LEARNING MODELS

PREPARED FOR:

MATHEMATICAL MODELLING.

6.337

STUDENT : WESLEY MIYAI.
STUDENT No : 5806240.
INSTRUCTOR : T. G. BERRY.
DATE : FRI., MARCH 31, 1989.

Much of our knowledge about <u>Classical Conditioning</u> (that is, when a neutral stimulus is paired with a stimulus that elicits a reflex response) is credited to the Nobel Prize winning Russian Physiologist Ivan P. Pavlov (1849-1936). Apparently, Pavlov was studying how the digestive system occured in dogs when he noticed, quite by accident, that the dogs would often salivate when he approached them with food. Pavlov knew that the taste of food would automatically promote a salivary reaction, but he thought it was strange for the dogs to salivate at the mere sight of food. Therefore, to test his hypothesis that any stimulus proceeding food would cause the dogs to salivate, Pavlov paired a tone with the presentation of food. Later, after a few such pairings, the dogs indeed began to salivate to the tone.

Let us now examine this situation through a probability model of learning.

<u>Objective</u>: To determine the probability of a correct response after a certain number of trials. That is, our objective is to trace the learning patterns of the subject dogs to be able to associate a tone with food. The <u>correct response</u> from the subject dogs, therefore, would be a <u>salivary reaction</u> from the tone.

Let p denote the probability of making a correct response. $(0 \le p \le 1)$

Let q denote the probability of making an incorrect respone. $(0 \le q \le 1)$

Let t denote the n^{TH} trial. $(n = 0, 1, 2, \ldots)$
~~trial number~~

<u>Note</u>: Instead of analyzing the performance level of each subject with respect to time, the analysis of each performance is taken with respect to the number of trials administered.

This is important since, as in the case of Pavlov's experiment, the dogs became conditioned to salivate after hearing a tone after several practice trials. Clearly, the length of time played a lesser role.

This brings us to our first assumption.

although the length of time between trials could play a role in memory of preceding trials

motivation?

<u>Assumption</u>: Let us assume that the experiment is carried out without any "large intervening gaps of time" between trials.

notation?

$p_t, q_t, p_{t+1}, q_{t+1}, \ldots$ etc shall all denote the probabilities taken with respect to the trial number.

Since $0 \le p \le 1$

$\Rightarrow q_t = 1 - p_t$. for $t = 0, 1, 2, \ldots$

<u>assumption</u>: We shall assume that q_t will decrease at a <u>constant rate</u> for each and every trial. In a physical sense, this means that the dogs ability to "learn" what the sound of the tone actually implies will increase at a constant rate, say α, between zero and one.

$$\therefore \; q_{t+1} = \alpha q_t \quad \text{for } t = 0, 1, 2, \ldots \quad 0 < \alpha < 1.$$

$$\therefore \; q_1 = \alpha q_0$$

$$q_2 = \alpha q_1 = \alpha(\alpha q_0) = \alpha^2 q_0$$

$$q_3 = \alpha q_2 = \alpha(\alpha^2 q_0) = \alpha^3 q_0$$

$$\vdots$$

$$q_t = \alpha^t q_0 . \quad \text{for } t = 0, 1, 2, \ldots$$

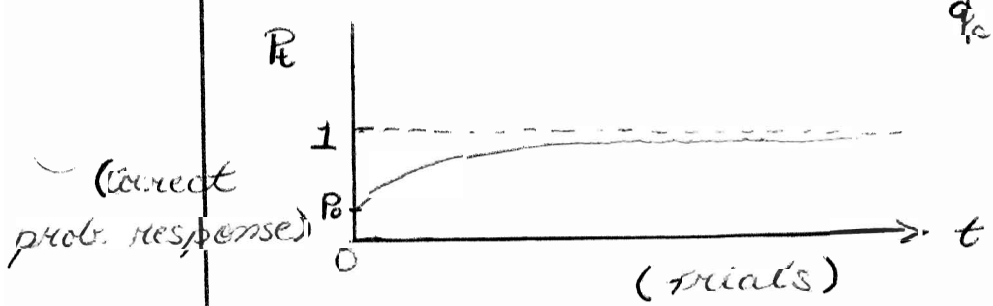
<u>Recall</u>: $q_t = 1 - p_t$ for $t = 0, 1, 2, \ldots$

$$\therefore \; p_t = 1 - q_t$$

or $p_t = 1 - \alpha^t q_0$ for $t = 0, 1, 2, \ldots$

& where $0 < \alpha < 1$

q_0 unknown.



Comments.

1) This is not only a very simple model to generate, its final result is quite significant. In a physical sense, the fact that there exists a gradual "learning" process over a period of t trials implies that the subject dog's, in Pavlov's experiment, should indeed at some point begin to associate a tone with an implication of food. This is what we have already seen to be true. Therefore, the mathematical analysis in our model correctly represents the physical nature in Pavlov's experiment.

2. However, before we can actually state that this model is a complete success, let us look at its model parameters 'q_i' and 'α'. Now although 'q_i' cannot actually be calculated directly, we can approximate 'q_i' by, say, taking the total number of incorrect ~~ex~~ responses in the first trial and then divide this quantity by the total number of trials administered. (This, of course, would be carried out with a "large" number of subjects). The determination of a learning rate (α), however, presents a much more difficult task. So, instead of attempting to derive some sort of imaginative technique to solve for α, let us go one step futher and modify our present model.

average

- Modified Model -

Let us first begin by declaring our variables.

∴ Let t denote the t^{TH} trial.

Let P_t denote the subjects performance level, which may increase or decrease in accordance to a reinforcing or inhibiting factor, in the t^{TH} trials. Once again, since P represents a probability, $0 \leq P \leq 1$.

<u>Question</u>: What exactly do we mean by a reinforcing and inhibiting factor?

Well, before we answer this question, let us make one very important assumption.
Let us assume that P_{t+1} is dependent only on the value of P_t. {This is known as the <u>Markov Property</u> where the likelihood of an outcome of one experiment depends only on the outcome of the immediately preceeding experiment.}

The level of performance in the t^{TH} trial is clearly P_t. However, in accordance with the above assumption, we can now make various observations with respect to the <u>$(t+1)^{ST}$ trial</u>.

1) The <u>maximum possible increase</u> in the subjects level of performance must be <u>$1 - P_t$</u> (recall $0 \leq P_t \leq 1$).

2) The <u>maximum possible decrease</u> in the subjects level of performance must be $\underline{-P_t}$.

3) <u>there exists a possibility of no change in the performance level from the t^{TH} to the $(t+1)^{ST}$ trial</u>.

So, by incorporating these three factors into an equation for P_{t+1}, we ~~claim~~ assume the following.

$$(*) \quad P_{t+1} = \underset{\text{case 1)}}{a(1-P_t)} + \underset{\text{case 2)}}{b(-P_t)} + \underset{\text{case 3)}}{P_t} \qquad : t = 0, 1, 2, \ldots$$

a, b, constants.

$0 \le a \le 1$, $0 \le b \le 1$. } fairly obvious but will be shown later

Now we may answer the question of what a reinforcing and inhibiting factor represent through the interpretation of the model parameters.

To begin with, $(1-P_t)$, as we have already seen, represents the maximum increase in the subject's level of performance from the t^{TH} to the $(t+1)^{ST}$ trial. Therefore, the term $a(1-P_n)$ allows 'a' to be interpreted as a <u>reinforcing factor</u> in (*). With respect to Pavlov's experiment, 'a' may be viewed (in a physical sense) as "The feeding of food to the subject dogs."

i.e. Tone → Dog's Salivate → Dog's are presented with food then fed.

? meaning becomes clear later

(Notice the learning process)

Similarly, since $-P_t$ represents the maximum decrease in a subjects performance level from the t^{TH} to the $(t+1)^{ST}$ trial, the term $b(-P_t)$ allows 'b' to be interpreted as an <u>inhibiting factor</u> in (*). Again, with respect to Pavlov's experiment, 'b' may be viewed (in a physical sense) as a "Presentation of food to the subject dog's, but not the actual feeding of food to the subject dog's."

i.e. Tone → Dog's fail to saliate → Dog's are presented with food but are not fed.

(Notice the learning process)

helps explain above

<u>Note</u>

1) If $P_t = 1$
then $P_{t+1} = 1 - b$.
but since $0 \leq P_{t+1} \leq 1$
$\Rightarrow 0 \leq 1 - b \leq 1$
or $0 \leq b \leq 1$.

2) If $P_t = 0$
then $P_{t+1} = a$
but since $0 \leq P_{t+1} \leq 1$
$\Rightarrow 0 \leq a \leq 1$.

<u>Summary</u>:

$$P_{t+1} = P_t + a(1 - P_t) - bP_t \quad : t = 0, 1, 2, \ldots$$

where $0 \leq a \leq 1$ ✓
$0 \leq b \leq 1$.

Notice that P_{t+1} may be rewritten in the following form:

$$P_{t+1} = P_t(1-a-b) + a \quad ; t = 0, 1, 2, \ldots$$

where $0 \leq a \leq 1$

$0 \leq b \leq 1$.

So, let us now study the effects of a general form of the above equation,

$$y_{k+1} = \alpha y_k + a \quad (\text{let } \alpha = 1-a-b)$$

(where we shall assume y_0 is prescribed.)

Thus, $y_1 = \alpha y_0 + a$

$$\begin{aligned} y_2 &= \alpha y_1 + a \\ &= \alpha[\alpha y_0 + a] + a \\ &= \alpha^2 y_0 + a(1+\alpha) \end{aligned}$$

$$\begin{aligned} y_3 &= \alpha y_2 + a \\ &= \alpha[\alpha^2 y_0 + a(1+\alpha)] + a \\ &= \alpha^3 y_0 + a(1+\alpha+\alpha^2) \end{aligned}$$

$$\begin{aligned} y_4 &= \alpha y_3 + a \\ &= \alpha[\alpha^3 y_0 + a(1+\alpha+\alpha^3)] + a \\ &= \alpha^4 y_0 + a(1+\alpha+\alpha^2+\alpha^3) \end{aligned}$$

$\vdots$

$$y_k = \alpha^k y_0 + a(1+\alpha+\alpha^2+\alpha^3+\cdots+\alpha^{k-1}).$$

where $k = \cancel{0}, 1, 2, 3, \ldots$

But, if we let $S_{\overset{k}{n}} = 1 + \alpha + \alpha^2 + \cdots + \alpha^{k-1}$,

then $\alpha S_{\underset{k}{n}} = \alpha + \alpha^2 + \alpha^3 + \cdots + \alpha^k$

$$\therefore \; S_n - \alpha S_n = 1 - \alpha^k$$

$$\text{or} \quad S_n = \frac{1-\alpha^k}{1-\alpha}, \quad \text{for } k = 0, 1, 2, \ldots$$

<u>Hence</u>,

$$1 + \alpha + \alpha^2 + \cdots + \alpha^{k-1} = \begin{cases} \dfrac{1-\alpha^k}{1-\alpha} & \text{if } \alpha \neq 1 \\ k & \text{if } \alpha = 1 \end{cases} \quad ; k = \cancel{0}, 1, \ldots$$

<u>This implies</u>

$$y_k = \begin{cases} \alpha^k y_0 + a\left(\dfrac{1-\alpha^k}{1-\alpha}\right) & \text{if } \alpha \neq 1 \\ y_0 + a k & \text{if } \alpha = 1 \end{cases}$$

linear for $k = \cancel{0}, 1, 2, \ldots$

<u>Note: For $\alpha \neq 1$</u>,

$$y_k = \alpha^k y_0 + a \frac{1-\alpha^k}{1-\alpha}$$

$$= \alpha^k y_0 + \frac{a}{1-\alpha} - \frac{a\alpha^k}{1-\alpha}$$

$$= \alpha^k \left\{ y_0 - \frac{a}{1-\alpha} \right\} + \frac{a}{1-\alpha}$$

$$y_k = \alpha^k (y_0 - Y) + Y$$

where $Y = \dfrac{a}{1-\alpha}$.

So, in the same sense of being able to convert

$$\boxed{y_{k+1} = \alpha y_k + a \quad \text{into} \quad y_k = \begin{cases} \alpha^k (y_0 - Y) + Y & : \alpha \neq 1 \\ y_0 + ak & : \alpha = 1 \end{cases} \quad \text{where } Y = \frac{a}{1-\alpha} \quad \text{for } k = 0, 1, 2, \ldots}$$

we are able to convert

$$(**) \quad \boxed{P_{t+1} = (1-a-b)^t P_t + a \quad \text{into} \quad P_t = \begin{cases} (1-a-b)^t (P_0 - P^*) + P^* & \text{for } a+b \neq 0, \\ P_0 & \text{for } a+b = 0 \end{cases} \quad \text{where } P^* = \frac{a}{a+b}. \quad \text{for } t = 0, 1, 2, \ldots}$$

<u>Note</u>. In both instances, we are assuming that the initial values y_0 and P_0 are prescribed.

<u>Question</u>: What kinds of results (conclusions?) can be made from (**)?

<u>Solution</u>

By examining equation (**), one should be able to see the term $(1-a-b)^t$ as playing the most significant role in determining any sequence of values for P_t.

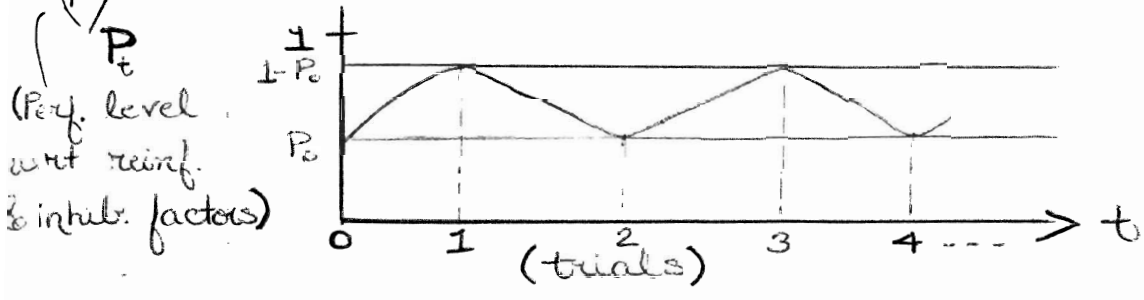
One should also be able to see that since $0 \leq a \leq 1$ and $0 \leq b \leq 1$, $1-a-b$ must range between values of 1 (when $a=0$ and $b=0$) and -1 (when $a=1$ and $b=1$). So, let us examine a few special cases of this observation.

<u>Case 1</u>: Where $a = 1$ and $b = 1$.

From (**), $P_t = (-1)^t (P_0 - P^*) + P^*$

where $P^* = \frac{a}{a+b} = \frac{1}{1+1} = \frac{1}{2}$, $t = 0, 1, 2, \ldots$

<u>Conclusion</u>: The sequence of values for P_t will oscillate between the values of P_0 (which we assumed to be prescribed) and $1-P_0$, over a period of t trials.

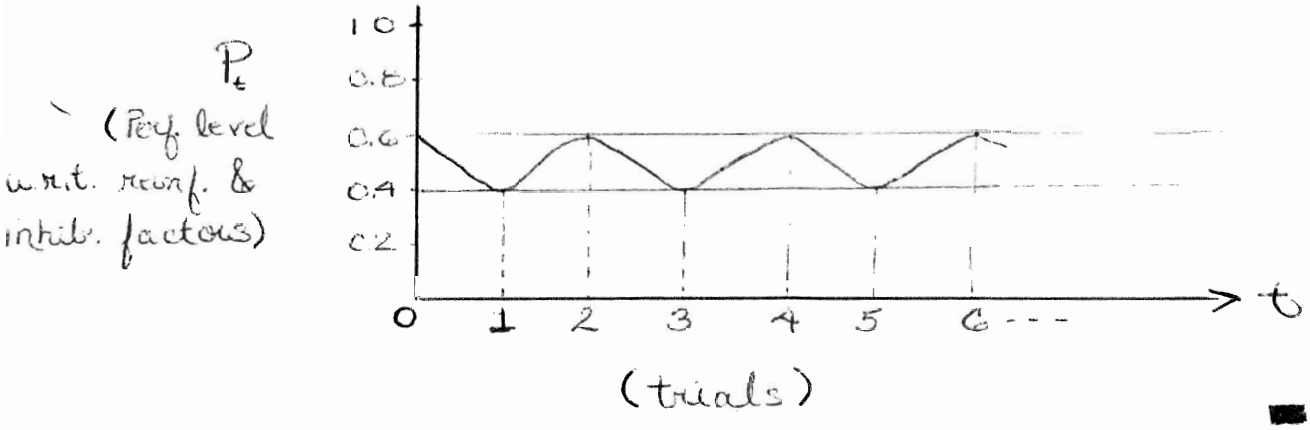


(eg) Let $P_0 = 0.6$, $a = 1$ and $b = 1$.

$$\Rightarrow P_t = (-1)^t (0.6 - 0.5) + 0.5$$
$$= (-1)^t (0.1) + 0.5 \quad \text{for } t = 0, 1, 2, \ldots$$

Hence, $\{P_t\} = \{0.6, 0.4, 0.6, 0.4, \ldots\}$

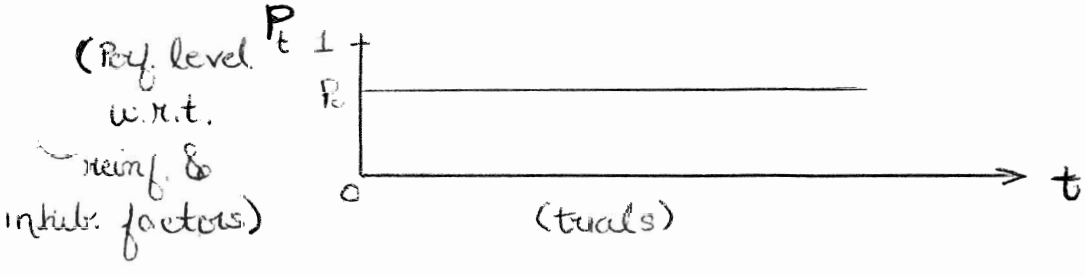
<u>interpretation?</u> — max reinforcement
max inhibition.



Case 2: Where $a = 0$ and $b = 0$

From (**), $P_t = P_0$, $t = 0, 1, 2, ---$

Conclusion: P_t will equal P_0 for trials of t.



interpretation?

Case 3: Where $a = 0$ and $b = 1$
or $a = 1$ and $b = 0$.

From (**), $P_t = P_0 = \begin{cases} 0 & \text{if } a = 0 \\ 1 & \text{if } b = 0 \end{cases}$; $t = 0, 1, 2, ---$

Conclusion: - Same as Case 2 -

Case 4: Where $0 < a < 1$ and $0 < b < 1$

From (**), $P_t = (1-a-b)^t (P_0 - P^*) + P^*$
where $P^* = \dfrac{a}{a+b}$, $t = 0, 1, 2, ---$

Here, $-1 < 1-a-b < 1$

$$\Rightarrow \lim_{t\to\infty} P_t = \lim_{t\to\infty}\left[(1-a-b)^t(P_0-P^*)+P^*\right] = P^*$$

not shown – fairly obvious but some explanation should be given

Conclusion:

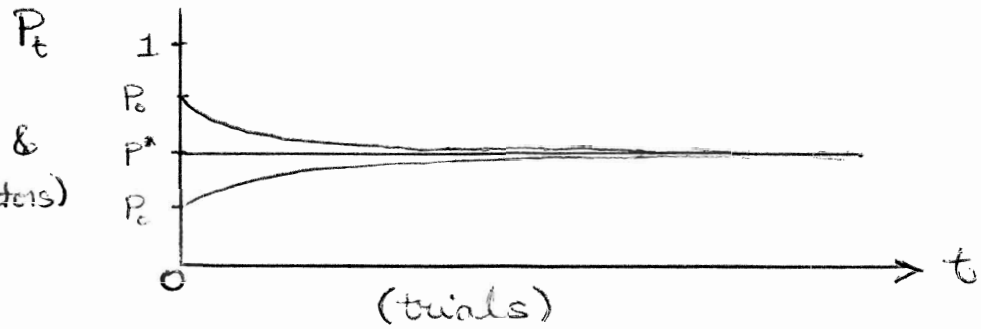
(a) $\{P_t\}$ is monotone increasing to P^* if $P_0 < P^*$,

(b) $\{P_t\}$ is monotone decreasing to P^* if $P_0 > P^*$,

not shown

illustrated below

(c) $\{P_t\}$ ~~is~~ remains constant if $P_0 = P^*$.



(eg.) Let $a = 0.3$ and $b = 0.2$.

$$\therefore P^* = \frac{a}{a+b} = \frac{0.3}{0.3+0.2} = 0.6$$

(i) Let $P_0 = 0.2$ $(P_0 < P^*)$

$$\therefore P_t = (1-0.3-0.2)^t(0.2-0.6)+0.6$$
$$= (0.5)^t(-0.4)+0.6$$

for $t = 0, 1, 2, \ldots$

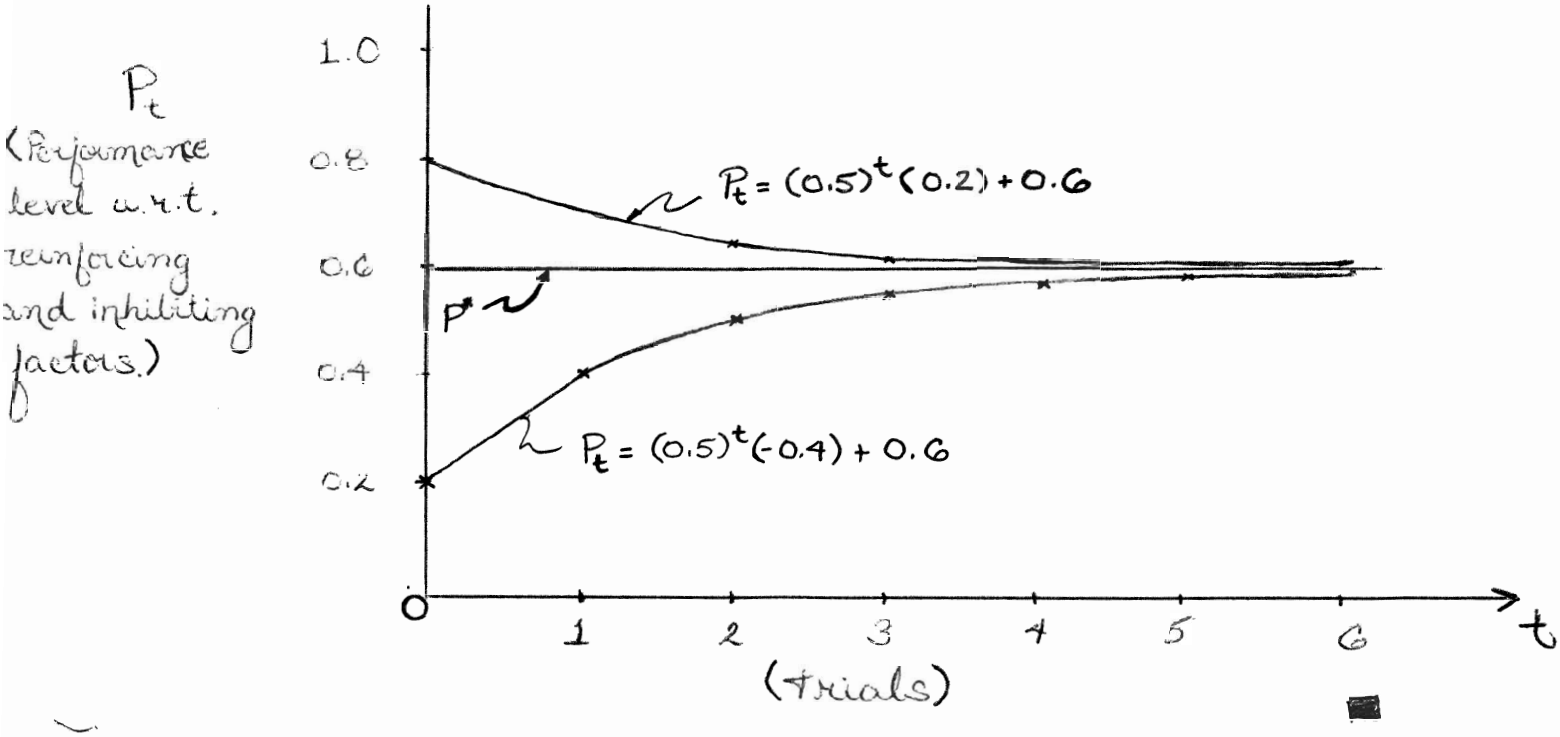
trials (t)	Performance level w.r.t reinforcing and inhibiting factors (P_t)
0	0.2000
1	0.4000
2	0.5000
3	0.5500
4	0.5750
5	0.5875
6	0.5934

(ii) Let $P_0 = 0.8$ $(P_0 > P^*)$

$$\therefore P_t = (1-0.3-0.2)^t(0.8-0.6)+0.6$$
$$= (0.5)^t(0.2)+0.6$$

for $t = 0, 1, 2, ---$

trials (t)	Performance level w.r.t. reinforcing and inhibiting factors (P_t).
0	0.8000
1	0.7000
2	0.6500
3	0.6250
4	0.6125
5	0.6063
6	0.6031


P_t
(Performance level w.r.t. reinforcing and inhibiting factors.)
1.0
0.8
0.6
0.4
0.2
$P_t = (0.5)^t(0.2) + 0.6$
P^*
$P_t = (0.5)^t(-0.4) + 0.6$
0
1
2
3
4
5
6
t
(trials)

Comments.

1. Pavlov's general conclusion does not compare with the results observed in either cases one, two, or three. However, we are not implying that Pavlov's conclusion is false, nor are we implying that our model is incorrect. It is simply that our mathematical analysis, in these three cases, is determining results from a set of unrealistic parameters.

In cases one, two, and three, it would be impossible to enforce any single combination of a maximum reinforcing or minimum reinforcing factor with a maximum inhibiting or minimum inhibiting factor. The reason is simply because it would be physically impossible to impose any such condition in any one trial.

For example, consider case one. Here, Pavlov would be required to carry out a maximum reinforcing and maximum inhibiting factor (since $a = 1$ and $b = 1$) in the experiment. In other words, Pavlov would have to, as a result of a salivary response, present food and feed the subject dog's, while simultaneously, as a result of a non-salivary response, present food but not feed the subject dog's. This would clearly be impossible because:

a) the subjects can only elicit one response to each stimuli in any one trial,

and b) Pavlov can only reply to each response with one conditioning factor in any one trial.

The same can be said in cases two and three where we again run into physically impossible situations.

Hence, we may eliminate any futher investigations associated with extream conditioning factors because of the physical implications they bring. Mathematically, however, we still must realize that these factors do exist.

2. Case four is capable of correctly modelling both the physical and mathematical nature of the experiment. In addition, notice the striking similarity between this case and the conclusion of the previous model. In both instances, there exists a gradual approachment to a carrying capacity. In our first model, that carrying capacity is the value one (which represents a precise response), and in our modified model, the carrying capacity is represented by P^*. P^*, however, is not simply restricted to equalling one, but instead is derived from other influences in the experiment itself; namely reinforcing and inhibiting factors. We thus are able to conclude that our modified model is an improvement over our initial model.

3 Recall that P_c is to be assumed to be prescribed in our modified model. So, the obvious question one may ask is "How do you go about solving that kind of value?"

<u>Solution</u> Let us once again realize that in case four, $P_t \to P^*$ as $t \to \infty$. Therefore, since (a) case four correctly represents both the physical and mathematical aspects in the experiment (as described in comment 2), and

(b) the results of case 4 clearly shows P_0 as having little significance in the determination of $\{P_t\}$ (especially if $t >> 0$), we can say that P_0 is actually an arbitrary value (as long as $0 \leq P_0 \leq 1$).

4. The next question one may ask is "How can you determine values for 'a' and 'b'?"
<u>Solution</u>: First recall that 'a' and 'b' are represented by:

<u>(1) the reinforcing factor</u>.

Tone → Dog Salivates → Dog is presented with food, then is fed.

and (2) <u>the inhibiting factor</u>:

Tone → Dog Does not Salivate → Dog is presented with food, but is not fed.

respectively.

Thus, our method of determining 'a' and 'b' will have to rely on the following relationship.

Stimuli → Response → Conditioning,

where we, the experimenters must evoke the stimuli, observe the response attributed to the stimuli, and then condition the subject accordingly.

Suggestion:

Let $a = \dfrac{\text{the number of adjacent trials resulting with a non-salivary then salivary response}}{\text{the number of trials}}$.

seems reasonable but what about $NS \to S \to S \to NS \to S \to S$?

Let $b = \dfrac{\text{the number of adjacent trials resulting with a salivary then non-salivary response}}{\text{the number of trials}}$

— here $a = \frac{2}{6}$, $b = \frac{1}{6}$ OK! — see below

The significance of this is

(1) $0 \le a \le 1$ and $0 \le b \le 1$ as required.

(2) Since we have already seen the important role that the Markov Property is playing in our modified model, why not take this property one step futher and study only adjacent trials that change from either a non-salivary to salivary response

or

salivary to non-salivary response.

(eg) Suppose we are required to determine the model parameters from the following data.

trial	Stimuli →	Response →	Conditioning
1	Tone	No Saliv.	PNF
2	Tone	No Saliv.	PNF
3	Tone	No Saliv.	PNF
4	Tone	Saliv.	PAF
5	Tone	No Saliv.	PNF
6	Tone	Saliv.	PAF
7	Tone	No Saliv.	PNF

8	Tone	Saliv.	PAF
9	Tone	Saliv.	PAF
10	Tone	Saliv.	PAF

where PAF denotes a presentation of food to the subjects and later feeding,
PNF denotes a presentation of food to the subjects but no feeding.

∴ no. of trials = 10
∴ no. of adjacent non-salivary → salivary responses = 3
∴ no. of adjacent salivary → non-salivary responses = 2.

$\Rightarrow a = \frac{3}{10} = 0.3$

$\Rightarrow b = \frac{2}{10} = 0.2$

<u>Note</u>

1) 'a' and 'b' would never both be equal to one (the equations thus eliminate case 1. ✓)

2) It would be unlikely for 'a' and 'b' to both be equal to zero as this would imply a single response throughout the entire experiment. (the equations thus "probably" eliminate case 2 ✓)

3) 'a' and 'b' would never equal zero and one nor one and zero respectively. (the equations thus eliminate case 3 ✓).

why

Suggestion for improvement.

As we have already seen, the "learning process" in both of our models is a gradual one which never actually realizes its full potential. Therefore, perhaps a future study could incorporate an "all or none" concept in the modelling procedure.

meaning?

References

1. Goldberg, S. (1958). A Probability Model for Learning. An Introduction To Difference Equations, 2.10, pp 103-106.

2. Roberts, F. S. (1976). Mathematical Models of Learning. Discrete Mathematical Models, 5.9, pp 331-333.

3. Wilson, D.W., & McGee, M.G. (1984). Classical and Instrumental Conditioning. Psychology Science and Application, 6, pp 133-135.