MARKOV LEARNING MODELS FOR MULTIPERSON SITUATIONS, I. THE THEORY

BY

PATRICK SUPPES AND RICHARD C. ATKINSON

TECHNICAL REPORT NO. 21

PREPARED UNDER CONTRACT Nonr 225(17)

(NR 171-034)

FOR

OFFICE OF NAVAL RESEARCH

BEHAVIORAL SCIENCES DIVISION

APPLIED MATHEMATICS AND STATISTICS LABORATORY

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

February 20, 1959

MARKOV LEARNING MODELS FOR MULTIPERSON SITUATIONS, I. THE THEORY[*/]

by

Patrick Suppes and Richard C. Atkinson

§1.1 Introduction. The analysis of learning behavior as a stochastic process began only within the last decade. It seems that the first serious article in this direction was W.K. Estes' "Towards a statistical theory of learning" in the 1950 volume of *Psychological Review*. Shortly thereafter Robert R. Bush and Frederick Mosteller also began publishing papers on stochastic models for learning.

Of slightly older vintage but still quite recent is the development of game theory. In spite of early work by Zermelo, Borel and others, John von Neumann's paper of 1928 is the first real landmark in the subject. The publication by von Neumann and Oskar Morgenstern in 1943 of their treatise *Theory of Games and Economic Behavior* introduced game theory to a much wider circle.

The present monograph is partly concerned with the task of bringing these two disciplines into closer alignment. More exactly, our aim has been to apply learning theory to simple two-person and three-person game situations. The present chapter describes the underlying theory used in our experiments. The second chapter is concerned with various methods,

---

*/ This report is Chapter 1 of a forthcoming monograph. This research was supported by the Group Psychology Branch of the Office of Naval Research, the Behavioral Sciences Division of the Ford Foundation, and the National Science Foundation.

statistical and otherwise, used in the analysis of data. Because the fundamental theory is probabilistic in character, the conceptual separation between the first two chapters is not absolutely clear. Roughly speaking, we proceed as follows. Theoretical quantities which do not depend on observed quantities are derived in the first chapter; examples are asymptotic mean probabilities of response and associated variances. Quantities which do depend on observed data are derived in the second chapter; a typical instance is the derivation of the maximum likelihood estimate of the learning parameter.

The remaining chapters are devoted to detailed presentation of the experiments. Chapter 3 is concerned with some simple zero-sum, two-person games, and Chapter 4 with some non-zero-sum, two-person games. Chapter 5 deals with the analysis of game-theoretical information from the standpoint of what learning theorists call discrimination theory. More particularly, in this chapter we study the effect of showing one player the responses or choices of the other player. Chapter 6 considers experiments concerned with a three-person, simple majority game. Chapter 7 describes some experiments in which the subjects were told various things about the pay-off matrix. Chapter 8 analyzes the effects of monetary pay-offs.

It is pertinent to remark why, before we embarked on our empirical investigations, we thought learning theory would predict the actual behavior of individuals in game situations. To begin with, we endorse the general characterization of learning given by Bush and Mosteller in the opening pages of their book [8, p.3]:

"We consider any systematic change in behavior to be learning whether or not the change is adaptive, desirable for certain purposes, or in accordance with any other such criteria. We consider learning to be 'complete' when certain kinds of stability--not necessarily stereotypy--obtain."

The general character of our experiments is to bring a pair of subjects into an interaction or game situation without telling them everything about the game. (The degree of information given them varies from experiment to experiment.) This restriction of information immediately makes subjects learners as much as players of a game. A subject's behavior naturally changes systematically as information accrues to him.

Readers oriented toward game theory might well wonder what was the point of restricting information so severely as not to show the subjects the payoff matrix, and even in some experiments (Chapter 3) not to tell subjects they were interacting at all. We chose this approach because statistical learning theory had already been tested successfully in a number of experimental studies. Our a priori confidence in learning theory was mainly based on the excellent quantitative results of many of these studies. Game theory, in contrast, was not originally formulated to predict behavior, but rather to <u>recommend</u> it. That is, as a theory it has been normative through and through. Yet many people familiar with game theory have hoped it might describe actual behavior of uninstructed but intelligent persons under certain stable and restricted conditions. Initially we tended to think of our experiments as a kind of competition between game theory and learning theory in their relative ability to predict behavior. However, when we turned to the actual design of experiments, it seemed obvious that the only reasonable thing to do was to begin by staying close to some thoroughly investigated learning setup and not worry about fidelity to the game-theoretic notion of a game. So our "competition" between theoretical approaches turned into the more constructive program of seeing how far learning theory can be extended to predict behavior in situations which correspond ever more closely to real games. As we shall see in Chapters 7 and 8, problems arise when this correspondence becomes very close.

Because statistical learning theory provides the theoretical background of our work, we would like to make certain general remarks about the status of this theory before entering into technical details. Although the theory is of recent origin, the concepts on which it is based have been current for a long period of time, and are the basic concepts of association and reinforcement psychology. These concepts are only three in number: <u>stimulus</u>, <u>response</u> and <u>reinforcement</u>. The great service, to this theoretical orientation, of earlier behavioristic psychologists like Watson, Thorndike, Guthrie, Tolman and Hull is to have developed such concepts in a scientific context, cleanly pruning away from them the tangled notions of common sense and of earlier philosophical psychology.

At this stage it would be a mistake to overemphasize the historical importance of statistical learning theory, for it is too early to evaluate its permanent significance. But it is possible to draw a historical analogy to the development of classical mechanics by Newton, and his successors in the eighteenth century (Euler, Lagrange, LaPlace and others). The qualitative, conceptual work of Descartes was a necessary preliminary for Newton. The virtue of Descartes was to view the world mechanically and thus to sweep aside the subtle distinctions of the Scholastics.

Descartes insisted that the physical world is nothing but matter in motion. In his Principia Philosophiae he explains everything and yet explains nothing. By this we mean he provides a way of looking at the world that is sound, but his explanations of any particular physical phenomena are hopelessly inadequate. To a certain extent the same is true of the earlier association psychologists, although they are definitely more empirically oriented than Descartes.[*]

---

*/ However, it may be said in defense of Descartes that he was not as rationalistic as contemporary opinion would have him. Parts I and II of the Principia are a priori and independent of experience, but Parts III and IV, in which he states his vortex hypothesis for explaining empirical details of the physical world, are regarded by him as hypothetical and in need of empirical support. Descartes is a worthy methodological predecessor to those psychologists (Freud, Tolman and Hull, for instance) who have advanced all-encompassing theories of behavior.

The contribution of statistical learning theory is to use the concepts of association psychology to develop a genuine quantitative theory of behavior. Earlier attempts at quantitative theory, notably Hull's, did not lead to a theory that was mathematically viable. That is to say, in Hull's theory it is impossible to make non-trivial derivations leading to new predictions of behavior. Our contention is that in statistical learning theory we have a theory which has the same sort of "feel" about it that theories in physics have. Non-trivial quantitative predictions can be made. Once we make an experimental identification of stimuli and reinforcing events, it is usually clear how to derive predictions about responses in a manner that is not ad hoc and is mathematically exact.

To a psychologist not familiar with game theory, it might seem more meaningful to say that we have been concerned with the application of learning theory to situations involving social interaction. From this standpoint a distinguishing feature of our work has been to ignore concepts like those of friendliness, cohesiveness, group pressure, opinion discrepancy, which have been important in recent investigations by many social psychologists. We have attempted instead to explain the detailed features of social interaction situations in terms of conditioning concepts. The exchange of information between players in a game, for instance, can be successfully analyzed in terms of an organism's ability to discriminate between stimuli, the important point being that this information may be treated in terms of the notion of stimulus in exactly the same way that perceptual stimuli in a one-person situation are handled. The social situation, qua social, does not require the introduction of new concepts. We do not claim that our experiments on highly structured game situations justify the inference that no new fundamental concepts are required to explain any social behavior. But we think it is important to demonstrate in empirical detail and with quantitative accuracy that no new concepts are needed for a substantial class of social situations.

In the stimulus sampling theory of learning outlined in the next section, as in many other learning theories, an experiment consists of a sequence of trials. Independent of particular theoretical assumptions, the course of events for a given subject on a trial may be roughly described as follows: (i) a set of stimuli is presented; (ii) a response is made by the subject; (iii) a reinforcement occurs. However, the empirical specification of what is meant by stimuli, responses and reinforcements is not a simple matter. The high degree of invariance in the identification of responses and reinforcements in the experiments reported in subsequent chapters arises from the fact that all of our experiments were performed with very similar apparatus and under relatively homogeneous conditions.

Experimental identification of the relevant stimuli is more complex. In fact, it is characteristic of many learning studies, including some of ours, that direct identification of stimuli is not possible. For example, in a simple learning situation for which the physical stimuli are constant from trial to trial, it is not clear how to enumerate the relevant stimuli and their conditioning relations. From a methodological standpoint the concept of stimulus would, for these experiments, seem to be best regarded as a theoretical construct which is useful in deriving experimental predictions.

On the other hand, in discrimination experiments, identification of the stimuli is often natural, and on the basis of such an identification successful predictions of behavior may be made. Nevertheless, as will be evident in subsequent chapters (see particularly Chapter 5), identification of the stimuli does not necessarily imply direct identification of the way in which the stimuli are conditioned to responses. In our analysis these relations of conditioning will turn out to be unobservable states of a Markov process.

As has already been remarked, the theory of behavior used in this book is based on the three concepts of stimulus, response and reinforcement. Clearly, an experiment testing the theory may fail for two reasons: the theory is wrong or the wrong experimental identification of the basic concepts has been made. Various philosophers of science have been fond of emphasizing, particularly in discussions of the general theory of relativity, that by sufficient distortion of the "natural" experimental interpretation of concepts any theory may be saved from failure. This viewpoint is usually discussed as the doctrine of conventionalism. It is not appropriate here to examine the functional role of this doctrine in the working practice of scientists, but it is our view that the issues of conventionalism are not highly relevant to a newly developing science. However beautiful the structure of a theory may be, if stable, non-artificial experimental interpretations leading to new empirical predictions cannot be found, the theory does not have significant empirical import. We believe that the interpretations of stimulus sampling theory of learning given in this book do have such a stable character.

§1.2 Stimulus Sampling Theory of Learning. The basic theory used in our experiments is a modification of stimulus sampling theory as first formulated by Estes and Burke [9], [11], [7], [10]. The exact way in which our theory deviates from theirs is indicated later, but certainly there is no deviation in basic ideas.

We begin our discussion of the general theory by formulating in a non-technical manner its fundamental axioms or assumptions. An exact mathematical formulation is to be found in Estes and Suppes [13].

The first group of axioms deals with the conditioning of sampled stimuli, the second group with the sampling of stimuli, and the third with responses.

CONDITIONING AXIOMS

C1. On every trial each stimulus element is conditioned to exactly one response.

C2. <u>If a stimulus element is sampled on a trial it becomes conditioned with probability θ to the response (if any) which is reinforced on that trial</u>.

C3. <u>If no reinforcement occurs on a trial there is no change in conditioning on that trial</u>.

C4. <u>Stimulus elements which are not sampled on a given trial do not change their conditioning on that trial</u>.

C5. <u>The probability θ of a sampled stimulus element being conditioned to a reinforced response is independent of the trial number and the outcome of preceding trials</u>.

SAMPLING AXIOMS

S1. <u>Exactly one stimulus element is sampled on each trial</u>.

S2. <u>If on a given trial it is known what stimuli are available for sampling, then no further knowledge of the subject's past behavior or of the past pattern of reinforcement will change the probability of sampling a given element</u>.

RESPONSE AXIOM

R1. <u>On any trial that response is made to which the sampled stimulus element is conditioned</u>.

There are a number of remarks we want to make about these axioms, including comparison with the Estes-Burke theory and modifications of it which have been proposed by other people.

To begin with, we may mention that the axioms assume there is a fixed number of responses and reinforcements, and a fixed set of stimulus elements for any specific experimental situation. (The formulation in [13] makes this obvious.) In all the experiments considered in this book the number of stimulus elements is assumed to be small. Because of the small number of stimulus elements we are able to consider explicitly the appropriate Markov process derivable from the theory. We return to this point in detail later.

Turning now to the first group of axioms, those for conditioning, we note to begin with that no use of Axiom C3 is made here, because none of our experiments involves non-reinforcement on any trial. We have included this axiom for completeness; experimental evaluations are reported by Anderson and Grant [1], Atkinson [2], and Neimark [23].

Readers familiar with the Estes-Burke stimulus sampling theory will recognize the basic modification we have incorporated in C1-C5. Namely, the conditioning process has itself been converted into a probabilistic process. For Estes and Burke, sampling of a stimulus element results in it becoming conditioned or connected, to the reinforced response with probability one, that is, deterministically. The experiments we consider lend themselves naturally to the assumption that exactly one stimulus element is presented on each trial and that the element is sampled with probability one. If the conditioning process were then assumed to be deterministic, we would be in the position of having a theory which predicts responses exactly if the sequence of stimulus presentations is known. This is, of course, too strong a theory. The model of conditioning used here is a rather drastic simplification of the actual state of affairs, and it is precisely the probabilistic element of the theory which provides our predictions with the right degree of definiteness.

If, as is natural for our discrimination experiments, we suppose that exactly one stimulus element is presented to the subject on each trial, we may assume that the element is then sampled with probability θ, and thus keep the Estes and Burke deterministic theory of conditioning. In this case, an additional assumption must be made concerning the probability of a response when no sample is drawn. Various assumptions may be introduced to cover the empty sample trials, but most of them seem to be either awkward to work with or ad hoc in character. It may be mentioned in this connection that Estes and Burke initially introduced their deterministic conditioning assumptions for situations in which it was natural to assume that a large number of stimulus elements were present. We have concentrated on a different type of situation, and this probably accounts for our changed emphasis: conditioning is probabilistic, sampling is required. It is possible to derive, for certain experimental situations, different numerical values for observable quantities in our theory and theirs, but in the cases we have looked at so far these differences are too small to make a direct test feasible. Consequently it seems best to consider our conditioning assumptions as being an obvious and not very deep generalization of the Estes-Burke theory. Further generalization may be obtained by introducing dependencies among the probabilities that a stimulus element will be conditioned.

Other kinds of generalization are possible. For example, Restle ([24], [25]) postulates two processes: conditioning and adaptation. Relevant stimuli are conditioned, irrelevant ones are adapted out. Atkinson introduces the notion of trace stimuli [3] and also observing responses [4] as determiners of the stimulus elements to be sampled. La Berge [17] weakens Axiom C1 and assumes that initially some stimulus elements are neutral in the sense that they are conditioned to no response. Independent of any assessment of the merits of these various generalizations, they, like the initial work of Estes and Burke, are aimed at models with a large number of stimulus elements. For reasons which will be evident before the end of this chapter such models are very unwieldy for analyzing our rather complicated experiments.

On the other hand, generalizations in the direction of giving a more complete account of motivation are highly pertinent to our work. Two tentative models of this sort are described in the final chapter.

Axioms C3 and S2 have not usually been explicitly formulated by statistical learning theorists, but they are necessary for strict derivation of the appropriate Markov process representing the course of learning. Axioms of this character are often called <u>independence of path assumptions</u>.

The theory formulated by our axioms would be more flexible and general if (i) Axiom S1 were replaced by the postulate that a fixed number of stimuli are sampled on each trial or that stimuli are sampled with independent probabilities and (ii) Axiom R1 were then changed to read: the probability of a response is the proportion of sampled stimulus elements conditioned to that response. However, for the set of experiments reported in this book it is scarcely possible experimentally to distinguish between S1 and R1 and their generalizations.

In the next three sections we consider detailed application of the learning theory formulated in the above axioms. Before turning to these applications, it will be useful to introduce some concepts and notations which we use throughout the rest of the book. Without going into complete technical detail, we want to introduce the basic sample space and the main random variables we define on this sample space. That the explicit use of random variables has been studiously avoided in most of the literature of statistical learning theory is not a serious argument for avoiding such use here, for the notion of a random variable is exactly the one we need to give precision to the probability assertions we want to make. We remind the reader that a random variable is simply a (measurable) function defined on the sample space, and the overall probability measure on the sample space induces a probability distribution on the values of the random variable.

Our sample space X for a given experiment is just the set of all possible outcomes x of the experiment. It is helpful to think of x as a given subject's behavior in a particular realization of the experiment (or in our situations, sometimes to think of x as the behavior of a particular pair or triple of subjects). For brevity we shall in the sequel speak of subject x, but we really mean "behavior x of the subject."

For simplicity, let us consider first notation for an experiment involving only single subjects, not pairs or triples of subjects. A subject is given a sequence of trials. On each trial the subject makes one of two responses, A_1 or A_2. Using boldface letters for random variables, we may thus define the <u>response random variable</u>:

$$\mathbf{A}_n(x) = \begin{cases} 1 & \text{if subject } x \text{ makes response } A_1 \text{ on trial } n, \\ 0 & \text{if subject } x \text{ makes response } A_2 \text{ on trial } n. \end{cases} \tag{1.2.1}$$

After x's response the correct response is indicated to him by appropriate means. Indication of the correct response constitutes reinforcement. On each trial exactly one of two reinforcing events, E_1 or E_2, occurs. The occurrence of E_i means that A_i (for $i = 1,2$) was the correct response. Thus we may define the <u>reinforcement random variable</u>:

$$\mathbf{E}_n(x) = \begin{cases} 1 & \text{if on trial } n \text{ reinforcement } E_1 \text{ occurred for subject } x, \\ 2 & \text{if on trial } n \text{ reinforcement } E_2 \text{ occurred for subject } x. \end{cases} \tag{1.2.2}$$

When experiments are considered which permit non-reinforcement on some trials, such a trial is called an E_0 trial, and the value of $\mathbf{E}_n$ is 0. However, as already remarked, no such experiments are reported here. The asymmetry between the values of $\mathbf{A}_n$ and $\mathbf{E}_n$ is justified by the fact that we want to sum the A_n's but not the E_n's . When no summation is involved, we prefer to use the values 1 and 2 , which usage readily generalizes to more than two responses.

The A and E notation is standard in the literature. Since the additional random variables we need have not been explicitly used, we introduce some new notation.

The enumeration of which stimulus elements are conditioned to which responses may be represented by a <u>state of conditioning random variable</u> $\underline{C}_n$. We use the term 'state' because the possible values of this random variable on a given trial correspond to the possible states of the fundamental Markov process we introduce in the next section. A general definition of the state of conditioning random variable is not feasible, since the definition depends on the number of stimulus elements in the experiment. For application in the next section, where only a single stimulus element s_1 is assumed, we define $\underline{C}_n$ as follows:

$$\underline{C}_n(x) = \begin{cases} 1 & \text{if on trial } n\text{ , } s_1 \text{ is conditioned} \\ & \text{to } A_1 \text{ for subject } x\text{ ,} \\ \\ 2 & \text{if on trial } n\text{ , } s_1 \text{ is conditioned} \\ & \text{to } A_2 \text{ for subject } x\text{ .} \end{cases} \tag{1.2.3}$$

Note that for an experiment with a single stimulus element, it follows from Axioms S1 and R1 that */

$$(1.2.4) \qquad \begin{cases} P(\underline{A}_n = 1 \mid \underline{C}_n = 1) = 1 \\ \\ P(\underline{A}_n = 1 \mid C_n = 2) = 0 \ . \end{cases}$$

We also need a <u>random variable</u> $\underline{F}_n$ <u>for effectiveness of conditioning</u>. The value of this random variable corresponds to the sampled stimulus element's becoming conditioned, or not, to the reinforced response. If it does the conditioning situation is <u>effective</u>, otherwise not. In view of the sampling Axiom S1, which says that exactly one element is sampled on each trial, the definition of $\underline{F}_n$ is simple.

$$(1.2.5) \qquad \underline{F}_n(x) = \begin{cases} 1 & \text{if on trial } n \text{, the stimulus element} \\ & \text{sampled is effectively conditioned,} \\ \\ 0 & \text{otherwise.} \end{cases}$$

---

*/ Notation like $P(A_n = 1)$ is standard. The explicit definition in terms of the sample space is:

$$P(\underline{A}_n = 1) = P(\{x: \ \underline{A}_n(x) = 1\}) \ ,$$

where $\{x: \ \underline{A}_n(x) = 1\}$ is the set of all experimental outcomes x which have response A_1 on trial n .

It follows immediately from Axiom C2 that

$$(1.2.6) \qquad \begin{cases} P(\underline{F}_n = 1) = \theta \\ \\ P(\underline{F}_n = 0) = 1 - \theta \quad . \end{cases}$$

We shall use F_1 and F_0 informally to indicate the event of the conditioning being effective or not. This usage is similar to that which we have adopted for A_1, A_2, E_1, and E_2 .

In most experimental studies response probabilities are computed by averaging over a block of trials as well as subjects. For this purpose we introduce the random variable $\overline{\underline{A}}_N$ for the sum of responses:

$$(1.2.7) \qquad \overline{\underline{A}}_N(x) = \sum_{n=1}^{N} \underline{A}_n(x) \quad .$$

Although this definition indicates that we sum from trial 1 of the experiment, ordinarily this is not our procedure. When the summation begins on trial m , and it is necessary to be explicit, the following notation will be used: */

$$(1.2.8) \qquad \overline{\underline{A}}_{m,N}(x) = \sum_{n=m+1}^{m+N} \underline{A}_n(x) \quad .$$

Obviously a troublesome problem is what to do about Definitions (1.2.1), (1.2.2), (1.2.3), (1.2.5), (1.2.7) and (1.2.8) when we turn

---

*/ In the experimental literature there has been a fair amount of confusion between the random variables $\underline{A}_n$ and $\overline{\underline{A}}_{m,N}$.

from experiments with individual subjects to pairs of subjects. The simplest systematic device is to add a superscript (1) or (2) to designate the first or second member of the pair. The practical difficulty is that this notation is rather cumbersome. Often we shall call one member of the pair of subjects <u>player</u> A and the other <u>player</u> B . We then define:

$$(1.2.9) \qquad \begin{cases} \underline{A}_n = \underline{A}_n^{(1)} \\ \underline{B}_n = \underline{A}_n^{(2)} \\ \alpha_n = P(\underline{A}_n = 1) \\ \beta_n = P(\underline{B}_n = 1) \\ \gamma_n = P(\underline{A}_n = 1, \underline{B}_n = 1) \end{cases}$$

Note that γ_n is the probability of a joint event, namely an A_1 response by player A and a B_1 response by player B . In a similar fashion we define sums of random variables and Cesáro mean probabilities:

$$(1.2.10) \qquad \begin{cases} \underline{\bar{A}}_N = \sum_{n=1}^{N} \underline{A}_n \\ \underline{\bar{B}}_N = \sum_{n=1}^{N} \underline{B}_n \\ \bar{\alpha}_N = \frac{1}{N} \sum_{n=1}^{N} \alpha_n \\ \bar{\beta}_N = \frac{1}{N} \sum_{n=1}^{N} \beta_n \\ \bar{\gamma}_N = \frac{1}{N} \sum_{n=1}^{N} \gamma_n \end{cases}$$

Summation starting at trial m rather than at trial 1 is defined in an analogous fashion. Finally, if the appropriate limits exist, we define:

$$(1.2.11)\quad \begin{cases} \alpha = \lim\limits_{n\to\infty} \alpha_n \\ \bar{\alpha} = \lim\limits_{N\to\infty} \bar{\alpha}_N \\ \beta = \lim\limits_{n\to\infty} \beta_n \\ \bar{\beta} = \lim\limits_{N\to\infty} \bar{\beta}_N \\ \gamma = \lim\limits_{n\to\infty} \gamma_n \\ \bar{\gamma} = \lim\limits_{N\to\infty} \bar{\gamma}_N \end{cases}$$

Because the other random variables are not often referred to explicitly in subsequent chapters we do not introduce an abbreviated notation for their use in the two-person experiments.

§1.3 Simple Example: Markov Model for Non-contingent Case. We shall now illustrate the method of deriving, from the learning axioms stated in the last section, the Markov learning process for a given experimental situation. The sequence of events on a trial is:

Stimulus sampled → Response made → Reinforcement occurs → Conditioning of sampled stimulus.

We consider what is from a theoretical standpoint one of the simplest cases: non-contingent reinforcement. This case is defined by the condition that the probability of E_1 on any trial is constant and independent of the subject's responses. It is customary in the literature to call this probability π . Thus

$$(1.3.1) \qquad \begin{cases} P(\underline{E}_n = 1) = \pi \\ P(\underline{E}_n = 2) = 1 - \pi \end{cases} .$$

We assume further that the set S of stimulus elements contains exactly one element which we label s_1 . The definition of the random variable $\underline{C}_n$ for the state of conditioning is thus defined by (1.2.3) of the previous section. It is possible to interpret s_1 as the ready signal for a trial, but a physical identification of s_1 is not necessary. Moreover, if two or more stimulus elements are postulated rather than one there is no obvious clear-cut physical interpretation of the stimuli.

What we may prove from our axioms is that the sequence of random variables $\underline{C}_1, \underline{C}_2, \underline{C}_3, \ldots, \underline{C}_n, \ldots$ is a Markov chain.*/ This means, among other things, that knowing the conditioning on trial n , the conditional probability

$$(1.3.2) \qquad P(\underline{C}_{n+1} = j \mid \underline{C}_n = i)$$

---

*/ For an exact definition of Markov chains, see Feller [14] or Kemeny and Snell [16]. For more complicated schedules of reinforcement this particular sequence of random variables may not be a Markov chain.

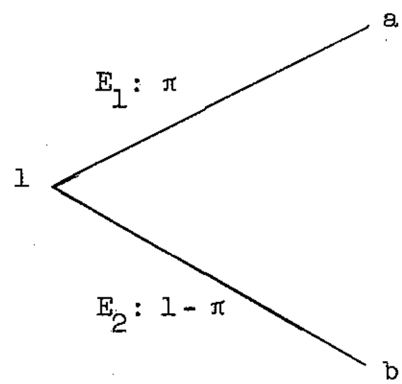
is unchanged by knowledge of the conditioning on any trials preceding n. This fact is characteristic of Markov processes. The process is a Markov chain when the transition probabilities (1.3.2) are independent of n, that is, constant over trials. When we have a chain, the transition probabilities may be represented by a matrix (p_{ij}); obviously the process is completely characterized by this matrix and the initial probabilities of a response. For explicitness we note:

$$p_{ij} = P(\underline{C}_{n+1} = j \mid \underline{C}_n = i) , \tag{1.3.3}$$

that is, p_{ij} is the probability that $\underline{C}_{n+1} = j$ given that $\underline{C}_n = i$. In the usual language of Markov processes, the values i and j of the random variable $\underline{C}_n$ are the states of the process. When there is but one stimulus element and two responses, there are only two states in the process, 1 and 2 (see (1.2.3)).

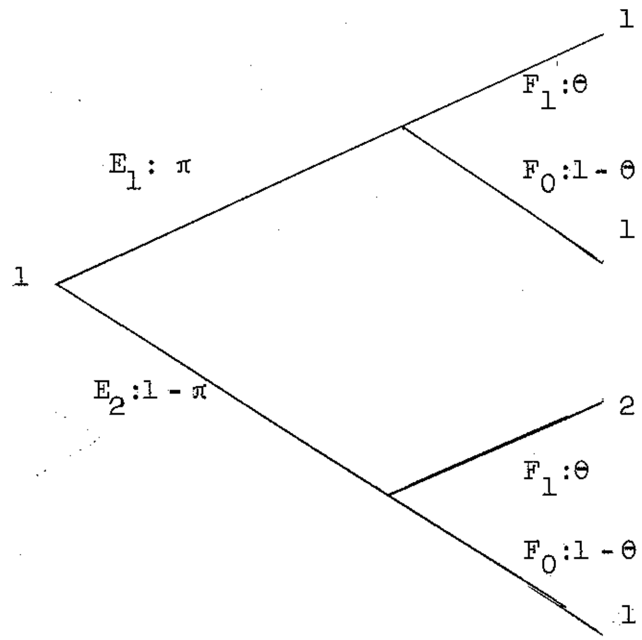
We now use the axioms of the preceding section and the particular assumptions for the non-contingent case to derive the transition matrix (p_{ij}). In making such a derivation it is convenient to represent the various possible events happening on a trial by a tree. Each set of branches emanating from a point must represent a mutually exclusive and exhaustive set of possibilities. Thus, suppose that at the end of trial n subject x is in state 1. At the beginning of trial $n+1$, x will make response A_1, then either E_1 or E_2 will occur, E_1 with probability π, and E_2 with probability $1-\pi$. We thus have this much of a tree:

(1.3.4)



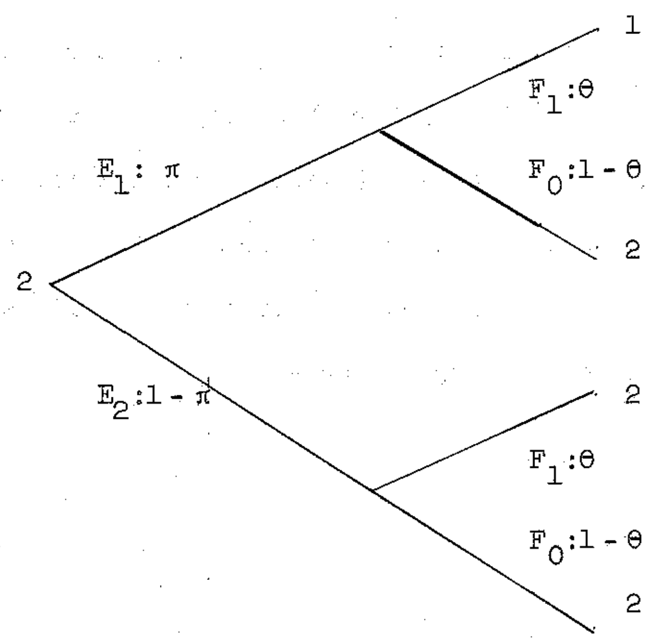
To complete the tree, we need to indicate the possibilities at a and b. At a, the stimulus s_1 will become conditioned to A_1 with probability θ (event F_1) and remain unchanged with probability $1-\theta$ (event F_0). But since x is already in state 1, this means there is only one possibility at a: stay in state 1. At b, the situation is different. With probability θ the stimulus s_1 will become conditioned to A_2 (event F_1) since the occurrence of E_2 reinforced A_2, that is, x will go from state 1 to state 2. And with probability $1-\theta$ event F_0 occurs, so that x will stay in state 1. We then have as the complete tree:

(1.3.5)



Assuming now that x is in state 2 at the end of trial n, we derive by exactly the same argument, the other tree (clearly we always need exactly as many trees as there are states in the process to compute the transition matrix (p_{ij})) .

(1.3.6)



Each path of a tree, from beginning point to a terminal point, represents a possible outcome on a given trial. The probability of each path is obtained by multiplication of conditional probabilities. Thus for the tree of (1.3.5) the probability of the top path of the four may be represented by:

$$P(\underline{E}_n = 1, \underline{F}_n = 1) = P(\underline{F}_n = 1 \mid \underline{E}_n = 1)P(\underline{E}_n = 1)$$

$$= \theta\,\pi \quad .$$

Of the four paths of (1.3.5), three of them lead from state 1 to state 1. Thus

$$\begin{aligned} p_{11} &= P(\underline{C}_{n+1} = 1 \mid \underline{C}_n = 1) \\ &= \theta\,\pi + (1-\theta)\pi + (1-\theta)(1-\pi) \\ &= \theta\,\pi + (1-\theta) \quad . \end{aligned}$$

Similarly

$$p_{12} = \theta(1-\pi) \quad .$$

Notice, of course, that

$$p_{11} + p_{12} = 1 \quad .$$

By similar computations for (1.3.6), we obtain:

$$p_{21} = \theta\,\pi$$

$$\begin{aligned} p_{22} &= (1-\theta)\pi + \theta(1-\pi) + (1-\theta)(1-\pi) \\ &= 1 - \theta\,\pi \quad . \end{aligned}$$

Combining the above results the transition matrix for the non-contingent case with one stimulus element is:

$$(1.3.7)\qquad \begin{array}{c|cc} & 1 & 2 \\ \hline 1 & \theta\pi + (1-\theta) & \theta(1-\pi) \\ 2 & \theta\pi & 1-\theta\pi \end{array}$$

Before examining some of the predictions which may be derived for the Markov chain represented by (1.3.7), some general remarks are pertinent concerning the relation of this process to the learning axioms of §1.2. The central problem may be illustrated by an example. Suppose subject x makes an A_1 response on trial n and reinforcing event E_1 then occurs. Assuming there is a single stimulus element, it follows that it must be conditioned to A_1 in order for A_1 to occur on trial n. But if E_1 then occurs, according to Axiom C2 the conditioning of s_1 cannot change and we predict, using Axioms S1 and R1, that response A_1 will occur on trial $n+1$ with probability one. This prediction may be represented by:

$$(1.3.8)\qquad P(\underline{A}_{n+1} = 1 \mid \underline{A}_n = 1, \underline{E}_n = 1) = 1 .$$

Equation (1.3.8) provides a very sharp test, in fact one which is sure to lead to rejection of the theory. On the other hand, (1.3.8) cannot be derived from the Markov process represented by (1.3.7).

The difficulty with the one-element model (meaning the assumption of one stimulus element) is that the fundamental theory laid down by

Axioms C1-C5, S1, S2, R1 is for this model deterministic in all but minor respects. In particular, from the response on trial $n+1$ we can always derive exactly what was, according to the theory, the conditioning on trial n. Effectively then, the random variables A_n, E_n, C_n and F_n are all observable, and the values of the dependent variables A_n, C_n and F_n can be predicted in a nearly deterministic fashion from experimental protocols for individual subjects.*/

On the other hand, the Markov process defined by (1.3.7) leads only to probabilistic predictions for the values of A_n and C_n, and no predictions about E_n and F_n. The assumed distributions on the latter two random variables are used in the derivation of (1.3.7) but are not observed in any direct fashion. Contrary to an opinion that seems to be widely held by psychologists, it is possible to compare the fit of (1.3.7) to the protocol of an individual subject. In fact, a standard goodness of fit test is available (see §2.1). Naturally a goodness of fit test can also be made for a sample of protocols drawn from a homogeneous population of subjects. Thus the Markov model defined by (1.3.7) is thoroughly testable, but it is in no respect deterministic.

---

*/ If A_n and E_n have the same value then the value of F_n cannot be observed, but it does not matter in this case. If the values of A_n and E_n differ on a trial, the value of F_n is determined uniquely by the value of A_{n+1}, and hence it is observable since A_{n+1} is observable.

If we assume that it seems too much to ask for a deterministic theory of learning at this stage of development, the above discussion does not really entail that the fundamental theory embodied in the axioms of §1.2 should be abandoned, or perhaps regarded as a "make believe" theory from which realistic stochastic processes like (1.3.7) may be derived. For the fundamental theory has a sharply deterministic character only when we assume there is but a single stimulus element. Notice that the axioms of §1.2 say nothing about the number of stimulus elements. When we assume there are two stimulus elements, say s_1 and s_2 in the non-contingent case, the random variables $\underline{C}_n$ and $\underline{F}_n$ are not observable, and few deterministic predictions can be made.

The one-element model has so many special features that it will be useful to discuss the two-element model in some detail. Let us introduce $\underline{S}_n$ as the sampling random variable defined for the two-element model by:

$$\underline{S}_n(x) = \begin{cases} 1 & \text{if } s_1 \text{ is sampled on trial } n , \\ 2 & \text{if } s_2 \text{ is sampled on trial } n . \end{cases} \tag{1.3.9}$$

The trial sequence beginning with $\underline{C}_n$ and ending with $\underline{C}_{n+1}$ may then be represented by:

$$\underline{C}_n \rightarrow \underline{S}_n \rightarrow \underline{A}_n \rightarrow \underline{E}_n \rightarrow \underline{F}_n \rightarrow \underline{C}_{n+1} \tag{1.3.10}$$

We also need to define the random variable $\underset{\sim}{C}_n$ representing the state of conditioning for the two-element model. For expository purposes we make the value of $\underset{\sim}{C}_n$ simply the <u>set</u> of stimulus elements conditioned to the A_1 response. Here and subsequently 0 is used to designate the empty set as well as the number zero.*/ Thus if

$$\underset{\sim}{C}_n(x) = 0$$

this means neither s_1 nor s_2 is conditioned to A_1.

$$(1.3.11) \quad \underset{\sim}{C}_n = \begin{cases} \{s_1, s_2\} & \text{if } s_1 \text{ and } s_2 \text{ are both conditioned to } A_1 \\ \{s_1\} & \text{if } s_1 \text{ is conditioned to } A_1 \text{ and } s_2 \text{ is conditioned to } A_2 \\ \{s_2\} & \text{if } s_2 \text{ is conditioned to } A_1 \text{ and } s_1 \text{ is conditioned to } A_2 \\ 0 & \text{if neither element is conditioned to } A_1 . \end{cases}$$

As (1.3.11) indicates the two-element model leads to a four state Markov process. In deriving this process we give here only the first two trees, that is, the two for which we are at the beginning of the trial in

---

*/ We use a familiar notation for sets. A set is described by writing the names of its members, separated by commas, and then enclosing the whole in braces. Thus $\{s_1, s_2\}$ is the set of two stimulus elements s_1 and s_2.

state $\{s_1,s_2\}$ or state $\{s_1\}$. Trees for the other two are in all essentials the same. Note that we have a set of branches for each of the intermediate steps in (1.3.10), that is, the successive branches correspond to the sequence

$$\underline{S}_n \rightarrow \underline{A}_n \rightarrow \underline{E}_n \rightarrow \underline{F}_n \quad . \tag{1.3.12}$$

But since by Axioms C1 and S1 the value of $\underline{A}_n$ is uniquely determined by the values of $\underline{C}_n$ and $\underline{S}_n$, that is, the response is uniquely determined by the state of conditioning and the single sampled stimulus element, we may reduce (1.3.12) to:

$$\underline{S}_n \rightarrow \underline{E}_n \rightarrow \underline{F}_n \quad . \tag{1.3.13}$$

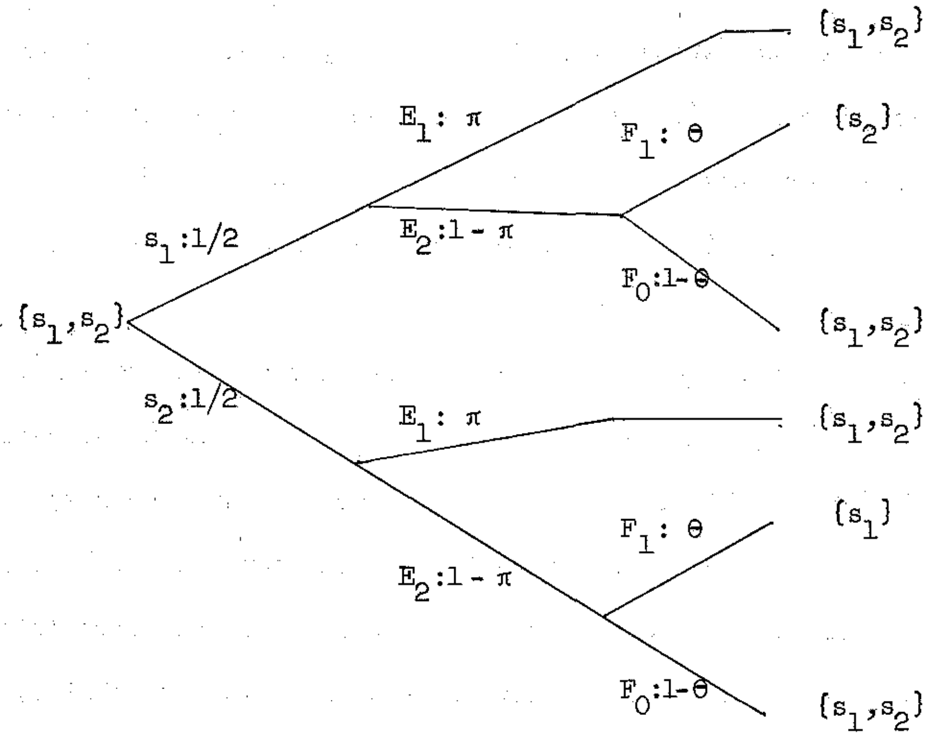
Each of the random variables in (1.3.13) has two possible values, which means then that there are eight possible paths in each of the four trees corresponding to the four possible states of conditioning. However, some reduction in the number of paths may be made by observing that if the sampled stimulus is conditioned to the response which is reinforced, the effectiveness of conditioning is irrelevant. For example,

$$\begin{aligned} &P(\underline{C}_{n+1} = \{s_1\} \mid \underline{C}_n = \{s_1\},\ \underline{S}_n = 1,\ \underline{E}_n = 1,\ \underline{F}_n = 1) = \\ &P(\underline{C}_{n+1} = \{s_1\} \mid \underline{C}_n = \{s_1\},\ \underline{S}_n = 1,\ \underline{E}_n = 1,\ \underline{F}_n = 0) = 1 \ . \end{aligned} \tag{1.3.14}$$

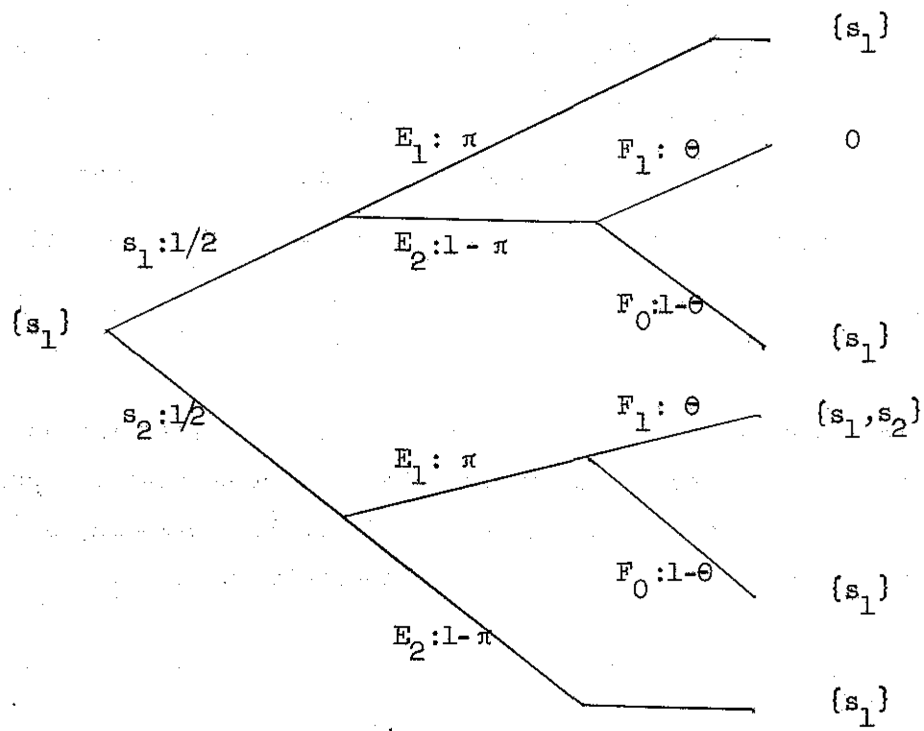
We take account of this irrelevancy to reduce the number of paths. Thus there are not eight but six paths in the tree of any state.

Given (1.3.13) and conditional probabilities like (1.3.14), it is straightforward to derive the trees:

(1.3.15)


{s_1,s_2}
s_1:1/2
s_2:1/2
E_1: π
E_2:1 - π
E_1: π
E_2:1 - π
F_1: Θ
F_0:1-Θ
F_1: Θ
F_0:1-Θ
{s_1,s_2}
{s_2}
{s_1,s_2}
{s_1,s_2}
{s_1}
{s_1,s_2}

(1.3.16)


{s_1}
s_1:1/2
s_2:1/2
E_1: π
E_2:1 - π
E_1: π
E_2:1- π
F_1: Θ
F_0:1-Θ
F_1: Θ
F_0:1-Θ
{s_1}
0
{s_1}
{s_1,s_2}
{s_1}
{s_1}

The tree for the state of conditioning represented by the empty set is symmetrical to (1.3.15), and the tree for $\{s_2\}$ is similar to (1.3.16). One assignment of probabilities in (1.3.15) and (1.3.16) which is not justified by any of the preceding discussion is that of the equi-probabilities of $1/2$ to s_1 or s_2 being sampled. Axiom S1 requires simply that exactly one stimulus element be sampled, and the special experimental conditions of the non-contingent case do not entail the probability with which s_1 or s_2 will be sampled. We make this additional assumption in its simplest form, but it is clear how the analysis could be carried through under the supposition that s_1 is sampled with probability ω and s_2 with probability $1-\omega$. When we come to discrimination experiments, where one player discriminates between known responses of the other player, we shall see that the probabilities corresponding to $1/2$ for s_1 and s_2 fall out in a natural manner from the theory (see Chapter 5).

From (1.3.15), (1.3.16) and the two additional trees not displayed, the transition matrix obtained for the two-element non-contingent model is as follows: */

*/ The possibility will be examined later of collapsing states $\{s_1\}$ and $\{s_2\}$ into a single state to yield a three-state process, where the states are designated 0, 1, 2, depending simply on the <u>number</u> of stimulus elements conditioned to A_1.

$$
(1.3.17)\qquad
\begin{array}{c|cccc}
 & 0 & \{s_1\} & \{s_2\} & \{s_1,s_2\} \\
\hline
0 & 1-\theta\pi & \theta\pi/2 & \theta\pi/2 & 0 \\
\{s_1\} & \theta(1-\pi)/2 & 1-\theta/2 & 0 & \theta\pi/2 \\
\{s_2\} & \theta(1-\pi)/2 & 0 & 1-\theta/2 & \theta\pi/2 \\
\{s_1,s_2\} & 0 & \theta(1-\pi)/2 & \theta(1-\pi)/2 & (1-\theta)+\theta\pi
\end{array}
$$

Detailed analysis is not required to see that the two-element model meets the criticisms mentioned above of the one-element model. The probabilistic character of the Markov chain represented by (1.3.17) is also part of the elementary process given by the trees (1.3.15) and (1.3.16). Knowing which response and which reinforcing event occurred on trial n does not permit us to make a deterministic prediction about the response on trial $n+1$. It should be noted that the states of conditioning are not observable; consequently deterministic predictions are generally not possible in the two-element model.

On the other hand, statistical analysis of the degree to which the Markov chain of (1.3.17) fits empirical data is more difficult and less satisfactory than the corresponding analysis for the Markov chain represented by (1.3.7). We return to this problem in the next chapter. If a two-element model were adopted for each subject then all of our two-person models would have at least sixteen states in the appropriate Markov chain. It is mainly to avoid the burden of this increased complexity that we have generally restricted ourselves to the one-element models. We believe the experimental predictions reported in subsequent chapters have been good enough to justify this restriction.

We now turn to the derivation of asymptotic probabilities of response for the Markov chains (1.3.7) and (1.3.17). That is, we want to find the quantity

$$\lim_{n \to \infty} P(\underset{\sim}{A}_n = 1) \tag{1.3.18}$$

if the appropriate limit exists. In most experiments estimates of (1.3.18) are obtained by averaging over subjects <u>and</u> a final block of trials. Thus, we could as well ask for the Cesáro mean asymptotic probability

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=m+1}^{m+N} P(\underset{\sim}{A}_n = 1) \ . \tag{1.3.19}$$

It is a well-known result that when both these limits exist, they are identical, although (1.3.19) may exist and (1.3.18) not. It is also well-known that for any finite-state Markov chain the limit (1.3.19) exists.

Some notation less awkward than (1.3.18) is useful. If (p_{ij}) is the transition matrix then $p_{ij}^{(n)}$ is the probability of being in state j at trial $r+n$ given that at trial r we were in state i. We define this quantity recursively:

$$\begin{aligned} p_{ij}^{(1)} &= p_{ij} \\ p_{ij}^{(n+1)} &= \sum_{\nu} p_{i\nu} p_{\nu j}^{(n)} \ . \end{aligned} \tag{1.3.20}$$

Moreover, if the appropriate limit exists and is independent of i, we set

$$u_j = \lim_{n \to \infty} p_{ij}^{(n)} \quad . \tag{1.3.21}$$

In particular, in the one-element model discussed above

$$u_1 = \lim_{n \to \infty} P(A_n = 1) \quad .$$

The limiting quantities u_j exist for any finite state Markov chain which is irreducible and aperiodic. A Markov chain is irreducible if there is no closed proper subset of states, that is, a proper subset of states such that once within this set the probability of leaving it is zero. For example, the chain whose transition matrix is:

	1	2	3
1	$\frac{1}{2}$	$\frac{1}{2}$	0
2	$\frac{1}{4}$	$\frac{3}{4}$	0
3	0	$\frac{1}{2}$	$\frac{1}{2}$

is reducible, because the set $\{1,2\}$ of states is a proper closed subset. A Markov chain is aperiodic if there is no fixed period for return to any state; that is, to put it the other way, a chain is periodic if a return to some state j having started in j is impossible except in $t, 2t, 3t, \ldots$ trials, for $t > 1$. Thus the chain whose matrix is:

	1	2	3
1	0	1	0
2	0	0	1
3	1	0	0

has period $t = 3$ for return to each state.

All of the Markov chains we consider in this book are irreducible and aperiodic. Moreover, since each has only a finite number of states, the limiting quantities u_j exist in all cases and are independent of the initial state on trial 1 . If there are r states, we call the vector $\underline{u} = (u_1, u_2, \ldots, u_r)$ the <u>stationary probability</u> vector of the chain. It may be shown (Feller [14], Frechet [15]) that the components of this vector are the solutions of the r linear equations

$$u_j = \sum_{\nu} u_\nu p_{\nu j} \qquad j = 1, \ldots, r \tag{1.3.22}$$

such that $\sum u_j = 1$.

Thus to find the asymptotic probabilities u_j of a state, we need only find the solutions of the r linear equations (1.3.22). The intuitive basis of this system of equations seems clear. Consider a two-state process. Then the probability p_{n+1} of being in state 1 is just:

$$p_{n+1} = p_{11}p_n + p_{21}(1 - p_n) \ ,$$

but at asymptote

$$p_{n+1} = p_n = u_1$$

$$1 - p_n = u_2 \ ,$$

whence

$$u_1 = p_{11}u_1 + p_{21}u_2 \ ,$$

which is the first of the two equations of the system (1.3.22) when $r = 2$.

We now find the asymptotic probabilities u_j for the one-element and two-element non-contingent models. Recalling (1.3.7), the transition matrix for the one-element model is:

$$\begin{array}{c|cc} & 1 & 2 \\ \hline 1 & \theta\pi + (1-\theta) & \theta(1-\pi) \\ 2 & \theta\pi & 1-\theta\pi \end{array}$$

The two equations given by (1.3.22) are:

$$(1.3.23) \qquad \begin{cases} u_1 = [\theta\pi + (1-\theta)]u_1 + \theta\pi u_2 \\ u_2 = \theta(1-\pi)u_1 + (1-\theta\pi)u_2 \ , \end{cases}$$

and the normalizing assumption is

$$(1.3.24) \qquad u_1 + u_2 = 1 \ .$$

Using (1.3.24) and the first equation of (1.3.23) we obtain:

$$u_1 = [\theta\pi + (1-\theta)]u_1 + \theta\pi(1-u_1) \ ,$$

and solving for u_1 we conclude

$$u_1 = \pi \quad . \tag{1.3.25}$$

For the two-element model, the five equations, including the normalizing assumption are, on the basis of (1.3.17):

$$\left\{\begin{array}{llllll} u_1 = & (1-\theta\pi)u_1 & + \frac{1}{2}\theta(1-\pi)u_2 & + \frac{1}{2}\theta(1-\pi)u_3 & \\ u_2 = & \frac{1}{2}\theta\pi u_1 & + (1-\frac{1}{2}\theta)u_2 & & + \frac{1}{2}\theta(1-\pi)u_4 \\ u_3 = & \frac{1}{2}\theta\pi u_1 & & + (1-\frac{1}{2}\theta)u_3 & + \frac{1}{2}\theta(1-\pi)u_4 \\ u_4 = & & \frac{1}{2}\theta\pi u_2 & + \frac{1}{2}\theta\pi u_3 & + [(1-\theta)+\theta]u_4 \\ u_1 + u_2 + u_3 + u_4 = 1 \quad . & & & & \end{array}\right. \tag{1.3.26}$$

To indicate that systems of equations like (1.3.26) may often be solved by a little insight rather than by applying routine methods which guarantee an answer, we proceed to solve this system. Note initially that the first three equations simplify to:

$$\pi u_1 = \frac{1}{2}(1-\pi)(u_2 + u_3) \tag{1}$$

$$u_2 = \pi u_1 + (1-\pi)u_4 \tag{2}$$

$$u_3 = \pi u_1 + (1-\pi)u_4 \quad . \tag{3}$$

We conclude from (2) and (3) that

(4) $$u_3 = u_2 \ ,$$

whence we may express u_1 and u_4 in terms of u_2, using (1) and (2)

(5) $$u_1 = \frac{1-\pi}{\pi} u_2$$

(6) $$u_4 = \frac{\pi}{1-\pi} u_2 \ .$$

Substituting from (4), (5) and (6) into the fifth equation of (1.3.26), we get:

(7) $$\left[\frac{1-\pi}{\pi} + 1 + 1 + \frac{\pi}{1-\pi}\right] u_2 = 1 \ ,$$

whence simplifying (7)

$$[(1-\pi)^2 + 2\pi(1-\pi) + \pi^2] u_2 = \pi(1-\pi)$$

but since the coefficient of u_2 is

$$(1-\pi+\pi)^2 = 1 \ ,$$

we infer

$$u_2 = \pi(1-\pi) \ ,$$

and by appropriate substitution

$$u_1 = (1-\pi)^2$$

$$u_3 = \pi(1-\pi)$$

$$u_4 = \pi^2 \ .$$

Since the states of conditioning cannot be observed in the two-element model, a fortiori the asymptotic probabilities of these states cannot. However, for any trial n we may relate the probabilities of states of conditioning to the probability of an A_1 response, namely:

$$(1.3.27) \qquad P(\underline{A}_n = 1) = P_n(\{s_1,s_2\}) + \frac{1}{2} P_n(\{s_1\}) + \frac{1}{2} P_n(\{s_2\}) \quad .$$

Note why the coefficient $\frac{1}{2}$ occurs. If the subject is in state s_1, then if s_1 is sampled on trial n he will make response A_1, but if s_2 is sampled, he will make response A_2. And by assumption the probability of sampling s_1 is $\frac{1}{2}$. From (1.3.27) and the asymptotic probabilities of the states $\{s_1,s_2\}$, $\{s_1\}$ and $\{s_2\}$, the following asymptotic probability of an A_1 response is obtained:

$$\begin{aligned} \lim_{n \to \infty} P(\underline{A}_n = 1) &= u_4 + \frac{1}{2} u_2 + \frac{1}{2} u_3 \\ &= \pi^2 + \frac{1}{2} \pi(1 - \pi) + \frac{1}{2} \pi(1 - \pi) \\ &= \pi \quad . \end{aligned}$$

This result agrees with the asymptotic probability of an A_1 response in the one-element model. Moreover, it is not difficult to show that this asymptotic probability of response will be obtained on the assumption of any number r of stimulus elements and any sampling probability distribution $\omega_1, \ldots, \omega_r$.

In the next chapter we shall be concerned with the general question of goodness of fit of a model to observed data, but the fundamental character of this general question should not obscure the importance of certain particular quantities like the asymptotic probabilities of response. It is generally agreed that the first and most gross test of a model in the area of statistical learning theory is its ability to predict observed asymptotic response probabilities. From this gross test we may move in either two directions: toward further asymptotic results or toward analysis of the rate of learning. Let us begin with the former.

It is often charged that statistical learning theory correctly predicts only the average behavior of a group of subjects and not the behavior of individual subjects. This misconception rests on some rather widespread misunderstandings of probabilistic theories of behavior. Rather than speak in generalities, we may illustrate what we mean by considering the variance of the sum of the random variables $\underline{A}_{m+1}, \ldots, \underline{A}_{m+n}$, at asymptote. Following (1.2.8), this quantity is defined as

$$\overline{\underline{A}}_{m,N} = \sum_{n=m+1}^{m+N} \underline{A}_n \quad . \tag{1.3.28}$$

In a given experiment consisting, say, of 240 trials for each subject, we may compute $\overline{\underline{A}}_{m,N}(x)$ for each subject x over the last 100 trials, that is, we compute $\overline{\underline{A}}_{140,100}$. By averaging over subjects we obtain the expectation $E(\overline{\underline{A}}_{m,N})$, which is predicted to be $N\pi$. But from a given sample of subjects drawn from a homogeneous population, we may also compute the variance of $\overline{\underline{A}}_{m,N}$, which can then be compared with the predicted variance. Moreover, the same variance is predicted for the corresponding number of blocks of trials of length N for a given subject at asymptote. In either case we have a probabilistic prediction about individual subjects. For example, if $N = 10$, and we consider the last 100 trials for a given subject at asymptote, then we may directly compare the variance of the ten blocks of ten trials each with the theoretically computed variance of $\bar{\underline{A}}_{m,N}$.

We now turn to the derivation of this variance. For notational simplicity we drop the m subscript, particularly since at asymptote it does not matter at which trial we begin.

Theorem. In the one-element, non-contingent model the variance of $\bar{\underline{A}}_N$ at asymptote is:

$$\text{(1.3.29)} \quad \operatorname{Var}(\bar{\underline{A}}_N) = N\pi(1-\pi)\frac{(2-\theta)}{\theta} - \frac{2\pi(1-\pi)(1-\theta)[1-(1-\theta)^N]}{\theta^2}$$

Proof: By the classical theorem for the sum of random variables (see Feller [13], p. 216):

$$\text{(1)} \qquad \operatorname{var}(\bar{\underline{A}}_N) = \sum_{n=1}^{N} \operatorname{var}(\underline{A}_n) + 2 \sum_{1 \leq j < k \leq N} \operatorname{cov}(\underline{A}_j \underline{A}_k) \quad .$$

(Note that for simplicity we sum from $n = 1$; this does not affect the result at asymptote.)

Now since at asymptote

$$E(\underline{A}_n) = \pi ,$$

the variance is:

$$\mathrm{var}(\underline{A}_n) = \pi(1-\pi)$$

and

$$\mathrm{cov}(\underline{A}_j\underline{A}_k) = E(\underline{A}_j\underline{A}_k) - E(\underline{A}_j)E(\underline{A}_k)$$

$$= E(\underline{A}_j\underline{A}_k) - \pi^2 \quad .$$

Furthermore

$$E(\underline{A}_j\underline{A}_k) = E(\underline{A}_k \mid \underline{A}_j)E(\underline{A}_j)$$

$$= \pi E(\underline{A}_k \mid \underline{A}_j)$$

$$= \pi p_{11}^{(k-j)} \quad .$$

Now for the transition matrix of the one-element model we may prove by induction that

$$p_{11}^{(n)} = \pi + (1-\pi)(1-\theta)^n \ ,$$

whence, combining the above results

$$\mathrm{cov}(\underline{A}_j\underline{A}_k) = \pi[\pi + (1-\pi)(1-\theta)^{k-j}] - \pi^2$$

$$= \pi(1-\pi)(1-\theta)^{k-j} \quad .$$

Finally, we need the following summation:

$$\sum_{1\leq j<k\leq N} (1-\theta)^{k-j} = (1-\theta)^{2-1} + [(1-\theta)^{3-2} + (1-\theta)^{3-1}]$$

$$+ [(1-\theta)^{4-3} + (1-\theta)^{4-2} + (1-\theta)^{4-1}]$$

$$+\ldots+ [(1-\theta)+\ldots+(1-\theta)^{N-1}]$$

$$= \frac{1-(1-\theta)^2}{\theta}+\ldots+\frac{1-(1-\theta)^N}{\theta} - (N-1)$$

$$= [\frac{N-1}{\theta} - (N-1)] - \frac{1}{\theta}[(1-\theta)^2+\ldots+(1-\theta)^N]$$

$$= \frac{N-1}{\theta}(1-\theta) - \frac{(1-\theta)}{\theta^2}[1-\theta-(1-\theta)^N]$$

$$= \frac{N}{\theta}(1-\theta) - \frac{1-\theta}{\theta^2}[1-(1-\theta)^N] \quad .$$

Using this result and substituting in (1) we then have:

$$\mathrm{var}(\underline{\bar{A}}_N) = N\,\pi(1-\pi) + 2\,\pi(1-\pi)\,\frac{N}{\theta}(1-\theta) - 2\,\pi(1-\pi)\,\frac{(1-\theta)}{\theta^2}[1-(1-\theta)^N]$$

$$= N\pi(1-\pi)\,\frac{(2-\theta)}{\theta} - \frac{2\,\pi(1-\pi)(1-\theta)[1-(1-\theta)^N]}{\theta^2} \quad . \qquad \text{Q.E.D.}$$

Later, (1.3.29) will be compared with some observed data. A nearly endless number of further interesting asymptotic quantities can be presented, but we delay additional computations until we consider some data for the non-contingent case in Chapter 7.

We conclude this section with a discussion of the rate of learning for the one-element non-contingent model. The learning rate may be represented by the absolute probability of an A_1 response on trial n , which, using a standard notation, is $a_1^{(n)}$ defined by:

$$a_1^{(n)} = \sum_{i=1}^{2} a_i p_{i1}^{(n-1)} , \tag{1.3.30}$$

where a_i is the initial probability of response A_i , that is, on trial 1. */

We observed in the proof of (1.3.29) that

$$p_{11}^{(n)} = \pi + (1-\pi)(1-\theta)^n$$

and we may easily prove that

$$p_{21}^{(n)} = \pi - \pi(1-\theta)^n \quad .$$

The absolute probability $a_1^{(n)}$ is then:

$$\begin{aligned} a_1^{(n)} &= a_1[\pi + (1-\pi)(1-\theta)^{n-1}] + a_2[\pi - \pi(1-\theta)^{n-1}] \\ &= \pi - (\pi - a_1)(1-\theta)^{n-1} \quad . \end{aligned} \tag{1.3.31}$$

For $\pi > a_1$ and $0 < \theta < 1$, it is clear from (1.3.31) that $a_1^{(n)}$ is a strictly increasing function of n .

---

*/ For explicitness, note that

$$a_1 = P(\underline{A}_1 = 1)$$

$$a_1^{(n)} = P(\underline{A}_n = 1) \quad .$$

The analogues of (1.3.29) and (1.3.31) for the two-element model are somewhat cumbersome to derive for arbitrary θ and π. They will not be pursued here, but the analogues in the two-person situations will.

At this point we stop our theoretical analysis of the non-contingent case, but we return to it briefly in §1.6 and again in Chapter 7. In Chapter 8 the effects of strengthened motivation (money payoffs) is analyzed by introducing another concept, namely that of memory.

§1.4 Markov Model for Zero-Sum Two-Person Games. In the preceding section stimulus sampling learning theory was applied to the simplest one-person experimental case. We now want to apply it to one of the simpler two-person cases (cf. [5]). The general experimental situation may be described as follows. On a given trial each of the two players (i.e., subjects) independently makes one of two responses. As indicated in §1.2, the players are designated as A and B, with A making response A_1 or A_2, and B making response B_1 or B_2. The probability that a given response for a given player will be reinforced on a particular trial depends on the actual responses made on the trial. It is by virtue of this dependence or contingency that the game aspect of the experiment arises. For example, if player A makes response A_1, and player B response B_1, then there is a probability a_1 that A_1 is reinforced and B_2 is reinforced, and a probability $1 - a_1$ that

A_2 is reinforced and B_1 is reinforced.*/ Thus with probability a_1 it turns out that A made a correct response and B an incorrect one; with probability $1 - a_1$ the situation is reversed. The zero-sum character of the game is due to the fact that whichever one of the four response pairs (A_1,B_1), (A_1,B_2), (A_2,B_1), (A_2,B_2) actually occurs, exactly one player wins in the sense of having made a correct response. Thus the reinforcement probabilities may be described by the following payoff matrix:

(1.4.1)

	B_1	B_2
A_1	a_1	a_2
A_2	a_3	a_4

In thinking of each trial as the play of a simple 2×2 game, it is to be noted that the payoff is not being correct or incorrect, but the probability a_i of being correct.**/ The experimenter's selection of

---

*/ Using the E_i notation introduced in §1.2, there is a probability a_1 that an E_1 event occurs for player A and an E_2 event for player B and a probability $1 - a_1$ that an E_2 occurs for player A and an E_1 for player B.

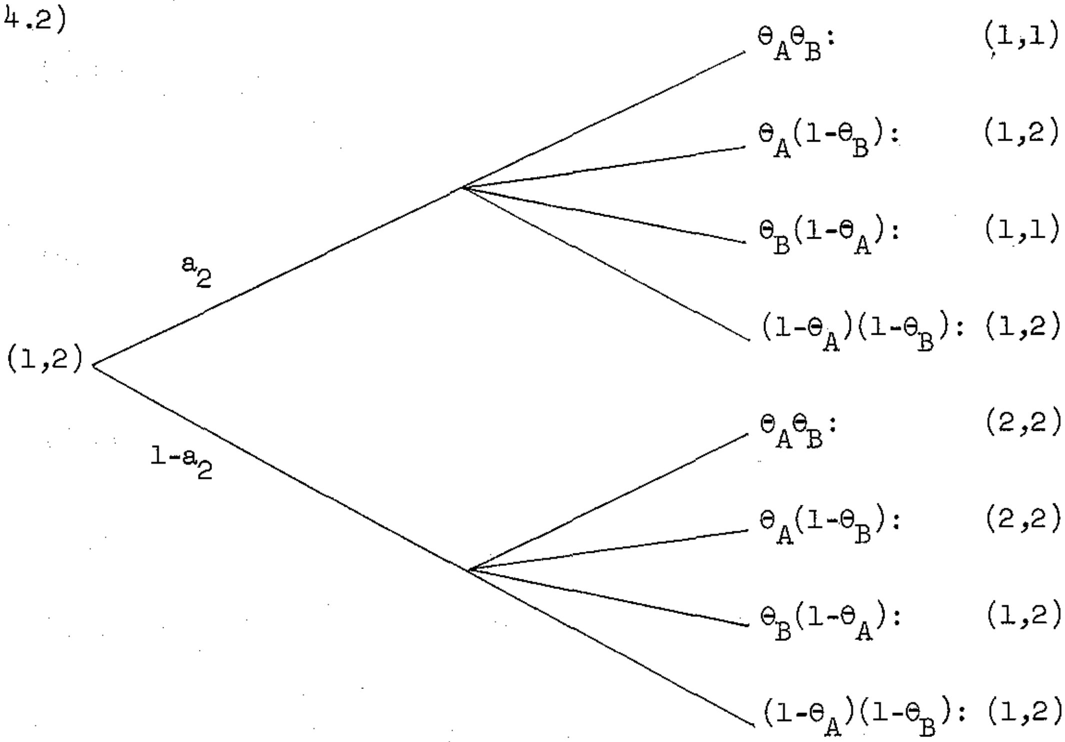
**/ Technically we then have a constant-sum game, which is strategically equivalent to the zero-sum game obtained by subtracting $1/2$ from each a_i .

actual reinforcements on the basis of the a_i's corresponds to a chance move by the referee in a game. Appropriate choice of a game strategy is in terms of the expected values of these chance moves.

Restricting ourselves to the assumption of one stimulus element for each player, the stipulation of (1.4.1) completely determines the derivation of the Markov process from the axioms of §1.2. We assume, of course, that both players satisfy the axioms. On each trial the single stimulus element of player A is conditioned to response A_1 or A_2, and the single element of player B is conditioned to response B_1 or B_2. There are, consequently, four possible states of conditioning, which will be represented by the four ordered pairs: (1,1), (1,2), (2,1), (2,2). The first member of each pair indicates the conditioning of player A's stimulus element, and the second member that of player B's. Thus the pair (1,2) represents the conditioning state defined by player A's element being conditioned to A_1 and player B's to B_2.

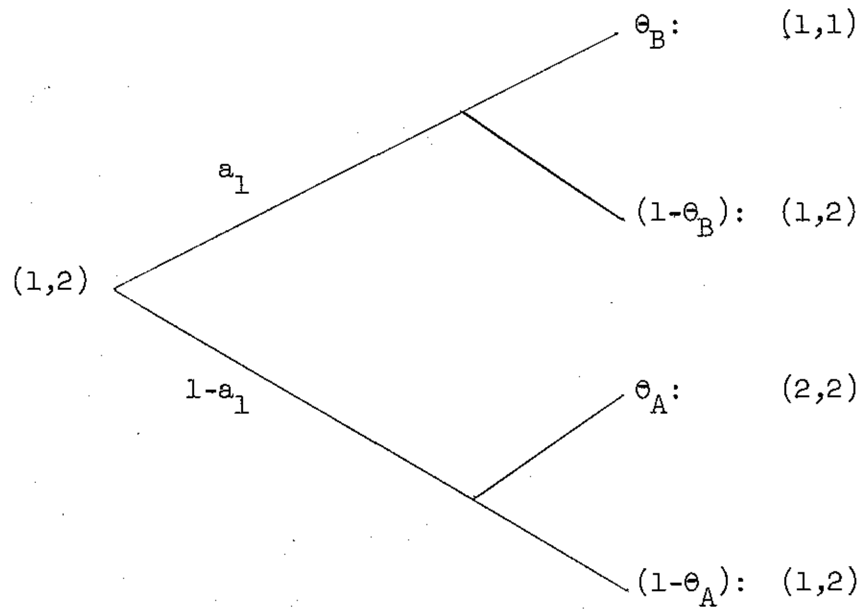
In view of the detailed derivations given in the preceding section, we restrict ourselves here to deriving only one of the four trees for the Markov process. Let us assume we are in the state (1,2). Then the tree looks like:

(1.4.2)



We want to make several remarks about this tree. First, in order to avoid a further multiplication of notation, the actual reinforcement and conditioning events have not been indicated on the tree, since these events are unequivocally specified by their probabilities. For example, a_2 on the upper main branch indicates that player A made a correct response and player B an incorrect one, that is, both players A and B received an E_1 reinforcing event. The probability $\theta_A\theta_B$ at the top indicates that conditioning was effective for both players. However, in the one-element model the effectiveness of conditioning is irrelevant when the response actually made is reinforced (a point already made in §1.3), and therefore by considering only relevant conditioning, (1.4.2) may be reduced as follows:

(1.4.3)



It should be clear from (1.4.2) and (1.4.3) that Θ_A is the probability of effective conditioning for player A, and Θ_B the corresponding probability for B . That is, using a superscript A or B to designate the random variable $\underline{F}_n$ for player A or B , we have as a generalization of (1.2.6):

$$\left\{ \begin{array}{l} P(\underline{F}_n^{(A)} = 1) = \Theta_A \\ \\ P(\underline{F}_n^{(B)} = 1) = \Theta_B \quad . \end{array} \right. \tag{1.4.4}$$

An important observation about (1.4.2) and (1.4.3) is that it is not possible in a single trial to go from the state (1,2) to the state (2,1). Analysis of the other three trees leads to similar conclusions: it is not possible in a single trial to go from (2,1) to (1,2), from (1,1) to (2,2), or from (2,2) to (1,1). This means that the anti-diagonal of the transition matrix must be identically zero. With this result we bring into the Markov process itself the kind of overly strong deterministic predictions which arise in the one-element stimulus model for the non-contingent case but not in the Markov process for that case, as represented by (1.3.7). In this respect then the situation is worse in the case under present discussion, for the Markov process itself yields deterministic predictions. As any experimenter would expect and as we shall see in Chapter 3, these predictions are not borne out by actual data. As in the case of the non-contingent case these difficulties may be met by passing from a one-element to a multi-element stimulus model for each player, the device which was also used for the simple contingent case in §1.3. The two-element model will be discussed in Chapter 3.

Construction of the other three trees like (1.4.3) yields the following transition matrix

$$
\begin{array}{c|cccc}
 & (1,1) & (1,2) & (2,1) & (2,2) \\
\hline
(1,1) & \begin{array}{l} a_1(\theta_A-\theta_B) \\ +\ (1-\theta_A) \end{array} & a_1\theta_B & (1-a_1)\theta_A & 0 \\
(1,2) & a_2\theta_B & \begin{array}{l} a_2(\theta_A-\theta_B) \\ +\ (1-\theta_A) \end{array} & 0 & (1-a_2)\theta_A \\
(2,1) & (1-a_3)\theta_A & 0 & \begin{array}{l} a_3(\theta_A-\theta_B) \\ +\ (1-\theta_A) \end{array} & a_3\theta_B \\
(2,2) & 0 & (1-a_4)\theta_A & a_4\theta_B & \begin{array}{l} a_4(\theta_A-\theta_B) \\ +\ (1-\theta_A) \end{array}
\end{array}
\tag{1.4.5}
$$

From the discussion in §1.3, it is clear that this matrix represents an irreducible, aperiodic case and thus the asymptotes exist and are independent of the initial probability distribution on the states. We are interested in establishing certain general conclusions about the asymptotic probabilities and consequently it is necessary to obtain the solution of (1.4.5). For this latter purpose we can, as easily as not, proceed by solving for the asymptotic probabilities of any four-state, irreducible and aperiodic Markov chain. To simplify notation the states will be numbered as follows: $1 = (1,1)$, $2 = (1,2)$, $3 = (2,1)$ and $4 = (2,2)$. Let (p_{ij}), for $i,j = 1,2,3,4$ be the transition matrix. Then we seek the numbers u_j such that

$$(1.4.6)\qquad \begin{aligned} u_j &= \sum_\nu u_\nu p_{\nu j} \ , \\ \sum u_j &= 1 \ . \end{aligned}$$

The general solution is given by */

$$(1.4.7)\qquad u_j = \frac{D_j}{D} \ , \qquad \text{for } \ j = 1, \ .., \ 4,$$

where

$$(1.4.8)\quad \begin{cases} D_1 = & [p_{21}\, p_{32}\, p_{43} + p_{31}\, p_{42}\, p_{23} + p_{41}(p_{22}-1)(p_{33}-1)] \\ & - [p_{41}\, p_{32}\, p_{23} + p_{21}\, p_{42}(p_{33}-1) + p_{31}(p_{22}-1)p_{43}] \\ D_2 = & -[(p_{11}-1)p_{32}\, p_{43} + p_{31}\, p_{42}\, p_{13} + p_{41}\, p_{12}(p_{33}-1)] \\ & + [p_{41}\, p_{32}\, p_{13} + (p_{11}-1)p_{42}(p_{33}-1) + p_{31}\, p_{12}\, p_{43}] \\ D_3 = & [(p_{11}-1)(p_{22}-1)p_{43} + p_{21}\, p_{42}\, p_{13} + p_{41}\, p_{12}\, p_{23}] \\ & - [p_{41}(p_{22}-1)p_{13} + (p_{11}-1)p_{42}\, p_{23} + p_{21}\, p_{12}\, p_{43}] \\ D_4 = & -[(p_{11}-1)(p_{22}-1)(p_{33}-1) + p_{21}\, p_{32}\, p_{13} + p_{31}\, p_{12}\, p_{23}] \\ & + [p_{31}(p_{22}-1)p_{13} + (p_{11}-1)p_{32}\, p_{23} + p_{21}\, p_{12}(p_{33}-1)] \\ D = & \sum D_j \ . \end{cases}$$

When the antidiagonal of (p_{ij}) is identically zero, as is the case with (1.4.5), the first four equations of (1.4.8) simplify to:

---

*/ We are indebted to Mr. Frank Krasne for making this computation.

$$(1.4.9)\begin{cases} D_1 = -[p_{21}\ p_{42}(p_{33}-1) + p_{31}(p_{22}-1)p_{43}] \\ D_2 = -p_{31}\ p_{42}\ p_{13} + [(p_{11}-1)p_{42}(p_{33}-1) + p_{31}\ p_{12}\ p_{43}] \\ D_3 = [(p_{11}-1)(p_{22}-1)p_{43} + p_{21}\ p_{42}\ p_{13}] - p_{21}\ p_{12}\ p_{43} \\ D_4 = -(p_{11}-1)(p_{22}-1)(p_{33}-1) + [p_{31}(p_{22}-1)p_{13} + p_{21}\ p_{12}(p_{33}-1)] \end{cases}$$

Applying (1.4.9) to (1.4.5) we obtain after some simplification:

$$(1.4.10)\begin{cases} D_1 = [a_2(1-a_3)(1-a_4) + (1-a_2)(1-a_3)a_4]\ \theta_A^2\theta_B + \\ \qquad [a_2a_3(1-a_4) + a_2(1-a_3)a_4]\ \theta_A\theta_B^2\ , \\ D_2 = [a_1(1-a_3)(1-a_4) + (1-a_1)a_3(1-a_4)]\ \theta_A^2\ \theta_B + \\ \qquad [a_1a_3(1-a_4) + a_1(1-a_3)a_4]\ \theta_A\theta_B^2\ , \\ D_3 = [(1-a_1)(1-a_2)a_4 + (1-a_1)a_2(1-a_4)]\ \theta_A^2\theta_B + \\ \qquad [(1-a_1)a_2a_4 + a_1(1-a_2)a_4]\ \theta_A\theta_B^2\ , \\ D_4 = [(1-a_1)(1-a_2)a_3 + a_1(1-a_2)(1-a_3)]\ \theta_A^2\theta_B + \\ \qquad [(1-a_1)a_2a_3 + a_1(1-a_2)a_3]\ \theta_A\theta_B^2\ . \end{cases}$$

Since D is the sum of the D_j's , and by virtue of (1.4.7)

$$u_j = \frac{D_j}{D}\ ,$$

we may infer from (1.4.10) that the asymptotic probability u_j of each state is a function only of the ratio θ_A/θ_B and the experimenter determined values a_i . To see this, first note that

$$u_j = \frac{c_j\theta_A^2\theta_B + d_j\theta_A\theta_B^2}{c\,\theta_A^2\theta_B + d\,\theta_A\theta_B^2} \tag{1.4.11}$$

where the coefficients c_j, d_j, c and d are functions of the a_i's given by equations (1.4.10), namely,

$$\left\{\begin{aligned}
c_1 &= a_2(1-a_3)(1-a_4) + (1-a_2)(1-a_3)a_4 \\
c_2 &= a_1(1-a_3)(1-a_4) + (1-a_1)a_3(1-a_4) \\
c_3 &= (1-a_1)(1-a_2)a_4 + (1-a_1)a_2(1-a_4) \\
c_4 &= (1-a_1)(1-a_2)a_3 + a_1(1-a_2)(1-a_3) \\
d_1 &= a_2a_3(1-a_4) + a_2(1-a_3)a_4 \\
d_2 &= a_1a_3(1-a_4) + a_1(1-a_3)a_4 \\
d_3 &= (1-a_1)a_2a_4 + a_1(1-a_2)a_4 \\
d_4 &= (1-a_1)a_2a_3 + a_1(1-a_2)a_3 \\
c &= c_1 + c_2 + c_3 + c_4 \\
d &= d_1 + d_2 + d_3 + d_4 \quad .
\end{aligned}\right. \tag{1.4.12}$$

Dividing numerator and denominator of (1.4.11) by $\theta_A\theta_B^2$ we obtain the desired result:

$$u_j = \frac{c_j(\theta_A/\theta_B) + d_j}{c(\theta_A/\theta_B) + d} \quad . \tag{1.4.13}$$

With (1.4.12) and (1.4.13) at hand it is a matter of elementary arithmetic to compute the asymptotic probabilities u_j for any fixed ratio θ_A/θ_B of the learning parameters. Moreover, for some experimental situations it is reasonable to expect that the rate of learning will be approximately the same for both players, and therefore to assume that

$$\theta_A = \theta_B \quad . \tag{1.4.14}$$

With this additional assumption a parameter-free prediction of asymptotic response probabilities can be made independent of, and prior to, any analysis of experimental data. The results of such predictions are reported in Chapter 3.

From (1.4.13) we can also draw some interesting conclusions about the relationship of the asymptotic response probabilities u_j to the ratio θ_A/θ_B . Setting $\rho = \theta_A/\theta_B$ and differentiating (1.4.13) with respect to ρ we obtain:

$$\frac{d}{d\rho} u_j = \frac{c_j d - c d_j}{(c\rho + d)^2} \quad . \tag{1.4.15}$$

If $c_j d \neq c d_j$ then u_j has no maximum for ρ in the open interval $(0,\infty)$, the permissible range of values for the ratio θ_A/θ_B . In fact, since the sign of the derivative is independent of ρ , u_j is either a monotonically decreasing or monotonically increasing function of ρ , strictly decreasing if $c_j d < c d_j$, and strictly increasing if $c_j d > c d_j$. Obviously both cases must obtain since

$$u_1 + u_2 + u_3 + u_4 = 1 \quad .$$

Moreover, because of the monotonicity of u_j in ρ, it is easy to compute the bounds of u_j from (1.4.13). Namely,

$$(1.4.16) \qquad \begin{cases} \lim\limits_{\rho \to \infty} u_j = \dfrac{c_j}{c.} \\ \\ \lim\limits_{\rho \to 0} u_j = \dfrac{d_j}{d} \end{cases} .$$

If u_j is an increasing function with respect to ρ, then $\frac{d_j}{d} < \frac{c_j}{c}$ and its values lie in the open interval $(\frac{d_j}{d}, \frac{c_j}{c})$; if decreasing, in the interval $(\frac{c_j}{c}, \frac{d_j}{d})$. Numerical values of these intervals are given in Chapter 3 for the sets of experimental parameters a_i actually used.

In connection with asymptotic response probabilities of players A and B we want to show that a certain linear relation obtains between $\alpha = \lim\limits_{n \to \infty} P(\underline{A}_n = 1)$ and $\beta = \lim\limits_{n \to \infty} P(\underline{B}_n = 1)$, which is independent of the ratio θ_A/θ_B.*/ (In §1.5 we show that this same linear relation obtains in the non-Markov linear model.) To begin with, note that

$$(1.4.17) \qquad \alpha = u_1 + u_2$$

since

$$P(\underline{A}_n = 1) = P(\underline{A}_n = 1, \underline{B}_n = 1) + P(\underline{A}_n = 1, \underline{B}_n = 0) ,$$

and correspondingly

$$(1.4.18) \qquad \beta = u_1 + u_3 .$$

*/ Note that γ as defined in (1.2.11) is simply u_1.

From (1.4.13), (1.4.17) and (1.4.18) it follows that

$$(c\rho + d)\alpha = (c_1 + c_2)\rho + d_1 + d_2$$

$$(c\rho + d)\beta = (c_1 + c_3)\rho + d_1 + d_3 \quad .$$

By elementary operations we may eliminate ρ from these two equations and obtain:

$$(c_1 + c_3 - c\beta)d\alpha - (c_1 + c_3 - c\beta)(d_1 + d_2) = (c_1 + c_2 - c\alpha)d\beta + (c_1 + c_2 - cd)(d_1 + d_3) \,.$$

The quadratic term $cd\alpha\beta$ cancels out and we have:

$$(c_1 d + c_3 d + cd_1 + cd_3)\alpha = (c_1 d + c_2 d + cd_1 + cd_2)\beta + (c_1 + c_2)(d_1 + d_3) + (c_1 + c_3)(d_1 + d_2) \quad . \tag{1.4.19}$$

In terms of the parameters a_i , we may derive from (1.4.12) and (1.4.19)

$$[(a_3 + a_4 - a_1 - a_2) + (a_1 a_2 - a_3 a_4)]\alpha = (a_1 a_3 - a_2 a_4)\beta + \tfrac{1}{2}(a_3 + a_4 - a_1 - a_2) + a_4(a_2 - a_3) \quad . \tag{1.4.20}$$

It would be pleasant to state that we have proved a theorem about the variance of the sum of A_1 or B_1 responses, corresponding to the variance theorem proved for the non-contingent case in the preceding section (see equation (1.3.29)). The analytical difficulties of proving such a theorem are not insurmountable, but the effort required to obtain the explicit higher transition probabilities $p_{ij}^{(n)}$ for (1.4.5) seemed considerably greater than the value of the result. The compromise we struck was to prove some relatively simple theorems about variances for arbitrary 4×4 irreducible, aperiodic Markov chains, and then proceed from these by numerical computation. The theorems are given in this section; the computations are reported in Chapters 3 and 4. The generalization of the theorems to $n \times n$ chains is obvious.

We first define the following random variables:

$$(1.4.21) \qquad \underline{X}_n(x) = \begin{cases} 1 & \text{if on trial } n \text{ state } 1 \text{ occurs,} \\ 0 & \text{otherwise.} \end{cases}$$

Thus, with reference to (1.4.5) if $\underline{X}_n = 1$ then state (1,1) occurred on trial n, that is, responses A_1 and B_1 were made. Note that here the proper interpretation of the sample space point x is as a <u>pair</u> of subjects, players A and B, not as a single subject.

$$(1.4.22) \qquad \underline{Y}_n(x) = \begin{cases} 1 & \text{if on trial } n \text{ state } 2 \text{ occurs,} \\ 0 & \text{otherwise.} \end{cases}$$

Again with reference to (1.4.5), state 2 corresponds to (1,2), that is, to the pair of responses A_1 and B_2.

$$(1.4.23) \qquad \underline{Z}_n(x) = \begin{cases} 1 & \text{if on trial } n \text{ state } 3 \text{ occurs,} \\ 0 & \text{otherwise.} \end{cases}$$

With respect to (1.4.5), state 3 corresponds to the state (2,1), that is, to the pair of responses A_2 and B_1.

Analogous to (1.4.17) and (1.4.18) we have the following two functional equalities for $\underline{A}_n$ and $\underline{B}_n$ in terms of $\underline{X}_n$, $\underline{Y}_n$ and $\underline{Z}_n$.

$$\begin{aligned} \underline{A}_n &= \underline{X}_n + \underline{Y}_n \\ \underline{B}_n &= \underline{X}_n + \underline{Z}_n \end{aligned} \qquad (1.4.24)$$

For the sum of any of these five random variables over a block of N trials we use the notation already introduced for $\underline{A}_n$ and $\underline{B}_n$, namely $\overline{\underline{X}}_N$, $\overline{\underline{Y}}_N$, $\overline{\underline{Z}}_N$, $\overline{\underline{A}}_N$ and $\overline{\underline{B}}_N$. Also, from previous notation the asymptotic probabilities of $\underline{X}_n$, $\underline{Y}_n$, $\underline{Z}_n$, $\underline{A}_n$ and $\underline{B}_n$ are u_1, u_2, u_3, α and β respectively.

The theorem with which we conclude this section is then:

Theorem. At asymptote we have the following variances for the sum of N random variables:

$$\operatorname{var}(\overline{\underline{X}}_N) = Nu_1(1 - Nu_1) + 2u_1 \sum_{j=1}^{N-1} (N-j)p_{11}^{(j)} , \qquad (1.4.25)$$

$$\operatorname{var}(\overline{\underline{Y}}_N) = Nu_2(1 - Nu_2) + 2u_2 \sum_{j=1}^{N-1} (N-j)p_{22}^{(j)} , \qquad (1.4.26)$$

$$\operatorname{var}(\overline{\underline{Z}}_N) = Nu_3(1 - Nu_3) + 2u_3 \sum_{j=1}^{N-1} (N-j)p_{33}^{(j)} , \qquad (1.4.27)$$

$$\begin{aligned} \operatorname{var}(\overline{\underline{A}}_N) = {} & \operatorname{var}(\overline{\underline{X}}_N) + \operatorname{var}(\overline{\underline{Y}}_N) + 2u_1 \sum_{j=1}^{N-1} (N-j)p_{12}^{(j)} \\ & + 2u_2 \sum_{j=1}^{N-1} (N-j)p_{21}^{(j)} - 2N^2u_1u_2 , \end{aligned} \qquad (1.4.28)$$

$$(1.4.29) \qquad \mathrm{var}(\overline{\underline{B}}_N) = \mathrm{var}(\overline{\underline{X}}_N) + \mathrm{var}(\overline{\underline{Z}}) + 2u_1 \sum_{j=1}^{N-1} (N-j)p_{13}^{(j)}$$

$$+ 2u_3 \sum_{j=1}^{N-1} (N-j)p_{31}^{(j)} - 2N^2 u_1 u_3 \quad .$$

<u>Proof</u>: We prove only (1.4.25) and (1.4.28); proofs of the other three cases are similar.

By the fundamental theorem for the variance of a sum of random variables (see Feller [14], p. 216), at asymptote

$$(1) \qquad \mathrm{var}(\overline{\underline{X}}_N) = N\ \mathrm{var}(\underline{X}_n) + 2 \sum_{1 \leq j < k \leq N} \mathrm{cov}(\underline{X}_j \underline{X}_k) \quad .$$

Now at asymptote

$$(2) \qquad \mathrm{var}(\underline{X}_n) = u_1(1 - u_1)$$

and

$$(3) \qquad \mathrm{cov}(\underline{X}_j \underline{X}_k) = E(\underline{X}_j \underline{X}_k) - E(\underline{X}_j)E(\underline{X}_k)$$

$$= E(\underline{X}_j \underline{X}_k) - u_1^2 \quad .$$

Moreover,

$$(4) \qquad E(\underline{X}_j \underline{X}_k) = E(\underline{X}_k \mid \underline{X}_j)E(\underline{X}_j) \qquad \text{for} \quad j < k$$

$$= u_1\ E(\underline{X}_k \mid \underline{X}_j)$$

$$= u_1 p_{11}^{(k-j)} \quad .$$

We need to evaluate one sum:

$$(5) \quad \sum_{1 \leq j < k \leq N} p_{11}^{(k-j)} = p_{11}^{(2-1)} + [p_{11}^{(3-2)} + p_{11}^{(3-1)}] +$$

$$[p_{11}^{(4-3)} + p_{11}^{(4-2)} + p_{11}^{(4-1)}] + \dots +$$

$$[p_{11}^{(N-(N-1))} + \dots + p_{11}^{(N-1)}]$$

$$= (N-1)p_{11}^{(1)} + (N-2)p_{11}^{(2)} + \dots + p_{11}^{(N-1)}$$

$$= \sum_{j=1}^{N-1} (N-j)p_{11}^{(j)} \quad .$$

Substituting (4) into (3) and applying the summation result (5), we infer that

$$(6) \quad 2 \sum_{1 \leq j < k \leq N} \operatorname{cov}(\underline{X}_j \underline{X}_k) = 2u_1 \sum_{j=1}^{N-1} (N-j)p_{11}^{(j)} - 2 \sum_{1 \leq j \leq k \leq N} u_1^2$$

$$= 2u_1 \sum_{j=1}^{N-1} (N-j)p_{11}^{(j)} - N(N-1)u_1^2 \quad .$$

Finally combining the last term on the right of (6) and (2) multiplied by N , we have:

$$(7) \quad Nu_1(1-u_1) - N(N-1)u_1^2 = Nu_1(1-Nu_1) \quad .$$

The term on the right of (7) and the first term on the right of (6) yield the desired result for $\operatorname{var}(\overline{\underline{X}}_N)$.

We turn now to (1.4.28). By virtue of (1.4.24) we have immediately:

$$(8)\qquad \mathrm{var}(\bar{\underline{A}}_N) = \mathrm{var}(\bar{\underline{X}}_N + \bar{\underline{Y}}_N)$$

$$= \mathrm{var}(\bar{\underline{X}}_N) + \mathrm{var}(\bar{\underline{Y}}_N) + 2\ \mathrm{cov}(\bar{\underline{X}}_N \bar{\underline{Y}}_N) \quad .$$

Now as before

$$(9)\qquad \mathrm{cov}(\bar{\underline{X}}_N \bar{\underline{Y}}_N) = E(\bar{\underline{X}}_N \bar{\underline{Y}}_N) - E(\bar{\underline{X}}_N)E(\bar{\underline{Y}}_N)$$

$$= E(\bar{\underline{X}}_N \bar{\underline{Y}}_N) - N^2 u_1 u_2 \quad .$$

We note that the product $\bar{\underline{X}}_N \bar{\underline{Y}}_N$ is a product of sums and for $1 \leq n \leq N$

$$E(\underline{X}_n \underline{Y}_n) = 0 \ ,$$

whence

$$(10)\quad E(\bar{\underline{X}}_N \bar{\underline{Y}}_N) = E(\underline{X}_1 \cdot \sum_{n \neq 1} \underline{Y}_n) + E(\underline{X}_2 \cdot \sum_{n \neq 2} \underline{Y}_n) + \ldots + E(\underline{X}_N \cdot \sum_{n \neq N} \underline{Y}_n)$$

$$= u_1 \sum_{j=2}^{N} p_{12}^{(j-1)} + [u_2 p_{21}^{(1)} + u_1 \sum_{j=3}^{N} p_{12}^{(j-2)}] +$$

$$[u_2 \sum_{j=1}^{2} p_{21}^{(j)} + u_1 \sum_{j=4}^{N} p_{12}^{(j-3)} + \ldots +$$

$$[u_2 \sum_{j=1}^{N-2} p_{21}^{(j)} + u_1 p_{12}^{(1)}] + u_2 \sum_{j=1}^{N-1} p_{21}^{(j)} \quad .$$

$$= u_1[(N-1)p_{12}^{(1)} + (N-2)p_{12}^{(2)} + \ldots + p_{12}^{(N-1)}] +$$

$$u_2[(N-1)p_{21}^{(1)} + (N-2)p_{21}^{(2)} + \ldots + p_{21}^{(N-1)}]$$

$$= u_1 \sum_{j=1}^{N-1} (N-j)p_{12}^{(j)} + u_2 \sum_{j=1}^{N-1} (N-j)p_{21}^{(j)} \quad .$$

Substituting the results of (9) and (10) into (8) yields the desired result. Q.E.D.

In comparing these variances with empirical data in Chapters 3 and 4 we shall divide them by N^2 to normalize to the relative frequencies of responses in N trials.

Models for more complicated experimental situations are presented in later chapters. And in Chapter 3 the two-stimulus-element model is discussed for the simple zero-sum game situation which has been the focus of this section. However, all of these models are derived on the basis of the fundamental axioms of §1.2.

§1.5 Alternative Linear Model. For those experiments in which the available stimuli are the same on all trials it is possible to use a model which dispenses with the concept of stimuli. In such a "pure" reinforcement model the only assumption is that the probability of a response on a given trial is a linear function of the probability of that response on the previous trial. A one-person experiment may be represented simply as a sequence $(\underline{A}_1, \underline{E}_1, \underline{A}_2, \underline{E}_2, \ldots, \underline{A}_n, \underline{E}_n, \ldots)$ of the response and reinforcement random variables defined by (1.2.1) and (1.2.2). Any sequence of values of these random variables represents a possible experimental outcome. (For analysis of an experiment in which more than two responses or reinforcements are possible, (1.2.1) and (1.2.2) need to be modified so that the value of the random variable $\underline{A}_n$ is a number j representing the response on trial n, and the value of $\underline{E}_n$ is a number k representing the reinforcing event on trial n. However, this modification is not necessary for purposes of this section.)

The linear theory is formulated for the probability of a response on trial $n+1$, _given_ the entire preceding sequence of responses and reinforcements.*/ For this preceding sequence we use the notation x_n. Thus, x_n is a sequence of length $2n$ with 0's and 1's in the odd positions indicating responses A_1 and A_2 and 1's and 2's in the even positions indicating reinforcing events E_1 and E_2. The axioms of the linear theory are as follows:

Axiom L1. _If_ $\underline{E}_n = 1$ _and_ $P(x_n) > 0$ _then_

$$P(\underline{A}_{n+1} = 1 \mid x_n) = (1-\theta)P(\underline{A}_n = 1 \mid x_{n-1}) + \theta \quad .$$

Axiom L2. _If_ $\underline{E}_n = 2$ _and_ $P(x_n) > 0$ _then_

$$P(A_{n+1} = 1 \mid x_n) = (1-\theta)P(A_n = 1 \mid x_{n-1}) \quad .$$

*/ In the language of stochastic processes, this means we have replaced the Markov chains of earlier sections by chains of infinite order.

Here, as usual, θ is to be thought of as the learning parameter.

For the non-contingent case of §1.3 we can derive from L1 and L2 the same asymptotic mean result as for the Markov model, namely,

$$\lim_{n\to\infty} P(\underline{A}_n = 1) = \pi \quad . \tag{1.5.1}$$

On the other hand, the expression for the variance of the Cesàro sum $\overline{\underline{A}}_N$ given by (1.3.29) is different for the linear model. We shall not derive it here, but in Estes and Suppes [12] it is shown to be:

$$\mathrm{var}(\overline{\underline{A}}_N) = \frac{\pi(1-\pi)}{(2-\theta)\theta}\{N\theta(4-3\theta) - 2(1-\theta)[1-(1-\theta)^N]\} \quad . \tag{1.5.2}$$

The following interesting result obtains.

Theorem. *In the non-contingent case, for every* $N \geq 2$ *and for every* θ *in the open interval* $(0,1)$, *the variance of* $\underline{A}_N$ *at asymptote is less in the linear model than in the one-element Markov model.*

Proof: We seek the conditions on N and θ under which the following inequality holds ($\pi(1-\pi)/\theta$ has been cancelled from both (1.3.29) and (1.5.2)):

$$\frac{N\theta(4-3\theta)}{2-\theta} - \frac{2(1-\theta)[1-(1-\theta)^N]}{2-\theta} < N(2-\theta) - \frac{2(1-\theta)[1-(1-\theta)^N]}{\theta}$$

This simplifies to:

$$1-(1-\theta)^N < N\theta \quad . \tag{1}$$

Now

$$(2) \qquad 1-(1-\theta)^N = \theta[1+(1-\theta)+(1-\theta)^2+\ldots+(1-\theta)^{N-1}] \ ,$$

whence from (1) and (2), cancelling θ , which by hypothesis is not 0 , we have:

$$1+(1-\theta)+(1-\theta)^2+\ldots+(1-\theta)^{N-1} < N \ ,$$

and clearly this strict inequality holds under the conditions of the hypothesis of the theorem. Q.E.D.

To extend the linear model to two-person situations, we assume that both subjects satisfy Axioms L1 and L2. For the study of higher moments, which will not be considered here (see Lamperti and Suppes [18]), we also need the assumption that, given the sequence of past responses and reinforcements, the probabilities of responses of the two players on trial n are statistically independent. It is shown in Estes and Suppes [12] that the following recursive equations hold for the two-person zero-sum situation defined in §1.4 (we use here the notation of (1.2.9)):

$$(1.5.3) \quad \begin{cases} \alpha_{n+1} = [1-\theta_A(2-a_2-a_4)]\alpha_n + \theta_A(a_4-a_3)\beta_n + \theta_A(1-a_4) \\ \qquad\qquad + \theta_A(a_1+a_3-a_2-a_4)\gamma_n \\ \beta_{n+1} = [1-\theta_B(a_3+a_4)]\beta_n + \theta_B(a_2-a_4)\alpha_n + \theta_B a_4 \\ \qquad\qquad + \theta_B(a_3+a_4-a_1-a_2)\gamma_n \ . \end{cases}$$

It is further shown in Lamperti and Suppes [18] that the limits α, β and γ exist, whence we obtain from (1.5.3) two linear relations, which are independent of θ_A and θ_B :

$$(1.5.4)\quad \begin{cases} (2 - a_2 - a_4)\alpha = (a_4 - a_3)\beta + (a_1 + a_3 - a_2 - a_4)\gamma + (1 - a_4) \\ (a_3 + a_4)\beta = (a_2 - a_4)\alpha + (a_3 + a_4 - a_1 - a_2)\gamma + a_4 \ . \end{cases}$$

By eliminating γ from these two equations we obtain the linear relation (1.4.20) in α and β . Unfortunately, this linear relationship represents one of the few quantitative results which can be directly computed when this model is applied to multiperson situations. Our relative neglect of the linear model in the sequel is due mainly to its mathematical intractability in comparison with the Markov models already discussed ([5]).

§1.6 Comparisons with Game Theory. As remarked in the first section, it is possible to view game theory as a descriptive, empirical theory of behavior, but in fact, this does not seem to be a very promising approach. Our *a posteriori* reason is that for our own experiments it did not make good predictions. It seems to us that there are several general reasons why one should not be surprised by the poor predictive success of game theory. In the classical sense of psychological theories, game theory is not a behavior theory. It does not provide an analysis of how the organism interacts with its environment, that is, of the way in which the organism receives cues or stimuli from its environment and then adjusts its behavior accordingly. Another way of stating the matter is that game theory does not provide, even in schematic form, a formulation of the elementary process which would lead an organism to select the appropriate game-theoretic choice of a strategy.

From a general methodological standpoint, the orientation of game theory is that of classical economics: the concern is with what should be the behavior of a rational man. This concern with the rational man is the basis of another strong bond between classical economics and game theory; namely, both are very much theories of equilibrium. The derivation of an equilibrium condition does not, in these disciplines, depend on any assumptions about the particular dynamic process by which the equilibrium point is reached. This static character no doubt accounts for the uneasiness with which many psychologists view the concept of utility. The economist and game theorist take the utility function of the individual consumer or player as given, whereas a psychologist immediately tends to inquire where the utility function came from and to seek the environmental factors controlling the process of acquisition of a particular set of values or utilities. (We do not imply by this last sentence that psychologists are yet able to propound a theory which will account in any detail for the actual preferences, tastes and values of organisms.)

Granted that game theory is an equilibrium theory, it may still be argued that it has predictive pertinence to our experiments. For it can be maintained that when individuals have reached what, from the learning standpoint, is described as an asymptotic level of behavior, then they will be in equilibrium with their environment (including the other players in a game situation) and the optimality concepts of game theory may well apply to their patterns of choice. For example, even when subjects are not shown the pay-off matrix, after a large number of trials they may have learned enough about the prospects of winning and losing to approximate an optimal game strategy.

For ready reference in reporting our experimental findings with regard to game theory, we define here the three concepts of optimality which later will be used for comparison with learning theory predictions.*/

In the first place, when it is applicable, the appealing sure-thing principle may be used to select an optimal strategy. A strategy satisfies the sure-thing principle if, no matter what your opponent does, you are at least as well off, and possibly better off, with this strategy in comparison to any other available to you. For example, consider the following 2×2 matrix of a two-person, zero-sum game.

---

*/ For a detailed presentation of game theory, see McKinsey [20], Luce and Raiffa [19], or Blackwell and Girshick [6].

	B_1	B_2
A_1	5	7
A_2	2	-1

On each play of the game player A chooses row A_1 or A_2 and player B chooses column B_1 or B_2. If A_1 and B_1 are chosen then A receives \$5.00 and B loses this amount. If A_2 and B_2 are chosen, A loses \$1.00 and B receives this amount, and similarly for the other two combinations of strategies. It is obvious that the choice of A_1 by player A is the selection of a strategy satisfying the sure-thing principle, for no matter what B does, A is better off with A_1 than A_2. However, the weakness of the sure-thing principle is exemplified by B's situation. Neither B_1 nor B_2 satisfies the sure-thing principle, and so, in his choice of a strategy, B is not in a position to apply this principle of optimality.

Intuitively, it seems clear what B should do, namely, always select B , since he will only lose \$5.00 rather than \$7.00 when A chooses A_1. The optimality principle which covers the selection of B_1 is von Neumann's famous minimax principle [26]. The idea is that B should minimize his maximum loss. If he chooses B_2 his maximum loss is \$7.00, and if he chooses B , it is \$5.00. So he should minimize this loss and choose B_1.

When the sure-thing principle applies, it agrees with the minimax principle. Thus, if A minimizes his maximum loss he will pick A_1. This is most easily seen by maximizing his minimum gain, which amounts to the same thing. If A picks A_1 his minimum gain is \$5.00 and if he picks A_2 his minimum "gain" is the loss of \$1.00. Thus to maximize his minimum gain he should pick A_1.

Let (a_{ij}), with $i = 1, \ldots, n$ and $j = 1, \ldots, m$, be the payoff matrix of an $n \times m$ zero-sum game. If

$$\max_i \min_j a_{ij} = \min_j \max_i a_{ij} = v ,$$

we call v the value of the game and any strategies i^* and j^* such that

$$\min_j a_{i^*j} = \max_i \min_j a_{ij}$$

and

$$\max_i a_{ij^*} = \min_j \max_i a_{ij}$$

are _pure_ minimax strategies, i^* for A and j^* for B. A _pure_ strategy is one which selects a given row or column with probability one.

Unfortunately, pure minimax strategies do not always exist. For instance, there are none for the following payoff matrix:

(1.6.1)

	B_1	B_2
A_1	-1	3
A_2	0	-2

because

$$\max_i \min_j \; a_{ij} = -1$$

but

$$\min_j \max_i \; a_{ij} = 0 \quad .$$

The implication of this situation is that in repeated plays of this game a fixed choice of A_1 or A_2, by A, or a fixed choice of B_1 or B_2, by B, will not be optimal against an intelligent opponent.

The insufficiency of pure strategies may be remedied by introducing probability <u>mixtures</u> of pure strategies. For instance, player A might choose A_1 with probability $1/3$ and A_2 with probability $2/3$. A probability mixture is called a <u>mixed strategy</u>. Such a strategy for A may be designated $\xi = (\xi_1, \xi_2)$, where ξ_i is the probability of choosing A_i, for $i = 1,2$.*/ Similarly, mixed strategies for B are designated $\eta = (\eta_1, \eta_2)$. The fundamental theorem of von Neumann

---

*/ If n strategies are available to A, then $\xi = (\xi_1, \dots, \xi_n)$ is an n-dimensional vector such that $\xi_i \geq 0$ for $i = 1, \dots, n$ and $\sum \xi_i = 1$.

is that <u>mixed</u> minimax strategies exist for any zero-sum, two-person game with a finite number of strategies available to each player. In other words, there are probability mixtures ξ^* and η^* such that

$$(1.6.2) \qquad \max_{\xi} \sum_{i,j} a_{ij}\, \xi_i\, \eta_j^* = \min_{\eta} \sum_{i,j} a_{ij}\, \xi_i^*\, \eta_j = v ,$$

and v is called the <u>value</u> of the game.*/ What (1.6.2) shows is that player A may assure himself of winning at least v by playing ξ^* and player B may assure himself of losing at most v by playing η^* .

It is not appropriate to discuss here general methods of finding the value of a game and its minimax strategies, but we can illustrate the simple technique for 2×2 games with payoff matrix (a_{ij}) . For simplicity, let $x = \xi_1$ and $y = \eta_1$. Then it may be shown that it is sufficient to consider x and y separately against the use of pure strategies by the other player. Thus we seek numbers x, y and v such that

$$(1.6.3) \qquad \begin{cases} x a_{11} + (1-x) a_{21} \geq v \\ x a_{12} + (1-x) a_{22} \geq v \\ y a_{11} + (1-y) a_{12} \leq v \\ y a_{21} + (1-y) a_{22} \leq v \\ 0 \leq x, y \leq 1 \quad . \end{cases}$$

---

*/ Note that $\sum_{i,j} a_{ij}\, \xi_i\, \eta_j$ is just the expectation of A's gain with respect to the two independent probability mixtures ξ and η .

In any numerical case solution of these inequalities is a simple matter. For (1.6.1), we obtain: $x = 1/3$, $y = 5/6$, $v = -1/3$. (In example (1.6.1), x and y are unique. This is not always the case; as a trivial instance, if $a_{ij} = 0$ for i and j , any ξ and any η is a mixed minimax strategy.)

The development of an adequate theory of optimal strategies for non-zero-sum, two-person games is a complicated matter which as yet does not have a satisfactory solution. (Recall that a two-person game is non-zero-sum when what one player receives at the end of a play is not the negative of what the other player receives.) A natural division of non-zero-sum games is into <u>cooperative</u> and <u>non-cooperative</u> games. In a cooperative game the players are permitted to communicate and bargain before selecting a strategy; in a non-cooperative game no such communication and bargaining is permitted.

Subsequent chapters devoted to non-zero-sum games are entirely concerned with those of the non-cooperative type. Probably the best concept of optimality yet proposed for such games is Nash's notion of an equilibrium point ([21], [22]). An <u>equilibrium point</u> is a set of strategies, one for each player, with the property that these strategies provide a way of playing the game such that if all the players but one follow their given strategies, the remaining player cannot do better by following any strategy other than one belonging to the equilibrium point. It was shown by Nash that every non-zero-sum, n-person game, in which each player has a finite number of strategies, has an equilibrium point among its mixed strategies. Consideration of techniques for finding the equilibrium point will be delayed until Chapter 4.

References

1. Anderson, N.H. & Grant, D.A. A test of a statistical learning theory model for two-choice behavior with double stimulus events. J. exp. Psychol., 1957, 54, 305-316.

2. Atkinson, R.C. An analysis of the effect of non-reinforced trials in terms of statistical learning theory. J. exp. Psychol., 1956, 52, 28-32.

3. Atkinson, R.C. A stochastic model for rote serial learning. Psychometrika, 1957, 22, 87-96.

4. Atkinson, R.C. A Markov model for discrimination learning. Psychometrika, 1958, 23, 309-322.

5. Atkinson, R.C. & Suppes P. An analysis of two-person game situations in terms of statistical learning theory. J. exp. Psychol., 1958, 55, 369-378.

6. Blackwell, D., & Girshick, M.A. Theory of games and statistical decisions. New York: Wiley, 1954.

7. Burke, C.J., and Estes, W.K. A component model for stimulus variables in discrimination learning. Psychometrika, 1957, 22, 133-145.

8. Bush, R.R., & Mosteller, F. Stochastic models for learning. New York: Wiley, 1955.

9. Estes, W.K. Toward a statistical theory of learning. Psychol. Rev., 1950, 57, 94-107.

10. Estes, W.K. Theory of learning with constant, variable, or contingent probabilities of reinforcement. Psychometrika, 1957, 22, 113-132.

11. Estes, W.K., & Burke, C.J. A theory of stimulus variability in learning. Psychol. Rev., 1953, 60, 276-286.

12. Estes, W.K. & Suppes, P. Foundations of statistical learning theory, I. The linear model for simple learning. Technical Report No. 16, Contract Nonr 225(17), Applied Mathematics and Statistics Laboratory, Stanford University, 1957.

13. Estes, W.K. & Suppes, P. Foundations of statistical learning theory, II. The stimulus sampling model for simple learning. Forthcoming as a technical report.

14. Feller, W. An introduction to probability and its applications, second edition. New York: Wiley, 1957.

15. Fréchet, M. Recherches théoriques modernes sur le calcul des probabilités, vol. 2, Paris, 1938.

16. Kemeny, J.G., & Snell, J.L. Finite Markov chains. Princeton: Van Nostrand, to be published in 1959.

17. LaBerge, D. A model with neutral elements. Ch.2, Studies in Mathematical Learning Theory, edited by R.R. Bush and W.K. Estes, Stanford University Press, to be published in 1959.

18. Lamperti, J. & Suppes, P. Chains of infinite order and their application to learning theory. Technical Report No. 18, Contract Nonr 225(17), Applied Mathematics and Statistics Laboratory, Stanford University, 1958.

19. Luce, R.D. & Raiffa, H. Games and decisions. New York: Wiley, 1957.

20. McKinsey, J.C.C. Introduction to the theory of games. New York: McGraw-Hill, 1952.

21. Nash, J.F. The bargaining problem. Econometrica, 1950, 18, 155-162.

22. Nash, J.F. Non-cooperative games. Annals of Mathematics, 1951, 54, 286-295.

23. Neimark, E.D. Effects of type of non-reinforcement and number of alternative responses in two verbal conditioning situations. J. exp. Psychol., 1956, 52, 209-220.

24. Restle, F. A theory of discrimination learning. Psychol. Rev., 1955, 62, 11-19.

25. Restle, F. A theory of selective learning with probable reinforcements. Psychol. Rev., 1957, 64, 182-191.

26. Von Neumann, J. & Morgenstern, O. Theory of games and economic behavior, second edition. Princeton University Press, Princeton, 1947.

STANFORD UNIVERSITY

Technical Reports Distribution List

Contract Nonr 225(17)

(NR 171-034)

Armed Services Technical
Information Agency
Arlington Hall Station
Arlington 12, Virginia 5

Commanding Officer
Office of Naval Research
Branch Office
Navy No. 100, Fleet Post Office
New York, New York 35

Director, Naval Research
Laboratory
Attn: Technical Information
Officer
Washington 25, D.C. 6

Office of Naval Research
Group Psychology Branch
Code 452
Department of the Navy
Washington 25, D.C. 5

Office of Naval Research
Branch Office
346 Broadway
New York 13, New York 1

Office of Naval Research
Branch Office
1000 Geary Street
San Francisco 9, Calif. 1

Office of Naval Research
Branch Office
1030 Green Street
Pasadena 1, California 1

Office of Naval Research
Branch Office
The John Crerar Library Building
86 East Randolph Street
Chicago 1, Illinois 1

Office of Naval Research
Logistics Branch, Code 436
Department of the Navy
Washington 25, D.C. 1

Office of Naval Research
Mathematics Division, Code 430
Department of the Navy
Washington 25, D.C. 1

Operations Research Office
6935 Arlington Road
Bethesda 14, Maryland
Attn: The Library 1

Office of Technical Services
Department of Commerce
Washington 25, D.C. 1

The Logistics Research Project
The George Washington University
707 - 22nd Street, N.W.
Washington 7, D.C. 1

The RAND Corporation
1700 Main Street
Santa Monica, California
Attn: Dr. John Kennedy 1

Library
Cowles Foundation for Research
in Economics
Box 2125, Yale Station
New Haven, Connecticut 1

Center for Philosophy of Science
University of Minnesota
Minneapolis 14, Minnesota 1

Institut für Math. Logik
Universitat
Schlossplatz 2
Münster in Westfalen
Germany 1

Professor Ernest Adams
Department of Philosophy
University of California
Berkeley 4, California 1

Professor Maurice Allais
15 Rue des Gates-Ceps
Saint-Cloud (S.-O.)
France 1

Professor Alan Ross Anderson
Department of Philosophy
Yale University
New Haven, Connecticut 1

Professor Norman H. Anderson
Department of Psychology
University of California
Los Angeles 24, California 1

Professor T. W. Anderson
Department of Statistics
Columbia University
New York 27, New York 1

Professor K. J. Arrow
Serra House
Stanford University 1

Professor Richard C. Atkinson
Department of Psychology
University of California
Los Angeles 24, California 1

Dr. R.J. Audley
Department of Psychology
University College
London WC 1, England 1

Dr. R. F. Bales
Department of Social Relations
Harvard University
Cambridge, Massachusetts 1

Professor Y. Bar-Hillel
Dept. for History and Philosophy of Science
Hebrew University
Jerusalem, Israel 1

Professor Alex Bavelas
Department of Psychology
Stanford University 1

Professor Gustav Bergman
Department of Philosophy
State University of Iowa
Iowa City, Iowa 1

Professor E.W. Beth
Bern, Zweerskade 23, I
Amsterdam, Z.,
The Netherlands 1

Professor Max Black
Department of Philosophy
Cornell University
Ithaca, New York 1

Professor David Blackwell
Department of Statistics
University of California
Berkeley 4, California 1

Professor Lyle E. Bourne, Jr.
Department of Psychology
University of Utah
Salt Lake City, Utah 1

Mr. Gordon Bower
Department of Psychology
Yale University
New Haven, Connecticut 1

Professor R. B. Braithwaite
King's College
Cambridge, England 1

Dr. John Brown
Department of Psychology
Birkbeck College
Malet Street
London WC 1, England 1

Professor C. J. Burke
Department of Psychology
Indiana University
Bloomington, Indiana 1

Professor R. R. Bush
106 College Hall
University of Pennsylvania
Philadelphia 4, Pennsylvania 1

Dr. Donald Campbell
Department of Psychology
Northwestern University
Evanston, Illinois 1

Dr. Violet Cane
Newham College
Cambridge, England 1

Mr. Merrill Carlsmith
Dept. of Social Relations
Harvard University
Cambridge 38, Massachusetts 1

Professor Rudolf Carnap
Department of Philosophy
University of California
Los Angeles 24, California 1

Professor Edward C. Carterette
Department of Psychology
University of California
Los Angeles 24, California 1

Professor C. West Churchman
School of Business Administration
University of California
Berkeley 4, California 1

Dr. Clyde H. Coombs
Department of Psychology
University of Michigan
Ann Arbor, Michigan 1

Dr. Donald Davidson
Department of Philosophy
Stanford University 1

Dr. Gerard Debreu
Cowles Commission
Box 2125, Yale Station
New Haven, Connecticut 1

Dr. Mort Deutsch
Bell Telephone Laboratories
Murray Hill, New Jersey 1

Professor Robert Dorfman
Department of Economics
Harvard University
Cambridge 38, Massachusetts 1

Dr. H.C.J. Duijker
Psychologisch Laboratorium
Keizergracht
Amsterdam 613/C,
The Netherlands 1

Dr. Ward Edwards
Department of Psychology
University of Michigan
Ann Arbor, Michigan 1

Dr. Jean Engler
Department of Statistics
Harvard University
Cambridge 38, Massachusetts 1

Professor W. K. Estes
Department of Psychology
Indiana University
Bloomington, Indiana 1

Professor Robert Fagot
Department of Psychology
University of Oregon
Eugene, Oregon 1

Dr. Leon Festinger
Department of Psychology
Stanford University 1

Dr. C. Flament
Laboratoire de Psychologie
Experimentale et Comparee
Ecole Pratique des Hautes Etudes
46 Rue St. Jacaues
Paris (Ve), France 1

Professor Merrill M. Flood
231 West Engineering
University of Michigan
Ann Arbor, Michigan 1

Professor Maurice Fréchet
Institut H. Poincaré
11 Rue P. Curie
Paris 5, France 1

Dr. Milton Friedman
Department of Economics
University of Chicago
Chicago 37, Illinois 1

Dr. Eugene Galanter
Center for Behavioral Sciences
202 Junipero Serra Boulevard
Stanford, California 1

Dr. Murray Gerstenhaber
Institute for Advanced Study
Dept. of Mathematics
Princeton, New Jersey 1

Dr. I. J. Good
25 Scott House
Cheltenham, England 1

Dr. Leo A. Goodman
Statistical Research Center
University of Chicago
Chicago 37, Illinois 1

Professor Nelson Goodman
Department of Philosophy
University of Pennsylvania
Philadelphia, Pennsylvania 1

Professor Harold Gulliksen
Educational Testing Service
20 Nassau Street
Princeton, New Jersey 1

Professor Louis Guttman
Israel Institute of Applied
Social Research
David Hamlech No. 1
Jerusalem, Israel 1

Mr. John Harsanyi
2/165 Victoria Road
Bellevue Hill, N.S.W.
Australia 1

Dr. T. T. ten Have
Social - Paed. Instituut
Singel 453
Amsterdam, Netherlands 1

Professor Carl G. Hempel
Department of Philosophy
Princeton University
Princeton, New Jersey 1

Dr. W. E. Hick
Psychology Laboratory
Downing Place
Cambridge, England 1

Professor Leonid Hurwicz
Serra House
Stanford University 1

Dr. A. R. Jonckheere
Dept. of Psychology
University College
London WC 1, England 1

Professor Lyle V. Jones
Department of Psychology
University of North Carolina
Chapel Hill, North Carolina 1

Professor Donald Kalish
Department of Philosophy
University of California
Los Angeles 24, California 1

Dr. Leo Katz
Department of Mathematics
Michigan State College
East Lansing, Michigan 1

Prof. Harold H. Kelley
Laboratory for Research
in Social Relations
University of Minnesota
Minneapolis 14, Minnesota 1

Professor John G. Kemeny
Department of Mathematics
Dartmouth College
Hanover, New Hampshire 1

Professor T. C. Koopmans
Cowles Foundation for Research
in Economics
Box 2125, Yale Station
New Haven, Connecticut 1

Professor W. Kruskal
Department of Statistics
Eckart Hall 127
University of Chicago
Chicago 37, Illinois 1

Dr. David La Berge
Department of Psychology
University of Minnesota
Minneapolis 14, Minnesota 1

Professor Douglas Lawrence
Department of Psychology
Stanford University 1

Mr. Bernhardt Lieberman
Emerson Hall
Cambridge 38, Massachusetts 1

Dr. Duncan Luce
Department of Social Relations
Harvard University
Cambridge 38, Massachusetts 1

Professor Robert McGinnis
Department of Sociology
University of Wisconsin
Madison 6, Wisconsin 1

Dr. W. G. Madow
Engineering Research Division
Stanford Research Institute
Menlo Park, California 1

Professor Jacob Marschak
Box 2125, Yale Station
New Haven, Connecticut 1

Dr. Samuel Messick
Educational Testing Service
Princeton University
Princeton, New Jersey 1

Professor G. A. Miller
Center for Behavioral Sciences
202 Junipero Serra Boulevard
Stanford, California 1

Professor Richard C. Montague
Department of Philosophy
University of California
Los Angeles 24, California 1

Dr. O. K. Moore
Department of Sociology
Box 1965, Yale Station
New Haven, Connecticut 1

Professor Sidney Morgenbesser
Department of Philosophy
Columbia University
New York 27, New York 1

Professor Oskar Morgenstern
Department of Economics
and Social Institutions
Princeton University
Princeton, New Jersey 1

Professor Frederick Mosteller
Department of Statistics
Harvard University
Cambridge 38, Massachusetts 1

Professor Ernest Nagel
Department of Philosophy
Columbia University
New York 27, New York 1

Dr. Theodore M. Newcomb
Department of Psychology
University of Michigan
Ann Arbor, Michigan

Professor Jerzy Neyman
Department of Statistics
University of California
Berkeley 4, California 1

Professor Hukukane Nikaido
The Institute of Social
and Economic Research
Osaka University
Toyonaka, Osaka, Japan 1

Professor A. G. Papandreou
Department of Economics
University of California
Berkeley 4, California 1

Dr. Juliette Popper
Department of Psychology
University of Kansas
Lawrence, Kansas 1

Dean G. Pruitt
Department of Psychology
University of Michigan
Ann Arbor, Michigan 1

Dr. Hilary Putnam
Department of Philosophy
Princeton University
Princeton, New Jersey 1

Professor Willard V. Quine
Center for Behavioral Sciences
202 Junipero Serra Boulevard
Stanford, California 1

Professor Roy Radner
Department of Economics
University of California
Berkeley 4, California 1

Professor Howard Raiffa
Department of Statistics
Harvard University
Cambridge 38, Massachusetts 1

Professor Nicholas Rashevsky
University of Chicago
Chicago 37, Illinois 1

Professor Philburn Ratoosh
Department of Psychology
University of California
Berkeley 4, California 1

Dr. Frank Restle
Department of Psychology
Michigan State University
East Lansing, Michigan 1

Professor David Rosenblatt
American University
Washington 6, D.C. 1

Professor Alan J. Rowe
Management Sciences Research
Project
University of California
Los Angeles 24, California 1

Professor Herman Rubin
Department of Mathematics
University of Oregon
Eugene, Oregon 1

Professor Richard S. Rudner
Department of Philosophy
Michigan State University
East Lansing, Michigan 1

Dr. I. Richard Savage
School of Business
University of Minnesota
Minneapolis, Minnesota 1

Professor L. J. Savage
Committee on Statistics
University of Chicago
Chicago, Illinois 1

Dr. Dana Scott
Eckart Hall
University of Chicago
Chicago 37, Illinois 1

Dr. C. P. Seitz
Special Devices Center
Office of Naval Research
Sands Point
Port Washington
Long Island, New York 1

Dr. Marvin Shaw
School of Industrial Management
Massachusetts Institute of
Technology
50 Memorial Drive
Cambridge 39, Massachusetts 1

Mr. R. L. Shuey
General Electric Research Lab.
Schenectady, New York 1

Dr. Sidney Siegel
Department of Psychology
Penn. State University
University Park, Pennsylvania 1

Professor Herbert Simon
Carnegie Institute of Technology
Schenley Park
Pittsburgh, Pennsylvania 1

Dr. Herbert Solomon
Department of Statistics
Stanford University 1

Professor K. W. Spence
Psychology Department
State University of Iowa
Iowa City, Iowa 1

Dr. F. F. Stephan
Box 337
Princeton University
Princeton, New Jersey 1

Mr. Saul Sternberg
Department of Social Relations
Emerson Hall
Harvard University
Cambridge 38, Massachusetts 1

Professor S. Smith Stevens
Memorial Hall
Harvard University
Cambridge 38, Massachusetts 1

Dr. Donald W. Stilson
Department of Psychology
University of Colorado
Boulder, Colorado 1

Dr. Dewey B. Stuit
108 Schadffer Hall
State University of Iowa
Iowa City, Iowa 1

Dr. N. S. Sutherland
Institute of Exper. Psychology
1 South Parks Road
Oxford, England 1

Professor Alfred Tarski
Department of Mathematics
University of California
Berkeley 4, California 1

Professor G. L. Thompson
Department of Mathematics
Ohio Wesleyan
Delaware, Ohio 1

Dr. Robert L. Thorndike
Teachers College
Columbia University
New York, New York 1

Professor R. M. Thrall
University of Michigan
Engineering Research Institute
Ann Arbor, Michigan 1

Dr. Masanao Toda
Department of Experimental
Psychology
Faculty of Letters
Hokkaido University
Sapporo, Hokkaido, Japan 1

Dr. E. Paul Torrance
Bureau of Educational Research
University of Minnesota
Minneapolis, Minnesota 1

Professor A. W. Tucker
Department of Mathematics
Princeton University, Fine Hall
Princeton, New Jersey 1

Dr. Ledyard R. Tucker
Educational Testing Service
20 Nassau Street
Princeton, New Jersey 1

Professor Edward L. Walker
Department of Psychology
University of Michigan
Ann Arbor, Michigan 1

Professor Morton White
Department of Philosophy
Harvard University
Cambridge 38, Massachusetts 1

Dr. John T. Wilson
National Science Foundation
1520 H Street, N. W.
Washington 25, D.C. 1

Professor Kellogg Wilson
Kent State University
Kent, Ohio 1

Professor J. Wolfowitz
Department of Mathematics
Cornell University
Ithaca, New York 1

Professor Robert J. Wolfson
Department of Economics
Michigan State University
East Lansing, Michigan 1

Professor O. L. Zangwill
Psychology Laboratory
Downing Place
Cambridge, England 1

Professor F. Zeuthen
Rungstend Kyst
Denmark 1

Additional copies for project
leader and assistants and reserve
for future requirements 25

Professor Davis H. Howes
Department of Economics
and Social Science
Massachusetts Institute
of Technology
Cambridge 39, Massachusetts 1