KYBERNETIKA — VOLUME 17 (1981), NUMBER 4

AN ALTERNATIVE METHOD FOR CONSTRUCTION OF OPTIMAL SEQUENTIAL QUESTIONNAIRES

RADIM JIROUŠEK

The paper presents an alternative procedure for construction of optimal sequential questionnaires. This procedure, based on often used Branch-and-Bound method, may be also used for implementation of suboptimal questionnaires when the problem under consideration is too large.

The sequential questionnaire is a formalized expression of a sequential decision scheme in which each test should depend on the results of the previous tests. In the paper the knowledge of a decision function which is to be realized is assumed.

INTRODUCTION

It has been shown in [5] that a concept of sequential questionnaire can be established in the frame of a classical model of statistical decision-making. Conditions under which the questionnaire realizing the given decision function exists are also given in [5]. From Theorem 3 of that paper follows that without a great loss of generality one may consider only finite questionnaires and a finite set of questions.

A measurable sample, or observation, space $(Y, \mathscr{Y})$ (where Y is a set of possible observations and $\mathscr{Y}$ is a σ-algebra of subsests of the set Y) and a set of decisions Z will be considered. Further, let P_Y be an a posteriori probability distribution defined on $(Y, \mathscr{Y})$.

Within this frame, a question is formalized as a partition of the space Y. In order to simplify some expressions, all questions are supposed to have the same number of answers.

Definition 1. Let α be an integer; $\alpha > 1$. The *question* $Q = \{q^k\}_{k=1}^{\alpha}$ is a $\mathscr{Y}$-measurable partition of Y (q^k represent answers). A *detector* $\mathscr{Q}$ is the finite set of questions $\mathscr{Q} = \{Q_j\}_{j=1}^{M}$, which is assumed to be fixed throughout this paper.

A σ-algebra generated by the system of all answers

$$\mathscr{B}_{\mathscr{Q}} = \{\{q_j^k\}_{k=1}^{\alpha}\}_{Q_j \in \mathscr{Q}}$$

will be denoted by the symbol $\mathscr{Y}_{\mathscr{Q}}$ (since $\mathscr{Q}$ is supposed to be finite, $\mathscr{Y}_{\mathscr{Q}}$ is an algebra). For the purposes of this paper, only $\mathscr{Y}_{\mathscr{Q}}$-measurable decision functions $d : Y \to Z$ will be considered.

Definition 2. The *questionnaire* is an ordered triplet $\Delta = (\mathbf{G}, \mathbf{g}, \mathbf{h})$ in which:

– $\mathbf{G}$ is a connected oriented graph with the following properties:

(i) There is one and only one node u_0 in $\mathbf{G}$ (it is referred to as the *root* of the questionnaire Δ) in which no edge terminates.
(ii) There are starts of exactly α edges in each nonterminal node.
(iii) All nodes of $\mathbf{G}$ distinct from u_0 are accessible from u_0 by unique path.

– $\mathbf{g}$ is such mapping from the set of all nonterminal nodes of $\mathbf{G}$ into the detector $\mathscr{Q}$ that its restriction to an arbitrary path from $\mathbf{G}$ is an injection into $\mathscr{Q}$.

– $\mathbf{h}$ is such mapping from the set of all edges into the system of all answers $\mathscr{B}_{\mathscr{Q}}$ that its restriction to the set of edges starting at u is a bijection into the set of answers to the question $\mathbf{g}(u)$.

V will denote the set of all nodes of the questionnaire under consideration. Similarly, U, W will denote internal, terminal nodes respectively.

The use of the questionnaire is quite clear. After the question $g(u_0)$ has been asked, then according to the answer obtained the appropriate son u is selected and the question $g(u)$ is asked. This procedure continues until a terminal node is reached. Our interest focuses on a questionnaire enabling unique decision after a terminal node has been reached (see Definition 4).

Now, the set $h(v) \subset Y$ will be defined for every node v of the questionnaire Δ. The set $h(v)$ will reflect the state of knowledge obtained in the node v.

For the root u_0 of the questionnaire the set $h(u_0)$ is defined as $h(u_0) = Y$. Consider an arbitrary node $v \neq u_0$ of the questipnnaire $\Delta = (\mathbf{G}, \mathbf{g}, \mathbf{h})$. According to the assumption (iii) of the definition of the questionnaire there exists unique path from u_0 to v. Denote the nodes of this path by

$$u_0 = v_0, v_1, \ldots, v_l = v\,.$$

Then $h(v)$ is defined by

$$h(v) = \bigcap_{k=0}^{l-1} \mathbf{h}(v_k v_{k+1})\,.$$

The value $\pi(v) = P_Y(h(v))$ is called the *probability of attaining* the node v.

The set of all questionnaires is partially ordered by the following relationship:

Definition 3. Let u_0, $\bar{u}_0$ be roots and V, $\bar{V}$ be sets of all nodes of questionnaires $\Delta = (\mathbf{G}, \mathbf{g}, \mathbf{h})$, $\bar{\Delta} = \bar{\mathbf{G}}(, \bar{\mathbf{g}}, \bar{\mathbf{h}})$ respectively. The questionnaire Δ is the *beginning* of the questionnaire $\bar{\Delta}(\Delta \prec \bar{\Delta})$ if there exists an injection $\nu : V \to \bar{V}$ with the following properties:

(i) $v(u_0) = \bar{u}_0$,
(ii) $\bar{\mathbf{g}}(v(u)) = \mathbf{g}(u)$ for all internal nodes u of $\mathbf{G}$,
(iii) If (uv) is an edge of $\mathbf{G}$ then $(v(u)\, v(v))$ is an edge of $\bar{\mathbf{G}}$ and $\bar{\mathbf{h}}(v(u)\, v(v)) = \mathbf{h}(uv)$.

Remark. In this paper Δ_0 will denote a questionnaire with only one node u_0 (i.e. the root of Δ_0 is simultaneously the terminal node). In this special case the set of internal nodes is empty, and it is easy to show that

$$\Delta_0 \prec \Delta$$

for any arbitrary questionnaire Δ.

The following definition expresses relationship between the questionnaires and decision functions.

Definition 4. The questionnaire Δ *realizes* the decision function d if the function d is constant on $h(w)$ for all terminal nodes w of the questionnaire Δ.

The questionnaire Δ *realizes almost everywhere* (a.e.) the decision function d if there exists a decision function d' realized by the questionnaire Δ and $P_Y\{y \in Y : : d(y) \neq d'(y)\} = 0$.

OPTIMALITY CRITERION

Let d be a $\mathcal{Y}_{\mathcal{Q}}$-measurable decision function which is to be realized by a questionnaire. Denote $\boldsymbol{T}_d$ the set of all questionnaires realizing a.e. the function d. The number of questionnaires of $\boldsymbol{T}_d$ may be estimated by

$$|\boldsymbol{T}_d| \geqq \prod_{i=0}^{M-1} (M - i)^{(\alpha^i)} . \tag{1}$$

(M is the number of questions).

Among this large number of questionnaires one may wish to select a questionnaire which is optimal from some viewpoint. In this paper, we will take a generalized average length as the optimization criterion.

Let us assume that a positive *cost function* c is defined on the detector

$$c : \mathcal{Q} \to R^+ ,$$

and denote

$$c_{\min} = \min_{1 \leqq j \leqq M} (c(Q_j)) .$$

Definition 5. *The generalized average length* of the questionnaire $\Delta = (\mathbf{G}, \mathbf{g}, \mathbf{h})$ is defined by

$$\bar{L}(\Delta) = \sum_{u \in U} \pi(u)\, c(\mathbf{g}(u)) .$$

The aim is to construct a questionnaire from $\boldsymbol{T}_d$ with the lowest possible generalized average length. It is obvious that the trivial algorithm examining all possible questionnaires from $\boldsymbol{T}_d$ cannot be used because of the inequality (1).

For the construction of the optimal questionnaire the Payne-Meisel algorithm ([3]) can be used. But this algorithm can be used only for small detectors. This algorithm is based on producing a sequence of all *Regular Lattices* which correspond, in the terminology used in this paper, to a system of special subsets of Y. The number of sets from this system is equal to

$$\sum_{i=0}^{M} \binom{M}{i} \alpha^{M-i} = (\alpha + 1)^M .$$

An alternative algorithm proposed in the presented paper is based on the widely used Branch-and-Bound method ([1], [2]). For reasons of practical applicability it is slightly modified.

Before describing the algorithm some useful properties should be presented.

PROPERTIES OF THE QUESTIONNAIRES FROM $\boldsymbol{T}_d$

Assume that $\bar{\Delta} = (\bar{\mathbf{G}}, \bar{\mathbf{g}}, \bar{\mathbf{h}}) \in \boldsymbol{T}_d$ and that $\Delta = (\mathbf{G}, \mathbf{g}, \mathbf{h})$ (not necessarily from $\boldsymbol{T}_d$) is the beginning of $\bar{\Delta}$, i.e. $\Delta \prec \bar{\Delta}$. The proposed algorithm takes advantage of the relationship between the generalized average length of the questionnaires Δ and $\bar{\Delta}$.

Let w be a node of the questionnaire Δ. Define $\mathrm{H}_d(w) = 0$ if $\pi(w) = 0$, otherwise define

$$\mathrm{H}_d(w) = - \sum_{z \in Z} P_d(z \mid w) \log P_d(z \mid w) ,$$

where

$$P_d(z \mid w) = \frac{1}{\pi(w)} P_Y(h(w) \cap d^{-1}(z)) ,$$

and

$$d^{-1}(z) = \{y \in Y : d(y) = z\} .$$

Theorem 1. If $\Delta \prec \bar{\Delta}$ and $\bar{\Delta} \in \boldsymbol{T}_d$, then

$$\bar{L}(\bar{\Delta}) \geqq \bar{L}(\Delta) + \frac{c_{\min}}{\log \alpha} \sum_{w \in W} \pi(w) \mathrm{H}_d(w) . \tag{2}$$

(W denotes the set of terminal nodes of Δ.)

The proof of the theorem is given in [4] and for the average length of a questionnaire (equal to the generalized average length when $c(Q) \equiv 1$) in [5].

Theorem 2. $\Delta = (\mathbf{G}, \mathbf{g}, \mathbf{h})$ realizes a.e. the decision function d *iff* $\mathrm{H}_d(w) = 0$ for all terminal nodes w.

Proof. Suppose

$$(\forall w \in W) \quad (\mathrm{H}_d(w) = 0) .$$

Then for each $w \in W$

$$-\sum_{z \in Z} P_d(z \mid w) \log P_d(z \mid w) = 0$$

or $\pi(w) = 0$.

But it is a well-known property of Shannon entropy that the previous sum is eoual to 0 *iff* one of the terms $P_d(z \mid w)$ is equal to 1 and the others are equal to 0 (recall that $\sum_{z \in Z} P_d(z \mid w) = 1$). Thus for some $z_0 \in Z$

$$P_d(z_0 \mid w) = \frac{1}{\pi(w)} P_Y(h(w) \cap d^{-1}(z_0)) = 1$$

and therefore

$$P_Y(h(w) \cap d^{-1}(z_0)) = \pi(w) = P_Y(h(w)) .$$

But from this follows that $d(y) = z_0$ for $y \in h(w)$ a.e. and thus $\Delta \in \boldsymbol{T}_d$.

On the other hand assume that for some terminal node w' $H_d(w') > 0$ and that $\Delta \in \boldsymbol{T}_d$. Since $\Delta \prec \Delta$ then, according to the preceding theorem,

$$\bar{L}(\Delta) \geqq \bar{L}(\Delta) + \frac{c_{\min}}{\log \alpha} \sum_{w \in W} \pi(w) H_d(w) \geqq \bar{L}(\Delta) + \frac{c_{\min}}{\log \alpha} \pi(w') H_d(w') > \bar{L}(\Delta)$$

which is a contradiction. □

The last theorem may be used to test questionnaires as to whether or not it realizes a.e. function d as the values $H_d(w)$ should be computed for all nodes when constructing the questionnaire according to the proposed algorithm.

Let $\Delta = (\mathbf{G}, \mathbf{g}, \mathbf{h})$ be a questionnaire which is not from $\boldsymbol{T}_d$. Then there must exist a terminal node w in the Δ from which

$$\mathrm{H}_d(w) > 0 .$$

We want to describe a new questionnaire which arises from Δ by assigning a question Q to the node w and by joining α new terminal nodes connected with w by edges. This new questionnaire will be denoted by $\Delta(w, Q)$. Now, a precise definition of this questionnaire is presented.

Definition 6. Let $\Delta = (\mathbf{G}, \mathbf{g}, \mathbf{h})$ be a questionnaire with a terminal node w and let a question Q be assigned to no internal node of the path from the root to w. $\Delta(w, Q) = (\mathbf{G}', \mathbf{g}', \mathbf{h}')$ will denote such questionnaire for which:

(i) $\Delta \prec \Delta(w, Q)$,

(ii) $|V| + \alpha = |V'|$,

(iii) $v(w)$ is internal node of $\Delta(w, Q)$,

(iv) $\mathbf{g}(v(w)) = Q$,

where $|V|$ and $|V'|$ are the numbers of nodes of the questionnaires Δ and $\Delta(w, Q)$ respectively, and v is the mapping from the definition of the relationship $\Delta \prec \Delta(w, Q)$.

Assume that $\Delta \prec \bar{\Delta} = (\bar{\mathbf{G}}, \bar{\mathbf{g}}, \bar{\mathbf{h}}) \in \boldsymbol{T}_d$ and that w is a terminal node of Δ such that $\mathrm{H}_d(w) > 0$. Let $\bar{v}$ be the mapping from the definition of $\Delta \prec \bar{\Delta}$. Since $\mathrm{H}_d(w) =$

$= \mathrm{H}_d(\bar{v}(w)) > 0$ (following from the assumption (iii) of Definition 3) the node $\bar{v}(w)$ must be (according to Theorem 2) an internal node of $\bar{\Delta}$. Denote $Q_j = \bar{g}(\bar{v}(w))$. It would be space consuming to present an exact proof, nevertheless, it is obvious that

$$\Delta(w, Q_j) \prec \bar{\Delta} .$$

Thus we have obtained the following assertion.

Theorem 3. Let $\Delta \prec \bar{\Delta} \in \boldsymbol{T}_d$ and w be a terminal node of Δ such that $\mathrm{H}_d(w) > 0$. Then there must exist a question $Q_j \in \mathscr{Q}$ such that $\Delta(w, Q_j) \prec \bar{\Delta}$.

BRANCH-AND-BOUND METHOD

We are still supposing that $\boldsymbol{T}_d$ denotes the set of all questionnaires realizing a.e. the given decision function d and that the function d is $\mathscr{Y}_{\mathscr{Q}}$-measurable for some fixed detector $\mathscr{Q} = \{Q_j\}_{j=1}^M$.

The fundamental idea of the Branch-and-Bound method is the following one: First define, in some way, the partition $\beta^1(\boldsymbol{T}_d)$ of the set $\boldsymbol{T}_d$. Then eliminate (if possible) from $\beta^1(\boldsymbol{T}_d)$ all those subsets $\tau \subset \boldsymbol{T}_d$ about which one may be sure that τ does not contain an optimal questionnaire. Define a new system $\beta^2(\boldsymbol{T}_d)$ of subsets of $\boldsymbol{T}_d$ which is a compound of some partition of the remaining subsets from $\beta^1(\boldsymbol{T}_d)$. Elimination of subsets from $\beta^i(\boldsymbol{T}_d)$ and generation of new systems $\beta^{i+1}(\boldsymbol{T}_d)$ continues till the system $\beta^n(\boldsymbol{T}_d)$ (for some n) contains one entry set $\{\Delta\} \in \beta^n(\boldsymbol{T}_d)$ and Δ is optimal.

An elimination of subsets from $\beta^i(\boldsymbol{T}_d)$ is made possible by lower an upper bounding rules (c.f. [1] and [2]).

A lower bounding rule is a function

$$b : \{\tau : (\exists n)\,(\tau \in \beta^n(\boldsymbol{T}_d))\} \to R$$

with the following properties:

(L 1) $$\Delta \in \tau \;\Rightarrow \bar{L}(\Delta) \geqq b(\tau) ,$$

(L 2) $$b(\{\Delta\}) = \bar{L}(\Delta) .$$

An upper bounding rule is a function

$$B : \{\beta^i(\boldsymbol{T}_d)\}_{i=0}^{\infty} \to R$$

with the following properties:

(U 1) $$B(\beta^n(\boldsymbol{T}_d)) \geqq \bar{L}(\Delta_{\mathrm{OPT}}), {}^{*})$$

(U 2) $$\{\Delta\} \in \beta^n(\boldsymbol{T}_d) \Rightarrow B(\beta^n(\boldsymbol{T}_d)) \leqq \bar{L}(\Delta) .$$

*) $\bar{L}(\Delta_{\mathrm{OPT}})$ denotes the generalized average length of an optimal questionnaire.

It is obvious that such $\tau \in \beta^i(\boldsymbol{T}_d)$ for which

$$b(\tau) > B(\beta^i(\boldsymbol{T}_d))$$

contains no optimal questionnaire and therefore may be eliminated from $\beta^i(\boldsymbol{T}_d)$.

In order to be able to define the lower bounding rule in an advantageous manner, the systems $\beta^i(\boldsymbol{T}_d)$ of subsets of $\boldsymbol{T}_d$ are defined to be of a special form; $\beta^i(\boldsymbol{T}_d)$ have to meet the following condition:

$$(\forall i)\,(\forall \tau \in \beta^i(\boldsymbol{T}_d))\,(\exists \Delta_\tau)\,(\tau = \{\Delta : \Delta_\tau \prec \Delta\} \cap \boldsymbol{T}_d)\,.$$

In this case the lower bounding rule may be defined according to (2)

$$b(\tau) = \bar{L}(\Delta_\tau) + \frac{c_{\min}}{\log \alpha} \sum_{w \in W_\tau} \pi(w)\,\mathrm{H}_d(w)\,.$$

The upper bounding rule may be defined

$$B(\beta^i(\boldsymbol{T}_d)) = \min_{\tau \in I(\beta^i(\boldsymbol{T}_d))} b(\tau) \quad \text{if} \quad I(\beta^i(\boldsymbol{T}_d)) \neq \emptyset\,,$$

$$B(\beta^i(\boldsymbol{T}_d)) = \infty \qquad \text{if} \quad I(\beta^i(\boldsymbol{T}_d) = \emptyset\,,$$

where

$$I(\beta^i(\boldsymbol{T}_d)) = \{\tau \in \beta^n(T) : \Delta_\tau \in \boldsymbol{T}_d\}\,.$$

The chosen special form of subsets $\tau \in \beta^i(\boldsymbol{T}_d)$ has another advantage. All subsets τ from $\beta^i(\boldsymbol{T}_d)$ may be represented only by its characteristic questionnaire Δ_τ and the whole system $\beta^i(\boldsymbol{T}_d)$ may be represented by a state of stack containing a set of questionnaires $\{\Delta_\tau\}_{\tau \in \beta^i(\boldsymbol{T}_d)}$. Moreover, when constructing a set $\tau \in \beta^i(\boldsymbol{T}_d)$ such that $\Delta_\tau \in \boldsymbol{T}_d$, then all the other questionnaires from τ may be omitted because of

$$\Delta_\tau \prec \Delta \Rightarrow \bar{L}(\Delta_\tau) \leqq \bar{L}(\Delta)\,.$$

Before describing the algorithm step by step, it should be mentioned how the branching rule is implemented, i.e. how the partitions of some subset τ are defined.

Let τ be a subset of $\boldsymbol{T}_d$ with its characteristic questionnaire Δ_τ. Next, suppose that $\Delta_\tau \notin \boldsymbol{T}_d$. If it is not the case, all the other questionnaires, except for Δ_τ, should be omitted, and thus, τ would contain only one questionnaire Δ_τ.

According to Theorem 2 there exists a terminal node w of Δ_τ such that $\mathrm{H}_d(w) > 0$. Let $\mathscr{Q}' \subset \mathscr{Q}$ be a subset of those questions from the detector $\mathscr{Q}$ which are assigned to no node on the path from the root to the node w. $\mathscr{Q}'$ is nonempty because of $\mathscr{Y}_{\mathscr{Q}}$-measurability of d and $\mathrm{H}_d(w) > 0$. Denote

$$\tau(Q) = \{\Delta \in \tau : \Delta_\tau(w, Q) \prec \Delta\}\,.$$

According to Theorem 3 to every questionnaire $\Delta \in \tau$ there exists a question $Q \in \mathscr{Q}'$ such that $\Delta_\tau(w, Q) \prec \Delta$ and, thus, $(\tau(Q))_{Q \in \mathscr{Q}'}$ is a partition of τ.

ALGORITHM

From the preceding discussion follows that the algorithm for the construction of the optimal questionnaire realizing a.e. the decision function d by the Branch-and-Bound method may be expressed in the following seven steps.

(A 1): Set $B \leftarrow \infty$;
Compute $b(\Delta_0)$;
Insert the pair $(\Delta_0, b(\Delta_0))$ into the stack;

(A 2): If the stack is empty go to (A 7);
Look up and remove the pair $(\Delta_S, b(\Delta_S))$ from the stack with the smallest value of $b(\Delta_S)$;
Set $\Delta \leftarrow \Delta_S$;
Set $b(\Delta) \leftarrow b(\Delta_S)$;

(A 3): If $b(\Delta) \geqq B$ go to (A 2);
If $\Delta \notin \boldsymbol{T}_d$ go to (A 4); *)
Set $\bar{\Delta} \leftarrow \Delta$;
Set $B \leftarrow b(\Delta)$;
Go to (A 2);

(A 4): Look up the terminal node w of Δ such that $H_d(w) > 0$;
Set $j \leftarrow 0$;

(A 5): Set $j \leftarrow j + 1$;
If $j > M$ go to (A 2);

(A 6): If the question Q_j is assigned to a node of the path from the root to w go to (A 5);
Construct $\Delta(w, Q_j)$;
Compute $b(\Delta(w, Q_j))$;
Insert the pair $(\Delta(w, Q), b(\Delta(w, Q_j)))$ into the stack;
Go to (A 5);

(A 7): Stop.
The questionnaire $\bar{\Delta}$ is the optimal questionnaire realizing a.e. the decision function d. Its generalized average length is equal to B.

ADDITIONAL REQUEST

The previous algorithm has a great disadvantage. Unfortunately, it is possible to construct a case when the algorithm works as the trivial one. That simply means that all questionnaires from $\boldsymbol{T}_d$ must be constructed before the algorithm finishes its work.

*) For testing whether $\Delta \in \boldsymbol{T}_d$ use Theorem 2.

It is very improbable that this case would occur in a practical application. But one must take into account that in some applications it might happen that a pretty large number of questionnaires will have to be constructed. But only a limited extent of the stack is available. For that reason we may want to alter the previous algorithm to obtain a new algorithm which would make use of only a limited extent of the stack.

It is easy to show that there is no possibility to modify the algorithm in the required manner without any influence on the results. If we want to have these modifications performed, we must permit situation that the algorithm will sometime offer us no guarantees that the resulting questionnaire would be the optimal one.

The requested modifications results in the following adaptable algorithm.

ADAPTABLE ALGORITHM

(AD 1): Set $B \leftarrow \infty$;
Set $BT \leftarrow \infty$;
Set $N \leftarrow$ available extent of the stack;
Set $k \leftarrow \frac{1}{2}$;
Compute $b(\Delta_0)$;
Insert the pair $(\Delta_0, b(\Delta_0))$ on top of the (pushdown) stack;

(AD 2): If the stack is empty go to (AD 13);
If there are more than $k \cdot N$ empty cells in the stack go to (AD 3), otherwise go to (AD 4);

(AD 3): Look up the pair $(\Delta', b(\Delta'))$ with the lowest value $b(\Delta')$ in the stack and set it on top of the stack;

(AD 4): Remove the top pair $(\Delta_S, b(\Delta_S))$ from the stack;
Set $\Delta \leftarrow \Delta_S$;
Set $b(\Delta) \leftarrow b(\Delta_S)$;

(AD 5): If $b(\Delta) \geqq B$ go to (AD 2);
If $\Delta \notin \boldsymbol{T}_d$ go to (AD 6);
Set $\bar{\Delta} \leftarrow \Delta$;
Set $B \leftarrow b(\Delta)$;
Go to (AD 2);

(AD 6): Look up the terminal node w of Δ such that $\mathrm{H}_d(w) > 0$;
Set $j \leftarrow 0$;

(AD 7): Set $j = j + 1$;
If $j > M$ go to (AD 2);

(AD 8): If the question Q_j is assigned to a node of the path from the root to w go to (AD 7);

Construct $\Delta(w, Q_j)$:
Compute $b(\Delta(w, Q_j))$;
Set $\Delta_D \leftarrow \Delta(w, Q_j)$;
Set $b(\Delta_D) \leftarrow b(\Delta(w, Q_j))$;

(AD 9): If $b(\Delta_D) \geqq B$ go to (AD 7);
If there is empty cell in the stack go to (AD 12);
Look up and remove from the stack the pair $(\Delta_S, b(\Delta_S))$ with the highest value of $b(\Delta_S)$,;

(AD 10): If $b(\Delta_S) \geqq B$ go to (AD 12);
If $\Delta_S \in \boldsymbol{T}_d$ go to (AD 11);
Set $BT \leftarrow b(\Delta_S)$;
Go to (AD 12);

(AD 11): Set $\bar{\Delta} \leftarrow \Delta_S$;
Set $B \leftarrow b(\Delta_S)$;

(AD 12): Insert the pair $(\Delta_D, b(\Delta_D))$ on top of the stack;
Go to (AD 7);

(AD 13): Stop.
The questionnaire $\bar{\Delta}$ is a questionnaire realizing a.e. the decision function d. Its generalized average length is equal to B. If $B \leqq BT$ then $\bar{\Delta}$ is optimal.

NOTES ON ADAPTABLE ALGORITHM

In addition to some initialization, the values of the two parameters k and N are defined in the step (AD 1) of the previous algorithm. The choice of these values may influence the behaviour of the algorithm.

If the avaliable extent of the stack is sufficient for the given task, then choosing $k = 0$ one can cause the adaptable algorithm to function as the nonadaptable one. If the extent of the stack is uncertain as to its sufficiency and, in spite of this fact, one wants to construct the optimal questionnaire then it is convenient to choose $k = 1$. In this case the expectancy of obtaining the optimal questionnaire will grow but this strategy is rather time consuming. On the other hand, by assigning a small value to $N (N \geqq 2)$, it is possible to shorten the time required but, of course, simultaneously, to make lower the probability of achieving the optimal questionnaire.

From the practical point of view there may be another reasonable requirement. Suppose we have available some amount of computing time and we want to construct a questionnaire realizing the given decision function. Naturally, we want to construct the optimal questionnaire but, if the specified time is insufficient, we want to obtain at least some suboptimal questionnaire. This may be very efficiently accomplished by taking advantage of the fact that the value of the constant k may be modified during the execution of the algorithm.

At the beginning of the algorithm (at the step (AD 1)) it is possible to set $k \leftarrow 1$ (or some $k' < 1$ which is close to 1) and during the execution of step (AD 5), when setting $\bar{\Delta} \leftarrow \Delta$ and $B \leftarrow b(\Delta)$, set also $k \leftarrow 0$ (or some nonnegative value close to 0). The algorithm modified in this way has a convenient property. Its primary concern is to construct any questionnaire from $\boldsymbol{T}_d$. Only after successful accomplishment of this mission it will start with looking for the optimal questionnaire.

So, when the previous algorithm is being programmed, it is advisable to follow a strategy that controls the algorithm by assigning different values to constants k and N.

CONCLUSIONS

An alternative method for the construction of the optimal questionnaire for a given decision function has been described. In comparison with the Payne-Meisel algorithm both of them have their own advantages:

The advantages of the Payne-Meisel algorithm:

(1) The method yields the optimal questionnaire.

(2) There is a simple way how to estimate the computing time.

The disadvantage of the Payne-Meisel algorithm:

(1) An exponential growth of time and space requirements with the number of questions.

The advantages of the adaptable Branch-and-Bound method:

(1) The method is sometimes usable also when Payne-Meisel algorithm is not usable for too high time requirement.

(2) When only suboptimal questionnaire is satisfactory, the method is usable for very large problems.

The disadvantage of the adaptable Branch-and-Bound method:

(1) The computing time is difficult to estimate.

(Received January 8, 1981.)

REFERENCES

[1] A. H. G. R. Kan: On Mitten's axioms for Branch-and-Bound. Operations Res. *24* (1976), 6, 1179—1178.

[2] L. G. Mitten: Branch-and-Bound methods. Operations Res. *18* (1970), 1, 24—34.

[3] H. J. Payne, W. S. Meisel: An algorithm for constructing optimal binary decision trees. IEEE Trans. Comp. *C-26* (1977), 9, 905—916.

[4] R. Jiroušek: Distinguishability of States by Economical Sequential Questionnaires (in Czech). Thesis, ÚTIA ČSAV, Praha 1978.

[5] R. Jiroušek: Bayes decision functions realized by sequential questionnaires. Trans. 8th Prague Conf. Inform. Theory, etc., Vol. C., Academia, Prague 1979, 187—202.

Radim Jiroušek, CSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 182 08 Praha 8. Czechoslovakia.