A Model of Decision-Making with Sequential Information-Acquisition (Part 2)

James C. MOORE and Andrew B. WHINSTON
Department of Computer Science and Krannert Graduate School of Management, Purdue University, West Lafayette, IN 47907, USA

While most real-life decision are of necessity made with less than perfect information, there is usually some opportunity to acquire additional information regarding the problem at hand before a final decision is made. It is, of course, the recognition of this fact which has led to the importance now attached to the field of Decision Support Systems. On the other hand, the formal analysis of the sort of decision problem for which Decision Support Systems can be useful appears to have lagged behind the developments in applications. In this paper we develope a model of decision-making in which there is available a variety of informational sources (experiments) which can reduce (though generally not eliminate) the uncertainty associated with the final decision. Since the informational sources are available only at some cost (either monetarily or in terms of time, or both), the decision-maker must solve two conceptually distinct problems: (1) developing an optimal information-gathering strategy, and (2) developing an optimal final decision strategy, conditional upon the information obtained during the information-gathering process. A theoretical framework is developed here for the analysis of this general problem, and fairly complete solutions are obtained for some interesting special cases; most notably the computer file search problem.

Keywords: Decision support systems, Decision theory, Dynamic programming, File management.

Andrew B. Whinston is on the faculty of the Krannert Graduate School of Management and the Department of Computer Science at Purdue University. His primary teaching interest is management information systems. His current research interests include data base management and applications of artificial intelligence to economics and management. He has also studied applied economics, regulatory economics, and accounting theory.
Among his numerous publications, he has co-authored two bool 'with C. Holsapple and R. Bonczek), *Foundations of Decision Support Systems* (Academic Press, 1981), and *Micro Database Management – Practical Techniques for Application Development* (Academic Press, 1985). He has been a consultant to various companies, governmental agencies, and international organizations on data processing questions.

North-Holland
Decision Support Systems 3 (1987) 47–72

4. The Categorization Problem

4.1. Basic Model

In Part 1 of this paper that appeared in the previous issue of the DSS Journal the basic model was presented and the structure of an efficient solution was analyzed. In Part 2 we apply the basic model to describe the categorization problem so that more specific results may be obtained.

The special case of Model I which we shall refer to as the 'categorization problem' is characterized by the following assumptions (in addition to those set forth in section 2).

(1) The final decision set, D, can be written in the form $D = \{0, 1, \dots, p\}$, where $p \geqslant 1$ is a positive integer.
(2) There exists a partition of X, $\{X_0, X_1, \dots, X_p\}$, such that ω takes the form [16]

$$\omega(x, d) = \begin{cases} \bar{\omega} > 0 & \text{if } x \in X_d \\ 0 & \text{otherwise} \end{cases}$$

for $(x, d) \in X \times D$.

We may think of $\{X_0, X_1, \dots, X_p\}$ as repre-

[16] Recall that we are supposing throughout the remainder of this paper that the payoff function is linear in monetary return, and thus can be written in the form $\omega^*[x, \delta(B), C(B)] = \omega[x, \delta(B)] - C(B)$.

James C. Moore is Professor of Economics in the Krannert School of Management, Purdue University, West Lafayette, Indiana. His research interests lie in the areas of Micro/mathematical Economics, Welfare Economics, and the Economics of Information and Organization. He received a B.A. in Economics from the University of Nebraska at Omaha in 1960, and his Ph.D. in Economics from the University of Minnesota in 1968. Prior to joining the faculty at Purdue University, he taught at the University of Missouri, Columbia. Dr. Moore is a member of the Americal Economics Association and the Econometrics Society.

senting an (exhaustive) set of categories (or classifications) to which the true state, $\hat{x}$, may belong. There is a constant (constant over categories) positive payoff, $\bar{\omega}$, if $\hat{x}$ is categorized correctly, and a payoff of zero is obtained if $\hat{x}$ is not correctly categorized (or classified). Some examples which seem to fit this formulation reasonably well are (a) chemical analysis of an unknown substance, (b) the game of 'twenty question', and (c) the computer file search problem.[17] We shall develop the computer file search problem in detail in the next subsection.

In dealing with the categorization problem, the function ψ: $\boldsymbol{P}(X) \to [0, 1]$ defined by

$$\psi(B) = \max\{\pi(B \cap X_d) \mid d \in D\} \quad \text{for} \quad B \subseteq X, \tag{1}$$

will be of particular interest.

4.1.1. Proposition. If $\sigma = \langle(\boldsymbol{B}_1, \alpha_1), \ldots, (\boldsymbol{B}_r, \alpha_r), (\boldsymbol{B}_{r+1}, \delta)\rangle$ *is an efficient strategy for* $\boldsymbol{D}$, *then*

$$\Omega(\sigma) = \bar{\omega} \sum_{B \in \boldsymbol{B}_{r+1}} \psi(B). \tag{1}$$

Proof. If σ is an efficient strategy for $\boldsymbol{D}$, then we must have

$$(\forall B \in \boldsymbol{B}_{r+1}): \delta(B) \in D^*(B).$$

From the form of $\omega(\cdot)$ it is clear, however, that for $B \subseteq X$, we will have $d^* \in D^*(B)$ if, and only if

$$(\forall d \in D): \pi(B \cap X_d) \le \pi(B \cap X_{d^*}) \equiv \psi(B).$$

Thus, for each $B \in \boldsymbol{B}_{r+1}$,

$$\begin{aligned}\sum_{x \in B} \phi(x)\omega[x, \delta(B)] &= \sum_{x \in B \cap X_{\delta(B)}} \phi(x)\bar{\omega} \\ &= \bar{\omega} \sum_{x \in B \cap X_{\delta(B)}} \phi(x) \\ &= \bar{\omega}\pi(B \cap X_{\delta(B)}) \\ &= \bar{\omega}\psi(B).\end{aligned}$$

Therefore

$$\begin{aligned}\Omega(\sigma) &= \sum_{B \in \boldsymbol{B}_{r+1}} \sum_{x \in B} \phi(x)\omega[x, \delta(B)] \\ &= \bar{\omega} \sum_{B \in \boldsymbol{B}_{r+1}} \psi(B).\end{aligned}$$

Q.E.D.

There are several special cases of the catogorization problem in which we shall have a particular interest. In order to define the first two (to which we shall not attach special names) we begin by defining

$$X = \{X_0, X_1, \ldots, X_p\}. \tag{2}$$

The following two alternative conditions then define the two basic sub-cases with which we shall be interested

$$X \ge \boldsymbol{B}^A, \quad \text{and} \tag{3}$$

$$\boldsymbol{B}^A \ge X. \tag{4}$$

In the first special case, that in which (3) holds, X is at least as fine as $\boldsymbol{B}^A$; that is

$$(\forall d \in D)(\exists B \in \boldsymbol{B}^A): X_d \subseteq B. \tag{5}$$

The second special case is nearly the opposite of the first; that is, (4) holds if and only if

$$(\forall B \in \boldsymbol{B}^A)(\exists d \in D): B \subseteq X_d \tag{6}$$

(it is, of course, possible for a decision problem to satisfy both (3) and (4); that is, we may have $\boldsymbol{B}^A = X$). Notice that in the situation in which (4) holds, we will have

$$(\forall B \in \boldsymbol{B}^A): \psi(B) = \pi(B). \tag{7}$$

There are also sub-cases of both (3) and (4) which we shall find particularly interesting: (a) the 'only correct guesses count' problem, which we define below, is a sub-case of (3), and (b) the computer file search problem, which we shall introduce in the next subsection, is a special case of (4).

4.1.2. Definition. If $\boldsymbol{D}$ is a categorization problem, we shall say that $\boldsymbol{D}$ has the *only correct guesses count form* iff

(i) $\boldsymbol{D}$ satisfies (3), above,

(ii) There exists a positive integer $m \ge p + 1$, satisfying

$$(\forall B \in \boldsymbol{B}^A): \psi(B) = 1/m, \tag{8}$$

[17] Some kinds of forecasting may fit this model fairly well also; in particular, when a forecaster is employed by a business firm. In connection with this observation, note that our formulation can allow for both qualitative forecasting (prices will go down sharply, prices will go down moderately, prices will remain the same, etc.) and quantitative forecasting as long as we agree that it is reasonable to bound the state space and to round off to the nearest decimal point at some level of significance (in particular, note that we do not rule out the possibility that $X \equiv \{X_0, X_1, \ldots, X_p\} = B^X$).

(iii) $\boldsymbol{D}$ has constant information cost, $c \geq 0$, i.e.,

$$(\forall a \in A_1): c(a) = c. \tag{9}$$

The simplest case in which condition (8) holds is that in which

$$\boldsymbol{X} = \boldsymbol{B}^X \tag{10}$$

In this case the catogories $X_0, X_1, \ldots, X_p$ are simply singleton sets, and the problem has the following interpretation: we know that the true state, $\hat{x}$, is one of $p+1$ possible states, $x_0, x_1, \ldots, x_p$; and we receive a reward of $\bar{\omega} > 0$ if we guess correctly which $x_i = \hat{x}$, and receive nothing if our guess is incorrect.[18] If, in addition to (10), we have

$$\boldsymbol{B}^A = \boldsymbol{B}^X,$$

then condition (8), above, amounts to the assumption that ϕ, the probability density function on X, is the uniform distribution, with

$$\phi(x) = 1/m = 1/(p+1) \quad \text{for each} \quad x \in X.$$

In the case where condition (3) holds, the expected gross payoff function takes on a particularly simple form, as follows.

4.1.3. Proposition. If $\boldsymbol{D}$ is a catogorization problem satisfying condition (3), *above* (*i.e.*, $\boldsymbol{X} \geq \boldsymbol{B}^A$) *then for any efficient strategy*, $\sigma = \langle \alpha, \boldsymbol{B}_{r+1}, \delta \rangle$, *we have*

$$\Omega(\sigma) = \left[\sum_{B \in \boldsymbol{B}_{r+1}} \max\{ \psi(B') \mid B' \in \boldsymbol{B}^A(B)\} \right] \bar{\omega}.$$

Proof. We have from Proposition 1 that

$$\Omega(\sigma) = \left[\sum_{B \in \boldsymbol{B}_{r+1}} \psi(B) \right] \bar{\omega}. \tag{11}$$

However, let $B \in \boldsymbol{B}_{r+1}$ be arbitrary, and let $d \in D$. Then we have, since $\boldsymbol{B}^A(B)$ is a partition of B,

$$\pi(B \cap X_d) = \sum_{B' \in \boldsymbol{B}^A(B)} \pi(B' \cap X_d). \tag{12}$$

However, since $\boldsymbol{X} \geq \boldsymbol{B}^A$, it follows that for all but at most one $B' \in \boldsymbol{B}^A(B)$, we have

$$\pi(B' \cap X_d) = 0,$$

and if $B' \in \boldsymbol{B}^A(B)$ is such that $\pi(B' \cap X_d) > 0$, we have

$$\pi(B' \cap X_d) \leq \psi(B').$$

It follows, therefore, that

$$\psi(B) = \max\{ \pi(B \cap X_d) \mid d \in D\}$$
$$= \max\{ \psi(B') \mid B' \in \boldsymbol{B}^A(B)\}$$

and our result follows. Q.E.D.

The following is an immediate consequence of Proposition 3.

4.1.4. Corollary. If $\boldsymbol{D}$ is of the only correct guesses count form and if $\sigma = \langle \alpha, \boldsymbol{B}_{r+1}, \delta \rangle$ is an efficient strategy for $\boldsymbol{D}$, then

$$\Omega(\sigma) = (\#\boldsymbol{B}_{r+1})\bar{\omega}/m.$$

Proposition 3 shows one reason why the case where condition (3) holds is of particular interest. In section 5.2 we shall be able to develop other results which exploit condition (3) to obtain a sharper characterization of the solution to this special case of the categorization problem; and in section 6.5 we shall also be able to obtain similar (though somewhat weaker) results for the case where condition (4) holds.

4.2. The Computer File Search Problem

The computer file search problem, as we shall develop it here, is also a special case of our categorization problem. We develop it as follows.

We suppose that there is some universal set, U, which is finite and linearly ordered, and that we are dealing with a non-empty subset,

$$S = \{b_1, b_2, \ldots, b_n\} \subseteq U, \tag{1}$$

with

$$b_i < b_{i+1} \quad \text{for} \quad i = 1, \ldots, n-1. \tag{2}$$

We suppose that there is a probability measure on U, so that the probabilities $\Pr(b < b_1)$, $\Pr(b > b_n)$, $\Pr(b_i < b < b_{i+1})$ for $i = 1, \ldots, n-1$, and $\Pr(b = b_i)$ for $i = 1, \ldots, n$, are well-defined, for b a random element of U.

The basic idea is that the elements of S correspond to a stored data set drawn from U. We consider the problem of searching the set in order to determine whether a randomly drawn element from U, b is in the set S or not; and if it is in S, to determine its location (i.e., for which i we have $b = b_i$). The available experiments can be denoted by

$$A = \{0, 1, \ldots, n\},$$

[18] It is in the case where (10) holds that the structure of this problem most closely resembles the game of twenty questions.

where for $a = 1, \dots, n$, the experiment a is interpreted as

'compare b with b_a'

(and $a = 0$ represents the null information experiment). Thus, for $a \in A_1$, the possible outcomes of the experiment are

$$b < b_a, \quad b = b_a \quad \text{or} \quad b > b_a. \tag{3}$$

For notational convenience, we shall represent the state space as

$X = Y \cup Z$,

where $Y = \{y_1, \dots, y_n\}$, with the interpretation

$x = y_i \Leftrightarrow b = b_i \quad \text{for} \quad i = 1, \dots, n,$

and $Z = \{z_0, z_1, \dots, z_n\}$, with the interpretation

$$x = z_j \Leftrightarrow \begin{cases} b < b_1 & \text{if} \quad j = 0 \\ b_j < b < b_{j+1} & \text{for} \quad j = 1, \dots, n-1 \\ b > b_n & \text{for} \quad j = n. \end{cases}$$

We then define

$p_i = Pr(x = y_i) = Pr(b = b_i) \quad \text{for} \quad i = 1, \dots, n,$

$q_0 = Pr(x = z_0) = Pr(b < b_1),$

$q_j = Pr(x = z_j) = Pr(b_j < b < b_{j+1})$

for $j = 1, \dots, n-1$, and

$q_n = Pr(x = z_n) = Pr(b > b_n).$

From (3) and our specification of X, we see that for each $a \in A_1$, the information structure for a, $\boldsymbol{M}_a$, can be written as

$\boldsymbol{M}_a = \{M_{a1}, M_{a2}, M_{a3}\}$, where

$M_{a1} = \{y_1, \dots, y_{a-1}\} \cup \{z_0, \dots, z_{a-1}\},$

$M_{a2} = \{y_a\}$, and

$M_{a3} = \{y_{a+1}, \dots, y_n\} \cup \{z_a, \dots, z_n\}.$

To complete our specification of the problem, we note that we can specify $\boldsymbol{D}$ as

$\boldsymbol{D} = \{0, 1, \dots, n\},$

with the following interpretation: $d = 0$ corresponds to the decision $b \notin S$ (i.e., $x \in Z$), $d = j$ corresponds to the decision $b = b_j$ (i.e., $x = y$) for $j = 1, \dots, n$. It also seems appropriate here to specify our gross payoff function ω: $X \times D \to \mathbb{R}$ as

$$\omega(x, d) = \begin{cases} \bar{\omega} > 0 & \text{if } d = 0 \text{ and } x \in Z, \\ & \quad \text{or if } d \in \{1, \dots, n\} \\ & \quad \text{and } x = y_d, \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

We also suppose that there exists some constant $c > 0$ such that

$c(a) = c \quad \text{for} \quad a = 1, \dots, n.$

Thus we see that, in the formulation developed here, the problem of finding an optimal for the solution of the computer file search problem is a special case of our categorization problem. Notice also that the file search problem satisfies condition (4) of the previous subsection, i.e., we have

$\boldsymbol{B}^A \geq \boldsymbol{X}.$

We shall examine the solution of the computer file search problem in some detail in section 6.5.

5. Expected Costs and the Number of Sets in the Final Information Structure

5.1. Preliminary Results

In this section we shall often be dealing with situations in which $\boldsymbol{D}$ satisfies one or both of the following conditions.

5.1.1. Definition. We shal say that $\boldsymbol{D}$ has *constant information cost c*, if $c \in \mathbb{R}_+$ is such that

$(\forall a \in A_1)$: $c(a) = c.$

5.1.2. Definition. We shall say that $\boldsymbol{D}$ is a *k-element information structure problem*, where $k \in \{2, 3, \dots\}$, iff

$(\forall a \in A_1)$: $n_a = \#\boldsymbol{M}_a = k,$

i.e., for each $a \in A_1$, $\boldsymbol{M}_a$ is of the form

$\boldsymbol{M}_a = \{M_{a1}, \dots, M_{ak}\}$

(*note*: If $k = 2$, we shall use the term '*binary*' in place of 'two-element'; and if $k = 3$, we shall similarly use the term 'trinary').

In this section we shall assume that $\boldsymbol{D}$ is a k-element information structure problem only where explicitly stated. However, *throughout this section we shall use k to denote the value of*

$$k = \max_{a \in A} \#\boldsymbol{M}_a = \max_{a \in A_1} n_a. \tag{1}$$

We shall also find the following to be convenient in characterizing the value of expected cost.

5.1.3. Definition. Given an efficient information-gathering strategy for $\boldsymbol{D}$, $\boldsymbol{\alpha} = \langle(\boldsymbol{B}_1, \alpha_1), \ldots, (\boldsymbol{B}_r, \alpha_r)\rangle$, and letting $\boldsymbol{B}_{r+1} = R(\boldsymbol{B}_r, \alpha_r)$, we define $\tau_\alpha: \boldsymbol{B}_{r+1} \to \{1,\ldots,r\}$ by

$$\tau_\alpha(B) = \max\{t \in \{1,\ldots,r\} \mid a(t, B) \neq 0\}.$$

5.1.4. Proposition. If the decision problem $\boldsymbol{D}$ has constant information cost, c, and $\boldsymbol{\alpha}$ is an efficient information-gathering for $\boldsymbol{D}$, then

$$\gamma(\boldsymbol{\alpha}) = \left[\sum_{B\in\boldsymbol{B}_{r+1}} \pi(B)\tau_\alpha(B)\right]c$$

$$= rc - \sum_{B\in\boldsymbol{B}_{r+1}} \pi(B)[r - \tau_\alpha(B)]c. \qquad (1)$$

Proof. We have

$$\gamma(\boldsymbol{\alpha}) = \sum_{B\in\boldsymbol{B}_{r+1}} \pi(B)C(B), \quad \text{where} \qquad (2)$$

$$C(B) = \sum_{t=1}^{r} c[a(t, B)]. \qquad (3)$$

However, since

$$c[a(t, B)] = \begin{cases} 0 & \text{if} \quad a(t, B) = 0, \\ c & \text{if} \quad a(t, B) \neq 0, \end{cases}$$

it follows (since $\boldsymbol{\alpha}$ is efficient) that

$$\sum_{t=1}^{r} c[a(t, B)] = \tau_\alpha(B)c. \qquad (4)$$

Therefore, by (2)–(4),

$$\gamma(\boldsymbol{\alpha}) = \sum_{B\in\boldsymbol{B}_{r+1}} \pi(B)\tau_\alpha(B)c$$

$$= \left[\sum_{B\in\boldsymbol{B}_{r+1}} \pi(B)\tau_\alpha(B)\right]c,$$

which establishes the first equality in (1). The second equality follows readily, using the fact that, since $\boldsymbol{B}_{r+1}$ is a partition of X,

$$\sum_{B\in\boldsymbol{B}_{r+1}} \pi(B) = 1.$$

Q.E.D.

Turning now to the question of characterizing the number of elements in $\boldsymbol{B}_{r+1}$, we note first that it is fairly obvious that if $\boldsymbol{\alpha}$ is a feasible information-gathering strategy for $\boldsymbol{D}$, and $\boldsymbol{B}_{r+1} = \boldsymbol{R}(\boldsymbol{B}_r, \alpha_r)$, then

$$\#\boldsymbol{B}_{r+1} \leq k^r.$$

However, with the use of the $\tau_\alpha(\cdot)$ function, we will be able to prove a somewhat sharper result; although we must first develop some supporting material, as follows.

Let $\boldsymbol{\alpha} = \langle(\boldsymbol{B}_1,\alpha_1),\ldots,(\boldsymbol{B}_r,\alpha_r)\rangle$ be an efficient information-gathering strategy for $\boldsymbol{D}$ such that $\alpha_1(X) \neq 0$, [19] and let

$$\boldsymbol{B}_{r+1} = R(\boldsymbol{B}_r, \alpha_r).$$

For $t = 1,\ldots, r$, we partition $\boldsymbol{B}_t$ into three subsets, as follows:

$$\boldsymbol{B}_1^1 = \boldsymbol{B}_1^0 = \boldsymbol{B}_2^0 = \phi,\ \boldsymbol{B}_1^2 = \boldsymbol{B}_1 = \{X\},$$

$$\boldsymbol{B}_2^1 = \{B \in \boldsymbol{B}_2 \mid \alpha_2(B) = 0\},$$

$$\boldsymbol{B}_2^2 = \{B \in \boldsymbol{B}_2 \mid \alpha_2(B) \neq 0\},$$

and, in general

$$\boldsymbol{B}_t^0 = \bigcup_{s=1}^{t-1} \boldsymbol{B}_s^1,$$

$$\boldsymbol{B}_t^2 = \{B \in \boldsymbol{B}_t \mid \alpha_t(B) \neq 0\}, \quad \text{and}$$

$$\boldsymbol{B}_t^1 = \boldsymbol{B}_t \setminus [\boldsymbol{B}_t^0 \cup \boldsymbol{B}_t^2] \quad \text{for} \quad t = 1,\ldots, r.$$

Note that for $t \in \{1,\ldots, r\}$, $\boldsymbol{B}_t^2$ might be called the *action set*; it is only if the previous actions (experiments) have shown that $\hat{x}$, the true state, is an element of some $B \in \boldsymbol{B}_t^2$ that an information-gathering activity is conducted at the tth step. The information sets in $\boldsymbol{B}_t \setminus \boldsymbol{B}_t^2$ all have the property that no new experimentation is to be performed at the tth step. However, the sets in $\boldsymbol{B}_t \setminus \boldsymbol{B}_t^2$ are of two types: (a) those on which no new experimentation was to have been performed at the $(t-1)$st step, and which were, therefore, also elements of $\boldsymbol{B}_{t-1}$, and (b) those sets $B \in \boldsymbol{B}_t$ for which $\alpha_t(B) = 0$, but which were not elements of $\boldsymbol{B}_{t-1}$ [thus for $B \in \boldsymbol{B}_t^1$, $\alpha_{t-1}[\beta_{t-1}(B)] \neq 0$].

Thus we have, for each $t \in \{1,\ldots, r\}$,

$$\boldsymbol{B}_t^0 \cup \boldsymbol{B}_t^1 = \{B \in \boldsymbol{B}_t \mid \alpha_t(B) = 0\},$$

and, since $\boldsymbol{\alpha}$ is efficient, it follows that

$$\boldsymbol{B}_t^0 \cup \boldsymbol{B}_t^1 \subseteq \boldsymbol{B}_{r+1} \quad \text{for} \quad t = 1,\ldots, r. \qquad (5)$$

In fact, for each $t \in \{1,\ldots, r+1\}$,

$$\boldsymbol{B}_t^1 = \{B \in \boldsymbol{B}_{r+1} \mid \tau_\alpha(B) = t-1\}, \quad \text{and} \qquad (6)$$

[19] In the following we shall speak of an efficient information-gathering strategy as being *non-trivial* if $\alpha_1(X) \neq 0$. Note that if $\boldsymbol{\alpha}$ is efficient, and $\alpha_1(X) = 0$, then $\boldsymbol{B}_{r+1} = \{X\}$.

$$\boldsymbol{B}_t = \boldsymbol{B}_t^2 \cup \left[\bigcup_{s=1}^{t} \boldsymbol{B}_s^1 \right] \quad \text{for} \quad t = 1, \ldots, r+1. \tag{7}$$

5.1.5. Lemma. If $\boldsymbol{\alpha} = \langle (\boldsymbol{B}_1, \alpha_1), \ldots, (\boldsymbol{B}_r, \alpha_r) \rangle$ is an efficient and non-trivial information-gathering strategy, and we define

$$\boldsymbol{B}_{r+1} = R(\boldsymbol{B}_r, \alpha_r),$$

then we have

$$\#\boldsymbol{B}_t \leq k^{t-1} - \sum_{s=1}^{t-1} (\#\boldsymbol{B}_s^1)(k^{t-s} - 1)$$

$$\text{for} \quad t = 2, \ldots, r+1. \tag{8}$$

Proof. For $t = 2$, we have $\boldsymbol{B}_2 = \boldsymbol{M}_{\alpha_1(X)}$, and obviously, $\#\boldsymbol{B}_2 = \#\boldsymbol{M}_{\alpha_1(X)} \leq k$, while $\sum_{s=1}^{1}(\#\boldsymbol{B}_s^1)(k^{2-s} - 1) = 0 \cdot (k-1) = 0$.

Suppose now that our formula (8) holds for $t = q (q \geq 2)$. then for $t = q+1$, we have

$$\begin{aligned}
\#\mathbf{B}_t = \#\boldsymbol{B}_{q+1} &= \#\left(\bigcup_{B \in \boldsymbol{B}_q} \iota\left[B, \alpha_q(B)\right] \right) \\
&= \sum_{B \in \boldsymbol{B}_q} \#\left(\iota\left[B, \alpha_q(B)\right]\right) \\
&= \sum_{B \in \boldsymbol{B}_q^0} \#\left(\iota\left[B, \alpha_q(B)\right]\right) \\
&\quad + \sum_{B \in \boldsymbol{B}_q^1} \#\left(\iota\left[B, \alpha_q(B)\right]\right) \\
&\quad + \sum_{B \in \boldsymbol{B}_q^2} \#\left(\iota\left[B, \alpha_q(B)\right]\right) \\
&\leq \sum_{s=1}^{q} \#\boldsymbol{B}_s^1 + \left(\#\boldsymbol{B}_q^2\right)k \\
&= \sum_{s=1}^{q} \#\boldsymbol{B}_s^1 + \left(\#\boldsymbol{B}_q - \sum_{s=1}^{q} \#\boldsymbol{B}_s^1 \right)k \\
&\leq \sum_{s=1}^{q} \#\boldsymbol{B}_s^1 + k\left[k^{q-1} - \sum_{s=1}^{q-1} (\#\boldsymbol{B}_s^1)(k^{q-s} - 1) \right. \\
&\qquad \left. - \sum_{s=1}^{q} \#\boldsymbol{B}_s^1 \right] \\
&= \sum_{s=1}^{q} \#\boldsymbol{B}_s^1 + k^q - \sum_{s=1}^{q} (\#\boldsymbol{B}_s^1)k^{q+1-s} \\
&= k^q - \sum_{s=1}^{q} (\#\boldsymbol{B}_s^1)(k^{q+1-s} - 1) \\
&= k^{t-1} - \sum_{s=1}^{t-1} (\#\boldsymbol{B}_s^1)(k^{t-s} - 1).
\end{aligned}$$

Q.E.D.

We can now prove the following.

5.1.6. Proposition. If $\boldsymbol{\alpha} = \langle (\boldsymbol{B}_1, \alpha_1), \ldots, (\boldsymbol{B}_r, \alpha_r) \rangle$ is an efficient information-gathering strategy for $\boldsymbol{D}$, and $\boldsymbol{B}_{r+1} = R(\boldsymbol{B}_r, \alpha_r)$, we have

$$\begin{aligned}
\#\boldsymbol{B}_{r+1} &\leq k^r - \sum_{t=1}^{r+1} (\#\boldsymbol{B}_t^1)(k^{r+1-t} - 1) \\
&= k^r - \sum_{B \in \boldsymbol{B}_{r+1}} [k^{r - \tau_\alpha(B)} - 1],
\end{aligned} \tag{9}$$

and thus

$$\sum_{t=1}^{r+1} (\#\boldsymbol{B}_t^1)k^{r+1-t} \leq k^r. \tag{10}$$

Proof. (*i*) We first prove the inequality in (9), as follows. If $\min\{\tau_\alpha(B) \mid B \in \boldsymbol{B}_{r+1}\} = 0$, then, since $\boldsymbol{\alpha}$ is efficient, $\boldsymbol{B}_{r+1} = \{X\}$, $\#\boldsymbol{B}_{r+1} = 1$, and

$$\sum_{t=1}^{r+1} (\#\boldsymbol{B}_t^1)(k^{r+1-t} - 1) = 1 \cdot (k^r - 1),$$

so that

$$1 = \#\boldsymbol{B}_{r+1} = k^r - \sum_{t=1}^{r+1} (\#\boldsymbol{B}_t^1)(k^{r+1-t} - 1),$$

as required.

Now suppose $\min\{\tau_\alpha(B) \mid B \in \boldsymbol{B}_{r+1}\} \geq 1$. Then $\boldsymbol{\alpha}$ is non-trivial, and by Lemma 5 we have

$$\#\boldsymbol{B}_{r+1} \leq k^r - \sum_{t=1}^{r} (\#\boldsymbol{B}_t^1)(k^{r+1-t} - 1). \tag{11}$$

Since

$$(\#\boldsymbol{B}_{r+1}^1)(k^{r+1-(r+1)} - 1) = 0,$$

we can extend (11) to

$$\#\boldsymbol{B}_{r+1} \leq k^r - \sum_{t=1}^{r+1} (\#\boldsymbol{B}_t^1)(k^{r+1-t} - 1), \tag{12}$$

which establishes the inequality in (9).

Now, we have

$$\boldsymbol{B}_{r+1} = \bigcup_{t=1}^{r+1} \boldsymbol{B}_t^1,$$

and, for all $B \in \boldsymbol{B}_t^1 (t = 1, \ldots, r+1)$, $\tau_\alpha(B) = t - 1$. Thus

$$\begin{aligned}
&k^r - \sum_{t=1}^{r+1} (\#\boldsymbol{B}_t^1)(k^{r+1-t} - 1) \\
&\quad = k^r - \sum_{B \in \boldsymbol{B}_{r+1}} [k^{r - \tau_\alpha(B)} - 1],
\end{aligned}$$

which establishes the equality in (9).

(*ii*) We have

$$\sum_{t=1}^{r+1} (\#\boldsymbol{B}_t^1)(k^{r+1-t}-1)$$
$$= \sum_{t=1}^{r+1} (\#\boldsymbol{B}_t^1)k^{r+1-t} - \sum_{t=1}^{r+1} \#\boldsymbol{B}_t^1$$
$$= \sum_{t=1}^{r+1} (\#\boldsymbol{B}_t^1)k^{r+1-1} - \#\boldsymbol{B}_{r+1}, \tag{13}$$

where the last equality follows from the fact $\{\boldsymbol{B}_1^1, \boldsymbol{B}_2^1, \ldots, \boldsymbol{B}_{r+1}^1\}$ is a partition of $\boldsymbol{B}_{r+1}$. Substitution of (13) into (12) yields inequality (10). Q.E.D.

5.1.7. Corollary. If, under the hypotheses of Proposition 6, *we have*

$$\#\boldsymbol{B}_{r+1} > k(k^{r-1}-1)+1, \qquad \text{then} \tag{14}$$

$$(\forall B \in \boldsymbol{B}_{r+1}):\ \tau_\alpha(B) = r, \tag{15}$$

or, equivalently, for each $t \in \{1, \ldots, r\}$, *we have*

$$(\forall B \in \boldsymbol{B}_t):\ \alpha_t(B) \neq 0, \qquad \text{or} \tag{16}$$

$$\#\boldsymbol{B}_t^1 = 0. \tag{17}$$

Proof. Since it is obvious that (15)–(17) are all equivalent, we shall prove that (14) implies (17). Accordingly, from (14) and (9) of Proposition 6, we have

$$k(k^{r-1}-1)+1 < k^r - \sum_{t=1}^{r+1} (\#\boldsymbol{B}_t^1)(k^{r+1-t}-1),$$

from which we obtain

$$\sum_{t=1}^{r+1} (\#\boldsymbol{B}_t^1)(k^{r+1-t}-1) < k-1. \tag{18}$$

If it were the case that for some $q \in \{1, \ldots, r\}$, we had $\#\boldsymbol{B}_q^1 \geq 1$, then we would obtain from (18) that

$$k-1 > (\#\boldsymbol{B}_q^1)(k^{r+1-q}-1) \geq k-1,$$

yielding a contradiction. Thus we conclude that

$$\#\boldsymbol{B}_t^1 = 0 \quad \text{for} \quad t = 1, \ldots, r.$$

Q.E.D.

5.2. Application to the Categorization Problem

Throughout this section, we shall suppose that **D** *is a categorization problem and that* **D** *has constant information cost*, c, i.e.,

$$(\forall a \in A_1):\ c(a) = c \geq 0. \tag{1}$$

We write

$$\boldsymbol{B}^A = \{B_1, \ldots, B_q\}, \tag{2}$$

and suppose, without loss of generality, that, defining

$$\theta_i = \psi(B_i) \quad \text{for} \quad i = 1, \ldots, q, \tag{3}$$

that we have

$$\theta_1 \geq \theta_2 \geq \cdots \geq \theta_q. \tag{4}$$

Finally, we shall once again let

$$k \equiv \max_{a \in A} \#\boldsymbol{M}_a. \tag{5}$$

Before turning to our first result of this section, it may be worthwhile to set the stage by considering some aspects of the principal new hypothesis we shall use in said result. In Theorem 1, below, we assume that there exists an efficient strategy, $\sigma^* = \langle \alpha^*, \boldsymbol{B}_{r+1}^*, \delta^* \rangle$ such that $\#\boldsymbol{B}_{r+1}^* = k^r$. If this is the case, then $q \geq k^r$ [where q is from (2), above]; and, since $\boldsymbol{B}^A \geq \boldsymbol{B}_{r+1}^*$, there is no loss in generality in supposing that we can label the sets in $\boldsymbol{B}^A$ in such a way that there exists a one-to-one and onto mapping,

$$\zeta\colon \{1, \ldots, k^r\} \to \boldsymbol{B}_{r+1}^*,$$

satisfying

$$B_i \subseteq \zeta(i) \quad \text{for} \quad i = 1, \ldots, k^r, \tag{6}$$

[each $B \in \boldsymbol{B}_{r+1}^*$ contains at least one $B' \in \boldsymbol{B}^A$; and, for each B, $B' \in \boldsymbol{B}_{r+1}^*$, $\boldsymbol{B}^A(B) \cap \boldsymbol{B}^A(B') = \emptyset$]. We shall suppose, however, that the labeling and the mapping ζ can be constructed in such a way as to satisfy (4) as well. This combination of conditions is, therefore, more restrictive [20] than the mere assumption that $\#\boldsymbol{B}_{r+1}^* = k^r$.

5.2.1. Theorem. If **D** *satisfies*

$$\boldsymbol{X} \geq \boldsymbol{B}^A, \tag{7}$$

$$(k-1)\bar{\omega}\theta \geq c, \tag{8}$$

where $\theta = \theta_{k^r}$, *and if there exists an efficient strategy for* **D**, $\sigma^* = \langle \alpha^*, \boldsymbol{B}_{r+1}^*, \delta^* \rangle$, *satisfying* $\#\boldsymbol{B}_{r+1}^* = k^r$, and there exists a one-to-one and onto mapping (9)

[20] Notice, however, that the labeling of $\boldsymbol{B}^A$ which produces (4) will generally not be unique; that is, it is unique if, and only if, all the inequalities in (4) are strict.

$\zeta:\{1,\ldots,k^r\} \to \boldsymbol{B}^*_{r+1}$, satisfying

$B_i \subseteq \zeta(i)$ for $i=1,\ldots,k^r$, then

(i) the expected net return from σ^* *is*

$$\Omega(\sigma^*) - \Gamma(\sigma^*) = \left(\sum_{i=1}^{k^r} \theta_i\right)\bar{\omega} - rc, \quad \text{and}$$

(ii) σ^* *is optimal for* $\boldsymbol{D}$.

Proof. (*i*) By Proposition 4.1.3, we have

$$\Omega(\sigma^*) = \left[\sum_{B\in \boldsymbol{B}^*_{r+1}} \max\{\psi(B') \mid B' \in \boldsymbol{B}^A(B)\}\right]\bar{\omega}. \tag{10}$$

However, if $B \in \boldsymbol{B}^*_{r+1}$ is such that

$\zeta^{-1}(B) = i \in \{1,\ldots,k^r\}$,

we have

$B_i \subseteq B$,

and, for $j<i$, $B \cap B_j = \emptyset$. Therefore, by (4),

$\max\{\psi(B') \mid B' \in \boldsymbol{B}^A(B)\} = \theta_i = \psi(B_i)$,

and, by (10) we then have

$$\Omega(\sigma^*) = \left(\sum_{i=1}^{k^r} \theta_i\right)\bar{\omega}. \tag{11}$$

Now, by Proposition 5.1.4,

$$\Gamma(\sigma^*) = \left[\sum_{B\in \boldsymbol{B}^*_{r+1}} \pi(B)\tau_{\alpha^*}(B)\right]c. \tag{12}$$

However, since $\#\boldsymbol{B}^*_{r+1} = k^r$, it follows from Corollary 5.1.7 that

$(\forall B \in \boldsymbol{B}^*_{r+1})$: $\tau_{\alpha^*}(B) = r$,

and thus we have from (12) that

$$\Gamma(\sigma^*) = \left[\sum_{B\in \boldsymbol{B}^*_{r+1}} \pi(B)\right] rc = rc, \tag{13}$$

where the second equality is by the fact that $\boldsymbol{B}^*_{r+1}$ is a partition of X.

(*ii*) Suppose $\sigma = \langle \alpha, \boldsymbol{B}_{r+1}, \delta\rangle$ is an efficient strategy for $\boldsymbol{D}$. Then we have by Proposition 4.1.3 and 5.1.4 that

$$\begin{aligned}\Omega(\sigma) - \Gamma(\sigma) &= \left[\sum_{B\in \boldsymbol{B}_{r+1}} \max\{\psi(B') \mid B' \in \boldsymbol{B}^A(B)\}\right]\bar{\omega} \\ &\quad - rc + \sum_{B\in \boldsymbol{B}_{r+1}} \pi(B)[r - \tau_\alpha(B)]c \end{aligned} \tag{14}$$

Now, by Proposition 5.1.6, we have

$$\#\boldsymbol{B}_{r+1} \le k^r - \sum_{B\in \boldsymbol{B}_{r+1}} [k^{r-\tau_\alpha(B)} - 1]. \tag{15}$$

Defining

$$\bar{p} = \sum_{B\in \boldsymbol{B}_{r+1}} [k^{r-\tau_\alpha(B)} - 1],$$

we see from (15), (4) and the form of $\Omega(\sigma)$ that we can write

$$\begin{aligned}\Omega(\sigma) &= \left[\sum_{B\in \boldsymbol{B}_{r+1}} \max\{\psi(B') \mid B' \in \boldsymbol{B}^A(B)\}\right]\bar{\omega} \\ &\le \left(\sum_{i=1}^{k^r-\bar{p}} \theta_i\right)\bar{\omega}. \end{aligned} \tag{16}$$

Thus from (11), (13), (14), and (16), we have

$$\begin{aligned}&\Omega(\sigma^*) - \Gamma(\sigma^*) - [\Omega(\sigma) - \Gamma(\sigma)] \\ &\quad \ge \left(\sum_{i=k^r-\bar{p}+1}^{k^r} \theta_i\right)\bar{\omega} - \sum_{B\in \boldsymbol{B}_{r+1}} \pi(B)[r-\tau_\alpha(B)]c \\ &\quad \ge [k^r - (k^r - \bar{p})]\theta\bar{\omega} \\ &\qquad - \sum_{B\in \boldsymbol{B}_{r+1}} \pi(B)[r-\tau_\alpha(B)]c \\ &\quad = \left[\sum_{B\in \boldsymbol{B}_{r+1}} (k^{r-\tau_\alpha(B)} - 1)\right]\theta\bar{\omega} \\ &\qquad - \sum_{B\in \boldsymbol{B}_{r+1}} \pi(B)[r-\tau_\alpha(B)]c, \end{aligned}$$

and thus,

$$\begin{aligned}&\Omega(\sigma^*) - \Gamma(\sigma^*) - [\Omega(\sigma) - \Gamma(\sigma)] \\ &\quad \ge \sum_{B\in \boldsymbol{B}_{r+1}} \{k^{r-\tau_\alpha(B)} - 1)\theta\bar{\omega} \\ &\qquad\qquad - \pi(B)[r-\tau_\alpha(B)]c\}. \end{aligned} \tag{17}$$

Now, let $B \in \boldsymbol{B}_{r+1}$ be arbitrary, and define

$$\chi(B) = (k^{r-\tau_\alpha(B)} - 1)\theta\bar{\omega} - \pi(B)[r-\tau_\alpha(B)]c. \tag{18}$$

We consider two cases.

(a) $\tau_\alpha(B) = r$. In this case, (18) becomes

$$\chi(B) = (k^0 - 1)\theta\bar{\omega} - \pi(B)\cdot 0\cdot c = 0. \tag{19}$$

(b) $\tau_\alpha(B) \in \{1,\ldots,r-1\}$. Since $r - \tau_\alpha(B) > 0$ in this case, we see that

$$\begin{aligned}\chi(B) \ge 0 \Leftrightarrow \theta\bar{\omega}[k^{r-\tau_\alpha(B)} - 1]/[r-\tau_\alpha(B)] \\ \ge \pi(B)c. \end{aligned} \tag{20}$$

However, it is easy to prove that, since $k \geq 2$,

$$(k^q - 1)/q \geq k - 1 \quad \text{for} \quad q = 1, 2, \ldots,$$

and thus

$$\theta\bar{\omega}[k^{r-\tau_\alpha(B)} - 1]/[r - \tau_\alpha(B)] \geq (k-1)\theta\bar{\omega}. \quad (21)$$

However, by assumption (8) and the fact that $B \subseteq X$, we have

$$(k-1)\theta\bar{\omega} \geq c \geq \pi(B)c, \quad (22)$$

and thus it follows from (20)–(22) that $\chi(B) \geq 0$ in this case as well.

From our analysis of the two possible cases, we see that the right-hand-side of (17) is a sum of non-negative terms, and it follows that

$$\Omega(\sigma^*) - \Gamma(\sigma^*) \geq \Omega(\sigma) - \Gamma(\sigma).$$

Q.E.D.

5.2.2. Corollary. If $\boldsymbol{D}$ *is of the only correct guesses count form, and satisfies*

$$(k-1)\bar{\omega}/m \geq c, \quad (23)$$

and $\sigma^* = \langle \alpha^*, \mathbf{B}^*_{r+1}, \delta^* \rangle$ *is an efficient strategy for* $\boldsymbol{D}$ *satisfying*

$\#\boldsymbol{B}^*_{r+1} = k^r$, then

(i) the expected net return from σ^* *is given by*

$\Omega(\sigma^*) - \Gamma(\sigma^*) = k^r\bar{\omega}/m - rc$, and

(ii) σ^* *is optimal for* $\boldsymbol{D}$.

Proof. We first note that, in the notation of Theorem 1 we have

$$\theta_i = 1/m \quad \text{for} \quad i = 1, \ldots, q.$$

Thus, since $\#\boldsymbol{B}^*_{r+1} = k^r$, it is clear that hypotheses (8) and (9) of Theorem 1 are satisfied. Since $\boldsymbol{D}$ is of the only correct guesses count form [and thus satisfies (7) as well], it then follows from Theorem 1 that

$$\Omega(\sigma^*) - \Gamma(\sigma^*) = \left(\sum_{j=1}^{k^r} \theta_i\right)\bar{\omega} - rc = k^r\bar{\omega}/m - rc,$$

and σ^* is optimal for $\boldsymbol{D}$. Q.E.D.

In the case where $\boldsymbol{D}$ has the structure assumed in Theorem 1, the following provides a simple condition for constructing admissible strategies for $\boldsymbol{D}$.

5.2.3. Proposition. Suppose $\boldsymbol{D}$ *satisfies*

$$X \geq \boldsymbol{B}^A, \quad \text{and} \quad (24)$$

$$\bar{\omega}\theta_q > c, \quad (25)$$

and that $\sigma = \langle(\boldsymbol{B}_1, \alpha_1), \ldots, (\boldsymbol{B}_r, \alpha_r), (\boldsymbol{B}_{r+1}, \delta)\rangle$ *is an efficient strategy for* $\boldsymbol{D}$ *such that for some* $t \in \{1, \ldots, r\}$ *and some* $B \in \boldsymbol{B}_t$ *we have*

$$B \notin \boldsymbol{B}^A \quad \text{and} \quad \alpha_t(B) = 0. \quad (26)$$

Then σ *is strictly dominated.*

Proof. If $B \in \boldsymbol{B}_t$ satisfies (26), it follows from Proposition 2.4.7 that there exists $\hat{a} \in A$ such that

$$\#\iota(B, \hat{a}) \geq 2.$$

Thus, using (the proof of) Proposition 4.1.3, we see that

$$\begin{aligned} &[1/\pi(B)]\left[\sum_{B' \in \iota(B, \hat{a})} \pi(B')\nu(B') \right. \\ &\qquad \left. - \sum_{x \in B} \phi(x)\omega[x, \delta(B)]\right] \\ &= [\bar{\omega}/\pi(B)] \\ &\quad \times \left[\sum_{B' \in \iota(B, \hat{a})} \max\{\psi(B'') \mid B'' \in \boldsymbol{B}^A(B')\}\right. \\ &\qquad \left. - \max\{\psi(B'') \mid B'' \in \boldsymbol{B}^A(B)\}\right]. \end{aligned} \quad (27)$$

However, since $\iota(B, \hat{a})$ is a partition of B, and each $B' \in \iota(B, \hat{a})$ is n-feasible,

$$\bigcup_{B' \in \iota(B, \hat{a})} \boldsymbol{B}^A(B') = \boldsymbol{B}^A(B).$$

Therefore, if

$$\max\{\psi(B'') \mid B'' \in \boldsymbol{B}^A(B)\} = \theta_h,$$

it follows from the fact that $\#\iota(B, \hat{a}) \geq 2$ that there exists $i \in \{1, \ldots, q\}$ such that

$$\begin{aligned} &[\bar{\omega}/\pi(B)] \\ &\quad \times \left[\sum_{B' \in \iota(B, \hat{a})} \max\{\psi(B'') \mid B'' \in \boldsymbol{B}^A(B')\}\right. \\ &\qquad \left. - \max\{\psi(B'') \mid B'' \in \boldsymbol{B}^A(B)\}\right] \\ &\geq [\bar{\omega}/\pi(B)][\theta_h + \theta_i - \theta_h] \\ &\geq \bar{\omega}\theta_q/\pi(B). \end{aligned} \quad (28)$$

However, it follows from (25), (27), (28), and the

fact that $0 < \pi(B) \le 1$, that

$$[1/\pi(B)]\left[\sum_{B'\in\iota(B,\hat{a})}\pi(B')\nu(B') - \sum_{x\in B}\phi(x)\omega[x,\delta(B)]\right] > c,$$

and it then follows from Proposition 3.2.3 that σ is strictly dominated. Q.E.D.

5.3. Balanced Strategies

In section 5.1 we presented a result (Proposition 5.1.6) which gave an upper bound on the number of sets in the final information structure. In some situations, however, we can provide a more exact characterization of the number of sets in the final information structure. One of these situations is defined by the following (note: in the following we shall use the $\boldsymbol{B}_t^0$, $\boldsymbol{B}_t^1$, $\boldsymbol{B}_t^2$ partition of $\boldsymbol{B}_t$ that was introduced in section 5.1.

5.3.1. Definition. Let $\boldsymbol{D}$ be a k-element information structure problem, and let $\boldsymbol{I} \in \{0, 1,\ldots, k-1\}$. We shall say that a feasible information-gathering strategy for $\boldsymbol{D}$,

$\alpha = \langle(\boldsymbol{B}_1, \alpha_1),\ldots,(\boldsymbol{B}_r, \alpha_r)\rangle$

is a *(k, I)-balanced (informational) strategy for* $\boldsymbol{D}$ iff α satisfies

(1) $\#\boldsymbol{B}_2 = k$ (and thus $\#\boldsymbol{B}_1^2 = 1$ and $\#\boldsymbol{B}_1^1 = 0$).
(2) $\#\boldsymbol{B}_{t+1} = k(\#\boldsymbol{B}_t^2) + \Sigma_{s=1}^t \#\boldsymbol{B}_s^1$ for $t = 1,\ldots, r$.
(3) $\#\boldsymbol{B}_t^1 = \boldsymbol{I}(\#\boldsymbol{B}_{t-1}^2)$ for $t = 2,\ldots, r$.

(We shall say that a feasible strategy, $\sigma = \langle\alpha, \boldsymbol{B}_{r+1}, \delta\rangle$ is a $(k, \boldsymbol{I})$-balanced strategy if α is a $(k, \boldsymbol{I})$-balanced strategy.)

Thus with a $(k, \boldsymbol{I})$-balanced strategy, $\alpha_t(B) \neq 0$ implies

$\#\iota[B, \alpha_t(B)] = k,$

for $B \in \boldsymbol{B}_t$, $t = 1,\ldots, r$ (experimentation is performed only when there are k-possible outcomes, given the information available beforehand). Furthermore, (on average) each time a non-null experiment is undertaken at $t \in \{1,\ldots, r-1\}$ on $B \in \boldsymbol{B}_t$, no new experimentation is undertaken on $\boldsymbol{I}$ of the sets in $\iota[B, \alpha_t(B)]$ at $t+1$. The most likely situation in which this latter condition will happen is when it is true that for each $B \in \boldsymbol{B}_t$ and each $a \in A$, $\iota(B, a)$ contains $\boldsymbol{I}$ sets from $\boldsymbol{B}^A$. We will find this sort of situation to obtain, for example, in the computer file search problem.

5.3.2. Proposition. If $\boldsymbol{D}$ is a k-element information structure problem, $\boldsymbol{I} \in \{0, 1,\ldots, k-1\}$, and

$\alpha = \langle(\boldsymbol{B}_1, \alpha_1),\ldots,(\boldsymbol{B}_r, \alpha_r)\rangle$

is a $(k, \boldsymbol{I})$-balanced strategy for $\boldsymbol{D}$, then

(i) $\#\boldsymbol{B}_t^2 = (k-\boldsymbol{I})^{t-1}$ *for* $t = 1,\ldots, r$,

(ii) $\#\boldsymbol{B}_t^1 = \boldsymbol{I}(k-\boldsymbol{I})^{t-2}$ *for* $t = 2,\ldots, r$,

(iii) $\sum_{s=1}^{t}(\#\boldsymbol{B}_s^1)$

$$= \begin{Bmatrix} \boldsymbol{I}\left[\dfrac{(k-\boldsymbol{I})^{t-1}-1}{k-\boldsymbol{I}-1}\right] \\ \quad\text{for}\quad \boldsymbol{I}\in\{0,\ldots,k-2\} \\ \boldsymbol{I}(t-1) \\ \quad\text{for}\quad \boldsymbol{I}=k-1 \end{Bmatrix}$$

for $t = 1,\ldots, r$,

(iv) $\#\boldsymbol{B}_t$

$$= \begin{Bmatrix} \dfrac{(k-1)(k-\boldsymbol{I})^{t-1}-\boldsymbol{I}}{k-\boldsymbol{I}-1} \\ \quad\text{for}\quad \boldsymbol{I}\in\{0,\ldots,k-2\} \\ \boldsymbol{I}(t-1)+1 \\ \quad\text{for}\quad \boldsymbol{I}=k-1 \end{Bmatrix}$$

for $t = 1,\ldots, r+1$,

where $\boldsymbol{B}_{r+1} = R(\boldsymbol{B}_r, \alpha_r)$.

Proof. (i) We have [see (7) of section 5.1],

$$\#\boldsymbol{B}_{t+1} = \#\boldsymbol{B}_{t+1}^2 + \sum_{s=1}^{t+1}\#\boldsymbol{B}_s^1 \quad\text{for}\quad t = 1,\ldots, r-1, \tag{1}$$

while, since α is a $(k, \boldsymbol{I})$-balanced strategy,

$$\#\boldsymbol{B}_{t+1} = k(\#\boldsymbol{B}_t^2) + \sum_{s=1}^{t}\#\boldsymbol{B}_s^1 \quad\text{for}\quad t = 1,\ldots, r, \tag{2}$$

$$\#\boldsymbol{B}_t^1 = \boldsymbol{I}(\#\boldsymbol{B}_{t-1}^2) \quad\text{for}\quad t = 2,\ldots, r, \quad\text{and} \tag{3}$$

$$\#\boldsymbol{B}_2 = k,\ \#\boldsymbol{B}_1^2 = 1,\ \#\boldsymbol{B}_1^1 = 0. \tag{4}$$

Substituting (3) into (1) and (2), and equating

the result, we obtain

$$\#\boldsymbol{B}_{t+1}^2 + \sum_{s=2}^{t+1} \boldsymbol{I}(\#\boldsymbol{B}_{s-1}^2)$$
$$= k(\#\boldsymbol{B}_t^2) + \sum_{s=2}^{t} \boldsymbol{I}(\#\boldsymbol{B}_{s-1}^2),$$

from which we have

$$\#\boldsymbol{B}_{t+1}^2 = (k-\boldsymbol{I})(\#\boldsymbol{B}_t^2) \quad \text{for} \quad t = 1, 2, \ldots, r-1. \tag{5}$$

From the initial condition (4), eq. (5), and a trivial induction argument, we obtain

$$\#\boldsymbol{B}_t^2 = (k-\boldsymbol{I})^{t-1} \quad \text{for} \quad t = 1, \ldots, r. \tag{6}$$

(ii) From (3) and (6) we obtain

$$\#\boldsymbol{B}_t^1 = \boldsymbol{I}(k-\boldsymbol{I})^{t-2} \quad \text{for} \quad t = 2, \ldots, r. \tag{7}$$

(iii) It follows at once from (7) that

$$\sum_{s=1}^{t} \#\boldsymbol{B}_s^1 = \boldsymbol{I} \sum_{s=2}^{t} (k-\boldsymbol{I})^{s-2} = \boldsymbol{I} \sum_{s=0}^{t-2} (k-\boldsymbol{I})^{s}$$
$$\text{for} \quad t = 2, \ldots, r. \tag{8}$$

From (8) we obtain two cases

$$\sum_{s=1}^{t} \#\boldsymbol{B}_s^1 = \boldsymbol{I}(t-1) \quad \text{for} \quad t = 1, \ldots, r,$$
$$\text{if} \quad \boldsymbol{I} = k-1, \tag{9}$$

and, using the formula for the partial sum of a geometric series

$$\sum_{s=1}^{t} \#\boldsymbol{B}_s^1 = \boldsymbol{I}\left[\frac{(k-\boldsymbol{I})^{t-1} - 1}{k-\boldsymbol{I}-1}\right] \quad \text{for} \quad t = 1, \ldots, r,$$
$$\text{if} \quad \boldsymbol{I} \in \{0, \ldots, k-2\}. \tag{10}$$

(iv) From (2), (6), (9), and (10), we obtain

$$\#\boldsymbol{B}_t = k(k-\boldsymbol{I})^{t-2} + \boldsymbol{I}(t-2) = \boldsymbol{I}(t-1) + 1$$
$$\text{for} \quad t = 2, \ldots, r, \quad \text{if} \quad \boldsymbol{I} = k-1, \quad \text{and}$$

$$\#\boldsymbol{B}_t = k(k-\boldsymbol{I})^{t-2} + \boldsymbol{I}\left[\frac{(k-\boldsymbol{I})^{t-2} - 1}{k-\boldsymbol{I}-1}\right]$$
$$= \left[k(k-\boldsymbol{I})^{t-1} - k(k-\boldsymbol{I})^{t-2} + \boldsymbol{I}(k-\boldsymbol{I})^{t-2} - \boldsymbol{I}\right] \times [k-\boldsymbol{I}-1]^{-1}$$
$$= \frac{(k-1)(k-\boldsymbol{I})^{t-1} - \boldsymbol{I}}{k-\boldsymbol{I}-1}$$
$$\text{for} \quad t = 2, \ldots, r+1, \quad \text{if} \quad \boldsymbol{I} \in \{0, \ldots, k-2\}.$$

Since the formulas

$$f(t) = \boldsymbol{I}(t-1) + 1, \qquad \text{and}$$
$$g(t) = \left[(k-1)(k-\boldsymbol{I})^{t-1} - \boldsymbol{I}\right]/[k-\boldsymbol{I}-1],$$

yield

$$f(1) = g(1) = 1 = \#\boldsymbol{B}_1,$$

our conclusion follows. Q.E.D.

It is entirely possible that the only interesting application of our next result is to the computer file search problem. However, since the hypotheses are satisfied in a much wider class of problems, we present the result here rather than postponing it to section 6.5.

5.3.3. Proposition. Suppose $\boldsymbol{D}$ *is a k-element information structure problem with constant information cost,* $c \geq 0$; *and suppose* $\alpha = \langle(\boldsymbol{B}_1, \alpha_1), \ldots, (\boldsymbol{B}_r, \alpha_r)\rangle$ *is a* $(k, \boldsymbol{I})$*-balanced strategy for* $\boldsymbol{D}$ *such that for each* $t \in \{1, \ldots, r\}$, *and each* $B \in \boldsymbol{B}_t^1$, *we have* $\pi(B) = p > 0$. *Then*

$$\gamma(\boldsymbol{\alpha}) = \begin{cases} (c/2)[2r - p\boldsymbol{I}r(r-1)] \\ \quad \text{if} \quad \boldsymbol{I} = k-1, \\ \{c/(k-\boldsymbol{I}-1)\} \\ \quad \cdot\{r(k-\boldsymbol{I}-1+p\boldsymbol{I}) \\ \qquad - p\boldsymbol{I}\left[(k-\boldsymbol{I})^r - 1\right] \\ \qquad /[k-\boldsymbol{I}-1]\} \\ \quad \text{if} \quad \boldsymbol{I} \in \{0, \ldots, k-2\}. \end{cases}$$

Proof. From Proposition 2.3.2 and the definition of $\boldsymbol{B}_t^2$, we have

$$\gamma(\boldsymbol{\alpha}) = \sum_{t=1}^{r} \sum_{B \in \boldsymbol{B}_t} \pi(B) c[\alpha_t(B)]$$
$$= c\left[\sum_{t=1}^{r} \sum_{B \in \boldsymbol{B}_t^2} \pi(B)\right]$$
$$= c\left(\sum_{t=1}^{r}\left[1 - \sum_{s=1}^{t} \sum_{B \in \boldsymbol{B}_s^1} \pi(B)\right]\right)$$
$$= c\left(\sum_{t=1}^{r}\left[1 - p\sum_{s=1}^{t}(\#\boldsymbol{B}_s^1)\right]\right)$$
$$= c\left(r - p\sum_{t=1}^{r}\sum_{s=1}^{t} \#\boldsymbol{B}_s^1\right). \tag{11}$$

Using Proposition 2, we now distinguish two cases.
(a) $I = k - 1$. Here we have

$$\sum_{s=1}^{t} \#B_s^1 = I(t-1). \tag{12}$$

Substituting (12) into (11), we then have

$$\begin{aligned}\gamma(\alpha) &= c\left[r - pI\sum_{t=1}^{r}(t-1)\right]\\ &= c\left(r - pI\left[\frac{r(r-1)}{2}\right]\right)\\ &= (c/2)[2r - pIr(r-1)].\end{aligned}$$

(b) $I \in \{0,\dots,k-2\}$. Here we have

$$\sum_{s=1}^{t} \#B_s^1 = I\left[(k-I)^{t-1} - 1\right]/(k-I-1). \tag{13}$$

Substituting (13) into (11), we then obtain

$$\begin{aligned}\gamma(\alpha) &= c\Bigg(r - pI/(k-I-1)\\ &\qquad \times \sum_{t=1}^{r}\left[(k-I)^{t-1} - 1\right]\Bigg)\\ &= c\Bigg[r + rpI/(k-I-1) - pI\\ &\qquad /(k-I-1)\sum_{t=1}^{r}(k-I)^{t-1}\Bigg]\\ &= c\left\{r + rpI/(k-I-1) - pI\left[(k-I)^r - 1\right]\right.\\ &\qquad \left./(k-I-1)^2\right\}\\ &= \{c/(k-I-1)\}\\ &\qquad \cdot\left\{r(k-I-1+pI) - pI\left[(k-I)^r - 1\right]\right.\\ &\qquad \left./(k-I-1)\right\}.\end{aligned}$$

Q.E.D.

The special case of the above result which will be pertinent to our analysis of the computer file search problem is the following; the proof of which is immediate.

5.3.4. Corollary. Under the hypotheses of 5.3.3, *and with* $k = 3$ *and* $I = 1$, *we have*

$$\gamma(\alpha) = c[r(1+p) - p(2^r - 1)].$$

While the results of this subsection have direct applications to both the 'only correct guesses count' problem and the computer file search problem, we shall postpone our discussion of these applications until after we have developed the theoretical results of the next section.

6. Binary and Trinary Information Structures with a Linear Ordering on the Set of Experiments

6.1. Introduction

Throughout this section we shall suppose that D is a k-element information structure problem, and that k is either two or three; that is, that information is either binary or trinary. Many of the concepts, and some of the results we shall develop here are valid for $k \geq 4$ as well; but the key results, those which make the material to follow of real use, do not hold (or hold only in trivial special cases) for $k \geq 4$.

By way of introducing the material to follow, suppose $k = 2$, and consider the binary relation $\gtrsim$ defined on A_1 by

$$a \gtrsim a' \Leftrightarrow M_{a1} \supseteq M_{a'1} \quad \text{for} \quad a,\ a' \in A_1. \tag{1}$$

It is easy to show that $\gtrsim$ is a partial order on A_1 (that is, it is reflexive, transitive, and antisymmetric), [21] and has the asymmetric part, '$\succ$ given by [22]

$$a \succ a' \Leftrightarrow M_{a'1} \subset M_{a1} \tag{2}$$

In section 6.4, below, we shall explore some implications of the assumption that $\gtrsim$ is total on A_1. Since $\gtrsim$ is a partial order on A_1 (and thus antisymmetric), it follows, however, that if $\gtrsim$ is total, then '$\succ$ is total [23] as well; and, conversely, if '$\succ$ is total, then $\gtrsim$ is total. Consequently, rather than assuming that $\gtrsim$ is total, we can equivalently (and shall) proceed by assuming that '$\succ$ is total on A_1.

In the trinary case ($k = 3$), we proceed by first

[21] In so far as the proof that $\gtrsim$ is antisymmetric is concerned, note that if $a \gtrsim a'$, then $M_{a1} = M_{a'1}$. But then, since both $\{M_{a1}, M_{a2}\}$ and $\{M_{a'1}, M_{a'2}\}$ are partitions of X, it follows that $M_{a2} = M_{a'2}$ as well. Consequently $M_a = M_{a'}$, and thus $a = a'$.

[22] Where we use the notation '$A \subset B$' to indicate that A is a *proper subset* of B; i.e., $A \subseteq B$ and $A \neq B$.

[23] That is, for all a, $a' \in A_1$, we have $a \succ a'$, $a' \succ a$, or $a = a'$.

defining '> on A_1 as a slight strengthening of (2)

$$a \,'\!> a' \Leftrightarrow M_{a'1} \cup M_{a'2} \subset M_{a1}. \tag{3}$$

It is easily shown that the relation '> defined in (3) is asymmetric and transitive. In section 6.2 we will explore the implications of the assumption that it is total as well.

In the next three subsections we shall treat the trinary and binary case separately. It would be possible to treat the two cases simultaneously [although we would have to introduce a new condition in place of (2) for the binary case], but this would require our introducing some new notation which would serve no purpose other than to allow us to treat the two cases simultaneously.

While we shall assume in both subsections that the pertinent ordering is total on A_1, it should be noted that in each case the results are applicable on any chain in A_1. That is, in section 6.2–6.4 we could replace A_1 with some (possibly proper) subset of A_1, call it A^*, on which '> is total. We shall discuss this extension very briefly in section 6.5.

6.2. The Linear Ordering in the Trinary Case

Throughout this section we shall assume that D is a trinary information structure problem. Moreover, defining '> on A_1 by

$$a \,'\!> a' \Leftrightarrow M_{a'1} \cup M_{a'2} \subset M_{a1}, \tag{1}$$

we assume that '> is total on A_1. Since '> is total on A_1, we can also suppose, without loss of generality that our experiments are numbered in such a way that '> coincides with the usual strict inequality for the real numbers. Thus

$$(\forall i, j \in A_1): i > j \Leftrightarrow M_{j1} \cup M_{j2} \subset M_{i1}. \tag{2}$$

It will be convenient in the following to adjoin to A_1 the two elements 0 and $n+1$, where we define

$$M_0 = \{M_{01}, M_{02}, M_{03}\} \quad \text{with}$$

$$M_{01} = M_{02} = \emptyset,\ M_{03} = X, \quad \text{and} \tag{3}$$

$$M_{n+1} = \{M_{n+1,1}, M_{n+1,2}, M_{n+1,3}\} \quad \text{with}$$

$$M_{n+1,1} = X,\ M_{n+1,2} = M_{n+1,3} = \emptyset. \tag{4}$$

Defining

$$\hat{A} = A_1 \cup \{0, n+1\} = \{0, 1, \ldots, n+1\}, \tag{5}$$

we can then prove the following.

6.2.1. Lemma. *With the definitions* (3)–(5), *above,* '> *is total on* $\hat{A}$ *and coincides with the usual strict inequality for the real numbers on* $\hat{A}$; *that is*

$$(\forall i, j \in \hat{A}): j > i \Leftrightarrow M_{i1} \cup M_{i2} \subset M_{j1}. \tag{6}$$

Moreover, we have

$$(\forall i, j \in \hat{A}):\ j > i \Leftrightarrow M_{j2} \cup M_{j3} \subset M_{i3}. \tag{7}$$

Proof. In order to establish (6) it obviously (given our previous reasoning) suffices to establish (6) for the special case where $\{i, j\} \cap \{0, n+1\} \neq \emptyset$. However, if $i = 0$ and $j \in \{1, \ldots, n+1\}$ then $M_{j1} \neq \emptyset$, while

$$M_{i1} \cup M_{i2} = \emptyset.$$

Thus, for $i = 0$ and $j \in \{1, \ldots, n+1\}$

$$\emptyset = M_{i1} \cup M_{i2} \subset M_{j1}.$$

Similarly, if $j = n+1$ and $i \in \{0, 1, \ldots, n\}$, then

$M_{j1} = X$, while

$M_{i3} \neq \emptyset$, so that

$X \neq X \setminus M_{i3} = M_{i1} \cup M_{i3}$, Therefore

$M_{i1} \cup M_{i2} \subset M_{j1}$

in this case as well.

Now let i, $j \in \hat{A}$ be such that $j > i$. Then by (6) we have

$M_{i1} \cup M_{i2} \subset M_{j1}$, so that

$$M_{j2} \cup M_{j3} = X \setminus M_{j1} \subset X \setminus (M_{i1} \cup M_{i2}) = M_{i3}.$$

Conversely, if i, $j \in \hat{A}$ such that

$M_{j2} \cup M_{j3} \subset M_{i3}$, then

$$M_{i1} \cup M_{i2} = X \setminus M_{i3} \subset X \setminus (M_{j2} \cup M_{j3}) = M_{j1},$$

and thus by (6), $j > i$. Q.E.D.

The following sets forth the basic facts regarding intersections of M_{ay}, $M_{a'y'}$.

6.2.2. Proposition. *For all* i, j, $k \in \hat{A}$, *we have*

(*i*) $M_{j1} \cap M_{i3} \neq \emptyset \Leftrightarrow j > i$,

(*ii*) $k \geq j \to M_{k2} \cap M_{j1} = \emptyset$,

(*iii*) $k \leq i \to M_{k2} \cap M_{i3} = \emptyset$,

(*iv*) $j \neq k \to M_{k2} \cap M_{j2} = \emptyset$,

(*v*) $i < k < j \to M_{j1} \cap M_{k2} \cap M_{i3} = M_{k2}$.

Proof. (i) If $j \leq i$, $M_{j1} \subseteq M_{i1}$. Since $M_{i1} \cap M_{i3} =$

Using (15), define

$$i = \max A_\eta^3 \quad \text{and} \quad j = \min A_\eta^1. \tag{16}$$

It then follows at once from (14) that

$$B \subseteq M_{j1} \cap M_{i3}, \tag{17}$$

and since $B \neq \emptyset$ (by the assumption that B is n-feasible), it follows from (i) of Proposition 2 that

$$j \geq i + 1. \tag{18}$$

Furthermore, by Lemma 1, we also have

$$\bigcap_{a \in A_\eta^1} M_{a1} = M_{j1}, \quad \text{and} \tag{19}$$

$$\bigcap_{a \in A_\eta^3} M_{a3} = M_{i3}. \tag{20}$$

From (14) and (18)–(20), we have, therefore

$j \in \{i+1, \dots, n+1\}$ and

$$B = M_{j1} \cap \left(\bigcap_{a \in A_\eta^2} M_{a2} \right) \cap M_{i3}. \tag{21}$$

Next, we note that, again using the fact that $B \neq \emptyset$, it follows at once from (21) and (iv) of Proposition 2 that there exists $k \in \{1, \dots, n\}$ such that

$$A_\eta^2 \subseteq \{k\}. \tag{22}$$

Thus we can distinguish two cases, which are mutually exclusive and exhaustive.

(a) $A_\eta^2 = \emptyset$. In this case,

$$\bigcap_{a \in A_\eta^2} M_{a2} = X,$$

and it follows from (21) that

$$B = M_{j1} \cap M_{i3} \quad \text{with} \quad i \in \{0, 1, \dots, n\} \quad \text{and}$$

$$j \in \{i+1, \dots, n+1\}. \tag{23}$$

(b) $A_\eta^2 = \{k\}$ for some $k \in \{1, \dots, n\}$. In this case we have from (21) that

$$B = M_{j1} \cap M_{k2} \cap M_{i3}. \tag{24}$$

However, from (ii) and (iii) of Proposition 2 we see that, since $B \neq \emptyset$, we must then have

$$i < k < j,$$

and it then follows from (24) and (v) of Proposition 2 that $B = M_{k2}$. Q.E.D.

We will make heavy use of the following definition in the remainder of this subsection.

6.2.4. Definition. For each $i \in \{0, 1, \dots, n\}$ and each $j \in \{i+1, \dots, n+1\}$, we define $B_{ij} \subseteq X$ by

$$B_{ij} = M_{j1} \cap M_{i3}.$$

Notice that with the use of the above definition, we can give an equivalent statement of Theorem 3 as

6.2.3′. Theorem. If $B \subseteq X$ is n-feasible, then exactly one of the following holds:
(a) $B = B_{ij}$ for some $i \in \{0, \dots, n\}$, $j \in \{i+1, \dots, n+1\}$ or
(b) $B = M_{k2}$ for some $k \in \{1, \dots, n\}$.

From Theorem 3 or 3′, we see that if $\alpha = \langle (B_1, \alpha_1), \dots, (B_r, \alpha_r) \rangle$ is a feasible information-gathering strategy for D, $B_{r+1} = R(B_r, \alpha_r)$, $t \in \{1, \dots, r\}$, and $B \in B_t$, then B is of either the form given by (a), or by (b) in Theorem 3′, above. Since this is the case, a question of obvious interest is, what is the form of $\iota(B, a)$ for $a \in A_1$? This is the subject matter of the following result.

6.2.5. Theorem. Suppose $i \in \{0, \dots, n\}$, $j \in \{i+1, \dots, n+1\}$, and $k \in \{0, 1, \dots, n+1\}$. Then we have:
(i) (a) $\#\iota(B_{ij}, k) > 1$ if, and only if, $k \in \{i+1, \dots, j-1\}$,
(b) $k \in \{i+1, \dots, j-1\}$ implies $\iota(B_{ij}, k) = \{B_{ik}, M_{k2}, B_{kj}\}$, and
(ii) $\iota(M_{i2}, k) = \{M_{i2}\}$.

Proof. (i) Suppose $i \in \{0, \dots, n\}$, $j \in \{i+1, \dots, n+1\}$, and $k \in \hat{A}$. Then

$$\iota(B_{ij}, k) = \{B_{ij} \cap M_{k1}, B_{ij} \cap M_{k2}, B_{ij} \cap M_{k3}\} \setminus \{\emptyset\}. \tag{25}$$

However, if $k < i$, then by Proposition 2,

$$B_{ij} \cap M_{k1} = (M_{j1} \cap M_{i3}) \cap M_{k1} = M_{j1} \cap (M_{k1} \cap M_{i3}) = M_{j1} \cap \emptyset = \emptyset,$$

and

$$B_{ij} \cap M_{k2} = M_{j1} \cap (M_{i3} \cap M_{k2}) = \emptyset, \quad \text{so that}$$

$$\iota(B_{ij}, k) = \{B_{ij} \cap M_{k3}\}, \quad \text{and}$$

$$B_{ij} \cap M_{k3} = M_{j1} \cap (M_{i3} \cap M_{k3}) = M_{i1} \cap M_{i3} = B_{ij}.$$

Since $M_{k2} \cap M_{k3} = \emptyset = M_{k1} \cap M_{k3}$, it is also clear that if $k = i$,

$$\iota(B_{ij}, k) = \{B_{ij}\}.$$

Similarly, it is easy to show that if $j \le k$, then

$\iota(B_{ij}, k) = \{B_{ij}\}$.

On the other hand, if $k \in \{i+1, \ldots, j-1\}$, then $M_{k1} \subseteq M_{j1}$, and

$$B_{ij} \cap M_{k1} = (M_{j1} \cap M_{i3}) \cap M_{k1} = (M_{j1} \cap M_{k1}) \cap M_{i3} = M_{k1} \cap M_{i3} = B_{ik}$$

and, since $M_{k3} \subseteq M_{i3}$,

$$B_{ij} \cap M_{k3} = M_{j1} \cap (M_{i3} \cap M_{k3}) = M_{j1} \cap M_{k3} = B_{kj};$$

while by (v) of Proposition 2,

$B_{ij} \cap M_{k2} = M_{k2}$.

Thus we see that if $\#\iota(B_{ij,k}) > 1$, $k \in \{i+1, \ldots, j-1\}$; while if $k \in \{i+1, \ldots, j-1\}$,

$\iota(B_{ij}, k) = \{B_{ik}, M_{k2}, B_{kj}\}$,

and thus

$\#\iota(B_{ij}, k) = 3 > 1$.

(ii) It is immediate that if $k = i$, then

$\iota(M_{i2}, k) = \{M_{i2}\}$.

Suppose now that $k \ne i$. Then

$i < k \to M_{i2} \subseteq M_{k1}$, so that $M_{i2} \cap M_{k2} = M_{i2} \cap M_{k3} = \emptyset$, while, by (7) of Lemma 1

$i > k \to M_{i2} \subseteq M_{k3}$, and thus $M_{i2} \cap M_{k1} = M_{i2} \cap M_{k2} = \emptyset$.

Therefore, in all possible cases, we have $\iota(M_{i2}, k) = \{M_{i2}\}$. Q.E.D.

The following two results are more or less immediate implications of the previous two theorems, and their proof will be left to the interested reader.

6.2.6. Corollary. If $\boldsymbol{\alpha} = \langle(\boldsymbol{B}_1, \alpha_1), \ldots, (\boldsymbol{B}_r, \alpha_r)\rangle$ is a feasible information-gathering strategy for $\boldsymbol{D}$, and we define

$\boldsymbol{B}_{r+1} = R(\boldsymbol{B}_r, \alpha_r)$,

then for every $t \in \{1, \ldots, r+1\}$ and every $B \in \boldsymbol{B}_t$, either

(i) there exists $i \in \{0, \ldots, n\}$ and $j \in \{i+1, \ldots, n+1\}$ such that $B = B_{ij}$, or

(ii) there exists $k \in \{1, \ldots, n\}$ such that $B = M_{k2}$. Furthermore, if $\boldsymbol{\alpha}$ is efficient, and t and B are such that $t \in \{1, \ldots, r\}$, $B \in \boldsymbol{B}_t$, and

$\alpha_t(B) \ne 0$,

then there exists $i \in \{0, \ldots, n-1\}$ and $j \in \{i+2, \ldots, n+1\}$ such that

$B = B_{ij}$ and $\alpha_t(B) \in \{i+1, \ldots, j-1\}$.

6.2.7. Corollary. The family $\boldsymbol{B}^A$ contains $2n+1$ elements, and is given by

$\boldsymbol{B}^A = \{M_{12}, M_{22}, \ldots, M_{n2}, B_{01}, B_{12}, \ldots, B_{n,n+1}\}$.

We will be able to use the results obtained here to develop a dynamic programming solution for $\boldsymbol{D}$ for the case where $r \ge n$; which solution will be developed in the next subsection. However, in some cases, most notably that where $\boldsymbol{D}$ has the 'only correct guesses count' form, finding an optimal solution essentially amounts to a matter of finding the feasible final information structure having the largest possible number of elements. In such a context, the following result will be of particular interest (see also 6.2.9, below).

6.2.8. Theorem. If r and n satisfy $2^r \le n+1$, then there exists a (3, 1)-balanced strategy for $\boldsymbol{D}$. Furthermore, if $\boldsymbol{\alpha} = \langle(\boldsymbol{B}_1, \alpha_1), \ldots, (\boldsymbol{B}_r, \alpha_r)\rangle$ is a (3, 1)-balanced strategy for $\boldsymbol{D}$ and $\boldsymbol{B}_{r+1} = R(\boldsymbol{B}_r, \alpha_r)$, then $\#\boldsymbol{B}_{r+1} = 2^{r+1} - 1$.

Proof. We begin be defining p as that unique integer satisfying

$$p \le \log_2(n+1) < p+1, \quad \text{and} \tag{25}$$

$$n^* = 2^p - 1, \tag{26}$$

and we note that

$$n^* \le n \quad \text{and} \quad r \le p. \tag{27}$$

Next, define N_t by

$$N_t = 2^{t-1} - 1 \quad \text{for} \quad t = 2, \ldots, r, \tag{28}$$

and the $r-1$ sequences, $\langle j_q^t \rangle$ by

$$j_q^t = \begin{cases} q2^{[p-(t-1)]} & \text{for} \quad q = 0, 1, \ldots, N_t \\ n+1 & \text{for} \quad q = N_t + 1 = 2^{t-1} \end{cases}$$

$$\text{for} \quad t = 2, \ldots, r. \tag{29}$$

(Since $j_{N_t+1}^t = n+1$ for each t, we shall generally simply write $n+1$ in place of $j_{N_t+1}^t$ in the following.)

We now define the (3, 1)-balanced strategy,

$\boldsymbol{\alpha} = \langle(\boldsymbol{B}_1, \alpha_1), \ldots, (\boldsymbol{B}_r, \alpha_r)\rangle$

inductively, as follows. First let

$$\alpha_1(X) = \frac{n^*+1}{2} = 2^p/2 = 2^{p-1} = j_1^2, \tag{30}$$

and obtain, by Theorem 5

$$\begin{aligned} \boldsymbol{B}_2 &= \{B_{0,j_1^2}, B_{j_1^2,n+1}, M_{j_1^2,2}\} \\ &= \{B_{j_0^2,j_1^2}, B_{j_1^2,n+1}, M_{j_1^2,2}\}, \quad \text{so that} \end{aligned} \tag{31}$$

$$\#\boldsymbol{B}_2 = 3. \tag{32}$$

Now suppose that after the $(t-1)$st step (i.e., after defining $\boldsymbol{B}_{t-1}$ and α_{t-1} for $t \geq 2$, we have obtained

$$\boldsymbol{B}_t = \{B_{j_0^t,j_1^t}, B_{j_1^t,j_2^t}, \ldots, B_{j_{N_t}^t,n+1}, M_{j_1^t,2}, \ldots, M_{j_{N_t}^t,2}\}. \tag{33}$$

Taking

$$\boldsymbol{B}_t^2 = \{B_{j_0^t,j_1^t}, \ldots, B_{j_{N_t}^t,n+1}\}, \quad \text{and} \tag{34}$$

$$\bigcup_{s=2}^{t} \boldsymbol{B}_s^1 = \{M_{j_1^t,2}, \ldots, M_{j_{N_t}^t,2}\}, \tag{35}$$

we define $\alpha_t : \boldsymbol{B}_t \to A$ by

$$\alpha_t(B) = 0 \text{ for } B \in \bigcup_{s=2}^{t} \boldsymbol{B}_s^1, \tag{36}$$

and on $\boldsymbol{B}_t^2$ we define α_t by

$$\alpha_t(B_{j_{q-1}^t,j_q^t}) = \begin{cases} (j_{q-1}^t + j_q^t)/2 & \text{for} \quad q = 1, \ldots, N_t, \\ j_{N_t}^t + 2^{p-t} & \text{for} \quad q = N_t + 1 = 2^{t-1}. \end{cases} \tag{37}$$

Now, for $q = 1, \ldots, N_t$, we have

$$\begin{aligned} \alpha_t(B_{j_{q-1}^t,j_q^t}) &= \frac{j_{q-1}^t + j_q^t}{2} = \frac{2q2^{p-(t-1)} - 2^{p-(t-1)}}{2} \\ &= (2q-1)2^{p-t} = j_{q-1}^t + (j_q^t - j_{q-1}^t)/2, \end{aligned} \tag{38}$$

and

$$j_q^t - j_{q-1}^t = 2^{p+1-t},$$

we see that, for $t \leq r \leq p$,

$$j_{q-1}^t + 1 \leq \alpha_t(B_{j_{q-1}^t,j_q^t}) \leq j_q^t - 1. \tag{39}$$

Furthermore, by (38)

$$\alpha_t(B_{j_{q-1}^t,j_q^t}) = (2q-1)2^{p-t} = j_{2q-1}^{t+1}. \tag{40}$$

Consequently, by (39), (40), and Theorem 5, we have

$$\begin{aligned} &\iota[B_{j_{q-1}^t,j_q^t}, \alpha_t(B_{j_{q-1}^t,j_q^t})] \\ &= \{B_{j_{q-1}^t,j_{2q-1}^{t+1}}, B_{j_{2q-1}^{t+1},j_q^t}, M_{j_{2q-1}^{t+1},2}\} \\ &\text{for} \quad q = 1, \ldots, N_t. \end{aligned} \tag{41}$$

Similarly

$$\begin{aligned} \alpha_t(B_{j_{N_t}^t,n+1}) = j_{N_t}^t + 2^{p-t} &= (2^{t-1}-1)2^{p-(t-1)} + 2^{p-t} \\ &= 2^p - 2^{p-(t-1)} + 2^{p-t} = 2^p - 2^{p-t} \\ &= (2^t - 1)2^{p-t} \\ &= N_{t+1}2^{p-t} = j_{N_{t+1}}^{t+1}, \end{aligned} \tag{42}$$

so that

$$\alpha_t(B_{j_{N_t}^t,n+1}) \geq j_{N_t}^t + 1,$$

and, since $t \leq r \leq p$,

$$\alpha_t(B_{j_{N_t}^t,n+1}) \leq 2^p - 1 \leq n$$

as well. Thus by Theorem 5,

$$\begin{aligned} &\iota[B_{j_{N_t}^t,n+1}, \alpha_t(B_{j_{N_t}^t,n+1})] \\ &= \{B_{j_{N_t}^t,j_{N_{t+1}}^{t+1}}, B_{j_{N_{t+1}}^{t+1},n+2}, M_{j_{N_{t+1}}^{t+1},2}\}. \end{aligned} \tag{43}$$

From (33)–(35), (41), and (43), we see that

$$\#\boldsymbol{B}_{t+1} = 3(\#\boldsymbol{B}_t^2) + \sum_{s=2}^{t} \#\boldsymbol{B}_s^1. \tag{44}$$

Furthermore, letting

$$\boldsymbol{B}_{t+1}^1 = \{M_{j_1^{t+1},2}, M_{j_3^{t+1},2}, \ldots, M_{j_{2q-1}^{t+1},2}, \ldots, M_{j_{N_{t+1}}^{t+1},2}\},$$

we see from (34), (38), and (41) that

$$\#\boldsymbol{B}_{t+1}^1 = \#\boldsymbol{B}_t^2. \tag{45}$$

Finally, we note that

$$j_{q-1}^t = (q-1)2^{p-(t-1)} = [2(q-1)]2^{p-t} = j_{2(q-1)}^{t+1}, \tag{46}$$

and

$$j_q^t = q2^{p-(t-1)} = 2q2^{p-t} = j_{2q}^{t+1} \tag{47}$$

for $q = 1, \ldots, N_{t-1} = 2^{t-1} - 1$; while

$$\begin{aligned} j_{N_t}^t = (2^{t-1}-1)2^{p-(t-1)} &= [(2^t - 1) - 1]2^{p-t} \\ &= j_{N_{t+1}}^{t+1}. \end{aligned} \tag{48}$$

Then, since also

$$N_{t+1} = 2^t - 1 = 2 \cdot (2^{t-1} - 1) + 1 = 2N_t + 1,$$

we see from (41), (43), and (46)–(48) that $\boldsymbol{B}_{t+1}$ has the form

$$\boldsymbol{B}_{t+1} = \{ B_{j_0^{t+1}, j_1^{t+1}}, B_{j_1^{t+1}, j_2^{t+1}}, \ldots, B_{j_{N_{t+1}}^{t+1}, n+1}, M_{j_1^{t+1}, 2}, \ldots, M_{j_{N_{t+2}}^{t+1}} \}.$$

Thus, if $t+1 < r$, we may define α_{t+1} on $\boldsymbol{B}_{t+1}$ by the formulas (34)–(37) (substituting $t+1$ for t). It then follows from (32), (44), and (45) that $\boldsymbol{\alpha}$ is a (3, 1)-balanced strategy. Q.E.D.

According to the above result, if $2^r \le n+1$, then there exists a (3, 1)-balanced strategy for $\boldsymbol{D}$,

$$\boldsymbol{\alpha}^* = \langle (\boldsymbol{B}_1^*, \alpha_1^*), \ldots, (\boldsymbol{B}_r^*, \alpha_r^*) \rangle,$$

and, defining

$$\boldsymbol{B}_{r+1}^* = R(\boldsymbol{B}_r^*, \alpha_r^*),$$

it then follows from Proposition 5.3.2 that

$$\#\boldsymbol{B}_{r+1}^* = 2^{r+1} - 1. \tag{49}$$

It will probably come as no surprise to the reader that no feasible strategy for $\boldsymbol{D}$ can result in a final information structure having a larger number of elements than $2^{r+1} - 1$; however, this statement would nonetheless appear to require a formal proof. In order to provide said proof, we begin by establishing the following, which is a result we shall find useful in other contexts. For purposes of this and the next result, we shall say that a feasible information-gathering process is efficient* if it satisfies (i) and (ii.a) of Definition 3.1.8.

6.2.9. Proposition. If $\boldsymbol{\alpha} = \langle (\boldsymbol{B}_1, \alpha_1), \ldots, (\boldsymbol{B}_r, \alpha_r) \rangle$ is non-trivial and efficient, and we define $\boldsymbol{B}_{r+1} = R(\boldsymbol{B}_r, \alpha_r)$, then*

$$\#\boldsymbol{B}_{t+1} = 2 \sum_{s=1}^{t} \#\boldsymbol{B}_s^2 + 1 \quad \text{for} \quad t = 1, \ldots, r. \tag{50}$$

Proof. We have for $t \in \{1, \ldots, r\}$

$$\begin{aligned} \boldsymbol{B}_{t+1} &= \bigcup_{B \in \boldsymbol{B}_t} \iota[B, \alpha_t(B)] \\ &= \bigcup_{B \in \boldsymbol{B}_t^2} \iota[B, \alpha_t(B)] \\ &\quad \cup \left[\bigcup_{s=2}^{t} \bigcup_{B \in \boldsymbol{B}_s^1} \iota[B, \alpha_t(B)] \right]. \end{aligned}$$

Therefore

$$\#\boldsymbol{B}_{t+1} = \sum_{B \in \boldsymbol{B}_t^2} \#\iota[B, \alpha_t(B)] + \sum_{s=2}^{t} \sum_{B \in \boldsymbol{B}_s^1} \#\iota[B, \alpha_t(B)]. \tag{51}$$

However, since $\boldsymbol{\alpha}$ is efficient,* we have

$$(\forall B \in \boldsymbol{B}_t^2)\colon \#\iota[B, \alpha_t(B)] \ge 2,$$

and it then follows from Theorem 5 that

$$(\forall B \in \boldsymbol{B}_t^2)\colon \#\iota[B, \alpha_t(B)] = 3. \tag{52}$$

Furthermore, by definition of $\boldsymbol{B}_s^1$ (and again using the fact that $\boldsymbol{\alpha}$ is efficient), we have

$$(\forall B \in \boldsymbol{B}_s^1)\colon \#\iota[B, \alpha_t(B)] = 1 \quad \text{for} \quad s = 2, \ldots, t. \tag{53}$$

Substituting (52) and (53) into (51), we then obtain

$$\#\boldsymbol{B}_{t+1} = 3(\#\boldsymbol{B}_t^2) + \sum_{s=2}^{t} \#\boldsymbol{B}_s^1. \tag{54}$$

Now, we also have

$$\#\boldsymbol{B}_{t+1} = \#\boldsymbol{B}_{t+1}^2 + \sum_{s=2}^{t+1} \#\boldsymbol{B}_s^1, \tag{55}$$

and equating (54) and (55), we have

$$\#\boldsymbol{B}_{t+1}^1 = 3(\#\boldsymbol{B}_t^2) - \#\boldsymbol{B}_{t+1}^2 \quad \text{for} \quad t = 1, \ldots, r \tag{56}$$

(for $t = r$, $\#\boldsymbol{B}_{t+1}^2$ is, of course, equal to zero). Since (56) holds for $t = 1, \ldots, r$, it follows that

$$\#\boldsymbol{B}_s^1 = 3(\#\boldsymbol{B}_{s-1}^2) - \#\boldsymbol{B}_s^2 \quad \text{for} \quad s = 2, \ldots, r,$$

and substituting into (54), we have

$$\begin{aligned} \#\boldsymbol{B}_{t+1} &= 3(\#\boldsymbol{B}_t^2) + \sum_{s=2}^{t} [3(\#\boldsymbol{B}_{s-1}^2) - \#\boldsymbol{B}_s^2] \\ &= 3(\#\boldsymbol{B}_t^2) = 3 \sum_{s=2}^{t} \#\boldsymbol{B}_{s-1}^2 - \sum_{s=2}^{t} \#\boldsymbol{B}_s^2 \\ &= 2 \sum_{s=2}^{t} \#\boldsymbol{B}_s^2 + 3(\#\boldsymbol{B}_1^2) \\ &= 2 \sum_{s=1}^{t} \#\boldsymbol{B}_s^2 + (\#\boldsymbol{B}_1^2). \end{aligned}$$

However, since $\boldsymbol{\alpha}$ is non-trivial, we have

$$\#\boldsymbol{B}_1^2 = 1,$$

and the above becomes

$$\#B_{t+1} = \left(2\sum_{s=1}^{t} \#B_s^2\right) + 1 \quad \text{for} \quad t = 1,\ldots,r.$$

Q.E.D.

6.2.10. Corollary. If $\alpha = \langle (B_1, \alpha_1),\ldots,(B_r, \alpha_r)\rangle$ is a feasible information-gathering strategy for D, and we define $B_{r+1} = R(B_r, \alpha_r)$, we have

$$\#B_{r+1} \le 2^{r+1} - 1. \tag{57}$$

Proof. If α is a feasible strategy, then it is easy to see that there exists an efficient* strategy

$$\alpha^* = \langle (B_1^*, \alpha_1^*),\ldots,(B_r^*, \alpha_r^*)\rangle$$

which is such that, defining $B_{r+1}^* = R(B_r^*, \alpha_r^*)$, we have

$$\#B_{r+1}^* \ge \#B_{r+1}.$$

Thus it suffices to prove that (57) holds for the case where α is a non-trival efficient* information-gathering strategy.

From Theorem 6.2.5 it follows that for each non-null information-gathering action taken at the tth step, we obtain at least one element of B^A. Consequently

$$\#B_{t+1}^1 \ge \#B_t^2 \quad \text{for} \quad t = 1,\ldots,r. \tag{58}$$

However, by (56) of the preceding proof, we then obtain

$$\#B_t^2 \le 3(\#B_t^2) - \#B_{t+1}^2, \qquad \text{or}$$

$$\#B_{t+1}^2 \le 2(\#B_t^2) \quad \text{for} \quad t = 1,\ldots,r. \tag{59}$$

Since

$$\#B_1^2 = 1,$$

it follows from (59) and a trivial induction argument that

$$\#B_t^2 \le 2^{t-1} \quad \text{for} \quad t = 1,\ldots,r. \tag{60}$$

Now, by Proposition 9, we have

$$\#B_{r+1} = 2\sum_{s=1}^{r} \#B_s^2 + 1, \tag{61}$$

and substituting (60) into (61), we obtain

$$\#B_{r+1} \le 2\left(\sum_{s=1}^{r} 2^{s-1}\right) + 1$$
$$= 2(2^r - 1) + 1 = 2^{r+1} - 1.$$

Q.E.D.

6.3. A Dynamic Programming Solution for the Trinary Case

In this subsection we shall retain all the assumptions of the previous subsection [(1) and (2) of section 6.2], *and assume, in addition*, that

$$r \ge n. \tag{1}$$

Our object here is to develop a dynamic programming solution for this case and to establish that the strategy obtained is, indeed, optimal. We proceed as follows.

(1) For each $i \in \{0, 1,\ldots, n\}$, we define $\Delta(B_{i,i+1})$ by

$$\Delta(B_{i,i+1}) = \pi(B_{i,i+1})\nu(B_{i,i+1}), \tag{2}$$

and, for each $k \in \{1,\ldots, n\}$, we define

$$\Delta(M_{k2}) = \pi(M_{k2})\nu(M_{k2}). \tag{3}$$

(2) For each $i \in \{0, 1,\ldots, n-1\}$, we calculate the following:

$$w(i+1) = \Delta(B_{i,i+1}) + \Delta(B_{i+1,i+2}) + \Delta(M_{i+1,2}) - \pi(B_{i,i+2})c(i+1), \quad \text{and} \tag{4}$$

$$w(0) = \pi(B_{i,i+2})\nu(B_{i,i+2}). \tag{5}$$

We then define

$$\Delta(B_{i,i+2}) = \max\{w(0), w(i+1)\}, \quad \text{and} \tag{6}$$

$$\hat{a}(i, i+2) = \min\{j \in \{0, i+1\} \mid w(j) = \Delta(B_{i,i+2})\}. \tag{7}$$

(3) For each $i \in \{0, 1,\ldots, n-2\}$, we calculate

$$w(j) = \Delta(B_{ij}) + \Delta(B_{j,i+3}) + \Delta(M_{j2}) - \pi(B_{i,i+3})c(j) \quad \text{for} \quad j = i+1, i+2, \tag{8}$$

$$w(0) = \pi(B_{i,i+3})\nu(B_{i,i+3}), \tag{9}$$

$$\Delta(B_{i,i+3}) = \max\{w(0), w(i+1), w(i+2)\}, \tag{10}$$

and

$$\hat{a}(i, i+3) = \min\{j \in \{0, i+1, i+2\} \mid w(j) = \Delta(B_{i,i+3})\}. \tag{11}$$

(4) Having found $\Delta(B_{i,i+I-1})$, $i = 0, 1,\ldots, n+1-(I-1)$, $I \ge 2$, we compute for $i \in \{0,\ldots, n+1$

$-I\}$

$$w(j)=\Delta(B_{ij})+\Delta(B_{j,i+I})+\Delta(M_{j2}) -\pi(B_{i,i+I})c(j)$$

for $j=i+1,\dots,i+I-1$ (12)

(note that for $j\in\{i+1,\dots,i+I-1\}$, both $j-i \le I-1$ and $i+I-j\le I-1$),

$$w(0)=\pi(B_{i,i+I})\nu(B_{i,i+I}), \tag{13}$$

$$\Delta(B_{i,i+I}) = \max\{w(j)\mid j\in\{0,\ i+1,\dots,i+I-1\}, \tag{14}$$

and

$$\hat{a}(i,\ i+I)=\min\{j\in\{0,\ i+1,\dots,i+I-1\} \mid w(j)=\Delta(B_{i,i+I})\}. \tag{15}$$

(5) Proceeding as above, we eventually obtain

$\Delta(B_{0,n+1})=\Delta(X)$ and $\hat{a}(0,\ n+1)$.

We then define the strategy $\sigma^*=\langle(B_1^*,\ \alpha_1^*),\dots,(B_r^*,\ \alpha_r^*),\ (B_{r+1}^*,\ \delta^*)\rangle$ by

$$\alpha_1^*(X)=\hat{a}(0,\ n+1)\equiv a^*. \tag{16}$$

By Theorem 6.2.5 we then obtain one of two cases.
(a) $\alpha_1^*(X)=0$ and $B_2^*=M_{a^*}=\{X\}$; in which case, we complete the definition of σ^* by defining

$$\alpha_t^*(X)=0 \quad \text{for} \quad t=2,\dots,r, \quad \text{and let} \tag{17}$$

$$d'\in D^*(X). \tag{18}$$

(b) $\alpha_1^*\in\{1,\dots,n\}$ and

$$B_2^*=M_{a^*}=\{B_{0a^*},\ M_{a^*2},\ B_{a^*,n+1}\}.$$

Here we define α_2^* by

$$\alpha_2^*(B_{0,a^*})=\hat{a}(0,\ a^*),\ \alpha_2^*(M_{a^*2}) =0,\ \alpha_2^*(B_{a^*,n+1})=\hat{a}(a^*,\ n+1). \tag{19}$$

Having obtained α_{t-1}^* and B_t^*, for $t\in\{3,\dots,r\}$, we have by Corollary 6.2.6 that each $B\in B_t^*$ is either of the form

$$B=B_{ij} \quad \text{for some} \quad i\in\{0,1,\dots,n\},$$

$$j\in\{i+1,\dots,n+1\}, \tag{20}$$

or is of the form

$$B=M_{k2} \quad \text{for some} \quad k\in\{1,\dots,n\}. \tag{21}$$

We then define α_t^* on B_t^* by

$$\alpha_t^*(B)=\begin{cases}\hat{a}(i,\ j) & \text{if } B \text{ is of the form (20)} \\ & \quad \text{with } j>i+1. \\ 0 & \text{otherwise.}\end{cases} \tag{22}$$

Proceeding in this fashion we eventually obtain B_{r+1}^*, and let $\delta^*(B)$ be an element of $D^*(B)$, for each $B\in B_{r+1}^*$. Q.E.D.

It will be an easy consequence of the following result that α^*, as defined in (17)–(22), above, is optimal for D.

6.3.1. Theorem. If $\sigma=\langle(B_1,\ \alpha_1),\dots,(B_r,\ \alpha_r),\ (B_{r+1},\ \delta)\rangle$ is a feasible strategy for D, $q\in\{1,\dots,r+1\}$, and $i\in\{0,1,\dots,n\}$ and $j\in\{i+1,\dots,n+1\}$ are such that $B_{ij}\in B_q$, then

$$\sum_{B\in B_{r+1}(B_{ij})}\ \sum_{x\in B}\phi(x)\omega[x,\ \delta(B)] - \sum_{t=q}^{r}\ \sum_{B'\in B_t(B_{ij})}\pi(B')c[\alpha_t(B')]\le\Delta(B_{ij}), \tag{23}$$

where for $q=r+1$, we define

$$\sum_{t=q}^{r}\ \sum_{B'\in B_t(B_{ij})}\pi(B')c[\alpha_t(B')]=0.$$

Proof. We distinguish two cases, based on the value of q.
(a) $q=r+1$. Here the left-hand-side of inequality (23) becomes

$$\sum_{x\in B_{ij}}\phi(x)\omega[x,\ \delta(B)], \tag{24}$$

and, since (24) is less than or equal to

$$\pi(B_{ij})\nu(B_{ij})\le\Delta(B_{ij}),$$

The desired inequality follows at once.
(b) $q\in\{1,\dots,r\}$. Here we establish our result for arbitrary $i\in\{0,1,\dots,n\}$ by induction on $I=j-i$, as follows.
(i) $I=1$ (and $j=i+1$). Here we have by Theorem 6.2.5 that

$$B_{r+1}(B_{ij})=\{B_{ij}\},$$

and hence the left-hand-side of (23) becomes

$$\sum_{x\in B_{ij}}\phi(x)\omega[x,\ \delta(B)]$$

$$-\sum_{t=q}^{r} \sum_{B' \in B_t(B_{ij})} \pi(B')c[\alpha_t(B')]$$
$$\leq \sum_{x \in B_{ij}} \phi(x)\omega[x, \delta(B)] \leq \pi(B_{ij})\nu(B_{ij})$$
$$= \Delta(B_{ij}).$$

(ii) Suppose the desired inequality holds for $j = i + I$, where $I \in \{1,\ldots,n-i\}$. Then for $j = i + I + 1$, we have three possible cases.

Case 1. $\alpha_q(B_{ij}) \equiv k \in \{i+1,\ldots,j-1\}$. Here it follows from Theorem 6.2.5 that we can write the left-hand-side of (23) as

$$\sum_{B \in B_{r+1}(B_{ik})} \sum_{x \in B} \phi(x)\omega[x, \delta(B)]$$
$$- \sum_{t=q+1}^{r} \sum_{B' \in B_t(B_{ik})} \pi(B')c[\alpha_t(B')]$$
$$+ \sum_{B \in B_{r+1}(B_{kj})} \sum_{x \in B} \phi(x)\omega[x, \delta(B)]$$
$$- \sum_{t=q+1}^{r} \sum_{B' \in B_t(B_{kj})} \pi(B')c[\alpha_t(B')]$$
$$+ \sum_{B \in B_{r+1}(M_{k2})} \sum_{x \in B} \phi(x)\omega[x, \delta(B)]$$
$$- \sum_{t=q+1}^{r} \sum_{B' \in B_t(M_{k2})} \pi(B')c[\alpha_t(B')] \qquad (25)$$
$$-\pi(B_{ij})c(k).$$

However, since $k \in \{i+1,\ldots,j-1\}$, and $j = i + I$, it follows that $k - i \leq I$ and $j - k \leq I$. Consequently, it follows from our inductive hypothesis that (25) is less than or equal to

$$\Delta(B_{ik}) + \Delta(B_{kj}) + \Delta(M_{k2}) - \pi(B_{ij})c(k)$$
$$\leq \Delta(B_{ij}). \qquad (26)$$

Case 2. $\alpha_q(B_{ij}) \equiv k \notin \{i+1,\ldots,j-1\}$ and $B_{r+1}(B_{ij}) = \{B_{ij}\}$.

Here it is immediate that the left-hand side of (23) is less than or equal to

$$\pi(B_{ij})\nu(B_{ij}) \leq \Delta(B_{ij})$$

Case 3. $\alpha_q(B_{ij}) \equiv k \notin \{i+1,\ldots,j-1\}$ and $B_{r+1}(B_{ij}) \notin \{B_{ij}\}$.

In this case, it follows at once from Theorem 6.2.5 that for some $t \in \{q+1,\ldots,r\}$ we have

$$\alpha_t(B_{ij}) = k' \in \{i+1,\ldots,j-1\}, \quad \text{and} \qquad (27)$$
$$\alpha_s(B_{ij}) \notin \{i+1,\ldots,j-1\}$$
$$\text{for} \quad s = q,\ldots,t-1, \qquad (28)$$

and thus

$$\sum_{B \in B_{r+1}(B_{ij})} \sum_{x \in B} \phi(x)\omega[x, \delta(B)]$$
$$- \sum_{t=q}^{r} \sum_{B' \in B_t(B_{ij})} \pi(B')c[\alpha_t(B')]$$
$$= \sum_{B \in B_{r+1}(B_{ij})} \sum_{x \in B} \phi(x)\omega[x, \delta(B)]$$
$$- \sum_{s=q}^{t-1} \pi(B_{ij})c[\alpha_s(B_{ij})]$$
$$- \sum_{s=t}^{r} \sum_{B' \in B_s(B_{ij})} \pi(B')c[\alpha_s(B')]$$
$$\leq \sum_{B \in B_{r+1}(B_{ij})} \sum_{x \in B} \phi(x)\omega[x, \delta(B)]$$
$$- \sum_{s=t}^{r} \sum_{B' \in B_s(B_{ij})} \pi(B')c[\alpha_s(B)],$$

and it follows from our analysis of the preceding cases that the right-hand side of (28) is less than or equal to $\Delta(B_{ij})$. Q.E.D.

6.3.2. Corollary. The strategy $\sigma^* = \langle (B_1^*, \alpha_1^*), \ldots, (B_r^*, \alpha_r^*), (B_{r+1}^*, \delta^*) \rangle$, defined in (17)–(22), above, is optimal for D.

Proof. It is an immediate consequence of our definition of σ^* (in particular of our definition of $\alpha_1^*(X)$ and δ^*) that

$$\Omega(\sigma^*) - \Gamma(\sigma^*) = \Delta(B_{0,n+1}) = \Delta(X).$$

Consequently, it follows from Theorem 1 that σ^* is optimal for D. Q.E.D.

6.4. The Linear Ordering in the Binary Case

This section parallels section 6.2; although here we suppose that D is a binary information problem. We also suppose, defining $'\succ$ on A_1 by

$$a \,'\!\!\succ a' \Leftrightarrow M_{a'1} \subset M_{a1}, \qquad (1)$$

that $'\succ$ is total on A_1. Thus we can suppose,

without loss of generality that

$$(\forall i,\ j \in A_1):\ j > i \Leftrightarrow M_{i1} \subset M_{j1}. \tag{2}$$

As in section 6.2, it will be convenient to adjoin to A_1 the two elements 0 and $n+1$, where

$$M_0 = \{M_{01}, M_{02}\} \quad \text{with} \quad M_{01} = \emptyset,\ M_{03} = X, \tag{3}$$

and

$$M_{n+1} = \{M_{n+1,1}, M_{n+1,2}\} \quad \text{with}$$
$$M_{n+1,1} = X,\ M_{n+1,2} = \emptyset. \tag{4}$$

As in section 6.2 (Lemma 6.2.1), we can then show that; defining $\hat{A} = A_1 \cup \{0, n+1\}$, we have

$$(\forall i,\ j \in \hat{A}):\ j > i$$
$$\Leftrightarrow \left[M_{i1} \subset M_{j1} \quad \text{and} \quad M_{j2} \subset M_{i2}\right]. \tag{5}$$

With the assumptions and definition set forth in the previous two paragraphs, we can obtain results for the binary case which parallel all the key results for the trinary case. We list the main results here without proof. The proof in all cases is similar to that for the corresponding result in the trinary case, except that it can generally be simplified slightly by virtue of having only two possible outcomes instead of three. In our list to follow, we have skipped some digits in our numbering in order to maintain the convention that 6.4.m always corresponds to 6.2.m.

6.4.3. Theorem. If $B \subseteq X$ is n-feasible, then there exists $i \in \{0,\ldots,n\}$, $j \in \{i+1,\ldots,n+1\}$ such that

$$B = M_{j1} \cap M_{i2}.$$

6.4.4. Definition. For each $i \in \{0,\ldots,n\}$ and each $j \in \{i+1,\ldots,n+1\}$, we define $B_{ij} \subseteq X$ by

$$B_{ij} = M_{j1} \cap M_{i2}.$$

6.4.5. Theorem. Suppose $i \in \{0,\ldots,n\}$, $j \in \{i+1,\ldots,n+1\}$, and $k \in \{0, 1,\ldots,n+1\}$. Then $\#\iota(B_{ij}, k) > 1$ if, and only if, $k \in \{i+1,\ldots,j-1\}$, in which case we have

$$\iota(B_{ij}, k) = \{B_{ik}, B_{kj}\}.$$

6.4.7. Corollary. The family $\boldsymbol{B}^A$ contains $n+1$ elements, and is given by

$$\boldsymbol{B}^A = \{B_{01}, B_{12},\ldots,B_{n,n+1}\}.$$

6.4.8. Theorem. If r and n satisfy $2^r \le n+1$, then there exists a (2, 0)-balanced strategy for $\boldsymbol{D}$. Furthermore, if $\boldsymbol{\alpha} = \langle(\boldsymbol{B}_1, \alpha_1),\ldots,(\boldsymbol{B}_r, \alpha_r)\rangle$ is a (2, 0)-balanced strategy for $\boldsymbol{D}$, and $\boldsymbol{B}_{r+1} = R(\boldsymbol{B}_r, \alpha_r)$, then $\#B_{r+1} = 2^r$.

6.4.9. Proposition. If $\boldsymbol{\alpha} = \langle(\boldsymbol{B}_1, \alpha_1),\ldots,(\boldsymbol{B}_r, \alpha_r)\rangle$ is a feasible information-gathering strategy for $\boldsymbol{D}$, and we define $\boldsymbol{B}_{r+1} = R(\boldsymbol{B}_r, \alpha_r)$, then $\#\boldsymbol{B}_{r+1} \le 2^r$.

We can also define a dynamic programming solution for the binary case in essentially the same fashion in which we proceeded in the trinary case (except that now we need only define $\Delta(B_{ij})$ for $i \in \{0,\ldots,n\}$, $j \in \{i+1,\ldots,n+1\}$; and it can be shown, as before that the resulting strategy is optimal for $\boldsymbol{D}$.

6.5. Applications to the Computer File Search Problem

As we noted earlier, all the maintained assumptions of section 6.2 and 6.3 are satisfied by our model of the computer file search problem (section 4.2); which we can demonstrate as follows. As developed in section 4.2, the information structures for the file search problem take the form

$$\boldsymbol{M}_a = \{M_{a1}, M_{a2}, M_{a3}\}, \qquad \text{where}$$
$$M_{a1} = \{y_1,\ldots,y_{a-1}\} \cup \{z_0,\ldots,z_{a-1}\},$$
$$M_{a2} = \{y_a\}, \qquad \text{and}$$
$$M_{a3} = \{y_{a+1},\ldots,y_n\} \cup \{z_a,\ldots,z_n\},$$

for $a \in \{1,\ldots,n\}$. Therefore

$$a > a' \Leftrightarrow M_{a'1} \cup M_{a'2} \subset M_{a1}, \tag{1}$$

and we see that (1) [and (2)] of section 6.2 is satisfied in this case. The results of sections 6.2 and 6.3 then yield a number of implications for the file search problem; the principal application being the following.

(1) If r and n satisfy

$$2^r \le n+1, \tag{2}$$

then there exists a (3, 1)-balanced strategy for $\boldsymbol{D}$, $\sigma^* = \langle\boldsymbol{\alpha}^*, \boldsymbol{B}^*_{r+1}, \delta^*\rangle$. Furthermore,

$$\#\boldsymbol{B}^*_{r+1} = 2^{r+1} - 1,$$

(Theorem 6.2.8).

(2) If, using the notation of section 6.2, we have

$$p_i = p > 0 \quad \text{for} \quad i = 1,\ldots,n,$$

then the cost of the above strategy, σ^*, is given by

$$\Gamma(\sigma^*) = [r(1+p) - p(2^r - 1)]c, \tag{3}$$

(Corollary 5.3.4).

(3) The dynamic programming solution developed in section 6.3 can be utilized to solve any particular realization of the computer file search problem.

However, we can supplement these results a bit, and at the same time obtain some results applicable to a more general special case of the categorization problem. Consider the following.

*6.5.1. Proposition. Suppose **D** is a categorization problem satisfying*

$$\theta\bar{\omega} > \bar{c}, \tag{4}$$

where $\theta > 0$ and $\bar{c}$ are defined by

$$\theta = \min\{\psi(B) \mid B \in \boldsymbol{B}^A\} \quad \text{and} \quad \bar{c} = \max_{a\in A} c(a),$$

*respectively, Then if $\sigma = \langle(\boldsymbol{B}_1, \alpha_1), \ldots, (\boldsymbol{B}_r, \alpha_r), (\boldsymbol{B}_{r+1}, \delta)\rangle$ is a feasible strategy for **D** such that for some $B^* \in \boldsymbol{B}_{r+1}$, some $q \in \{1, \ldots, r\}$ and some $\bar{a} \in A$ we have*

$$a(q, B^*) = 0 \quad \text{and for some} \quad B^\tau \in \iota(B^*, \bar{a})$$

we have

$$B^\tau \cap X_{\delta(B^*)} = \emptyset, \tag{5}$$

then σ is strictly dominated.

Proof. We have

$$\sum_{B\in\iota(B^*,\bar{a})} \pi(B)\nu(B)$$

$$= \bar{\omega}\left[\sum_{B\in\iota(B^*,\bar{a})\setminus\{B^\tau\}} \psi(B) + \psi(B^\tau)\right]$$

$$\geq \bar{\omega}\left(\sum_{B\in\iota(B^*,\bar{a})\setminus\{B^\tau\}} \pi\left[B \cap X_{\delta(B^*)}\right]\right) + \bar{\omega}\psi(B^\tau)$$

$$= \sum_{x\in B^*} \phi(x)\omega[x, \delta(B^*)] + \bar{\omega}\psi(B^\tau), \tag{6}$$

where the last equality is by the fact that

$$B^\tau \cap X_{\delta(B^*)} = \emptyset.$$

Moreover, by (4), the definition of θ, and the fact that $\pi(B^*) \leq 1$ we have

$$\bar{\omega}\psi(B^\tau) \geq \theta\bar{\omega} > \bar{c} \geq \pi(B^*)c(\bar{a}). \tag{7}$$

From (6) and (7) we see that

$$\sum_{B\in\iota(B^*,\bar{a})} \pi(B)\nu(B) - \sum_{x\in B^*} \phi(x)\omega[x, \delta(B^*)] > \pi(B^*)c(\bar{a}),$$

and it then follows from Proposition 3.2.3 that σ is strictly dominated, Q.E.D.

*6.5.2. Corollary. Suppose **D** is a categorization problem satisfying* (4) *of Proposition* 1, *and that, in addition **D** satisfies*

$$\text{for all n-feasible } B \subseteq X, \text{ if } B \notin \boldsymbol{B}^A, \text{ then for each } d \in D^*(B), \text{ there } \exists\ a^* \in A \text{ (possibly depending on } d) \text{ and } B^* \in \iota(B, a^*) \text{ such that } B^* \cap X_d = \emptyset. \tag{8}$$

*Their if $r \geq n$, and $\sigma = \langle(\boldsymbol{B}_1, \alpha_1), \ldots, (\boldsymbol{B}_r, \alpha_r), (\boldsymbol{B}_{r+1}, \delta)\rangle$ is optimal for **D**, we have $\boldsymbol{B}_{r+1} = \boldsymbol{B}^A$.*

Proof. Suppose $\sigma = \langle(\boldsymbol{B}_1, \alpha_1), \ldots, (\boldsymbol{B}_r, \alpha_r), (\boldsymbol{B}_{r+1}, \delta)\rangle$ is optimal for **D**, let $B \in \boldsymbol{B}_{r+1}$ be arbitrary and suppose, by way of obtaining a contradiction, that $B \notin \boldsymbol{B}^A$. Since σ is optimal for **D**, σ must be efficient, and thus

$$\delta(B) \in D^*(B), \tag{9}$$

and, since $r \geq n$, there must exist $q \in \{1, \ldots, r\}$ such that

$$a(q, B) = 0. \tag{10}$$

However, by (9) and hypothesis (8), there exists $\bar{a} \in A$ and $B^* \in \iota(B, \bar{a})$ such that

$$B^* \cap X_{\delta(B)} = \emptyset,$$

and it then follows from Proposition 1 that σ is strictly dominated; contradicting the assumption that σ is optimal for **D**. Q.E.D.

While condition (8) of the preceding Corollary may look a bit odd, and possibly somewhat artificial, it is satisfied in the computer file search problem; as we shall demonstrate in the proof of the following.

*6.5.3. Proposition. If **D** is a computer file search problem satisfying (in the notation of section* 4.2),

$$\bar{\omega}\min\{p_1, \ldots, p_n, q_0, \ldots, q_n\} > c, \tag{11}$$

*and $r \geq n$, then if $\sigma = \langle(\boldsymbol{B}_1, \alpha_1), \ldots, (\boldsymbol{B}_r, \alpha_r), (\boldsymbol{B}_{r+1}, \delta)\rangle$ is optimal for **D**, we must have*

$$\boldsymbol{B}_{r+1} = \boldsymbol{B}^A = X,$$

and thus $\Omega(\sigma) = \bar{\omega}$.

Proof. The reader will readily see that in order to prove this result it suffices to establish the fact that the computer file search problem satisfies (8) of Corollary 2. To do this, we note that if $B \subseteq X$ is n-feasible, and $B \notin \boldsymbol{B}^A$, then it follows from Theorem 6.2.3′ and Corollary 6.2.7 that there exists $i \in \{0, \dots, n-1\}$ and $j \in \{i+2, \dots, n+1\}$ such that

$$B = B_{ij}. \tag{12}$$

Furthermore, if $d^* \in D^*(B)$, then either $d^* = 0$, or $d^* \in \{i+1, \dots, j-1\}$. Letting $k = i+1$, we have from Theorem 6.2.5, we have that $\iota(B, k) = \{B_{ik}, M_{k2}, B_{kj}\}$. We can thus distinguish two cases.

(a) $d^* = i+1$. In this case

$$B_{ik} \cap X_{d^*} = \emptyset \qquad \left[\text{and } B_{kj} \cap X_{d^*} = \emptyset \text{ as well}\right].$$

(b) $d^* \in \{0, i+2, \dots, j-1\}$. Here we have

$$M_{k2} \cap X_{d^*} = \emptyset.$$

Thus we see that $\boldsymbol{D}$ satisfies condition (8) of Corollary 2, and our conclusion is then an immediate implication of that result. Q.E.D.

It is an immediate implication of the above result that if in the computer file search problem one can reasonably assume that

$$\bar{\omega} \min\{p_1, \dots, p_n, q_0, \dots, q_n\} > c, \tag{11'}$$

and that $r \geq n$, then one can proceed as follows to solve the problem. First find the collection of efficient strategies, call it Σ^*, given by

$$\Sigma^* = \left\{\sigma = \langle(\boldsymbol{\alpha}, \boldsymbol{B}_{r+1}, \mathfrak{I}\rangle \in \Sigma^e \mid \boldsymbol{B}_{r+1} = \boldsymbol{B}^A\right\}.$$

Secondly, find $\sigma^* \in \Sigma^*$ satisfying

$$(\forall \sigma \in \Sigma^*): \Gamma(\sigma) \geq \Gamma(\sigma^*). \tag{13}$$

if $\sigma^* \in \Sigma^*$ satisfies (13), then σ^* will be optimal for $\boldsymbol{D}$. Of course the algorithm developed in section 6.3 will find a σ^* satisfying (13); however, it is interesting to note that, while we have approached the problem from a very different point of view that which has been followed in the computer science literature, the simple (and fairly plausible) condition (11′) and the assumption that $r \geq n$ together imply that the solution obtained here is perfectly consistent with the approach followed in computer science (cf. Aho, Hopcroft, and Ullman [1974, pp. 113–123]).

In connection with this latter point, however, we would have to note that it appears to us at this

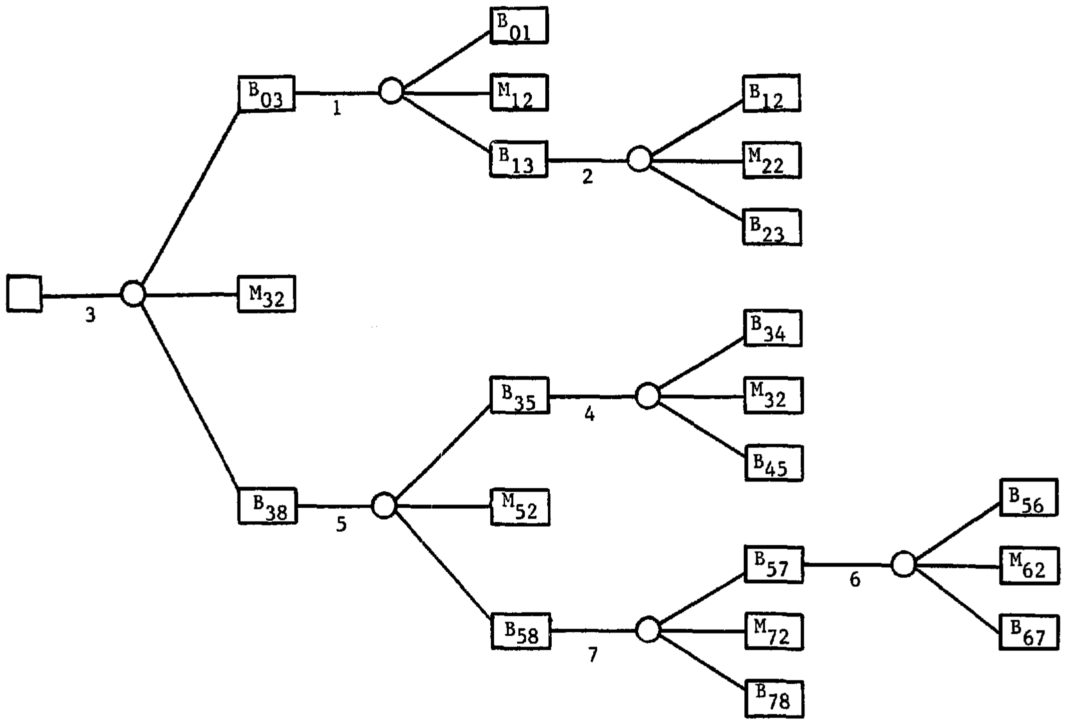

Fig. 5.

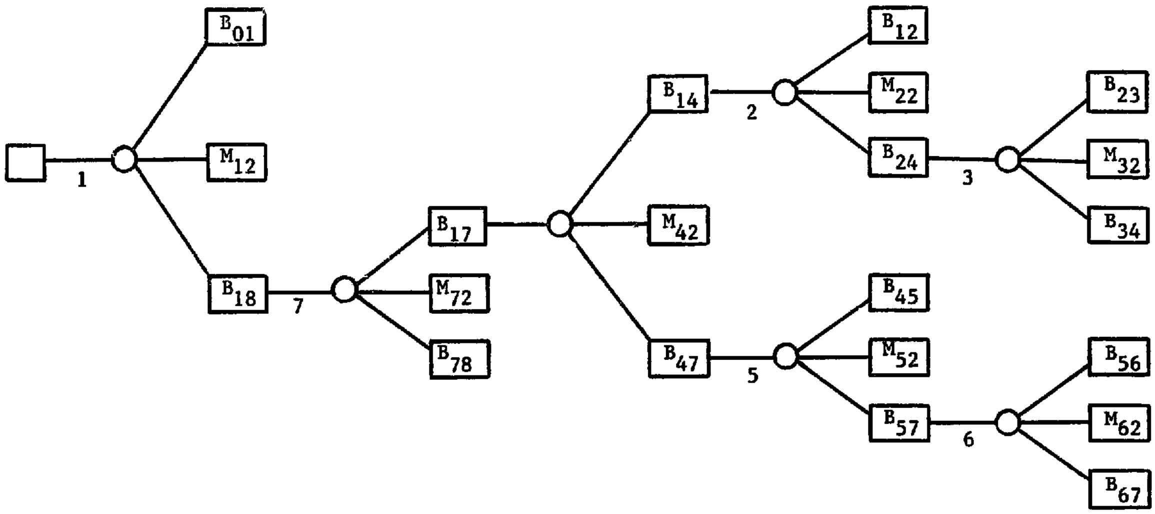

Fig. 6.

juncture that there may have been a bit too much emphasis in the computer science literature on the 'balanced tree' solution to the search problem. This may come about because of a certain ambiguity in that literature as to whether what is most important as a criterion for solution is to (a) minimize 'worst-case' time, or (b) minimize expected cost. The 'balanced tree' solution, which is equivalent to the (3, 1)-balanced strategy solution developed in the proof of Theorem 6.2.8, does minimize the worst-case time of obtaining a solution with certainty (in our terms $\boldsymbol{B}_{r+1} = \boldsymbol{X}$), as is shown by Theorem 6.2.8 and 6.2.9. However, that solution may not minimize the expected cost of obtaining such a solution.[24] At this point, we believe that the following is true:[25] if, in the notation used here and in section 4.2 we have

$p_i = p > 0$ for $i = 1, \ldots, n$, and

$q_i = q > 0$ for $i = 0, 1, \ldots, n$,

(and, of course, $nq + (n+1)q = 1$) then the (3, 1)-balanced strategy of Theorem 6.2.8 statisfies (13), above; and thus is optimal for any value of $r \geq \log_2(n+1)$. On the other hand, with $n = 7$, and

$p_i = p$ for $i = 1, \ldots, 7$, and

$$q_i = q > 0 \quad \text{for} \quad i = 1, \ldots, 6, \tag{14}$$

it can be shown on the basis of the algorithm of Section 6.3 [or on the basis of the algorithm on pp. 119–123 of Aho, Hopcroft, and Ullman, for that matter] that the following is true.

(1) If we also have $q_0 \leq 2q + p$ and $q_7 \leq 2q + p$ then the best (i.e., least) expected costs of strategies which begin with the various possible first steps are as in the table

If the first experiment is	least expected cost is
1	$(q_0 + 3q_7 + 24q + 21p)c$
2	$(2q_0 + 3q_7 + 21q + 19p)c$
3	$(2q_0 + 3q_7 + 20q + 18p)c$
4	$(3q_0 + 3q_7 + 18q + 17p)c = (3 - 4p)c$
5	$(3q_0 + 2q_7 + 20q + 18p)c$
6	$(3q_0 + 2q_7 + 21q + 19p)c$
7	$(3q_0 + q_7 + 24q + 21p)c$

In this case it is easy to see that the best (least-) expected cost obtainable when one begins with experiment 4 (i.e., first compares b with b_4) is no higher than that obtainable with any other possible first step; and it can be shown that the (3, 1)-

[24] We should note, incidentally, that Aho, Hopcroft and Ullman are probably well aware of this fact. See their [1974] work, pp. 119–123.

[25] We have constructed a (formally incomplete) argument to establish this, which is based on the algorithm developed in section 6.3. While we believe that said argument could be refined to provide a formal proof, our best estimate at this point is that it would be about a twenty-page proof, and we decided not to pursue this line of argument further at this point (there really must be a simpler and thus more enlightening way of approaching a proof of this conjecture).

balanced strategy is optimal in this case.[26] However, it is interesting to note that if $q_0 = 2q + p$, then the strategy which begins with experiment 3 has exactly the same expected cost as that which begins with experiment 4. Moreover, we can get somewhat strange-looking solutions in this general situation, as is shown by the following two cases.

(2) Suppose (14) holds and that

$$2q + p < q_0 \leq 3q + p \quad \text{and} \quad q < q_7 \leq 2q + p.$$

Then it can be shown that the search strategy in fig. 5 is optimal (if $r \geq 4$).

(3) Suppose (14) holds and that

$$4q + 3p < q_7 \leq q_0.$$

Then it can be shown that the search strategy in fig. 6 is optimal, if $r \geq 5$.

The reason that we consider cases 2 and 3 to be of some significance is that in general one would suppose that in many, if not most, realizations of a file search problem we would likely have

$$q_0 > q_i \quad \text{for} \quad i = 1, \ldots, 6 \quad \text{and}$$
$$q_7 > q_i \quad \text{for} \quad i = 1, \ldots, 6,$$

(i.e., it is more likely that a randomly-drawn element from U is outside the range of $\{b_1, \ldots, b_7\}$ than that is a 'gap' between some b_i and b_{i+1}). Where this is the case the unbalanced strategy of case 3 may well have a lower expected cost than the balanced strategy of case 1. To take a more specific example, suppose our data set consists of seven integers

$$b_1 = 114,\ b_2 = 126,\ b_3 = 138,\ b_4 = 150,\ b_5 = 162,$$
$$b_6 = 174,\ b_7 = 186.$$

In this case if we take the universal set, U, to be given by

$$U = \{101, \ldots, 199\},$$

and take the probability distribution on U to be uniform; then, multiplying all probabilities by 99, for the sake of convenience, we will have

$$q_0 = q_7 = 13,\ q_2 = q_3 = \cdots = q_6 = 11, \quad \text{and}$$
$$p_i = p = 1 \quad \text{for} \quad i = 1, \ldots, 7.$$

Thus, if we denote the balanced strategy of case 1, above, by σ^b and the strategy of case 3 by σ^e, we have, respectively

$$\Gamma(\sigma^b) = (297 - 4)c = 293c, \quad \text{and}$$
$$\Gamma(\sigma^e) = 371c.$$

Thus in this case, the expected cost of the balanced strategy, σ^b, is significantly lower than that of the extremes/balance strategy, σ^e. However, suppose we change the example by taking

$$U = \{1, \ldots, 250\},$$

retaining our assumption of a uniform distribution on U. If we again normalize (this time multiply by 250) the values of p and $q_1, \ldots, q_6$ remain as before. Thus the expected cost of σ^b is given by

$$\Gamma(\sigma^b) = 746c.$$

On the other hand, we have

$$q_0 = 113,\ q_7 = 64, \quad \text{and}$$
$$\Gamma(\sigma^e) = 573c.$$

The extreme-balance strategy σ^e thus has a significantly lower expected cost than that for σ^b with the second universal set.[27] The optimal expected cost solution for the computer file search problem appears, therefore, to be fairly sensitive to the specification of the universal set U. Whether this is a fact which should be of great concern to computer scientists is, however, a question upon which we shall not speculate futher.

In closing this section, we believe that it is of interest to note that all of the general results obtained here hold if $\mathbf{D}$ is a general search problem, defined as follows.

6.5.4. Definition. We shall say that a decision problem, $\mathbf{D}$ is a general search problem iff $\mathbf{D}$:

(1) is a categorization problem,

(2) satisfies the assumptions of section 6.2, so that we can write

$$\mathbf{B}^A = \{M_{12}, \ldots, M_{n2}, B_{01}, B_{12}, \ldots, B_{n,n+1}\},$$

(3) satifies the condition: for each $d \in \{0, 1, \ldots, p\}$, and each $i \in \{1, \ldots, n\}$,

$$M_{i-1,i} \cap X_d \neq \emptyset \rightarrow M_{i2} \cap X_d = \emptyset,$$

(4) has constant information cost, $c > 0$.

Reference

Aho, A.V., J.E. Hopcroft, and J.D. Ullman, The Design and Analysis of Computer Algorithms, Addison-Wesley, Reading, MA (1974).

[26] Notice that the expected cost for the balanced strategy, as set forth in the table, is given by $(3q_0 + 3q_7 + 18q + 17p)c = [3(q_0 + q_7 + 6q + 7p) - 4p]c = (3 - 4p)c$, which agrees with the formula stated earlier in this section (and which was derived from Corollary 5.3.4), since with $r = 3$, $[r(1 + p) - p(2^r - 1)]c = (3 + 3p - 7p)c = (3 - 4)c$.

[27] The condition of Case 3 is satisfied in this case, and thus σ^e is the optimal solution for this example.