

A Mathematical Study of the Dynamics of Conscious Acquiring of Knowledge through Reading and Cramming and the Process of Losing Information from the Brain by Natural Forgetting of Facts

Sudipto Roy¹, Priyadarshi Majumdar²

¹Department of Physics, Calcutta Institute of Engineering and Management, Kolkata, India; ²Jyotinagar Bidyasree Niketan H.S. School, Kolkata, India.

Email: sr_ciem@rediffmail.com, majumdar_priyadarshi@yahoo.com

Received May 29th, 2010; revised June 17th, 2010; accepted August 23rd, 2010.

ABSTRACT

We model the conscious learning process of human brain with a dynamical equation (cramming dynamics) by considering both the data entry and loss of data simultaneously. We show the analytical solution of the differential equation in some special cases. We define some indexes like memory index, merit index, utilization index etc. Using them we can measure the corresponding memory functions. Applications of this model have also been discussed. More general numerical and analytical results are also presented at the end.

Keywords: Switching Function, Cramming Dynamics, Memory Index, Merit index, Utilization Index, Relative Performance Index

1. Introduction

The memory retention mechanism in human brain is quite a complicated process indeed. It is just like information processing for a huge database. This database is filled with all sorts of information that we use to go about our everyday lives. The information is stored and retrieved as needed. No matter what we are doing, at any instant or other, at anywhere, the memory is involved in an active fashion. This complex system or network of data processing is located in different parts of the brain like the hippocampus and cortex. As these parts work in tandem, memory begins to process and interacts with the environment and its surroundings.

After understanding how we process memory, why do we sometimes lose memories? It has been proven that we forget simply because of a problem with encoding, storage, retrieval, or a combination of any of these. The first significant study in this field was carried out by Ebbinghaus [1]. He studied the memorisation of nonsense syllables. By repeatedly testing himself after va-

rious time periods and recording the results, he was the first to describe the shape of the forgetting curve. On the other hand according to Eichenbaum [2], most forgetting occurs very soon after learning. However, when meaningful material is used, the forgetting curve is not so steep. Memory also fades and becomes less reliable with time and aging. An overflow of information may also cause certain information to be forgotten as a result of competition. Benfenati [3] in his work examined the cellular and molecular mechanisms that contribute to the various forms of memory including short and long-term memory, unconscious and conscious memory etc. Crovitz *et al.* [4] suggested that a memory measurement (M) can be expressed as a power function of time t as, $M = \beta t^{-\alpha}$. Anderson [5] mentioned that the experimental power function curves are related to the mean taken over all the subjects. Wixted [6] opposed the idea of Anderson by showing that the power function also fits better than the exponential function when data from the individual subjects are fitted. Other important models of the forgetting process are CHARM, due to Metcalfe [7],

Chappell [8], Matrix Model, due to Humphreys [9] and MINERVA II due to Hintzman [10]. The neural network model predicted by Hopfield [11] was latter modified by sikstrom [12] using the concept of bounded weights and a distribution of learning rates. In a recent communication Stepanov [13] proposed a new model of memorization dynamics using exponential functions.

In our work we have attempted to build up a physical model of the memorizing process of the human brain (of a learner in particular) undergoing a course of study with a fixed duration of time. The act of learning is considered here as a process of data storage in the brain. It is assumed that one accumulates data while studying a subject consciously and there is a continuous process of data loss, caused by several physiological and psychological factors such as mental stress and fatigue etc. As this memorizing and forgetting processes are continuous in the said time interval hence we can express this process as a dynamical equation to be explained in more details in the section to follow. The dynamical equation involves different parameters quantifying the capacity of the brain from different aspects e.g. memorization ability, grasping power etc. The solution to that differential equation of learning process gives us a clear picture of how data are being stored and lost continuously from the brain. We have obtained the analytical solution of the said equation in some particular limits and the numerical solution has also been obtained as a result of system simulation using MATLAB in Intel platform. The corresponding results are mentioned in the sections to follow.

The idea behind choosing this particular model of the brain and the basic assumptions are very simple and based upon our common experiences. We have assumed that the rate of data storage (accumulation rate) in the brain at any instant can be calculated simply by subtracting the rate of data loss from the rate of data entry. Obviously the actual dynamics of the brain may not be exactly the same as we have assumed but for all practical purposes our work can mimic the activity of the brain up to a certain extent that will be cleared from our next analysis. To support this claim we have shown that the work of Ebbinghaus [1] will come as a special case of our model. Also none of the previous models as mentioned earlier are as so much simple like ours and can give a complete mathematical description of both the learning and forgetting processes simultaneously using the concept of switching process of the brain.

2. Mathematical Modeling

Let R_S , R_L and R_E be the rate of data storage, the rate of data loss and the rate of data entry in the brain at any instant respectively. Let $x(t)$ be the amount of data or information already stored in the brain at any time t , hence, the rate of storage at that moment is given by

$$R_S = \frac{dx}{dt} = R_E - R_L \quad (1)$$

The rate of data entry can be enhanced by the factors like intelligence, concentration, the ability of a person to cope with the stressed situations etc. Experience tells us that as we go on acquiring more and more knowledge and thereby store more and more data, the rate of data entry becomes slower due to some mental stress or brain fatigue, as we generally perceive. As the accumulated data increases in volume in the brain, the rate of data entry must decrease. Hence, to give it a very simple mathematical form, one can safely assume that at any point of the learning process we have

$$R_E = S(t) R_E^{on} \quad (2)$$

where $S(t)$ is a time dependent switching function (to be discussed elaborately later on) whose value toggles between 0 and 1 for its *OFF* and *ON* stages respectively. During a conscious effort of cramming, this switch remains *ON* and otherwise it is *OFF*. R_E^{on} is actually the rate of data entry for the *ON* state of the switch $S(t)$. Thus, during the *ON* state, we have $R_E = R_E^{on}$. During the *OFF* state of $S(t)$ we have $R_E = 0$. We define R_E^{on} as Equation (3).

Here, C_1 quantifies one's intelligence, concentration, eagerness, urgency of learning etc. Thus, it is a measure of the traits of the learner that causes faster entry of data. C denotes the maximum storage capacity, hence $x \leq C$. The quantity x/C is the fraction of memory occupied by information and therefore $(1 - x/C)$ is the fraction of storage space still available for data entry. Common experience tells us that, larger the value of x/C , greater will be the difficulty in the further storage of data. The parameter α is a positive quantity that may be called the *index of brain fatigue* related to memorization (taking care of the non-linearity). Since $\left(1 - \frac{x}{C}\right)$ is generally a fraction, R_E^{on} is smaller for larger values of α .

$$R_E^{on} = \frac{\text{measure of grasping power, concentration, IQ and urgency of learning}}{\text{measure of factors impeding accumulation of data}} = \frac{C_1}{\left[\frac{1}{1 - x/C}\right]^\alpha} = C_1 \left(1 - \frac{x}{C}\right)^\alpha \quad (3)$$

It is a common experience that one forgets information more rapidly when the amount of accumulated data is large. In mathematical terms, the larger the value of x/C , greater will be the amount of data loss per unit time. As the storage becomes higher the rate of loss becomes more and more pronounced, possibly due to the limitation of retention ability and the stress caused by the load of already accumulated data. One may simply write an expression for rate of data loss (R_L), at any stage of learning, as a function of the data (x) already stored in the brain in the following way

$$R_L = \frac{\text{stress caused by data storage or data load}}{\text{memory retention ability}} = \frac{(x/C)^\beta}{C_2} \quad (4)$$

The parameter C_2 is a measure of one's ability of memory retention or memorizing ability. The term $(x/C)^\beta$ may be regarded as a measure of stress caused by the accumulated data where the parameter β is introduced to take care of the natural non-linearity of the process. Like α , it is also a positive quantity. Since (x/C) is a fraction, R_L is smaller for higher values of β . The parameter β may be called the *stress endurance index*. A person with a larger value of β feels less stressed by the accumulated data (*i.e.* less internal anxiety about the necessity of retention of accumulated data). With substitutions in (1), from (2, 3) and (4) we have

$$\frac{dx}{dt} = S(t)C_1 \left(1 - \frac{x}{C}\right)^\alpha - \frac{(x/C)^\beta}{C_2} \quad (5)$$

Let us define a dimensionless variable as $X = x/C$ (with X varying from 0 to 1). Hence, in terms of X , (5) can be expressed as

$$\frac{dX}{dt} = \frac{S(t)C_1}{C} (1-X)^\alpha - \frac{X^\beta}{CC_2} \quad (6)$$

3. Analysis

For any arbitrary learner the parameters C_1 and C_2 are defined below

$$C_1 = f_1 C_1^{\max}, \text{ where } 0 < f_1 \leq 1, \quad (7)$$

$$C_2 = f_2 C_2^{\max}, \text{ where } 0 < f_2 \leq 1, \quad (8)$$

Here, C_1^{\max} and C_2^{\max} are the values of C_1 and C_2 for the best possible learner (ideal, being the most intelligent, enthusiastic, diligent and having the strongest memory and greatest zeal for learning). Here, we define the dimensionless parameters f_1 and f_2 as the *merit index* and the *memory index* respectively of any arbitrary learner, quantifying one's IQ and MRA (memory retention ability) respectively, relative to the best learner. For

the learner of the highest calibre or merit we have $f_1 = f_2 = 1$. Using (7) and (8) in (6) we obtain

$$\frac{dX}{dt} = \frac{S(t)f_1 C_1^{\max}}{C} (1-X)^\alpha - \frac{X^\beta}{C f_2 C_2^{\max}} \quad (9)$$

For convenience of calculation (without any loss of generality) we may choose $C_1^{\max} = C_2^{\max} = 1$. If a system of units can be defined for these quantities, they can always be chosen to satisfy this equality. Hence (9) can be expressed as

$$\frac{dX}{dt} = \frac{S(t)f_1}{C} (1-X)^\alpha - \frac{X^\beta}{C f_2} \quad (10)$$

To determine the variation of X as a function of time, we need to solve (10) for different functional forms of $S(t)$ (**Figures 1-3**). Where $S(t)$ is a function determining the duration for which the data entry channel remains open.

As long as one maintains a conscious learning effort, $S(t)$ remains 1, zero otherwise. During the time when $S(t) = 0$ we have from (10)

$$\frac{dX}{dt} = - \frac{X^\beta}{C f_2} \quad (11)$$

Let T_m be the duration for which one maintains a conscious memorizing effort without any break. Hence we may write

$$S(t) = 1 \text{ for } 0 \leq t \leq T_m \quad (12)$$

$$S(t) = 0 \text{ for } t > T_m \quad (13)$$

The above criteria can very well be approximated by a tan-hyperbolic function as given below

$$S(t) = 1 - \frac{1}{2} [\tanh \varepsilon(t - T_m) + 1] \quad (14)$$

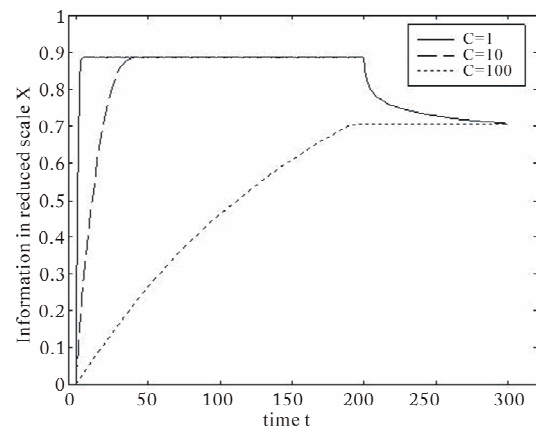


Figure 1. Variation of X with t for three different values of C , with $f_1 = 0.6, f_2 = 0.6, \varepsilon = 10^6, T_m = 200, \alpha = 0.9, \beta = 25$ and $S(t)$ follows (14).

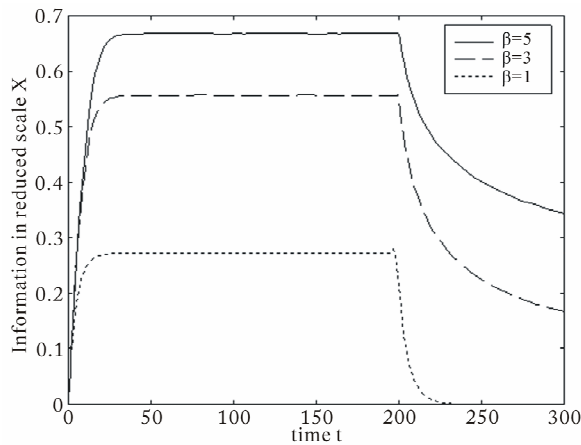


Figure 2. Variation of X with t for three different β , with $f_1 = 0.6, f_2 = 0.6, \varepsilon = 10^6, T_m = 200, \alpha = 0.9, C = 10$ and $S(t)$ follows (14).

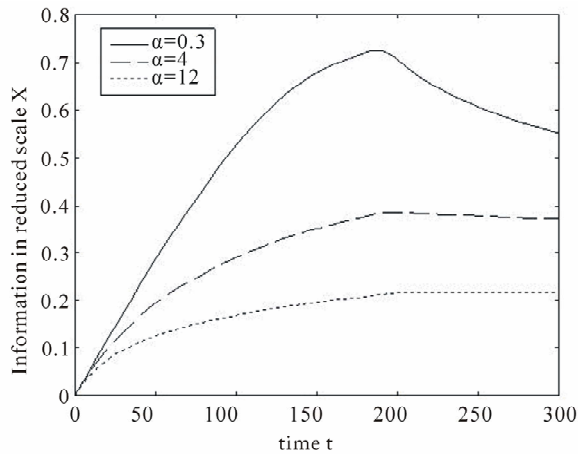


Figure 3. Variation of X with t for three different α , with $f_1 = 0.6, f_2 = 0.6, \varepsilon = 10^6, T_m = 200, C = 100, \beta = 5$ and $S(t)$ follows (14).

For a sufficiently large positive value of ε , this function behaves almost exactly like (12, 13).

There may be another situation where the data entry channel remains open intermittently. The learner can alternately open and close the channel in a periodic fashion. Let T be the interval of time over which it remains open and, for the next phase of the same duration it remains closed. In this case, the plot of $S(t)$ vs. t should have the pattern of a square wave, varying between 0 and 1. For numerical calculations in this case, $S(t)$ can be approximated by a function of the following form (the case is illustrated in **Figure 4**).

$$S(t) = \frac{1}{2} \left[\tanh \varepsilon \left(\sin \frac{\pi t}{T_m} \right) + 1 \right] \quad (15)$$

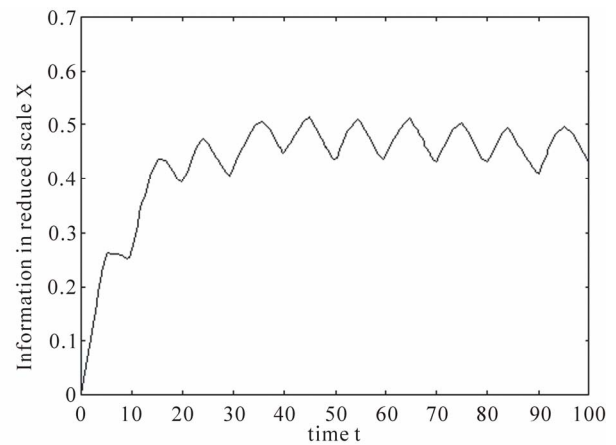


Figure 4. X - t variation with $f_1 = 0.6, f_2 = 0.6, \varepsilon = 10^6, T_m = 5, C = 10, \beta = 3, \alpha = 0.9$ and $S(t)$ follows (15).

The sharpness of the square wave pattern of $S(t)$ increases with higher values of ε . For a sufficiently large positive value of ε , this function behaves in a manner such that

$$S(t) = 1 \text{ for } 2nT_m \leq t \leq (2n+1)T_m, \quad (16)$$

$$n = 0, 1, 2, 3, \dots$$

$$S(t) = 0 \text{ for } (2n+1)T_m < t < (2n+2)T_m, \quad (17)$$

$$n = 0, 1, 2, 3, \dots$$

4. Analytical Solution for a Particular Case

For $\alpha = \beta = 1$, both growth and decay processes mentioned earlier are exponential in nature (as suggested by (10, 11)), similar to that obtained by Ebbinghaus [1]. Under the boundary condition that, at $t = 0, X = X_0$, we have from (10)

$$X = \frac{f_1 f_2}{1 + f_1 f_2} - \left(\frac{f_1 f_2}{1 + f_1 f_2} - X_0 \right) \exp \left(-\frac{t}{\tau} \right), \quad (18)$$

$$\text{With } \tau = \frac{C f_2}{1 + f_1 f_2}$$

It suggests that as $t \rightarrow \infty, X \rightarrow \frac{f_1 f_2}{1 + f_1 f_2}$. Hence we

define $X_{\max} = \frac{f_1 f_2}{1 + f_1 f_2}$ and $x_{\max} = C X_{\max} = \frac{C f_1 f_2}{1 + f_1 f_2}$, as the saturation values of X and x respectively. For the best learner, having $f_1 = f_2 = 1$, these quantities can be expressed as $X_{\max}^b = \frac{1}{2}$ and $x_{\max}^b = C X_{\max}^b = \frac{C}{2}$. Here,

$\tau = \frac{C f_2}{1 + f_1 f_2}$ is the characteristic time constant determining the rapidity with which X reaches its saturation

value X_{\max} . For smaller values of τ , X varies at a faster rate with time.

For $f_1 = f_2 = 1$, we have $\tau_{\max} = \frac{C}{2}$. The duration (T_m) of the interval, for which one continues a conscious cramming effort, may be expressed as a multiple of τ_{\max} as $T_m = N\tau_{\max} = \frac{NC}{2}$. Hence,

$$\tau = \frac{2 T_m f_2}{N(1 + f_1 f_2)} \quad (19)$$

Let us now define a dimensionless variable n such that

$$t = nT_m \quad (20)$$

Combining (19) and (20) we obtain

$$\frac{t}{\tau} = \frac{nN(1 + f_1 f_2)}{2f_2} \quad (21)$$

Incorporating all these results in (18) we get the following expression of X representing the data storage process up to the time of $t = T_m$ we are getting

$$X(t \leq T_m) = X_{\max} - (X_{\max} - X_0) \exp\left(-\frac{nNf_1}{2X_{\max}}\right) \quad (22)$$

While studying the behaviour of X beyond $t = T_m$, we need to solve (10) for $\beta = 1$ and $S(t) = 0$ under the boundary condition that, at $t = T_m$, $X = X_M$ where X_M is the value of $X(t \leq T_m)$ at $t = T_m$ or equivalently at $n = 1$. The corresponding solution holds only for $t \geq T_m$ and is given by

$$X(t \geq T_m) = X_M \exp\left(\frac{T_m - t}{Cf_2}\right) \quad (23)$$

X_M can be determined from (18) by using $t = T_m$. $= NC/2$. Then, by using (20) the above expression takes the form of Equation (24).

It would be reasonable to express X_0 as a fraction of X_{\max} for any learner. Taking into account the expression of X_{\max} we may have

$$X_0 = \lambda X_{\max} = \frac{\lambda f_1 f_2}{1 + f_1 f_2}, \text{ with } 0 \leq \lambda < 1 \quad (25)$$

Now using (24) and (25) we obtain

$$X(t \geq T_m) = X_{\max} \left[1 - (1 - \lambda) \exp\left(-\frac{Nf_1}{2X_{\max}}\right) \right] \exp\left[\frac{N(1 - n)}{2f_2}\right] \quad (26)$$

The behavior of X , as a function of n , is obtained from (22) and (26) for the ranges $0 \leq n \leq 1$ and $n > 1$ respectively (see **Figures 5-7**).

5. Applications

Let us now consider the learning behaviour for a group of students, preparing for a certain examination process. The time allotted for preparation before the examination is T_1 . Let X_h be the information gathered by the best learner. Hence using $f_1 = f_2 = 1$, $t = T_1$ and $\tau = \tau_{\max}$ ($= C/2$) again.

$$X_h = \frac{1}{2} - \left(\frac{1}{2} - X_0\right) \exp\left(-\frac{T_1}{\tau_{\max}}\right) \quad (27)$$

A student may make some delay while starting the process of learning. Let μT_1 be the time utilized by any arbitrary learner where $0 \leq \mu \leq 1$. The delay in starting the learning process is $(1 - \mu)T_1$. Let us define μ as the *utilization index*. For the most sincere student $\mu = 1$. Let X_a be the amount of knowledge acquired by any arbitrary learner before the commencement of examination.

Thus for $X = X_a$ at $t = \mu T_1$ with $\tau = \frac{Cf_2}{1 + f_1 f_2}$ we have

$$X_a = \frac{f_1 f_2}{1 + f_1 f_2} - \left(\frac{f_1 f_2}{1 + f_1 f_2} - X_0\right) \exp\left(-\frac{\mu T_1}{\tau}\right) \quad (28)$$

The performance of any arbitrary student in an examination, relative to the best learner, can be defined as

$$P_r = \frac{X_a}{X_h} = \frac{\frac{f_1 f_2}{1 + f_1 f_2} - \left(\frac{f_1 f_2}{1 + f_1 f_2} - X_0\right) \exp\left(-\frac{\mu T_1}{\tau}\right)}{\frac{1}{2} - \left(\frac{1}{2} - X_0\right) \exp\left(-\frac{T_1}{\tau_{\max}}\right)} \quad (29)$$

Let us call P_r the *relative performance index*. We can always express T_1 as a multiple of τ_{\max} as

$$T_1 = \eta \tau_{\max} = \frac{\eta C}{2}, \text{ with } \eta > 0 \quad (30)$$

Assuming $X_0 = 0$ and substituting for T_1 and τ we get see (**Figures 8-13**)

$$X(t \geq T_m) = \left[\frac{f_1 f_2}{1 + f_1 f_2} - \left(\frac{f_1 f_2}{1 + f_1 f_2} - X_0\right) \exp\left(-\frac{N(1 + f_1 f_2)}{2f_2}\right) \right] \exp\left[\frac{N(1 - n)}{2f_2}\right]. \quad (24)$$

$$P_r = \frac{X_a}{X_h} = \frac{\frac{2f_1f_2}{1+f_1f_2} - \left(\frac{2f_1f_2}{1+f_1f_2}\right) \exp\left[-\frac{\mu \eta (1+f_1f_2)}{2f_2}\right]}{1 - \exp(-\eta)} \quad (31)$$

Defining $K = \frac{2f_1f_2}{1+f_1f_2}$ we obtain

$$P_r = \frac{K \left[1 - \exp\left(-\frac{\mu \eta f_1}{K}\right)\right]}{1 - \exp(-\eta)} \quad (32)$$

When the time available for study is sufficiently long, for a sincere learner (with $\mu = 1$), the exponential terms in the last equation become negligible and it will reduce to

$$P_r = K = \frac{2f_1f_2}{1+f_1f_2} \quad (33)$$

Using experimental results one can determine C and f_2 .

Let us suppose somebody performs two tests of

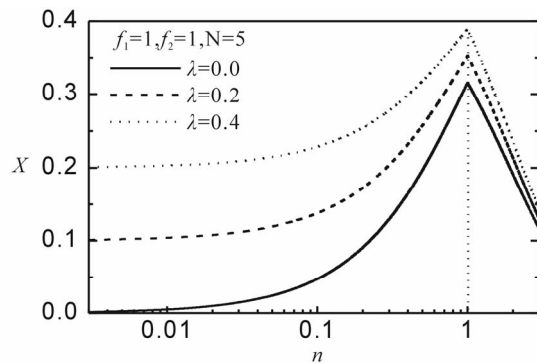


Figure 5. Graphical representation of (22), (23) and (24) with parametric variations.

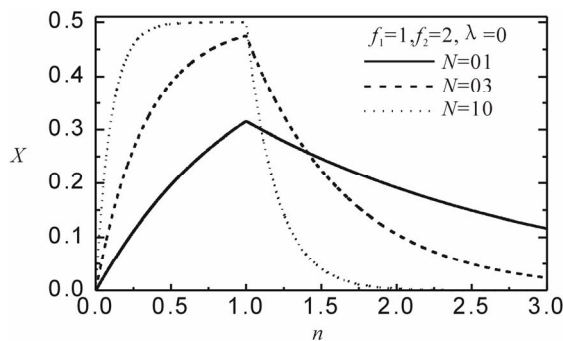


Figure 6. Graphical representation of (22), (23) and (24) with parametric variations.

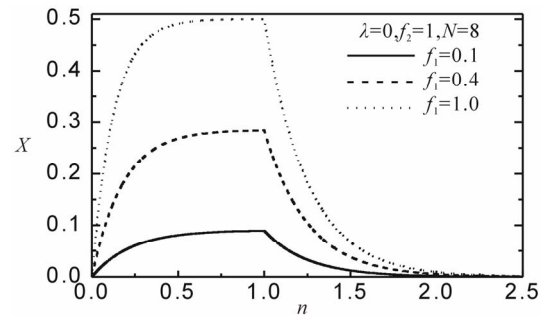


Figure 7. Graphical representation of (22), (23) and (24) with parametric variations.

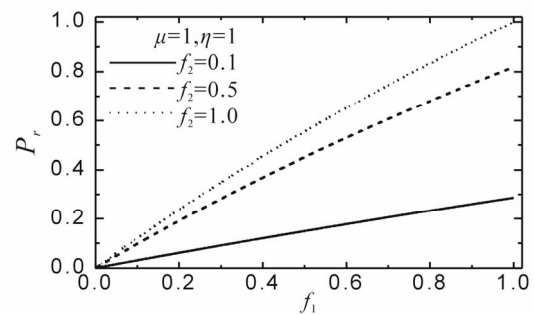


Figure 8. $P_r - f_1$ variation for three different f_2 , drawn on the basis of (31)

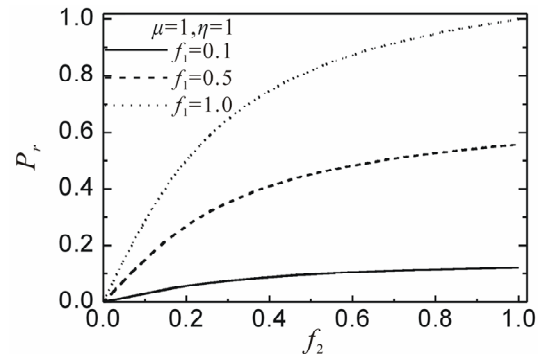


Figure 9. $P_r - f_2$ variation for three different f_1 , drawn on the basis of (31).

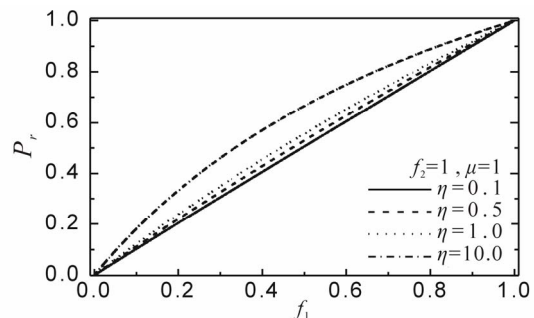


Figure 10. $P_r - f_1$ variation for three different η , drawn on the basis of (31).

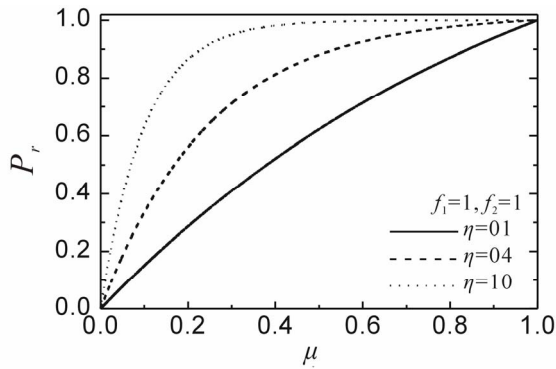


Figure 11. $P_r - \mu$ variation for three different η , drawn on the basis of (31).

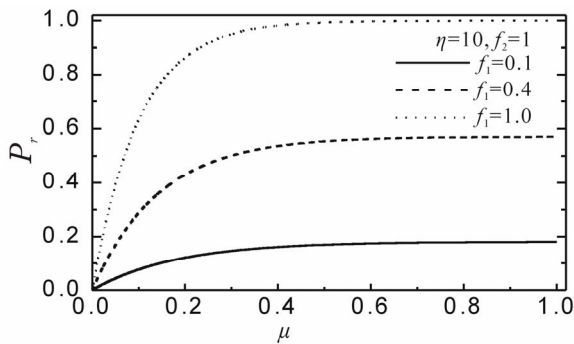


Figure 12. $P_r - \mu$ variation for three different f_1 , drawn on the basis of (31).

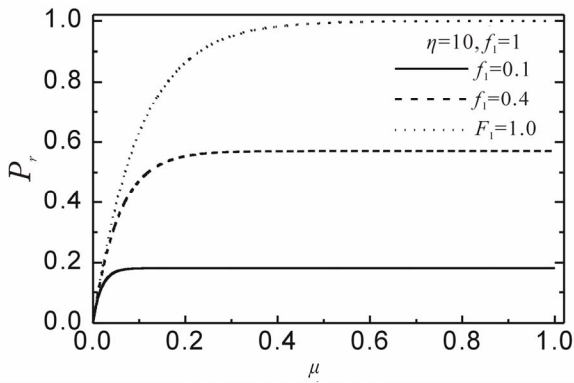


Figure 13. $P_r - \mu$ variation for three different f_2 , drawn on the basis of (31).

knowledge at $t = T$ and $t = 2T$, resulting in $X = X_T$ and $X = X_{2T}$ respectively. One can determine the ratio X_{2T} / X_T as

$$\ln(X_{2T} / X_T) = -T / C f_2 \quad (34)$$

For the best learner (with $f_2 = 1$), let the corresponding X 's be X_T^b and X_{2T}^b respectively. Hence

$$\ln(X_{2T}^b / X_T^b) = -T / C \quad (35)$$

While constructing (34) and (35) the underlying assumption is that, the maximum data storage capacity C of the brain does not vary from person to person. It is a constant for all human beings in the adulthood. Again as

$$f_2 = \frac{\ln(X_{2T}^b / X_T^b)}{\ln(X_{2T} / X_T)} \quad (36)$$

One can determine the *memory index* (f_2) of any learner using the grade points scored by him and by the best learner also, in the two examinations held at $t = T$ and $t = 2T$. Now from (35)

$$C = -\frac{T}{\ln(X_{2T}^b / X_T^b)} = \frac{T}{\ln X_T^b - \ln X_{2T}^b} \quad (37)$$

Using (33) and (36) we get

$$f_1 = \frac{P_r \ln(X_{2T} / X_T)}{(2 - P_r) \ln(X_{2T}^b / X_T^b)} \quad (38)$$

One can determine the *merit index* f_1 of any individual from the grade points scored by him and the best learner in two examination held at $t = T$ and $t = 2T$.

The validity of the above set of equations are ensured if the data entry channel remains closed during the time between $t = T$ and $t = 2T$. It means that these equations hold good only if there is no conscious learning effort during this interval of time, on the part of both learners.

6. More Generalised Analysis

The common experience of learning tells us that if we sustain the learning process (by keeping $S(t) = 1$) for an indefinite period of time, the memory X will continue to increase but with a gradually decreasing rate. The value of X will asymptotically approach a definite saturation level (X_{\max}), which is likely to depend on parameters ($f_1, f_2, C, \alpha, \beta$) of the model given in (10).

Therefore, at $S(t) = 1$ and $X = X_{\max}$, (10) is expected to show the following behavior

$$\frac{dX}{dt} = \frac{f_1}{C} (1 - X_{\max})^\alpha - \frac{X_{\max}^\beta}{C f_2} = 0 \quad (39)$$

Hence

$$(f_1 f_2)^{-1/\alpha} X_{\max}^{\beta/\alpha} + X_{\max} - 1 = 0 \quad (40)$$

The parameter $f_1 f_2$ is generally a fraction. Hence we can write $f_1 f_2 = 1/\kappa$ where $\kappa > 1$. Substituting this in (40) we have

$$(\kappa)^{1/\alpha} X_{\max}^{\beta/\alpha} + X_{\max} - 1 = 0 \quad (41)$$

The real solutions of this polynomial are the intersections of the curve $y = (\kappa X_{\max}^{\beta})^{1/\alpha}$ with straight line $y = 1 - X_{\max}$. For $X_{\max} = 1$, the curve will never have $y = 0$ as obtained for the straight line. Thus, we always have $X_{\max} < 1$ which means that a learner will always have $x < C$. Theoretically, the highest possible value of $X (\equiv x/C)$ is 1. Here we have to consider the particular real and positive X_{\max} (we definitely have some positive roots as all the parameters are positive) which is less than 1. For an individual learner, it is always desirable to have a saturation value (X_{\max}) that is as close to 1 as possible. To ensure it, the curve $y = (\kappa X_{\max}^{\beta})^{1/\alpha}$ must be very flat. Since $0 < X_{\max} < 1$, for a sufficiently large value of β we must have $\kappa X_{\max}^{\beta} < 1$. Hence, for this large value of β , one can increase the flatness of the curve by reducing the value of α . From the definitions of α and β it is clear that, persons with larger values of the ratio β/α must make better learning performances. Here we see that larger values of this ratio increase the flatness of the curve $y = (\kappa X_{\max}^{\beta})^{1/\alpha}$, making X_{\max} closer to 1. In **Figure 14**, we have a plot of X_{\max} as a function of β/α for three different values of β . Any of these curves shows that as the ratio β/α increases, X_{\max} increases. At any fixed value of β/α , the value of X_{\max} increases as β increases.

It is important to calculate the time required for reaching the saturation level. We know that, as $t \rightarrow \infty$, we have $X \rightarrow X_{\max}$, where X_{\max} is the saturation value. One can calculate the time required to reach a certain fraction of X_{\max} . Let us define *optimum time* (T_{opt}) as the time required to reach 99% of the saturation value. Putting $S(t) = 1$ in (10) we find an analytical expression for T_{opt} as

$$T_{opt} = Cf_2 \int_{x_{in}}^{0.99X_{\max}} \frac{dX}{f_1 f_2 (1-X)^{\alpha} - X^{\beta}} \quad (42)$$

Here, X_{in} indicates the initial amount of memory at $t = 0$. (42) shows that the integrand diminishes for smaller values of α and larger values of β since $0 < X < 1$. Therefore, the ratio β/α may be expected to play a significant role here. This integrand also decreases for any increase in the value of the quantity $f_1 f_2$. For a definite set of numerical values of the model-parameters we can estimate T_{opt} numerically from (42) and can compare it with T_m . For any learner it is desirable to maximize X_{\max} and minimize T_{opt} .

Let us consider a situation that allows one to continue the conscious learning process up to the time of $t = T_m$.

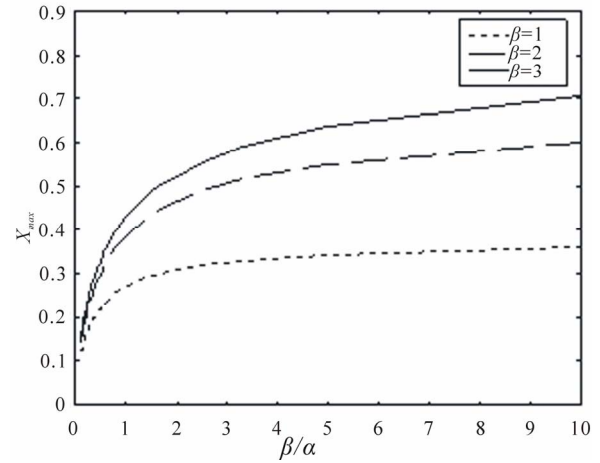


Figure 14. Variation of X_{\max} as a function of β/α for different values of β for $C = 10$, $f_1 = 0.6$, $f_2 = 0.6$, $\varepsilon = 1000000$, $T = 200$, $X_0 = 0$, $S(t) = 1$.

Thus, from $t = T_m$ onwards we have $S(t) = 0$. Here, T_m may be the time given for preparation before any test of learning. Comparing T_m with T_{opt} , let us discuss different possible cases below.

a) $T_m > T_{opt}$: In this case the learner gets the perception of reaching the saturation level before having to stop the learning process.

b) $T_m \approx T_{opt}$: Here, the learner just reaches the saturation level at the end of the time interval given for learning. It is worse compared to the last case.

c) $T_m < T_{opt}$: This situation is the most undesirable one. Here the learner has to stop the process before attaining the saturation level of learning.

To explore the role played by the ratio β/α , we may define average learning speed (S_{av}) as follows

$$S_{av} = \frac{0.99X_{\max}}{T_{opt}} \quad (43)$$

In **Figure 15**, we have a plot of S_{av} as a function of the ratio β/α for three different values of β . Each curve has an increasing trend with a gradually decreasing slope. Although these curves have intersections with each other at lower values of β/α , at higher values of this ratio, S_{av} becomes larger for greater values of β . Let us now analyze the situation after one stop the learning process at $t = T_m$. The system now follows the (11). Since we have already discussed cases with $\beta = 1$ earlier, we should now consider the cases with $\beta \neq 1$. Considering the boundary condition that $X = X_k$ at $t = T_m$, we may write

$$X^{1-\beta} = (1-\beta) \left(\frac{T_m - t}{Cf_2} + \frac{X_k^{1-\beta}}{1-\beta} \right) \quad (44)$$

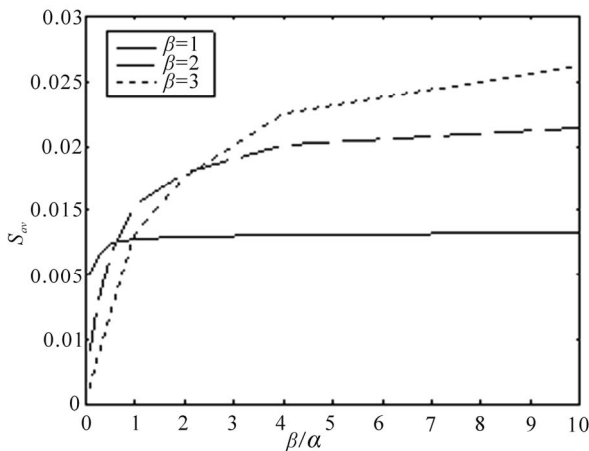


Figure 15. Variation of average learning speed (S_{av}) as a function of β/α for different values of β for $C = 10$, $f_1 = 0.6$, $f_2 = 0.6$, $\varepsilon = 1000000$, $T = 200$, $X_0 = 0$, $S(t) = 1$.

Here we have to take into account the physical reality that, for $t > T_m$ we must always have $X \geq 0$ and therefore $X^{1-\beta} \geq 0$. (44) clearly shows that, $X^{1-\beta} \geq 0$ at any point of time beyond $t = T_m$ only if $\beta > 1$. Thus, the present study shows that, the model is logically acceptable only for $\beta \geq 1$. Unlike exponential decay, which was obtained for the case with $\beta = 1$, we have the following solution for $\beta > 1$

$$X = \left[(1-\beta) \left(\frac{T_m - t}{Cf_2} + \frac{X_k^{1-\beta}}{1-\beta} \right) \right]^{1/(1-\beta)} \quad (45)$$

(45) suggests that X decreases with time but it remains greater than zero (since $t > T_m$ and $\beta > 1$). We never lose our entire memory in the decay process.

Common experience tells us that when one is engaged in a conscious learning process (*i.e.* $S(t) = 1$), we must have $dx/dt \geq 0$. Then, from (10) we can write

$$1 - X \geq (\kappa X^\beta)^{1/\alpha}, \quad \text{where } \kappa = 1/f_1 f_2 \quad (46)$$

The values of α and β should be such that (46) is satisfied for the highest possible values of κ and X .

7. Conclusions

We have proposed and analyzed a simple model of learning process. Some numerical results including simulations are also presented. Learning and memorizing are two most essential features of human brain. The model can further be improved by considering more complicated growth and decay process. These can be achieved if we express the features of memorizing process as men-

tioned in (3) and (4), e.g. grasping power, stress, IQ etc. separately in clear mathematical form with proper explanations.

REFERENCES

- [1] H. Ebbinghaus, "Memory: A Contribution to Experimental Psychology," translated by H. A. Ruger and C. E. Bussenius, published by Teachers College, Columbia University, New York City, 1913.
- [2] H. Eichenbaum, "Conscious Awareness, Memory and the Hippocampus," *Nature Neuroscience*, Vol. 2, No. 9, 1999, pp. 775-776.
- [3] F. Benfenati, "Synaptic Plasticity and the Neurobiology of Learning and Memory," *Acta Biomed*, Vol. 78, Suppl. 1, 2007, pp. 58-66.
- [4] H. F. Crovitz and H. Schiffman, "Frequency of Episodic Memories as a Function of their Age," *Bulletin of the Psychonomic Society*, Vol. 4, No. 5B, 1974, pp. 517-518.
- [5] R. B. Anderson and R. D. Tweney, "Artifactual Power Curves in Forgetting," *Memory and Cognition*, Vol. 25, No. 5, 1997, pp. 724-730.
- [6] J. T. Wixted and E. B. Ebbesen, "Genuine Power Curves in Forgetting: A Quantitative Analysis of Individual Subject Forgetting Functions," *Memory and Cognition*, Vol. 25, No. 5, 1997, pp. 731-739.
- [7] J. Metcalfe, "Recognition Failure and the Composite Memory Trace in Charm," *Psychological Review*, Vol. 98, No. 4, 1991, pp. 529-553.
- [8] M. Chappell and M. S. Humphreys, "An Auto-Associative Neural Network for Sparse Representations: Analysis and Application to Models of Recognition and Cued Recall," *Psychological Review*, Vol. 101, No. 1, 1994, pp. 103-128.
- [9] M. S. Humphreys, J. D. Bain and R. Pike, "Different ways to Cue a Coherent Memory System: A Theory for Episodic, Semantic and Procedural Tasks," *Psychological Review*, Vol. 96, No. 2, 1989, pp. 208-233.
- [10] D. L. Hintzman, "Judgments of Frequency and Recognition Memory in a Multiple-Trace Memory Model," *Psychological Review*, Vol. 95, No. 4, 1988, pp. 528-551.
- [11] J. J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities," *Proceedings of the National Academy of Sciences*, Vol. 79, No. 8, 1982, pp. 2554-2558.
- [12] S. Sikstrom, "Power Function Forgetting Curves as an Emergent Property of Biologically Plausible Neural Network Models," *International Journal of Psychology*, Vol. 34, No. 5-6, 1999, 460-464.
- [13] I. I. Stepanov and C. I. Abramson, "A New Mathematical Model for Assessment of Memorization Dynamics," *The Spanish Journal of Psychology*, Vol. 8, No. 2, 2005, pp. 142-156.