A MATHEMATICAL ANALYSIS OF MAZE LEARNING*

SHIGERU ONO

Hiroshima University

Numerous researches have revealed that in learning a maze which involves several choice points, a white rat does not learn all choice points uniformly. That is to say, the relative difficulty in a maze varies with the position of choice points. We shall be able to understand it most closely, when we compare the learning curves of respective choice point with each other. Then, we shall possibly find one curve improves more quickly but another shows no progress, sometimes even a drop. Several factors that bring such a variety of learning curves have been already identified by many experimenters and theorists (for this subject, see e.g., a detailed exposition by Munn (12, p. 225–266, p. 376–396)). But the functional relation between those factors and learning curves at each choice point has not yet been mathematically formulated. It is a purpose of this study to deduce its equation from some basic assumptions.

I Basic Theory

Though at present many types of maze serve experiments in laboratory, it will be desirable that we proceed to analyse from a relatively simpler situation to more complex one. Therefore, we shall limit our following discussion to the analysis of a linear maze or its modified type which seems to be simpler.

Consider that a white rat learns a $N$-unit maze. On every trial, the rat shall acquire some quantity of correct response tendency at each choice point. Such an increment of correct response tendency added by a trial is designated as $\Delta e$. Here $\nu$ is the ordinal number of choice points which are numbered from entrance to goal. In other words, $\nu$ is serial position and $\nu=1, 2, \ldots N$. The direction of responses in $\Delta e$ varies with a given correct response sequence. For example, let the sequence be RRLL ($R$=right, $L$=left), and then the directions of $\Delta e_1$ and $\Delta e_2$ are right while those of $\Delta e_3$ and $\Delta e_4$ are left. If equal correct response tendency increases on each trial, the tendency on trial $n$, $e^{(n)}$, is represented by the equation

$$e^{(n)} = n \Delta e, \quad n = 1, 2, \ldots$$

(I. 1)
This set of tendencies \( \epsilon_1^{(n)}, \epsilon_2^{(n)}, \ldots, \epsilon_N^{(n)} \) we term the \textit{basic gradient}. For the number of trials, conveniently the first trial is called the 0th trial, the next trial the 1st trial, and so on. For a while we shall consider only the case where \( n \geq 1 \) and we shall later refer to the case where \( n = 0 \).

Inferring from the results of past analysis on maze learning, it will be adequate to consider \( \Delta \epsilon \), as composed of two components (see Fig. 1). The first is goal gradient indicated by Hull (9), Spence (16) and others. The second is entrance gradient found by Carr (4) or Haslerud (7). If we accept these two gradients, we can write \( \Delta \epsilon \) as

\[
\Delta \epsilon = F_{\epsilon} + G_{\epsilon},
\]

where \( F_{\epsilon} \) is entrance gradient and \( G_{\epsilon} \) is goal gradient. Though on the shape of gradients the researchers differ in opinion, it will be most plausible to consider that both gradients are represented by exponential function. In fact, according to Haslerud (7), the shape of entrance gradient is exponential. Goal gradient is also formulated in exponential type by Hull (11)*. If we follow their formulation, we can obtain

\[
F_{\epsilon} = F_{\epsilon} \lambda^{n-1}, \quad (I.3)
\]

\[
G_{\epsilon} = G_{\epsilon} \mu^{N-n}, \quad (I.4)
\]

where \( \lambda \) and \( \mu \) are the parameters determining the slope of entrance or goal gradient.

As Hull pointed out (11, p. 159-162), correct response tendency of the \( v \)th choice point on trial \( n \), \( \epsilon_v^{(n)} \), will generalize toward every other choice point. Here, generalization toward the front is called anticipatory tendency, while generalization toward the rear is perseverative tendency. And, trial by trial the animal discriminates each choice situation more and more. As the result, the slope of generalization gradient will become steeper (see Fig. 2). These qualitative considerations do not immediately lead to any special mathematical function. However, they give us a general framework

* As independent variable of his function, Hull adopts delay of reinforcement (11, p. 158). Therefore, strictly speaking Hull’s equation differs a little from ours.
under which we formulate it. As one of equations consistent with the above-mentioned statement on generalization and discrimination, we tentatively formulate as below. By \( a_{m'}^{(n)} \), we denote the tendency to anticipate correct response of the \( v \)th choice point at the front with remoteness \( m' \) on trial \( n \). Then

\[
\epsilon_{m'}^{(n)} = \epsilon_v^{(n)} e^{-m' \beta - u(n-1)} \quad (I.5)
\]

where \( \beta \) is the slope of anticipation gradient and \( u \) is speed of discrimination. For the definition of the remoteness, see Fig. 2. Similarly, let \( \gamma_{n'}^{(n)} \) be the tendency to persevere the correct response of the \( v \)th choice point at the rear with remoteness \( n' \) on trial \( n \), and then

\[
\gamma_{n'}^{(n)} = \gamma_v^{(n)} e^{-n' \delta - u(n-1)} \quad (I.6)
\]

where \( \delta \) is the slope of perseveration gradient.

According to the above hypotheses, obviously the responses in each choice point are determined by both basic and generalization gradients. Moreover, the position preference will bear on the choice response, too. On the summation of these response tendencies Hull (11, p. 162) assumed the "behavioral summation". However, it seems that this assumption proposed by Hull does not improve fitness of theory to data but merely complicates the theory so far as maze learning is concerned. Thus we assume that the right or left response tendency in each choice point is determined by simple addition of the above-mentioned three tendencies. Let us illustrate it by an example. Suppose that a rat learns 4-unit maze with sequence RRLL (see Fig. 3). Denoting the right or left total response tendency in the \( v \)th choice point on trial \( n \) by \( h_R^{(n)} \) or \( h_L^{(n)} \), we have

\[
h_R^{(n)} = j_R + \epsilon_1^{(n)} + \alpha_1^{(n)} = j_R + \epsilon_1^{(n)} + \alpha_1^{(n)},
\]

\[
h_L^{(n)} = j_L + \epsilon_2^{(n)} + \alpha_2^{(n)} = j_L + \epsilon_2^{(n)} + \alpha_2^{(n)},
\]

\[
h_3^{(n)} = j_R + \gamma_1^{(n)} + \epsilon_1^{(n)} + \alpha_2^{(n)} + \alpha_3^{(n)},
\]

\[
h_4^{(n)} = j_R + \gamma_2^{(n)} + \gamma_3^{(n)} + \epsilon_1^{(n)} + \epsilon_2^{(n)} + \epsilon_4^{(n)}.
\]

where \( j_i \) (i = R, L) shows right or left response tendency owing to the position preference.

The above set of equations can be represented in a simple and general form if we use matrix and vectors. By \( E^{(n)} \) we denote a square matrix that has basic gradient at a diagonal, anticipation gradient at its upper right, and perseveration gradient at its lower left:

\[
E^{(n)} = \begin{pmatrix}
\epsilon_1^{(n)} & \alpha_1^{(n)} & \cdots & \alpha_2^{(n)} \\
\alpha_2^{(n)} & \epsilon_2^{(n)} & \cdots & \alpha_3^{(n)} \\
& \alpha_3^{(n)} & \cdots & \alpha_{N,1}^{(n)} \\
& & \cdots & \epsilon_{N-1}^{(n)} \\
& & & \epsilon_N^{(n)}
\end{pmatrix} \quad (I.7)
\]
This matrix is termed the gradient matrix, for it is composed of basic anticipation and perseveration gradients.

Thus far we did not refer to case where \( n=0 \). Now we must examine the gradient on trial 0. Since no habit acts on the initial trial, we get

\[
E^{(0)} = 0.
\]

(I. 8)

The symbol \( O \) is zero matrix.

Next, we designate the correct response sequence as a column vector \( r \):

\[
\begin{bmatrix}
\vdots \\
R \\
\vdots \\
L
\end{bmatrix}
\]

(I. 9)

Put \( R=1 \) and \( L=0 \) in this vector \( r \), and then we have a vector, which is denoted by \( r_R \). On the other hand, by \( r_L \), we denote a vector obtained by putting \( R=0 \) and \( L=1 \) in \( r \).

Using these matrix and vectors, we can write the above-mentioned hypotheses on additiveness of response tendencies as follows. Let \( h_i^{(n)} \) be the column vectors whose components are \( h_{ii}^{(n)} \), \( h_{ii+1}^{(n)} \), \( \ldots h_{NN}^{(n)} \) \((i=\text{R, L})\). (We may call it habit vector when \( j=0 \).) Then,

\[
\begin{align*}
\hat{h}_R^{(n)} &= j_R e + E^{(n)} r, \\
\hat{h}_L^{(n)} &= j_L e + E^{(n)} r_L
\end{align*}
\]

(I. 10)

(I. 11)

In the above equations,

\[
e = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}
\]

(I. 12)

Though both right and left response tendencies take place in the choice situation, the response that occurs really is either one. The probability that either response occurs will depend on the relative magnitude of two response tendencies \( \hat{h}_R^{(n)} \) and \( \hat{h}_L^{(n)} \). For calculating probability of response, we must introduce a hypothesis on probability distribution of response tendencies. Let \( X_R^{(n)} \) and \( X_L^{(n)} \) be the random variables corresponding to right and left response tendencies in the \( n \)th choice point on trial \( n \). Here we assume that both \( X_R^{(n)} \) and \( X_L^{(n)} \) have the normal distribution with the means and the variances as follow*.

\[
\begin{align*}
E(X_R^{(n)}) &= h_R^{(n)}, \\
E(X_L^{(n)}) &= h_L^{(n)}, \\
\sigma^2(X_R^{(n)}) &= \sigma^2(X_L^{(n)}) = \sigma^2.
\end{align*}
\]

(I. 13)

That is to say, the means are given by the vectors, \( h_R^{(n)} \) and \( h_L^{(n)} \), and the variance has no bearing on the number of trials, serial position and direction of choices. Moreover, it is assumed that \( X_R^{(n)} \) and \( X_L^{(n)} \) are mutually independent.

Then, designating probability of right response of the \( n \)th choice point on trial \( n \) as \( p_R^{(n)} \), we have

\[
p_R^{(n)} = \Pr\{X_R^{(n)} > X_L^{(n)}\}
\]

\[
= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{h_R^{(n)} - h_L^{(n)}}{\sigma}} \exp\left(-\frac{1}{2}y^2\right) dy.
\]

(I. 14)

Though we can represent left response probability, \( p_L^{(n)} \), by the similar equation as Eq. (I. 14), it is not necessary to do so. Because

\[
p_R^{(n)} + p_L^{(n)} = 1.
\]

(I. 15)

We get the value of \( p_R^{(n)} \) easily, if only the value of \( p_L^{(n)} \) is given.

* Hull's calculation of the percent of choices of the act associated with the shorter or longer delay of reinforcement (10, p. 162-163) is based on a similar consideration as ours.
Eq. (I. 14) is well-known probability integral and their values are shown in ordinary statistical tables. Therefore we can obtain the value of $p_R^{(n)}$ when a vector, $h_R^{(n)} - h_L^{(n)}$, is known. This vector can be calculated as follows.

Substitute Eq. (I. 5) and Eq. (I. 6) for the right side of Eq. (I. 7), and then we get

$$E^{(n)} = n\{A + e^{-u(n-1)}B\}.$$  \hfil (I. 16)

Here

$$A = \begin{pmatrix} \varepsilon_1^{(1)} & 0 \\ \varepsilon_2^{(1)} & 0 \\ \vdots & \vdots \\ 0 & \varepsilon_N^{(1)} \end{pmatrix}.$$  \hfil (I. 17)

This is called the basic matrix, since its diagonal is the basic gradient. And

$$B = \begin{pmatrix} 0 & \varepsilon_2^{(1)}e^{-\delta} & \ldots & \varepsilon_N^{(1)}e^{-(N-1)\delta} \\ \varepsilon_1^{(1)}e^{-\delta} & 0 & \ldots & \varepsilon_N^{(1)}e^{-(N-2)\delta} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_1^{(1)}e^{-(N-1)\delta} & \varepsilon_2^{(1)}e^{-(N-1)\delta} & \ldots & 0 \end{pmatrix}.$$  \hfil (I. 18)

We call $B$ the generalization matrix.

Let $r^+_R$ be a vector obtained by putting $R=1$ and $L=-1$ in the correct response sequence $r$. And put

$$j_R^+ = j_R - j_L.$$  \hfil (I. 19)

Then, according to Eq. (I. 10), Eq. (I. 11) and Eq. (I. 16)

$$h_R^{(n)} - h_L^{(n)} = j_R^+e + E^{(n)}r_R^+ = j_R^+e + n\{Ar_R^+ + e^{-u(n-1)}Br_R^+\}.$$  \hfil (I. 20)

Owing to Eq. (I. 14) and Eq. (I. 20), we can obtain the value of $p_R^{(n)}$. From this the probability of correct response is calculated immediately. If the correct response of the choice point is right, $p_R^{(n)}$ becomes the probability of correct response just as it is. Conversely, if the correct response is left, $1 - p_R^{(n)}$ is the probability according to Eq. (I. 15). Formally the probability of correct response is expressed by following equation. Let $p^{(n)}$ be a column vector whose elements are the $N$ probabilities of correct response on trial $n$. Then

$$p^{(n)} = r_L + D_Rp^{(n)}.$$  \hfil (I. 21)

In the above equation,

$$p^{(n)} = \begin{pmatrix} p_{1R}^{(n)} \\ p_{2R}^{(n)} \\ \vdots \\ p_{NR}^{(n)} \end{pmatrix}.$$  \hfil (I. 22)

And $D_R$ is a diagonal matrix which is obtained by putting $R=1$ and $L=-1$ in a diagonal matrix $D$ whose diagonal is correct response sequence.

Eq. (I. 21) is the equation of learning curves for which we have searched. From this we can get theoretical learning curves of each choice point in a given maze.

II THE APPLICATION OF THEORY

As represented in the last section, the correct response sequence is one of the important experimental variables that determine the relative difficulty of each choice point in a given maze. So, at first we shall classify the correct response sequences into several types and thereafter on each type of sequences we shall examine whether our theory fits to the experimental data.

The types of correct response sequence

When all $N$ correct responses in a
sequence are of the same direction such as in a sequence RRRR or LLLL, we call it a homogeneous sequence. In contrast to this, the sequence which is composed of two kinds of correct response, R and L, is called a heterogeneous sequence. This involves several special subtypes as follow. (1) Double-homogeneous sequence. In this sequence, the number of units, \( N \), is even and its first and second half form respectively the homogeneous sequences which are different in the direction. For instance, RRRLLL is a double-homogeneous sequence. (2) Alternation sequence. This is a sequence that right and left responses appear alternatively such as RLRLRL. (3) Isolated-correct-response sequence. This is a homogeneous sequence which involves a single heterogeneous response. For example, RRLRRR is so. Though other kinds of heterogeneous sequence than these are, of course, conceivable, only the above ones shall be analysed in this article.

In connection with the types of correct response sequences, we will refer to the reversed sequence. When each component of correct response, R or L, in an original sequence \( r \) is exchanged for the opposed turn, we call such a sequence the reversed one and we designate it as \( r' \). For example, the reversed sequence of RLRL is LRLR. If the number of subjects is large enough and both sequences, \( r \) and \( r' \), are used in experiment, position preference may be offset (that is \( j'_R = 0 \)) by pooling the data. It is clear that probabilities of correct responses in \( r \), \( p^{(m)} \), are equal to that of \( r' \), \( p'^{(m)} \), provided that there is no position preference. Therefore, owing to either sequence we may obtain the same values of probability of correct responses, when both \( r \) and \( r' \) are used and consequently the preference is cancelled.

In the following pages, we shall begin with consideration on each subtype of heterogeneous sequences. Later on we shall refer to the homogeneous one.

Double-homogeneous sequence

Double-homogenous sequences were used by Woodbury (20). In his experiment, the linear maze which has four, six or eight units \( (N=4, 6, 8) \) was learned by three groups of rats fifty trials. Though he subdivided each group into two groups according to which direction of the first half of sequence had been preferred, we will lump these together and show the learning curves of each choice point in Fig. 4. Empirical values of each curve are respectively the average data of thirty rats on trials 0-4, 5-9, 10-14, 15-19, 20-29, 30-39 and 40-49 from left to right in the figure.

Theoretical values of probability \( p^{(m)} \) shown in the right half in Fig. 4 are calculated from Eq. (I. 14) and Eq. (I. 20) using eight parameter-values of the first row in Table 1. Among these parameter-values, those of \( F_1 \), \( G_N \) and \( j'_R \) are the values that standard deviation of \( X^{(n)}_R \) or \( X^{(n)}_L \), \( \sigma \), is taken as a unit. Hence the probabilities must be computed putting \( \sigma = 1 \) in Eq. (I. 14). Not only in this sequence but also in others, the above three parameters are measured by unit of \( \sigma \).

Let us observe Fig. 4 more carefully. Then we shall find it is the characteristic phenomenon in the double-homogeneous sequence that learning of the \( N/2 \)th choice point retards remarkably. Owing to our theory it arises on account that the rat anticipates intensely the following inverse turn in the \( N/2 \)th choice point.

Comparing empirical curves with theoretical ones individually, it is found that each theoretical curve does not fit to experimental data so well. However, in the respect of general character
of curve families both are similar. Accordingly, we may consider that though our theory is incomplete, it plays the role of stepstone toward a better theory.

**Alternation sequence**

Using the 6-unit alternation sequence Ruch (15) trained 36 rats for twenty
trials under two conditions of low and high motivation. Though he reported each result under both conditions, it is not our present question to analyse motivational factor. Therefore, pooling these data together, we will compute the values of $p^{(m)}$ (see Fig. 5). The empirical values of each curve are the average on trial 0-4, 5-9, 10-14 and 15-19.

Learning curves of alternation sequence have two characteristics. In the first place when $\nu$ is an even number, the curves increase monotonously. On the contrary, if $\nu$ is an odd number, the curves drop at early stage of learning and thereafter they increase gradually. (When $N$ is odd, the above-mentioned relation of even or odd becomes inverse.)

As the second characteristics, we can mention that the relative difficulty of learning rises and falls alternately varying with position. Seeing learning curves we can understand it. But, to know it more intuitively, it is desirable to compute the mean errors per a subject and to draw the serial position curve. Designating the mean errors of a given choice point as $Q$, we have

$$Q = \sum_{n=0}^{\xi} (1 - p^{(m)}),$$  \hspace{1cm} (II. 1)

where $p^{(m)}$ is the probability of correct response in a given choice point on trial $n$, and $\xi$ shows the number of trials trained. (Deduction of this equation is explained in e.g., Ono (13, p. 5-6)). In Fig. 6 is showed the comparison between theoretical values of $Q$ and Ruch's data. Both empirical and theoretical values of this figure were calculated from the curves of Fig. 5.

Buel's experimental result (3) shows the intrinsically similar tendency, too.

Fig. 5. Learning curves in alternation sequence. Empirical curves were plotted according to Ruch (15).

Fig. 6. Mean errors of each position in alternation sequence. Solid circles show the values calculated from Ruch's data. Smooth curve is theoretical one.
He used 8-unit alternation sequence and trained twenty-seven rats 225 trials. In this series of trials the eighth unit was removed after the 19th trial. Since the effect of removal of unit is not our problem at present, we will omit consideration on it. Furthermore he did not report the experimental result on trial 0. So we show the mean errors on trials 1–19 (see Fig. 7). Observation of Fig. 7 reveals that values of mean errors represent the similar swing to Ruch's data.

These phenomena, according to our theory, occur because in the odd choice points the animals anticipate the immediately following inverse turn more intensely.

**Isolated-correct-response sequence**

Ruch (14) trained twenty rats 30 trials using 6-unit sequence RRLRRR and its reversed one. His experimental result is represented in Fig. 8. The empirical values of each curve are respectively the average data on trials 0–4, 5–9, 10–14, 15–19, 20–24, and 25–29. As shown in the figure the learning of isolated correct response is most difficult. However, if the isolated response is at the end of sequence, we shall obtain a different result from the above one.

Using 8-unit sequence RRRRRRRRL, Spragg (18) trained four rats 100 trials. In Fig. 9 are shown his data. The empirical values are the average on trials 0–9, 10–24, 25–39, 40–54, 55–69, 70–84, and 85–99. The apparatus used by Spragg is a maze that a linear one is doubled. In this type of maze, it will be possible that there act some factors neglected in our theory such as goal orientation (17) * . But we tentatively consider these factors as negligible.

* Goal orientation shall be ascribable to goal gradient and anticipation in our theory, if apparent goal orientation is really determined by the direction of the last turn as Brogden pointed out (2. p. 601).
Fig. 9. Learning curves in isolated-correct-response sequence which has a heterogeneous turn at the last choice point. Empirical curves were drawn owing to Spragg (18).

Fig. 10. Mean errors in homogeneous sequence. Each solid circle was plotted from De Montpellier (6). Bow curve was computed theoretically.

Homogeneous sequence

Homogeneous sequence was used by De Montpellier (6). In his study the rats learned the linear diamond maze of 6-unit sequence RRRRRR until they could run perfectly three successive trials. At Fig. 10 we represent the mean errors of thirteen rats who learned his Maze III.

Since the theoretical values of $p^{(n)}$ in this sequence converge quickly to unity as $n$ becomes greater, the theoretical value of $Q$ shall not accompany so great error even if $\infty$ is substituted for $\xi$ in Eq. (II. 1). Hence, putting $\xi = 20$ which is a sufficiently great value, we computed $Q$-value of Fig. 10.

In homogeneous sequence, the subjects learn the maze very easily, and consequently there occur few errors. This is because the directions of all response tendencies—correct response tendency, anticipation and perseveration—are same. Here we recall Dashiell and Bayroff's experiment (5) in which the difficulty was compared between alternation sequence and homogeneous one using 6-unit U-maze. The result was that the learning of homogeneous sequence was much easier. This finding was explained by them in terms of forward-going tendency or,
speaking more adequately, centrifugal swing. But, beside it, we think we should take account of the summation of the above-mentioned response tendencies.

Homogeneous sequence has another characteristic that the middle of the sequence is learned most easily. This fact has been already pointed out by Hull (11, p. 162).

**Some further implications**

Thus far we have examined fitness of our theory to various experimental data. Next, we will refer to the parameter-values involved in theoretical curves. In Table 1 are represented the values of parameters estimated from experimental data. When we compare these values with each other in the respective column, we find they are different considerably. This shall occur owing to the difference of various conditions for which no regard have been paid in this paper such as hereditary character, age, past experience, drive condition, structure of maze, extra-maze cues, amount of incentive, distribution of practice, etc.

However, if all values in Table 1 are surveyed, we shall know these values are never random but ruled by noticeable regularity. It can be shown by the following inequalities.

\[ F_1 < G_N, \quad \lambda > \mu, \quad \beta < \delta. \]  \hspace{1cm} (II. 2)

By the first and second inequalities we can see that goal gradient is greater than entrance gradient at each head point (\( v = N \) or 1) and on the respect of their slopes the former's one is steeper than the latter's (see Fig. 1). This accounts for the reason why the existence of goal gradient has been often pointed out by many researchers while entrance gradient has been apt to be overlooked. The third inequality means that anticipation is greater than perseveration (see Fig. 2). This statement accords with Hull's inference (11, p. 161).

At the end of this section, we should make reference to centrifugal swing. It is confirmed enough by several experiments that centrifugal swing plays an important role in maze behavior (1, 19). Nevertheless our theory has taken no thought of it because of the following two reasons. (a) Centrifugal swing does not always occur in maze behavior.

<table>
<thead>
<tr>
<th>Investigator</th>
<th>Sequence of correct responses</th>
<th>Values of parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Woodbury (1950)</td>
<td>Double-homogeneous</td>
<td>( F_1 )</td>
</tr>
<tr>
<td>Ruch (1935)</td>
<td>Alternation</td>
<td>.064</td>
</tr>
<tr>
<td>Buehler (1934)</td>
<td>Alternation</td>
<td>.026</td>
</tr>
<tr>
<td>Ruch (1935)</td>
<td>Isolated-correct-response</td>
<td>.028</td>
</tr>
<tr>
<td>Spragg (1933)</td>
<td>Isolated-correct-response</td>
<td>.009</td>
</tr>
<tr>
<td>De Montpellier (1933)</td>
<td>Homogeneous</td>
<td>.041</td>
</tr>
<tr>
<td>Hill (1939)</td>
<td>Diagonal (Four-choice maze)</td>
<td>.063</td>
</tr>
</tbody>
</table>
behavior. Depending on the pattern of maze, sometimes it does not appear at all or scarcely does. For example, in Woodbury-type maze we cannot expect that centrifugal swing arises, because any explicit turning movement will hardly occur before choice. In these circumstances we need not pay regard to centrifugal swing. (b) Even under the situations in which we cannot ignore it, centrifugal swing shall act mainly at the early stage of learning and its importance shall decrease trial by trial, for learning factor shall become dominant as training proceeds. Therefore, centrifugal swing will be considered as important only under the limited cases. So that, we constructed our theory disregarding it as first approximation.

III Generalization of Theory to Multiple-Choice Situation

Thus far we have considered the situation in which the alternatives at any choice point are two. In this section we proceed a step to generalize our theory to the case of $k$ alternatives, where $k \geq 2$.

At first, we will devise how to represent the correct response sequence in multiple-choice situation. Though in the case where $k=2$ we represented it as a column vector, it is convenient to use a matrix in multiple-choice situation. Let $r_{\nu \kappa}$ ($\nu = 1, 2, \ldots, N$, $\kappa = 1, 2, \ldots, k$) be the $\kappa$th alternative of the $\nu$th choice point (we define $\kappa$ to be numbered from the right end of alternatives, looking toward the goal of maze), and arrange $r_{\nu \kappa}$ in a form of matrix:

\[
\begin{pmatrix}
    r_{11} & r_{12} & \cdots & r_{1k} \\
    r_{21} & r_{22} & \cdots & r_{2k} \\
    \vdots & \vdots & \ddots & \vdots \\
    r_{N1} & r_{N2} & \cdots & r_{Nk}
\end{pmatrix}
\]

In the above matrix we let correct alternatives be unity and wrong ones be zero. This new matrix we designate as $R$ and call response sequence matrix or shortly sequence matrix. Let us illustrate it by an example. If correct response sequence of 4-unit three-choice maze is RLML ($M=\text{middle}$), its sequence matrix is

\[
\begin{pmatrix}
    1 & 0 & 0 \\
    0 & 0 & 1 \\
    0 & 1 & 0 \\
    0 & 0 & 1
\end{pmatrix}
\]

That is to say, $R$ is a matrix obtained in such a way that disposing the start of maze at the upper and the goal at the lower we substitute unity or zero for correct or wrong alternatives. Usually each choice situation in a given maze has a single correct alternative and consequently any row of sequence matrix involves only one element whose value is unity.

Now, we may consider that six equations from Eq. (I. 1) to Eq. (I. 6) hold in the multiple-choice situation just as it is. In other words, correct response tendency, $\Delta_{\nu \kappa}$ ($\nu = 1, 2, \ldots, N$), which consists of both entrance and goal gradients is formed every trial; this tendency generalizes to the front and rear choice points; each choice situation is discriminated as learning proceeds.

Next, we assume additiveness of correct response tendency, anticipation, perseveration and position preference as we did in two-choice situation. As an example, consider that a rat learns 4-unit three-choice sequence RLML. Then, the total response tendency of the $\kappa$th alternative in the $\nu$th choice point on trial $n$, $h_{\nu \kappa}^{(n)}$, is represented as follows.
A Mathematical Analysis of Maze Learning

where \( j_i \) is the position preference to alternative \( i \) \( (i = 1, 2, 3) \). Using matrices we can show the above set of equations as given below.

\[
\begin{pmatrix}
\Delta_{11}^{(n)} & \Delta_{12}^{(n)} & \Delta_{13}^{(n)} \\
\Delta_{21}^{(n)} & \Delta_{22}^{(n)} & \Delta_{23}^{(n)} \\
\Delta_{31}^{(n)} & \Delta_{32}^{(n)} & \Delta_{33}^{(n)} \\
\Delta_{41}^{(n)} & \Delta_{42}^{(n)} & \Delta_{43}^{(n)}
\end{pmatrix}
\begin{pmatrix}
\Delta_1^{(n)} \\
\Delta_2^{(n)} \\
\Delta_3^{(n)} \\
\Delta_4^{(n)}
\end{pmatrix}
= \begin{pmatrix}
\Delta_1^{(n)} \\
\Delta_2^{(n)} \\
\Delta_3^{(n)} \\
\Delta_4^{(n)}
\end{pmatrix}
\]

The matrix of left side in the above equation we call the total response tendency matrix generally and we designate it as \( H^{(n)} \) \( (H^{(n)} \) may be called the habit matrix if no position preference occurs).

Rewriting the above matrix equation in general form, we have relation

\[
H^{(n)} = J + E^{(n)} R,
\]

where

\[
J = \begin{pmatrix}
\Delta_{11}^{(n)} & \Delta_{12}^{(n)} & \Delta_{13}^{(n)} \\
\Delta_{21}^{(n)} & \Delta_{22}^{(n)} & \Delta_{23}^{(n)} \\
\Delta_{31}^{(n)} & \Delta_{32}^{(n)} & \Delta_{33}^{(n)} \\
\Delta_{41}^{(n)} & \Delta_{42}^{(n)} & \Delta_{43}^{(n)}
\end{pmatrix}
\begin{pmatrix}
\Delta_1^{(n)} \\
\Delta_2^{(n)} \\
\Delta_3^{(n)} \\
\Delta_4^{(n)}
\end{pmatrix}
+ \begin{pmatrix}
\Delta_{11}^{(n)} & \Delta_{12}^{(n)} & \Delta_{13}^{(n)} \\
\Delta_{21}^{(n)} & \Delta_{22}^{(n)} & \Delta_{23}^{(n)} \\
\Delta_{31}^{(n)} & \Delta_{32}^{(n)} & \Delta_{33}^{(n)} \\
\Delta_{41}^{(n)} & \Delta_{42}^{(n)} & \Delta_{43}^{(n)}
\end{pmatrix}
\begin{pmatrix}
\Delta_1^{(n)} \\
\Delta_2^{(n)} \\
\Delta_3^{(n)} \\
\Delta_4^{(n)}
\end{pmatrix}
\]

If \( R = I \), Eq. (III. 1) leads to the equation

\[
H^{(n)} = J + E^{(n)}.
\]

Next problem is to calculate the response probabilities of each alternative from \( H^{(n)} \). Here we will conveniently fix the number of trials, \( n \), and the serial position, \( \nu \), and omit the suffixies \( n \) and \( \nu \) to be appended to symbols.

Let \( X_1, X_2, \ldots X_k \) be random variables corresponding to the response tendencies of alternatives at a given position on a given trial. As we did in two-choice situation we assume that these random variables are mutually independent and that their probability density function, \( f_r(x) \), are given as below.

\[
f_r(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-h)^2}{2\sigma^2}\right\}
\]

By \( Y_{ij} \) we denote the difference between \( x_i \) and \( x_j \) which are taken out from the above \( k \) variables \( X_1, X_2, \ldots X_k \).

That is,

\[
Y_{ij} = x_i - x_j,
\]

\( i = 1, 2, \ldots k \), \( j = 1, 2, \ldots k \).

Let \( Y_\nu \) be the set of these \( k-1 \) random variables:

\[
Y_\nu = (Y_{1\nu}, Y_{2\nu}, \ldots Y_{k-1\nu}, Y_k).
\]

When the distribution function of this random variable, \( Y_\nu \), is designated as
the response probability of the $k$th alternative, $p_k$, takes the following form.

$$p_k = \Pr \left\{ Y_{r_1} > 0, \ Y_{r_2} > 0, \ldots \ Y_{r_{i-1}} > 0, \ Y_{r_i+1} > 0, \ldots \ Y_{r_k} > 0 \right\}$$

$$= \int_{0}^{\infty} \cdots \int_{0}^{\infty} dG_k(y_{r_1}, y_{r_2}, \ldots, y_{r_{i-1}}, y_{r_{i+1}}, \ldots, y_{r_k})^*.$$

(III.6)

Here $Y_r$ has $(k-1)$-variate normal distribution and its mean, variance and correlation coefficient are given as below.

$$E(Y_{r_i}) = h_i - h_*$$

$$\sigma^2(Y_{r_i}) = 2\alpha^2$$

$$\rho(Y_{r_i}, Y_{r_i'}) = \frac{1}{k},$$

$$i \neq i', i' = 1, 2, \ldots k - 1, k + 1, \ldots k.$$

From Eq. (III.6), we have

$$\sum_{i=1}^{k} p_i = 1.$$

(III.8)

In Eq. (III.6), we can obtain the response probabilities of any alternative.

Though an expression of theory, generalized in this section, is different a little from that of preceding one, it is clear that the former is the natural generalization of the latter. The reason is as follows. If $k = 2$, we get

$$H^{(m)} = [h^m_R, h^m_L],$$

$$J = [j_R e, j_L e],$$

$$R = [r_R, r_L].$$

Therefore Eq. (III.1) is the form in which both Eq. (I.10) and Eq. (I.11) are summed up. Putting $k = 2$ in Eq. (III.6), we have

$$p_k = \frac{1}{2\sqrt{\pi\sigma^2}} \int_{h_* - h_i}^{\infty} e^{-\frac{(y_{r_i} - (h_* - h_i))^2}{4\sigma^2}} dy.$$
Fig. 11. Learning curves of correct responses in 4-unit four-choice maze. Empirical curves were drawn from Hill (8). For the method to compute theoretical values, see appendix.

\[
\begin{pmatrix}
1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 \\
1 & 0 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0
\end{pmatrix}
\]

the number of reversed sequence is 3, 9, 9 respectively. In Hill’s sequence matrix there are 23 (=4!-1) reversed sequences since no column is equal. Hill used all these reversed sequences in addition to an original one.

The position preference will be cancelled, that is, \( J = 0 \) when all reversed sequences are used in experiment and their results are pooled together. Here we pool the results in the way as follows. Consider that we obtain one of the reversed sequences, \( \tilde{R} \), applying a special operation \( A \) (e.g., an operation that the 1st column is exchanged for the 3rd one)* to an original sequence, \( R \), and that by \( R \) we have a probability matrix, \( P \), whose components are \( p_{iv} \) and by \( \tilde{R} \) we have \( \tilde{P} \). At the time we pool both \( P \) and \( A \tilde{P} \). According to Hill’s experimental results of pretraining period in which no door of maze is blocked, the middle alternatives (\( \kappa = 2, 3 \)) tend to be more chosen than both ends and this tendency is more remarkable in the neighbourhood of goal (8, p. 571, his Table 1). This finding is not the position preference in our usual term but a kind of goal-depending behavior. It is, however, apparently the same phenomenon as the position preference. Consequently, it is cancelled by means of summing up both experimental results of original and reversed sequences, even if the above tendency occurs in learning period. Every empirical values of probability cited in the following was based on data pooled in the above-mentioned way by Hill (8, p. 577, his Table IV). Therefore, in Hill’s experiment we consider \( J = 0 \).

We will proceed to Hill’s main data. In the first place, we observe the learning curves of correct response in each choice situation. The left half of Fig. 11 shows the empirical curves and the points in any curve are the average on trials 0–9, 10–19, 20–29, 30–39, 40–49 respectively. As represented in Fig. 11, learning of the fourth choice point

* Strictly speaking, \( A \) is an operation which is as follows. By \( I' \) denote a matrix obtained in such a way that columns of unit matrix \((k \times k)\) are exchanged in a given manner. Then, \( A \) is defined as multiplying \( I' \) from the right.
nearest the goal is easiest while the learning score of middle positions is worse. These phenomena occur because the form of basic gradient is concave (see Fig. 1). In the right half of Fig. 11 are showed the theoretical curves.

When we calculated the theoretical values of response probabilities, we had to devise some approximate method. Because in two-choice situation the probabilities can be obtained easily from $h^{(m)}_R - h^{(m)}_L$ if we consult statistical table, but in multiple-choice situation
we have no table to consult except the case where $k=3$. (When $k=3$, we can utilize the table of bivariate normal distribution with correlation coefficient $1/2$.) It requires a great amount of labor to compute probability of 3-variate normal distribution. So, to approximate it we assumed uniform distribution$^*$. 

Next, let us observe the data in details. Though in Fig. 11 are re-

* This approximate method is based on a suggestion of Mr. M. Okamoto (Osaka Univ., Dept of Mathematics). The detail of this method is explained in the appendix at the end of this paper.
S. Ono

presented only probabilities of correct responses, error responses also occur in the process of learning and their frequency is ruled by some regularity. Fig. 12 shows probabilities of both correct and wrong responses in various stages of learning. Each figure indicates the average results on trial 0-9, 10-19, 20-29, 30-39 and 40-49 respectively. Consequently, probabilities of correct responses are the same values as those of Fig. 11. Theoretical curves in the right halves of Fig. 12 are computed, putting \( n = 4.5, 14.5, 24.5, 34.5, 44.5 \). In Fig. 12 the values of error probabilities closely follow the degree of anticipation or perseveration. Because the sequence matrix used by Hill is diagonal and consequently summation of correct response tendency, anticipation and perseveration does not occur. Seeing Fig. 12 we can find that both anticipation and perseveration gradients arise explicitly throughout all stages of learning.

Computation of theoretical values in Fig. 11 and Fig. 12 is based on parameter-values at the last row of Table 1. By these values of parameters we can know that three inequalities (II. 2) stated previously hold also in multiple-choice situation. Perhaps these inequalities may be satisfied in every kind of maze learning.

**APPENDIX**

The method to compute probability in multiple-choice situation

As stated in the text, computation of probability in multiple-choice situation is accompanied with great difficulty. So we need some approximate methods. In this paper, we assume the uniform distribution instead of normal one. That is to say, as a substitute for Eq. (III. 3) we put

\[
f_\varepsilon(x) = \begin{cases} 
1, & h_\varepsilon - \sqrt{3} \sigma \leq x_\varepsilon \leq h_\varepsilon + \sqrt{3} \sigma, \\
0, & h_\varepsilon - \sqrt{3} \sigma > x_\varepsilon, h_\varepsilon + \sqrt{3} \sigma < x_\varepsilon, \\
\varepsilon = 1, 2, \ldots, k. 
\end{cases} 
\] (0.1)

Then, we have

\[
E(X_\varepsilon) = h_\varepsilon, \\
\sigma^2(X_\varepsilon) = \sigma^2(X_\varepsilon) = \ldots = \sigma^2(X_k) = \sigma^2 
\] (0.2)

The response probability of the \( \varepsilon \)th alternative, \( p_\varepsilon \), is obtained in such a way as follows. At first, we rearrange \( h_\varepsilon \) in the order of magnitude and change the number of subscript appended to \( h \) as

\[
h_1 \leq h_2 \leq \ldots \leq h_k.
\]

By \( \kappa' \) we denote this new number of subscript. Corresponding to subscripts of \( h \), we change subscripts of random variables \( X \). After these preparations, we put

\[
\frac{h_\varepsilon' - h_1}{2 \sqrt{3} \sigma} = t_\varepsilon', \quad \kappa' = 1, 2, \ldots, k. \] (0.3)

Then, if

\[
0 \leq t_2 \leq t_3 \leq \ldots \leq t_k \leq 1 \] (0.4)

we have

\[
p_\varepsilon' = \Pr(X_\varepsilon' = \text{max})
\]

\[
= \sum_{l=1}^{h} \prod_{m=l}^{h} \frac{1}{t_m+1} \int_{m+1}^{h} (x-t_m)dx , \] (0. 5)

where we put \( t_0+1 = t_k \). In Eq. (0.5), we have

\[
\sum_{\varepsilon'=1}^{k} p_\varepsilon' = 1 \] (0. 6)

In particular when \( k = 4 \) as in Hill's experiment, according to Eq. (0.5) the response probabilities are given as below.

\[
p_1 = \int_{t_4}^{1} (x-t_2)(x-t_3)(x-t_4)dx,
\]

\[
p_2 = \int_{t_4}^{1} x(x-t_3)(x-t_4)dx + \int_{t_4}^{t_3+1} (x-t_4)dx,
\]

\[
p_3 = \int_{t_4}^{1} x(x-t_2)(x-t_4)dx + \int_{t_4}^{t_3+1} x(x-t_4)dx.
\]
A Mathematical Analysis of Maze Learning

In the above matrix, we will assume the minimum values of each row are $a^{(1)}_{34}$, $a^{(1)}_{23}$, $r^{(1)}_{21}$, and $r^{(1)}_{34}$ respectively. And by $L$ we denote a diagonal matrix which has these minimum values as its diagonal, and by $M$ we show a matrix which has these values in each row except diagonal. That is,

\[
L = \begin{pmatrix}
  a^{(1)}_{34} & 0 \\
  a^{(1)}_{23} & r^{(1)}_{21} \\
  0 & r^{(1)}_{34}
\end{pmatrix},
\]

\[
M = \begin{pmatrix}
  a^{(1)}_{34} & a^{(1)}_{32} & a^{(1)}_{24} \\
  0 & a^{(1)}_{23} & a^{(1)}_{21} \\
  r^{(1)}_{34} & r^{(1)}_{23} & 0 \\
  r^{(1)}_{21} & r^{(1)}_{12} & r^{(1)}_{13} & 0
\end{pmatrix}.
\]

Then, if we take account of the conditions, $J=0$ and $R=1$, we obtain

\[
T^m = \frac{n}{2\sqrt{3}a} \left[ (A - e^{-u(n-1)}L) + e^{u(n-1)}(B - M) \right],
\]

where we put $\sigma=1$, since the values of $F_1$ and $G_N$ have been measured by $\sigma$-unit. In each row of $T^m$ we newly number its components in order of magnitude, and then we get $t_m^{(n)}$. From these values of $t_m^{(n)}$ and Eq. (0.5), theoretical values in Fig. 11 and Fig. 12 were computed.

REFERENCES


(Received May 5, 1959)