Model-Based Shared Control of Human-Machine Systems with Unknown Dynamics

Doctoral Thesis Proposal

Alexander Broad
Department of Electrical Engineering
and Computer Science
Northwestern University
alex.broad@u.northwestern.edu

December 9, 2017

Contents

1 Introduction

Robot autonomy offers great promise as a tool by which we can enhance, or restore, the natural abilities of a human partner. For example, in the fields of assistive and rehabilitative medicine, devices such as exoskeletons and powered wheelchairs can be used to assist a human who has severely diminished motor capabilities. However, many assistive devices can be extremely difficult to control. This can be due to the inherent complexity of the system, the required fidelity in the control signal, or the physical limitations of the human partner. We can, therefore, further improve the efficacy of these devices by offloading challenging aspects of the control problem to an autonomous partner. In doing so, the human operator is freed to focus their mental and physical capacities on important high-level tasks like path planning and interaction with the environment. This idea forms the basis of *shared control*, a paradigm that aims to produce joint human-machine systems that are more capable than either the human or machine on their own.

In my proposed thesis, I plan to develop novel methods of shared control for human-machine systems. The underlying motivation of my work is a desire to develop autonomous agents that are capable of assisting human operators to control mechanical systems. These systems can range from relatively safe and simple (e.g., powered wheelchairs) to complex and dangerous (e.g., aircraft). In this work I am particularly inspired by devices that are designed to assist people overcome physical limitations due either to injury or disease. This use case suggests a few key desirable attributes that help guide the development of our shared control framework. For example, a growing consensus in the literature [39, 46, 52, 68] suggests that users of assistive devices prefer to retain as much control as possible. Therefore, instead of developing methods that cede full control authority to the autonomous partner (e.g., Level 4 & 5 in SAE's taxonomy of self-driving cars [1]), we focus on methods that can *dynamically adjust the level of assistance* given to a human partner to account for suboptimal user input that may otherwise destabilize, or endanger, the joint system. Additionally, due to the extremely large number of possible human-robot partners, we focus on *platform- and user-independent techniques* that can be used to define a shared control framework for any joint human-machine system.

To achieve the stated goal of assisting human operators to control mechanical systems in a general framework, I plan to research *data-driven approaches to shared control*. Using techniques from machine learning, we can develop methods that are well founded and are viable options even when we do not know the system or control dynamics *a priori*. As suggested above, I will also focus on methods that can be used to dynamically adjust the amount of assistance provided. We can then ensure that the human operator retains a significant portion of control authority. In support of the overarching goal of this work, we must address three main questions which relate to *system modeling*, *autonomous policy generation* and *dynamic shared control*.

The first open question (**RQ1**) that I will address in this work is *how to best model the human and robot system*. The main reason that we are interested in answering this question is to ensure that our shared control framework is applicable to any pair of human and robot. We therefore assume no *a priori* knowledge or model of either partner. These models promise to provide a better understanding of how the dynamic system and user interact with the environment. In this work, we are particularly interested in using the model in a Model Predictive Control [79, 93] framework to generate autonomous policies (see RQ2). Instead of using pre-defined models (which are unlikely to be available for many real-world applications), we use data-driven techniques to learn models of the *human-robot system directly through observation*. Our recent work focused on modeling joint human-machine systems [21, 24] is one of the key insights that motivates this thesis. Specifically, we posit that it is easier to both model and regulate the joint system than it would be to model each system independently. That is, by using the learned representation of the joint system, we can pro-actively regulate the mechanical device to enforce safety and stability constraints. While many assistive devices are considered stable in most configurations (e.g., powered wheelchairs), others can be significantly more complex (e.g., exoskeletons) and can easily enter non-stable configurations. For this reason, our exploration of this space will be based on the development of machine learning techniques that can be used to build generalizable and actionable models of *any* joint human-robot system, regardless of the complexity.

The second open question (**RQ2**) that I will address in this work is *how to best generate polices for a controlled dynamic system*. As described above, our goal is to develop an autonomous agent that can both improve the natural ability of a user to achieve a given task and ensure desirable constraints such as the stability of the joint system. Therefore the autonomous aide must be capable of controlling the mechanical system on it's own. We can then use the policy generated by the autonomous partner to enhance, or compensate, for the natural abilities of a human operator (see RQ3). In this work we explore methods capable of developing efficient control policies. Due to the wide range of human-robot pairings and the requirement of real-time interaction, we focus on methods that are both generalizable and data-efficient. In particular, we plan to use information from the learned dynamics and control models discussed in RQ1 to produce control policies that are capable of operating robots that are defined in continuous, high-dimensional state spaces. However, developing autonomous control policies does not describe how to best allocate control between the human and robot partners.

Therefore, the third and final open question (**RQ3**) we address in this work is *how to best share control between the human and robot*. So far we have described an autonomous agent that can learn to control a dynamic system from observation using a combination of techniques from machine learning and optimal control. In this work, though, we are interested in developing a shared control system that improves the abilities of the human operator while enforcing pertinent constraints such as safety. For this reason, we must develop methods that can dynamically allocate control authority to each partner depending on the skill of a given user and the robot's trust in the operator's abilities. Shared control is a relatively new area of robotics research and is evolving quickly. In the proposed thesis I will focus on methods that encourage the user to command as much control authority as they would like and only provide assistance in times of need. I plan to explore numerous approaches to shared control including ideas such as partitioning the control space into a hierarchy as well as geometric and probabilistic comparisons of the human and autonomy's input. In each case, the shared control algorithm will be designed to enhance the user's abilities while considering constraints such as safety and stability.

In the proposed thesis, I plan to research solutions to the questions presented above (RQ1, RQ2, and RQ3) to address how autonomy can be used to assist humans in operating unknown mechanical systems. I am motivated by the desire to increase human capability while ensuring desirable traits of the joint human-machine system. By combining the relative advantages of the human and robotic partner we can improve the efficacy of the joint system while simultaneously ensuring the safety and stability of the human and robot. This work is the first to consider shared control of mechanical systems with unknown dynamics using data-driven approaches. By drawing on ideas from machine learning and optimal control we can ensure that our methods can be applied to any human-robot system and are valid for wide range of tasks. In particular, through shared

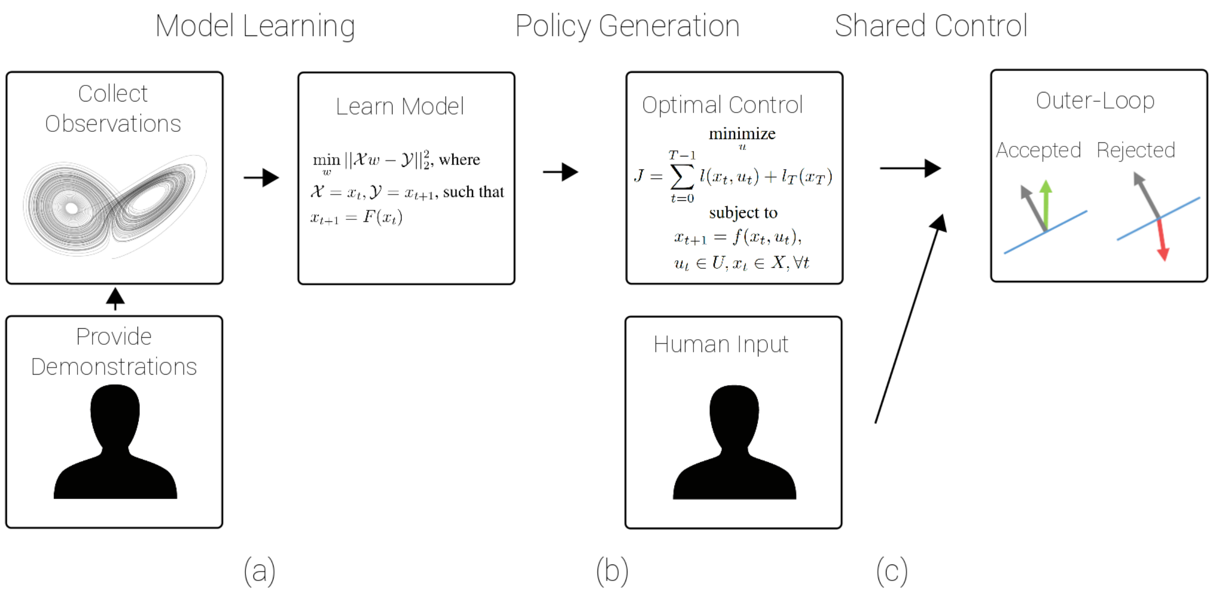


Figure 1: Pictorial depiction of shared control framework defined and evaluated in the proposed thesis work.

control paradigms we can develop autonomous agents that are capable of dynamically assisting a user to control any mechanical system. A pictorial depiction of our defined data-driven shared control framework can be seen in Figure 1.

2 Model Learning from Demonstration Data

In the proposed work, I am interested in studying the shared control of dynamic systems by a human and robot partner. To effectively account for the influence of both the human and robot in the shared control of the system, we must consider how to model each partner.

Computational models of dynamic systems have long been used to evaluate important attributes of mechanical devices, such as the strength of the system and regions of stability [20]. More recently, such models have been used to develop autonomous behaviors for robotic devices [87, 111, 112]. These techniques can be used to generate an optimal control policy for a robot to achieve a specific goal (see RQ2); however, there remain challenges in this area. One important aspect of model-based optimal control is that the optimality (and, more broadly, the success of) the computed policy is highly dependent on the accuracy of the model. Even small inaccuracies in this representation of the system can result in unstable and unsafe control policies. To account for these challenges, I will focus on techniques that aim to reduce errors in the system model by learning the representation directly from observation data. Additionally, to account for potential changes in the underlying model, I plan to explore the use of online learning techniques that can update the learned representation in real-time.

Computational models of human behavior have also proven useful in many areas such as psychology, operations research [38], and human-computer interaction [44]. In this work, we are particularly interested in modeling the specific aspects of human cognition that relate to controlling and interacting with a dynamic system (e.g., communication, decision making, control response to stimuli, etc...); however, even simplified notions of these phenomena can quickly grow past our ability to model and comprehend the original mental process. For that reason, in this work, I will focus on techniques that further reduce the complexity of this modeling problem by incorporating our knowledge of each partner into a single model of the joint human-machine system. The idea that a model of the joint system will be simpler than treating each partner separately may seem counter-intuitive, however, importantly, our approach solves this problem by implicitly enforcing constraints that are inherent to the joint system due to the control interface. That is, we can restrict our modeling technique to *only consider the impact of the human on the dynamic system through a chosen interface.* This approach greatly reduces the complexity of the joint system model by restricting the impact of the user to a set of known observables. Additionally, this choice allows us to effectively regulate the joint system and ensure desirable properties like safety and stability.

A key aspect to this work is the degree to which the described approach is generalizable to any human-robot system. In particular, we use data-driven methods to learn a model of the joint human-machine system and therefore require no *a priori* knowledge of the mechanical device or human partner. Instead, we collect demonstrations of the human and robot interacting and then train a joint model based on the observations. Additionally, due to the constraints of real-time interaction, we focus on methods that can learn quickly and scale easily to high dimensional systems.

2.1 Related Work

Our approach to modeling the human and robot partners is to use data-driven methods to learn a single, joint model that describes how both partners interact with, and control a specific dynamic system. There are numerous viable approaches to learning general models of dynamic systems for autonomous control. In this field, Nguyen et al. [87] make a distinction between direct and indirect model learning methods, where direct refers to the process of learning a mapping when one has access to the system state [76], whereas indirect refers to the process of learning the system model when one only has access to an error measure (e.g., the feedback error [57, 85]). Researchers have explored single-step prediction model, which predict only a single step into the future [62], as well as multi-step prediction models, which predict a sequence of steps into the future [43]. From a methodological standpoint, Neural Networks [112], Gaussian Processes [88], and Gaussian Mixture Models [58] have all shown great promise in this area. There are, of course, practical differences between the various aforementioned modeling and function approximation techniques. Gaussian Mixture Models learn local approximations to system dynamics over various regions of the state space and can therefore fail to generalize to portions of the state space that are not covered by the mixture. Gaussian Processes are similar to Gaussian Mixture Models, but instead of learning a distribution over the state space, they learn a distribution over functions of the state space. This approach is very data efficient, but scales poorly in large data regimes. Neural networks are able to learn models using complex loss functions, and while they can perform poorly in low data regimes, their performance improves consistently with larger amounts of training data. A survey of learning models for control in robotics can be found in [87] and additional information on model-based control can be found in [99].

In the proposed work, I will focus specifically on learning models of *joint human-robot systems*. That is, unlike the previously cited work, we aim to learn a single model that describes the human and machine. In particular, I plan to learn this model through the use of data-driven approaches which can produce a finite approximation to the Koopman Operator (K) [113].

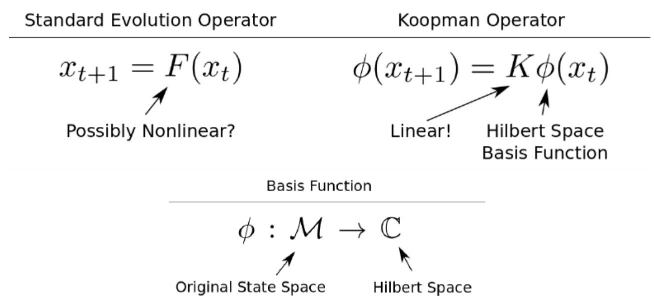


Figure 2: Pictorial depiction of how the *linear* Koopman operator can be used to model the dynamics of *nonlinear* systems.

The Koopman operator is an infinite-dimensional linear operator that can capture all relevant information about the evolution of nonlinear dynamical systems. This is possible because the operator describes a linear mapping between sequential *functions of states* instead of the state itself (see Figure 2). In particular, the Koopman operator acts on an infinite dimensional Hilbert space representation of the state. To define the Koopman operator, let us consider a discrete time dynamic system $(\mathcal{X}, t, F)$:

$$x_{t+1} = F(x_t) \tag{1}$$

where $\mathcal{X} \subseteq \mathbb{R}^N$ is the state space, $t \in \mathbb{R}$ is time and $F : \mathcal{X} \rightarrow \mathcal{X}$ is the state evolution operator. We also define ϕ, a nominally infinite dimensional observation function

$$y_t = \phi(x_t) \tag{2}$$

where $\phi : \mathcal{X} \rightarrow \mathbb{C}$ defines the transformation from the original state space into the Hilbert space representation that the Koopman operator acts on. The Koopman operator K is defined as the composition of ϕ with F, such that

$$K\phi = \phi \circ F. \tag{3}$$

By acting on the Hilbert state representation, the *linear* Koopman operator is able to capture the complex, nonlinear dynamics described by the evolution operator.

The Koopman operator has been studied significantly less then other model learning techniques, however, recent work has demonstrated it's efficacy in modeling a wide range of dynamic systems included uncontrolled systems [108, 28], controlled systems [6], and shared control systems [24]. Exactly how the Koopman operator compares to other model learning techniques is still an open question. However, we do note, that it demonstrates numerous desirable properties. For example, the Koopman operator scales well to high-dimensional systems because the computational complexity of the model grows with the size of the basis function and not the number of data points (as is the case with Gaussian Processes [6, 87]). For this same reason, the training process is fast and data-efficient, as is evident from it's efficacy in online learning paradigms [21]. Additionally, the Koopman operator has also been demonstrated to appropriately model non-smooth dynamics such as contact dynamics [6]. Finally, a benefit of learning a model of the joint human-machine system from user demonstration is the fact that experimental evidence suggests that common data-augmentation techniques, such as persistent excitation [87, 86], are not necessary. We believe this is due to the fact that, during demonstration, the user control provides a wide enough range of inputs to active the important parts of the dynamic spectrum for the learning process. These properties are important when modeling system and control dynamics for the shared control of mechanical systems as the autonomous system must be capable of interacting with the human operator in real-time.

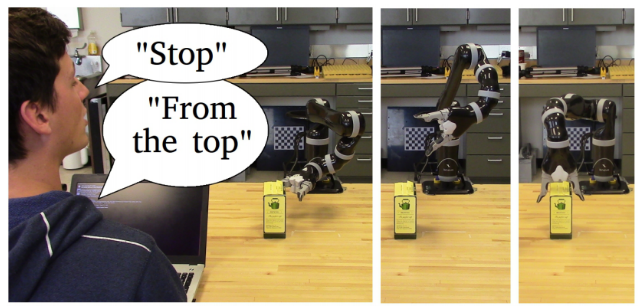
2.2 Contributions

My prior work in the field of machine learning has focused on developing data-driven models for areas as varied as robotic perception [I,II], natural language understanding [III] and a robot's trust in a human partner [IV]. In [I], we demonstrated how roboticists can combine techniques from deep learning with known geometric features of the real-world to produce real-time (12 Hz) 3D object detectors. This idea is well suited to robotic applications as, contrary to standard vision-based object detectors which only operate in 2 dimensions, our solution produces 3 dimensional results which can then be incorporated into the robot's model of the world. In [II], we developed a novel neural network architecture to produce a real-time (50 Hz) object detection framework that improves upon standard single-frame detectors by incorporating temporal features extracted from multiple sequential frames in a video sequence. Again, this work is well suited for robotics applications as, unlike standard image-based techniques, we explicitly include temporal information that is available using the vision sensors deployed on robotic systems. While multiple sequential frames are not always available in real-world applications like web-searches, it is fair to assume that this information is available on any robotic device that has a camera.

In [III], we learned a model that describes how human partners can communicate desired corrections to control a robotic manipulator in real-time using natural language. A particular focus of this work is offloading the challenging aspects of the control problem to an autonomous partner, thereby freeing the physical and mental capacities of a human partner to focus on solving higher level problems for the joint system. Our assumption was that, in the majority of scenarios, the human operator does not care how the robot achieves the goal; however, in the trials in which they do care, we provide the user with the ability to provide corrections which are applied in real-time. In this work, we used a probabilistic graphical model (i.e., a Distributed Correspondence Graph) to learn a model that can ground natural language in the robot's environment (see Figure 3).

Our more recent work has focused on applying machine learning techniques specifically to the problem I am interested in studying in this thesis : modeling joint human-machine systems [V, VI]. In our most recent work, we have demonstrated the power of the Koopman operator in learning both linear [24] and nonlinear [21] models of a joint human machine system. In particular, we developed a Koopman operator model-based optimal control algorithm for general, nonlinear systems identification and autonomous policy generation. We also validated a shared control framework that uses the Koopman-based, autonomously generated optimal control policy as an outer-loop stabilization technique to improve the performance of a joint human-machine system. Finally, in this work, we began to explore the individuality of the learned models. That is, given that our models are learned from demonstration data, we investigated how the source of the demonstration data affected defined system. Contrary to our initial expectation, we found that the efficacy of the shared control system proved to be independent of the source of the data.

(a) Initial command.



(b) Corrective instruction.

Figure 3: Pictorial depiction of natural language interface for providing corrections to an assistive robotic manipulator.

I Broad, A., Argall, B. Geometry-Based Region Proposals for Real-Time Robot Detection of Tabletop Objects. AURO (Submitted). 2018. arXiv:1703.04665.

II Broad, A., Lee, T., Jones, M. Recurrent Multi-frame Single Shot Detector for Video Object Detection. CVPR (Submitted). 2018.

III Broad, A., Arkin, J., Ratliff, N., Howard, T., Argall, B. Real-Time Natural Language Corrections for Assistive Robotic Manipulators. International Journal of Robotics Research (IJRR). 2017.

IV Broad, A., Derry, M., Schultz, J., Murphey, T., Argall, B. Trust Adaptation Leads to Lower Control Effort in Shared Control of Crane Automation. Robotics and Automation Letters (RA-L). 2016. Also presented at the Conference on Automation Science and Engineering (CASE). 2016.

V Broad, A., Murphey, T., Argall, B. Learning Models for Shared Control of Human-Machine Systems with Unknown Dynamics. Robotics: Science and Systems (RSS). 2017.

VI Broad, A., Abraham, I. Murphey, T., Argall, B. Koopman-Based Assistance for Dynamically Shared Control of Human-Machine Systems from User Demonstration. IJRR (Submitted). 2018.

3 Policy Generation and Control of Dynamic Systems

The second research question that I will address in my thesis work, is *how to generate autonomous control policies for dynamic systems*. As mentioned in Section 1, our overarching goal is to produce a shared control framework for generic human-machine systems. Under this paradigm, the autonomous system may need to do more or less work depending on the skill of the operator. On one end of this spectrum, the human partner may be highly skilled in controlling the device. In this case, it is our desire to develop an autonomy that will lie mostly dormant and only exert its influence at specific moments to account for suboptimal user input. On the other extreme of this spectrum, the human partner may have significantly limited abilities in controlling the assistive device—this may be due to injury, disease or simply the inherent complexity of the robot. In this case, the autonomous agent must be capable of controlling the entire system on its own. Therefore, to span the entire spectrum of assistance levels, our approach must be capable of developing fully autonomous policies that can be used to control any dynamic system. Additionally, similar to RQ1, a key motivating factor for this work is to develop generalizable techniques that can be applied to *any* human-robot pairing.

To achieve this goal, we plan to explore novel combinations of learned system models (as described in RQ1) with techniques from Optimal Control. In particular, we develop a learning-based optimal control algorithm that embeds the learned system model into a Model Predictive Control (MPC) framework. The idea of merging concepts from these two fields is an open research area that shows great promise [11, 19, 29, 112]. In particular, one can compute autonomous policies for generic dynamic systems by incorporating the learned model into an MPC framework and then solving the resulting online optimization problem. This method is promising in the Human-Robot Interaction domain because it is fast enough to run in real time, an important feature for systems that interact with a human partner.

In the proposed thesis, we plan to extend our work developing Koopman-operator model-based optimal control algorithms [24, 21]. Specifically, we incorporate the model of a joint human-machine system (learned via a finite approximation to the Koopman operator (see RQ1)) with real-time optimal control methods like Linear Quadratic Regulators [64] and Sequential Action Control [9]. Notably for our purposes, this approach is requires no *a priori* knowledge of the system or human operator. It is also extremely data-efficient, works in continuous state spaces, and runs in real-time. For these reasons, it is well suited for application within our target domain which spans numerous robotic devices that can be used to assist human operators.

3.1 Related Work

In this work, I have proposed the use of model-based Optimal Control (OC) as a method of developing control policies using learned dynamics models. This approach has been explored previously, for example Atkeson et al. [13] described a trajectory-optimization based approach to developing control policies for simulated hoppers and walkers. Morimoto et al. [83] also demonstrated the efficacy of using a learned system model in conjunction with Differential Dynamic Programming to produce autonomous control policies for bipedal robots. Abbeel et al. [2] demonstrated that a similar approach could be used to compute policies for complex, aerobatic helicopter flight on real hardware. A description of how demonstration-based training data can be used to learn system models is explored in further detail in [4].

Reinforcement Learning (RL) [56, 100] is an alternative method for learning control policies directly from data. In fact, Sutten et al. [101] describe Reinforcement Learning as *direct adaptive optimal control*, referencing the fact that RL is distinct from OC because computing an optimal policy does not require an online optimization process. Ernst et al. [40] provide an empirical comparison of these two approaches as they apply to nonlinear power systems problems.

As mentioned above, unlike Optimal Control methods, Reinforcement Learning is a data-driven approach to computing control policies. For that reason, RL is well known for it's data-intensive requirements, particularly in high-dimensional and continuous control domains. One method researchers have developed to help reduce the complexity of this problem is termed Learning from Demonstration (LfD). In one flavor of Learning from Demonstration [10], autonomous agents learn control policies that closely mirror example trajectories provided by a human partner. By forcing the learned policy to resemble the demonstrations (e.g., by ensuring that the learned policy distribution is close to the distribution defined by the examples [82, 69]), researchers have been able to develop policies that can successfully achieve specific tasks, even in high-dimensional state spaces (e.g., [70, 71, 89]).

However, forcing the learned policies to stay within a small deviation of the example trajectories can also introduce some negative features. In particular, by explicitly forcing policies to resemble the example data, large parts of the state space can be left unexplored. For example, in LfD it is common to only use data from successful trials. While this approach can produce policies that are capable of succeeding in the task, they may be poor at recovering from mistakes and deviations from the example trajectories. This is know as the *data-mismatch problem* [94]. To solve this problem, Ross et al. [94] proposed DAgger, a human-in-the-loop solution to that aims to reduce data scarcity in unexplored parts of the state space. By tasking humans with labeling important data that would otherwise be left out of the learning process, they are able to improve the generalizability of the learned models. Building on this work, Laskey et al. [65, 66] described an approach that shifts the burden of solving the data mismatch problem from a human supervisor to an outer-loop supervisory model.

In this work, we also make use of demonstration data, however, instead of learning control policies that mimic human (or algorithmic) demonstration, we focus solely on learning a model of the joint human-machine system. Our approach learns a global model of the system and control dynamics from user interaction. This idea is more akin to model-based reinforcement learning, a technique that is similar to the approach we describe in this work and again fits within the Learning from Demonstration literature [10]. In particular, model-based RL is a paradigm that explicitly learns a model of the system dynamics in addition to learning an effective control policy. Early work in model-based reinforcement learning includes the research of Barto et al. [16] and Kaelbling et al. [56]. Atkeson et al. [14] provide a comparison of model-free and model-based RL techniques and demonstrate that model-based RL is more data-efficient, produces better trajectories, and is more robust to goal configuration.

There are of course, challenges in this domain as well. Specifically, global models can be difficult to define appropriately, particularly if there is state-dependent information that is not properly sampled during the learning process. More recent model-based RL, such as PILCO [33], attempts to reduce model bias in a principled manor. Through improved modeling techniques, Tassa et al. [102] and Lillicrap et al. [75] have demonstrated the efficacy of model-based RL even in high-dimensional and continuous state spaces. In other related work, model-based reinforcement learning has even been shown to be successful when the initial model is inaccurate by allowing for continuous updates to the underlying model [5]. Most closely related to the policy generation approach we describe in this thesis is that of Williams et al. [112]. In this work, the authors describe a model-based optimal control algorithm that learns a model from observation data and then incorporates the model into a Model Predictive Control framework. Our work differs from theirs in the choice of modeling technique, the choice of optimal control algorithm, and most importantly, our focus on shared control instead of fully autonomous control.

3.2 Contributions

In prior work, I have explored numerous policy generation techniques. For example, in [I] we mapped corrective language to topologically distinct paths as defined on a Riemannian manifold (see Figure 4). In this work, users provided real-time natural language corrections to the motion of an assistive manipulator. To compute the initial trajectory (as well as the updated, constrained trajectory), we used geometric methods to compute the most direct trajectory that adhered to the desired constraints. These trajectories were computed on a Riemannian manifold to account for obstacles in the environment. In [II], we allowed users to define a desired path and then used a trajectory optimization technique to track the user-defined motion. This approach constrained the user's inputs based on a computational model of the robot's trust in the human's understanding of the system dynamics. The trajectory generation approaches described in [I] and [II] both use an *a priori* model of the system dynamics to develop control policies.

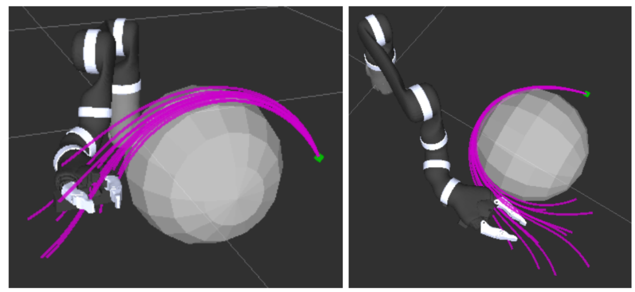


Figure 4: Sample trajectories generated by reimannian motion optimization under various constraints provided by the human operator through natural language.

In our most recent work [III,IV], we begin to explore the idea of using learned models in conjunction with Model Predictive Control algorithms to produce optimal control policies. This idea is important because it allows us to compute autonomous control policies even when we don't have *a priori* knowledge of the system dynamics. In these works, we choose to learn a model of the system and control dynamics through an approximation to the Koopman operator. We then integrate this model into an optimal control algorithm to compute a policy which can provide outer-loop stabilization around the human input. This replaces the policy learning formulation standard in reinforcement learning with techniques from optimal control, similar to [112]. This approach helps produce more generalizable policies that can correct for errors in the model in a structure manor, though the online optimization procedure.

I Broad, A., Arkin, J., Ratliff, N., Howard, T., Argall, B. Real-Time Natural Language Corrections for Assistive Robotic Manipulators. International Journal of Robotics Research (IJRR). 2017.

II Broad, A., Derry, M., Schutlz, J., Murphey, T., Argall, B. Trust Adaptation Leads to Lower Control Effort in Shared Control of Crane Automation. Robotics and Automation Letters

(RA-L). 2016. Also presented at the Conference on Automation Science and Engineering (CASE). 2016.

III Broad, A., Murphey, T., Argall, B. Learning Models for Shared Control of Human-Machine Systems with Unknown Dynamics. Robotics: Science and Systems (RSS). 2017.

IV Broad, A., Abraham, I. Murphey, T., Argall, B. Koopman-Based Assistance for Dynamically Shared Control of Human-Machine Systems from User Demonstration. IJRR (Submitted). 2018.

4 Assistive Robotics through Shared Control

The overarching goal in my proposed thesis is to use ideas from autonomous robots to improve the capabilities of a human operator through the intelligent incorporation of a dynamic system. This goal can be addressed using a variety of mechanical systems to solve numerous problems in different application domains. For example, robot autonomy has been incorporated into assistive and rehabilitative medicine [55], manufacturing [50], construction [18], and aviation [42]. In each of these domains, mechanical devices can be used to extend, enhance or restore the physical abilities of a human operator.

I am especially motivated to by application domains in which the benefit provided to the human partner (through the incorporation of robot autonomy) is particularly meaningful. For example, in assistive and rehabilitative medicine, robotic systems can be developed specifically to enhance the *quality of life* of a human operator. Non-robotic devices have long been employed for just such a purpose. Consider how the progression from a simple cane to an assistive walker to a wheelchair to a powered wheelchair has expanded our ability to replace lost functionality to a wider and wider group of people. In this work, we hope to develop technologies that can be used to further enhance a human's abilities by incorporating robot autonomy into the joint human-machine system. I will particularly focus on approaches that intelligently *share control* between the human and robot partners.

To achieve this goal, we build off the work described in RQ1 (learning models of dynamic systems) and RQ2 (generating optimal control policies). Namely, in the final section of the proposed thesis I will describe various techniques for integrating the learned models and autonomous policies with a human operator to define a shared control system. As mentioned above, we plan to explore shared control methodologies that can enhance the joint system performance by reducing the control burden of the human operator.

The question of how to best share control between a human operator and an autonomous agent depends on a large number of variables that can be challenging to define. For example, the optimal method of sharing control between a human and machine will likely depend on the specific user (e.g., their skill, attention, comprehension), the specific robot (e.g., its capabilities, functionality, trust), and the specific task. We therefore explore numerous techniques and paradigms that alter how to partition the decision space and share control authority (e.g., [39]). In the proposed work, we are particularly interested in developing autonomous systems that are capable of interacting with a human partner in real time and adjusting to dynamic environments. For this reason, we focus on approaches that can *dynamically adjust* how much control authority is granted to each partner online.

In addition to restoring functionality to a human partner, we are also concerned with enforcing desirable constraints like safety and stability [67]. Therefore, our shared control methodology will consider both task performance and system safety when deciding how to dynamically allocate control authority. While many assistive devices are dynamically stable, there invariably remain safety concerns due to dynamic environments (e.g. planning in complex, dynamic environments [105], crowd navigation [106], and general obstacle avoidance [104]). Additionally, more complicated assistive machines (e.g., exoskeletons) are not dynamically stable. Therefore, to ensure that our shared control methodology is valid for all pairs of human and robot partners, we plan to experimentally evaluate our approach using systems with varying levels of inherent stability.

Finally, in this section, we place a particular emphasis on how we evaluate the efficacy of our experimental system. Specifically, we analyze the effectiveness of various shared control methodologies through data collected from human subjects studies. During these studies, we collect information that can be used to empirically validate our defined shared control method. In addition to analyzing metrics that relate directly to the efficacy of our system, we also employ user questionnaires to query the general acceptance of the system. While, some of the defined shared control methods can be examined without a user study, we are particularly interested in the human partner's response to the control paradigm. Our analysis will be based on the results of statistical tests that compare baseline user-only control methods with various shared control paradigms. We expect the analysis of both performance and user response metrics to vary widely between system configurations and even between users [39].

4.1 Related Work

Shared control is an area of research that aims to produce joint human-machine systems that are more capable than either the human or machine on their own. If done intelligently, and with appropriate knowledge of the individual capabilities of each team member, one can improve the overall efficiency, stability and safety of the joint system.

This idea is a promising research area in a number of domains. For example, researchers have explored shared control as it applied to autonomous vehicles such as self-driving cars [31, 96]. Sadigh et al. [96] explore the effect of a human operator's driving style on the behavior of an autonomous driver. De Winter et al. [31] explore the dangers of shared control when conflict of authority arises, demonstrating a precise need for dynamic allocation methods that ensure system safety. Similarly, researchers have explored shared control as it applies to the teleoperation of larger mobile robots and human-machine systems, such as aircraft [78]. Matni et al. [78] describe a shared control system aircraft landing mechanism which is used primary to ensure the safety of the joint system, a common concern when dealing with systems of this size.

Shared control has also been explored as a method of improving joint human-machine performance in manufacturing and construction. For example, using a hierarchical breakdown of the task, Tellex et al [103] describe a natural language interface that allows human operators to impart desired actions for a robotic forklift. Search and Rescue is a particularly active area in shared control research. Nourbakhsh et al. [90] describe how shared control can benefit search and rescue missions at a high-level. Bruemmer et al. [25, 26] describe an empirically designed interface for shared control to accomplish search and detection tasks. They follow this work with an exploration of how and when to allocate control between the human and autonomous partner [27]. In related work, the results of the DARPA robotics challenge demonstrate that some of the most successful teams incorporate the use of a skilled human operator to solve certain challenging problems [12, 32, 54, 116].

There is also prior work on shared control in our motivating domains, assistive and rehabilitation robotics. In these domains, researchers have explored the effect of shared control on the teleoperation of smart wheelchairs [72, 73, 107]. Early work in this field includes the development of NavChair, an off-the-shelf powered wheelchair, outfitted with a computer and some ultrasonic sensors [72], and an autonomous navigating walking cane [7]. More recently, Li et al. [73] have explored how to share control between an autonomous wheelchair and individual human partners. Trieu et al. [107] describe how shared control of wheelchairs can be used to help avoid dynamic obstacles. Researchers have also explored how shared control can be used with robotic manipulators [59]. Kim et al. [59] describe a shared control system in which a human operator provides control input via a Brain-Machine Interface to control a robotic manipulator. Similarly, Jain et al. [51] describe a shared control system in which a human operator provides control input via a Body-Machine Interface. Kofman et al. [60] develop a shared control interface for robotic manipulators based on computer vision. Cipriana et al. [30] describe an EMG-based control interface for shared control of a prosthetic hand. Nudehi et al. [91] describe how haptic shared control can be used during training for minimally invasive surgical procedures.

Independent of the application domain, researchers have also explored numerous methodological choices that describe different shared control paradigms. For example, one can share control by partitioning the task into high-level (e.g., path planning, goal selection, etc...) and low-level (e.g., motor control) requirements and then allocate control authority in line with the abilities of the human and autonomous agent [84]. Another option is to define a shared control paradigm that allows users to operate in the same state space (e.g., input to a joystick) as the autonomy. In this case, one needs to define a blending mechanism that decides when and how to incorporate information from both partners [41]. These paradigms span a spectrum of interaction that ranges from task-level approaches like joint action [48] to low-level, control blending approaches [41]. A survey of control sharing techniques can be found in [84].

The chosen communication interface also has a great affect on the result of the shared control system [84, 39]. When partitioning the shared control problem into sub-tasks (or when the human operator acts in a supervisory roll), common interfaces include graphical user interfaces (GUIs) [27] and language [23, 37, 49, 61, 103]. When the human and robot act in the same control space, common control interfaces include steering wheels [95] and joysticks [39]. The choice of control interface has an effect on the joint system due to it's impact on the controllability of the mechanical device [84]. On one hand, the interface may constrain the ease of control from the human operators perspective either due to a mismatch in the dimensionality of the control interface and the dynamic system [22], or due to reduced motor functionality of the user [39]. It may also impact the manner in which the human an autonomous partners interact with the system if specific actions are more or less difficult for a human operator to perform given the specific interface [22, 45].

In any shared control paradigm, it appears clear that the level of assistance will be greater if the control authority allocated to each partner is not static over the course of the lifetime of the human and robot partners [74]. For that reason, we also make a distinction between approaches that rely on offline computation and pre-define set levels of control allocation [63] and those that operate online to produce dynamic allocation of control [36, 53]. In this thesis, I am particularly interested in developing a methodology that allows us to dynamically adjust the amount of control authority given to the robot and human partners [47].

Improved system stability and safety are not the only desirable traits of shared control systems. Another important aspect of these systems is user acceptance. That is, if the human partner is more accepting of the help provided by the autonomous system, there is a greater likelihood that they will adopt the system into their lives. For this reason, researchers begun evaluating methods for deciding how and went the autonomy should offer help during a joint human-machine task [39, 15].

4.2 Contributions

In prior work, I have explored numerous aspects of shared control. For example, in [I] we developed a planning algorithm that specifically accounts for difficulties that arise in sharing control of dynamic systems due to the user interface. Our approach reduces errors in blending the user and autonomy commands that stem from constraints imposed by the control interface. A standard concept in the shared control literature is the idea of *linear blending* [41]. This approach takes the average of the control input provided by a human operator and the autonomous robot and has been applied in many domains. However, there are known issues with this approach – consider for example a human operating a power wheelchair. In such a scenario there are often two equally valid paths around an obstacle. If the human picks one direction and the autonomy picks the other, a linear blend of their signals will produce an unsafe path that takes the wheelchair directly through the object [105]. Similarly, problems can arise when the control interface changes the relative ease of achieving particular paths. Our methodology therefore develops paths based on a cost function that better matches the human operators, thereby increasing the validity of linear blending.

In other work we explore the idea of developing a hierarchical partitioning of the control requirements for the human and robot partner. For example in [II] we first allow the autonomous robotic system to compute multi-step trajectories. We then allow the human to provide control inputs that signal how fast robot performs each segment and at what point to switch between segments. This work was validated with a one participant with spinal cord injury (SCI) and 4 healthy control participants at the Rehabilitation Institute of Chicago. The users were able to provide control inputs through a series of sensors placed on a vest that they controlled via residual upper body movement.

Similarly, in [III] we allow the robotic device to come up with an initial trajectory under the assumption that the user likely does not care much about how the robot achieves the desired goal. We then allow the human operator to provide real-time corrections to the robot through natural language. This partitioning offloads the challenging aspect of the control problem to the autonomous agent, but allows the human operator to provide high-level constraints to the executed trajectory. These constraints can impart important environmental information (e.g., to stay away from dangerous areas), or simply personal preference.

In [IV] we again develop a shared control methodology that is based on partitioning the control space. In this work, the human operator provides a desired trajectory (for a simulated crane) by simply drawing a two dimensional path. The autonomous agent then provides the low-level control to achieve the defined trajectory, thereby off-loading the challenge of controlling the highly dynamic system to the autonomous agent. In this work, we demonstrated that we can improve the efficacy of the joint system by constraining the user input using a computed notion of the robot's trust in the human operator.

Most recently, in [V,VI] we allow both the human operator and robotic system to provide control inputs in the same input space to move a simulated lunar lander from it's starting point to a desired goal location. In this work, we use utilize Maxwell's Demon Algorithm (MDA) [110] to autonomously partition the control space. MDA is described below in Algorithm 1,

Algorithm 1 Maxwell's Demon Algorithm (MDA)

if $\langle u_h, u_a \rangle \geq 0$ **then**
  $u = u_h$;
else
  $u = 0$;
end if

where u_h is the control input from the human operator, u_a is the control produced by the autonomy, and u is applied to the dynamic system. Depending on the specific criteria we use to decide which user signals are let through to the system, we can provide guarantees on the stability and likelihood of task success. In [V,VI] we restrict the user input to the system to be in the same half-plane as the optimal control solution and place no other limitations on the human-machine interaction. This allows the human operator to retain a great deal of control of the system. In these works we also study the individuality of learned system and control representation.

I Broad, A., Argall, B. Path Planning under Interface-Based Constraints for Assistive Robotics. International Conference on Automated Planning and Scheduling (ICAPS). 2016.

II Jain, S., Farshchiansadegh, A., Broad, A., Abdollahi, F., Mussa-Ivaldi, F., Argall, B. Assistive Robotic Manipulation through Shared Autonomy and a Body-Machine Interface. IEEE International Conference on Rehabilitation Robotics. 2015.

III Broad, A., Arkin, J., Ratliff, N., Howard, T., Argall, B. Real-Time Natural Language Corrections for Assistive Robotic Manipulators. International Journal of Robotics Research (IJRR). 2017.

IV Broad, A., Derry, M., Schutlz, J., Murphey, T., Argall, B. Trust Adaptation Leads to Lower Control Effort in Shared Control of Crane Automation. Robotics and Automation Letters (RA-L). 2016. Also presented at the Conference on Automation Science and Engineering (CASE). 2016.

V Broad, A., Murphey, T., Argall, B. Learning Models for Shared Control of Human-Machine Systems with Unknown Dynamics. Robotics: Science and Systems (RSS). 2017.

VI Broad, A., Abraham, I. Murphey, T., Argall, B. Koopman-Based Assistance for Dynamically Shared Control of Human-Machine Systems from User Demonstration. IJRR (Submitted). 2018.

5 Extensions and Future Work

In the proposed thesis, I plan to continue to study shared control of human-machine systems with unknown dynamics. My high-level goal is to develop technologies that can be used to enhance the efficacy of joint human-robot systems by offloading challenging aspects of control to an autonomous partner. To achieve this goal, we define shared control paradigms that are capable of *dynamically adjusting the amount of assistance* given to a *particular user* in order to achieve a desired task. I will now detail a few studies that I plan to carry out during the course of my thesis work that will continue to investigate this idea.

I am interested in exploring methods that expand the functionality of the Koopman operator model-based shared control algorithm described in [21, 24]. In particular, like many optimal control-based methods, our approach relies on a pre-defined notion of the desired task goal. In this work, I am interested in demonstrating the applicability of koopman-based assistance for shared control to scenarios where the goal is not known *a priori*. To indicate the validity of this claim, I plan to again rely on the human operator to provide demonstrations that illustrate their desired goal. For more information on this study, see Section 5.1.

I am also interested in extending the level of control that the human partner has in the joint system. Under the paradigm defined in our prior work [21, 24] the user control input is checked against the Maxwell's Demon Algorithm (see Algorithm 1) criteria at each instant. This choice was made to give the human operator a large degree of freedom in controlling the dynamic system (because the MDA criteria is quite lenient); however, it does restrict the input to the system to be near the optimal solution at each time-step. We believe we can further extend the freedom of the user's control by only incorporating the control of the autonomous agent when the autonomous agent recognizes that it would not be able to recover given the current trajectory. This concept builds on our work presented in [8] and extends the robot's notion of trust based on how frequently the automation decides that it must assist the human operator. More information on this study is presented in Section 5.2.

I also plan to extend the online learning version of the Koopman operator model-based shared control algorithm described in [21] to define a model learning paradigm that provides explicit safety constraints during the learning process. In particular, we know that there are many real-world problems that are difficult to solve using data-driven methods because there will always be an inherent trade-off between exploring the state space (to improve the model and/or control policy) and ensuring the safety of the robot [17]. This is particularly important in shared control systems when we must consider the safety of both the human and robot. More information on this study is presented in Section 5.3

Finally, I also plan to extend our analysis of the online learning version of our shared control paradigm to compare how human demonstrations relate to strategies that maximize information during the model learning process. This is a collaboration with Todd Murphey's student Ian Abraham. More information on this study is presented in Section 5.4.

5.1 Learning User-Defined Goals from Demonstration under Safe Shared Control of Dynamic Systems

5.1.1 Introduction

This work is motivated by a desire to explore how well the Koopman operator model-based shared control algorithm (described in [21]) generalizes to scenarios in which the task goal is not known *a priori*. I am interested in this idea because, while certain constraints (e.g., stability) may remain desirable across numerous tasks, many other properties of the desired goal will not remain consistent across trials, particularly in real-world scenarios. Therefore, instead of assuming that we have prior knowledge of the human partner's desired goal, we will allow the user to provide demonstrations. The motivation here is similar to that of Inverse Reinforcement Learning (IRL) [3] or Learning from Demonstration (LfD) [98]. While we have previously shown the ability to learn a model of the joint human-machine system from demonstration data, this study will expand the information extracted from the human operator to include the overall goal that they would like to achieve. The idea of providing demonstrations under a shared control paradigm is a novel and unexplored concept. This is particularly challenging in scenarios where the dynamic system is not inherently stable and therefore must exert control effort to remain stable. For this reason, it is not clear how to properly allocate control authority during the task demonstration phase of the learning process. In this work, we plan to explore efficacy of learning user-defined goals from demonstration during stable shared control of the joint system.

5.1.2 Approach

In this work, we do not assume *a priori* knowledge of the cost function, nor do we assume a constant cost function. The second point (a time-varying cost function) is particularly important to our exploration of the concepts discussed in the introduction. Namely, we would like to design a system that allows users to provide demonstrations that represent the goal of their task, while simultaneously enforcing stability constraints. Therefore, during the first phase of the study we define an initial cost function designed only to improve system stability. During this initial phase, users can then provide demonstrations that represent how they would like to solve a given task. Then, over time, we adapt the defined cost function to assist the user in achieving the task they have demonstrated, as well as accounting for system stability. To assist the user in achieving the task they demonstrated, we must also define a metric to describe how closely the autonomously generated trajectories match the user demonstrations.

To achieve this goal, we use an ergodic metric to specify desirable trajectories in space [77, 80, 81]. We choose this metric as it allows us to compare the motion of the dynamic system over time to a distribution defined by the demonstrations provided by a user. The benefit of this type of analysis is that, unlike many Learning from Demonstration techniques which use the final system configuration as the definition of the goal, we can instead incorporate information from the user over the course of the entire trajectory. This allows the user to both provide information about *how* they would like to achieve the goal as well as demonstrate more complex tasks that interact with the environment.

5.1.3 Experimental Design

Similar to our prior work [21, 24], we will begin by learning a model of the joint human-machine system from user demonstration. In this work, the demonstration data can come from *any user*, as our prior work [21, 24] showed that the source of the demonstration data had no effect on the ability of the shared control system to succeed in a task. We then augment our prior experimental design to incorporate information from the user demonstration to define the desired task goal.

To provide continuous assistance to the user in achieving the desired task goal, we then define a dynamic cost function. Initially the cost function is specified to enforce stability constraints during the shared control of the system. As the user provides additional demonstration data, we build a model of the desired motion of the system using an ergodic metric [77, 80, 81]. In particular, we learn the user's goal by defining a distribution in trajectory space. As we collect this data, the defined cost function shifts to consider both the stability of the system and the desired goal as demonstrated by the user.

In this work, the model of the joint human-machine system remains constant throughout the entire experiment. However, the joint system is still individualized to each user as the user can now impart personal control preference during the task demonstration phase. As we collect this additional data, we define a second feature of the cost function that represents the desired trajectory distribution. This information is incorporated into the cost function by a computation of the ergodicity of the current trajectory with respect to the distribution of trajectories provided by user demonstration.

5.1.4 Evaluation

We plan to evaluate this idea with an analysis of the defined methodology for numerous stable goal points in the lunar lander environment (see Figure 5). We will expand this environment to include a large state space to allow users to explore regions that are further from the initial position. We will then provide random goal locations (either completely random locations that are sufficient distinct from the initial configuration, or randomly chosen from a set of pre-specified goals) which the users will be asked to navigate to. While we pre-specify the goals in this work, this is simply a way to ensure that users remember the final configuration of the system and can provide repeated demonstrations. As the user provides demonstrations of how they would like to maneuver to the goal, we will adapt the desired cost function to include this information.

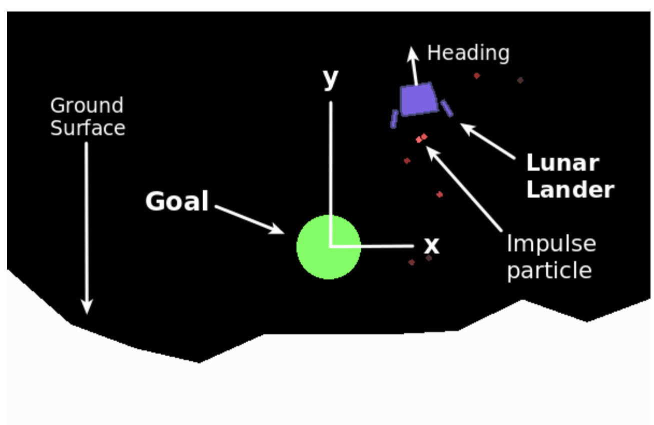


Figure 5: Pictorial depiction of our experimental system—a simulated lunar lander.

To evaluate the efficacy of the system described in this work, we plan to perform a human subjects study similar to the studies performed in [8, 24]. This study will demonstrate (1) the validity of the described approach, and (2) that the assistance is useful in successfully achieving the goal when compared with a user only control paradigm. The analysis will come from a description of how the dynamic cost function operates and also from metrics relevant to how the shift occurs. Some example metrics include—does the system remain stable in all trials? How does the initial stability criteria effect the operator's ability to provide demonstration data? How do the demonstration trajectories evolve over time?

5.2 Robot Trust in a Human Partner as Defined by Frequency of Assistance during Shared Control

5.2.1 Introduction

In this work, we are motived by growing research in the area of human-robot trust. Within the shared-control literature, a subset of works model trust in an effort to improve human-robot team performance [34, 92, 115]. In these works, the formulation of trust represents the trust that the human has *in the autonomous system*, so that the automated system can use the model to choose actions to maximize trust [114]. Additionally, trust has been studied in the context of robot-robot teams [92]. Our proposed system is differentiated in that we are instead proposing a formulation of trust that represents the trust that the autonomous system has *in the operator* [8, 97].

5.2.2 Approach

Prior work related to a robot's trust in the human operator has either relied on a notion of trust that is defined using pre-defined performance parameters [1] or based on a notion of the human's understanding of the system dynamics [22]. In this work, we instead *define trust as being inversely related to the frequency of required assistance by the autonomous agent*. This definition is motivated by the idea that trust grows as a function of time and degrades with evidence that a partner is incapable of successfully achieving the task.

To study this notion of trust, we alter our shared control system so that assistance if provided to the user only when the autonomous controller recognizes that it would be unable to recover from the state and actions of the human partner. In this way, we can ensure that the autonomy only acts when absolutely necessary. The longer the human operator can control the dynamic system on their own, the greater the robot's trust in their abilities.

5.2.3 Experimental Design

From a methodological standpoint, our shared control approach requires two major changes to the techniques presented in [21]. The first change is that we replace Sequential Action Control (SAC) [9] with iterative Sequential Action Control (iSAC) [109], an iterative version of SAC that keeps track of the open-loop control solution and can therefore continue to reduce the associated cost with each iterative step. The second change is that at each time step in the simulation (for a given number of optimization iterations), we check to see if the autonomous controller (iSAC) will be able to recover to a stable point from the given state. If so, we let the user input through to the system; if not, we use the input from the optimal control algorithm to re-stabilize the system. This second change allows the human operator to remain in control for extended periods of time, instead of using the MDA formulation which requires that the operator remain within a window of optimality at each time-step. This shared control paradigm is also inline with growing evidence in the literature that suggests that human operators of shared control systems desire to retain as much control as possible [39, 46, 52, 68].

In this work, we do not learn a unique model of the human-machine system from each individual participant due our recent findings which suggest that the source of the demonstration data does not have a material effect on the suitability of the learned joint system model [21, 24]. Instead, we use a model learned from the interaction of a single user and then use the same model for each of the participants in the experiment. We then individualize the shared control system to a particular task, in which the autonomous agent only provides assistance when the human operator moves into a portion of the state space from which they cannot recover.

5.2.4 Evaluation

To evaluate this concept, we will run a user study using the simulated lunar lander environment depicted in Figure 5. The study will compare our trust-based shared control paradigm with other shared control variants. One reference point will be our prior, MDA-based shared control paradigm [21, 24] that does not explicitly compute a notion of trust. We can also consider comparing this method with other trust-based approaches. For example, in related work Williams et al. [112] demonstrate how sampling based optimal control methods can be used to control dynamic systems with unknown dynamics. Therefore, a related notion of trust can be defined by computing how trust-worthy the user's control signal is based on the mean and variance of the computed autonomous control samples.

To compare our defined shared control technique with other methods, we will begin by evaluating how the trajectories compare under each control paradigm and analyze any resulting differences. We expect that the new shared control paradigm may produce trajectories that act near the "edge" of stability more often then those produced by other methods. We also plan to include a user questionnaire to analyze preference. This questionnaire will investigate whether users prefer additional freedom in controlling the device or a general notion of success and stability. We can also evaluate these different shared control systems using a worst case analysis. Finally, we also plan to highlight particular moments when the system transfers control authority to the autonomous system. We believe this paradigm will produce switching dynamics which are easily interpretable upon inspection. That is, at each point it should be clear to human observers why the autonomous agent took control.

5.3 Online Model Learning for Shared Control with Safety Constraints

5.3.1 Introduction

In this work, we plan to further explore the efficacy of the online Koopman operator model-based assistance discussed in [21] with a specific focus on domains in which system stability is particularly important. Safe exploration of the environment to learn appropriate models is a well known issue in the reinforcement learning community, and remains challenging even in model-based RL [17]. An important distinction between this work and our own is the fact that we consider a shared control system. The motivation for safe exploration and model learning is even more important in these scenarios as there are many cases in which the human and robot are co-located so both partners may be at risk of injury. In shared control systems, the human operator may also be able to enforce safety constraints, however there are many cases in which this is not true due to the complexity of the system or simply due to loss in physical capabilities due to injury or disease. This work is also motivated by complex assistive devices like exoskeletons, which exist in high-dimensional control spaces and can easily enter unstable configurations.

5.3.2 Approach

In this work, we define a shared control system that simultaneously learns a model of the joint human-machine system while it attempts to provide stability constraints. The human operator will begin in full control, during which time we will begin to learn a model of the joint human-machine system. This model can then be used to provide outer-loop stabilization during runtime. That is, while the user provides control inputs to the dynamic system, we collect the data and use it as a demonstration to build a model of the system and control dynamics. Then, similar to the concept described in Section 5.1, we plan to use a dynamic cost function to define the outer loop controller. The outer loop control algorithm will grow proportionally more lenient as our confidence in the learned model grows. We will measure our confidence in the learned system and control dynamics as a function of the variance in the model over time.

5.3.3 Evaluation

To evaluate the efficacy of our approach, we will select a dynamic system that exists in a high-dimensional control space. Importantly, it must still be possible for human operators to provide demonstrations of how to operate the mechanical system. Certain high-dimensional systems are easier for users to provide demonstrations for (e.g., robotic manipulators, as the task space is lower dimensional), while others are more challenging (e.g., humanoids) as it becomes unclear how a user can provide demonstration data if they are unable to successfully control the full system on their own. Here we note that not every user needs to be able to provide demonstration during the model learning process which is important in assistive and rehabilitation robotics as not all users are equally able to provide demonstrations. Additionally, given our motivation domain of assistive and rehabilitative medicine, we believe that the majority of target human-machine systems fall into the first category (easy to provide demonstrations for). For example, exoskeletons [35] are high dimensional systems that are difficult to control, however, by incorporating a human operator into the mechanical system, people are able to provide sufficient control inputs to maneuver the cyber-physical system. We plan to evaluate our shared control system in a higher dimensional state space through the use of more complex dynamic systems.

5.4 Information Maximization during User Exploration of a Dynamic System

This will be joint work with Todd Murphey's student Ian Abraham.

5.4.1 Introduction

In this work, we plan to further evaluate the online Koopman operator model-based assistance discussed in [21]. In particular, we are interested in analyzing *how* users interact with the system and provide demonstrations during the model learning process. The main question we want to answer is, during the learning phase, do users provide input in a principled manner, or do are the inputs essentially drawn from a random distribution? One hypothesis is that users provide inputs that are inline with information maximization techniques. This study will help elucidate how and why the user interacts with the dynamic system during the model learning process. It will also allow us to compare the efficacy of our approach using demonstration data comes from a human partner with data that is produced by another autonomous system in a principled manner.

5.4.2 Approach

Our online Koopman operator model-based shared control algorithm [21] uses a model that is naively initialized (i.e., no prior information about the system) and then updated continuously after every set of observation. This is distinct from standard model-based learning algorithms because of the data that we choose to use to learn the model. Instead of using all of the observations collected from the user demonstrations to learn the model. However, in the online learning paradigm we only update the model when the user input is admitted by the Maxwell's Demon Algorithm (see Algorithm 1) outer-loop controller.

5.4.3 Evaluation

During the study described in [21], we collected a significant amount of data from human's interacting with the lunar lander shared control system. In this work, we plan to further evaluate that data by compring the user inputs with information theoretic approaches. Our hypothesis is that users provide inputs that highly correlate with the control signals that would be generated by an information maximization algorithm. If this is true, it suggests that users act in a rational manner to maximize the autonomous agent's ability to learn the system and control dynamics. If this hypothesis proves false, it suggests that user's may care less about the autonomous system's ability to learn the system dynamics and instead are either (1) continuing to act as if the autonomous agent already knows the correct system and control dynamics or (2) they are simply providing random inputs.

6 Timeline

6.1 Plan for completion of the research

Table 1 shows my plan for completion of the research.

Timeline	Work	Progress
Dec. 2017	IJRR RSS Invitation	submitted
Dec. 2017	Thesis Proposal	ongoing
Jan-March 2018	Section 5.1 Paper Submission	planning
June 2018	Section 5.2 Paper Submission	planning
Sept-Dec 2018	Section 5.3 Paper Submission	planning
Jan. 2019	Section 5.4 Paper Submission	planning
Jan. 2019	Begin thesis writing	
Apr. 2019	Submit draft of thesis	
May. 2019	Revise thesis & defend!	

Table 1: Timeline for the completion of my doctorate and thesis writing/defense.

References

[1] Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. SAE J3016, SAE International, Warrendale, PA, 2016.

[2] Pieter Abbeel, Adam Coates, Morgan Quigley, and Andrew Y Ng. An application of reinforcement learning to aerobatic helicopter flight. In *Advances in Neural Information Processing Systems*, pages 1–8, 2007.

[3] Pieter Abbeel and Andrew Y Ng. Apprenticeship learning via inverse reinforcement learning. In *Proceedings of the International Conference on Machine Learning*, page 1. ACM, 2004.

[4] Pieter Abbeel and Andrew Y Ng. Exploration and apprenticeship learning in reinforcement learning. In *International Conference on Machine Learning*, pages 1–8. ACM, 2005.

[5] Pieter Abbeel, Morgan Quigley, and Andrew Y Ng. Using inaccurate models in reinforcement learning. In *International Conference on Machine Learning*, pages 1–8. ACM, 2006.

[6] Ian Abraham, Gerardo De La Torre, and Todd D Murphey. Model-based control using koopman operators. 2017.

[7] Peter Aigner and Brenan McCarragher. Shared control framework applied to a robotic aid for the blind. In *International Conference on Robotics and Automation*, volume 1, pages 717–722. IEEE, 1998.

[8] Matthew Derry Todd Murphey Alexander Broad, Jarvis Schultz and Brenna Argall. Trust adaptation leads to lower control effort in shared control of crane automation, 2016.

[9] Alexander R Ansari and Todd D Murphey. Sequential action control: Closed-form optimal control for nonlinear and nonsmooth systems. *Transactions on Robotics*, 32(5):1196–1214, 2016.

[10] Brenna D Argall, Sonia Chernova, Manuela Veloso, and Brett Browning. A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, 57(5):469–483, 2009.

[11] Anil Aswani, Humberto Gonzalez, S Shankar Sastry, and Claire Tomlin. Provably safe and robust learning-based model predictive control. *Automatica*, 49(5):1216–1226, 2013.

[12] Christopher G Atkeson, Benzun P Wisely Babu, Nandan Banerjee, Dmitry Berenson, Christoper P Bove, Xiongyi Cui, Mathew DeDonato, Ruixiang Du, Siyuan Feng, Perry Franklin, et al. No falls, no resets: Reliable humanoid behavior in the darpa robotics challenge. In *International Conference on Humanoid Robots*, pages 623–630. IEEE, 2015.

[13] Christopher G Atkeson and Jun Morimoto. Nonparametric representation of policies and value functions: A trajectory-based approach. In *Advances in Neural Information Processing Systems*, pages 1643–1650, 2003.

[14] Christopher G Atkeson and Juan Carlos Santamaria. A comparison of direct and model-based reinforcement learning. In *International Conference on Robotics and Automation*, volume 4, pages 3557–3564. IEEE, 1997.

[15] Jimmy Baraglia, Maya Cakmak, Yukie Nagai, Rajesh PN Rao, and Minoru Asada. Efficient human-robot collaboration: When should a robot take initiative? *The International Journal of Robotics Research*.

[16] Andrew G Barto, Steven J Bradtke, and Satinder P Singh. Learning to act using real-time dynamic programming. *Artificial Intelligence*, 72(1-2):81–138, 1995.

[17] Felix Berkenkamp, Matteo Turchetta, Angela Schoellig, and Andreas Krause. Safe model-based reinforcement learning with stability guarantees. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, pages 908–919. 2017.

[18] Thomas Bock. Construction robotics. *Autonomous Robots*, 22(3):201–209, 2007.

[19] Patrick Bouffard, Anil Aswani, and Claire Tomlin. Learning-based model predictive control on a quadrotor: Onboard implementation and experimental results. In *International Conference on Robotics and Automation*, pages 279–284. IEEE, 2012.

[20] William E Boyce and Richard C DiPrima. *Elementary Differential Equations and Boundary Value Problems*, volume 9. Wiley, 1969.

[21] Alexander Broad, Ian Abraham, Todd Murphey, and Brenna Argall. Koopman-based assistance for dynamically shared control of human-machine systems from user demonstration. In *International Journal of Robotics Research*, 2018.

[22] Alexander Broad and Brenna Argall. Path planning under interface-based constraints for assistive robotics, 2016.

[23] Alexander Broad, Jacob Arkin, Nathan Ratliff, Thomas Howard, and Brenna Argall. Real-time natural language corrections for assistive robotic manipulators. *The International Journal of Robotics Research*, 36(5-7):684–698, 2017.

[24] Alexander Broad, Todd Murphey, and Brenna Argall. Learning models for shared control of human-machine systems with unknown dynamics. In *Robotics: Science and Systems*, 2017.

[25] David J Bruemmer, Ronald L Boring, Douglas A Few, Julie L Marble, and Miles C Walton. " i call shotgun!": An evaluation of mixed-initiative control for novice users of a search and rescue robot. In *International Conference on Systems, Man and Cybernetics*, volume 3, pages 2847–2852. IEEE, 2004.

[26] David J Bruemmer, Donald D Dudenhoeffer, and Julie L Marble. Dynamic-autonomy for urban search and rescue. In *AAAI*, pages 33–37, 2002.

[27] David J Bruemmer, Douglas A Few, Ronald L Boring, Julie L Marble, Miles C Walton, and Curtis W Nielsen. Shared understanding for collaborative control. *IEEE Transactions on Systems, Man, and Cybernetics*, 35(4):494–504, 2005.

[28] Marko Budišić, Ryan Mohr, and Igor Mezić. Applied koopmanism. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 22(4), 2012.

[29] Girish Chowdhary, Maximilian Mühlegg, Jonathan P How, and Florian Holzapfel. Concurrent learning adaptive model predictive control. In *Advances in Aerospace Guidance, Navigation and Control*, pages 29–47. Springer, 2013.

[30] Christian Cipriani, Franco Zaccone, Silvestro Micera, and M Chiara Carrozza. On the shared control of an emg-controlled prosthetic hand: Snalysis of user–prosthesis interaction. *IEEE Transactions on Robotics*, 24(1):170–184, 2008.

[31] Joost CF de Winter and Dimitra Dodou. Preparing drivers for dangerous situations: A critical reflection on continuous shared control. In *International Conference on Systems, Man, and Cybernetics*, pages 1050–1056, 2011.

[32] Mathew DeDonato, Velin Dimitrov, Ruixiang Du, Ryan Giovacchini, Kevin Knoedler, Xianchao Long, Felipe Polido, Michael A Gennert, Taşkın Padır, Siyuan Feng, et al. Human-in-the-loop control of a humanoid robot for disaster response: A report from the darpa robotics challenge trials. *Journal of Field Robotics*, 32(2):275–292, 2015.

[33] Marc Deisenroth and Carl E Rasmussen. Pilco: A model-based and data-efficient approach to policy ssearch. In *Proceedings of the International Conference on Machine Learning (ICML)*, pages 465–472, 2011.

[34] Munjal Desai. *Modeling Trust to Improve Human-robot Interaction*. PhD thesis, University of Massachusetts Lowell, 2012.

[35] Ye Ding, Ignacio Galiana, Alan T Asbeck, Stefano Marco Maria De Rossi, Jaehyun Bae, Thiago Ribeiro Teles Santos, Vanessa Lara de Araujo, Sangjun Lee, Kenneth G Holt, and Conor Walsh. Biomechanical and physiological evaluation of multi-joint assistance with soft exosuits. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(2):119–130, 2017.

[36] Anca Dragan and Siddhartha Srinivasa. A policy blending formalism for shared control. *International Journal of Robotics Research*, 32(7):790–805, May 2013.

[37] Felix Duvallet, Matthew R Walter, Thomas Howard, Sachithra Hemachandra, Jean Oh, Seth Teller, Nicholas Roy, and Anthony Stentz. Inferring maps and behaviors from natural language instructions. In *Experimental Robotics*, pages 373–388. Springer, 2016.

[38] Mica R Endsley. Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37(1):32–64, 1995.

[39] Ahmetcan Erdogan and Brenna Argall. The effect of robotic wheelchair control paradigm and interface on user performance, effort and preference: An experimental assessment. *Robotics and Autonomous Systems*, 94:282–297, 2017.

[40] Damien Ernst, Mevludin Glavic, Florin Capitanescu, and Louis Wehenkel. Reinforcement learning versus model predictive control: A comparison on a power system problem. *IEEE Transactions on Systems, Man, and Cybernetics*, 39(2):517–529, 2009.

[41] C. Ezeh, P. Trautman, L. Devigne, V. Bureau, M. Babel, and T. Carlson. Probabilistic vs linear blending approaches to shared control for wheelchair driving. In *International Conference on Rehabilitation Robotics*, pages 835–840, July 2017.

[42] Eric Frew, Tim McGee, ZuWhan Kim, Xiao Xiao, Stephen Jackson, Michael Morimoto, Sivakumar Rathinam, Jose Padial, and Raja Sengupta. Vision-based road-following using a small autonomous aircraft. In *Aerospace Conference*, volume 5, pages 3006–3015. IEEE, 2004.

[43] Agathe Girard, Carl Edward Rasmussen, Joaquin Quinonero Candela, and Roderick Murray-Smith. Gaussian process priors with uncertain inputs application to multiple-step ahead time series forecasting. In *Advances in Neural Information Processing Systems*, pages 545–552, 2003.

[44] Michael A Goodrich and Alan C Schultz. Human-robot interaction: A survey. *Foundations and Trends in Human-Computer Interaction*, 1(3):203–275, 2007.

[45] Deepak Gopinath and Brenna Argall. Mode switch assistance to maximize human intent disambiguation. In *Robotics: Science and Systems*, 2017.

[46] Deepak Gopinath, Siddarth Jain, and Brenna D Argall. Human-in-the-loop optimization of shared autonomy in assistive robotics. *Robotics and Automation Letters*, 2(1):247–254, 2017.

[47] Thomas Hoeniger. Dynamically shared control in human-robot teams through physical interactions. In *International Conference on Intelligent Robots and Systems*, pages 744–749, 1998.

[48] Guy Hoffman and Cynthia Breazeal. Collaboration in human-robot teams. In *Intelligent Systems Technical Conference*, pages 6434–6452, 2004.

[49] Thomas M Howard, Stefanie Tellex, and Nicholas Roy. A natural language planner interface for mobile manipulators. In *International Conference on Robotics and Automation*, pages 6652–6659. IEEE, 2014.

[50] Jacob Huckaby and Henrik I Christensen. A taxonomic framework for task modeling and knowledge transfer in manufacturing robotics. In *Workshops at 26th AAAI*, 2012.

[51] Siddarth Jain, Ali Farshchiansadegh, Alexander Broad, Farnaz Abdollahi, Ferdinando Mussa-Ivaldi, and Brenna Argall. Assistive robotic manipulation through shared autonomy and a body-machine interface. In *International Conference on Rehabilitation Robotics*, volume 14. IEEE, 2015.

[52] Shervin Javdani, Siddhartha Srinivasa, and J. Andrew Bagnell. Shared autonomy via hindsight optimization. In *Robotics: Science and Systems*, Rome, Italy, 2015.

[53] A. W. Johnson, C. M. Oman, T. B. Sheridan, and K. R. Duda. Dynamic task allocation in operational systems: Issues, gaps, and recommendations. In *2014 IEEE Aerospace Conference*, pages 1–15, March 2014.

[54] Matthew Johnson, Brandon Shrewsbury, Sylvain Bertrand, Tingfan Wu, Daniel Duran, Marshall Floyd, Peter Abeles, Douglas Stephen, Nathan Mertins, Alex Lesman, et al. Team ihmc's lessons learned from the darpa robotics challenge trials. *Journal of Field Robotics*, 32(2):192–208, 2015.

[55] Michelle Johnson, Silvestro Micera, Takanori Shibata, and Eugenio Guglielmelli. Rehabilitation and assistive robotics. *IEEE Robotics and Automation Magazine*, 15(3):16–110, 2008.

[56] Leslie Pack Kaelbling, Michael L Littman, and Andrew W Moore. Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4:237–285, 1996.

[57] Mitsuo Kawato. Feedback-error-learning neural network for supervised motor learning. *Advanced Neural Computers*, 6(3):365–372.

[58] S Mohammad Khansari-Zadeh and Aude Billard. Learning stable nonlinear dynamical systems with gaussian mixture models. *IEEE Transactions on Robotics*, 27(5):943–957, 2011.

[59] Hyun K Kim, J Biggs, W Schloerb, M Carmena, Mikhail A Lebedev, Miguel AL Nicolelis, and Mandayam A Srinivasan. Continuous shared control for stabilizing reaching and grasping with brain-machine interfaces. *IEEE Transactions on Biomedical Engineering*, 53(6):1164–1173, 2006.

[60] Jonathan Kofman, Xianghai Wu, Timothy J Luu, and Siddharth Verma. Teleoperation of a robot manipulator using a vision-based human-robot interface. *IEEE Transactions on Industrial Electronics*, 52(5):1206–1219, 2005.

[61] Thomas Kollar, Stefanie Tellex, Deb Roy, and Nicholas Roy. Toward understanding natural language directions. In *International Conference on Human-Robot Interaction*, pages 259–266. IEEE, 2010.

[62] Marek Kopicki, Jeremy Wyatt, and Rustam Stolkin. Prediction learning in robotic pushing manipulation. In *International Conference on Advanced Robotics*, pages 1–6. IEEE, 2009.

[63] David Kortenkamp, Debra Keirn-Schreckenghost, and R Peter Bonasso. Adjustable control autonomy for manned space flight. In *Aerospace Conference*, volume 7, pages 629–640. IEEE, 2000.

[64] Huibert Kwakernaak. *Linear Optimal Control Systems*, volume 1. Wiley Interscience, 1972.

[65] Michael Laskey, Jonathan Lee, Caleb Chuck, David Gealy, Wesley Hsieh, Florian T Pokorny, Anca D Dragan, and Ken Goldberg. Robot grasping in clutter: Using a hierarchy of supervisors for learning from demonstrations. In *International Conference on Automation Science and Engineering*, pages 827–834. IEEE, 2016.

[66] Michael Laskey, Sam Staszak, Wesley Yu-Shu Hsieh, Jeffrey Mahler, Florian T Pokorny, Anca D Dragan, and Ken Goldberg. Shiv: Reducing supervisor burden in dagger using support vectors for efficient learning from demonstrations in high dimensional state spaces. In *International Conference on Robotics and Automation*, pages 462–469. IEEE, 2016.

[67] Przemyslaw A Lasota, Terrence Fong, Julie A Shah, et al. A survey of methods for safe human-robot interaction. *Foundations and Trends in Robotics*, 5(4):261–349, 2017.

[68] Martin Lawitzky, Melanie Kimmel, Peter Ritzer, and Sandra Hirche. Trajectory generation under the least action principle for physical human-robot cooperation. In *IEEE International Conference on Robotics and Automation*, pages 4285–4290. IEEE, 2013.

[69] Sergey Levine and Pieter Abbeel. Learning neural network policies with guided policy search under unknown dynamics. In *Advances in Neural Information Processing Systems*, pages 1071–1079, 2014.

[70] Sergey Levine, Chelsea Finn, Trevor Darrell, and Pieter Abbeel. End-to-end training of deep visuomotor policies. *Journal of Machine Learning Research*, 17(39):1–40, 2016.

[71] Sergey Levine, Nolan Wagener, and Pieter Abbeel. Learning contact-rich manipulation skills with guided policy search. In *International Conference on Robotics and Automation*, pages 156–163. IEEE, 2015.

[72] Simon P Levine, David A Bell, Lincoln A Jaros, Richard C Simpson, Yoram Koren, and Johann Borenstein. The navchair assistive wheelchair navigation system. *IEEE Transactions on Rehabilitation Engineering*, 7(4):443–451, 1999.

[73] Qinan Li, Weidong Chen, and Jingchuan Wang. Dynamic shared control for human-wheelchair cooperation. In *International Conference on Robotics and Automation*, pages 4278–4283. IEEE, 2011.

[74] Yanan Li, Keng Peng Tee, Wei Liang Chan, Rui Yan, Yuanwei Chua, and Dilip Kumar Limbu. Continuous role adaptation for human–robot shared control. *IEEE Transactions on Robotics*, 31(3):672–681, 2015.

[75] Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. *International Conference on Representation Learning*, 2016.

[76] Lennart Ljung. System identification: Theory for the user. In *Signal Analysis and Prediction*, pages 163–173. Springer, 1998.

[77] George Mathew and Igor Mezić. Metrics for ergodicity and design of ergodic dynamics for multi-agent systems. *Physica D: Nonlinear Phenomena*, 240(4):432–442, 2011.

[78] Nikolai Matni and Meeko Oishi. Reachability-based abstraction for an aircraft landing under shared control. In *American Control Conference*, pages 2278–2284, 2008.

[79] David Q Mayne and Hannah Michalska. Receding horizon control of nonlinear systems. *IEEE Transactions on Automatic Control*, 35(7):814–824, 1990.

[80] Lauren M Miller and Todd D Murphey. Trajectory optimization for continuous ergodic exploration. In *American Control Conference*, pages 4196–4201, 2013.

[81] Lauren M Miller, Yonatan Silverman, Malcolm A MacIver, and Todd D Murphey. Ergodic exploration of distributed information. *IEEE Transactions on Robotics*, 32(1):36–52, 2016.

[82] Igor Mordatch and Emo Todorov. Combining the benefits of function approximation and trajectory optimization. In *Robotics: Science and Systems*, July 2014.

[83] Jun Morimoto, Garth Zeglin, and Christopher G Atkeson. Minimax differential dynamic programming: Application to a biped walking robot. In *International Conference on Intelligent Robots and Systems*, volume 2, pages 1927–1932. IEEE, 2003.

[84] Selma Musić and Sandra Hirche. Control sharing in human-robot team interaction. *Annual Reviews in Control*, 2017.

[85] Jun Nakanishi and Stefan Schaal. Feedback error learning and nonlinear adaptive control. *Neural Networks*, 17(10):1453 – 1465, 2004.

[86] Kumpatis Narendra and Anuradham Annaswamy. A new adaptive law for robust adaptation without persistent excitation. *Transactions on Automatic control*, 32(2):134–145, 1987.

[87] Duy Nguyen-Tuong and Jan Peters. Model learning for robot control: A survey. *Cognitive processing*, 12(4):319–340, 2011.

[88] Duy Nguyen-Tuong, Jan R Peters, and Matthias Seeger. Local gaussian process regression for real time online model learning. In *Advances in Neural Information Processing Systems*, pages 1193–1200, 2009.

[89] Scott Niekum, Sarah Osentoski, George Konidaris, and Andrew G Barto. Learning and generalization of complex tasks from unstructured demonstrations. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5239–5246, 2012.

[90] Illah R Nourbakhsh, Katia Sycara, Mary Koes, Mark Yong, Michael Lewis, and Steve Burion. Human-robot teaming for search and rescue. *Pervasive Computing*, 4(1):72–79, 2005.

[91] Shahin S Nudehi, Ranjan Mukherjee, and Moji Ghodoussi. A shared-control approach to haptic interface design for minimally invasive telesurgical training. *IEEE Transactions on Control Systems Technology*, 13(4):588–592, 2005.

[92] Charles Pippin and Henrik Christensen. Trust modeling in multi-robot patrolling. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2014.

[93] James B Rawlings and Kenneth R Muske. The stability of constrained receding horizon control. *IEEE Transactions on Automatic Control*, 38(10):1512–1516, 1993.

[94] Stéphane Ross, Geoffrey J Gordon, and Drew Bagnell. A reduction of imitation learning and structured prediction to no-regret online learning. In *International Conference on Artificial Intelligence and Statistics*, pages 627–635, 2011.

[95] Dorsa Sadigh and Ashish Kapoor. Safe control under uncertainty with probabilistic signal temporal logic. In *Proceedings of Robotics: Science and Systems*, 2016.

[96] Dorsa Sadigh, S. Shankar Sastry, Sanjit A. Seshia, and Anca D. Dragan. Planning for autonomous cars that leverage effects on human actions. In *Proceedings of Robotics: Science and Systems*, 2016.

[97] H Saeidi, F McLane, B Sadrfaidpour, E Sand, S Fu, J Rodriguez, JR Wagner, and Y Wang. Trust-based mixed-initiative teleoperation of mobile robots. In *American Control Conference*, pages 6177–6182. IEEE, 2016.

[98] Stefan Schaal. Learning from demonstration. In *Advances in Neural Information Processing Systems*, pages 1040–1046, 1997.

[99] Stefan Schaal and Christopher G Atkeson. Learning control in robotics. *IEEE Robotics and Automation Magazine*, 17(2):20–29, 2010.

[100] Richard S Sutton and Andrew G Barto. *Reinforcement Learning: An Introduction*, volume 135. MIT Press, 1998.

[101] Richard S Sutton, Andrew G Barto, and Ronald J Williams. Reinforcement learning is direct adaptive optimal control. *IEEE Control Systems*, 12(2):19–22, 1992.

[102] Yuval Tassa, Tom Erez, and William D Smart. Receding horizon differential dynamic programming. In *Advances in Neural Information Processing Systems*, pages 1465–1472, 2008.

[103] Stefanie Tellex, Thomas Kollar, Steven Dickerson, Matthew R Walter, Ashis Gopal Banerjee, Seth Teller, and Nicholas Roy. Understanding Natural Language Commands for Robotic Navigation and Mobile Manipulation. In *Proceedings of the American Association for the Advancement of Artificial Intelligence (AAAI)*, 2011.

[104] C. Ton, Z. Kan, and S. S. Mehta. Obstacle avoidance control of a human-in-the-loop mobile robot system using harmonic potential fields. *Robotica*, page 121, 2017.

[105] Pete Trautman. Assistive planning in complex, dynamic environments: A probabilistic approach. In *IEEE International Conference on Systems, Man, and Cybernetics*, pages 3072–3078. IEEE, 2015.

[106] Peter Trautman and Andreas Krause. Unfreezing the robot: Navigation in dense, interacting crowds. In *International Conference on Intelligent Robots and Systems (IROS)*, pages 797–803. IEEE, 2010.

[107] Hoang T Trieu, Hung T Nguyen, and Keith Willey. Shared control strategies for obstacle avoidance tasks in an intelligent wheelchair. In *International Conference on Engineering in Medicine and Biology Society*, pages 4254–4257, 2008.

[108] Jonathan H Tu, Clarence W Rowley, Dirk M Luchtenburg, Steven L Brunton, and J Nathan Kutz. On dynamic mode decomposition: Theory and applications. *Journal of Computational Dynamics*, 1:391–421, 2014.

[109] Emmanouil Tzorakoleftherakis and Todd Murphey. Iterative sequential action control for stable, model-based control of nonlinear systems. *arXiv preprint arXiv:1706.08932*, 2017.

[110] Emmanouil Tzorakoleftherakis and Todd D Murphey. Controllers as filters: Noise-driven swing-up control based on maxwells demon. In *IEEE Conference on Decision and Control (CDC)*, 2015.

[111] Xin Wang and Thomas G Dietterich. Model-based policy gradient reinforcement learning. In *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*, pages 776–783, 2003.

[112] Grady Williams, Nolan Wagener, Brian Goldfain, Paul Drews, James M Rehg, Byron Boots, and Evangelos A Theodorou. Information theoretic mpc for model-based reinforcement learning. In *International Conference on Robotics and Automation*, 2017.

[113] Matthew O. Williams, Ioannis G. Kevrekidis, and Clarence W. Rowley. A data–driven approximation of the koopman operator: Extending dynamic mode decomposition. *Journal of Nonlinear Science*, 25(6):1307–1346, 2015.

[114] Anqi Xu and Gregory Dudek. Trust-driven interactive visual navigation for autonomous robots. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2012.

[115] Anqi Xu and Gregory Dudek. OPTIMo: Online probabilistic trust inference model for asymmetric human-robot collaborations. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2015.

[116] Holly A Yanco, Adam Norton, Willard Ober, David Shane, Anna Skinner, and Jack Vice. Analysis of human-robot interaction at the darpa robotics challenge trials. *Journal of Field Robotics*, 32(3):420–444, 2015.