On the Separation of Estimation and Control in Discrete-Event Systems [Version with proofs of IEEE CDC 2000 paper with same title and by same authors] [1]

George Barrett george.barrett@jhuapl.edu
System and Information Sciences Group, The Johns Hopkins University Applied Physics Laboratory, 11100 Johns Hopkins Road, Laurel, MD 20723-6099

Stéphane Lafortune stephane@eecs.umich.edu
Department of Electrical Engineering and Computer Science, The University of Michigan, 1301 Beal Avenue, Ann Arbor, MI 48109–2122

Abstract

The discussion of this paper concerns both centralized and decentralized control of logical discrete-event systems. Of interest are the maximal information sets of the centralized or decentralized supervisors and the potential for control and estimation policy independence. We show that there exists a form of a centralized supervisor's maximal information sets that is independent of the supervisor's control policy. We also show that this method of separation is not generally applicable in decentralized settings. These results are consistent with the literature in stochastic control; however, supervisory control potentially presents a more simple framework in which to explore these concepts.

Keywords: discrete-event systems, decentralized control, communicating controllers, policy independence

1 Introduction

The term "separability", which refers to the capability of designing control and estimation policies independently, is well known in stochastic control [9], [12]. Separability, in stochastic systems, results from the policy independence of conditional expectations, and this independence is available for systems with particular types of information structures [10], [11].

In this paper, we demonstrate that the "maximal information sets" available in the centralized supervisory control problem have a form which can be determined without knowing the supervisor's control policy. We also demonstrate that this technique does not generally extend to the decentralized supervisory control problem (with or without communication).

2 Preliminaries

We consider a discrete-event system to be modeled by an automaton G with associated language $\mathcal{L}(G)$. Recall that associated with the system is a set of events that can be disabled Σ_c, and there is a set of events, Σ_o, that can be observed. The set of all events is denoted by Σ. The sets of uncontrollable and unobservable events are denoted by $\Sigma_{uc} = \Sigma \setminus \Sigma_c$ and $\Sigma_{uo} = \Sigma \setminus \Sigma_o$ respectively. To control the system, there is a finite set of coordinating controllers represented by

[1]This research is supported in part by internal research and development funding at The Johns Hopkins University Applied Physics Laboratory and by the DDR&E MURI on Low Energy Electronics Design for Mobile Platforms and managed by ARO under grant ARO DAAH04-96-1-0377 .

$Z = \{1, 2, \ldots, n\}$. Centralized control is characterized by $|Z| = 1$. Each controller $i \in Z$ has an associated set of events $\Sigma_{c,i} \subseteq \Sigma_c$ that it can disable, a set Ξ_i of symbols that it is allowed to communicate to other controllers, and a set of events $\Sigma_{o,i} \subseteq \Sigma_o$ that it can directly observe. The events that are unobservable to each controller are given by $\Sigma_{uo,i} = \Sigma \setminus \Sigma_{o,i}$. To represent the fact that controllers have only partial observations of traces in $\mathcal{L}(G)$, a projection operator $\mathbb{P}_i : \Sigma^* \to \Sigma_{o,i}^*$ is used. Recall that $\mathbb{P}_i(\sigma) = \sigma$ if $\sigma \in \Sigma_{o,i}$ otherwise $\mathbb{P}_i(\sigma) = \varepsilon$, and $(\forall s\sigma \in \Sigma^*)$ $\mathbb{P}_i(s\sigma) = \mathbb{P}_i(s)\mathbb{P}_i(\sigma)$. If a subscript is not given, e.g. $\mathbb{P}$, then it is assumed the codomain is Σ_o^*. The inverse projection of $\mathbb{P}_i$ is the mapping $\mathbb{P}_i^{-1} : \Sigma_{o,i}^* \to 2^{\Sigma^*}$ defined as $\mathbb{P}_i^{-1}(\omega) = \{s \in \Sigma^* | \mathbb{P}_i(s) = \omega\}$.

Regarding control, numerous voting schemes can be used to combine the control actions of the controllers in Z (see [7], [13]); however, regardless of how the control signals are combined, the closed-loop language of the plant, G, under control of the set of supervisors, Z, is denoted by $\mathcal{L}(Z/G)$. This basic framework has been extended for the case of communication between controllers in [1], [2], [3]; however, our discussion here will omit controller communication.

3 Separability and the Concept of State

Centralized Control

The situation where only one supervisor ($|Z| = 1$) is to be used to control a plant, G, is called centralized supervisory control and has been well studied. In this scenario the supervisor partially observes, via a projection operator $\mathbb{P}$, a sequence of events generated by the plant. In this case, the dynamics of the closed-loop control system (plant and controller) can be completely determined from the pair

$$(s, \mathbb{P}(s)) \in \mathcal{L}(Z/G) \times \mathbb{P}(\mathcal{L}(Z/G)). \tag{1}$$

The maximal information available to the supervisor to generate a control decision following trace s is denoted here by $\Psi(s)$ and is represented by the set of all traces in the closed-loop behavior that "look" like s to the supervisor:

$$\Psi(s) = \mathbb{P}^{-1}(\mathbb{P}(s)) \cap \mathcal{L}(Z/G). \tag{2}$$

An alternative description of $\Psi(s)$ is the set of all sequences the supervisor "infers" could have occurred given its partial observation of the occurrence of s. For languages containing an infinite number of traces, both $\mathcal{L}(Z/G)$ and (2) may be infinite sets; hence, the designer must handle infinite objects to determine optimal control laws. When the desired closed-loop behavior and plant are both represented by finite automata H and G, respectively, a finite-state estimator for $\mathcal{L}(Z/G)$ can be constructed, and an optimal control law can be based on a finite number of finite-state estimates. Of course, this does not imply that (2) is a finite set nor does it imply that there are a finite number of information sets of the form taken by (2). The capability of using a finite number of finite-state estimates for control means that the maximal information sets (2) contain more information that what is needed for optimal control, and so they can be "compressed" into a useful, finite set of states.

The key properties required to achieve a desired behavior $\mathcal{L}(H)$ are language controllability and observability, and when both conditions are satisfied then constructing an optimal control law and finite-state estimator is performed without issue. However, if $\mathcal{L}(H)$ does not possess the observability property and some (unknown, maximal) "closest" approximation to $\mathcal{L}(H)$ is to be achieved, then synthesis difficulties become apparent. The difficulties arise because the maximal information set (2) used by the state estimator depends upon the control policy implicit in $\mathcal{L}(Z/G)$, the control actions depend upon state estimates, and $\mathcal{L}(Z/G)$ is *not* known prior to the synthesis of the control policy. Basically, this cyclic dependency means that in order to generate a finite-state estimator based on (2), the designer needs to know the control policy, but in order to generate even a single optimal control decision the designer needs to know what dynamics are possible; hence, the set of state/state-estimate pairs from (1) that are reachable needs to be known. Based on (2) control and estimation are not separable, and this policy dependence can be viewed as a primary reason a unique (supremal) controllable and observable sublanguage generally does not exist.

Note that this difficulty does not arise when it is known *a priori* that $\mathcal{L}(H)$ *will be* the language generated by the closed-loop system, i.e., $\mathcal{L}(H) = \mathcal{L}(Z/G)$ is observable and controllable. When it is known a priori that $\mathcal{L}(H)$ will be achieved, then all elements of (2) are known, and the regularity of $\mathcal{L}(H)$ and $\mathcal{L}(G)$ provide the designer with a useful, finite set of states that provide as much information needed *for the purposes of control* as the maximal information sets themselves.

One method that has been repeatedly used to bypass this difficulty is to assume that all controllable events are observable, $\Sigma_c \subseteq \Sigma_o$ which then produces maximal information sets $\mathbb{P}^{-1}(\mathbb{P}(\cdot)) \cap \mathcal{L}(G)$ which do not depend on the control policy (compare to $\mathbb{P}^{-1}(\mathbb{P}(\cdot)) \cap \mathcal{L}(Z/G)$). The result is that only *normal* languages are considered. Another method [4] prioritizes control actions to provide enough information about the control policy so that state-estimates can be generated: The resulting control policies are guaranteed to be maximal. Both of these methods place some type of assumption on the control policies to enable synthesis. Here, we show that no assumptions are needed to establish a separation.

Theorem 3.1 *The centralized supervisor's maximal information sets, $\Psi(\cdot)$, have a form which do not depend on the supervisor's control policy.*

Proof of Theorem 3.1 The proof is straightforward, but we include it for completeness.

Let $\gamma : \Sigma_o^* \to 2^{\Sigma_c}$ be the control policy (disablement map) of the supervisor, and denote the k-th control *action* by $\gamma^k \subseteq 2^{\Sigma_c}$, that is, for $\sigma^k \in \Sigma_o$:

$$\begin{aligned}
\gamma^0 &= \gamma(\epsilon) \\
\gamma^1 &= \gamma(\sigma^1) \\
\gamma^2 &= \gamma(\sigma^1\sigma^2) \\
&\vdots \\
\gamma^j &= \gamma(\sigma^1\sigma^2\ldots\sigma^j).
\end{aligned}$$

We also use the standard definition of $L(G, \gamma(\mathbb{P}(s)))$ to be the set of traces allowed by the control policy following the occurrence of $s \in \mathcal{L}(G)$. The idea behind the proof given here is to show there is a map, $\varphi : (2^{\Sigma_c} \times \Sigma_o)^* 2^{\Sigma_c} \to 2^{\Sigma}$ such that if $s \in \mathcal{L}(G)$ (with $\mathbb{P}(s) = \sigma^1\sigma^2\ldots\sigma^k$) is the trace that has actually occurred in the closed-loop system, then

$$\varphi(\gamma^0\sigma^1\gamma^1\sigma^2\gamma^2\ldots\sigma^k\gamma^k) = \Psi(s) = \mathbb{P}^{-1}(\mathbb{P}(s)) \cap \mathcal{L}(Z/G),$$

that is, the information set $\mathbb{P}^{-1}(\mathbb{P}(s)) \cap \mathcal{L}(Z/G)$ can be "recovered" using only observations and control *actions* and without knowledge of the control *policy*. The proof is by induction on the length of observed sequences.

Let s be the test sequence. For $|\mathbb{P}(s)| = 0$ we define

$$\varphi(\gamma^0) = (\Sigma_{uo} \setminus \gamma^0)^* \cap \mathcal{L}(G). \quad (3)$$

The maximal information set is

$$\begin{aligned} \Psi(s) &= \mathbb{P}^{-1}(\mathbb{P}(s)) \cap \mathcal{L}(Z/G) \\ &= \mathbb{P}^{-1}(\varepsilon) \cap L(G, \gamma(\varepsilon)) \\ &= \Sigma_{uo}^* \cap [(\Sigma_{uo} \setminus \gamma(\varepsilon))^* \cap \mathcal{L}(G)] \\ &= (\Sigma_{uo} \setminus \gamma^0)^* \cap \mathcal{L}(G) \\ &= \varphi(\gamma^0). \end{aligned}$$

So, our assertion about the map φ is true for the base case, and the induction hypothesis is that it holds for $|\mathbb{P}(s)| = k$. Now, we define the "update rule" for φ to be

$$\varphi(\gamma^0\sigma^1\gamma^1 \dots \sigma^k\gamma^k\sigma^{k+1}\gamma^{k+1}) = \left[\varphi(\gamma^0\sigma^1\gamma^1 \dots \sigma^k\gamma^k)\sigma^{k+1}(\Sigma_{uo} \setminus \gamma^{k+1})^*\right] \cap \mathcal{L}(G).$$

For the case $|\mathbb{P}(s)| = k + 1$ we represent s as $w\sigma^{k+1}w'$ where $|\mathbb{P}(w)| = k$ and $\sigma^{k+1} \in \Sigma_o$. Now, the induction hypothesis can be used to yield:

$$\begin{aligned} \varphi(\gamma^0\sigma^1\gamma^1 \dots \sigma^k\gamma^k\sigma^{k+1}\gamma^{k+1}) &= \left[\Psi(w)\sigma^{k+1}(\Sigma_{uo} \setminus \gamma^{k+1})^*\right] \cap \mathcal{L}(G) \\ &= \left[\Psi(w)\sigma^{k+1}\Sigma_{uo}^*\right] \cap L(G, \gamma(\sigma^1\sigma^2 \dots \sigma^{k+1})) \\ &= \left[\Psi(w)\sigma^{k+1}\Sigma_{uo}^*\right] \cap \mathcal{L}(Z/G) \\ &= \left[\left[\mathbb{P}^{-1}(\mathbb{P}(w)) \cap \mathcal{L}(Z/G)\right]\sigma^{k+1}\Sigma_{uo}^*\right] \cap \mathcal{L}(Z/G) \\ &= \mathbb{P}^{-1}(\mathbb{P}(w\sigma^{k+1}\Sigma_{uo}^*)) \cap \mathcal{L}(Z/G) \\ &= \mathbb{P}^{-1}(\mathbb{P}(w\sigma^{k+1}w')) \cap \mathcal{L}(Z/G) \\ &= \Psi(w\sigma^{k+1}w') \\ &= \Psi(s), \end{aligned}$$

and the proof is complete. ⊣

Theorem 3.1 also has an additional benefit in the DEDS framework that is revealed in its proof. If the plant and desired behaviors are regular, then the sufficient statistic $\varphi(\gamma^0\sigma^1\gamma^1 \dots \sigma^k\gamma^k)$ has a finite-state representation. Furthermore, only a record of $(\sigma^{k+1}\gamma^{k+1})$ is required versus a record of every observation and control action. Hence, policy separation is possible in the centralized case using a finite statistic when the behaviors are regular.

Decentralized Control

Now, the discussion will shift focus to the situation where two controllers ($|Z| = 2$) are to act in concert on a plant as shown in Figure 1. The dynamics of the closed-loop control system can be completely determined from the three-tuple

$$(s, \mathbb{P}_1(s), \mathbb{P}_2(s)) \in \mathcal{L}(Z/G) \times \mathbb{P}_1(\mathcal{L}(Z/G)) \times \mathbb{P}_2(\mathcal{L}(Z/G)), \quad (4)$$

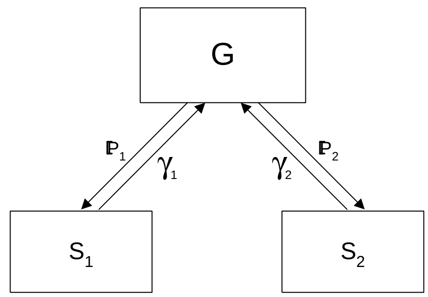

Figure 1: Multiple controller supervision without communication.

where, for each controller, we have basically the same form as in (1) above.

If it is known a priori (via a polynomial test for coobservability [8] and controllability) that the desired language $\mathcal{L}(H)$ is achievable, then the maximal information sets available to each controller following some trace s are

$$\mathbb{P}_1^{-1}(\mathbb{P}_1(s)) \cap \mathcal{L}(H), \text{ and } \mathbb{P}_2^{-1}(\mathbb{P}_2(s)) \cap \mathcal{L}(H). \tag{5}$$

Where, again, all components of (5) are known and, when coupled with the regularity of $\mathcal{L}(H)$ and $\mathcal{L}(G)$, provide the designer with two useful, finite sets of states for the controllers. Although the controllers "cooperate" to achieve $\mathcal{L}(H)$, the design of each controller is completely decoupled from the design of the other.

If the test for coobservability of $\mathcal{L}(H)$ fails and some (unknown, maximal) "closest" approximation to $\mathcal{L}(H)$ is to be achieved, then synthesis difficulties are encountered. Synthesis in the decentralized case suffers from the compounded difficulties that the maximal information set of Controller 1 depends upon both the control policies of Controller 1 and Controller 2 and vice versa. Existing algorithms ([5], [6]) for finding maximal equilibrium solutions are not guaranteed to terminate.[2] To date, no useful concept of state has been found to assist synthesis of decentralized controllers when $\mathcal{L}(Z/G)$ is not known a priori.

Given our previous comment that for each controller $i \in Z$ both (4) and (5) have the same forms as (1) and (2), it is natural to ask whether Theorem 3.1 generalizes to the multiple controller case. The answer is generally negative. Attempting to write an analogous Eq(3) for the decentralized controllers reveals the problem immediately:

$$\begin{aligned} \varphi_i(\gamma_i^0) &= (\Sigma_{uo,i} \setminus \gamma_i^0)^* \cap \mathcal{L}(G) \cap \\ & \left[\varphi_j(\gamma_j^0) \bigcup_{\sigma^1 \in \Sigma_{o,j} \setminus \Sigma_{o,i}} \varphi_j(\gamma_j^0 \sigma^1 \gamma_j^1) \cdots \bigcup_{\sigma^1, \sigma^2, \ldots, \sigma^k \in \Sigma_{o,j} \setminus \Sigma_{o,i}} \varphi_j(\gamma_j^0 \sigma^1 \gamma_j^1 \ldots \sigma^k \gamma_j^k) \cdots \right] \end{aligned} \tag{6}$$

The "additional" components of Eq(6) that Eq(3) does not have represent an apparently inescapable dependence of controller i's information sets on controller j's control policy. Basically, controller i needs to know all of controller j's contingency plans for all event sequences that controller i does not disable and/or does not observe.

[2]The idea behind these algorithms is to hold one controller fixed and to optimize the other, then fix the second controller and optimize the first, repeating until convergence.

References

[1] G. Barrett and S. Lafortune, *A novel framework for decentralized supervisory control with communication*, Proceedings of IEEE International Conference on System, Man and Cybernetics (San Diego, CA), 1998, pp. 617–620.

[2] ______, *On the synthesis of communicating controllers with decentralized information structures for discrete-event systems*, Proceedings of the 37th IEEE Conference on Decision and Control (Tampa, FL), 1998, pp. 3281–3286.

[3] ______, *Decentralized supervisory control with communicating controllers*, IEEE Trans. Automatic Control (2000), To Appear.

[4] N. Ben Hadj-Alouane, S. Lafortune, and F. Lin, *Centralized and distributed algorithms for online synthesis of maximal control policies under partial observation*, Discrete Event Dynamic Systems: Theory and Applications **6** (1996), no. 4, 379–427.

[5] G. J. Olsder, *Comments on a numerical procedure for the solution of differential games*, IEEE Transactions on Automatic Control **20** (1975), 704–705.

[6] A. Overkamp and J. H. van Schuppen, *A characterization of maximal solutions for decentralized discrete event control problems*, Proceedings of the International Workshop on Discrete Event Systems (Edinburgh, UK), August 1996, pp. 278–283.

[7] J. H. Prosser, M. Kam, and H. G. Kwatny, *Decision fusion and supervisor synthesis in decentralized discrete-event systems*, Proceedings of the 1997 American Control Conf., June 1997, pp. 2251–2255.

[8] K. Rudie and J. C. Willems, *The computational complexity of decentralized discrete-event control problems*, IEEE Transactions on Automatic Control **40** (1995), no. 7, 1313–1318.

[9] C. Striebel, *Sufficient statistics in the optimum control of stochastic systems*, Journal of Mathematical Analysis and Applications **12** (1965), 576–592.

[10] H. S. Witsenhausen, *Separation of estimation and control for discrete time systems*, Proceedings of the IEEE **59** (1971), no. 11, 1557–1566.

[11] ______, *On policy independence of conditional expectations*, Information and Control **26** (1975), 65–75.

[12] W. M. Wonham, *On the separation theorem of stochastic control*, SIAM J. Control **6** (1968), no. 2, 312–326.

[13] T-S. Yoo and S. Lafortune, *A new architecture for decentralized supervisory control*, 1999, Preprint.