Int. J. Robust and Nonlinear Control, to appear 2004. Special Issue on
"Robust Control Design from Data: Direct and Model Based Approaches."

LEARNING ABOUT DYNAMICAL SYSTEMS VIA UNFALSIFICATION OF HYPOTHESES

Paul B. Brugarolas
Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Dr, Pasadena, CA 91109

Michael G. Safonov
Dept. Electrical Engineering – Systems, University of Southern California
3740 McClintock Ave, Los Angeles, CA 90089-2563

Abstract

This paper examines the problem of learning behaviors of a dynamical system from experimental data via unfalsification of hypotheses within the behavioral approach to system theory of Willems. Behaviors of the dynamic systems are postulated as hypotheses and then tested against experimental data. A simple and concise condition for falsification of hypotheses by experimental data in terms of a kernel is presented. The approach is applicable both to learning models for a plant and to adapting controllers to satisfy performance and robustness goals.

Keywords: control, system identification, behavioral systems, adaptive control, model validation, set membership, falsification, experimental data.

1. INTRODUCTION

A variety of techniques have been proposed for adaptively identifying robust control laws from experimental data [1-5,10-14,17,19-23,29]. There are two main approaches, indirect and direct. In indirect approaches, one attempts to learn the plant and its uncertainty and then use robust control design techniques to synthesize a robust controller. In direct approaches, one tries to learn the robust controllers directly.

In this paper we examine both direct and the indirect approaches to problem of learning from experimental data from the perspective of scientific discovery, where families of alternative hypotheses are conjectured and then tested against experimental data. The key idea is the concept of experimental falsification, which concerns the logic by which evolving experimental data winnows candidate hypotheses. This is related to Popper's "criteria of falsifiability" for determining the scientific status of a theory [18]. A hypothesis could be falsified using experimental data, but it cannot be validated because future data may show that it was false. For this reason, the term unfalsified is used to indicate that a hypothesis has not been falsified by the available experimental data. Using the behavioral approach to dynamical systems [16,25-28], we examine what is experimental data and how to learn the behavior of a dynamical system via falsification of hypotheses. We apply these ideas both to learn models and to adapt controllers to achieve robustness goals (unfalsify adaptive controllers).

In fact, this approach is a generalization of the unfalsified control approach [23] to learning about dynamical systems. Its advantages are its generality, its capability to make maximum use of the data (all available information in the data is fully used), and its assumption free formulation. Moreover, it offers a framework to evaluate other learning methods and to analytically assess the consequences of relaxing assumptions associated with these methods.

In this paper we base our formulation in the behavioral framework of Willems [16,25-28] which provides an elegant framework that permits deep insight into the nature of dynamic systems and their representation. Our results are related to the indirect adaptive control method of Polderman [15] for identifying an unfalsified model for a plant within a prescribed model set. Our approaches differ in that we make no assumptions on the plant, and we do not assume prior knowledge that the true plant lies in a given model set.

This paper is organized as follows. Section 2 presents the basic concepts. It briefly reviews Willems' behavioral approach to system theory, it also gives a definition for experimental data, and it presents a mathematical formulation of the hypothesis unfalsification framework for learning. Section 3 applies this framework to control and system identification problems. Section 4 ties this framework to our previous work in [23]. Sections 5 and 6 present discussion and conclusions respectively. This paper contains and expands the results of [4, 20].

2. BASIC CONCEPTS

2.1 Background on the Behavioral Approach to Mathematical Systems Theory

This section provides an informal review of the basic concepts behind Willems' behavioral approach to mathematical systems theory [16,25-28], which will be used to develop the hypothesis-learning framework.

Consider a phenomenon. Assume that this phenomenon produces outcomes in a set $\boldsymbol{Z}$, called the universum. Consider the subset of the universum made of the possible outcomes of the phenomenon, called the behavior $\boldsymbol{B} \subset \boldsymbol{Z}$. Then a mathematical model of the phenomenon is defined to be the pair $(\boldsymbol{Z}, \boldsymbol{B})$. Mathematical models could be given by multiple representations. Often, behaviors are given as the solution set to behavioral equations. Hence, a behavioral equation representation of the mathematical model $(\boldsymbol{Z}, \boldsymbol{B})$ is $(\boldsymbol{Z}, \boldsymbol{E}, \boldsymbol{f})$, where $\boldsymbol{f} : \boldsymbol{Z} \to \boldsymbol{E}$ where $\boldsymbol{E}$ is the equating space ($\boldsymbol{Null}(\boldsymbol{E}) = 0$), and $\boldsymbol{B} = \{z \in \boldsymbol{Z} | \boldsymbol{f}(z) = 0\} = \ker(\boldsymbol{f})$, that is the kernel of the behavioral equation, not to be confused with the image representation [16] which we do not use in the present paper. This behavioral equation representation and their associated kernels will be the preferred representation in this paper.

In this approach, dynamical systems are viewed as mathematical models. These mathematical models have behaviors that evolve with time and as such their outcomes are signals. These signals live in the signal space $\boldsymbol{Z}$, which plays the role of the universum. Hence $\Sigma = (\boldsymbol{Z}, \boldsymbol{B})$. Behaviors for dynamic systems could be represented in different ways. Often they are represented by differential or difference equations, in which case the behavioral equation is of the form $\boldsymbol{f}(z) = \boldsymbol{R}\left(\frac{d}{dt}\right)z = 0$. This type of representation is called a kernel representation. For example, most traditional input-output and state space representations can be cast in this form [14]. Moreover, interconnections between dynamical systems that live in the same signal space $\boldsymbol{Z}$ are described in this approach as follows. Given two dynamical systems $\Sigma_1 = (\boldsymbol{Z}, \boldsymbol{B}_1)$ and $\Sigma_2 = (\boldsymbol{Z}, \boldsymbol{B}_2)$, then the interconnection of Σ_1 and Σ_2 is the following dynamical system $\Sigma_1 \wedge \Sigma_2 = (\boldsymbol{Z}, \boldsymbol{B}_1 \cap \boldsymbol{B}_2)$. This means that the interconnected system consists of the trajectories consistent with both behaviors, $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$.

2.2 Experimental data

This section explains what is understood by experimental data from a dynamical system and provides a formal definition in the behavioral approach. Experimental data from a system usually consists of command signals to the actuators, and measurement signals from the sensors as shown in Figure 1. For this reason the concept of an extended plant is given as follow.

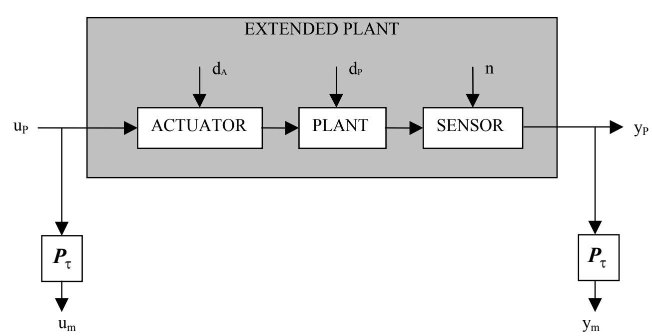

Figure 1: Extended plant

Definition 2.1 *The **extended plant** is defined to contain the plant, sensors and actuators.*

The experimental data is gathered by means of an observation process. This process will be defined using an observation operator, P_τ, which definition follows.

Consider the signal space Z, which contains the extended plant input-output space $U_P \times Y_P \subset Z$.

Definition 2.2 *An **observations operator** P_τ is a mapping $P_\tau : Z \to Z$.*

Examples of observation operators are the time-truncation operator $P_\tau(x) = \begin{cases} x(t) & \text{if} \quad t \le \tau \\ 0 & \text{if} \quad t > \tau \end{cases}$, and the time-sampling operators $P_\tau(x) = \{x_1, x_2, ..., x_\tau\}$.

Then, the measurements are $(u_m, y_m) = P_\tau(u_P, y_P)$, the observed inputs and outputs of the extended plant.

Definition 2.3 *A **data interpolant set** of an extended plant is a subset $D(\tau) \subset Z$ where $D(\tau) = \{z \in Z | P_\tau(u, y) = (u_m, y_m)\}$.*

For example, if the observations operator were the time-sampling operator, then the data interpolant set $D(\tau)$ will consist of all possible signals $z \in Z$ that interpolate the observed samples.

2.3 Learning via Hypotheses Falsification

This section presents the mathematical framework for learning from experimental data via the hypothesis testing principle. It is developed in the behavioral framework presented in the previous section. Experimental data from a system is used to test a hypothesis about the system. If the hypothesis passes the test then we will say that it has been validated or, more precisely, unfalsified. The term unfalsified stresses the tentative nature of experimental validation: future data may later falsify a hypothesis that had been validated by past data.

Consider a signal space Z. Consider a data interpolant set from the extended plant $D \subset Z$. Consider a hypothesis, a mathematical model that defines a relation between signals in Z, which is formally defined as follows.

Definition 2.4 *A **hypothesis** is a dynamical system (Z, E, h) with behavior $H = \{z \in Z | h(z) = 0\} = \ker(h)$ where $h : Z \to E$.*

This hypothesis will be tested against the experimental data by the following condition.

Definition 2.5 *A hypothesis (Z, E, h), is said to be **falsified** by a data interpolant set $D \subset Z$ if, and only if,*

$$\ker(h|D) = \varnothing \tag{1}$$

Where $h|D : D \to E$ denotes restriction of the domain of the behavioral operator h to $D \subset Z$. Otherwise, is said to be unfalsified.

Remarks:

- Condition (1), $\ker(h|D) = \{z \in D | h(z) = 0\} = \varnothing$, means that does not exist any behavior for which the hypothesis is consistent with the data. In some cases, this condition could be written as set intersection $\boldsymbol{H} \cap \boldsymbol{D} = \varnothing$. In practice, this is the case for hypotheses that describe a model or a controller, but not for hypotheses that describe a performance specification because often require a constraint to be met for any possible reference signal, unless we work in a truncated space. For example, a performance specification for a closed loop regulation control system may be $\|\boldsymbol{y} - \boldsymbol{r}\| < \varepsilon, \forall \boldsymbol{r}$.

- Condition (1) differs from $\boldsymbol{D} \subseteq \boldsymbol{H}$ [26], because our definition of the data interpolant set (viz. Def. 2.3) contains all possible extensions of the measurements. Hence, for a hypothesis to be falsified we require that this hypothesis is not consistent with any of the interpolant of the measurements. However, in practice it is useful to restrict the problem to a truncated subspace (truncated by the observations operator) such there is no need to generate all possible extensions of the measurements and simpler conditions could be used. This work was reported in [3].

- Implicit in Definition 2.3, there is a decision rule $\boldsymbol{d} : \boldsymbol{D} \to \boldsymbol{F}$. A mapping from the data to one of two possible outcomes, unfalsified or falsified, viz.

$$\boldsymbol{d}(\boldsymbol{D}) = \begin{cases} \textit{\textbf{unfalsified}}, & \textit{\textbf{iff}} \quad \ker(\boldsymbol{h}|\boldsymbol{D}) \neq \varnothing \\ \textit{\textbf{falsified}}, & \textit{\textbf{otherwise}} \end{cases}$$

- This method could be applied to learning multiple behaviors of a plant simultaneously by combining them in one hypothesis.

- Moreover, it could also be applied to a set of hypotheses. Doing so, it is possible to obtain the subset of hypotheses that are unfalsified by the data. And if in addition, a performance cost, which provides an ordering, is defined then we can learn the optimal hypothesis with respect to that cost.

- Furthermore, it could be applied to hierarchical learning. Hypotheses about hypotheses can be tested sequentially.

3. APPLICATION TO CONTROL AND SYSTEM IDENTIFICATION PROBLEMS

In this section, we apply the learning framework to control and system identification problems. As in [3], we restrict ourselves to a class of problems were the controller is causally-left-invertible (CLI class of controllers) and the performance specification is defined at the observations instances (OI class of specs). This class of problems has the advantage that limits the computational growth. In one hand the causally-left-invertibility property of the controller will guarantee that given an output from the controller, it could only have been generated by a unique controller input signal. On the other hand, the performance specification given at the observations instances allows one to work in a truncated space, as explained in [3]. The restrictions associated with this class of problems are met in most applications.

3.1 Controller Unfalsification Problem

As we said in the introduction, this approach to learning sprang from the unfalsified control concept. For this reason, we will review the key results of the unfalsified control concept [23].

The unfalsified control concept provides a way to evaluate if a controller could meet a desired closed loop performance criterion based on data from the plant, without necessarily having to put the controller in the loop to run tests. The main idea was to create a "conceptual experiment" where a 'fictitious reference signal' was engineered for a controller such if it were used to drive the controller in closed loop with the plant, then it will have produced the measured plant input-output data. Then, the closed loop performance associated with this conceptual experiment was used to evaluate the performance of the controller. As such it is a very powerful tool to screen controllers. Furthermore, it formulated an adaptive control algorithm, where data from a plant with a controller in closed loop was used to evaluate a set of candidate controllers, and if the controller in loop is determined by be falsified then it switched for the 'best' controller from the subset of as yet unfalsified candidate controllers. Recently, related results on this topic have been reported in [13] wherein the fictitious reference signal is called the virtual reference signal. In sections 3.1 and 3.3 we recast some of these results under this new framework.

We consider the case in which the universum $\boldsymbol{Z}$, in addition to containing the extended plant input-output space, also contains the controllers and performance specification input-output spaces. That is: $\boldsymbol{Z} = \boldsymbol{U_P} \times \boldsymbol{Y_P} \times \boldsymbol{U_K} \times \boldsymbol{Y_K} \times \boldsymbol{U_G} \times \boldsymbol{Y_G}$.

Definition 3.3 *A* ***controller*** *is a mathematical model* $(\boldsymbol{Z}, \boldsymbol{E}, \boldsymbol{k})$, *with behavior* $\boldsymbol{K} = \{z \in \boldsymbol{Z} | \boldsymbol{k}(\boldsymbol{u_K}, \boldsymbol{y_K}) = 0\}$.

For example in a typical unity feedback control system the controller input is $\boldsymbol{u_K} = \boldsymbol{r} - \boldsymbol{y_P}$ where r is the reference signal, and the controller output is the input to the plant $\boldsymbol{y_K} = \boldsymbol{u_P}$. Hence $\boldsymbol{U_K} = \boldsymbol{R} \times \boldsymbol{Y_P}$ and $\boldsymbol{Y_K} = \boldsymbol{U_P}$.

Definition 3.4 *A* ***performance specification*** *is a mathematical model* $(\boldsymbol{Z}, \boldsymbol{E}, \boldsymbol{g})$, *with behavior* $\boldsymbol{G} = \{z \in \boldsymbol{Z} | \boldsymbol{g}(\boldsymbol{u_G}, \boldsymbol{y_G}) \le \gamma, \forall \boldsymbol{u_{Gi}} \in \boldsymbol{U_{Gi}}\}$.

Typically $\boldsymbol{g}$ is a cost function, a mapping $\boldsymbol{g} : \boldsymbol{Z} \to \boldsymbol{E}$. A signal $z \in \boldsymbol{G}$ is said to satisfy the performance specification if the cost $\boldsymbol{g}(\boldsymbol{u_G}, \boldsymbol{y_G})$ is smaller or equal than a threshold $\gamma \in \Re$ for all possible values of $\boldsymbol{u_{Gi}}$.

We note that performance specifications like mixed sensitivity or model tracking can be cast in this formulation. Furthermore, in general, inputs to the cost function are $\boldsymbol{u_G} = (\boldsymbol{u_P}, \boldsymbol{y_P}, \boldsymbol{r})$, and the result of evaluating the cost function is $\boldsymbol{y_G} = \gamma$. Note that $\boldsymbol{r}$ plays the role of the $\boldsymbol{u_{Gi}}$, the performance specification should be met for all reference signals. Hence in this example $\boldsymbol{U_G} = \boldsymbol{U_P} \times \boldsymbol{Y_P} \times \boldsymbol{R}$ and $\boldsymbol{Y_G} = \Re$.

We now formalize the controller unfalsification problem defined in [23].

Problem 3.1 ***Controller unfalsification*** *[23]: Given experimental data* $\boldsymbol{D}(\tau)$, *a candidate controller* (Z, E, k) *in the CLI class, and a closed loop performance specification* $(\boldsymbol{Z}, \boldsymbol{E}, \boldsymbol{g})$ *in the OI class, determine whether the controller is unfalsified by* $\boldsymbol{D}(\tau)$, *that is determine if the performance will be violated if the controller were in the loop.*

Note that unfalsification in the definition of the problem is with respect to the control problem. The hypothesis to be tested is if the controller could satisfy the performance specification if it were in closed loop with the plant that generated the data. This hypothesis contains the controller and performance specification behaviors as follows

$$(\boldsymbol{Z},\boldsymbol{E},\boldsymbol{h}) \equiv \left(\boldsymbol{Z},\boldsymbol{E},\begin{bmatrix}\boldsymbol{k}\\ \boldsymbol{g}\end{bmatrix}\right) = (\boldsymbol{Z},\boldsymbol{E},\boldsymbol{k}) \wedge (\boldsymbol{Z},\boldsymbol{E},\boldsymbol{g}) = (\boldsymbol{Z},\boldsymbol{K}\cap\boldsymbol{G})$$

where $\boldsymbol{h} \equiv \begin{bmatrix}\boldsymbol{k}\\ \boldsymbol{g}\end{bmatrix} : \boldsymbol{Z} \to \boldsymbol{E}$ is the behavioral equation and the corresponding behavior is:

$$\boldsymbol{H} = \left\{ z \in \boldsymbol{Z} \middle| \boldsymbol{k}(\boldsymbol{u_K}, \boldsymbol{y_K}) = 0, \boldsymbol{g}(\boldsymbol{u_G}, \boldsymbol{y_G}) \le \gamma, \forall \boldsymbol{u_{Gi}} \in \boldsymbol{U_{Gi}} \right\}.$$

The following proposition casts the solution to this problem by using definition 2.3.

Proposition 3.1 *Given experimental data* $\boldsymbol{D}(\tau)$*, a candidate controller* $(\boldsymbol{Z},\boldsymbol{E},\boldsymbol{k})$ *in the CLI class, and a closed loop performance specification* $(\boldsymbol{Z},\boldsymbol{E},\boldsymbol{g})$ *in the OI class, then the composite hypothesis* $(\boldsymbol{Z},\boldsymbol{E},\boldsymbol{h})$ *is said to be falsified if, and only if,*

$$\ker\left(\boldsymbol{h}\middle|\boldsymbol{D}(\tau)\right) = \varnothing \qquad (2)$$

where $(\boldsymbol{Z},\boldsymbol{E},\boldsymbol{h}) \equiv \left(\boldsymbol{Z},\boldsymbol{E},\begin{bmatrix}\boldsymbol{k}\\ \boldsymbol{g}\end{bmatrix}\right)$ *is the composite mathematical model that contains the hypothesis and performance specification behaviors. Otherwise, is said to be unfalsified.*

Remark: Note that when evaluating composite hypothesis, it is useful to evaluate if the problem is well-posed in the sense that the controller and performance specification as individual hypothesis are unfalsified by the data, and that the composite hypothesis is feasible (i.e., $\ker\left(\boldsymbol{h}\middle|\boldsymbol{D}(\tau)\right) = \varnothing$ *where* $\boldsymbol{h} = \begin{bmatrix}\boldsymbol{k}\\ \boldsymbol{g}\end{bmatrix}$). Some of these concepts were also presented in [3].

3.2 Optimal Unfalsified Control Problem

Now we consider the problem of given a set of candidate controllers, a closed loop performance criteria, and experimental data, find the candidate controller which produces the best performance.

Consider a parameterized set of candidate controllers.

Definition 3.5 *A* ***set of candidate controllers*** *is a set of hypothesis* $(\boldsymbol{Z},\boldsymbol{E},\boldsymbol{k}(\theta))$ *where* $\theta \in \Theta$ *is a parameter vector, such that* $\boldsymbol{K}(\theta) = \left\{ z \in \boldsymbol{Z} \middle| \boldsymbol{k}(\boldsymbol{u_K}, \boldsymbol{y_K}, \theta) = 0 \right\}$.

Problem 3.2 ***Optimal Unfalsified Control:*** *Given experimental data* $\boldsymbol{D}(\tau)$*, a set of candidate controllers* $(\boldsymbol{Z},\boldsymbol{E},\boldsymbol{k}(\theta))$ *in the CLI class where* $\theta \in \Theta$*, a closed loop performance specification* $(\boldsymbol{Z},\boldsymbol{E},\boldsymbol{g})$ *in the OI class. Find the best unfalsified controller.*

The solution to this problem is given in the following proposition.

Proposition 3.2 *Given experimental data* $\boldsymbol{D}(\tau)$*, a set of candidate controller* $(\boldsymbol{Z},\boldsymbol{E},\boldsymbol{k}(\theta))$ *in the CLI class, and a closed loop performance specification* $(\boldsymbol{Z},\boldsymbol{E},\boldsymbol{g}(\gamma))$ *in the OI class, then the optimal unfalsified controller is*

$$\theta* = \arg\min \gamma$$

$$\text{subject to } \ker\left(\boldsymbol{h}(\theta)\middle|\boldsymbol{D}(\tau)\right) \ne \varnothing$$

where $\boldsymbol{h}(\theta) = \begin{bmatrix}\boldsymbol{k}(\theta)\\ \boldsymbol{g}\end{bmatrix}$.

Remark: This optimization problem could have no solution if all the candidate controllers happen to be falsified, or it could have multiple solutions if several unfalsified candidate controllers happen to give the same performance.

3.3 Model Unfalsification Problem

The model validation problem, more precisely unfalsification, consists on verifying that a model fits the data given a certain criteria. As in the controller unfalsification problem, the model and the fitting criteria are mathematical models. To solve the problem, the combined hypothesis is constructed and tested using definition 2.3. In this case, the criteria could be of the prediction error type and in a sense it represents the knowledge on the noise about the system. Hence, we test the hypothesis about the model being able to predict the output of the plant modulo de noise given in the data.

3.4 Unfalsified Parametric System Identification Problem

The parametric system identification problem consists on given a parametric model identifying the parameter set that fits the data for a given fitting criteria. Given a set of parametric models, identification of the best parameters could be achieved by applying the model unfalsification approach to every element of the set as it was shown for the control case. Then, picking the one that gives the smallest cost in the performance specification.

4. APPLICATION TO THE GENERAL PROBLEM

In the previous section we restricted our attention to problems where the controllers were causally-left-invertible (CLI class of controllers) and the performance specification were defined at the observations instances (OI class of specs), which are the more interesting problems from a practical point of view since real problems fall into that category. In this section, we are going to consider the general problem as it was considered in the original unfalsified control formulation [23]. We are going to show that those general results could be explained in this framework.

4.1 Review of unfalsified control definition [23]

This section is a brief summary of the condition on controller unfalsification given in [23].

Let $\boldsymbol{Z} = \boldsymbol{R} \times \boldsymbol{U} \times \boldsymbol{Y}$ be a signal space.

Consider

- the experimental data set $\boldsymbol{P}_{data} = \mathrm{P}_\tau \boldsymbol{P}$, where P_τ is the time truncation operator, and $\boldsymbol{P} = \{(\boldsymbol{r}, \boldsymbol{u}, \boldsymbol{y}) | \boldsymbol{P}_\tau(\boldsymbol{u} - \boldsymbol{u}_{data}) = 0, \boldsymbol{P}_\tau(\boldsymbol{y} - \boldsymbol{y}_{data}) = 0\} \subset \boldsymbol{Z}$ where $(\boldsymbol{u}_{data}, \boldsymbol{y}_{data})$ are measurements from a plant,
- the controller $\boldsymbol{K} = \{(\boldsymbol{r}, \boldsymbol{u}, \boldsymbol{y}) | \boldsymbol{k}(\boldsymbol{r}, \boldsymbol{u}, \boldsymbol{y}) = 0\} \subset \boldsymbol{Z}$; and
- the cost function set $\boldsymbol{T}_{spec} = \{(\boldsymbol{r}, \boldsymbol{u}, \boldsymbol{y}) | \boldsymbol{J}(\boldsymbol{r}, \boldsymbol{u}, \boldsymbol{y}) \geq 0\} \subset \boldsymbol{Z}$.

Consider the unfalsified control problem of determining if the controller ability to meet the specification $(\boldsymbol{r}, \boldsymbol{u}, \boldsymbol{y}) \in \boldsymbol{T}_{spec} \quad \forall \boldsymbol{r} \in \boldsymbol{R}$ is not falsified by the experimental data.

Theorem 4.1 [23, Theorem 1] The control system is unfalsified if, for each triple $(\boldsymbol{r}_0, \boldsymbol{u}_0, \boldsymbol{y}_0) \in \boldsymbol{P}_{data} \cap \boldsymbol{K}$, there exists at least one pair $(\boldsymbol{u}_1, \boldsymbol{y}_1)$ such that $(\boldsymbol{r}_0, \boldsymbol{u}_1, \boldsymbol{y}_1) \in \boldsymbol{P}_{data} \cap \boldsymbol{K} \cap \boldsymbol{T}_{spec}$.

4.2 Formulation in present framework

In this section we are going to show by construction that the present formulation permits a useful alternative description of the results of [23].

In the notation used in this paper $\boldsymbol{P}_{data} \equiv \boldsymbol{D}(\tau)$, the data interpolate set.

The controller is a hypothesis defined by an equation, $\boldsymbol{k}(\boldsymbol{r}, \boldsymbol{u}, \boldsymbol{y}) = 0$, with behavior $\boldsymbol{K} = \{(\boldsymbol{r}, \boldsymbol{u}, \boldsymbol{y}) | \boldsymbol{k}(\boldsymbol{r}, \boldsymbol{u}, \boldsymbol{y}) = 0\} \subset \boldsymbol{Z}$.

The cost function is a hypothesis defined by an inequality, $\boldsymbol{J}(\boldsymbol{r}, \boldsymbol{u}, \boldsymbol{y}) \geq 0$, with behavior $\boldsymbol{T} = \{(\boldsymbol{r}, \boldsymbol{u}, \boldsymbol{y}) | \boldsymbol{J}(\boldsymbol{r}, \boldsymbol{u}, \boldsymbol{y}) \geq 0\} \subset \boldsymbol{Z}$.

The condition for the controller to be consistent with the data interpolant set is equivalent to the condition of the controller hypothesis being unfalsifed. By applying definition 2.5, the controller is unfalsified by the interpolant data set if, and only if, $\ker(\boldsymbol{k} | \boldsymbol{D}) \neq \emptyset$. This means that there exists feasible closed loop systems, that is the controller could have produced the data if it were in closed loop with the plant at the time the data was gathered. In fact, there could be multiple feasible closed loops since the data interpolant set represents all possible interpolation of the plant measurements, that is a set of possible input-output signals and the given controller could be consistent with several of these input-output pairs.

Now, consider each of the feasible closed loops and evaluate if they are consistent with the closed loop cost function. That is, as shown earlier, first construct the composite mathematical model that contains the hypothesis and performance specification behaviors $(\boldsymbol{Z}, \boldsymbol{E}, \boldsymbol{h}) \equiv \left(\boldsymbol{Z}, \boldsymbol{E}, \begin{bmatrix} \boldsymbol{k} \\ \boldsymbol{g} \end{bmatrix} \right)$, and then evaluate it this model is unfalsified by the data, hence the model is unfalsified if, and only if, $\ker(\boldsymbol{h} | \boldsymbol{D}(\tau)) \neq \emptyset$. This corresponds to the controller unfalsification problem given in section 3.1.

Please note that the unfalsified control problem defined in [23], as shown in section 4.1 is more general. It requires that there exist at least one feasible closed loop system for each possible reference signal in the data interpolant set, because the control performance specification is: $\boldsymbol{J}(\boldsymbol{r}, \boldsymbol{u}, \boldsymbol{y}) \geq 0 \quad \forall \boldsymbol{r}$. Note these reference signals have been named fictitious or virtual in the literature since they are obtained from a conceptual experiment.

If we were to consider controllers in the causally-left-invertible class (CLI class) and the performance specification in the observations instances class (OI class), then there is only one feasible closed loop system and therefore a unique fictitious reference signal.

Let's now take the subset of good performance feasible closed loop systems, and evaluate if there exist at least one good performing closed loop system for each fictitious reference signal. If it does, then it is said the controller is unfalsified according to the definition in [23].

To address this problem in the present formulation, we consider the behavior of spanning all fictitious reference signals that lead to a feasible closed loop. Mathematically, $\forall \boldsymbol{r} \in \boldsymbol{P_R}(\boldsymbol{k}(\boldsymbol{r}, \boldsymbol{u}, \boldsymbol{y}) = 0))$ where $\boldsymbol{P_R}$ is a projection operator which extracts the r component of a signal $z=(\boldsymbol{r},\boldsymbol{u},\boldsymbol{y})$, that is $\boldsymbol{P_R}(z)=\boldsymbol{r}$. Allowing for a slight abuse of notation, we may write this as $\boldsymbol{span}(\boldsymbol{P_R}(\boldsymbol{k}(\boldsymbol{r},\boldsymbol{u},\boldsymbol{y})=0))$. This behavior could be added to the behavior of the composite model. That is, $(\boldsymbol{Z},\boldsymbol{E},\boldsymbol{h})\equiv\left(\boldsymbol{Z},\boldsymbol{E},\begin{bmatrix}\boldsymbol{k}(z)=0\\ \boldsymbol{J}(z)\geq 0\\ \boldsymbol{span}(\boldsymbol{P_R}(\boldsymbol{k}(z)=0))\end{bmatrix}\right)$. This augmented or composite hypothesis, which describes the unfalsified control problem, will be unfalsified if, and only if, $\ker(\boldsymbol{h}|\boldsymbol{D}(\tau))\neq\varnothing$. This corresponds to the controller being unfalsified according to the definition in [23].

5. DISCUSSION

In this paper we have taken Willems behavioral perspective to mathematical system theory. This behavioral perspective provided a fertile ground because of its elegance and generality; where the behavior is a simple but powerful concept for describing and analyzing dynamical systems.

In this behavioral setting we have looked at the problem of learning from data. We have studied the problem of learning properties about a dynamical system via falsification of hypotheses against experimental data. The key result is the concept of hypothesis unfalsification. This concept provides a concise and simple definition to learning from experimental data via the kernel of the behavioral operator associated with the hypothesis.

This hypothesis unfalsification concept has been applied to study control and system identification problems. We have formulated families of candidate hypotheses about the behaviors that we wish to learn, models for a plant or controllers for a closed loop system. In doing so, we have shown that complicated learning problems like the unfalsified control problem can be analyzed using a simple kernel condition involving an augmented hypothesis containing all the desired behaviors. This augmentation corresponds to certain interconnections of dynamical systems in the behavioral setting.

The hypothesis unfalsification concept arising in system identification for robust control is similar to that that arises in unfalsified control theory. The theory provides a unifying view to learning from data, including system identification, model validation, direct adaptive control and controller validation. This unified perspective provides a useful framework for comparing and evaluating different learning methodologies.

6. CONCLUSIONS

This paper provides a simple and concise formulation for learning from experimental data via falsification of hypotheses in Willems' behavioral setting. At the core of our formulation is the condition for falsification. This condition states that a property of a dynamical system, conjectured as a hypothesis, is false if the kernel of the hypothesis operator restricted to the experimental data set is empty. This condition is then used to analyze system identification and adaptive control problems via augmentation of hypotheses. This formulation offers a strong theoretical link between Willems' behavioral approach to systems theory and unfalsified control, from which our work is derived. Like unfalsified control, this formulation offers a unifying view to learning from data and provides a framework for comparing and evaluating different data-driven approaches in the system identification and control fields.

7. ACKNOWLEDGEMENTS

Paul B. Brugarolas's research supported in part by JPL.

Michael Safonov's research supported in part by AFOSR Grant F49620-01-1-0302.

8. REFERENCES

[1] T.F. Brozenec and M.G. Safonov, "Controller invalidation," in *Proc. IFAC World Congress*, San Francisco, CA, July 19-23,1996. Oxford, England: Elsevier, pp. 224-233.

[2] P. B. Brugarolas, V. Fromion, and M. G. Safonov "Robust switching missile autopilot," in *Proc. American Control Conference*, Philadelphia, PA, June 22-26, 1998.

[3] P.B. Brugarolas and M.G. Safonov, "A canonical representation for unfalsified control in truncated spaces," in *Proc. IEEE CACSD*, Hawaii, August 22-27, 1999.

[4] P.B. Brugarolas and M.G. Safonov, " A Data-Driven Approach to Learning Dynamical Systems," to appear in *Proc. IEEE CDC* 2002.

[5] F.B. Cabral and M.G. Safonov "A falsification perspective on Model Reference Adaptive Control" in *Proc. IEEE Conf. on Decision and Control*, Kobe, Japan, December 11-13, 1996.

[6] J.D. Glover and F.C. Schweppe, "Control of linear dynamic systems with set constrained disturbances," *IEEE Trans. on Automatic Control*, vol. 16, pp. 411-423, 1971.

[7] J. Helmicki, C. Jacobson, and C. Nett, "Control oriented system identification: a worst-case/deterministic approach in H-infinity," *IEEE Trans. on Automatic Control*, vol. 36, no. 10, 1991.

[8] J. Helmicki, C. Jacobson, and C. Nett, "Worst-case/deterministic identification in H-infinity: the continuous-time case, " *IEEE Trans. on Automatic Control*, vol. 37, no. 5, 1992.

[9] C. Jacobson and C. Nett, "Worst-case system identification in l1: optimal algorithms and error bounds," *System and Control Letters*, no. 19, pp. 419-424, 1992.

[10] R.L. Kosut. "Uncertainty Model Unfalsification for Robust Adaptive Control" in *Proc. IFAC Workshop on Adaptive Systems in Control & Signal P*rocessing, University of Strathclyde, Glasgow, Scotland, UK August 26-28, 1998.

[11] R.L. Kosut, M.K. Lau, and S.P. Boyd, "Set-membership identification of systems with parametric and nonparametric uncertainty," *IEEE Trans. on Automatic Control,* vol. 37, no.7, pp. 929-941, July 1992.

[12] S. Morse, "Supervisory control of families of linear set-point controllers-Part II: Robustness," in *Proc. IEEE Conf. Decision and Control*, New Orleans, LA, Dec. 13-15, 1995. New York, pp. 1750-1755.

[13] E. Mosca and T. Agnoloni, "Inference of candidate loop performance and data filtering for switching supervisory control" in *Automatica* 37 (2001) 527-534.

[14] Narendra, K.S.; Balakrishnan, J, "Adaptive control using multiple models," *IEEE Trans. on Automatic Control*, vol 42, no.2, pp171 -187

[15] J.W. Polderman, "Sequential continuous time adaptive control: A behavioral approach," in *Proc. IEEE Conf. on Decision and Control*, Sydney, Australia, Dec. 12-15, 2000.

[16] J.W. Polderman and J.C. Willems, *Introduction to Mathematical Systems Theory: A Behavioral Approach*. Springer-Verlag New York. 1998.

[17] K. Poolla, P. Khargonekar, A. Tikku, J. Krause and K. Nagpal, "A time-domain approach to model validation," *IEEE Trans. on Automatic Control*, vol. 39, no. 5, pp. 951-959, May 1994.

[18] K.R. Popper. *Logic of Scientific Discovery*. 1934. (English edition: London: Routledge, 1959)

[19] M.G. Safonov. Focusing on the Knowable: Controller Invalidation and Learning, in *Control Using Logic-Based Switching*, A.S. Morse, Ed. Berlin: Springer-Verlag, 1996, pp. 224-233.

[20] M.G. Safonov, "Unfalsified Control: A Behavioral Approach to Learning and Adaptation," in *Proc. IEEE Conf. on Decision and Control*, Orlando, FL, Dec. 4-7, 2001.

[21] M.G. Safonov and F.B. Cabral "Fitting Controllers to Data", *Systems and Control Letters* 43(2001) 299–308.

[22] M.G. Safonov and F.B. Cabral, "Robustness-oriented controller identification", *Proc. IFAC Symp. on Robust Control Design ROCOND'97*, Budapest, Hungary, June 25-27, 1997.

[23] M.G. Safonov and T.-C. Tsao, "The unfalsified control concept and learning," *IEEE Trans. on Automatic Control*, 42(6):843-847, June 1997.

[24] R. S. Smith and M. Dahleh, Eds., The modeling of uncertainty in control system, Proc. of the 1992 Santa Barbara Workshop, Springer-Verlag, New York, 1994.

[25] J. C. Willems, "*The analysis of feedback systems*", MIT Press, Cambridge, MA 1969.

[26] J.C. Willems, "Paradigms and puzzles in the theory of dynamical systems," *IEEE Trans. on Automatic Control*, 36(3): 259-294, March 1991.

[27] J.C. Willems, "On Interconnections, Control, and Feedback," *IEEE Trans. on Automatic Control*, 42(2): 326 - 339, 1997.

[28] J.C. Willems and H.L. Trentelman "The Behavioral Approach to Systems and Control" *CDC 2000* Short Course. Sydney, Australia, Dec. 2000.

[29] B.R. Woodley, R.L. Kosut and J.P. How, "Uncertainty model unfalsification with simulation," in *Proc. American Control Conference*, Philadelphia, PA, June 22-26, 1998.