A Model for Learning and Imprinting with Finite and Infinite Memory Range

E. Pfaffelhuber
Institut für Informationsverarbeitung, Universität Tübingen, FRG

Received: January 26, 1973

Abstract

A generalization of the Bush-Mosteller learning model, in the sense of a response strength rationale, can simulate both a finite and an infinite memory range and describe learning as well as imprinting processes, the latter being characterized by the property that earlier observations enter with more weight into the system's response tendencies than later ones. The resulting difference equation for the response probabilities is no longer time invariant. Optimality properties of the model are discussed, in particular with respect to probability learning. Carnap's inductive probabilities are shown to provide a least mean square estimate for a stationary stochastic environment.

I. Introduction

It is well known that the linear learning models introduced by Bush and Mosteller (Bush and Mosteller, 1951, 1953, 1955; see also Pfaffelhuber, 1973; Pfaffelhuber and Damle, 1972) describe a vast number of data from learning experiments sufficiently well. Nevertheless, these models possess some general drawbacks with respect to certain applications, namely the fact, that, roughly speaking, the influence of a learning system's previous observation or experience upon its response tendencies decreases exponentially with increasing time distance between this experience and the response. As a consequence, the BM (Bush-Mosteller) models are, by construction, finite memory models, so that the interesting and not unrealistic case of a long term memory with infinite memory span (Underwood, 1964; Peterson, 1966) where earlier experiences are stored with the same weight as later ones and which clearly has certain optimality properties, cannot be dealt with; for the same reason, these models are unable to simulate processes where an organism's earlier observations enter with more weight into its response tendencies than later ones.

This is, e.g., a familiar phenomenon in the "imprinting" process (Hess, 1958; Eibl-Eibesfeldt, 1969, pp. 235ff.) of animals as well as, in some sense, in animal instinct behavior (and perhaps even in biological evolution in general) if, instead of learning, we talk about the acquisition of behavioral abilities, and, instead of consecutive learning trials, we consider consecutive generations. These drawbacks of the BM models are not overcome by the stimulus sampling models (Atkinson, Bower and Crothers, 1965, pp. 343ff.; Estes, 1959; Feichtinger, 1971) as is already indicated by the fact that in both types of models the mean learning curves coincide (Hilgard and Bower, 1966, pp. 366).

We show in Section II, by elaborating more explicitly the reasons for these drawbacks mentioned, that there is a natural extension of the BM scheme. This EBM (extended Bush-Mosteller) model possesses a response strength interpretation in the sense of Hull (Hull, 1943) and Luce (Luce, 1959, pp. 91ff.), and has the following properties, not shared by the conventional BM model:

1. The relative weight with which the learning system's consecutive experiences enter into its response tendencies may be less than, equal to, or greater than 1, so that both proper learning as well as the characteristic features of imprinting processes may be described. A truly infinite memory range can thus be simulated.

2. The BM scheme can be considered as a special case of the extended model since it can be obtained from the latter by imposing a certain relation between the initial total response strength and the (finite) memory range, so that, with respect to data fitting, the extended model is at least as good as the conventional one.

3. While the first order difference equation governing the time behavior of the response strengths is time invariant, viz. all coefficients are time independent (which seems to be a reasonable requirement for a biologically relevant learning model), the resulting difference equation for the response probabilities, though still path independent (Bush and Mosteller, 1955, pp. 17), contains a time varying learning param-
eter sequence of a characteristic form whose two parameter describe essentially the system’s memory range and the total initial response strength. As a consequence, difference equations for response probabilities which are themselves not time invariant, may nevertheless possess biological significance.

4. For a stationary environment and a special choice of parameters, corresponding to vanishing initial response strengths, the system’s response probabilities are, for all learning trials, unbiased estimates for certain environmental parameters; in the case of probability learning (Atkinson et al., 1965, pp. 187ff.; Feichtinger, 1971) where the system’s response probabilities can be identified with its “subjective” probabilities (Pfaffelhuber, 1972; Belis, 1968; Weiss, 1967) of the occurrence of certain events, these environmental parameters coincide with the “objective” i.e. experimenter fixed probabilities (Bush and Mosteller, 1955, pp. 68) of these events.

5. For the limiting case of an infinite memory span, the resulting response probabilities represent, for a stationary environment, optimal estimates, in a least mean square sense, of the environmental parameters mentioned above. In the case of probability learning, the system’s subjective probabilities, possessing then, also these optimality properties, are identical with Carnap’s inductive probabilities (Carnap, 1950; Carnap and Stegmüller, 1958).

These five statements are proved in Section III, where also some results on the moments of the response probabilities and their asymptotic behavior are given.

II. The BM Model and its Extension

The BM learning model is based upon a linear difference equation for the vector \( q \) whose components \( q^{(j)} \) represent the probabilities of the learning system’s possible response alternatives \( r^{(j)} = 1, \ldots, R \), which are assumed to be mutually exclusive and exhaustive, so that

\[
\sum_{j=1}^{R} q^{(j)} = 1 .
\]

If the “combining classes condition” (Bush and Mosteller, 1955, pp. 37ff.) is satisfied then this difference equation reads, for the “equal-alpha case” (Bush and Mosteller, 1955, pp. 106ff.)

\[
q_n - q_{n-1} = (1 - \alpha) (\lambda(e_n) - q_{n-1}) , \quad n = 1, 2, \ldots
\]

where

\[
0 \leq \alpha < 1 .
\]

In (2), \( n \) labels consecutive learning trials, \( q_n \) is the response probability vector after the \( n \)th trial, \( e_n \in \hat{E} = \{ e^{(j)} ; j = 1, \ldots, E \} \) is the event (reinforcement, penalty etc.) occurring during the \( n \)th learning trial, thus altering the learning system’s response tendencies, and, for \( e \in \hat{E} \), \( \lambda(e) \) is a probability vector representing the limiting value of \( q_n \) for \( n \to \infty \) if event \( e \) is presented consistently during almost all learning trials. Thus, denoting by \( p_n(e) \) the probability of the occurrence of event \( e \in \hat{E} \) during the \( n \)th learning trial,

\[
p_n(e) = Pr(e_n = e) , \quad n = 1, 2, \ldots
\]

we have

\[
\lim_{n \to \infty} q_n = \lambda(e) \quad \text{if} \quad p_n(e) = 1 \quad \text{for almost all} \quad n .
\]

The quantity \( 1 - \alpha \) in (2) is the so-called learning parameter, characterizing the system’s tendency to adjust its response probabilities according to its new observations of the events \( e \in \hat{E} \). A large resp. small \( \alpha \)-value indicates that this tendency is small resp. large. For \( \alpha = 1 \), \( q_n \) would be equal to the initial \( q_0 \) for all \( n \), so that no learning would occur at all, while for \( \alpha = 0 \) the system’s response probability vector \( q_n \) coincides with the vector \( \lambda(e_n) \) describing the system’s very last observation, so that the system does not take into account any previous experience.

It is easy to show that the solution of (2) is given by

\[
q_n = \alpha^n q_0 + (1 - \alpha) \sum_{m=0}^{n-1} \alpha^m \lambda(e_{n-m})
\]

(where \( \alpha^0 = 1 \) also for \( \alpha = 0 \)). It follows that the vector \( \lambda(e_{n-m}) \) stemming from the system’s observation \( m \) time steps before the \( n \)th learning trial, enters into its response probability \( q_n \) linearly, with a weight factor given by

\[
w_m = (1 - \alpha) \alpha^m , \quad 0 \leq m < n .
\]

It is natural, therefore, to define the learning system’s memory range \( M \) as the first moment of these (normalized) weight factors in the limit of a large number of learning trials, viz.

\[
M = \lim_{n \to \infty} \frac{\sum_{m=0}^{n-1} m w_m}{\sum_{m=0}^{n-1} w_m} .
\]

We obtain

\[
M = \frac{\alpha}{1 - \alpha}.
\]