Contents lists available at ScienceDirect

Computers and Chemical Engineering

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

Process control using finite Markov chains with iterative clustering

Enso Ikonen*, István Selek, Kaddour Najim

Systems Engineering, University of Oulu, PO Box 4300, FIN-90014, Finland

ARTICLE INFO

Article history:
Received 14 January 2016
Received in revised form 9 May 2016
Accepted 13 May 2016
Available online 1 July 2016

Keywords:
Markov decision processes
Bayesian state estimation
Model-based process control
Model predictive control
Clustering
Process simulation

ABSTRACT

Finite state Markov Decision Processes (MDP) for process control are considered. MDP provide robust tools to perform optimization in closed-loop, and their finite state description enables an easy implementation of Bayesian state estimation. An approach to tackle the curse of dimensionality problem, yet retaining the benefits of the finite state MDP in control and estimator design, is proposed. The suggested approach uses iterative re-discretization based on clustering of closed-loop data. An efficient modification of the k-means clustering technique is proposed. The performance of the approach is demonstrated using a challenging benchmark from chemical engineering, the van der Vusse continuous stirred tank reactor control problem. It is shown that the requirements of the benchmark are met, and that the suggested iterated clustering significantly improves the performance. It is concluded that the finite state MDP approach is a viable alternative for small-to-medium scale problems of practical process control and state estimation.

1. Introduction

Finite Markov chains and dynamic programming are powerful tools for design and analysis of dynamic systems. There is a breathtaking literature on finite Markov chains, dynamic programming, and Markov Decision Processes (MDP): (Kemeny and Snell, 1960; Puterman, 1994; Bertsekas, 2005; Poznyak et al., 2000; Hsu, 1987) just to mention a few. MDP have found a lot of applications in the area of robotics, computer sciences and operations research (Powell, 2010; Lee and Lee, 2004, 2006; White, 1989). Markov Transition Models (MTM) have been proposed for process engineering and biochemistry (Tamir, 1998; Pande et al., 2010; Berthiaux and Mizonov, 2004) and process control (see *e.g.* Ikonen and Najim, 2009; Negenborn et al., 2005), but the MDP have largely been considered unpractical by the process control community due to that the computational and storage requirements with almost all problems of practical interest have remained unwieldy (Lee and Lee, 2004). Instead, Model Predictive Control (MPC) has been the main paradigm in advanced process control during the past decades.

Indeed, process control and state estimation problems have characteristics which distinguish them from the decision making problems in many other fields (Lee and Lee, 2004; Ikonen and Najim, 2002):

- multivariate problems with high dimensional state spaces,
- continuous state, action and measurement spaces, with unmeasurable internal states,
- mild non-linearities (smooth or appear as few discontinuities),
- heavily constrained systems,
- long delays,
- availability of rough process models, with a strong physical background,
- expensive experimentation (large-scale systems running in production),
- redundancy in measurements,
- low sampling rates, multiple sampling rates,
- substantial on-site tuning due to uniqueness of local conditions and products, and
- competing sets of performance objectives.

Clearly, this type of characteristics differ from those encountered, *e.g.*, in the field of economics (lack of reliable models), robotics (very precise models are available, trial-and-error learning is feasible to some extent), consumer electronics (mass production of low cost products), telecommunication (extensive use of test signals, fast sampling), or academic toy problems (complex multimodal deterministic test functions).

MPC relies on open loop optimization that is conducted online. The optimization minimizes a cost function defined over a finite future horizon, and results in an optimal sequence of future control actions. Following the receding horizon principle, only the first (immediate) control action from the sequence is applied and

* Corresponding author.
E-mail address: Enso.Ikonen@oulu.fi (E. Ikonen).

http://dx.doi.org/10.1016/j.compchemeng.2016.05.007
0098-1354/© 2016 Elsevier Ltd. All rights reserved.

the optimization is repeated at the next control instant. The overwhelming success of MPC in practical process control can largely be addressed to the possibility to solve dynamic multivariable constrained optimization problems, taking into account disturbances in a feed forward fashion. Adding the possibility to deal with long delays (due to the use of the prediction horizon), and having an option to express the costs in terms of economical quantities, the MPC indeed provides a tempting methodology for dealing with practical problems of process control.

The controlled finite Markov chains (aka Markov Decision Processes) are close relatives to MPC. Both MDP and MPC are based on numerical optimization, and their success relies fundamentally on the availability of good process models. It is well known that the MDP suffer from curse-of-dimensionality (Powell, 2010). For a vast majority of problems, the state of a system is a vector, and the size of the state space grows very quickly as the number of dimensions grow. For example, if each continuous process state variable x_i of a state vector $\mathbf{x} = [x_i] \in \Re^{n_x}$ would be approximated by Z different values, the number of state cells S in a finite state system would grow exponentially with n_x, $S = Z^{n_x}$. However, MDP also provide remedies for the fundamental problems of MPC (Lee and Wong, 2010). In particular, the optimization in MDP concerns closed-loop performance, and can naturally take into account stochastic uncertainties in various system components. In addition, the design is conducted off-line, no on-line optimization of control sequences is required.

Some of the above-mentioned characteristics of process control problems favor the use of Markov chains, while others do not. The size of the finite state and action spaces is a good measure for the main limiting constraints, the memory requirements and computational complexity. The need to work with multivariate problems in high dimensional state spaces is certainly discouraging; the requirement to deal with continuous spaces seems impossible with MDP unless one turns into approximate dynamic programming. The cost of real-life experimentation rules out any large-scale use of machine learning techniques for learning the required mappings from data. On the other hand, constraints can easily be dealt with via penalization and long delays and other complicating dynamics present *a priori* no problem as they can be included in the plant models. Availability of plant models for many industrial processes of practical significance is a strong argument favoring the use of Markov chains, as these techniques enable to take into account not only plant state and measurement noises but also modeling errors. Redundancy in measurements, again, favors the use of techniques taking the uncertainties seriously. The possibility to use Bayesian state estimation makes a maximal benefit from redundant measurements, and an exact implementation of Bayesian filters is possible with finite state descriptions (Ungarala et al., 2006; Kaelbling et al., 1996). As most of the computations are conducted off-line, fast sampling rates can be considered.

Approximate dynamic programming (ADP) (see Lee and Lee, 2004, 2006; Lee and Wong, 2010; Powell, 2010, 2009; Selek et al., 2013 and references therein) is a loose group of techniques pursuing large-scale applications. The basic idea is to rely on the dynamic programming, but to avoid excessive simulations by using sample paths and forward dynamic programming. Parametrized or nonparametric function approximations are used to directly approximate the cost-to-go's at states not evaluated by simulation. In fact, one can build function approximation from continuous domains and avoid the discretization and explicit building of the MTM transition probabilities altogether. At first sight, this looks very tempting. However, a significant fraction of the power of the MDP methodology is lost, which can be addressed to the lack of the transition probability matrices. The finite state Markov chain representation provides a powerful model of propagation of the system uncertainties, as well as means for analysis of the nonlinear dynamics of both open and closed-loop systems. In many ADP approaches, learning by simple counting is traded to function approximation which easily makes model identification and validation cumbersome. In addition, ADP focuses on solving the optimal control problem, so state estimation does not come with the methods but needs to be designed separately. However, the ADP enables the use of dynamic programming for large-scale optimization problems (in the sense of large number of finite state cells, and/or approximation of high dimensional continuous spaces), and has been reported to provide useful practical results in reasonable time in a number of applications.

This paper focuses on the use of finite state Markov chains for process control. In order to apply MDP for problems of practical interest, a simple but powerful idea from ADP is borrowed: focusing only in state space of interest in closed-loop control (Lee and Lee, 2006). In order to find this part of the state space, it is suggested that clustering of closed-loop data is performed iteratively during the design. Clustering is used for discretization of the continuous state space of the process, and the reminder of the controller and state estimator design and analysis is based on finite state/action MDP techniques. The approach is termed Finite ADP (FADP). The idea of using clustering for discretization (partitioning) is well-known and widely used in the literature of artificial intelligence, including Markov models (Pande et al., 2010) even if fixed grids (rectangular, hexagonal) are much more commonly used with MTM. The suggested approach contains several components similar to those discussed by Yang and Lee (2010), in particular in that clustering is used with MDP in order to solve a process control problem. However, Yang and Lee consider small MTM (less than 10^2 state cells in their examples) and consequently are able to propose a more complex clustering of simulated/historical data series, based on Kohonen's self organizing map and 'outlier treatment'. The FADP aims at a much larger number of state cells (of the order of $10^4 \ldots 10^5$) using a modified k-means algorithm for clustering, and independent one-step-ahead simulations for building the MTM. Most importantly, Yang and Lee aim at constructing an additional RTO-level above linear MPC controllers, while this paper develops these ideas toward a systematic method for practical model-based design of optimal process controllers, emphasizing off-line optimization of the underlying stochastic system's closed-loop performance.

Section 2 introduces modeling, control and state estimation with finite state Markov chains. The main contributions of the paper are given in Section 3, which proposes a novel Finite ADP (FADP) algorithm for process control and state estimation design. The FADP requires a fast clustering algorithm, and a hierarchical k-means algorithm is proposed as a secondary contribution of the paper. The Appendix provides a brief analysis of the properties of this clustering technique. Section 4 gives a thorough illustration of the approach using a well known and easily reproducible MPC benchmark example, the van der Vusse CSTR control problem (Chen et al., 1995). The paper ends with discussion and conclusions.

2. Process modeling, state estimation and control in finite spaces using Markov chains

This section reviews modeling, state estimation and control with finite state controlled Markov chains, to the extent needed in Section 3 to describe the suggested MDP with iterative re-discretization. Control with Markov chains is based on a state space description of the plant

$$\mathbf{x}(k+1) = f(\mathbf{x}(k), \mathbf{u}(k), \mathbf{w}(k)) \tag{1}$$

$$\mathbf{y}(k) = h(\mathbf{x}(k), \mathbf{v}(k)) \tag{2}$$