British Journal of Mathematical and Statistical Psychology (1988), 41, 1–23 *Printed in Great Britain*

A class of stochastic procedures for the assessment of knowledge

J.-Cl. Falmagne
Department of Psychology, New York University, 6 Washington Place, New York, NY 10003, USA

J.-P. Doignon
Université Libre de Bruxelles, Belgium

The knowledge state of an individual with respect to a particular body of information is conceptualized as the set of all the questions that this individual is capable of solving. The goal of an assessment procedure is to identify, by a sequence of appropriately chosen questions, the individual's state among all possible ones. A deterministic procedure is conceivable, but not realistic, in that it does not account for possible inconsistencies in the observed responses. Such inconsistencies may arise from careless errors or lucky guesses from the subject, but may also be of a more fundamental character. A stochastic framework is developed here, in which an individual 'state' is formalized as a distribution on the set of all possible knowledge states. On each trial, the assessor has a likelihood function on the set of knowledge states which provides the basis for selecting the question to be asked on that trial. The response is assumed to depend on the individual 'state'. The question asked and the response observed are used to update the likelihood function. Several examples of *questioning rules* and of *updating rules* are discussed, which lead to Markov processes. A central problem is to describe conditions ensuring that the latent distribution corresponding to the subject's 'state' can be estimated. We show that, under general conditions, the 'state' of a subject is uncoverable if the latent distribution is concentrated on a particular knowledge state.

A central concept of this paper is a *knowledge structure*, which we define as a finite set $\mathcal{Q}$ of questions or problems, equipped with a distinguished family κ of subsets (cf. Degreef, Doignon, Ducamp & Falmagne, 1986; Doignon & Falmagne, 1985, in press). Any $K \in \kappa$ is a possible *knowledge state* representing the subset of all the questions in $\mathcal{Q}$ which could be answered correctly by some individual. The notation κ_q stands for the set of all states containing a question q. We also write $\kappa_{\bar{q}}$ for $\kappa - \kappa_q$.

As an illustration, consider a set $\mathcal{Q}$ containing five questions, labelled a, b, c, d and e, and suppose that there are 12 knowledge states:

$$\kappa = \{\mathcal{Q}, \varnothing, \{e\}, \{c\}, \{e, c\}, \{e, c, d\}, \{e, b\}, \{e, b, c\},$$
$$\{e, b, d, c\}, \{e, b, a\}, \{e, b, a, c\}, \{e, c, d, a\}\}. \tag{1}$$

In this example, we assume thus that an individual may be able to solve question e without knowing the correct response to any of the other questions. On the other hand, it is not possible to solve question a without also knowing either the responses to e and b, or the responses to e, c and d. We have, for instance,

$$\kappa_{\bar{c}} = \{\varnothing\{e\}, \{e, b\}, \{e, b, a\}\}, \tag{2}$$

$$\kappa_{b} = \{Q, \{e, b\}, \{e, b, c\}, \{e, b, d, c\}, \{e, b, a\}, \{e, b, a, c\}\}, \tag{3}$$

$$\kappa_{a} = \{Q, \{e, b, a\}, \{e, b, a, c\}, \{e, c, d, a\}\}. \tag{4}$$

The problem at hand is to efficiently assess the knowledge state of an individual, from his or her answers to a sequence of questions.

A simple idea is as follows. Suppose that some individual has provided a correct answer to the first question asked, say q_1, but has failed the second question q_2. Then it could be inferred that this individual's state belongs to

$$\kappa_{q_1} \cap \kappa_{\bar{q}_2}.$$

Further questions $q_3, \ldots, q_j$ can then be asked until a single state K_0 has been isolated. That is (assuming that the subject has failed q_j, the last question asked) we would have

$$\kappa_{q_1} \cap \kappa_{\bar{q}_2} \cap \ldots \cap \kappa_{\bar{q}_j} = \{K_0\}.$$

In the case of the above example, the knowledge state $\{e, b, a\}$ could be identified by asking first, say question c and verifying that the student fails to solve it, and then observing that b, and then a, are solved. Indeed, from equations (2), (3) and (4), we have

$$\kappa_{\bar{c}} \cap \kappa_b \cap \kappa_a = \{\{e, b, a\}\}.$$

(This procedure is illustrated in Fig. 1.)

In principle, the questions may be chosen successively so as to minimize, in some appropriate sense, the expected cost (for instance, the average, or the maximal number of questions required to identify a knowledge state). The foundations of this procedure have been discussed by Degreef *et al.* (1986) and Doignon & Falmagne (1985). A practical implementation has been carried out by M. Pavel (personal communication, 1983) concerning the mastery of a standard computer operating system. However appealing the simplicity of this procedure, serious objections can be raised concerning its efficiency and reliability.

In general, a strong case can be made for adopting a probabilistic framework. For example, it is sensible that the choice of a question takes into account the *a priori* distribution of the knowledge states probabilities. For a particular population of subjects, this *a priori* distribution either may be obtained empirically, or may represent the uncertainty of the assessor at a particular time of the examination. The latter viewpoint will be taken, but this choice has no bearing on our developments.

In the deterministic procedure outlined above, a subject's response to a question results in the final elimination of a number of knowledge states. (For instance, an incorrect response to question c in the example eliminates all the states but those in $\kappa_{\bar{c}}$.) This is rather brutal, in that it does not consider the possibility of various chance factors, which will be considered in a moment. In this paper, we shall investigate a

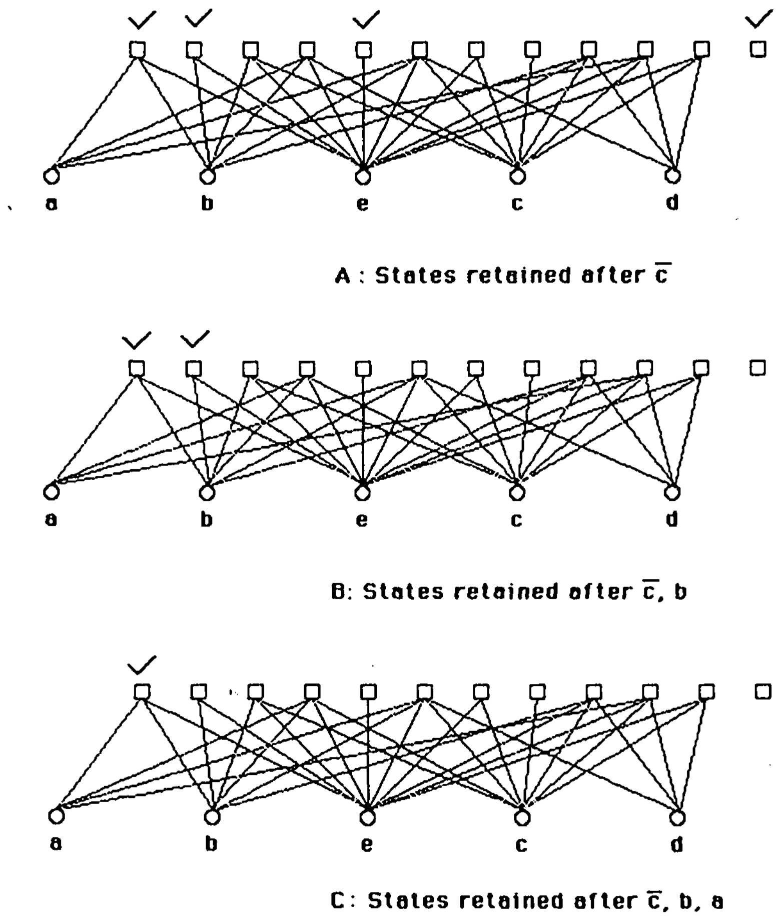

Figure 1. Selection of the state $\{e, b, a\}$ by the questions c, b and a. The knowledge states and the questions are indicated by the squares and the circles, respectively. A linkage between a square and a circle indicates that the corresponding state contains the corresponding question. We denote by $\bar{q}$ the observation of an incorrect response to question q.

large class of procedures in which the succession of questions and responses results in a gradual decrease of the uncertainty of the assessor, as measured (for instance) by a decrease of the entropy of the distribution of the probabilities of the knowledge states in the course of the examination. In other words, a knowledge state will not suddenly cease to be considered. Rather, its likelihood will decrease. Needless to say, there are many ways of implementing this idea, which is evoked by Fig. 2, and several will be discussed.

A stochastic version of the notion of a 'knowledge state' will be used. A basic concept of our earlier papers is that, if the subject is in some state $K_0 \in \kappa$, then any question $q \in K_0$ will be answered correctly, and any question $q \in Q - K_0$ will be answered incorrectly. A generalization of this idea will be considered here: we shall assume that the subject's responses to the questions in the set Q are governed by some

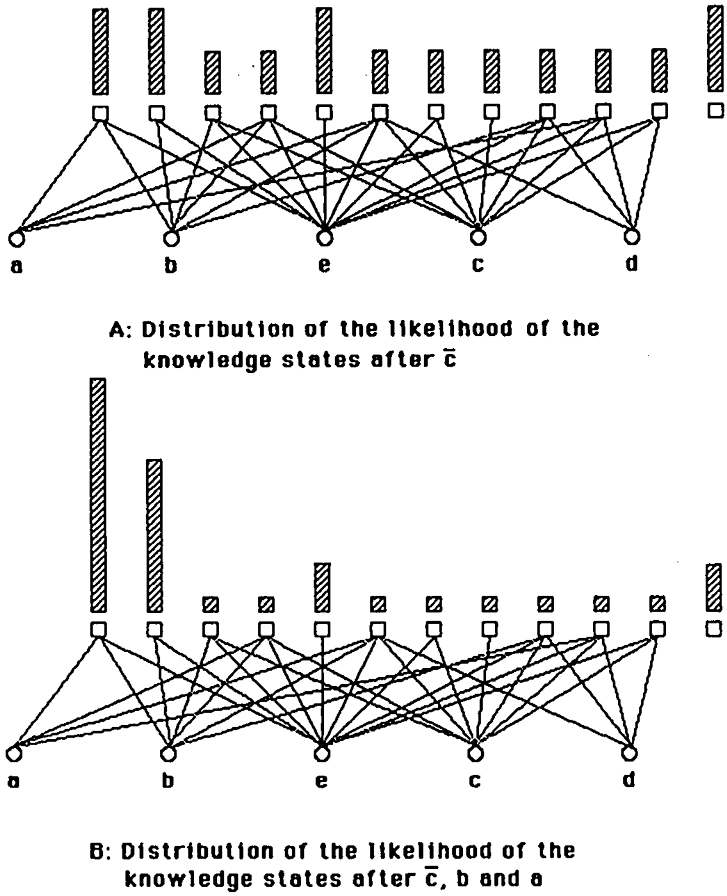

Figure 2. Illustration of a gradual increase of the likelihood of the knowledge state $\{e, b, a\}$ resulting from observing the subject's responses to the questions c, b and a.

unknown probability distribution π on the set κ of knowledge states. Thus, the probability $\rho(q)$ of a correct response to a question $q \in Q$ is given by the equation

$$\rho(q) = \sum_{K \in \kappa_q} \pi(K), \tag{5}$$

a quantity that does not vary in the course of the examination. This probability distribution π, which will sometimes be called *latent*, is what has to be assessed or estimated. Two non-contradictory interpretations of this concept are possible:

(1) The 'knowledge state' of a subject is not stable, but may oscillate from one trial to another. This may explain some 'errors' or apparent inconsistencies in the subject's responses.
(2) The knowledge state of a subject will never be assessed with perfect accuracy. The best that one can hope is to construct a likelihood function on the set of all states. The latent probability distribution π is then estimated by that likelihood.

As the reader will soon realize, the assumption that the response probabilities depends on some latent distribution π on the set of states is tantamount to opening a Pandora box. There are indeed many cases in which π cannot be estimated, no matter which assessment procedure is used. Several examples will be given. We believe however that a non-trivial family of latent distributions can be defined which, on the one hand, offers a realistic modelling of a useful class of empirical situations, and on the other hand, such that any latent distribution π in the family can be assessed empirically by some appropriate method. The problem of specifying such a family appears difficult however, and will not be dealt with here.

Another possible source of noise in the data arises in that a correct response may result from a lucky guess, for example in a multiple choice format. Similarly, an incorrect response may be due to a careless error, say, in a calculation. To simplify our developments, we will assume in this paper that such factors play a minor role and can be neglected (see, however, 3.2 and Falmagne & Doignon (in press).

We remark in passing that even though our efforts are currently focused on the assessment of knowledge, we are in fact implicitly dealing with a considerably more general situation. Examples abound of problems formally similar to that considered here, ranging from medical diagnostic, to discovering the causes of a malfunction of a machine, such as a computer. Our work may thus be regarded as a contribution to the development of a general theory of the assessment of systems. Even though these extensions will not be envisaged in this paper, the reader should keep such potential extensions in mind when pondering our axioms.

Our purpose here is to define a Markovian class of stochastic assessment procedures, and to begin investigating the properties of such a system. It will be shown that, when the latent distribution π has its mass concentrated on one set K_0 in κ—thus, when $\pi(K_0) = 1$—then, under fairly general conditions, this set can be uncovered. (Such a result may at first seem trivial, but a close look at our axioms should quickly dispell this impression. In the stochastic framework developed here, to allow for the possibility of errors of various sorts, a subject's response does not necessarily result in a definitive elimination of a subset of the states.) A couple of special cases will also be discussed in detail, which differ in the method used to select the question asked, and in the rule used to aggregate the information collected in the successive trials of the test.

A different type of knowledge assessment procedure was investigated by Falmagne & Doignon (in press) in the framework of discrete Markov chains. Note that careless errors and lucky guesses were explicitly modelled in that paper.

A knowledge assessment procedure based on the concept of a knowledge state relies critically on the empirical validity of the model. This raises the question of the practical testing of a hypothesized knowledge structure. A stochastic approach suited for this purpose was developed by Falmagne (in press).

1. Notation and basic concepts

We begin by adapting our former terminology to this probabilistic framework. Any probability distribution p on the set κ will be called a *stochastic knowledge state*, which we shall often abbreviate as *'state'*. (The use of the single quotes is meant to eliminate a possible confusion between the knowledge states and the states of the Markov processes which we are about to describe.) The set of all such probability distributions will be denoted by Λ. When the mass of some 'state' $p \in \Lambda$ is concentrated on a particular $K_0 \in \kappa$, we shall say that p is a *pure 'state'*. It will be convenient to identify any $K \in \kappa$ with the pure 'state' $p \in \Lambda$ such that $p_K = 1$.

Any application of the assessment procedure to be described will be a realization of a stochastic process. Each trial starts with an updated estimate of the likelihood of the various $K \in \kappa$. We shall denote by $\mathbf{L}_{n,K}$ the likelihood, *for the observer*, that the subject examined is in the pure 'state' $K \in \kappa$ on trial n. In the framework of the stochastic process, $\mathbf{L}_{n,K}$ is a random variable. More generally, we write

$$\mathbf{L}_{n,\Psi} = \sum_{K \in \Psi} \mathbf{L}_{n,K}$$

for every $\Psi \subseteq \kappa$. The random function $\mathbf{L}_n : K \to \mathbf{L}_{n,K}$ mapping κ into the interval $[0, 1]$ specifies thus a probability distribution on the set κ on trial n of the procedure.

Next, a question must be selected. We write $\mathbf{Q}_n$ for the particular question asked on trial n. Thus $\mathbf{Q}_n$ is a random variable taking its values in the set $\mathcal{Q}$ of all questions. In general, the choice of a question at trial n of the procedure is governed by a function $(q, \mathbf{L}_n) \to \psi(q, \mathbf{L}_n)$ mapping $\mathcal{Q} \times \Lambda$ into the interval $[0, 1]$, and specifying the probability that $\mathbf{Q}_n = q$. The function ψ will be referred to as the *questioning rule*. Two special cases of the function ψ will be considered, which will be discussed later on (see Section 3).

The response observed on trial n will be denoted by $\mathbf{R}_n$. By convention, only two cases will be considered: (i) a correct response, which is coded as $\mathbf{R}_n = 1$; (ii) an error, which is coded as $\mathbf{R}_n = 0$. As indicated, the probability $\rho(q)$ of a correct response to a question q is equal to the probability that the subject is in a pure 'state' containing q. This quantity, which does not vary over trials and only depends upon the latent distribution π, has been specified by equation (5). The function ρ will be called the *response rule*.

A Markovian transition property is at the core of the procedure, which states that, with probability one, the likelihood function $\mathbf{L}_{n+1}$ on trial $n+1$ takes on a value which only depends upon the likelihood function $\mathbf{L}_n$, the question asked on that trial and the response given by the subject. In general, this property is formalized by the equation

$$\mathbf{L}_{n+1} \overset{\text{a.s.}}{=} u(\mathbf{R}_n, \mathbf{Q}_n, \mathbf{L}_n),$$

in which u is a function mapping the Cartesian product $\{0, 1\} \times \mathcal{Q} \times \Lambda$ into Λ. The function u will be referred to as the *updating rule*. Two special cases of updating rules will be considered (see Section 3).

The diagram below summarizes these transitions:

$$(\mathbf{L}_n \to \mathbf{Q}_n \to \mathbf{R}_n) \to \mathbf{L}_{n+1}.$$

The Cartesian product $\Gamma = \{0, 1\} \times \mathcal{Q} \times \Lambda$ is thus the state space of the process, each trial n being characterized by a triple $(\mathbf{R}_n, \mathbf{Q}_n, \mathbf{L}_n) \in \Gamma$. We denote by Ω the sample space, that is, the set of all sequences of points in Γ. The complete random history of the process from trial 1 to trial n will be denoted by

$$\mathbf{W}_n = ((\mathbf{R}_n, \mathbf{Q}_n, \mathbf{L}_n), \ldots, (\mathbf{R}_1, \mathbf{Q}_1, \mathbf{L}_1)).$$

The subset smallest $\Psi \subseteq \kappa$ such that $\pi(\Psi) = 1$ will be called the *support* of π. A special attention will be given to the situation in which $\pi(K_0) = 1$ for some $K_0 \in \kappa$. Thus, the subject is almost surely in the pure 'state' K_0. In such a case, $\{K_0\}$ (or K_0, by abuse of notation)will be called the *unit support* of π. In general, the assessment problem consists in uncovering the support of the latent distribution π, and if possible, to estimate π itself as accurately as possible.

If π has a unit support K_0, uncovering this pure 'state' has a natural formalization in terms of the condition

$$\mathbf{L}_{n,K_0} \xrightarrow{\text{a.s.}} 1.$$

When this condition holds for some particular assessment procedure, we shall sometimes say that K_0 is *uncoverable* (by that procedure).

In the case where π does not have a unit support, the goals of the assessment procedure can be formulated by various conditions. In general, let $\Psi \subseteq \kappa$ be the support of the latent distribution π. A desirable condition would require that

$$\mathbf{L}_{n,K} \xrightarrow{\text{a.s.}} 0 \qquad \text{iff} \quad K \notin \Psi.$$

Thus, only the support of π is uncoverable, but not necessarily π itself. A much stronger condition would be

$$E(\mathbf{L}_{n,K}) \to \pi(K), \qquad \text{for all} \quad K \in \kappa.$$

Obviously, other types of convergence are conceivable that might lead to useful results.

We recall that the indicator function ι_S on a set S is defined by

$$\iota_S(x) = \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{if } x \notin S. \end{cases}$$

2. Axioms

Three axioms specify our general framework.

[U] *Updating rule*. There is a fixed probability distribution l_1 on κ, with $0 < l_1 < 1$, such that $\mathbb{P}\{\mathbf{L}_1 = l_1\} = 1$, and a function u mapping the interior of Γ into the interior of Λ such that, for all positive integers n and all measurable sets $B \subseteq \Lambda$,

$$\mathbb{P}\{\mathbf{L}_{n+1} \in B | \mathbf{W}_n\} = \iota_B[u(\mathbf{R}_n, \mathbf{Q}_n, \mathbf{L}_n)].$$

Writing u_K for the coordinate of u associated with the pure 'state' K, we have thus

$$\mathbf{L}_{n+1,K} \overset{\text{a.s.}}{=} u_K(\mathbf{R}_n, \mathbf{Q}_n, \mathbf{L}_n).$$

Moreover, the function u satisfies the following conditions:

$$u_K(\mathbf{R}_n, \mathbf{Q}_n, \mathbf{L}_n) = \begin{cases} > \mathbf{L}_{n,K} & \text{if } \iota_K(\mathbf{Q}_n) = \mathbf{R}_n, \\ < \mathbf{L}_{n,K} & \text{if } \iota_K(\mathbf{Q}_n) \neq \mathbf{R}_n. \end{cases}$$

[Q] *Questioning rule.* For all $q \in \mathcal{Q}$,

$$\mathbb{P}\{\mathbf{Q}_n = q | \mathbf{L}_n, \mathbf{W}_{n-1}\} = \psi(q, \mathbf{L}_n);$$

where ψ is a function mapping $\mathcal{Q} \times \Lambda$ into $[0, 1]$.

[R] *Response rule.* There is a 'state' π such that

$$\mathbb{P}\{\mathbf{R}_n = 1 | \mathbf{Q}_n, \mathbf{L}_n, \mathbf{W}_n\} = \rho(\mathbf{Q}_n),$$

with the function ρ defined from π as in (5).

To avoid trivialities, we suppose that $\mathcal{K}$ contains at least two pure 'states'. We shall refer to any process $(\mathbf{R}_n, \mathbf{Q}_n, \mathbf{L}_n)$ satisfying these three axioms as a *stochastic assessment procedure (parametrized by* π, l_1, u, ψ, and ρ). As mentioned earlier, we shall occasionally refer to the functions u, ψ and ρ as the *updating rule*, the *questioning rule* and the *response rule* respectively.

Axiom [R] is straightforward. Axioms [Q] and [U] govern, respectively, the choice of a question $\mathbf{Q}_n$ on trial n, and the reallocation of the mass of $\mathbf{L}_n$ on trial $n+1$ depending on the values of $\mathbf{Q}_n$ and $\mathbf{R}_n$. Axiom [U] states, in effect, that if $\mathbf{L}_n = p$, $\mathbf{Q}_n = q$, and $\mathbf{R}_n = r$, then the value of $\mathbf{L}_{n+1}$ is almost surely equal to $u(r, q, p)$. As a general scheme, this seems reasonable, since we want our procedure to specify exactly the likelihood of each of the possible pure 'states' at each trial of the procedure. This axiom ensures moreover no pure 'state' will ever have a likelihood zero, and that the likelihood of any 'state' K will increase whenever either a correct response is observed to a question $q \in K$, or an incorrect response to a question $q \notin K$, and decrease in the two remaining cases. Note that the first two axioms describe the assessment procedure *per se*, while the third axiom specifies a hypothetical mechanism governing a student's responses.

It is easy to show that each of $(\mathbf{R}_n, \mathbf{Q}_n, \mathbf{L}_n)$, $(\mathbf{Q}_n, \mathbf{L}_n)$, and $(\mathbf{L}_n)$, is a Markov process (Theorem 4.2). An important question concerns conditions on the functions u and ψ under which $\mathbf{L}_n$ converges to some random vector $\mathbf{L}$ independent of the initial distribution l_1. These aspects of the process will not be investigated here, however. As indicated, we shall focus on the problem of defining a useful, general class of procedures capable of uncovering a pure 'state'. Such a system will be discussed in Section 5.

The class of processes defined by axioms [U], [Q] and [R] is very large. Useful special cases will be obtained by specializing the questioning rule and the updating rule.

3. Special cases

The value of the initial likelihood l_1 plays a relatively minor role. In the absence of information at the beginning of the procedure, it seems reasonable to set

$$l(K) = \frac{1}{|\kappa|} \tag{6}$$

for all pure 'states' $K \in \kappa$. The functions ψ and u of axioms [Q] and [U] provide a natural classification of the special cases of the system defined in Section 2.

3.1. Two examples of updating rules u

Suppose that question q is presented at trial n and that the subject's response is correct; thus, $\mathbf{Q}_n = q$ and $\mathbf{R}_n = 1$. Axiom [U] requires that the likelihood of all those pure 'states' containing q should almost surely increase, and that the likelihood of all those pure 'states' not containing q should almost certainly decrease. If the response is incorrect, the opposite pair of results should be obtained. Some questions may be judged more informative than others. For instance, it may be argued that, since a correct response to a multiple choice question may be due to a lucky guess, it should not be given as much weight as, say, a correct numerical response resulting from a computation. Moreover, the response itself may be taken into account: a correct numerical response may signify the mastery of the question, but an error does not necessarily imply complete ignorance. This concept will be implemented into two rather different exemplary updating rules, in which the reallocation of the mass of $\mathbf{L}_n$ on trial $n+1$ will be governed by a parameter which may depend on the question asked and on the response given on trial n.

The updating rule u of Axiom [U] will be called *convex with parameters* $0 < \theta_{q,r} < 1$ ($q \in \mathcal{Q}$, $r = 0, 1$) iff the function u of Axiom [U] satisfies the following condition. For all $K \in \kappa$, and with $\mathbf{R}_n = r$, $\mathbf{Q}_n = q$,

$$u_K(r, q, \mathbf{L}_n) = (1 - \theta_{q,r})\mathbf{L}_{n,K} + \theta_{q,r} g_K(r, q, \mathbf{L}_n), \tag{7}$$

where

$$g_K(r, q, \mathbf{L}_n) = \begin{cases} r\dfrac{\mathbf{L}_{n,K}}{\mathbf{L}_{n,\kappa_q}}, & \text{if } K \in \kappa_q, \\[2ex] (1-r)\dfrac{\mathbf{L}_{n,K}}{\mathbf{L}_{n,\kappa_{\bar{q}}}}, & \text{if } K \in \kappa_{\bar{q}}. \end{cases}$$

Thus, the right member of equation (7) specifies a convex combination between the current likelihood, and a conditional one, obtained from discarding all those pure 'states' eliminated by the observed response.

One objection to this particular form of the assessment rule is that it is not 'commutative'. One could maintain that the likelihood on trial $n+1$ should not depend, as it does in (7), on the order of the pairs of questions and responses up to that trial. Consider for example, the two cases

(1) $(\mathbf{Q}_{n-1} = q, \mathbf{R}_{n-1} = r)$, $(\mathbf{Q}_n = q', \mathbf{R}_n = r')$,

(2) $(\mathbf{Q}_{n-1} = q', \mathbf{R}_{n-1} = r')$, $(\mathbf{Q}_n = q, \mathbf{R}_n = r)$.

For a given value of the likelihood $\mathbf{L}_{n-1} = l$, the likelihood on trial $n+1$ should be the same in these two cases since they convey the same information, or so it could be argued. Slightly changing our notation by setting $\xi = (r, q)$, $\xi b = (r', q')$, and $F(l, \xi) = u(r, q, l)$ this translates into the condition

$$F[F(l, \xi), \xi'] = F[F(l, \xi'), \xi]. \tag{8}$$

In the functional equation literature, an operator F satisfying (8) is called *permutable* (see Aczél, 1966, p. 270). In some special cases, permutability greatly reduces the possible forms of such an operator. However, the side conditions used in Aczél (1966) are too strong to be used in our context. We do not know whether these results have been extended in a manner suitable for our purpose (see Luce (1964), and Marley (1967), in this connection). In any event, this concept is of obvious interest, and we call *permutable* an updating rule u with an operator F satisfying (8).

An example of a permutable updating rule is given below. The assessment rule is said to be *multiplicative with parameters* $1 < \eta_{q,r}$, $(q \in Q,\ r = 0, 1)$ iff the function u of Axiom [U] satisfies the condition: with $\mathbf{Q}_n = q$, $\mathbf{R}_n = r$, and

$$\eta^K_{q,r} = \begin{cases} 1 & \text{if } \iota_K(q) \neq r, \\ \eta_{q,r} & \text{if } \iota_K(q) = r, \end{cases}$$

we have

$$u_K(r, q, \mathbf{L}_n) = \frac{\eta^K_{q,r}\mathbf{L}_{n,K}}{\sum_{K' \in \kappa} \eta^{K'}_{q,r}\mathbf{L}_{n,K'}}. \tag{9}$$

The permutability of this assessment rule follows readily from the fact that, for any $n > 1$, and $K \in \kappa$, $\mathbf{L}_{n,K}$ is proportional to

$$\eta^K_{\mathbf{Q}_{n-1},\mathbf{R}_{n-1}}\eta^K_{\mathbf{Q}_{n-2},\mathbf{R}_{n-2}} \cdots \eta^K_{\mathbf{Q}_1,\mathbf{R}_1}\mathbf{L}_{1,K}.$$

Other updating rules applicable in similar, but different, situations are reviewed by Landy & Hummel (1986).

3.2. Remarks

It was pointed out to us by Mathieu Koppen (personal communication), that the multiplicative updating rule could be given an interpretation in a Bayesian framework. Fixing question q, and changing notation slightly, we write

$P_q(K)$ for the *a priori* probability of state K;

$P_q(K|r)$ for the *a posteriori* probability of state K after observing response r;

with a similar interpretation for $P_q(r|K)$. From Bayes Theorem:

$$P_q(K|r) = \frac{P_q(r|K)P_q(K)}{\sum_{K' \in \kappa} P_q(r|K')P_q(K')}. \tag{10}$$

We see that (9) and (10) have a similar form; but $\eta^K_{q,r}$ cannot be regarded as a conditional probability. In particular, we do not have

$$\eta^K_{q,1} + \eta^K_{q,0} = 1.$$

Assume, however, that

$$\eta_{q,1} = \eta_{q,0}.$$

Defining then $\eta_q = \eta_{q,r}$ and

$$\delta\delta^K_{q,r} = \frac{\eta^K_{q,r}}{1+\eta_q},$$

we see that

$$0 \leqslant \delta^K_{q,r} \leqslant 1;$$

for all q, r and K and

$$\delta^K_{q,1} + \delta^K_{q,0} = 1,$$

for all q and K. Since $\delta^K_{q,r}/\eta^K_{q,r} = 1/(1+\eta_q)$, independent of K, we may write

$$u_K(r, q, \mathbf{L}_n) = \frac{\delta^K_{q,r}\mathbf{L}_{n,K}}{\sum\limits_{K' \in \kappa} \delta^{K'}_{q,r}\mathbf{L}_{n,K'}}$$

in which $\delta^K_{q,r}$ may be interpreted as the conditional probability of response r on question q, given state K. In particular, note that

$$\delta^K_{q,r} = \begin{cases} \eta_q/(1+\eta_q) & \text{if } q \in K, r = 1 \\ 1/(1+\eta_q) & \text{if } q \in K, r = 0 \\ 1/(1+\eta_q) & \text{if } q \notin K, r = 1 \\ \eta_q/(1+\eta_q) & \text{if } q \notin K, r = 0 \end{cases}.$$

Thus,

$$\beta_q = \frac{1}{1+\eta_q}$$

may be interpreted as the probability of committing a 'careless error' in responding to question q, while

$$\gamma_q = \frac{1}{1+\eta_q}$$

may be interpreted as the probability of a 'lucky guess' correct response to question q. The multiplicative updating rule can thus be regarded as a generalization of a Bayesian updating rule. Bayesian updating rules in intelligent tutoring systems have been discussed by Kimball (1982).

The two examples of updating rules discussed here were inspired by some operators used in learning theory (Atkinson, Bower & Crothers, 1965; Bush & Mosteller, 1955;

Norman, 1972). In particular, the convex updating rule is closely related to a Bush and Mosteller learning operator (Bush & Mosteller, 1955), while the multiplicative updating rule is germane to an operator used in the so-called beta learning model of Luce (1959).

3.3. *Two examples of questioning rules* ψ

A simple idea for the questioning rule is to choose, on any trial n, a question q so as to partition κ into two subsets κ_q and $\kappa_{\bar{q}}$ with a mass $\mathbf{L}_n$ as equal as feasible; that is, so as to have $\mathbf{L}_{n,\kappa_q}$ close to $1-\mathbf{L}_{n,\kappa_q}$. Note in this connection that any distribution $\mathbf{L}_n$ defines a set $E_n \subseteq \mathcal{Q}$ containing all those questions q minimizing the quantity

$$|2\mathbf{L}_{n,\kappa_q} - 1|.$$

The process requires that $\mathbf{Q}_n \in E_n$ with a probability equal to one. The questions in the set E_n are then chosen with equal probability. We have thus

$$\psi(q, \mathbf{L}_n) = \frac{\iota_{E_n}(q)}{|E_n|}. \tag{11}$$

This particular form of the questioning rule will be called *half-split*.

Another method may be used, which is computationally more demanding and seems at first blush more exact. The uncertainty of the examiner on trial n of the procedure may be evaluated by the entropy of the likelihood on that trial, that is, by the quantity

$$H(\mathbf{L}_n) = -\sum_{K\in\kappa} \mathbf{L}_{n,K} \log_2 \mathbf{L}_{n,K}.$$

It seems sensible to select a question so as to reduce this entropy as much as possible. For $\mathbf{R}_n = r$, $\mathbf{Q}_n = q$ and $\mathbf{L}_n = l$, the expected value of the entropy on trial $n+1$ is given by

$$\rho(q)H[u(1, q, l)] + [1-\rho(q)]H[u(0, q, l)]. \tag{12}$$

But the probability $\rho(q)$ of a correct response is unknown, since it is defined by the latent distribution π. Thus, (12) cannot be computed. However, it is natural to replace, in the evaluation of (12), the probability $\rho(q)$ by the likelihood $\mathbf{L}_{n,\kappa_q}$ of a correct response to question q. The idea is thus to minimize, for a particular $l \in \Lambda$, the quantity

$$\tilde{H}(q, l) = \mathbf{L}_{n,\kappa_q} H[u(1, q, l)] + \mathbf{L}_{n,\kappa_{\bar{q}}} H[u(0, q, l)] \tag{13}$$

over all possible $q \in \mathcal{Q}$. Let $I_n \subseteq \mathcal{Q}$ be the set of all those questions q minimizing $\tilde{H}(q, l)$. This form of the questioning rule ψ, which is specified by the equation

$$\psi(q', \mathbf{L}_n) = \frac{\iota_{I_n}(q')}{|I_n|}, \tag{14}$$

will be referred to as *informative*. Note that, in equation (14), the choice of a question varies with the updating rule. This is not the case for equation (11). Both of these questioning rules are reasonable. Surprisingly enough, for the convex updating rule with a constant parameter θ, the half-split and the informative questioning rule induce the same drawing of questions. Our proof of this fact relies on computations which are collected in the Appendix of this paper (see also 5.7).

4. General results

4.1. Convention

In the rest of this paper, we shall assume that $(\mathbf{Q}_n, \mathbf{R}_n, \mathbf{L}_n)$ is a stochastic assessment procedure parametrized by π, l_1, u, ψ and ρ. Special cases of this process will be specified whenever appropriate.

4.2. Theorem

The process $(\mathbf{L}_n)$ *is Markovian. That is, for any* $n \geqslant 1$ *and any measurable set* $B \subseteq \Lambda$,

$$\mathbb{P}\{\mathbf{L}_{n+1} \in B | \mathbf{L}_n, \ldots, \mathbf{L}_1\} = \mathbb{P}\{\mathbf{L}_{n+1} \in B | \mathbf{L}_n\}. \tag{15}$$

A similar Markovian property holds for the two processes $(\mathbf{R}_n, \mathbf{Q}_n, \mathbf{L}_n)$ *and* $(\mathbf{Q}_n, \mathbf{L}_n)$.

Proof. Using successively axioms [U], [R] and [Q], we have, for any $K \in \kappa$ and $n \geqslant 1$,

$$\begin{aligned}
&\mathbb{P}\{\mathbf{L}_{n+1} \in B | \mathbf{L}_n, \ldots, \mathbf{L}_1\} \\
&= \sum_{(\mathbf{R}_n, \mathbf{Q}_n)} \int\{\mathbf{L}_{n+1} \in B | \mathbf{R}_n, \mathbf{Q}_n, \mathbf{L}_n, \mathbf{L}_{n-1}, \ldots, \mathbf{L}_1\} \mathbb{P}\{\mathbf{R}_n, \mathbf{Q}_n | \mathbf{L}_n, \ldots, \mathbf{L}_1\} \\
&= \sum_{(\mathbf{R}_n, \mathbf{Q}_n)} \iota_B[u(\mathbf{R}_n, \mathbf{Q}_n, \mathbf{L}_n)] \mathbb{P}\{\mathbf{R}_n, \mathbf{Q}_n | \mathbf{L}_n, \ldots, \mathbf{L}_1\} \\
&= \sum_{(\mathbf{R}_n, \mathbf{Q}_n)} \iota_B[u(\mathbf{R}_n, \mathbf{Q}_n, \mathbf{L}_n)] \mathbb{P}\{\mathbf{R}_n | \mathbf{Q}_n \mathbf{L}_n, \ldots, \mathbf{L}_1\} \mathbb{P}\{\mathbf{Q}_n | \mathbf{L}_n, \ldots, \mathbf{L}_1\} \\
&= \sum_{(\mathbf{R}_n, \mathbf{Q}_n)} \iota_B[u(\mathbf{R}_n, \mathbf{Q}_n, \mathbf{L}_n)] \mathbb{P}\{\mathbf{R}_n | \mathbf{Q}_n\} \psi(\mathbf{Q}_n, \mathbf{L}_n),
\end{aligned}$$

which only depends upon B and $\mathbf{L}_n$. We leave the two other cases to the reader. ■

In general, a stochastic assessment procedure is not necessarily capable of uncovering a unit support K_0. The next theorem gathers some simple, but very general results in this connection.

4.3. Theorem

If π *has unit support* K_0, *then, for all positive integer* n, *all* $0 < \varepsilon < 1$, *and all pure 'states'* $K \neq K_0$,

$$\mathbb{P}\{\mathbf{L}_{n+1, K_0} > \mathbf{L}_{n, K_0}\} = 1\,; \tag{16}$$

$$\mathbb{P}\{\mathbf{L}_{n+1, K_0} \geqslant 1-\varepsilon\} > \mathbb{P}\{\mathbf{L}_{n, K_0} \geqslant 1-\varepsilon\}\,; \tag{17}$$

$$\mathbb{P}\{\mathbf{L}_{n+1, K} < \mathbf{L}_{n, K}\} = \mathbb{P}\{\mathbf{Q}_n \in K \Delta K_0\} \tag{18}$$

(where Δ denotes the symmetric difference between sets). Moreover,

$$\lim_{n\to\infty} \mathbb{P}\{\mathbf{L}_{n+1,K_0} \geqslant 1-\varepsilon > \mathbf{L}_{n,K_0}\} = 0. \tag{19}$$

Equation (17) *implies that the sequence* $\mathbb{P}\{\mathbf{L}_{n+1,K_0} \geqslant 1-\varepsilon\}$ *converges.*

Proof. Equation (16) is an immediate consequence of axioms [U] and [R]. Notice that

$$\begin{aligned}\mathbb{P}\{\mathbf{L}_{n+1,K_0} \geqslant 1-\varepsilon\} &= \mathbb{P}\{\mathbf{L}_{n+1,K_0} \geqslant 1-\varepsilon | \mathbf{L}_{n,K_0} \geqslant 1-\varepsilon\}\mathbb{P}\{\mathbf{L}_{n,K_0} \geqslant 1-\varepsilon\}\\ &\quad+\mathbb{P}\{\mathbf{L}_{n+1,K_0} \geqslant 1-\varepsilon | \mathbf{L}_{n,K_0} < 1-\varepsilon\}\mathbb{P}\{\mathbf{L}_{n,K_0} < 1-\varepsilon\} \\ &\geqslant \mathbb{P}\{\mathbf{L}_{n,K_0} \geqslant 1-\varepsilon\},\end{aligned} \tag{20}$$

since, by (16),

$$\mathbb{P}\{\mathbf{L}_{n,K_0} \geqslant 1-\varepsilon | \mathbf{L}_{n,K_0} \geqslant 1-\varepsilon\} = 1. \tag{21}$$

This establishes (17). Writing $\mu(\varepsilon) = \lim_{n\to\infty} \mathbb{P}\{\mathbf{L}_{n,K_0} \geqslant 1-\varepsilon\}$, and taking limits on both sides of (20), we obtain, using (21) again,

$$\mu(\varepsilon) = \mu(\varepsilon) + \lim_{n\to\infty} \mathbb{P}\{\mathbf{L}_{n+1,K_0} \geqslant 1-\varepsilon > \mathbf{L}_{n,K_0}\},$$

yielding (19).

The left member of (18) can be decomposed into

$$\begin{aligned}&\mathbb{P}\{\mathbf{L}_{n+1,K} < \mathbf{L}_{n,K} | \mathbf{Q}_n \in K\,\Delta K_0\}\mathbb{P}\{\mathbf{Q}_n \in K\,\Delta K_0\}\\ &\qquad+\mathbb{P}\{\mathbf{L}_{n+1,K} < \mathbf{L}_{n,K} | \mathbf{Q}_n \in \overline{K_0\,\Delta K}\}\mathbb{P}\{\mathbf{Q}_n \in \overline{K_0\,\Delta K}\}.\end{aligned}$$

By axioms [R] and [U], the last term vanishes, and the factor $\mathbb{P}\{\mathbf{L}_{n+1,K} < \mathbf{L}_{n,K} | \mathbf{Q}_n \in (K\,\Delta K_0)\}$ in the first term is equal to one. Thus, equation (18) follows. ■

5. Uncovering a pure state

We assume here that some $K_0 \in \kappa$ is the unit support of π. We show that, under some fairly general conditions on the updating and the questioning rule, the pure 'state' K_0 can be uncovered. These general conditions include the cases in which the updating rule is convex or multiplicative, and the questioning rule is half-split. We first consider an example using a convex updating rule with a constant parameter $\theta > 0$ and the half-split questioning rule.

5.1. Example.

Let $\mathcal{Q} = \{a, b, c\}$ and

$$\kappa = \{\varnothing\{a\}, \{b, c\}, \{a, c\}, \{a, b, c\}\},$$

with $\pi(\{b, c\}) = 1$ and $\mathbf{L}_{1,K} = 0.2$ for all $K \in \kappa$. Since the questioning rule is half-split, and

$$|2\mathbf{L}_{1,\kappa_x} - 1| = 0.2$$

for all $q \in \mathcal{Q}$, we have $E_1 = \{a, b, c\}$. On trial one, the questions are selected in E_1 with equal probability. Notice that

$$\mathbf{L}_{1,\kappa_a} = \mathbf{L}_{1,\kappa_b} = 0.4$$

while

$$\mathbf{L}_{1,\kappa_c} = 0.6.$$

For the likelihood of the pure 'state' $K_0 = \{b, c\}$ on trial two, we obtain thus, but the convex updating rule

$$\mathbf{L}_{2,K_0} = \begin{cases} (1-\theta)0.2+\theta 0.2/0.4 & \text{with prob. } 1/3 \ (a \text{ is chosen}) \\ (1-\theta)0.2+\theta 0.2/0.4 & \text{with prob. } 1/3 \ (b \text{ is chosen}) \\ (1-\theta)0.2+\theta 0.2/0.6 & \text{with prob. } 1/3 \ (c \text{ is chosen}) \end{cases}$$

This implies that

$$\mathbb{P}\{\mathbf{L}_{2,K_0} > \mathbf{L}_{1,K_0}\} = 1,$$

by 4.3. In fact, the following theorem shows that K_0 is uncoverable.

In the case where π does not have a unit support, one cannot assert in general that this support is uncoverable, as shown by the example below.

5.2. *Example*

Suppose that

$$\kappa = \{\varnothing \{a, b\}, \{a, c\}, (c, d\}, \{b, d\}\},$$

with

$$\pi(\{a, b\}) = \pi(\{c, d\}) = \tfrac{1}{2}.$$

Some reflexion will surely suffice to convince the reader that the 'state' π will not be uncovered by any stochastic assessment procedure in the sense of this paper. Such a procedure cannot discriminate between the 'states' π and π', with $\pi'(\{a, c\}) = \pi'(\{b, d\}) = \frac{1}{2}$. One might dismiss this example as 'pathological', in that it involves a stochastic knowledge state in which the subject 'oscillates' between disjoint pure 'states'. The next example is more plausible, but suffers from similar difficulties, at least if the half-split questioning rule is used.

5.3. *Example*

Let $\kappa = \{\varnothing \{a, b\}, (b, c\}, \{a, c\}\}$. We suppose that $\pi\{\mathrm{a}, b\} = \pi\{a, c\} = \frac{1}{2}$. Asking question a will increase the probability of the set $\Psi = \{\{a, b\},\{a, c\}\}$. We shall have, if this question is asked,

$$\mathbf{L}_{n+1,\Psi} = (1-\theta)\mathbf{L}_{n,\Psi} + \theta.$$

However, if the half-split questioning rule is used, question a will cease to be chosen as soon as $\mathbf{L}_{n,\Psi}$ gets large. When b or c are chosen, the expected likelihood of $\{b, c\}$ increases. One can show that the support cannot be uncovered on the basis of the half-split questioning rule and the convex updating rule.

We now turn to a positive result of convergence, based on some strengthening of the assumptions on a stochastic assessment procedure.

5.4. Definition

An updating rule u is said to be *regular* if there is a non-increasing, real valued function $\lambda > 1$, defined on the open interval]0, 1[, such that:

(i) $$u_K(\mathbf{R}_n, \mathbf{Q}_n, \mathbf{L}_n \geqslant \lambda(\mathbf{L}_{n,K_{\mathbf{Q}_n}})\mathbf{L}_{n,K}, \qquad \text{if } \iota_K(\mathbf{Q}_n) = \mathbf{R}_n = 1;$$

(ii) $$u_K(\mathbf{R}_n, \mathbf{Q}_n, \mathbf{L}_n) \geqslant \lambda(\mathbf{L}_{n,K_{\bar{\mathbf{Q}}_n}})\mathbf{L}_{n,K}, \qquad \text{if } \iota_K(\mathbf{Q}_n) = \mathbf{R}_n = 0.$$

5.5. Proposition

Both the convex and the multiplicative updating rules are regular.

Proof. For the convex updating rule, if $\iota_K(\mathbf{Q}_n) = \mathbf{R}_n = 1$, equations (7) can be rewritten as

$$u_K(\mathbf{R}_n, \mathbf{Q}_n, \mathbf{L}_n) = \left[\theta_{\mathbf{Q}_n,\mathbf{R}_n}\left(\frac{1}{\mathbf{L}_{n,K_{\mathbf{Q}_n}}} - 1\right)\right]\mathbf{L}_{n,K},$$

Setting $\theta = \min\{\theta_{q,r} | q \in \mathcal{Q}, r \in \{0, 1\}\}$, we get

$$u_K(\mathbf{R}_n, \mathbf{Q}_n, \mathbf{L}_n) \geqslant \left[1 + \theta\left(\frac{1}{\mathbf{L}_{n,K_{\mathbf{Q}_n}}} - 1\right)\right]\mathbf{L}_{n,K},$$

In the case $\iota_K(\mathbf{Q}_n) = \mathbf{R}_n = 0$, we obtain similarly

$$\begin{aligned} u_K(\mathbf{R}_n, \mathbf{Q}_n, \mathbf{L}_n) &= \left[1 + \theta_{\mathbf{Q}_n,\mathbf{R}_n}\left(\frac{1}{\mathbf{L}_{n,K_{\mathbf{Q}_n}}} - 1\right)\right]\mathbf{L}_{n,K} \\ &\geqslant \left[1 + \theta\left(\frac{1}{\mathbf{L}_{n,K_{\mathbf{Q}_n}}} - 1\right)\right]\mathbf{L}_{n,K}. \end{aligned}$$

Thus, (i) and (ii) in definition 5.4 are satisfied, with

$$\lambda(t) = 1 + \theta\left(\frac{1}{t} - 1\right) > 1$$

for $t \in]0, 1[$.

For the multiplicative updating rule, we have, in the case $\iota_K(\mathbf{Q}_n) = \mathbf{R}_n = 1$,

$$u_K(\mathbf{R}_n, \mathbf{Q}_n, \mathbf{L}_n) = \frac{\eta_{\mathbf{Q}_n,\mathbf{R}_n}}{\eta_{\mathbf{Q}_n,\mathbf{R}_n}\mathbf{L}_{n,\kappa_{\mathbf{Q}_n}} + \mathbf{L}_{n,\kappa_{\bar{\mathbf{Q}}_n}}}\mathbf{L}_{n,K}$$

$$= \frac{\eta_{\mathbf{Q}_n,\mathbf{R}_n}}{1 + (\eta_{\mathbf{Q}_n,\mathbf{R}_n} - 1)\mathbf{L}_{n,\kappa_{\mathbf{Q}_n}}}\mathbf{L}_{n,K},$$

and similarly, if $\iota_K(\mathbf{Q}_n) = \mathbf{R}_n = 0$,

$$u_K(\mathbf{R}_n, \mathbf{Q}_n, \mathbf{L}_n) = \frac{\eta_{\mathbf{Q}_n,\mathbf{R}_n}}{\mathbf{L}_{n,\kappa_{\mathbf{Q}_n}} + \eta_{\mathbf{Q}_n,\mathbf{R}_n}\mathbf{L}_{n,\kappa_{\bar{\mathbf{Q}}_n}}}\mathbf{L}_{n,K}$$

$$= \frac{\eta_{\mathbf{Q}_n,\mathbf{R}_n}}{1 + (\eta_{\mathbf{Q}_n,\mathbf{R}_n} - 1)\mathbf{L}_{n,\kappa_{\bar{\mathbf{Q}}_n}}}\mathbf{L}_{n,K}.$$

Each of the functions $\lambda_{q,r}$,

$$\lambda_{q,r}(t) = \frac{\eta_{q,r}}{1 + (\eta_{q,r} - 1)t}$$

is decreasing and >1 for $t \in]0, 1[$. Thus $\lambda = \min\{\lambda_{q,r} | q \in \mathcal{Q}, r \in \{0, 1\}\}$ satisfies (i) and (ii) in Definition 5.4. ■

As far as the questioning rule is concerned, it is intuitively clear that, from the standpoint of the observer, it would not be efficient to choose a question q with a likelihood $\mathbf{L}_{n,\kappa_q}$ of a correct response close to zero or one. Actually, it makes good sense to choose, as in the half-split questioning rule, a question q with $\mathbf{L}_{n,\kappa_q}$ as far as possible from zero or one. A much weaker form of this idea is captured by the following definition.

5.6. *Definition*

The questioning rule ψ is said to be *inner* if it satisfies the following conditions. There is a real valued function v defined on the open interval $]0, 1[$, with two numbers $\gamma, \delta > 0$, such that

(i) v is strictly decreasing on $]0, \gamma[$;
(ii) v is strictly increasing on $]1-\delta, 1[$;
(iii) $v(t) > v(t')$ whenever $\gamma \leqslant t' \leqslant 1-\delta$ and $t < \gamma$ or $t > 1-\gamma$.

Defining

$$S_n = \{q \in \mathcal{Q} | v(\mathbf{L}_{n,\kappa_q}) \leqslant v(\mathbf{L}_{n,\kappa_{q'}}) \quad \text{for all } q' \in \mathcal{Q}\}$$

one has

$$\psi(q, \mathbf{L}_n) = \frac{\iota_{S_n}(q)}{|S_n|}.$$

5.7. Proposition

The half-split questioning rule is inner.

Proof. This is obvious from the definitions using $\nu(t) = |\frac{1}{2}t - 1|$. ■

5.8. Proposition

When the updating rule is convex with a constant parameter θ, the informative questioning rule is inner.

The proof will be given in the Appendix.
We now state our main result.

5.9. Theorem

Let $(\mathbf{R}_n, \mathbf{Q}_n, \mathbf{L}_n)$ be a stochastic assessment procedure parametrized by π, l_1, u, ψ and ρ, with u regular and ψ inner. Suppose that $\pi(K_0) = 1$ for some $K_0 \in \kappa$, that is, π has unit support K_0. Then K_0 is uncoverable:

$$\mathbf{L}_{n,K_0} \overset{\text{a.s.}}{\rightarrow} 1\,.$$

Proof. Define $\tilde{\Omega}$ to be the set of all realizations ω for which, on every trial n,

(i) $\mathbf{Q}_n \in S_n$, with S_n defined as in 5.6;
(ii) $\mathbf{R}_n = \iota_{K_0}(\mathbf{Q}_n)$.

Notice that $\tilde{\Omega}$ is a measurable subset of the sample space Ω, and that $\mathbb{P}\{\tilde{\Omega}\} = 1$. Thus, we only have to establish

$$\lim_{n \to \infty} \mathbf{L}_{n,K_0}(\omega) = 1$$

for any point $\omega \in \tilde{\Omega}$. Take $\omega \in \tilde{\Omega}$ arbitrarily. It follows readily from the conditions that $\mathbf{L}_{n,K_0}(\omega)$ is non-decreasing, and thus converges. Henceforth, it suffices to show that $\mathbf{L}_{n_i,K_0}(\omega) \to 1$ for at least one subsequence $s = (n_i)$ of the positive integers. Since κ is finite, and $\mathbf{L}_{n,K} \in]0, 1[$ we can take $s = (n_i)$ such that $\mathbf{L}_{n_i,K}(\omega)$ converges for all $K \in \kappa$. In the rest of this proof, we consider a fixed subsequence s satisfying those conditions.

We define a function $f_{\omega,s}: \mathcal{Q} \to [0, 1]$ by

$$f_{\omega,s}(q) = \begin{cases} \lim\limits_{i \to \infty} \mathbf{L}_{n_i,K_q}(\omega) & \text{if } q \in K_0\,; \\ \lim\limits_{i \to \infty} \mathbf{L}_{n_i,K_{\bar{q}}}(\omega) & \text{if } q \notin K_0\,. \end{cases}$$

We also set

$$\tilde{\mathcal{Q}}_{\omega,s} = \{q \in \mathcal{Q} \mid f_{\omega,s}(q) < 1\}\,.$$

If $\tilde{\mathcal{Q}}_{\omega,s}$ is empty, the fact that $\mathbf{L}_{n_i,K_0}(\omega) \to 1$ follows readily from Lemma 3 (see below).

The core of the proof consists in establishing that $\tilde{\mathcal{Q}}_{\omega,s} = \varnothing$, which is achieved by Lemma 2.

Lemma 1

If $f_{\omega,s}(q) < 1$, *then* $\{i | \mathbf{Q}_{n_i}(\omega) = q\}$ *is a finite set.*

Proof. Assume $q \in K_0$. Since $f_{\omega,s}(q) < 1$, there is $\varepsilon > 0$ such that $f_{\omega,s}(q) + \varepsilon < 1$. If $\mathbf{Q}_{n_i}(\omega) = q$ and i is large enough to ensure that $\mathbf{L}_{n_i,K_q} \leqslant f_{\omega,s}(q) + \varepsilon$, we derive

$$\begin{aligned}\mathbf{L}_{n_{i+1},K_0}(\omega) &= \mathbf{L}_{n_i+1,K_0}(\omega)\\ &\geqslant \lambda[\mathbf{L}_{n_i,K_q}(\omega)]\mathbf{L}_{n_i,K_0}(\omega)\\ &\geqslant \lambda[f_{\omega,s}(q) + \varepsilon]\mathbf{L}_{n_i,K_0}(\omega).\end{aligned}$$

Since $\lambda[f_{\omega,s}(q) + \varepsilon] > 1$ does not depend on i, and $\mathbf{L}_{n_i,K_0}(\omega)$ remains $\leqslant 1$, we may have $\mathbf{Q}_{n_i}(\omega) = q$ for at most a finite number of values of i. The proof is similar if $q \notin K_0$. ■

Lemma 2

$\tilde{\mathcal{Q}}_{\omega,s} = \{q \in \mathcal{Q} | f_{\omega,s}(q) < 1\} = \varnothing$

Proof. We proceed by contradiction. If $f_{\omega,s}(q) < 1$, we can assert the existence of some positive integer j and some $\varepsilon > 0$ such that whenever $i > j$,

$$0 < l_{1,K_0} \leqslant \mathbf{L}_{n_i,K_q}(\omega) < f_{\omega,s}(q) + \varepsilon < 1 \qquad \text{if} \quad q \in K_0,$$

and

$$0 < l_{1,K_0} \leqslant \mathbf{L}_{n_i,K_{\bar{q}}}(\omega) < f_{\omega,s}(q) + \varepsilon < 1 \qquad \text{if} \quad q \notin K_0.$$

This means that for $i > j$, both $\mathbf{L}_{n_i,K_q}(\omega)$ and $\mathbf{L}_{n_i,K_{\bar{q}}}(\omega)$ remain in some interval $]\gamma'_q, 1 - \delta'_q[$ with $0 < \gamma'_q, \delta'_q$. The above argument applies to all $q \in \tilde{\mathcal{Q}}_{\omega,s}$. In view of the finiteness of $\tilde{\mathcal{Q}}_{\omega,s}$, the index q may be dropped in γ'_q, δ'_q. Moreover, referring to 5.6, we can assert the existence of $\bar{\gamma}$ and $\bar{\delta}$ such that $0 < \bar{\gamma} < \gamma$, $0 < \bar{\delta} < \delta$ and $\mathbf{L}_{n_i,K_q}(\omega) \in]\bar{\gamma}, 1 - \bar{\delta}[$ for $i > j$ and $q \in \tilde{\mathcal{Q}}_{\omega,s}$.

Since $\tilde{\mathcal{Q}}_{\omega,s}$ is finite, Lemma 1 may be invoked to infer the existence of k such that $\mathbf{Q}_{n_i}(\omega) \notin \tilde{\mathcal{Q}}_{\omega,s}$ whenever $i > k$. Note that, by definition of $\tilde{\mathcal{Q}}_{\omega,s}$, $f_{\omega,s}(q') = 1$ for all $q' \notin \tilde{\mathcal{Q}}_{\omega,s}$. Since $\mathbf{L}_{n_i,K_{q'}}(\omega)$ converges, there is $i^* > j, k$ such that neither $\mathbf{L}_{n_i,K_{q'}}(\omega)$ no $\mathbf{L}_{n_i,K_{\bar{q}'}}(\omega)$ are points of $]\bar{\gamma}, 1 - \bar{\delta}[$ for all $i > i^*$ and $q' \notin \tilde{\mathcal{Q}}_{\omega,s}$. By definition of $\tilde{\Omega}$, we must have $\mathbf{Q}_{n_i}(\omega) \in \tilde{\mathcal{Q}}_{\omega,s}$, contradicting $i > k^*$. ■

Define, for any $K \in \kappa$

$$\mathcal{A}_{\omega,s}(K) = \{q \in K | f_{\omega,s} = 1\}.$$

Lemma 3

Suppose that, for some $K \in \kappa$, $\mathcal{A}_{\omega,s}(K) \neq \mathcal{A}_{\omega,s}(K_0)$. *Then*

$$\lim_{i\to\infty} \mathbf{L}_{n_i,K}(\omega) = 0.$$

Proof. Assume that there is some $q \in K_0 - K$ such that $f_{\omega,s}(q) = 1$; that is $\lim_{i\to\infty} \mathbf{L}_{n_i,\kappa_q}(\omega) = 1$. This implies $\lim_{i\to\infty} \mathbf{L}_{n_i,\kappa_{\bar{q}}}(\omega) = 0$ and the thesis, since $K \in \kappa_{\bar{q}}$. The other case, $q \in K - K_0$, is similar. ■

By Lemma 2, $K \neq K_0$ implies $A_{\omega,s}(K) \neq A_{\omega,s}(K_0)$. Thus, for all $K \notin K_0$, $\lim_{i\to\infty} L_{n_i,K}(\omega) = 0$, yielding $L_{n_i,K_0}(\omega) \to 1$.

This concludes the proof of the proposition. ■

5.10. Remark

A careful study of the proof shows that the assumption of an inner questioning rule can be replaced by the following one.

For any $\gamma, \delta \varepsilon]0, 1[$, denote by $E_{n,q}(\gamma, \delta)$ the event that $\gamma < \mathbf{L}_{n,\kappa_q} < 1-\delta$, and let

$$E_n(\gamma, \delta) = \bigcup_{q \in Q} E_{n,q}(\gamma, \delta).$$

The condition states that there exist $\sigma > 0$, such that, for all $\gamma, \delta \in]0, \sigma[$

$$\mathbb{P}\{\mathbf{Q}_n = q' | \overline{E_{n,q'}(\gamma, \delta)} \cap E_n(\gamma, \delta)\} = 0.$$

In words, and somewhat loosely: no question q will be chosen with $\mathbf{L}_{n,\kappa_q}$ in a neighbourhood of one or zero when this can be avoided.

5.11. Corollary

A unit support of the latent distribution is uncoverable by a stochastic assessment procedure with an updating rule which is either convex or multiplicative, and a questioning rule which is half-split.

Proof. This results immediately from 5.5, 5.7, 5.9 and the definitions. ■

5.12. Corollary

A unit support of the latent distribution is uncoverable by a stochastic assessment procedure with an updating rule which is convex with a constant parameter θ, and a questioning rule which is informative.

Proof. This results immediately from 5.5, 5.8, 5.9 and the definitions. ■

At this stage, our results are thus incomplete. In particular, we do not have a proof of the almost sure convergence $\mathbf{L}_{n,K_0} \to 1$ for the fourth special case, namely, a multiplicative updating rule paired with an informative questioning rule.

Acknowledgements

This work was supported by grant No. IST-8418860 to New York University. We are grateful to Mathieu Koppen and Michael Landy for their comments on a previous draft, and especially to Landy for uncovering an error in one of our statements.

References

Aczél, J. (1966). *Lectures on Functional Equations and their Applications*. New York: Academic Press.

Atkinson, R. C., Bower, G. H. & Crothers, E. J. (1965). *An Introduction to Mathematical Learning Theory*. New York: Wiley.

Bush, R. R. & Mosteller, F. (1955). *Stochastic Models for Learning*. New York: Wiley.

Degreef, E., Doignon, J.-P., Ducamp, A. & Falmagne, J.-Cl. (1986). Languages for the assessment of knowledge. *Journal of Mathematical Psychology*, **30**, 243–256.

Doignon, J.-P. & Falmagne, J.-Cl. (1965). Spaces for the assessment of knowledge, *International Journal of Man–Machine Studies*, **23**, 175–196.

Doignon, J.-P. & Falmagne, J.-Cl. (in press). Parametrization of knowledge structures. *Discrete Applied Mathematics*.

Falmagne, J.-Cl. & Doignon, J.-P. (in press). A Markovian procedure for assessing the state of a system. *Journal of Mathematical Psychology*.

Falmagne, J.-Cl. A latent trait theory via a stochastic learning theory for a knowledge space. *Mathematical Studies in Perception and Cognition*, 87-2, Department of Psychology, New York University. Submitted to *Psychometrika*.

Kimball, R. (1982). A self-improving tutor for symbolic integration. In D. Sleeman & J. S. Brown (Eds), *Intelligent Tutoring Systems*. Computer and People series. London/New York: Academic Press.

Landy, M. S. & Hummel, R. A. (1986). A brief survey of knowledge aggregation methods. *Proceedings of the International Conference on Pattern Recognition, Paris, 27–30 October*, pp. 248–252.

Luce, R. D. (1959). *Individual Choice Behavior*. New York: Wiley.

Luce, R. D. (1964). Some one-parameter families of communicative learning operators. In R. C. Atkinson (Ed.), *Studies in Mathematical Psychology*. Stanford: Stanford University Press.

Marley, A. A. J. (1967). Abstract one-parameter families of commutative learning operations. *Journal of Mathematical Psychology*, **4**, 414–429.

Norman, M. F. (1972). *Markov Processes and Learning Models*. New York: Academic Press.

Received 19 March 1987; revised version received 1 October 1987

Appendix

Following up on 3.2 and equation (13), we have to minimize the quantity

$$\tilde{H}(q, l) = l_{\kappa_q} H[u(1, q, l)] + l_{\kappa_{\bar{q}}} H[u(0, q, l)], \tag{22}$$

where

$$l_{\kappa_q} = \sum_{K \in \kappa_q} l_K, \qquad l_{\kappa_{\bar{q}}} = \sum_{K \in \kappa_{\bar{q}}} l_K.$$

This quantity depends upon the updating rule u. We assume here that the updating rule is convex with a constant parameter θ. We obtain, from (7)

$$\begin{aligned} H[u(1, q, l)] = &- \sum_{K \in \kappa_q} l_K \left[(1-\theta) + \frac{\theta}{l_{\kappa_q}}\right]\left[\log l_K + \log\left(1 - \theta + \frac{\theta}{l_{\kappa_q}}\right)\right] \\ &- \sum_{J \in \kappa_{\bar{q}}} l_J (1-\theta)[\log l_J + \log(1-\theta)] \end{aligned} \tag{23}$$

and

$$H[u(0,q,l)] = -\sum_{J\in\kappa_{\bar q}} l_J\left[(1-\theta)+\frac{\theta}{l_{\kappa_{\bar q}}}\right]\left[\log l_J+\log\left(1-\theta+\frac{\theta}{l_{\kappa_{\bar q}}}\right)\right]$$
$$-\sum_{K\in\kappa_{\bar q}} l_K(1-\theta)[\log l_K+\log(1-\theta)]. \tag{24}$$

Using (22), (23) and (24) and grouping appropriately, leads to

$$\tilde H(q,l) = H(l)-2l_{\kappa_q}l_{\kappa_{\bar q}}(1-\theta)\log(1-\theta)$$
$$-l_{\kappa_q}[(1-\theta)l_{\kappa_q}+\theta]\log\left(1-\theta+\frac{\theta}{l_{\kappa_q}}\right)$$
$$-l_{\kappa_{\bar q}}[(1-\theta)l_{\kappa_{\bar q}}+\theta]\log\left(1-\theta+\frac{\theta}{l_{\kappa_{\bar q}}}\right).$$

That is, with $l_{\kappa_q}=t$, and for $t\in\,]0,1[$

$$g(t) = -2t(1-t)(1-\theta)\log(1-\theta)-t[(1-\theta)t+\theta]\log\left(1-\theta+\frac{\theta}{t}\right)$$
$$-(1-t)[(1-\theta)(1-t)+\theta]\log\left(1-\theta+\frac{\theta}{1-t}\right), \tag{25}$$

we have

$$\tilde H(q,l) = H(l)+g(t). \tag{26}$$

To prove both the assertion at the end of 3.2, and Proposition 5.7, it suffices to establish that the function g is convex on $]0,1[$ and has an extremum at $\frac{1}{2}$. (Thus, g will serve as the function ν in 5.6.) Since g is symmetric around $\frac{1}{2}$, that is, $g(t)=g(1-t)$ for $t\in\,]0,1[$, we only have to show that the second derivative $g''(t)$ is positive on $]0,\frac{1}{2}[$. We shall derive this fact from $g'''(t)<0$ for $t\in\,]0,\frac{1}{2}[$, together with $g''(\frac{1}{2})>0$. To compute the derivatives, we simplify the expression of g. First, notice that it can be written as

$$g(t) = -\log(1-\theta)+\frac{1}{1-\theta}\Bigg\{-[(1-\theta)t][(1-\theta)t+\theta]\log\frac{(1-\theta)t+\theta}{(1-\theta)t}$$
$$-[(1-\theta)(1-t)][(1-\theta)(1-t)+\theta]\log\frac{(1-\theta)(1-t)+\theta}{(1-\theta)(1-t)}\Bigg\} \tag{27}$$

Setting $a(t)=(1-\theta)t$ and $b(t)=(1-\theta)t+\theta$, and

$$f(t) = -a(t)b(t)\log\frac{b(t)}{a(t)},$$

(27) becomes

$$g(t) = -\log(1-\theta)+\frac{1}{1-\theta}[f(t)+f(1-t)]. \tag{28}$$

Using $a'(t)=b'(t)=1-\theta$, we obtain, for the derivatives of f,

$$f''(t) = (1-\theta)^2\left[\theta\frac{a(t)+b(t)}{a(t)b(t)}-2\log\frac{b(t)}{a(t)}\right],$$
$$f'''(t) = -\frac{(1-\theta)^3\theta^3}{a(t)^2b(t)^2},$$

which yields

$$g'''(t) = -\frac{(1-\theta)^2\theta^3}{[(1-\theta)t]^2[(1-\theta)t+\theta]^2}+\frac{(1-\theta)^2\theta^3}{[(1-\theta)(1-t)]^2[(1-\theta)(1-t)+\theta]^2}.$$

For $t \in]0, 1[$, we have $t < 1 - t$, and thus $g'''(t) < 0$. On the other hand,

$$g''(\tfrac{1}{2}) = 4(1-\theta)^2 \left[\frac{2\theta}{1-\theta^2} - \log \frac{1+\theta}{1-\theta} \right].$$

Setting

$$h(\theta) = \frac{2\theta}{1-\theta^2} - \log \frac{1+\theta}{1-\theta},$$

we get

$$\lim_{\substack{\theta \to 0 \\ >}} h(\theta) = 0,$$

and

$$\frac{\mathrm{d}h}{\mathrm{d}\theta}(\theta) = \left(\frac{2\theta}{1-\theta^2} \right)^2 > 0.$$

Thus,

$$g''(\tfrac{1}{2}) = 4(1-\theta)^2 h(\theta) > 0.$$