INTRODUCTION

Stimulus sampling models have enjoyed increasingly wide application in learning theory. Although the representation of a stimulus situation as a set of elements is a familiar aspect, at a verbal level, of the classical identical-elements conception of transfer (Thorndike, 1914), this notion was apparently first formalized in a theory developed to account for the effects of repetition in the acquisition process (Estes, 1950). Natural extensions of this theory have led to interpretations of discrimination, generalization, temporal processes, and even motivational phenomena. New developments reported during the past year include interpretations of stimulus encoding in memory (Bower, 1972), the partial reinforcement effect (Koteskey, 1972), and aspects of communication systems (Estes, 1971). At a more abstract level, Suppes (1969) has demonstrated that finite automata can be given representations within stimulus sampling theory. Thus we may anticipate efforts to deal with the still more challenging problem of extending the theory to interpret the acquisition of language.

The development of stimulus sampling theory has proceeded on a piece-
meal basis, instigated and directed largely by empirical considerations. Consequently it is not surprising that the literature includes many variations even in the treatment of simple learning. Always it is assumed that the subject in a learning experiment samples a population of stimuli, or 'cues,' on each trial, that his probability of making a given response depends on the proportion of sampled stimuli that are 'conditioned,' or 'connected,' to the response, and that the connections between stimuli and response change as a result of reinforcement and nonreinforcement during learning. However, some investigators have assumed independent sampling of stimulus elements on the part of the subject; some have assumed that the sample size is fixed; some have used a 'fluid model' in which the distinction does not appear. For the most part, earlier work it was always assumed that both population and sample are large; consequently differences among the varying treatments were glossed over by 'large numbers' approximations. More recently, the advantage of dealing with sampling models as finite Markov processes has been recognized and the consequences of different assumptions about the sampling process have begun to come to view (Kemeny & Snell, 1957; Suppes & Atkinson, 1960).

An article (Estes, 1959b) organizing many of the results obtained in the mathematical aspects of stimulus sampling theory to date has clarified the differences among the particular variants and at the same time has brought out the need for a systematic treatment of the subject.

In this report we give some of the principal results of our analysis of the foundations of stimulus sampling models. The investigation has been conducted in the same spirit as our treatment of linear models (Estes & Suppes, 1959a). It has proved to be a more extensive enterprise, however, and more of the results are new from a psychological viewpoint. The novelty arises primarily from the generality of our methods, not from any intentional introduction of novel concepts or hypotheses. Where we define new concepts or state new assumptions, it is only for the purpose of making explicit distinctions that have not been made in previous work. One of our objectives has been to discover just what assumptions are necessary to justify the various results, learning functions, asymptotic predictions, etc., that are generally accepted as belonging to stimulus sampling theory.

The axioms for stimulus sampling presented in this chapter make explicit the independence of path properties and the trial invariance that have been tacitly assumed in nearly all contemporary theories. Concerning the nature of the sampling process, these axioms are less restrictive than those of any one extant theory; they suffice to generate as special cases both the independent sampling model and the fixed sample-size model.

The two principal sources of variation in stimulation are sharply distinguished in our analysis. As in the various published formulations of statistical learning theory, the total population of stimuli ever available during a given experiment is represented by a set $S$ of stimulus elements, or 'cues.' The subset of cues that are available for sampling on a particular trial will be designated the presentation set, denoted by the letter $T$ (with superscripts if needed to distinguish among two or more different presentation sets). Ordinarily the different presentation sets defined for a given experiment correspond to stimuli whose probabilities are under the control of the experimenter, as under a classical discrimination procedure. However, the probabilities of presentation sets may also vary with time, as in fluctuation models for retention, spacing effects, and other time-dependent phenomena. The subset of cues that affect the subject's behavior on a particular trial will be designated the sample, denoted $s$. The probability of a particular sample given the presentation set is assumed to be fixed for a given subject and experimental situation and, in particular, does not vary during the course of learning. Owing to this property, it might be expected that changes in the conditioned status of conditioned elements from trial to trial might be represented as transitions among states of a Markov chain. This is indeed the case, fortunately for calculational purposes, under certain circumstances; not always, however, for the probabilities of presentation sets and reinforcing events may, in general, depend upon outcomes of any number of preceding trials. One of our general results is a theorem prescribing the way in which, for any given experimental routine, states may be defined so as to permit interpretation of the process as a finite Markov chain.

The basic differences between stimulus sampling models and linear models can be explicated in terms of the sample space underlying each. To characterize a single trial in a stimulus sampling model, we require six items of information: the way in which the stimulus population is partitioned into subsets of elements conditioned ($C$) to the various possible responses; the presentation set ($T$); the sample ($s$); the response made by the subject ($A$); the experimenter-defined outcome, reward, punishment, etc. ($O$); and the reinforcing event ($E$). An experiment comprises a sequence of such trials; therefore, the sample space must contain a point corresponding to each possible sequence.

In the linear model of Bush and Mosteller (1955), a trial is characterized by only two of these items, $A$ and $O$, both of which are observable; a third class of events is provided for, but does not actually enter into any applications of the Bush and Mosteller model that have appeared to date.

The linear model investigated by Estes and Suppes (1959a) requires three items of information to specify a trial: $A$, $O$, and $E$. The first two are, again, observable responses and experimenter-defined outcomes, but the third, $E$, is a purely hypothetical 'reinforcing event' that represents the conditioning...
effect of the trial outcome. Event $E_i$ is said to have occurred when the outcome of a trial is such as to increase the probability of the response $A_i$ in the presence of the given stimulus (provided, of course, that this probability is not already at its maximum value). Strictly speaking, an item corresponding to the presentation set $T$ should appear in the linear models, but it has been possible conveniently to suppress it because the linear models have been developed primarily for the case of simple learning in which the same presentation set occurs on all trials. The learning axioms of linear models prescribe linear transformations of response probabilities; the parameters of the 'linear operators' applied on any trial depend, in general, on the response and the outcome of the trial but otherwise have no interpretation in terms of more primitive notions. The learning axioms of stimulus sampling models prescribe how the conditioned status of stimuli in the trial sample changes. Given the reinforcing event of a trial, the changes prescribed by these axioms are strictly deterministic in character. Linear transformations of response probability occur in certain cases of stimulus sampling modek, but they are outcome of the trial but otherwise have no interpretation in terms of more.

In most experiments, $r$ and $t$ are completely determined by the experimenter.

The fourth primitive notion is the sample space $X$. Each element $x$ of the sample space represents a possible experiment, that is, an infinite sequence of trials. In the present theory, each trial may be described by an ordered sextuple $(C, T, s, i, j, k)$, where $C =$ the conditioning function; $T =$ the subset of stimuli presented to the subject on that trial; $s =$ the sampled subset of $T; i =$ the response made on that trial by the subject, $(1 \leq i \leq r); j =$ the outcome of the trial, $(0 \leq j \leq i);$ and $k =$ the reinforcing event occurring on that trial, $(0 \leq k \leq r)$.

Our fifth primitive notion is a probability measure $P$ on the Borel field $\mathcal{B}(X)$ of cylinder sets of $X$. All probabilities must be defined in terms of the measure $P$.

We want to make a number of remarks about this sextuple description of a trial. Neither the conditioning function $C$ nor the sampled set $s$ is ordinarily observable; on the other hand, the response $i$ and the outcome $j$ are directly observable by the experimenter. The situation is more complicated for the presentation set $T$ of stimuli. For the analysis of simple learning, we identify $T$ and $S$ (i.e., $T = S$). In the case of discrimination learning, the set $T$ will vary from trial to trial; its complete observability on a given trial depends on the exact character of the discrimination experiment. We return to this point later. The $r$ responses associated with a sample space $X$ are mutually exclusive and exhaustive. The component $j$ on any trial is to be interpreted as the operationally defined outcome of a trial, e.g., the reward, omission of reward, back beyond the trial on which the reinforcement occurs. The kind of theoretical application to finite automata developed in Suppes (1969) already requires that the reinforcement on a given trial be partially dependent on the response on the preceding trial. Additional work in this direction seems likely. The general Markov theorem given here provides a foundation for such work.

**FORMAL STATEMENT OF THEORY**

**Primitive and Defined Notions**

The theory is based on five primitive notions, each of which has a simple psychological interpretation. The first notion is the set $S$ of stimuli. It is to be emphasized from the beginning that in general the set $S$ is not directly observable. The second and third primitive notions are, respectively, the number $r$ of responses available and the number $t$ of possible trial outcomes. In most experiments, $r$ and $t$ are completely determined by the experimenter.

The fourth primitive notion is the sample space $X$. Each element $x$ of the sample space represents a possible experiment, that is, an infinite sequence of trials. In the present theory, each trial may be described by an ordered sextuple $(C, T, s, i, j, k)$, where $C =$ the conditioning function; $T =$ the subset of stimuli presented to the subject on that trial; $s =$ the sampled subset of $T; i =$ the response made on that trial by the subject, $(1 \leq i \leq r); j =$ the outcome of the trial, $(0 \leq j \leq i);$ and $k =$ the reinforcing event occurring on that trial, $(0 \leq k \leq r)$.

Our fifth primitive notion is a probability measure $P$ on the Borel field $\mathcal{B}(X)$ of cylinder sets of $X$. All probabilities must be defined in terms of the measure $P$.

We want to make a number of remarks about this sextuple description of a trial. Neither the conditioning function $C$ nor the sampled set $s$ is ordinarily observable; on the other hand, the response $i$ and the outcome $j$ are directly observable by the experimenter. The situation is more complicated for the presentation set $T$ of stimuli. For the analysis of simple learning, we identify $T$ and $S$ (i.e., $T = S$). In the case of discrimination learning, the set $T$ will vary from trial to trial; its complete observability on a given trial depends on the exact character of the discrimination experiment. We return to this point later. The $r$ responses associated with a sample space $X$ are mutually exclusive and exhaustive. The component $j$ on any trial is to be interpreted as the operationally defined outcome of a trial, e.g., the reward, omission of reward,
unconditioned stimulus, or knowledge of results. Ordinarily, as already remarked, the outcome is under the control of the experimenter, and always the outcome is observable by the experimenter. The component \( k \), customarily referred to as a 'reinforcing event,' will determine which one of the three learning axioms (see following section) is to be applied on the trial. We shall follow the usual notational convention that the \( k \)th reinforcing event (\( 1 \leq k \leq r \)) corresponds to the learning axiom which prescribes reinforcement of (an increase in probability of) the \( k \)th response, and that the interpretation of \( k = 0 \) is that no response was reinforced on the trial. Frequently, but not always, there is a one-to-one correspondence between outcomes and reinforcing events; when this is the case, it is customary in the literature to ignore the distinction between experimentally defined outcomes and reinforcing events and to speak of the latter as though they were experimental operations (see, e.g., Bush & Mosteller, 1955; Estes, 1959b; Estes & Suppes, 1959a). In general, however, the number \( i \) of observable outcomes may be either smaller or larger than the number of reinforcing events. It is a consequence of our axioms that the probability of a reinforcing event depends only on the response and the outcome of the given trial.

The conditioning function \( C \) is defined over the set of first \( r \) positive integers, and \( C_i \), (i.e., the value of the function \( C \) for the argument \( i \)) is the subset of \( S \) conditioned, or connected, to response \( i \) on the given trial.\(^3\) A conditioning function partitions the set \( S \) of stimuli, assigning thereby each element of \( S \) to exactly one response.\(^4\) Thus if \( r = 2 \) and

\[
S = \{a, b, c\},
\]

then a possible \( C \) is

\[
C_1 = \{a, b\},
\]

\[
C_2 = \{c\}.
\]

To summarize, the sample space \( X \) consists of all possible sequences of trials consisting of sextuples \((C, T, s, i, j, k)\). We call \( X \) the sample space relative to \( S, r, \) and \( t \). It should also be mentioned that the order of events in the sextuple representation \((C, T, s, i, j, k)\) reflects the assumption that this is the temporal order of events on each trial. In other words, the temporal order on any trial is (state of conditioning at beginning of trial) \(\rightarrow\) (presentation of stimuli) \(\rightarrow\) (response) \(\rightarrow\) (outcome) \(\rightarrow\) (reinforcing event) \(\rightarrow\) (state of conditioning at beginning of new trial).

It is also convenient to define the notion of a sampling function \( s \). The domain of \( s \) is the set of responses, i.e., the first \( r \) positive integers; then \( s \) is the subset of \( S \) sampled and conditioned to response \( i \) on the trial which has sampling function \( s \), and of course \( s = \bigcup_{i \leq r} s_i \) is the total sample on that trial.\(^5\) For example, if \( r = 2 \), \( S = \{a, b, c\} \), \( C_1 = \{b\} \), \( C_2 = \{a, c\} \), and \( s = \{a, b\} \), then \( s_1 = \{b\} \) and \( s_2 = \{a\} \); and the function \( s \) is the set of ordered couples \((i, b), (2, a)\). Given a sextuple \((C, T, s, i, j, k)\) then \( s \) is uniquely determined by \( C \) and \( s \) for each \( s_i = C_i \cap s \).

We turn now to a number of defined notions. The first definitions are of certain important subsets, i.e., events, of the sample space \( X \). First, \( A_{i, n} \) is the event: response \( i \) on trial \( n \) (i.e., the set of all possible experimental realizations (elements of \( X \)) having \( i \) as the response component on the \( n \)th trial). Similarly, \( O_{i, n} \) is the event: outcome \( j \) on trial \( n \); and \( E_{k, n} \) is the event: reinforcement \( k \) on trial \( n \). (Informally, we have referred to \( k \) as a 'reinforcing event,' but from the standpoint of the sample space \( X \) it is \( E_{k, n} \) that is an event, i.e., a subset of \( X \). The point is a minor technical one and should cause no confusion.)

By attaching a subscript \( n \) to a presentation set \( T \), we denote as \( T_n \) the set of experimental realizations that have \( T \) as the presentation set on trial \( n \), and similarly for other sets or functions constructed from the set \( S \) of stimuli. Thus, for instance, \( s_{i, n} \) is the set of all elements of \( X \) having \( s \) as the subset of \( S \) sampled and connected to response \( i \) on trial \( n \). Of course, supposing that in a simple discrimination experiment with \( r = 2 \) and \( S = \{a, b, c\} \), the experimenter makes the subset \( T = \{a, b\} \) available for sampling on trials when sequence \( A_1 \) is to be reinforced and \( T' = \{b, c\} \) available when \( A_2 \) is to be reinforced. Then \( T_n \) would be the event that subset \( \{a, b\} \) is available for sampling on trial \( n \) and \( T'_n \) the event that subset \( \{b, c\} \) is available for sampling on trial \( n \). Ordinarily the two conditions are equiprobable, so we would have \( P(T_n) = P(T'_n) = \frac{1}{2} \).

It is also necessary to have a notation for the number of elements conditioned to a given response on trial \( n \), the number of elements sampled, etc. In general, if \( A \) is any set, we designate the cardinality of \( A \) by \( N(A) \). Following usage in learning theory, we let \( N = N(S) \). Moreover, \( N(s) \) is the set of all sequences \( x \) in \( X \) that on the \( n \)th trial have exactly \( N(s) \) sampled stimulus elements. In like fashion we use the notation \( N(T) \), \( N(T_n) \), \( N(C) \), \( N(C_{i, n}) \), \( N(s, C) \), \( N(s) \), \( N(s_{i, n}) \). The distinction between \( N(s) \) and \( N(s_{i, n}) \), and so forth, needs explicit emphasis. \( N(s) \) is an integer standing for the finite number of elements

\(^{3}\) We use 'conditioned' and 'connected' as synonyms.

\(^{4}\) More formally, we may define \( C \) as a conditioning function if, and only if,

(i) \( C \) is a function whose domain is the set of first \( r \) positive integers,
(ii) \( \bigcup C_i = S \),
(iii) if \( i \neq i' \) then \( C_i \cap C_{i'} = 0 \).

\(^{5}\) From a mathematical standpoint it would perhaps be simpler and more natural to let \( s \) be the sampling function and let \( s \) be the sampled subset, but two arguments weigh against this notation. First, the notation for the sampled subset is used more frequently than that for the sampling function. Second, usage in the literature is strongly on the side of the notation \( s \) for the sampled subset.
in $s$, a subset of $S$. In contrast, $N(s_n)$ is an infinite set of sample space points.

Finally, we note that we continually use

$$T, T', T'', \ldots$$
for presentation sets of stimuli,

$$C, C', C'', \ldots$$
for conditioning functions,

$$s, s', s'', \ldots$$
for sampled sets of stimuli.

### General Axioms

Our general axioms for stimulus sampling theory of learning divide naturally into two classes. First there are deterministic axioms that are concerned wholly with the sample space $X$ itself independent of any probability considerations. They may be regarded as *learning* axioms since they deal with deterministic changes from one trial to the next.

The second class consists of the probabilistic axioms, involving the probability measure $P$. This class is naturally partitioned into three *sampling* axioms, three *response* axioms, two *reinforcement* axioms, and one *presentation* axiom. The reinforcement and presentation axioms make explicit a restriction on applicability of the model to situations in which probabilities of stimulus presentations and trial outcomes are set by a person (normally the experimenter) or environmental agency that has no direct access to the subject's internal state. It *may be of interest to consider learning situations in which stimulus presentations or trial outcomes are under the control of the subject (other than by conditionalization on his responses)*, but these cases would not be interpretable within the class of stimulus sampling models here defined.

Our axioms require the notion of an *experimenter's partition*, already introduced in Estes and Suppes (1959a). Here we shall give only a verbal, somewhat loose definition of such a partition.

An experimenter's partition $H(n)$ for trial $n$, with elements $\eta, \eta', \eta'', \ldots$, is a partition of the sample space $X$ such that each element $\eta$ of $H(n)$ is defined only in terms of events $T_{n}, A, A', O_{j, n}, \ldots$, with $1 \leq n', n'' \leq n$ and $1 \leq n''' < n$. In the following axioms that are concerned with ‘independence of path’ assumptions, there is continual reference to $n$- or $(n - 1)$-dimensional cylinder sets, which are simply sets defined in terms of possible trial outcomes through trial $n$ or $n - 1$, respectively. And it should be clear that the elements of an experimenter's partition $H(n)$ are simply special sorts of $n$-dimensional cylinder sets.

Formally, we *say that a quintuple $\mathcal{X} = (S, r, t, X, P)$ is a STIMULUS SAMPLING MODEL* if the following five groups of axioms are satisfied.

### Learning axioms. The three learning axioms describe how the conditioning of stimuli changes from trial to trial as a function of reinforcement and sampling.

For every conditioning function $C$, every sampled subset $s$, and every integer $k$, for $0 \leq k \leq r$, there is a unique conditioning function $C'$ such that for every $n$ the following three axioms are satisfied.

L1. The occurrence of a given conditioning function $C$, a given reinforcement $E_n$, and a given sample $s$ on trial $n$ imply the occurrence of a unique conditioning function $C'$ on trial $n + 1$.

$$C_n \cap E_{n,n} \cap s_n \subseteq C'_{n+1}.$$  

L2. If reinforcement occurs, i.e., $k \neq 0$, all stimuli sampled become conditioned to the reinforced response.

$$C'_{k} = C_k \cup s,$$

and

$$C'_j = C_j \sim s_j \text{ for } j \neq k.$$  

L3. If no reinforcement occurs, i.e., $k = 0$, then there is no change in the state of conditioning.

$$C' = C.$$  

### Sampling axioms.

S1. The probability of a given sample is independent of the trial number.

If $P(T_n) > 0$ and $P(T_s) > 0$ then

$$P(s_n \mid T_n) = P(s_n \mid T_s).$$

S2. Samples of equal size are sampled equally often.

If $s \cup s' \subseteq T$ and $N(s) = N(s')$ and $P(T_n) > 0$ then

$$P(s_n \mid T_n) = P(s'_n \mid T_n).$$

S3. Sampling on any trial is independent of events on previous trials.

If $W_{n-1}$ is an $n - 1$ cylinder set and $Y \subseteq W_{n-1} \cap C_n$ and $P(Y \cap T_n) > 0$ then

$$P(s_n \mid T_n \cap Y) = P(s_n \mid T_n).$$

### Response axioms.

R1. The probability of any response $A_i$ is the ratio of sampled elements connected to $A_i$ to the total number of sampled elements.

If $s \neq 0$ and $P(s_n) > 0$ then

$$P(A_{i,n} \mid s_{n}) = \frac{N(s_i)}{N(s)}.$$
R2. If no stimulus elements are sampled on a trial, the probability of response $A_i$ is equal to some number which does not depend on the conditioning function but may depend on the trial number. If $s = 0$ and $P(s_n) > 0$ then there is a number $\rho_{i,n}$ such that

$$P(A_i | s_n) = \rho_{i,n}.$$  

R3. Given the sample on trial $n$, no additional information about events of previous trials affects the conditional probability of any response $A_i$.

If $W_{n-1}$ is an $n-1$ cylinder set and $Y \subseteq W_{n-1} \cap C_n \cap T_n$ and $P(Y \cap s_n) > 0$ then

$$P(A_{i,n} | s_n \cap Y) = P(A_{i,n} | s_n).$$

Reinforcement axioms.

E1. The probability of an outcome event depends only on previous observables, namely, preceding presentation sets of stimuli, preceding responses, and preceding outcomes.

There is an experimenter's partition $H(n)$ such that if $W_{n-1}$ is any $n-1$ cylinder set, $Y \subseteq W_{n-1} \cap C_n \cap s_n$, $\eta$ is in $H(n)$, and $P(Y \cap \eta) > 0$ then

$$P(O_{j,n} | Y \cap \eta) = P(O_{j,n} | \eta).$$

E2. The probability of a reinforcing event depends only on the response $A_{i,n}$ and the outcome $O_{i,n}$ of the same trial.

If $W_{n-1}$ is an $n-1$ cylinder set and $Y \subseteq W_{n-1} \cap C_n \cap T_n \cap s_n$ and $P(Y \cap A_{i,n} \cap O_{i,n}) > 0$, then

$$P(E_{k,n} | Y \cap A_{i,n} \cap O_{i,n}) = P(E_{k,n} | A_{i,n} \cap O_{i,n}).$$

Presentation axiom.

P1. The probability of a stimulus presentation set depends only on previous observables.

There is an experimenter's partition $H(n-1)$ such that if $W_{n-1}$ is any $n-1$ cylinder set, $Y \subseteq W_{n-1} \cap C_n$, $\eta$ is in $H(n-1)$ and $P(Y \cap \eta) > 0$, then

$$P(T_n | Y \cap \eta) = P(T_n | \eta).$$

Preliminary Theorems

In this section we give some theorems needed for the general Markov theorem that is the subject of the following section.

We first state an independence-of-path result for conditional probability of a response. The gist of it is that if we are given the presentation set $T$ and the set of conditioned stimuli $C_n$ then no further knowledge about conditioning of stimuli to other responses or about events on past trials will affect the probability of response $i$. (In Theorem 1 and hereafter the intersection symbol is suppressed and juxtaposition is used to denote intersection in the interest of simplifying notation. Thus we write $T_n C_{i,n} Y$ instead of $T_n \cap C_{i,n} \cap Y$, etc.)

**Theorem 1.** If $W_{n-1}$ is an $n-1$ cylinder set, $Y \subseteq W_{n-1} C_n$, and $P(T_n C_{i,n} Y) > 0$, then

$$P(A_{i,n} | T_n C_{i,n} Y) = P(A_{i,n} | T_n C_{i,n}).$$

It may be noted that as a special case of the theorem just stated we have

$$P(A_{i,n} | T_n C_n) = P(A_{i,n} | T_n C_{i,n}),$$

provided $P(T_n C_n) > 0$.

We next give expressions for the probability of arriving at any given conditioned subset $C_i$ given the reinforcing event, presentation set, and conditioning function (or merely the conditioned subset $C_i$) on the preceding trial. The first of these expressions refers to the change in the subset of elements connected to a given response on a trial when that response is reinforced.

**Theorem 2.** If $P(E_{i,n} T_n C_{i,n}) > 0$, then

$$P(C_{i,n+1} | E_{i,n} T_n C_{i,n}) = \sum_s P(E_{i,n} | s_n T_n C_{i,n}) P(s_n | T_n),$$

where the summation runs over all $s$ such that $s \cup C_i = C_i$ and $s = T$.

**Proof:** It is easily seen that

$$P(C_{i,n+1} E_{i,n} T_n C_{i,n}) = P(C_{i,n+1} E_{i,n} s_n T_n C_{i,n})$$

$$= \sum_s P(E_{i,n} s_n T_n C_{i,n}).$$

Then, making the appropriate conditionalizations, the desired result is obtained where $C_{i,n}$ is eliminated from $P(s_n | T_n C_{i,n})$ by appropriate use of the sampling axioms.

One might have expected that this basic expression for the change in the conditioning function on a reinforced trial would involve only the factor representing probability of a particular type of sample and not also the factors representing probabilities of the reinforcing event. However, it is possible
for the reinforcing event to depend on the response made by the subject on the given trial, and this in turn depends on the nature of the sampled subset of stimuli. Thus information about the reinforcing event of trial \( n \) may convey information about the stimulus sample. Under some particular reinforcement schedules, e.g., the classical discrimination paradigm or simple non-contingent reinforcement, the factors involving \( E_{n} \) cancel out.

The next two theorems deal similarly with the change in the subset of elements connected to a given response on trials when some alternative response is reinforced or when the neutral event \( E_{0} \) occurs, respectively. The proofs follow that of Theorem 2 closely and are thus omitted.

**THEOREM 3.** If \( k \neq 0 \) and \( k \neq i \) and \( P(E_{x,n}T_{n}C_{i,n}) > 0 \), then

\[
P(C_{i,n+1} \mid E_{i,n}T_{n}C_{i,n}) = \sum_{s} \frac{P(E_{i,n} \mid s_{n}T_{n}C_{i,n})P(s_{n} \mid T_{n})}{P(E_{i,n} \mid T_{n}C_{i,n})},
\]

where the summation now runs over all samples \( s \) such that \( s \subseteq T \) and \( s_{i} \cup C_{i} = C_{i} \).

**THEOREM 4.** If \( P(E_{0,n}T_{n}C_{i,n}) > 0 \), then

\[
P(C_{i,n+1} \mid E_{0,n}T_{n}C_{i,n}) = \begin{cases} 1, & \text{if } C_{i} = C_{r}, \\ 0, & \text{otherwise.} \end{cases}
\]

**MARKOV CHAIN PROPERTY**

With a substantial number of independence-of-path theorems now established, we turn to the proof of what is probably the most important general theorem of stimulus sampling theory, namely, that under very broad conditions an appropriately chosen sequence of events is a Markov chain. The significance of this theorem lies in the fact that the mathematical theory of finite state Markov chains is simpler and more complete than that of nearly any other class of stochastic processes.\(^6\)

In principle, response probabilities can always be computed, given the probabilities of the states of the chain. Except when the number of states is very small, it is usually impractical to give explicit formulas for response probabilities as functions of \( n \), but in many experimentally important cases, asymptotic response probabilities turn out to have simple, closed expressions in terms of the asymptotic state probabilities.\(^6\)

Some Examples

Before turning to formal developments, we want to describe intuitively in somewhat more detail the character of the results in this section and to give a sketch of some examples that show that the results are in a certain sense the best possible. By this we mean that the definition of state in the Markov chain cannot be essentially simplified without destroying the Markovian character of the process.

A result much used in the experimental literature is that the sequence \( \langle C_{1}, C_{2}, \ldots, C_{n}, \ldots \rangle \) of conditioning random variables that take possible conditioning functions as their values forms a Markov chain for simple non-contingent and contingent reinforcement schedules. Superficially it might be thought that the sequence of events (or random variables) that forms a Markov chain would be quite different in discrimination and simple learning experiments. However, as we shall see, this difference is not critical at all. The crucial question is always, how many trials back do the probability dependencies in the reinforcement and presentation schedules extend?

**Double contingent reinforcement with constant presentation set.** For example, let us first consider one of the most direct generalizations, within simple learning, of the simple contingent case, namely, the double contingent case for which the probability of an outcome or reinforcement on trial \( n \) depends on the responses of that trial and the previous trial. For the double contingent case the sequence of conditioning random variables \( \langle C_{1}, C_{2}, \ldots, C_{n}, \ldots \rangle \) is not a Markov chain. Now it might be thought that this is not surprising, but that if we define the states as pairs \( C_{n-1}C_{n} \) of conditioning functions, reflecting the fact that the contingency extends over two trials, then the sequence \( \langle C_{1}C_{2}, C_{2}C_{3}, \ldots, C_{n-1}C_{n}, \ldots \rangle \) would be a Markov chain. However, this conclusion is not true except under very special restrictions. Consider, for example, the following model. The set \( S \) of stimuli has exactly one element that is sampled with probability \( \theta \), \( 0 < \theta < 1 \), on every trial. To show that we do not have a Markov chain we need to show that for some \( n, C, C', C'', C''' \),

\[
P(C_{n+2}C_{n+1} \mid C_{n+1}C_{n-1}) = P(C_{n+2}C_{n+1} \mid C_{n+1}C_{n}). \quad (1)
\]

If we take \( r = t = 2 \),

\[
P(E_{x,n} \mid O_{k,n}) = 1,
\]

\[
P(O_{1,n} \mid A_{j,n}A_{j',n-1}) = \pi_{j',j}
\]

and \( n = 2 \), then it is tedious and lengthy, but not difficult, to establish Equation 1. Rather than give this computation, we believe it will be more instruc-

---

\(^6\) We define a Markov chain to be a Markov process whose transition probabilities are independent of \( n \).
FIGURE 1.
The tree of paths leading from the pair $C_{n-1}C_i$ terminating on trial $n$ to the possible new pairs terminating on trial $n + 1$.

The difficulty with the tree is that we do not have as part of it the responses on trial $n-1$, and thus we are not able to indicate the probabilities $x$ and $1-x$ of outcomes $O_{1,n}$ and $O_{2,n}$ in the appropriate branches. This problem suggests as the second possibility a tree that retains the same definition of state but includes the responses on trial $n-1$ among the branching possibilities.

In the case of this second tree, the difficulty is that the probability of an $A_i$ response on trial $n-1$ is disturbed by the knowledge that the conditioning state on trial $n$ is $C^i$. Without this knowledge, the probability of an $A_i$ response on trial $n-1$ given $C^i_{n-1}$ is simply $\theta + (1 - \theta)p$, where $p$ is the fixed probability of an $A_i$ response when no sample is drawn. With this additional information we cannot compute the probability of $A_{i,n-1}$ without also considering the reinforcement on trial $n-1$, which immediately requires knowledge of the response on trial $n-2$, and thereby invalidates the Markovian character of the tree. An appropriate definition of state for this double contingent case of reinforcement is the response on trial $n-1$ and the conditioning function on trial $n$. A typical tree is shown in Figure 2. Because the probabilities of the various branches are independent of $n$, the trial number subscripts have been dropped.

The critical thing in defining the states of the Markov chain is to include all those events on which the reinforcement probabilities depend, as well as the conditioning function on the given trial. Before proceeding to the general theory, we give some additional examples. In all these examples we assume the same possible conditioning function and the same sampling condition as in the double contingent instance just discussed.

**Noncontingent, Markov reinforcement schedule with constant presentation set.** Let

$$P(O_{i,n+1} | O_{j,n}) = \pi_{ij},$$

be the transition probabilities of a one-stage Markov chain in the trial outcomes. Then the appropriate definition of a state in the Markov learning process is the reinforcement on trial $n-1$ and the conditioning function on trial $n$.

**Noncontingent presentation and contingent reinforcement with lag.** Let $T^m, m = 1, 2$, be the two presentation sets of stimuli, and let the reinforcement schedule be defined by

$$P(O_{j,n} | T^m_{i,n}, A_{i,n-1}) = \pi_{ijm}.$$

The Markov chain then has eight states $A_i C^i T^m$ for $i,j,m = 1, 2$; and it should be clear from the previous examples how to construct the tree with each state, and thereby the transition matrix of the process, once we specify noncontingent probabilities for $T^1$ and $T^2$, say, $\tau_1$ and $\tau_2$.

**Contingent presentation and contingent reinforcement.** To show that it is not sufficient just to consider the past dependence of the reinforcement schedule, a simple discrimination setup will suffice, where the reinforcement is dependent on the presentation sets and where the probability of a presentation set itself depends on the response on the preceding trial. That is, $P(O_{i,n} | T^m_{i,n}) = $
\( \pi_{ni} \) and \( P(T_{n}^{m} | A_{i,n-1}) = \pi_{im} \). For this situation, we may take as the states of the Markov process \( A_{i}C_{i} \), for \( i, j = 1, 2 \). In this particular instance we may also take as states the conditioning functions and the presentation sets, i.e., \( C'T^{m} \). However, to simplify the statement of general results we shall restrict ourselves to states which have the conditioning function as the last event.

**Experimenter’s Schedule**

We turn now to the general theory. We shall mean by an experimenter’s schedule the set of conditional probabilities that determine the schedule of outcome events and the schedule of presentation sets. By virtue of Reinforcement Axiom EI and Presentation Axiom PI this dependence is restricted to preceding outcome events, responses, and presentation sets. We say that an experimenter’s schedule has finite character \( \mathcal{K} \) if there is a collection of finite sequences \( \mathcal{K} = \{ (X_{i1}, \ldots, X_{in}) \} \) with each \( X_{ij} \) an \( O_{j_1} \), an \( A_{i} \), or a \( T^{m} \) and such that, for each \( n \),

\[
P(O_{j,n} | A_{i,n}T_{n}^{m}K_{n-1}W_{n-1}) = P(O_{j,n} | A_{i,n}T_{n}^{m}K_{n-1})
\]

and

\[
P(T_{n}^{m} | K_{n-1}W_{n-1}) = P(T_{n}^{m} | K_{n-1})
\]

for any response \( A_{i} \), presentation set \( T^{m} \), sequence \( K \) in \( \mathcal{K} \), and \( W_{n-1} \) and \( n - 1 \) cylinder set such that

\[
P(A_{i,n}T_{n}^{m}K_{n-1}W_{n-1}) = P(A_{i,n}T_{n}^{m}K_{n-1}) > 0
\]

for Equation 2 and

\[
P(K_{n-1}W_{n-1}) = P(K_{n-1}) > 0
\]

for Equation 3. The notation \( K_{n-1} \) denotes a sequence belonging to \( \mathcal{K} \) that terminates on trial \( O_{n-1} \). For example, if \( \mathcal{K} = \{ (O_{1}, A_{1}), (O_{2}, A_{2}), (O_{3}, A_{3}) \} \) and \( K = (O_{1}, A_{1}) \), then \( K_{n-1} \) is the event \( O_{1,n-2}A_{2,n-1} \). The superscript \( t \) indicates the lag behind trial \( n \). Thus \( t \) = 2 refers to an event on trial \( n - 2 \). A more elaborate and explicit definition and notation could be used, but the sense of the present notation should be clear, and it is adequate for our purposes. We do need to add the provision that for any \( n \) the probabilities of the event sequences \( K_{n-1} \) for all \( K \) in \( \mathcal{K} \) add to one, in order for the conditional probabilities of presentation and reinforcement to be completely fixed. Moreover, it is also understood from the definition of finite character that the conditional probabilities of Equations 2 and 3 are independent of \( n \).

Unfortunately, the sets \( C_{n}K_{n-1} \), when \( K \) is in \( \mathcal{K} \) and \( C \) is a conditioning function, cannot by themselves serve as the states of the Markov chain when the dependencies in the experimenter’s schedule extend back more than one trial, as is illustrated by the example of contingent reinforcement with lag 2 and constant presentation set. Because this is a case of simple learning, that is, the full set \( S \) of stimuli is presented on each trial, the experimenter’s sched-
where $\sum_{s}$ is over all samples $s$ such that $C_k = C_s \cup s_k$ and, for $t \neq k$, $s_t \cup C_t = C_t$, provided if $s$ is empty $N(s) \cap C_t)/N(s)$ is replaced by $p_{t,i,n}$ and

$$\begin{align*}
K_{i,n-1}^{*2} &= O_{i,n-1}A_{i,n-1}T_{n-1}^{m_i} \\
\cdots &= O_{i,n-1}A_{i,n-1}T_{n-1}^{m_{i-1}}O_{i,n-1}A_{i,n-1}T_{n-1}^{m_{i-2}};
\end{align*}$$

(b) $P(C_{i,n}K_{i,n}^{*2} | C_{i,n-1}K_{i-1,n}^{*2}) = 0$ provided $K_{i,n}^{*2} = 0$.

For convenient statement of the general Markov theorem we may define for each $n$ the random variable $K_{i,n}$ that takes as its values the events $C_{i,n}K_{i,n}^{*2}$ with $K$ in $\mathcal{K}$ and $C$ a possible conditioning function.

**THEOREM 6.** (General Markov Theorem). If a stimulus sampling model has an experimenter's schedule of finite character $\mathcal{K}$, then the sequence of random variables $(K_i, K_2, \ldots, K_{i-1}; \ldots)$ is a finite-state Markov chain.

**Proof:** To simplify notation, we shall prove only a special case of the theorem, but the method of proof required is exactly that needed for the general case, and it will be perfectly apparent how to extend the proof in a routine manner. Let

$$\mathcal{K} = \{\{O_j, T_i, m_i, A_i\}\};$$

that is, the experimenter's schedule is determined by the conditional probabilities

$$P(O_j,n \mid O_{j',n-1}T_{n-1}^{m_i}A_{i,n-1}) = \pi(j, j', m, i)$$

and

$$P(T_{n}^{m'} \mid O_{j',n-1}T_{n-1}^{m_i}A_{i,n}) = \pi(m', j', m, i)$$

for $0 \leq j, j' \leq t$, $1 \leq m, m' \leq M$, $1 \leq i \leq r$, where $M$ is the total number of presentation sets.

To establish the theorem for this special case we need to extend $K$ to $K^*$ and prove that—for every $n, i, j, j', m, m', i, i'$, and $i''$, every pair of conditioning functions $C$ and $C'$, and every $n - 2$ cylinder set $W_{n-2}$ which is an intersection of $C_{n}K_{i,n}^{*2}$ sets with $n' < n - 1$ such that $P(C_{n-1}K_{i,n}^{*2}W_{n-2}) > 0$—we have

$$P(C_{n}K_{i,n}^{*} \mid C_{n-1}K_{i,n}^{*2}W_{n-2}) = P(C_{n}K_{i,n}^{*} \mid C_{n-1}K_{i,n}^{*2}),$$

where

$$C_{n}K_{i,n}^{*2} = C_{n}O_{j,n-1}A_{i,n-1}T_{n-1}^{m_i}A_{i',n-2}$$

and

$$C_{n}K_{i,n}^{*} = C_{n}K_{i,n}^{*2} = C_{n}O_{j,n-1}A_{i,n-1}T_{n-1}^{m_i}A_{i',n-2},$$

Now by elementary probability theory

$$P(C_{n}K_{i,n}^{*} \mid C_{n-1}K_{i,n}^{*2}W_{n-2}) = \frac{\sum_{s} P(C_{n}K_{i,n}^{*} \mid C_{n-1}K_{i,n}^{*2}W_{n-2})}{P(C_{n-1}K_{i,n}^{*2}W_{n-2})},$$

In analyzing the various conditional probabilities obtained from the resulting expression, let us use $Y$ for the set of remaining events that are not needed on the basis of independence-of-path results already established and that occur before the event whose conditional probability is being considered. Thus in Equation 10 below,

$$Y = O_{j,n-1}A_{i,n-1}T_{n-1}^{m_i}A_{i',n-2}T_{n-2}^{m'}A_{i',n-3}W_{n-2}.$$
which simply indicate the number of stimuli conditioned to each response, provided the presentation set is constant over trials. Let us define for each \( n \) the random variable \( L_n \) which takes as its values the events \( N(Cn) \alpha K - 1 \) with \( K \) in \( \alpha \) and \( C \) a possible conditioning function. We omit proof of the following theorem.

**THEOREM 7.** If a stimulus sampling model has an experimenter’s schedule of finite character \( \alpha' \), and if \( S \) is the presentation set on all trials, then the sequence of random variables \( (L_1, L_2, \ldots, L_n, \ldots) \) is a finite-state Markov chain.

**ACKNOWLEDGMENTS**

This chapter is a revised and shortened version of a longer technical report (Estes & Suppes, 1959b). The original research was supported by the Office of Naval Research under contract NR 171-034 with Stanford University. Work on the present revised version has been partially supported by the National Science Foundation under grant NSF-GJ-443X to Stanford. We are indebted to David Rogosa for a number of useful comments on this revised version.

**REFERENCES**


