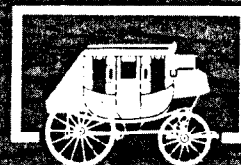# PROCEEDINGS OF THE
32nd IEEE CONFERENCE ON
DECISION AND CONTROL

DECEMBER 15-17, 1993
MARRIOTT RIVERCENTER
SAN ANTONIO, TEXAS USA

VOLUME 4 OF 4

93CH3307-6

IEEE
CONTROL
SYSTEMS
SOCIETY

IEEE

CSS

Implementation of a Dynamic Game Controller for Partially Observed Discrete-Time Nonlinear Systems

C. A. Teolis
University of Maryland *

M. R. James
Australian National University †

J. S. Baras
University of Maryland ‡

Abstract

In this paper we describe an implementation of a certainty equivalence controller for a discrete time partially observed dynamic game problem.

Key words: Output feedback robust control, Nonlinear systems, Certainty equivalence control.

1. Introduction

In [1] a controller is derived for the finite-horizon partially observed dynamic game problem for discrete-time nonlinear systems. The controller is inherently infinite dimensional. It is described by two dynamic programming equations which are nonlinear infinite dimensional recursions: the first, the information state, evolving forward in time, is a function of the state variable, and the second, the value function, evolving backward in time and determining the optimal control policy, is a function of the information state, a variable which takes values in an infinite-dimensional space. The second equation is doubly infinite dimensional which makes direct implementation impractical in terms of computational effort required. A suboptimal alternative is to make use of a *Certainty Equivalence Principle* (CEP) [1, 2, 3]. When the CEP is valid, the value function and optimal control policy may be computed using a simpler recursion which determines the upper value of the corresponding completely observed dynamic game. The state is estimated using the *minimum stress estimate* of Whittle [3], i.e. the state which maximizes the sum of the information state and the upper value.

The conditions under which the CEP holds are often difficult to check. Because of computational and practical considerations, focus is placed on implementation of the certainty equivalence controller, even if the CEP fails to hold. In general the certainty equivalence controller is suboptimal. In spite of this we empirically demonstrate that the implementation of the certainty equivalence controller, although suboptimal, can be stabilizing and robust to noise for many nonlinear systems.

In section 2 the dynamic game problem is stated and the certainty equivalence controller given. In section 3 implementation issues are discussed and some examples are given.

2. Controller

Consider the discrete-time partially observed dynamic game for the system

$$\begin{cases} x_{k+1} &= b(x_k, u_k) + w_k, \\ y_{k+1} &= h(x_k) + v_k \end{cases}$$

on the finite time interval $k = 0, 1, 2, \ldots M$, with cost

$$\mathcal{J}^\mu(u, w, v) = \sup_{x_0 \in \mathbb{R}^n} \left\{ \sum_{l=0}^{M-1} (|x_l|^2 + |u_l|^2) - \frac{1}{\mu} \sum_{l=0}^{M-1} (|w_l|^2 + |v_l|^2) \right\}.$$

The risk sensitivity parameter, μ, is related to the H_∞ norm bound γ, by $\mu = 1/\gamma^2$. For convergence of the algorithm μ must be chosen *small enough*.

In the context of a game there are two players: the control system designer and nature. The designer's objective is to minimize the cost, while nature, acting in direct opposition to that objective, strives to disturb the system so as to maximize the cost. The game is played as follows:

(i) The initial condition $x(0) = x_0$ is unknown, and $y(0) = 0$.

(ii) Player 1 (designer) selects a U-valued control u_k which may be any non-anticipating function of the observation path y. This selection is designed to minimize the cost. Let $\mathcal{O}_{0,M-1}$ denote the class of all such controllers.

(iii) Player 2 (nature) selects a disturbance (w_k, v_k), which is a square summable open loop sequence. Nature's selection is assumed to be made to maximize the cost.

More precisely, if we let $J^\mu(u)$ denote the effect of nature's selection so that

$$J^\mu(u) = \sup_{(w,v) \in \ell_2[0,M-1]} \mathcal{J}^\mu(u, w, v),$$

then the partially observed dynamic game problem is to find an admissible sequence $u \in \mathcal{O}_{0,M-1}$ such that

$$J^\mu(u^*) = \inf_{u \in \mathcal{O}_{0,M-1}} J^\mu(u).$$

The certainty equivalence controller [1, 2, 3] is given by the following two infinite dimensional dynamic programming equations. The sequence of information states $\{p_k\}$ is given recursively by the dynamic programming equation

$$\begin{cases} p_k^\mu &= \Lambda^\mu(u_{k-1}, y_k) p_{k-1}^\mu \\ p_0^\mu &= 0 \end{cases},$$

where $\Lambda^\mu(u, y) p(z)$ is

$$\Lambda^\mu(u, y) p(z) \triangleq \sup_{\xi \in \mathbb{R}^n} \left\{ p(\xi) - \frac{1}{\mu} \left[|h(\xi)|^2 - 2h(\xi)y \right] + |\xi|^2 + |u|^2 - \frac{1}{\mu} |z - b(\xi, u)|^2 \right\}.$$

The sequence of upper values $\{\bar{f}_k^\mu\}$ of the fully observed dynamic game is given recursively by the dynamic programming equation

$$\begin{cases} \bar{f}_k^\mu(x) &= \inf_{u \in U} \sup_{w \in \mathbb{R}^n} \{\bar{f}_{k+1}^\mu(b(x,u)+w) \\ &\quad + |x|^2 + |u|^2 - \frac{1}{2\mu}|w|^2\} \\ \bar{f}_M^\mu(x) &= 0 \end{cases} .$$

If $\widetilde{u}_k^*(x)$ achieves the minimum upper value for step k, then $u_k^* = \widetilde{u}_k^*(x_k)$ is an optimal feedback policy for the completely observed game. The minimum stress estimate $\bar{x}_k$ is given by

$$\bar{x}_k \in \mathop{\arg\max}_{x \in \mathbb{R}^n} \{p_k^\mu(x) + \bar{f}_k^\mu(x)\} \triangleq \widehat{x}_k^\mu$$

where $\widehat{x}_k^\mu$ set valued. The certainty equivalence controller is defined by

$$u_k^* = \widetilde{u}_k^*(\bar{x}_k)$$

and if the CEP holds then this controller is an optimal policy for the partially observed game.

3. Implementation and Examples

The sequence of upper values $\{\bar{f}_k^\mu\}$, and thus the optimal control policy for the fully observed game, is computable off line. The information state, playing a role similar to an observer, is dependent on the current output and control of the system. The information state is part of the controller dynamics which must be computed on line.

In our examples we have found that convergence of the value function and optimal control to steady state is achieved after a relatively small number of iterations. Thus, in order to reduce the computational effort as well as the memory required for storage of the value function and optimal control, our implementation uses the steady state optimal control and value function for all simulations.

The state space, the control space, and the disturbance space, are each truncated to compact sets centered at the origin, e.g., the cube L^n where n is the dimension of the original space and L is the length of a side. Excursions from the respective truncated spaces are are handled by projection onto the boundary of the truncated space. These spaces are each discretized to N points in each dimension uniformly with sampling interval $\Delta \approx 0.01$.

Plotted for each example are the state, the output, and minimum stress estimate trajectories obtained using the certainty equivalence output feedback controller. The systems were subject to additive Gaussian noise in the state update and observation equations with mean 0 and variance 0.05. A risk sensitivity value of $\mu = 0.1$ was used for both examples. Note that in Example 2 (figure 2) the output function is *discontinuous*.

4. Conclusions and Work in Progress

For the two examples presented in this paper we have demonstrated that the certainty equivalence controller is stabilizing and robust to noise. Work in progress includes further investigations for one dimensional as well as multi-dimensional nonlinear systems, e.g. the inverted pendulum. For multi-dimensional systems, however, the computational complexity involved is a clear obstacle. Thus a primary objective of our future work is to investigate methods for the reduction of the computational complexity. Two avenues we shall pursue for such a reduction
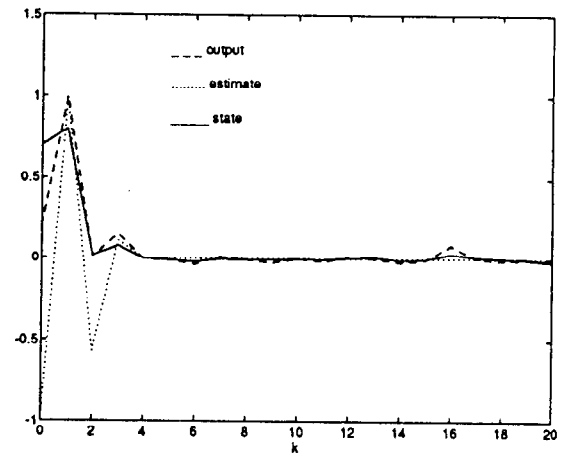

Figure 1: $\begin{aligned} x_{k+1} &= x_k^3 + \sin(u_k) + w_k \\ y_{k+1} &= \sin(2x_k) + v_k \end{aligned}$
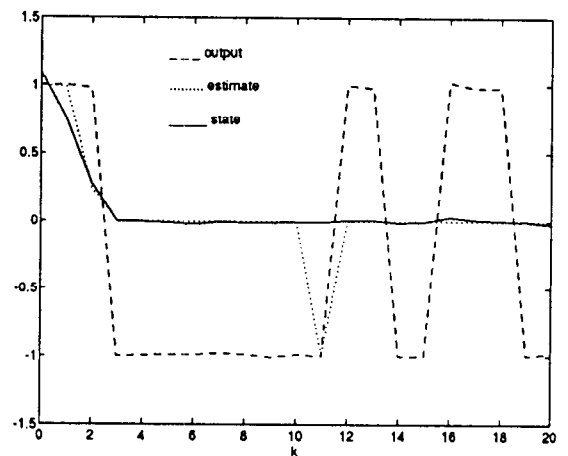

Figure 2: $\begin{aligned} x_{k+1} &= x_k^3 - x_k(x_k - 0.5)u_k + w_k \\ y_{k+1} &= \operatorname{sgn}(x_k) + v_k \end{aligned}$

are (i) approximations and (ii) algorithm redesign. Alternatively, we shall investigate implementation on high performance computational platforms to decrease computational time. We are currently looking at an implementation on the Connection Machine.

We hope that numerical experiments of the type that we have presented here will provide insights into the nature of nonlinear robust control, and ultimately lead to practically useful controllers for a wide class of nonlinear systems.

References

[1] M.R. James, J.S. Baras, and R.J. Elliott. Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems. *To appear: IEEE Transactions on Automatic Control.*

[2] T. Başar and P. Bernhard. H_∞ *Optimal Control and Related Minimax Design Problems.* Birkhauser, Boston, 1991.

[3] P. Whittle. Risk-sensitive linear/quadratic/gaussian control. *Adv. Appl. Prob.*, 13:764–777, 1981.