Automatica 79 (2017) 115–126

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Data-driven and model-based verification via Bayesian identification and reachability analysis[☆]

Sofie Haesaert[a], Paul M.J. Van den Hof[a], Alessandro Abate[b]

[a] Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands
[b] Department of Computer Science, University of Oxford, Oxford, United Kingdom

ARTICLE INFO

Article history:
Received 16 November 2015
Received in revised form
21 August 2016
Accepted 12 December 2016

Keywords:
Temporal logic properties
Bayesian inference
Linear time-invariant models
Model-based verification
Reachability analysis
Data-driven validation
Statistical model checking

ABSTRACT

This work develops a measurement-driven and model-based formal verification approach, applicable to dynamical systems with partly unknown dynamics. We provide a new principled method, grounded on Bayesian inference and on reachability analysis respectively, to compute the confidence that a physical system driven by external inputs and accessed under noisy measurements verifies a given property expressed as a temporal logic formula. A case study discusses the bounded- and unbounded-time safety verification of a partly unknown system, encompassed within a class of linear, time-invariant dynamical models with inputs and output measurements.

1. Introduction

The design of complex, high-tech, safety-critical systems such as autonomous vehicles, intelligent robots, and cyber-physical infrastructures, demands guarantees on their correct and reliable behaviour. Correct functioning and reliability over models of systems can be attained by the use of formal methods. Within the computer sciences, the formal verification of software and hardware has successfully led to industrially relevant and impactful applications (Clarke, 2008). Carrying the promise of a decrease in design faults and implementation errors and of correct-by-design synthesis, the use of formal methods, such as model checking (Clarke, 2008), has become a standard in the avionics, automotive, and railway industries (Vardi, 2009). Life sciences (Belta, Habets, & Kumar, 2002; Del Vecchio & Sontag, 2009) and robotic applications (Belta et al., 2007; Burdick et al., 2007) have also recently benefited by the application of these successful techniques from the computer sciences: this has required a shift from finite-state to physical and cyber-physical models, which are of practical use in nowadays science and technology (Lee, 2008; Tabuada, 2009).

The strength of formal techniques, such as model checking, is bound to the fundamental requirement of having access to a given model, obtained from the knowledge of the behaviour of the underlying system of interest. In practice, for most physical systems the dynamical behaviour is known only in part: this holds in particular with biological systems (Abate, Hillen, & Wahl, 2012) or with classes of engineered systems where, as a consequence, the use of uncertain control models built from data is a common practice (Hjalmarsson, 2005). As an example consider a battery cell to be placed in a car, of which we have only a partial model but know the demand limits that will be raised while in operation. Before installing the battery we can probe and measure its dynamics, and wish to verify that the battery will never heat up excessively under the demanded operational limits.

Only limited work within the formal methods community deals with the verification of models with partly unknown dynamics. Classical results (Batt, Belta, & Weiss, 2007; Henzinger & Wong-Toi, 1996) consider verification problems for non-stochastic models described by differential equations with bounded parametric uncertainty. Similarly, but for continuous-time *probabilistic*

[☆] The research of Sofie Haesaert is supported by DISC (no. 022.002.025) through a personal grant from the NWO graduate program. The material in this paper was partially presented at the 2015 American Control Conference, July 1–3, 2015, Chicago, IL, USA and at the 54th IEEE Conference on Decision and Control, December 15–18, 2015, Osaka, Japan. This paper was recommended for publication in revised form by Associate Editor Antonio Vicino under the direction of Editor Torsten Söderström.

E-mail addresses: s.haesaert@tue.nl (S. Haesaert), P.M.J.Vandenhof@tue.nl (P.M.J. Van den Hof), alessandro.abate@cs.ox.ac.uk (A. Abate).

http://dx.doi.org/10.1016/j.automatica.2017.01.037

models, (Bortolussi & Sanguinetti, 2014; Brim, Češka, Dražan, & Šafránek, 2013) explore the parameter space with the objective of model verification (respectively statistical or probabilistic). Whenever full state measurements of the system are available, Statistical Model Checking (SMC) (Legay, Delahaye, & Bensalem, 2010; Sen, Viswanathan, & Agha, 2004b) replaces numerical model-based procedures with empirical testing of formalised properties. SMC is limited to fully observable stochastic systems with little or no non-determinism, and may require the gathering a large set of measurements. Extensions towards the inclusion of non-determinism have been studied in Henriques, Martins, Zuliani, Platzer, and Clarke (2012) and Legay and Sedwards (2013), with preliminary steps towards Markov decision processes. Related to SMC techniques, but bound to finite state models, Chen and Nielsen (2012), Mao and Jaeger (2012) and Sen, Viswanathan, and Agha (2004a) assume that the system is encompassed by a finite-state Markov chain and efficiently use data to learn the corresponding model and to verify it. Similarly, Bartocci, Bortolussi, and Sanguinetti (2013) and Bortolussi and Sanguinetti (2013) employ machine learning techniques to infer finite-state Markov models from data over given logical formulae.

An alternative approach, allowing both partly unknown dynamics over uncountable (continuous) variables and noisy output measurements, is the usage of a Bayesian framework relating the confidence in a formal property to the uncertainty of a model built from data. When applied on nonlinearly parameterised, linear time invariant (LTI) models this approach introduces heavy computational issues, which can only be mitigated via statistical methods (Gyori, Paulin, & Palaniappan, 2014). Instead, in order to obtain reliable and numerical solutions, we propose the use of linearly parameterised model sets defined through orthonormal basis functions to represent these partially unknown systems. This is a broadly used framework in system identification (Heuberger, Van den Hof, & Wahlberg, 2005; Hjalmarsson, 2005): while maintaining the beneficial computational aspects of linear parameterisations, the choice of orthonormal basis functions allows for the incorporation of prior knowledge on the system behaviour. Practically, this has been widely used for the modelling of physical systems, such as the thermal dynamics of buildings (Virk & Loveday, 1994).

This work investigates the verification of temporal logic properties over partially unknown systems, using both prior modelling knowledge and data drawn from the system in a Bayesian setting. Building on Haesaert, Van den Hof, and Abate (2015a,b), we provide a complete framework and newly extend the modelling class in Haesaert et al. (2015a) to multi-input multi-output models. The focus of this work is further set apart from Haesaert et al. (2015b), which explored the design of experiments to ameliorate the data-driven verification procedure.

2. General framework and problem statement

In this section we overview a new methodology to assess the confidence in whether a system **S** satisfies a given specification ψ, formulated in a suitable temporal logic, by integrating the partial knowledge of the system dynamics with data obtained from a measurement setup around the system.

Let us further clarify this framework. Let us denote with **S** a physical system, or equivalently its associated dynamical behaviour. A signal input $u(t) \in \mathbb{U}, t \in \mathbb{N}$, captures how the environment acts on the system. Similarly, an output signal $y_0(t) \in \mathbb{Y}$ indicates how the system interacts with the environment, or alternatively how the system can be measured. Note that the input and output signals are assumed to take values over continuous domains. The system dynamics can be described via mathematical models, which quantify the behavioural relation between its inputs and outputs. The knowledge of the behaviour of the system is often limited or uncertain, making it impossible to analyse its dynamics by means of a "true" model. In this case, a-priori available knowledge allows to construct a model set $\mathcal{G}$ with elements $\mathbf{M} \in \mathcal{G}$: this model class encompasses the uncertainty on the underlying system by means of a parameterisation $\theta \in \Theta, \mathcal{G} = \{\mathbf{M}(\theta) | \theta \in \Theta\}$. The unknown "true" model $\mathbf{M}(\theta^0)$ representing **S**, is assumed to be an element of $\mathcal{G}$, namely $\theta^0 \in \Theta$. Model sets $\mathcal{G}$ obtained through first principles and with unknown parameters adhere to this standard setup.

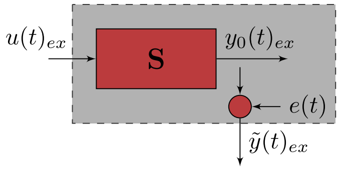

Fig. 1. System (smaller red box) and measurement setup (grey box). In the measurement setup the output $\tilde{y}(t)_{ex}$ includes the system output $y_0(t)_{ex}$ and the measurement noise $e(t)$. Data collected from experiments comprises the input $u(t)_{ex}$ and the measured output $\tilde{y}(t)_{ex}$ signals.

Samples can be drawn from the underlying physical system via a measurement setup, as depicted in Fig. 1. An experiment consists of a finite number (N_s) of input–output samples drawn from the system, and is denoted by $Z^{N_s} = \{u(t)_{ex}, \tilde{y}(t)_{ex}\}_{t=1}^{N_s}$, where $u(t)_{ex} \in \mathbb{U}$ (in general a continuous domain) is the input for the experiment and $\tilde{y}(t)_{ex}$ is a (possibly noisy) measurement of $y_0(t)_{ex}$. In general, the measurement noise can enter non-additively and be a realisation of a stationary stochastic process.[1] We assume that at the beginning of the measurement procedure (say at $t = 0$), the initial condition of the system, encompassed by the initial state $x(0)_{ex}$ of models in $\mathcal{G}$, is either known, or, when not known, has a structured uncertainty distribution that is based on the knowledge of past inputs and/or outputs. As reasonable, we implicitly consider only well-defined problems, such that for any model $\mathbf{M}(\theta)$ representing the system, given an input signal $u(t)_{ex}$ and an (uncertainty distribution for) $x(0)_{ex}$, the probability density distribution of the measured signal can be fully characterised.

The end objective is to analyse the behaviour of system **S**. We consider properties encoded as specifications ψ and expressed in a temporal logic of choice (to be detailed shortly). Let us remark that the behaviour of **S** to be analysed is bound to a set of operating conditions that are pertinent to the verification problem and that will be indexed by "*ver*": this comprises the set of possible input signals $u(t)_{ver}$ (e.g., a white or coloured noise signal, or a non-deterministic signal $u(t)_{ver} \in \mathbb{U}_{ver} \subseteq \mathbb{U}$), and of the set of initial states $x(0)_{ver} \in \mathbb{X}_{ver}$ for the mathematical models $\mathbf{M}(\theta)$ reflecting past inputs and/or outputs of the system. The system satisfies a property if the "true" model representing the system satisfies the property, namely $\mathbf{S} \vDash \psi$ if and only if $\mathbf{M}(\theta^0) \vDash \psi$.

In this work we consider the satisfaction of a property $\mathbf{M}(\theta) \vDash \psi$ as a *binary-valued mapping* from the parameter space Θ. More generally, when in addition to the measurements of the system also its internal transitions are disturbed by stochastic noise (known as process noise), then property satisfaction is a mapping from the parameter space Θ to the interval [0, 1], and quantifies the probability that the model $\mathbf{M}(\theta)$ satisfies the property. This mapping generalises the definition of the satisfaction function discussed in Bortolussi and Sanguinetti (2014), and is now stated as follows.

[1] Notice that the operating conditions of the experiment, that is the input signal $u(t)_{ex}$, the initial state $x(0)_{ex}$, and the measurements $\tilde{y}(t)_{ex}$, have been indexed with "*ex*" to distinguish them from the conditions of interest for verification ("*ver*"), to be discussed shortly.