Computing Minimax Decisions with Incomplete Observations

Thijs van Ommen T.VANOMMEN@UVA.NL
Universiteit van Amsterdam
Amsterdam (The Netherlands)

Abstract

Decision makers must often base their decisions on incomplete (coarse) data. Recent research has shown that in a wide variety of coarse data problems, minimax optimal strategies can be recognized using a simple probabilistic condition. This paper develops a computational method to find such strategies in special cases, and shows what difficulties may arise in more general cases.

Keywords: coarse data; incomplete observations; minimax decision making.

1. Introduction

Suppose that we are faced with a decision where the loss we will incur depends on the outcome x of a random experiment. While the distribution of x is known, our observation of x is incomplete: we only have access to a *coarse* observation y, a set that we know includes x but may also include other elements. An infamous example of this problem is the Monty Hall puzzle (Selvin, 1975).

Example 1 *In a game show, a car is hidden uniformly at random behind one of three doors. The contestant picks a door; we will assume the middle one. But now the quizmaster steps in and opens one of the two remaining doors, revealing a goat behind it. Should the contestant switch to the remaining door, or stick with his initial guess?*

We will make the standard assumptions that the quizmaster always opens a door, always with a goat behind it. Then this is an instance of the incomplete data problem, where we will either observe $y = \{\text{left}, \text{middle}\}$ (if the quizmaster opens the rightmost door) or $y = \{\text{middle}, \text{right}\}$ (if he opens the leftmost door). It is well known—but quite surprising—that it is wrong to conclude the remaining doors now each have probability $1/2$ of hiding the car. But then what probability distribution (or set of distributions) should we use to base our decision on?

A key issue here is that we do not know the *coarsening mechanism*, the random process that maps the true outcome x to the set y we observe. A common assumption about this mechanism is *coarsening at random (CAR)*, which says that for each set y, the probability that the coarsening mechanism reports y is the same no matter which outcome $x \in y$ is the true outcome (Heitjan and Rubin, 1991). But this is a strong assumption that often fails to hold in practice; in fact, in the Monty Hall puzzle, it can never hold (Grünwald and Halpern, 2003; Gill and Grünwald, 2008).

An approach that avoids any assumptions on the coarsening mechanism is to model the problem using the credal set $\mathcal{P}$ of all joint distributions P on (x, y) that are (a) consistent with the known distribution of x, and (b) satisfy $P(x, y) = 0$ for $x \notin y$. This (convex) set $\mathcal{P}$ represents both our aleatory uncertainty about x and our epistemic uncertainty about its relation with y. To then incorporate an observation y, the generalized Bayes rule can be used; De Cooman and Zaffalon (2004) apply this approach to coarse data problems. The resulting posterior on the outcomes exhibits *dilation* (Seidenfeld and Wasserman, 1993): the prior was a precise distribution, but the posterior may be a large set of distributions. If we want to be sure that the true distribution of x given y is included in this set, then this phenomenon is unavoidable. However, it may lead to suboptimal decisions, as described by Augustin (2003), Grünwald and Halpern (2011), and others.

By formulating a strategy before making an observation, the effect of dilation on decisions can be avoided (Seidenfeld, 2004). This approach has been investigated for coarse data problems by Van Ommen et al. (2016), who found that for many situations, minimax strategies are characterized by the *RCAR condition* (which looks like the CAR condition, but with x and y reversed).

To apply these results in practice, we would like efficient computational methods to find RCAR strategies. How difficult this is depends largely on the family of possible observations. In this paper, we describe a computational method for a restricted class of such families. This reveals a relation between minimax optimal strategies and statistical independence. We also point out the various computational difficulties that may occur in larger classes of families: there finding an exact solution may involve a combinatorial search, or solving polynomial (rather than linear) equations.

This paper is structured as follows. In Section 2, we summarize the relevant results of Van Ommen et al. (2016). Section 3 introduces the main tool: homogeneous induced colourings. These may not exist for all families of possible observations, which leads to a categorization of such families. A computational procedure, and its limitations, are described in Section 4. Section 5 interprets this procedure for the families where it is guaranteed to work. Section 6 concludes.

The contents of Sections 3 to 5 are adapted from Chapter 7 of PhD thesis (Van Ommen, 2015).

2. Optimality of RCAR Strategies

This section summarizes the main results from Van Ommen et al. (2016). We consider coarse data decision problems with finite outcome space $\mathcal{X}$. The decision maker must pick an action based on a coarse observation $y \subseteq \mathcal{X}$, which we call a *message*. The choice of action will depend not only on the received message, but on the entire set of messages that the coarsening mechanism might produce: the *message structure* $\mathcal{Y} \subset 2^{\mathcal{X}}$ (in the Monty Hall example, this is $\mathcal{Y} = \{\{\text{left}, \text{middle}\}, \{\text{middle}, \text{right}\}\}$). It will also depend on the (known) distribution p of the outcomes, which we assume to be nowhere zero, and on the loss function $L : \mathcal{X} \times \mathcal{A} \to [0, \infty]$.[1] We assume throughout that L satisfies the technical conditions in (Van Ommen et al., 2016, Theorem 3); these are in particular satisfied for all finite L, and also for logarithmic loss $L(x, Q) = -\log Q(x)$.

The decision problem is modelled as a zero-sum game between a quizmaster and a contestant. The (imaginary) quizmaster picks as strategy a joint distribution P on $\mathcal{X} \times \mathcal{Y}$ from the credal set $\mathcal{P} = \{P \mid \sum_y P(x, y) = p_x \text{ for all } x, P(x, y) = 0 \text{ for all } y \in \mathcal{Y}, x \notin y\}$. Simultaneously, the contestant picks as strategy a function $A : \mathcal{Y} \to \mathcal{A}$. The two players seek to maximize resp. minimize the expected loss

$$\sum_{x,y} P(x, y) L(x, A(y)) = \sum_{x} p_x \sum_{y \in \mathcal{Y}, y \ni x} P(y \mid x) L(x, A(y)),$$

where the second expression reflects that the quizmaster's influence is limited to $P(y \mid x)$, with x always sampled from the fixed marginal p. Strategies achieving this maximum/minimum are called *worst-case optimal*. If the action space is rich enough, this game has a Nash equilibrium; then

1. In (Van Ommen et al., 2016), the decision maker's action space $\mathcal{A}$ is always taken to be the set of distributions on $\mathcal{X}$, and the actions are interpreted as probability updates. But due to the generality of the loss functions allowed, the same theory can be applied for arbitrary action spaces.

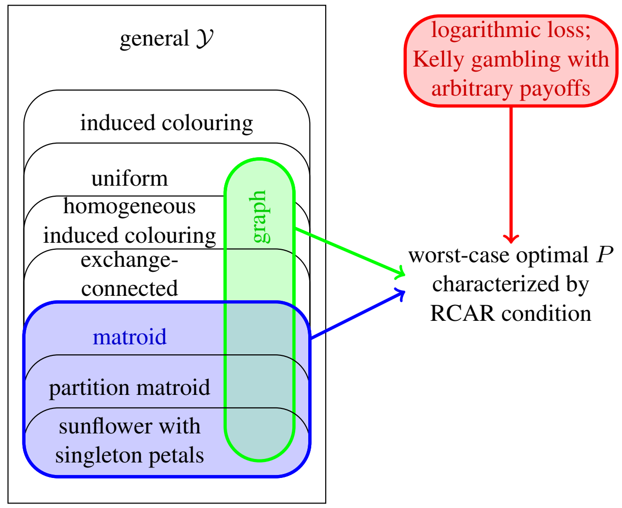

Figure 1: Overview of classes of message structures. As shown by Van Ommen et al. (2016), the RCAR condition characterizes worst-case optimal P for many games, including all graph and matroid games. The other classes shown in this figure are defined and explored in Sections 3 and 5.

given a worst-case (maximin) optimal P for the quizmaster, we can easily determine a worst-case (minimax) optimal A for the contestant.[2] Thus we focus on finding a worst-case optimal P.

For several classes of games, a strategy P is worst-case optimal if, for some vector $q \in [0,1]^{\mathcal{X}}$, it satisfies the *RCAR condition*:

$$\begin{aligned} q_x = P(x \mid y) \text{ for all } y \in \mathcal{Y} \text{ with } P(y) > 0 \text{ and all } x \in y, \text{ and} \\ \sum_{x \in y} q_x \leq 1 \text{ for all } y \in \mathcal{Y}. \end{aligned} \tag{1}$$

Such a strategy is called an *RCAR strategy*, and the vector q is called an *RCAR vector*. The three classes of games where this holds are illustrated in Figure 1: if L is logarithmic loss or an affine transformation of it (this represents Kelly gambling games with arbitrary payoffs); if $\mathcal{Y}$ is a graph (each message consists of two outcomes); and if $\mathcal{Y}$ is a matroid (defined in (3) below). In the latter two cases, there is also a symmetry condition on the loss function L. What is surprising here is that, as long as we are in one of these cases, we can find a worst-case optimal P without knowing what the loss function is, because the RCAR condition is purely probabilistic and does not depend on L. The rest of this paper deals with the problem of computing an RCAR strategy for a given game.

3. Induced Colourings

Fix a set $\mathcal{Y}' \subseteq \mathcal{Y}$ with $\bigcup_{y \in \mathcal{Y}'} y = \mathcal{X}$, and assume that an RCAR strategy P exists with support $\mathcal{Y}_P := \{y \in \mathcal{Y} \mid P(y) > 0\}$ equal to $\mathcal{Y}'$. (For example, we may in many cases take $\mathcal{Y}' = \mathcal{Y}$.) We will now consider different properties of $\mathcal{Y}_P$ that may help us find P. The classes of message

2. In many cases, $A(y)$ is simply the optimal response to $P(\cdot \mid y)$ for each $y \in \mathcal{Y}$. This is not always well-defined; a general solution is given by Theorem 7 of (Van Ommen et al., 2016), using Theorem 3 to determine λ^*.

structures defined by these properties, and the inclusion relations between them that we establish here, are shown graphically in Figure 1, and examples are given in Figure 2.

Consider the system of linear equations

$$\sum_{x \in y} q_x = 1 \text{ for all } y \in \mathcal{Y}'. \tag{2}$$

If an RCAR strategy $P \in \mathcal{P}$ exists with support $\mathcal{Y}'$ and RCAR vector q, then q is positive and satisfies (2). The converse is not true: if $\mathcal{Y}' \neq \mathcal{Y}$, then (1) additionally imposes inequalities on messages $y \in \mathcal{Y} \setminus \mathcal{Y}'$. We will start our search for RCAR strategies by examining the solutions of (2). (A similar system is studied in the CAR literature, where it plays a role in characterizing message structures that admit a CAR coarsening mechanism; see Grünwald and Halpern (2003); Jaeger (2005); Gill and Grünwald (2008). Since we study RCAR rather than CAR, the roles of outcomes and messages are reversed here.)

Define a *colouring* as a partition of $\mathcal{X}$. We say a colouring is *induced* by a set of messages $\mathcal{Y}'$ if the system of linear equations (2) has at least one solution q with $q_x > 0$ for all x, and x, x' are in the same class of the colouring ('have the same colour') if and only if $q_x = q_{x'}$ for all such solutions to that system (in other words, the colour classes are the equivalence classes of this relation on $\mathcal{X}$). If the system has at least one positive solution, then the colouring induced by $\mathcal{Y}'$ is unique; otherwise, there is no induced colouring.

We say a colouring is *homogeneous on* $\mathcal{Y}'$ if the number of outcomes of each colour is the same for every message in $\mathcal{Y}'$ (for example, if each message consists of one 'red' and two 'blue' outcomes). This is only possible if $\mathcal{Y}'$ is *uniform*: all messages in $\mathcal{Y}'$ have the same size. We are interested in $\mathcal{Y}'$ whose induced colouring is homogeneous. One class of such $\mathcal{Y}'$ is defined in terms of pairs of messages y_1, y_2 that differ by the *exchange* of one outcome, meaning that $|y_1 \backslash y_2| = |y_2 \backslash y_1| = 1$. We call $\mathcal{Y}'$ *exchange-connected* if, for each pair of messages $y^{(a)}, y^{(b)} \in \mathcal{Y}'$, there exists a sequence of messages $y_1, y_2, \ldots, y_\ell \in \mathcal{Y}'$ (an *exchange-path*) with $y_1 = y^{(a)}$ and $y_\ell = y^{(b)}$ whose adjacent messages differ by the exchange of one outcome. Finally, $\mathcal{Y}'$ is a *matroid* if it satisfies the *basis exchange* property: for all $y_1, y_2 \in \mathcal{Y}'$ and $x_1 \in y_1 \setminus y_2$,

$$(y_1 \setminus \{x_1\}) \cup \{x_2\} \in \mathcal{Y}' \text{ for some } x_2 \in y_2 \setminus y_1. \tag{3}$$

Figure 2 illustrates these definitions with a few examples. Each table represents a message structure $\mathcal{Y}$ as an incidence matrix: each row represents a message, and (coloured) stars mark the outcomes it contains.

The message structure shown in Figure 2a has no induced colouring: any solution of (2) must have $q_{x_3} = 1 - q_{x_4} = q_{x_5} = 1 - q_{x_1}$ and thus $q_{x_2} = 0$, so there is no positive solution, and it follows that no RCAR strategy P exists with $P(y) > 0$ for all $y \in \mathcal{Y}'$. On the other hand, any uniform game has an induced colouring, because there is at least one solution to (2):

$$q_x = 1/k \quad \text{for all } x \in \mathcal{X}, \tag{4}$$

where k is the size of the game's messages.

Figures 2b and 2c are examples of message structures that do have an induced colouring, but one that is not homogeneous. In both these examples, all outcomes have different colours in the induced colouring, because no pair of outcomes necessarily has the same value of q in a solution of (2). The message structure shown in Figure 2c will be revisited in Example 2 in the next section.

	x_1	x_2	x_3	x_4	x_5
y_1	$*$	$*$	$*$	—	—
y_2	—	—	$*$	$*$	—
y_3	—	—	—	$*$	$*$
y_4	$*$	—	—	—	$*$

(a) No induced colouring.

	x_1	x_2	x_3	x_4
y_1	$*$	$*$	—	—
y_2	—	$*$	$*$	$*$

(b) Induced colouring but not uniform.

	x_1	x_2	x_3	x_4	x_5	x_6
y_1	$*$	$*$	$*$	—	—	—
y_2	—	—	$*$	$*$	$*$	—
y_3	$*$	—	—	—	$*$	$*$

(c) Uniform but induced colouring not homogeneous.

	x_1	x_2	x_3	x_4	x_5	x_6
y_1	$*$	$*$	$*$	—	—	—
y_2	—	—	$*$	$*$	$*$	—
y_3	$*$	—	—	—	$*$	$*$
y_4	—	$*$	—	$*$	—	$*$

(d) Homogeneous induced colouring but not exchange-connected.

	x_1	x_2	x_3	x_4	x_5
y_1	$*$	$*$	$*$	—	—
y_2	—	$*$	$*$	$*$	—
y_3	—	—	$*$	$*$	$*$

(e) Exchange-connected but not a matroid.

	x_1	x_2	x_3	x_4	x_5
y_1	$*$	$*$	$*$	—	—
y_2	$*$	$*$	—	$*$	—
y_3	$*$	—	$*$	$*$	—
y_4	$*$	—	$*$	—	$*$
y_5	$*$	—	—	$*$	$*$

(f) Matroid.

Figure 2: Examples of messages structures and their induced colourings.

The three remaining message structures do have homogeneous induced colourings. Figure 2d shows that it is possible for a message structure to have a homogeneous induced colouring without being exchange-connected. In this message structure, which adds the message y_4 to the structure in Figure 2c, each pair of messages differs by *two* exchanges. Yet the added message changes the induced colouring: for example, $q_{x_1} = q_{x_4}$ follows because by the equalities from (2) on y_1 and y_3, $1 - q_{x_1} = q_{x_2} + q_{x_3} = q_{x_5} + q_{x_6}$, and by y_2 and y_4, $1 - q_{x_4} = q_{x_3} + q_{x_5} = q_{x_2} + q_{x_6}$; thus $2 - 2q_{x_1} = 2 - 2q_{x_4} = q_{x_2} + q_{x_3} + q_{x_5} + q_{x_6}$.

The message structure shown in Figure 2e is exchange-connected. For such structures, it easy to determine the induced (homogeneous) colouring: if messages y_1, y_2 differ by the exchange of one outcome (x_1 for x_2), then any solution of (2) must satisfy $q_{x_1} = q_{x_2}$, so such x_1, x_2 must be the same colour. Any vector q that satisfies all these equalities and satisfies $\sum_{x \in y} q_x = 1$ for any one message $y \in \mathcal{Y}'$ satisfies (2) for all messages in $\mathcal{Y}'$, so this determines the induced colouring. This colouring is clearly homogeneous on any pair of message that differ by the exchange of one outcome; because exchange-paths exist between all pairs of messages, it follows that the induced colouring of an exchange-connected game is homogeneous.

Finally, the class of matroid games is a subclass of exchange-connected games: (3) requires the existence of not just one, but possibly many different exchange-paths between any pair of messages. Figure 2f gives an example. The structure in Figure 2e is not a matroid: there is no outcome in $y_3 \setminus y_1$ that can be added to $y_1 \setminus \{x_2\} = \{x_1, x_3\}$ to make a message.

The following lemma gives two alternate characterizations of the induced colouring of a matroid. The first of these is in terms of a concept from matroid theory: the colour classes of the induced colouring coincide with the *2-connected components* of the matroid. (We refer to Oxley (2011) for the definition.) We observed above (when discussing Figure 2e) that if messages exist that differ in the exchange of one outcome, then the outcomes being exchanged must be the same colour. The second characterization shows that for matroids, the converse also holds.

Lemma 1 (Matroid colouring) *Given a matroid $(\mathcal{X}, \mathcal{Y})$ and two elements $x_1, x_2 \in \mathcal{X}$, the following statements are equivalent:*

1. *x_1 and x_2 are in the same colour class of the induced colouring of $\mathcal{Y}$;*

2. *x_1 and x_2 are in the same 2-connected component of $(\mathcal{X}, \mathcal{Y})$;*

3. *There exist $y_1, y_2 \in \mathcal{Y}$ such that $y_1 \setminus y_2 = \{x_1\}$ and $y_2 \setminus y_1 = \{x_2\}$.*

4. A Computational Procedure for Finding RCAR Strategies

Consider the case that $\mathcal{Y}'$ induces a homogeneous colouring, and assume as before that an RCAR strategy P exists with $\mathcal{Y}_P = \mathcal{Y}'$. Then the corresponding RCAR vector q must be a solution of the linear system (2). Additionally, P must agree with the marginal p. These constraints allow us to compute the vector q directly.

Let S be the set of all outcomes with a particular colour. Then there is some value q_S such that $P(x \mid y) = q_x = q_S$ for all $y \in \mathcal{Y}$, $x \in S \cap y$. Let $k_S = |S \cap y|$ (this is independent of y by homogeneity). We must have

$$k_S q_S = k_S q_S \sum_{y} P(y) = \sum_{y} k_S P(y) q_S = \sum_{y} \sum_{x \in S \cap y} P(y) q_S = \sum_{x \in S} \sum_{y \ni x} P(y) P(x \mid y) = \sum_{x \in S} p_x,$$

so that q_S can be computed by

$$q_S = \frac{1}{k_S} \sum_{x \in S} p_x. \tag{5}$$

A simple case is when the induced colouring assigns the same colour to all outcomes: then we see that as in (4), we get $q_x = 1/k$ for all $x \in \mathcal{X}$, where k is the size of the messages. When a colour consists of just one outcome x (which must then be an element of every message for the colouring to be homogeneous), we find $q_x = p_x$.

If an RCAR strategy P exists with $\mathcal{Y}_P = \mathcal{Y}'$ where $\mathcal{Y}'$ induces a homogeneous colouring, then P must have the vector q determined by (5) as its RCAR vector. However, it may be the case that no such strategy exists. To find P if it exists, we still need to determine the $P(y)$'s. We can find a nonnegative solution or determine that no nonnegative solution exists by solving the following linear programming problem (which we can do in polynomial time):

$$\begin{aligned} \text{maximize} \quad & \sum_{y \in \mathcal{Y}} r_y \\ \text{subject to} \quad & \sum_{y \ni x} r_y \leq \frac{p_x}{q_x} \quad \text{for all } x \in \mathcal{X}, \end{aligned} \tag{6}$$

with $r \in \mathbf{R}^{\mathcal{Y}}_{\geq 0}$. If a vector achieving $\sum_{y \in \mathcal{Y}} r_y = 1$ is found, we have a strategy P with r as the marginal on messages ($P(x, y) = q_x r_y$ for all $x \in y$). If no vector r achieves the value 1, there is no RCAR strategy P satisfying the assumption $\mathcal{Y}_P = \mathcal{Y}'$.

Now we may want to apply this procedure in practice to find an RCAR strategy for a given game. (By Lemma 11 from Van Ommen et al. (2016), such a strategy always exists.)

When doing so we encounter two problems: we need to provide the procedure with an $\mathcal{Y}'$ such that $\bigcup \mathcal{Y}' = \mathcal{X}$, and even if we have an idea about what $\mathcal{Y}'$ to take, it may not have a homogeneous induced colouring. Still, let us investigate what happens if we just guess an $\mathcal{Y}'$. We will then encounter one of the cases 1, 2a-2c which we now describe. Briefly, in case 1, the procedure cannot be used because it cannot determine q, and in case 2a and 2b it gives an inconclusive result; in case 2c we have success. We now consider each case in detail.

1. *$\mathcal{Y}'$ has no homogeneous induced colouring.*

In this case, the procedure is not applicable. Indeed, finding an RCAR vector may be a more difficult type of problem, as illustrated by the following example which uses the message structure from Figure 2c. (This example is a uniform game; the class of uniform games is the smallest class among those identified in the previous section that strictly contains the class of games with a homogeneous induced colouring.)

Example 2 (Irrational RCAR vector) *Consider the problem with $\mathcal{X} = x_1, \ldots, x_6$, $\mathcal{Y} = \{y_1 = \{x_1, x_2, x_3\}, y_2 = \{x_3, x_4, x_5\}, y_3 = \{x_1, x_5, x_6\}\}$, and marginal p and strategy P given by the following table:*

P	x_1	x_2	x_3	x_4	x_5	x_6
y_1	$1/10$	$1/10$	$\frac{3}{10} - \frac{1}{10}\sqrt{5}$	–	–	–
y_2	–	–	$\frac{1}{10}\sqrt{5} - \frac{1}{10}$	$1/5$	$\frac{1}{10}\sqrt{5} - \frac{1}{10}$	–
y_3	$1/10$	–	–	–	$\frac{3}{10} - \frac{1}{10}\sqrt{5}$	$1/10$
q_x	$\frac{1}{4} + \frac{1}{20}\sqrt{5}$	$\frac{1}{4} + \frac{1}{20}\sqrt{5}$	$\frac{1}{2} - \frac{1}{10}\sqrt{5}$	$\frac{1}{5}\sqrt{5}$	$\frac{1}{2} - \frac{1}{10}\sqrt{5}$	$\frac{1}{4} + \frac{1}{20}\sqrt{5}$
p_x	$1/5$	$1/10$	$1/5$	$1/5$	$1/5$	$1/10$

The strategy P is RCAR, with the vector q that is also shown in the table. We see that the RCAR strategy P and RCAR vector q (both of which are unique) contain irrational numbers, while the marginal p was rational. The solution techniques used in this section (the formula (5) for q and linear optimization for (6)) do not yield irrational results when given rational inputs, so this example shows that these techniques will not suffice in general for games that do not have a homogeneous induced colouring. (General-purpose convex optimization techniques could be used here instead.)

Conclusion: in this case, an RCAR strategy P with $\mathcal{Y}_P = \mathcal{Y}'$ may exist, but it may be not be easy to find. So in general, for such $\mathcal{Y}'$, we do not know how to efficiently determine if such a P exists.

2. *$\mathcal{Y}'$ does have a homogeneous induced colouring.*

In this case, we can use (5) to compute a candidate q for the RCAR vector. We distinguish three subcases:

2a. *If $\mathcal{Y}' \neq \mathcal{Y}$, there may be a message $y \in \mathcal{Y} \setminus \mathcal{Y}'$ for which $\sum_{x \in y} q_x > 1$.*

This may happen because the described procedure ignores the existence of messages not in $\mathcal{Y}'$. However, the RCAR condition (1) puts an inequality constraint on $\sum_{x \in y} q_x$ even for messages y with $P(y) = 0$. If the vector q computed by (5) does not satisfy this constraint, then q is not an RCAR vector: we chose the wrong $\mathcal{Y}'$.

2b. *No solution r of (6) achieves $\sum_{y \in \mathcal{Y}} r_y = 1$.*

This also means that our choice of $\mathcal{Y}'$ was incorrect.

2c. *Otherwise, q is an RCAR vector, and together with r determines an RCAR strategy P.*

In this case, we can report success.

In cases 2a and 2b, $\mathcal{Y}'$ has a homogeneous induced colouring but we find that no RCAR strategy P exists with $\mathcal{Y}_P = \mathcal{Y}'$. Then we may face two problems. First, it is not clear how we might choose a different $\mathcal{Y}'$ on which to try the procedure next. For small message structures, it may be feasible to try all candidates. For larger structures, the number of possible choices grows exponentially, and a more efficient way of searching would be needed.

The second problem is that in general, $\mathcal{Y}'$ might not induce a homogeneous colouring even though $\mathcal{Y}$ does. For example, if $\mathcal{Y}$ is the message structure shown in Figure 2e, but there is no RCAR strategy P with $\mathcal{Y}_P = \mathcal{Y}$ for our marginal p, we have to conclude that the RCAR strategy must have $\mathcal{Y}_P = \{y_1, y_3\}$ (because this is the only other choice of $\mathcal{Y}'$ that satisfies $\bigcup \mathcal{Y}' = \mathcal{X}$). However, this message structure is no longer exchange-connected, and in fact does not have a homogeneous induced colouring, so that we end up in case 1.

In Section 5, we will see a subclass of matroid games for which the procedure is guaranteed to succeed for the choice $\mathcal{Y}' = \mathcal{Y}$. So for that class of inputs, the procedure discussed here is an efficient algorithm for finding an RCAR strategy (which is worst-case optimal for any loss function by the results of Van Ommen et al. (2016)).

Two efficient algorithms, for graph games and for matroid games, are given in Van Ommen (2015, Chapter 8). These algorithms can also be viewed as instances of the computational procedure in this section: both essentially compute q and r as we did here; then, if $\sum_{y \in \mathcal{Y}} r < 1$, they pick a new set $\mathcal{Y}'$, guided by properties of the linear optimization problem (6). The choice of $\mathcal{Y}'$ is such that each new $\mathcal{Y}'$ is a subset of the previous $\mathcal{Y}'$ (i.e. no backtracking is needed), and such that case 2a will never occur.

Case 1 will never occur either for these algorithms: the chosen $\mathcal{Y}'$ will always have a homogeneous induced colouring. This happens for different reasons for the two cases of graph and matroid games. These reasons shed light on what makes graphs and matroids special among more general message structures, so we conclude this section by giving brief explanations.

For graphs: Any connected component of a graph is also exchange-connected, and thus induces a homogeneous colouring. While some choices of $\mathcal{Y}'$ may produce a disconnected graph $(\mathcal{X}, \mathcal{Y}')$, each component of this graph will have a homogeneous induced colouring, and the algorithm can be applied to each of these components recursively.

For matroids: On a matroid game, for any RCAR strategy P, $\mathcal{Y}_P$ determines a homogeneous colouring. (This colouring is not induced in the usual sense, but is uniquely determined by the equalities on $\mathcal{Y}_P$ combined with inequalities for $\mathcal{Y} \setminus \mathcal{Y}_P$; see (Van Ommen et al., 2016, proof of Theorem 19) for details.) The conditional probabilities $P(x \mid y)$ respect this colouring.

5. Partition Matroids

We now describe a class of games for which a worst-case optimal strategy can be completely computed using the procedure from the previous section, because regardless of the marginal p, we can take $\mathcal{Y}' = \mathcal{Y}$ and the procedure will succeed in finding an RCAR strategy.

A message structure $\mathcal{Y}$ is called a *partition matroid* if $\mathcal{X}$ can be partitioned into nonempty sets $S_1, \ldots, S_k$ such that $\mathcal{Y}$ consists of all subsets of $\mathcal{X}$ that take one element from each of the sets S_i (Oxley, 2011). This class forms a subclass of matroids, so if $\mathcal{Y}$ is a partition matroid, it induces a homogeneous colouring. Using Lemma 1, it is easy to see that this colouring is given by the sets S_i. An example of a partition matroid is given in Figure 3a; the matroid we saw in Figure 2f is not a partition matroid.

	x_1	x_2	x_3	x_4	x_5
y_1	$*$	—	$*$	—	—
y_2	$*$	—	—	$*$	—
y_3	$*$	—	—	—	$*$
y_4	—	$*$	$*$	—	—
y_5	—	$*$	—	$*$	—
y_6	—	$*$	—	—	$*$

(a) Partition matroid but not a sunflower.

	x_1	x_2	x_3	x_4	x_5
y_1	$*$	$*$	$*$	—	—
y_2	$*$	$*$	—	$*$	—
y_3	$*$	$*$	—	—	$*$

(b) Sunflower with singleton petals.

Figure 3: More examples of messages structures and their induced colourings.

As an illustration, suppose a shopkeeper sells items of brands x_1 and x_2, in colours x_3, x_4 and x_5, and customers buy items based on a preference for either a brand or a colour. The shopkeeper observes a customer buying an item, but would like to know the underlying preference for recommendation purposes. This coarse data problem corresponds to the partition matroid in Figure 3a.

Because a partition matroid induces a homogeneous colouring, we can carry out the procedure described in the previous section to find for each x that $q_x = \sum_{x' \in S_i} p_{x'}$, where S_i is the set containing x. Now a solution for the $P(y)$'s that satisfies $\sum_{y \ni x} P(y) q_x = p_x$ always exists:

$$P(y) = \prod_{x \in y} \frac{p_x}{q_x}.$$

In words, this means that given the true outcome x, it is worst-case optimal for the quizmaster to choose a message by randomly sampling an outcome from each set $S_i \not\ni x$ according to the marginal probabilities conditioned on S_i, and give the message consisting of x and these outcomes. The existence of this strategy shows that, for partition matroid games, the procedure always succeeds in finding a worst-case optimal strategy for the choice $\mathcal{Y}' = \mathcal{Y}$.

Example 1 (continued) *The message structure $\mathcal{Y} = \{\{\mathit{left}, \mathit{middle}\}, \{\mathit{middle}, \mathit{right}\}\}$ in the Monty Hall puzzle is a partition matroid with sets $\{\mathit{left}, \mathit{right}\}$ and $\{\mathit{middle}\}$. For an arbitrary prior p on the three doors, the RCAR strategy and vector are given by*

$$q_{\mathit{left}} = q_{\mathit{right}} = p_{\mathit{left}} + p_{\mathit{right}}; \qquad q_{\mathit{middle}} = p_{\mathit{middle}};$$
$$P(\{\mathit{left}, \mathit{middle}\}) = p_{\mathit{left}}/(p_{\mathit{left}} + p_{\mathit{right}}); \qquad P(\{\mathit{middle}, \mathit{right}\}) = p_{\mathit{right}}/(p_{\mathit{left}} + p_{\mathit{right}}).$$

What does a message Y generated by this strategy tell the contestant about the true (random) outcome X? Clearly, it means that *if* $X \in S_i$ for some i, then X must be the unique outcome in $Y \cap S_i$. Of course, the contestant does not know which of these sets contains X. Write I for the (random) index of the set containing X. Does Y tell the contestant anything about I? The answer is no: For each index i, regardless of whether $I = i$, the outcome in $Y \cap S_i$ will be randomly distributed according to the marginal p conditioned on S_i, independently of $Y \cap S_j$ for $j \neq i$. This implies that Y is independent of I. Then for each outcome $x \in Y$, the probability that $X = x$ given message Y equals the probability that $I = i$, where i is the index of the set containing x. These are exactly the probabilities that appear in the RCAR vector q. We know from (Van Ommen et al., 2016, Theorem 19) that the same is true also if the quizmaster is using a worst-case optimal strategy different from the one described above.

In more general message structures, there may be a message that must be excluded from $\mathcal{Y}'$, so that the worst-case optimal P cannot be computed so easily:

Theorem 2 *If a game induces a homogeneous colouring but is not a partition matroid, then there exist a marginal p and a message $y \in \mathcal{Y}$ such that $P(y) = 0$ for all RCAR strategies P.*

We distinguish one subclass of the class of partition matroid games. A message structure in which the intersection of any two messages is constant is called a *sunflower* (Jukna, 2001). The common intersection is called the *core*, and each set difference between a message and the core is called a *petal*. An example of a *sunflowers with singleton petals* is shown in Figure 3b. The Monty Hall game itself (Example 1) is another example.

If a message structure is a sunflower with singleton petals, it is a partition matroid: each outcome in the core forms a (singleton) class of the partition, and another class contains all the petals. Among partition matroids, sunflowers can be recognized by the property that all of its colour classes except one are singleton outcomes. For this class of games, the strategy P described above is the *unique* RCAR strategy: a strategy P' with $P'(y) \neq P(y)$ for some $y \in \mathcal{Y}$ would disagree with the unique RCAR vector.

The message structure shown in Figure 3a is a partition matroid, but not a sunflower. Because at least two of its colour classes are not singletons, such a message structure contains a cycle of four messages in which neighbouring messages differ by the exchange of one outcome, but the pairs of messages on opposite sides of the cycle differ by two outcomes. (In Figure 3a, there are three such cycles; one is (y_1, y_2, y_5, y_4).) For this type of game, the strategy P found above can be modified by increasing $P(y)$ for two messages at opposite sides of the cycle, and decreasing it by the same amount for the other two, leaving the conditionals unchanged. Thus P is not the unique RCAR strategy. In fact, RCAR strategies exist with $P(y) = 0$ for some $y \in \mathcal{Y}$. For such a strategy P, we have $\mathcal{Y}_P \subsetneq \mathcal{Y}$, but we do still have $\sum_{x \in y} q_x = 1$ even for messages y with $P(y) = 0$.

6. Conclusion

We have presented an efficient algorithm for finding the minimax optimal strategy in a coarse data problem where the message structure is a partition matroid. While this problem could also be solved using general-purpose convex optimization algorithms, this would be much less efficient. We have also seen how RCAR strategies may be qualitatively different beyond partition matroids, suggesting that in the general case, exact computation of these strategies may be a harder problem.

Acknowledgments

I thank Peter Grünwald and Wouter Koolen for the enjoyable collaboration which led to this paper, and the anonymous reviewers for their valuable feedback. This research was supported by Vici grant 639.073.04 from the Netherlands Organization for Scientific Research (NWO).

Appendix A. Proofs

Proof [Lemma 1] ($2 \Leftarrow 3$) Two elements $x_1 \neq x_2$ of $\mathcal{X}$ are in the same 2-connected component if and only if there is a *circuit* (minimal dependent set) containing both (Oxley, 2011). Since a *basis* $y \in \mathcal{Y}$ is a maximal independent set, $y_1 \cup y_2$ is dependent. Find a circuit $C \subseteq y_1 \cup y_2$; this circuit contains both x_1 and x_2, as otherwise it would be contained in a basis and thus independent.

($2 \Rightarrow 3$) Let C be a circuit with $\{x_1, x_2\} \subseteq C$; our goal is to find the bases y_1, y_2, which we will do iteratively. Let y_1 be a basis containing the independent set $C \setminus \{x_2\}$, and y_2 a basis containing $C \setminus \{x_1\}$. While $y_1 \setminus \{x_1\} \neq y_2 \setminus \{x_2\}$, pick any $x'_1 \in y_1 \setminus (y_2 \cup \{x_1\})$ and use basis exchange to find a basis $y' = (y_1 \setminus \{x'_1\}) \cup \{x'_2\}$ for some $x'_2 \in y_2 \setminus y_1$. Note that $x'_2 \neq x_2$, as that would result in $C \subseteq y'$. Replace y_1 by y' and repeat until $y_1 \setminus \{x_1\} = y_2 \setminus \{x_2\}$. This process terminates, as the set difference becomes smaller with each step.

($1 \Leftrightarrow 3$) For exchange-connected message structures, the colour classes are the equivalence classes of the transitive reflexive closure of the relation on $\mathcal{X}$ stated in point 3. For matroids, the equivalence of points 2 and 3 shows that this relation is already transitive. Thus for all $x_1 \neq x_2$, points 1 and 3 are equivalent. ■

Proof [Theorem 2] We will construct a marginal p with the required property by first finding a vector q that is the RCAR vector for some game with the given message structure. We distinguish two cases. If there exists $y' \subset \mathcal{X}$ that is consistent with the homogeneous induced colouring but $y' \notin \mathcal{Y}$, then pick $0 < \epsilon < 1/(k(k-1))$ and set initial values for q as

$$q_x = \begin{cases} \frac{1}{k} + \epsilon & \text{for } x \in y'; \\ \frac{1}{k} - (k-1)\epsilon & \text{otherwise.} \end{cases}$$

Each message contains at least one outcome with the smaller q_x, so $\sum_{x \in y} q_x \leq 1$ for all $y \in \mathcal{Y}$.

Otherwise, if $\mathcal{Y}$ is not a partition matroid there must exist a colour class $C \subseteq \mathcal{X}$ for which the number of outcomes of this colour occurring in a message is at least two. Then pick any $x^+ \in C$ and $0 < \epsilon < 1/k$, and initialize q according to

$$q_x = \begin{cases} \frac{1}{k} + \epsilon & \text{for } x = x^+; \\ \frac{1}{k} - \epsilon & \text{for } x \in C \text{ but } x \neq x^+; \\ \frac{1}{k} & \text{otherwise.} \end{cases}$$

Again we see $\sum_{x \in y} q_x \leq 1$ for all $y \in \mathcal{Y}$.

Starting from the values of q determined above, we apply a greedy algorithm that repeatedly increases q_x for some x until none can be increased further, maintaining $\sum_{x \in y} q_x \leq 1$ for all $y \in \mathcal{Y}$. For the resulting vector q, let P be the joint distribution on x, y with $P(y)$ uniform on

$\{y \in \mathcal{Y} \mid \sum_{x \in y} q_x = 1\}$, $P(x \mid y) = q_x$ for all $x \in y$, and $P(x, y) = 0$ elsewhere. This P is an RCAR strategy for the game with marginal $p_x = \sum_{y \ni x} P(x, y)$, and q is the unique RCAR vector.

In the first case, there must exist some $x^- \in \mathcal{X}$ with $q_{x^-} \leq 1/k$. Let C be the colour class containing x^-, and let x^+ be the unique outcome in $C \cap y'$. In the second case, there must exist some $x^- \in C$ with $q_{x^-} \leq 1/k$. Thus in either case, we have two outcomes x^- and x^+ of the same colour C but with $q_{x^-} \leq 1/k < 1/k + \epsilon \leq q_{x^+}$. Because this contradicts the definition of an induced colouring, there must be a message for which q violates the equality (2). This message must have $P(y) = 0$ in any RCAR strategy for the game with message structure $\mathcal{Y}$ and marginal p. ■

References

T. Augustin. On the suboptimality of the generalized Bayes rule and robust Bayesian procedures from the decision theoretic point of view: A cautionary note on updating imprecise priors. In *Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications (ISIPTA)*, pages 31–45, 2003.

R. D. Gill and P. D. Grünwald. An algorithmic and a geometric characterization of coarsening at random. *The Annals of Statistics*, 36:2409–2422, 2008.

P. D. Grünwald and J. Y. Halpern. Updating probabilities. *Journal of Artificial Intelligence Research*, 19:243–278, 2003.

P. D. Grünwald and J. Y. Halpern. Making decisions using sets of probabilities: Updating, time consistency, and calibration. *Journal of Artificial Intelligence Research*, 42:393–426, 2011.

D. F. Heitjan and D. B. Rubin. Ignorability and coarse data. *The Annals of Statistics*, 19:2244–2253, 1991.

M. Jaeger. Ignorability in statistical and probabilistic inference. *Journal of Artificial Intelligence Research*, 24:889–917, 2005.

S. Jukna. *Extremal combinatorics: With applications in computer science*. Springer, Berlin, 2001.

T. van Ommen. *Better predictions when models are wrong or underspecified*. PhD thesis, Mathematical Institute, Faculty of Science, Leiden, 2015.

T. van Ommen, W. M. Koolen, T. E. Feenstra, and P. D. Grünwald. Robust probability updating. *International Journal of Approximate Reasoning*, 74:30–57, 2016.

J. Oxley. *Matroid theory*. Oxford University Press, New York, second edition, 2011.

T. Seidenfeld. A contrast between two decision rules for use with (convex) sets of probabilies: Γ-maximin versus E-admissibility. *Synthese*, 140:69–88, 2004.

T. Seidenfeld and L. Wasserman. Dilation for sets of probabilities. *The Annals of Statistics*, 21: 1139–1154, 1993.

S. Selvin. A problem in probability. *The American Statistician*, 29:67, 1975. Letter to the editor.

G. de Cooman and M. Zaffalon. Updating beliefs with incomplete observations. *Artificial Intelligence*, 159(12):75–125, 2004. ISSN 0004-3702.