Information Structures, the Witsenhausen Counterexample, and Communicating Using Actions

Pulkit Grover, Carnegie Mellon University

Abstract

The concept of "information-structures" in decentralized control is a formalization of the notion of "who knows what and when do they know it." Even seemingly simple problems with simply stated information structures can be extremely hard to solve. Perhaps the simplest of such unsolved problem is the celebrated Witsenhausen counterexample, formulated by Hans Witsenhausen in 1968. This article discusses how the information structure of the Witsenhausen counterexample makes it hard, and how an information-theoretic approach, that explores the knowledge-gradient due to a non-classical information pattern, has helped obtain insights into the problem.

Keywords: team decision theory; decentralized control; information theory; implicit communication

1 Introduction

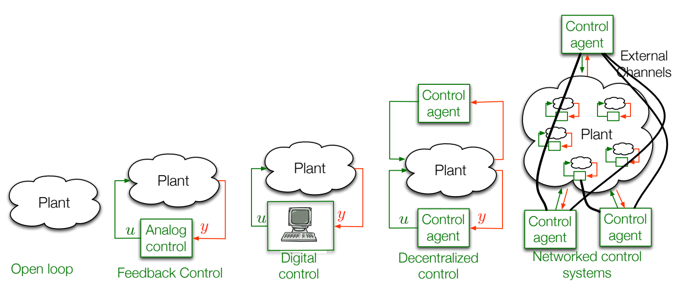

Figure 1: The evolution of control systems. Modern "networked control systems" (also called "cyber-physical systems") are decentralized and networked using communication channels.

Modern control systems often comprise of multiple decentralized control agents that interact over communication channels. What characteristic distinguishes a centralized control problem from a decentralized one? One fundamental difference is a "knowledge gradient": agents in a decentralized team often observe, and hence know, different things. This observation leads to the idea of *information patterns* [1], a formalization of the notion of "who knows what and when do they know it" [2, 3].

The information pattern is said to be *classical* if all agents in the team receive the same information, and have perfect recall (so they do not forget it). What is so special about classical information patterns? For these patterns, the presence of external communication links has no effect on the optimal costs! After all, what could the agents use the communication links for, when there is no knowledge gradient? More interesting, therefore, are the problems for which the information-pattern is *non*-classical. These problems sit at the intersection of communication and control: *communication between agents* can help reduce the knowledge-differential that exists between them, helping them perform the *control* task. Intellectually and practically, the concept of non-classical information patterns motivates a lot of formulations at the control-communication intersection. Many of these formulations — including some discussed in this Encyclopedia (e.g. [4–8]) — intellectually ask the question: for a realistic channel that is constrained by noise, bandwidth, and speed, what is the optimal communication *and* control strategy?

One could ask the question of optimal control strategy even for decentralized control problems where no external channel is available to bridge this knowledge gradient. Why could these problems be of interest? First, these problems are limiting cases of control with communication constraints. Second, and perhaps more importantly, they bring out an interesting possibility that can allow the agents to "communicate", *i.e.*, exchange information, *even when the external channel is absent.* It is possible to use control actions to *communicate* through changing the system state! We now introduce this form of communication using a simple toy example.

2 Communicating using actions: an example

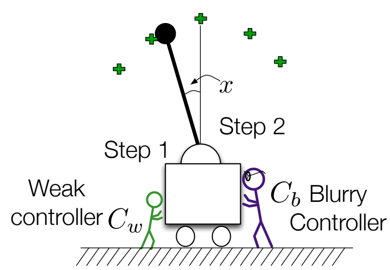

Figure 2: Two controllers, with their respective strengths and weaknesses, attempting to bring an inverted pendulum close to the center. Also shown (using green '+' signs) are possible quantization points chosen by the controllers for a quantization-based control strategy.

To gain intuition into when communication using actions could be useful, consider the inverted pendulum example shown in Fig. 2. The goal of the two agents is to bring the pendulum as close to the origin as possible. Both controllers have their strengths and weaknesses. The "weak" controller C_w has little energy, but has perfect state observations. On the other hand, the "blurry" controller C_b has infinite energy, but noisy observations. They act one after the other, and their goal is to move the pendulum close to the center from any initial state. The information structure of the problem is nonclassical: the C_w, but not C_b, knows the initial state of the pendulum, and C_w does not know the precise (noisy) observation of C_b using which C_b takes actions.

A possible strategy: A little thought reveals an interesting strategy: the weak controller, having perfect observations, can move the state to the closest of some pre-decided points in space, effectively *quantizing* the state. If these quantization points are sufficiently far from each other, they can be estimated accurately (through the noise) by the blurry controller, which can then use its energy to push the pendulum all the way to zero. In this way, the weak controller expends little energy, but is able to "communicate" the state through the noise to the blurry controller, by making it take values on a finite set. Once the blurry controller has received the state through the noise, it can use its infinite energy to push the state to zero.

3 The Witsenhausen counterexample

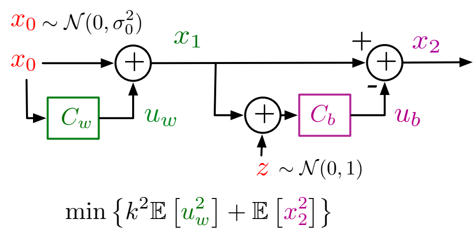

Figure 3: The Witsenhausen counterexample is a deceptively simple two-time-step two-controller decentralized control problem. The weak and the blurry controllers, C_w and C_b act in a sequential manner.

The above two-controller inverted-pendulum example is, in fact, motivated by what is now known as "the Witsenhausen counterexample," formulated by Hans Witsenhausen is 1968 [9] (see Fig. 3). In the counterexample, two controllers (denoted here by C_w for "weak" and C_b for "blurry") act one after the other in two time-steps to minimize a quadratic cost function. The system state is denoted by x_t, where t is the time-index. u_w and u_b denote the inputs generated by the two controllers. The cost function is $k^2\mathbb{E}\left[u_w^2\right] + \mathbb{E}\left[x_2^2\right]$ for some constant k. The initial state x_0 and the noise z at the input of the blurry controller are assumed to be Gaussian distributed and independent, with variances σ_0^2 and 1 respectively. The problem is a "Linear-Quadratic-Gaussian" (LQG) problem, *i.e.*, the state evolution is linear, the costs are quadratic, and the primitive random variables are Gaussian.

Why is the problem called a "counterexample"? The traditional "certainty-equivalence" principle [10] shows that for all centralized LQG problems, linear control laws are optimal. Witsenhausen [9] provided a nonlinear strategy for the Witsenhausen problem which outperforms all linear strategies. Thus, the counterexample showed that the certainty-equivalence doctrine does not extend to decentralized control.

What is this strategy of Witsenhausen that outperforms all linear strategies? It is, in fact, a quantization-based strategy, as suggested in our inverted-pendulum story above. Further, it was shown by Mitter and Sahai [3] that multi-point quantization strategies can outperform linear strategies by an arbitrarily large factor! This observation, combined with the simplicity of the counterexample, makes the problem very important in decentralized control. This simple two-time-step two-controller LQG problem needs to be understood to have any hope of understanding larger and more complex problems.

While the optimal costs for the problem are still unknown[1], there exists a wealth of understanding of the counterexample that has helped address more complicated problems. A body of work, starting with that of Baglietto, Parisini, and Zoppoli [11], numerically obtained solutions that could be close-to-optimal (although there is no mathematical proof thereof). All these solutions have a consistent form (illustrated in Fig. 4), with slight improvements in the optimal cost. Because the discrete

[1] Even though it is known that an optimal strategy exists [9, 14].

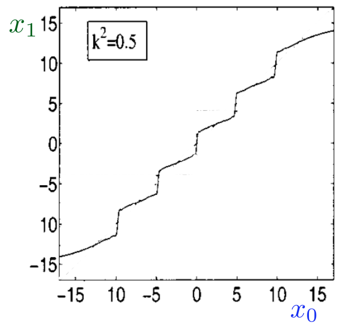

Figure 4: The optimization solution of Baglietto *et al.* [11] for $k^2 = 0.5$, $\sigma_0^2 = 5$. The information-theoretic strategy of "dirty-paper coding" [12] also yields the same curve [13].

version of the problem, appropriately relaxed, is known to be NP-complete [15], this approach cannot be used to understand the entire parameter space, and hence has focused on one point: $k^2 = 0.5, \sigma_0^2 = 5$. Nevertheless, the approach has proven to be insightful: a recent information-theoretic body of work shows that the strategies of Fig. 4 can be thought of as information-theoretic strategies of "dirty-paper coding" [12] that is related to the idea of embedding information in the state. The question here is: how do we embed the information about the state *in the state itself*?

An information-theoretic view of the counterexample: This information-theoretic approach, that culminated in [16], also obtained the first approximately-optimal solutions to the Witsenhausen counterexample, as well as its vector extensions. The result is established by analyzing information-flows in the counterexample that work towards minimizing the knowledge-gradient, effectively an information-pattern in which C_w can predict the observation of C_b more precisely. The analysis provides *an information-theoretic lower bound on cost that holds irrespective of what strategy is used*. For the original problem, this characterizes the optimal costs (with associated strategies) within a factor of 8 for all problem parameters (i.e., k, and σ_0^2). For any finite-length extension, uniform-finite-ratio approximations also exist [16]. The asymptotically infinite-length extension has been solved *exactly* [17].

The problem has also driven delineation of decentralized LQG control problems with optimal linear solutions and those with nonlinear optimal solutions. This led to development and understanding of many variations of the counterexample [2, 18–20], and understanding that can extend to larger decentralized control problems. More recent work shows that the promise of the Witsenhausen counterexample was not a misplaced one: the information-theoretic approach that provides approximately-optimal solutions to the counterexample [16] yields solutions to other more complex (e.g. multi-controller, more time-steps) problems as well [21, 22].

4 Summary and future directions

Even simple problems with non-classical information structures can be hard to solve using classical techniques, as is demonstrated by the Witsenhausen counterexample. However, non-classical information pattern for some simple problems — starting with the counterexample — have recently been explored via an information-theoretic lens, yielding the first optimal or approximately-optimal solutions to these problems. This approach is promising for larger decentralized control problems as well. It is now important to explore what is the simplest decentralized control problem that can not be solved (exactly or approximately) using ideas developed for the counterexample. In this manner, the Witsenhausen counterexample can provide a platform to unify the more modern (*i.e.*, external-channel centric approaches, see [4–8] in the encyclopedia) with the more classical decentralized LQG problems, leading to enriching and useful formulations.

References

[1] H. S. Witsenhausen, "Separation of estimation and control for discrete time systems," *Proceedings of the IEEE*, vol. 59, no. 11, pp. 1557–1566, Nov. 1971.

[2] Y. C. Ho, M. P. Kastner, and E. Wong, "Teams, signaling, and information theory," *IEEE Trans. Autom. Control*, vol. 23, no. 2, pp. 305–312, Apr. 1978.

[3] S. K. Mitter and A. Sahai, "Information and control: Witsenhausen revisited," in *Learning, Control and Hybrid Systems: Lecture Notes in Control and Information Sciences 241*, Y. Yamamoto and S. Hara, Eds. New York, NY: Springer, 1999, pp. 281–293.

[4] G. Nair, "Quantized control and data-rate constraints," *Encyclopedia of Systems and Control*, 2014.

[5] C. Kawan, "Data-rate of nonlinear control systems and feedback entropy," *Encyclopedia of Systems and Control*, 2014.

[6] L. Bushnell, "Networked control systems: architecture and stability issues," *Encyclopedia of Systems and Control*, 2014.

[7] J. Hespanha, "Networked control systems: estimation and control over lossy networks," *Encyclopedia of Systems and Control*, 2014.

[8] S. Yuksel, "Information and communication complexity of networked control systems," *Encyclopedia of Systems and Control*, 2014.

[9] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal on Control*, vol. 6, no. 1, pp. 131–147, Jan. 1968.

[10] D. Bertsekas, *Dynamic Programming.* Belmont, MA: Athena Scientific, 1995.

[11] M. Baglietto, T. Parisini, and R. Zoppoli, "Nonlinear approximations for the solution of team optimal control problems," *Proceedings of the IEEE Conference on Decision and Control (CDC)*, pp. 4592–4594, 1997.

[12] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[13] P. Grover and A. Sahai, "Vector Witsenhausen counterexample as assisted interference suppression," *Special issue on Information Processing and Decision Making in Distributed Control Systems of the International Journal on Systems, Control and Communications (IJSCC)*, vol. 2, pp. 197–237, 2010.

[14] Y. Wu and S. Verdú, "Witsenhausen's counterexample: A view from optimal transport theory," in *IEEE Conference on Decision and Control and European Control Conference (CDC-ECC),*, 2011, pp. 5732–5737.

[15] C. H. Papadimitriou and J. N. Tsitsiklis, "Intractable problems in control theory," *SIAM Journal on Control and Optimization*, vol. 24, no. 4, pp. 639–654, 1986.

[16] P. Grover, S.-Y. Park, and A. Sahai, "Approximately-optimal solutions to the finite-dimensional Witsenhausen counterexample," *IEEE Trans. Autom. Control*, vol. 58, no. 9, p. 2189, Sep. 2013.

[17] C. Choudhuri and U. Mitra, "On Witsenhausen's counterexample: The asymptotic vector case," in *IEEE Information Theory Workshop (ITW)*, 2012, pp. 162–166.

[18] R. Bansal and T. Başar, "Stochastic teams with nonclassical information revisited: When is an affine control optimal?" *IEEE Trans. Autom. Control*, vol. 32, pp. 554–559, Jun. 1987.

[19] M. Rotkowitz, "Linear controllers are uniformly optimal for the Witsenhausen counterexample," *Proceedings of the 45th IEEE Conference on Decision and Control (CDC)*, pp. 553–558, Dec. 2006.

[20] T. Başar, "Variations on the theme of the Witsenhausen counterexample," *Proceedings of the 47th IEEE Conference on Decision and Control (CDC)*, pp. 1614–1619, 2008.

[21] S. Y. Park and A. Sahai, "It may be easier to approximate decentralized infinite-horizon LQG problems," in *51st IEEE Conference on Decision and Control (CDC)*, 2012, pp. 2250–2255.

[22] P. Grover, "Actions can speak more clearly than words," Ph.D. dissertation, UC Berkeley, Berkeley, CA, 2010.