Proceedings of the Forty-Second Annual Allerton Conference
on Communication, Control and Computing

Decentralized Stochastic Decision Problems and Polynomial Optimization

Randy Cogill Sanjay Lall
Stanford University
{rcogill, lall}@stanford.edu

Abstract

In this paper we consider the problem of determining optimal decentralized decision rules in discrete stochastic decision problems. Here we consider a static single-stage problem. It has been shown in [8] that the static problem is $\mathcal{NP}$-hard, even for the case of two decision makers. We show that this problem has an equivalent formulation as minimization of a bilinear polynomial subject to linear constraints. We then form a relaxation of this polynomial optimization problem, from which we can compute suboptimal decentralized decision rules as well as bounds on the optimal achievable value. The methods are illustrated by an example of decentralized detection.

1 Introduction

Decentralized decision problems are optimization problems in which a collection of decisions are made in response to a set of observations with the goal of minimizing some cost. The complicating factor is that decisions can only be made to depend on some specified subset of the observations. That is, the complete set of observations can be thought of as the state of the environment. Each decision is made on the basis of an incomplete observation of the state, although the cost incurred depends on the entire state and set of decisions. Such problems are common in areas such as engineering and economics. Much of the early work on team decision problems was motivated by economic problems [2]. In certain engineering problems, such as the design of distributed detection schemes and distributed data transmission protocols, the key difficulty lies in the design of good rules for interacting decision makers to follow.

Here we consider a fairly general discrete version of this problem, where the sets of possible observations and decisions are finite. The problem considered is a *static* decision problem, where a single set of decisions is made in response to a single set of observations. Given the probabilities of all sets of observations, the goal is to choose decentralized decision rules which minimize the expected cost. This problem is shown in [8] to be $\mathcal{NP}$-hard, even for the case of two decision makers. Therefore, the goal of this paper is to determine effective methods of computing good *suboptimal* solutions to this problem. Here we show that this problem can be equivalently formulated as a minimization of a polynomial subject to linear constraints. Relaxations of this polynomial optimization problem can then be efficiently solved. From these relaxations, we obtain lower bounds on the minimum achievable value for the original problem, as well as suboptimal decision rules. The combination of lower bounds together with suboptimal solutions is powerful, since this gives us a way to put a bound on how suboptimal the best known decision rules are.

2 Previous work

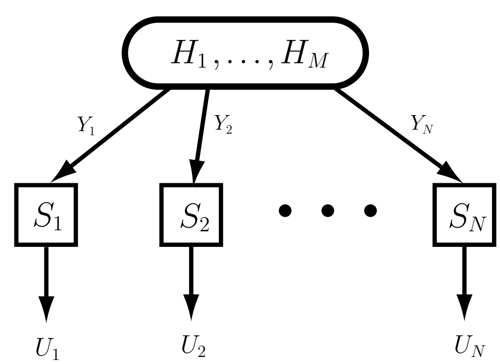
Much of the previous work on decentralized decision problems can be roughly categorized as complexity results, tractable special cases, and applications. Some of the earliest work on decentralized decision problems is the work of Radner and Marschak [4, 2]. Along with introducing the general static decentralized decision problem, they have shown that for certain convex quadratic costs and continuous decision variables, person-by-person optimality is sufficient for global optimality. A nice survey of the early work in the field of decentralized decision problems, including extensions to dynamic problems, can be found in [1]. In [8], it is shown that the general static decentralized decision problem with finite state and action spaces is $\mathcal{NP}$-hard.

A great deal of work on the static decentralized decision problem has been done for the application of decentralized detection. The decentralized detection problem was introduced in [6], where it was shown that under certain independence assumptions, optimal decentralized detection rules take the form of likelihood ratio tests. However, it is shown in [8] that the problem of decentralized detection, a special case of the decentralized decision problem, is also $\mathcal{NP}$-hard. Therefore, most approaches to the problem of decentralized detection focus on determining person-by-person optimal detection rules. Surveys of the field of decentralized detection can be found in [7] and the book [9].

3 Motivating example

In this section we motivate the study of the general problems discussed in this paper by a specific application. The problem of *decentralized detection* is an example of a decentralized stochastic decision problem. Here we present a very brief overview of this subject. Detailed surveys can be found in [7] and [9].

In a detection problem, we have several hypotheses on the underlying state of our environment, and we would like use measurements of our environment to decide which hypothesis is true.



Classical detection methods assume all measurements are available to a single detector, which estimates the true hypothesis based on all measurements. Such a detection scheme is called *centralized*. Optimal decision rules in centralized schemes are given by the well-known MAP (maximum a-posteriori probability) detector. In a *decentralized* detection scheme, each sensor is responsible for making a decision based only on its own measurement. The goal is to choose decision rules for all sensors which are optimal with respect to some system-wide cost function.

For example, suppose we have a collection of sensors each monitoring various elements of some industrial process. We would like the sensors to sound an alarm when some part of the process is malfunctioning. In this case we may wish to maximize the probability that the alarm sounds when there is a malfunction and does not sound when there is no malfunction. One option is to transmit all sensor measurements to a central location, where a decision to sound the alarm is made on the basis of all measurements. An alternative is to equip each sensor with its own decision rule and the ability to sound the alarm. When the loss of performance associated with employing the second alternative is small, such a scheme is preferable due to the reduced implementation complexity associated with the elimination of the communication requirements.

One might initially assume that good decentralized decision rules can be obtained by allowing each sensor to use a MAP detection rule. While this is true in some special cases, it is not true in general. Unlike the centralized case, the general problem of computing optimal decentralized detection rules is $\mathcal{NP}$-hard [8]. While centralized decision rules are described by Also, decentralized decision rules can appear considerably more complex than their centralized counterparts. For example, optimal decentralized decision rules typically involve hedging among the sensors, a strategic element which is not present when simply using MAP rules at each detector.

Due to the complexity of this problem, most existing methods for computing decentralized detection rules produce equilibrium policies. Such policies are said to be *person-by-person optimal*; for a set of such decision rules, no improvement can be obtained by adjusting the decision rule for any given sensor while leaving the others fixed. In general, a single problem instance may have many equilibrium policies. The globally optimal policy is clearly an equilibrium policy. However, for any given equilibrium policy, we have no way of knowing how this policy relates to the globally optimal policy. In particular, we have no way of knowing how much improvement we could obtain by using the globally optimal policy. In the next section we will show by a simple example that an equilibrium policy can perform arbitrarily poorly compared to the optimal policy. The methods that we present for in this paper are *relaxations*. In addition to generating an equilibrium policy, they return a lower bound on the minimum achievable cost by any decentralized policy. When the bound is exact, we have a proof that our computed policy is globally optimal. Even when the bound is not exact, we still have a measure of the suboptimality of the computed policy.

4 Main results

4.1 Formulation and complexity

Decentralized decision problems are optimization problems in which a collection of decisions are made in response to a set of observations with the goal of maximizing some payoff. The complicating factor is that each decision can only be made to depend on some specified subset of the observations. In this section we consider a general *static* decentralized decision problem, also commonly referred to as a *team decision problem* [4, 1]. For notational simplicity, we only discuss problems involving two decision makers. Extensions of all results to the general case of N decision makers is straightforward. The specific problem under consideration is the following:

Decentralized Decision Problem: *Given finite sets* Y_1, Y_2, U_1, U_2, *a probability mass function* $p : Y_1 \times Y_2 \to \mathbb{R}$, *and a cost function* $c : Y_1 \times Y_2 \times U_1 \times U_2 \to \mathbb{R}$, *find policies* $\gamma_i : Y_i \to U_i$, $i = 1, 2$ *which minimize the expected cost*

$$J(\gamma_1, \gamma_2) = \sum_{Y_1 \times Y_2} c(y_1, y_2, \gamma_1(y_1), \gamma_2(y_2))p(y_1, y_2)$$

It was shown in [8] that this problem is $\mathcal{NP}$-hard. Unless $\mathcal{P} = \mathcal{NP}$, we cannot hope to find an efficient algorithm capable of always producing globally optimal policies. Methods for computing policies must aim to find good suboptimal solutions.

One way to formulate the static decentralized decision problem involves expressing policy i as a $|Y_i| \times |U_i|$ Boolean matrix for each i:

$$K^i_{y_i u_i} = \begin{cases} 1 & \text{if } \gamma_i(y_i) = u_i \\ 0 & \text{otherwise} \end{cases}$$

Similarly, we can express the system cost as a $|Y_1||Y_2| \times |U_1||U_2|$ matrix:

$$\begin{aligned} C_{yu} &= c(y, u)p(y) \\ &= c(y_1, y_2, u_1, u_2)p(y_1, y_2) \end{aligned}$$

Here, the matrix is indexed according to a lexicographic order on the pairs (y_1, y_2) and (u_1, u_2). The static decentralized decision problem can be equivalently formulated as

$$\begin{array}{rll} \text{minimize:} & \sum_{y,u} C_{yu} K_{yu} & \\ \text{subject to:} & K_{yu} = K^1_{y_1 u_1} K^2_{y_2 u_2} & \\ & K^i \geq 0 & i = 1, 2 \\ & K^i \mathbf{1} = \mathbf{1} & i = 1, 2 \\ & K_{yu} \in \{0, 1\} & \text{for all } y, u \end{array} \tag{1}$$

This problem is clearly a nonconvex optimization problem due to the Boolean constraints and the bilinear constraint. However, we can eliminate the Boolean constraints and show that the resulting problem is equivalent to (1):

Theorem 1. *The optimization problem*

$$\begin{array}{rll} \textit{minimize:} & \sum_{y,u} C_{yu} K_{yu} & \\ \textit{subject to:} & K_{yu} = K^1_{y_1 u_1} K^2_{y_2 u_2} & \\ & K^i \geq 0 & i = 1, 2 \\ & K^i \mathbf{1} = \mathbf{1} & i = 1, 2 \end{array} \tag{2}$$

always has an optimal solution satisfying $K_{yu} \in \{0, 1\}$ *for all* y, u.

Proof: Suppose $\overline{K^1}$ and $\overline{K^2}$ are optimal for (2). Note that $\overline{K^1}$ and $\overline{K^2}$ may have non-integer entries. The problem

$$\begin{array}{rl} \text{minimize:} & \sum_{y_1,u_1}\left(\sum_{y_2,u_2} C_{yu}\overline{K^2}_{y_2,u_2}\right) K^1_{y_1,u_1} \\ \text{subject to:} & K^1 \geq 0 \\ & K^1\mathbf{1} = \mathbf{1} \end{array}$$

is a linear program in the variable K^1. An optimal solution $\widehat{K^1}$ to this LP satisfies

$$\sum_{y,u} C_{yu}\widehat{K^1}_{y_1u_1}\overline{K^2}_{y_2u_2} \leq \sum_{y,u} C_{yu}\overline{K^1}_{y_1u_1}\overline{K^2}_{y_2u_2}$$

Also, it is clear that $\widehat{K^1}$ can be chosen to have 0-1 entries. Now consider the linear program

$$\begin{array}{rl} \text{minimize:} & \sum_{y_2,u_2}\left(\sum_{y_1,u_1} C_{yu}\widehat{K^1}_{y_1,u_1}\right) K^2_{y_2,u_2} \\ \text{subject to:} & K^2 \geq 0 \\ & K^2\mathbf{1} = \mathbf{1} \end{array}$$

Again, an optimal solution $\widehat{K^2}$ satisfies

$$\sum_{y,u} C_{yu}\widehat{K^1}_{y_1u_1}\widehat{K^2}_{y_2u_2} \leq \sum_{y,u} C_{yu}\widehat{K^1}_{y_1u_1}\overline{K^2}_{y_2u_2}$$

and can be chosen to have 0-1 entries. Therefore, $\widehat{K^1}$, $\widehat{K^2}$, and $\widehat{K} = \widehat{K^1} \otimes \widehat{K^2}$ constitute an optimal 0-1 solution to (2). ■

When the optimal solution in (2) is not unique, there may be a mixed optimal solution. However, the above theorem shows that there is always a Boolean solution which achieves the same objective value.

Although we were able to eliminate the Boolean constraints, finding a globally optimal solution to (2) is still a difficult problem. The most common approach for handling this problem is to employ an iterative scheme for finding a *person-by-person* optimal solution [4, 1, 9]. This type of scheme starts by initially choosing an arbitrary pair of policies. Policies are then modified by alternately optimizing each policy while leaving the other policy fixed (as in the proof of Theorem 1). Since there are a finite number of policies, and each step never results in a decrease in the objective, this method leads to an equilibrium solution in a finite number of steps. The problem with such methods is that problems may have many equilibria, and it not clear if any given equilibrium solution is necessarily a good one. In fact, we can show by a simple example that an equilibrium policy can perform arbitrarily poorly compared to the optimal policy.

Consider the case where $Y_1 = Y_2 = U_1 = U_2 = \{1, 2\}$ and we have the cost function

$$c(y_1, y_2, u_1, u_2) = \begin{cases} 1 & \text{for } u_1 = u_2 = 1 \text{ and all } y_1, y_2 \\ \rho & \text{for } u_1 = u_2 = 2 \text{ and all } y_1, y_2 \\ \rho + 1 & \text{for } u_1 \neq u_2 \text{ and all } y_1, y_2 \end{cases}$$

Consider the decentralized policy where $\gamma_1(y_1) = 2$ for all y_1 and $\gamma_2(y_2) = 2$ for all y_2. For any probability distribution on $Y_1 \times Y_2$, this policy achieves an expected cost of

$J(\gamma_1, \gamma_2) = \rho$. Leaving one decision rule fixed while changing the other always achieves an expected cost greater than ρ. Therefore, this policy is an equilibrium policy. However, the optimal decentralized policy in this case achieves an expected cost of $J(\gamma_1^*, \gamma_2^*) = 1$. Since ρ is arbitrary, we can choose its value so that a suboptimal equilibrium policy achieves an expected cost arbitrarily worse than the optimal cost.

We consider an alternate approach to searching for equilibrium policies in this paper. We treat this problem as a polynomial optimization problem, and apply lifting methods to obtain convex relaxations. Such methods either produce a globally optimal solution, or produce a suboptimal solution along with a bound on its suboptimality.

4.2 A relaxation for the polynomial problem

Here we will discuss a specific low-order relaxation of the problem (2). This relaxation is formed by adding *valid constraints* to the problem, then removing all nonconvex constraints [5]. Valid constraints are constraints which, when added to a problem, do not change its feasible set. For example, sums or products of existing constraints produce valid constraints.

Consider the problem (2) with additional valid constraints added by taking products of the original linear constraints:

$$\begin{array}{rll} \text{minimize:} & \sum_{y,u} C_{yu} K_{yu} & \\ \text{subject to:} & K_{yu} = K^1_{y_1 u_1} K^2_{y_2 u_2} & \\ & \sum_{u_1} K_{yu} = K^2_{y_2 u_2} & \text{for all } y_1 \in Y_1 \\ & \sum_{u_2} K_{yu} = K^1_{y_1 u_1} & \text{for all } y_2 \in Y_2 \\ & K^i \mathbf{1} = \mathbf{1} & i = 1, 2 \\ & K \geq 0 & \end{array}$$

The additional constraints $K^i \geq 0$ and $K\mathbf{1} = \mathbf{1}$ are implied by the *linear* constraints, so they are left out for brevity. By dropping the bilinear constraint, we obtain the LP relaxation:

$$\begin{array}{rll} \text{minimize:} & \sum_{y,u} C_{yu} K_{yu} & \\ \text{subject to:} & \sum_{u_1} K_{yu} = K^2_{y_2 u_2} & \text{for all } y_1 \in Y_1 \\ & \sum_{u_2} K_{yu} = K^1_{y_1 u_1} & \text{for all } y_2 \in Y_2 \\ & K^i \mathbf{1} = \mathbf{1} & i = 1, 2 \\ & K \geq 0 & \end{array} \tag{3}$$

Solving this linear program produces a lower bound on the minimum value achievable by a decentralized policy, as well as suboptimal policies described by K^1 and K^2. When the relaxation is not exact, the policies K^1 and K^2 may not be person-by-person optimal. However, we can always apply the iterative scheme described at the end of Section 4.1 using these policies as a starting point to obtain improved deterministic equilibrium policies.

5 Numerical example

Here we illustrate some of the concepts discussed in this paper with a numerical example. Consider a decentralized detection problem with four hypotheses and two detectors. Let H denote the current hypothesis. The a-priori probabilities for each hypothesis are given by:

$$\mathbf{Prob}\{H = h_i\} = \begin{cases} 0.39 & \text{for } i = 1 \\ 0.31 & \text{for } i = 2 \\ 0.16 & \text{for } i = 3 \\ 0.14 & \text{for } i = 4 \end{cases}$$

The measurements M_1 and M_2 are made by each detector are each quantized to one of ten measurements. The conditional probabilities of each possible pair of measurements given each hypothesis are illustrated by the figure below.

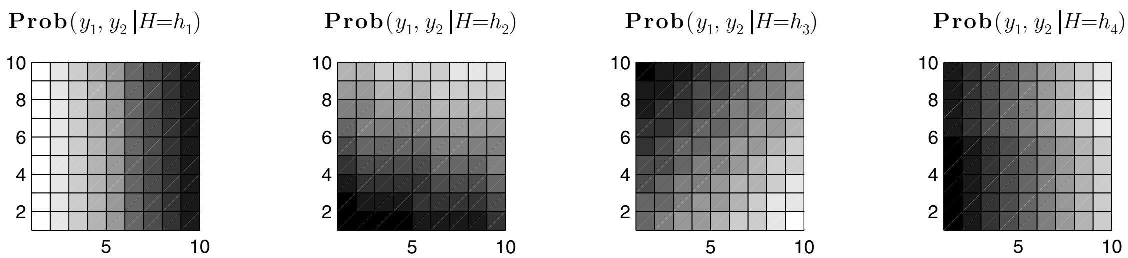


Figure 1: Conditional probabilities of each pair of measurements given each hypothesis. Dark areas on the plots represent low probabilities.

Each detector will estimate the hypothesis based only on its own observation. We would like to find decentralized detection rules which maximize the probability that *at least one detector is correct.*

We can formulate a relaxation of this problem as the linear program (3). In this case, the costs are

$$C_{yu} = \mathbf{Prob}\{(H \neq u_1) \cap (H \neq u_2) \cap (M_1 = y_1) \cap (M_2 = y_2)\}$$

Solving the relaxation, we obtain the globally optimal detection rules:

$$\gamma_1^*(y_1) = \begin{cases} 3 & \text{for } y_1 \leq 3 \\ 2 & \text{otherwise} \end{cases} \qquad \gamma_2^*(y_2) = \begin{cases} 1 & \text{for } y_2 \leq 7 \\ 4 & \text{otherwise} \end{cases}$$

The optimal strategy achieves $\mathbf{Prob}\{\text{at least one correct}\} = 0.77$.

It is interesting to compare the optimal strategy to the one obtained by using a maximum a-posteriori detection rule for each detector. For each detector, the a-posteriori probabilities of the true hypothesis given each measurement are shown below

We can observe from the plots that the MAP strategy is identical for both detectors and is given by

$$\gamma_i^{\text{MAP}}(y_i) = \begin{cases} 1 & \text{for } y_i \leq 6 \\ 2 & \text{otherwise} \end{cases}$$

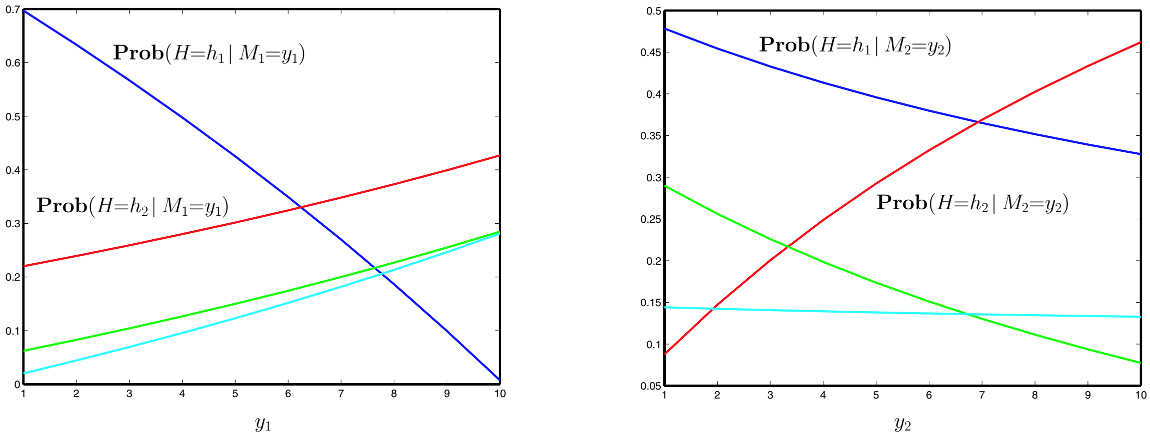


Figure 2: A-posteriori probabilities of the truth of each hypothesis given each measurement. The most probable hypotheses (1 and 2) are labelled.

The MAP strategy achieves **Prob**{at least one correct} $= 0.62$. The key difference between the optimal strategy and the MAP strategy is the element of *hedging* employed by the optimal strategy. That is, the first hypothesis is the most likely, and it is most reliably detected by the second detector. In the optimal strategy, the first detector never guesses the first hypothesis. This is done to maximize the probability of guessing the correct hypothesis when the second detector guesses incorrectly. In the MAP strategy, both detectors are often both guessing the first hypothesis. When one is incorrect, the other is likely to be incorrect as well.

6 Conclusions

In this paper, we considered the problem of determining optimal decentralized decision rules in stochastic decision problems. It was shown that a general discrete decision problem has an equivalent formulation as a polynomial optimization problem. We obtain a relaxation of this polynomial optimization problem which can be used to compute suboptimal policies as well as bounds on the optimal achievable value.

Here we restricted our treatment to the problem of two decision makers. This was done mostly to simplify notation, and all results discussed can readily be extended to cases with more than two decision makers. When considering cases with more than two decision makers, we may want to consider information structures more general than perfectly decentralized structures. Arbitrary information structures can be handled by methods like those discussed, although we have not discussed this issue here.

References

[1] Y.C. Ho. Team decision theory and information structures. *Proceedings of the IEEE*, 68(6):644–654, 1980.

[2] J. Marschak and R. Radner. *The economic theory of teams.* Yale University Press, New Haven, CT., 1972.

[3] C. Papadimitriou and J. Tsitsiklis. The complexity of Markov decision processes. *Mathematics of operations research*, 12(3):441–450, 1987.

[4] R. Radner. Team decision problems. *Annals of Mathematical Statistics*, 33(3):857–881, 1962.

[5] H. Sherali and W. Adams. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM Journal on Discrete Mathetmatics*, 3:555–555, 1990.

[6] R. Tenney and N. Sandell. Detection with distributed sensors. *IEEE Trans. on Aerospace and Elec. Sys.*, 17(4):501–510, 1981.

[7] J. Tsitsiklis. *Decentralized Detection*, volume 2, pages 297–344. JAI Press, 1993. H.V. Poor and J.B. Thomas, editors.

[8] J. Tsitsiklis and M. Athans. On the complexity of decentralized decsion making and detection problems. *IEEE Trans. Automatic Control*, 30(5):440–446, 1985.

[9] P. Varshney. *Distributed detection and data fusion.* Springer, New York, 1997.