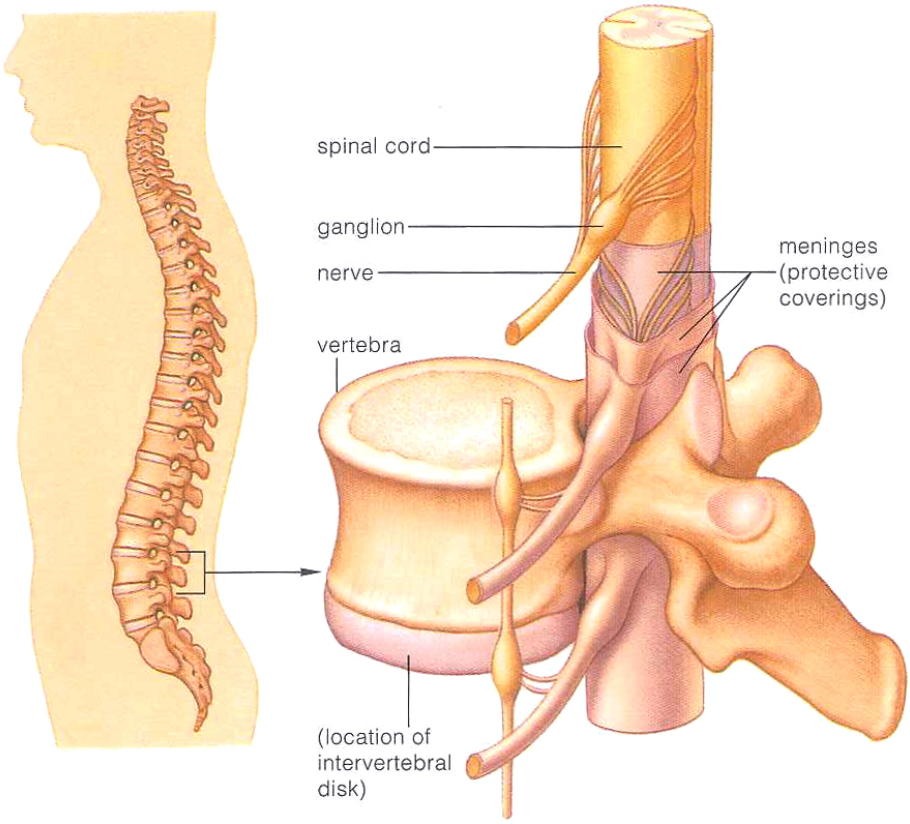

Figure 1.1: The vertebral column and spinal cord. The vertebral column, shown at left, both provides articulating support for the body and protects the spinal cord. At right, the spinal cord is shown with the associated vertebra, dorsal and spinal roots, dorsal ganglia meninges, and sympathetic nerve ganglia (the chain-like structure parallel to the spinal cord, not labeled). The dura is the outermost of the three meninges; the middle is the arachnoid and the innermost is the pia. Note that the spinal cord is contained within the vertebral canal, such that it is surrounded by fat (not shown) and bone which respectively cushion and protect it. From Starr/Taggart. Biology: The Unity & Diversity of Life w/ CD & InfoTrac, 10E. © 2004 Wadsworth, a part of Cengage Learning, Inc. Reproduced by permission. www.cengage.com/permissions

1.1 Spinal Cord Injury

For the purposes of this dissertation, a *spinal cord injury* (SCI) is defined as a traumatic injury to the spinal cord resulting in a loss of function, especially of the arms or legs. This differentiates SCI from many other disease-related losses in spinal cord or nervous function, such as amyotropic lateral sclerosis (ALS), poliomyelitis, or multiple sclerosis (MS).

Briefly, drawing from the text by Kandel et al. (2012), the spinal cord is both the major connection between the brain and most of the body and a local processing system for many reflexes, as well as more complex behaviors such as locomotion. The spinal cord extends from the brain stem (at the base of the skull) to the first lumbar vertebrae, and has enlargements at the levels which innervate the arms and legs, called the cervical and lumbar enlargements, respectively. Within it, the spinal

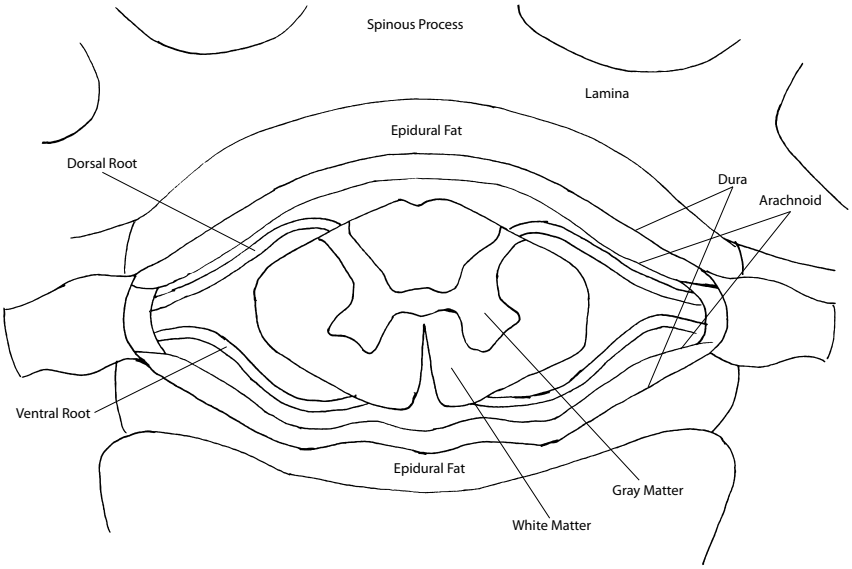

Figure 1.2: Schematic diagram of the the spinal cord in cross-section. The dorsal surface of the body is upward in this figure. The spinal cord is composed of the gray matter and white matter regions in the center of the figure. The epidural electrode array used for EES (see Section 1.2) is implanted in the epidural space, flush with the outside of the dura; this location is approximately where the dorsal of the two labels reading "Epidural Fat" is located. Figure produced with reference to Figure 4.2 of Watson et al. (2008).

cord contains both *gray matter*, composed primarily of the cell bodies of neurons, and *white matter*, composed of tightly packed axons carrying information rostro-caudally; due to extensive myelination of these ascending and descending axons, the white matter does in fact appear lighter in color than the gray matter. When viewed in cross-section, the gray matter is organized in a butteryfly-shape; the dorsal "wings" are referred to as the *dorsal horns*, and the corresponding structures on the ventral side of the cord are called the *ventral horns*. Packed around the periphery of the spinal cord are the well-organized tracts of white matter (see Figure 1.1). The spinal cord, like the brain, is organized into laminae, which are numbered from I (most dorsal) to X (most ventral). Sensory neurons from the body project into the spinal cord through the dorsal roots, and may terminate in a stereotyped fashion in one or more laminae. For the purposes of this dissertation, the most relevant of these projections are those by Aα sensory neurons into or close to the ventral horn, which hosts the motoneurons which control skeletal muscles. Large populations of inhibitory spinal interneurons also reside in the spinal cord and regulate motor activities through interaction with the motoneurons and each other. The motoneurons in turn project out to the body through the ventral roots, which merge with the incoming dorsal roots to form the spinal nerves, which thus carry a mix of afferent and efferent fibers. A cross section of the spinal cord is shown in Figure 1.2. The spinal nerves are designated with the name (e.g., L1) of the vertebra which is below them (cervical) or above them (thoracic, lumbar, and sacral nerves) as they exit the vertebral column, and this name

is conferred upon the spinal segments which give rise to the nerve; since spinal nerves remain in the vertebral column for some distance caudal to their origin in the spinal cord, the designation of spinal segements is thus shifted with respect to the designations of the vertebrae. These spinal nerves also innervate well-defined regions of the body called *dermatomes* (for sensation) or *myotomes* (for motor control) (Kandel et al., 2012, pp. 338-340, 357-359, and 488-490). Another thorough treatment of the organization, function, and dysfunction of the spinal cord, can be found in the text and atlas by Watson et al. (2008).

Patients with SCI present different clinical symptoms depending on the location of the injury within the spinal cord, including a variety of syndromes which are symptomatic of damage to different structures within the spinal cord. Sufficiently severe damage to the spinal cord can result in the loss of voluntary control (frequently accompanied by loss of sensation as well) of the legs (paraplegia) or the legs and arms (quadraplegia). The severity of a patient's injury is most commonly assessed on the ASIA (American Spinal Injury Association) scale, as well as by the neurological level of the injury in the spinal cord, diagnosed via the affected dermatomes and myotomes, which correspond in a fixed fashion to spinal levels.

1.2 Epidural Electrostimulation

One technique for SCI therapy is Epidural electrostimulation (EES). EES involves electrically stimulating the spinal cord via an electrode or multi-electrode array placed in the epidural space (see Figure 1.2). Historically, spinal electrostimulation has been applied for a number of purposes, including the the alleviation of chronic pain (Shealy et al., 1967a,b). EES has also been used for the treatment of motor deficits, such as cerebral palsy; this field drove a push toward more complex and capable stimulating devices (for an insider's view of this early history, see Waltz, 1997). Recent leads and electrostimulators (such as the Specify 5-6-5 and RestoreAdvanced, Medtronic, Minneapolis, MN) provide great flexibility in terms of which of many electrodes are active and what stimuli are applied with these electrodes. These increased capabilities allow complex stimuli to be customized post-implantation. A single device can thus accommodate changes in the stimuli as the optimal parameters change with time, as well as variations in surgical placement, injury, and patient-specific needs for symptom alleviation.

Mechanistically, SCI therapy by EES is intended to promote activity, particularly closed-loop activity, of the spinal cord below the site of the injury. This is accomplished by applying a tonic electrical stimulus to activate specific networks and structures in the spinal cord. This stimulus is typically not intended to drive the desired activity directly; rather, stimuli enable the patient's native neural circuits to regulate motor activity according to the sensory environment of the patient, such that the responses, e.g., muscle contractions, are appropriate to the environment and behavior

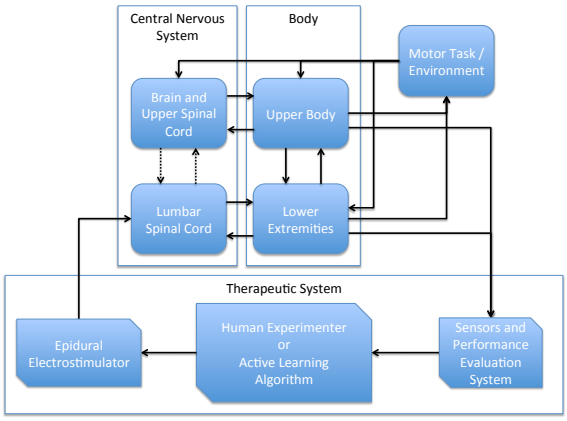

Figure 1.3: A schematic view of SCI therapy through EES. Arrows show the flow of information or action through the composite system. The deficit due to SCI is primarily in terms of disrupted communication between the upper and lower central nervous system (dashed arrows). The intervention employing EES is a modulation of the activity of the lower spinal cord in order to produce better performance in a desired behavior.

of the patient, e.g., weight shifts during standing. Such circuitry does in fact remain intact, if quiescent, below the site of the injury; an autonomously rhythm-generating structure known as a central pattern generator (CPG) is known to exist in a variety of species, including rats and cats, and is thought to exist in humans (Gerasimenko et al., 2008). These neural circuits drive and control complex motor behaviors such as stepping, even in the absence of input from the brain. EES has been applied to stimulate these networks of neurons, enabling stepping and standing after SCI (Harkema et al., 2011; van den Brand et al., 2012). From a control-theoretic perspective, the EES system is not intended to be the controller to the body's plant or process; rather, the EES system modifies the activity of the intrinsic spinal controller or (partially) replaces the absent supraspinal control signal. This scheme is shown in Figure 1.3. In order to make the EES system a higher-level controller for the spinal cord and lower body system, it is necessary to measure the performance of the spinal cord and body (the inner-loop controller) and use these experimental measurements to make decisions about how to change the EES parameters. As the number of electrodes and free parameters increases, however, it is necessary to develop more advanced methods of selecting the stimuli delivered by EES arrays. The motivating problem of this dissertation is optimizing the stimulus patterns for the complex arrays available now and in the near future (e.g., the Medtronic RestoreAdvanced/Specify 5-6-5 system and the parylene-based microelectrodes of Gad et al., 2013, pictured in Figure 4.1).

1.3 Active Learning

Active learning techniques are much as the name indicates; they are algorithms for actively, rather than passively, attempting to learn about a system. In the traditional formulation, the active learner is interacting with an *oracle*, a system which accepts experimental interrogations, each single one of which is called a *query* and returns observations which correspond to the queries. In this fashion, the learner gradually acquires information about expected the response to any query. The element which makes this interaction active, rather than passive, is that the active learner has choices, most typically the choice of which query to submit to the oracle at each opportunity. The active learner makes these choices on the basis of a model of the oracle, constructed based upon the data. Using the choice of which queries to submit, rather than waiting passively for whatever data happens to arrive, the active learner is thus able to acquire the desired information (or simply more information) about the oracle in fewer observations than a passive learner. The text by Settles (2012) gives a thorough treatment of a variety of active learning techniques.

Active learning can be targeted to particular pieces of information about the oracle. One important example is the problem of Bayesian optimization (BO). An interesting approach to this problem is taken by Hennig and Schuler (2012), who also make interesting points regarding the appropriate algorithm design philosophy for this setting. In this case, the active learner is given a finite (but possibly unknown) budget of queries and is tasked with spending these queries to find the action within the available set which yields the maximum value of the *reward*, a measure of the desirability of the oracle's output. After the learning process is complete, the (apparent) optimal action will be chosen and the algorithm will receive this reward. In order to be effective (i.e., probabilistically obtain a high reward), an algorithm must observe the reward values which would be associated with the queries it has so far submitted and then choose future queries which are likely to decrease its uncertainty about where the optimum lies. Choosing queries on-line, rather than *a priori*, allows the algorithm to target these queries to the regions of the set of possible actions which appear promising on the basis of the data being acquired.

An important and closely related problem is the exploitation-exploration tradeoff. If the algorithm receives reward for each and every single query, rather than having distinct search phase in which no rewards are obtained, followed by an exploit phase (as in the BO setting), it is important to choose these queries not simply to learn about the best rewards which may be obtained, but also to obtain high reward at this very moment. This is particularly appropriate in the SCI therapeutic setting, in which each EES stimulus and each interval of training time is valuable and should be spent intelligently. Algorithms for solving this problem have traditionally be explored under the framework of so-called bandits (Robbins, 1952). These algorithms make sequential decisions by trading off exploitation of actions known to yield high reward action with exploration of novel or

poorly-understood actions. If these competing imperatives are properly balanced, it can sometimes be demonstrated that the algorithm will converge to the optimal action (i.e., the rate of sub-optimal actions approaches zero) with high probability in the limit of infinite time. Recent work in this field has brought bandits and Bayesian optimization together, yielding algorithms which seek to explore and exploit over very large decision sets, using models of the response function (e.g., the GP-UCB algorithm of Srinivas et al., 2010, which uses Gaussian processes to model the reward funciton). This dissertation continues this line of work, developing the novel algorithms GP-BUCB and GP-AUCB.

1.4 Objective Statement: Major Problem

In order to make a practical, fully-implantable system to deliver epidural electrostimulation which is highly effective for SCI therapy, it is necessary to create, implement, and test a class of active learning algorithms which can:

- Exploit the structure of the epidural spinal stimulation problem, i.e., the anatomical and neurophysiological knowledge of the spinal cord and the lower limbs, as well as the capabilities and construction of the stimulating device, to learn the responses of the patient's spinal cord and muscle activity to epidural electrostimulation;
- Use such a model to choose queries or experimental actions in a way which enables the response function(s) to be learned in a query-efficient manner, due to the expense of individual queries and the combinatorially vast extent of the search space (e.g., 10^7 or more possibilities);
- Also use such a model to perform effective therapy for the patient, as measured by metrics of success available on a per-trial basis; and
- Choose actions and accept observations in batches, or with substantial delay between action initiation and the receipt of the experimental results.

The first property is necessary because, in order for the learning process to be both efficient and tractable, and thus to provide an effective therapy, prior information must be combined with measurements taken for this particular individual. This prior information, largely invariant, structural, and qualitative in nature, is the result of many years of neurophysiological studies and clinical experience and represents a tremendous resource for exploitation by an automated agent. Since it is desirable for this agent to accomplish the same (very difficult) tasks as would normally be performed by experienced experimentalists and clinicians, incorporating this prior information is a crucial first step. The "budget" of experiments, constrained principally by the time required to perform the desired measurements, but also the monetary expense of doing so, will often be several orders of magnitude smaller than the number of potential stimuli; thus, stimuli which will likely be ineffective

must be rapidly eliminated from consideration, such that experimental effort is concentrated on stimuli which are more likely to be therapeutically useful. This motivates the second requirement. The third property is required by the fact that the calibration sessions in which the algorithm is run will also constitute a substantial part of the patient's therapy, and indeed, are arguably the most therapeutically useful sessions available due to the (very expensive) presence of highly trained clinicians and therapists. Optimally, all stimuli ever administered (including those delivered by the stimulator as the patient undertakes the tasks of daily living) should be evaluated in terms of their functional performance, such that an algorithm which takes full advantage of this opportunity for experimentation and learning (i.e., is always on) may be preferable. If the algorithm operates continuously, it must treat the therapeutic effectiveness of the stimuli delivered as a substantial component of its decision-making if an effective therapy is to be applied. Further, poor stimuli (i.e., those which produce low reward values, indicative of poor therapeutic performance) may produce high fatigue or confound the results of later experiments. Hence, poor stimulus choices destroy much of the utility of the experimental or therapeutic training session. The fourth property is important, and provides the motivation for much of the theory developed in this dissertation, because it allows much greater flexibility in applications; the requirement of algorithms like GP-UCB that all observations be available before the next action can be selected, and thus that only one action can be pending at any time can prove to be a substantial encumbrance. In the SCI therapy setting, the data processed into the performance metric used by the algorithm are often complex and time-consuming to calculate, resulting in substantial delays between the performance of an experiment and the availability of the assessed performance on that experiment. Motion capture, for example, may take extensive hand annotation to analyze fully, and multi-channel EMG may take several minutes to process into a useful form. However, it is most efficient to assemble an experimental session which consists of an unbroken sequence of requested stimuli; this necessitates either a batch procedure or delayed selection of stimuli.

Further, it is highly desirable that an active learning system have the following additional properties:

- It has rigorous guarantees of behavior, at least under some conditions;
- It is sufficiently modular to enable adaptation to different experimental conditions, e.g. by the revision of structural assumptions, the inclusion or exclusion of stimuli within the decision set, and possibly even modification of the decision rule;
- The predictions made by and the assumptions encoded within the algorithm are human-interpretable; and
- The computations comprising the modeling and action selection steps of the algorithm should

be as efficient as possible, with an eye toward deployment on systems with limited computational power, i.e., miniaturized fully implantable devices.

These secondary specifications also describe important capabilities. Guarantees of performance are an important requirement, as the algorithm's practical performance may be easier to understand in light of these guarantees. Modularity is highly desirable because various components can be interchanged to suit the particular problem at hand. From a practical perspective, modularity is also useful because it enables the re-use of computer code between similar experiments, as well as potentially allowing rigorous comparisons of different modules, e.g., model selection on the Gaussian process kernel functions. The third desire, human-interpretable predictions, is important for both contributing to the body of clinical and neuroscientific literature on the spinal cord, as well as verification of these models by clinical observation and experience. Finally, computational efficiency is important, as the long-range goal of a fully-implantable, autonomous device which administers a dynamic, data-driven therapy requires algorithms which can be run with extremely limited computational resources.

1.5 Contributions

Motivated by the above considerations, this thesis:

- Introduces the GP-BUCB and GP-AUCB algorithms (see Chapter 3), which enable batch or parallel selection of stimuli for epidural electrostimulation, as well as for other general problems;
- Derives theoretical bounds on the regret of a general class of algorithms including the GP-BUCB and GP-AUCB algorithms and also shows that with an easily implemented initialization with a set of finite size, the regret of the GP-BUCB algorithm can be split into two additive terms, such that the time-scaling of the regret is independent of the batch size (also in Chapter 3);
- Successfully tests a derivative of GP-BUCB in a closed-loop SCI therapy therapy setting in a rat SCI model, seeking to maximize a measure of spinal cord interneuronal activity (Chapter 4); and
- Suggests a number of crucial extensions to these algorithms for human studies, as well as further animal experimentation (Chapters 5 and 6).

These contributions demonstrate substantial satisfaction of the major components of the problem specification described in Section 1.4; in particular, the novel algorithms presented here are structured to flexibly incorporate expert knowledge about the structure of the response function over the set of possible stimuli, and were able to elicit a desired motor behavior from four complete

SCI rats in a fashion which, under the chosen reward metric, was at least as effective as the parallel performance of an expert human experimenter. As compared to the existing GP-UCB algorithm (Srinivas et al., 2010), GP-BUCB and GP-AUCB can select batches of experiments, or use knowledge of pending experimental observations to aid in the selection of future experiments. Further, these algorithms make predictions which are readily human-interpretable, are highly modular, and can be computed in closed form, thus requiring (comparatively) little computation.

1.6 Organization

A review of background material relevant to this dissertation, including SCI therapy, Gaussian processes (the major modeling tool used throughout the work), kernel functions (the heart of problem-specific Gaussian process models), and active learning algorithms, follows in Chapter 2. The theoretical properties of GP-BUCB and GP-AUCB are examined in Chapter 3; these results are presented with a series of computational experiments comparing these algorithms with several others. Chapter 4 presents the primary application study of this work, a series of experiments in complete SCI rats. Chapter 5 describes work done toward future experiments in humans, including pilot experiments, along with some suggestions for further improvements. Chapter 6 makes some final conclusions with regard to the present work and also discusses potential extensions of this dissertation's results.