

Research Article

A Simple Mathematical Model of Cyclic Circadian Learning

J. Šimon¹ and M. Bulko²

¹ Department of Applied Mathematics, Faculty of Mechanical Engineering, University of Žilina, Univerzitná 8215/1, 010 26 Žilina, Slovakia

² Department of Nuclear Physics and Biophysics, Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, Slovakia

Correspondence should be addressed to J. Šimon; jan.simon@fstroj.uniza.sk

Received 7 November 2013; Accepted 30 January 2014; Published 18 March 2014

Academic Editor: Sabri Arik

Copyright © 2014 J. Šimon and M. Bulko. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper deals with the derivation of a simple mathematical model of cyclic learning with a period of 24 hours. Various requirements are met with an emphasis and approach which relies on simple mathematical operations, the prediction of measurable quantities, and the creation of uncomplicated processes of calibration. The presented model can be used to answer questions such as the following. *Will I be able to memorize a given set of information? How long will it take to memorize information? How long will I remember the information that was memorized?* The model is based on known memory retention functions that are in good agreement with experimental results. By the use of these functions and by formalism of differential equations, the concurrent processes of learning and forgetting are described mathematically. The usability of this model is limited to scenarios where logical bonds (connections to prior learning) are not created and mnemonic devices cannot be utilized during the learning process.

1. Introduction

In both professional and personal lives, one is sometimes confronted with situations in which a large amount of information must be memorized and remembered at a later date. Demonstration of acquired knowledge is required from students at any educational level, as well as from members of professions that involve specific intellectual skills in the workplace.

Learning is a process that requires concentration, mental wellbeing, focused intellectual work, and time. Before an individual starts this process, it is a good idea to ask questions like “*Will I be able to memorize a given set of information? How long will it take to memorize this piece of information? After a week, will I remember at least 50% of what I have learned?*” The answers to these questions may play a crucial role in the selection of the right educational institution or job. This group attempts to formulate a simple mathematical model that could be used to provide answers to these questions. The model should preferably be *simple* because if the mathematical operations are too complicated, the theory would only be

useful for mathematicians. An emphasis was placed on the model's ability to provide *analytical predictions of measurable quantities* as well as its ability to be *calibrated for a particular user. This calibration needs to be as simple as possible.* These assumptions rule out the use of advanced mathematical methods that require strong computer hardware, complicated algorithms, or results in numerical format. Models with a large amount of free parameters are also not suitable due to their difficult nature in the process of calibration.

With these requirements in mind, this paper shows how one can derive the formulas that allow a student to estimate the time necessary for memorizing a certain amount of information and the theoretical capacity of their memory. The steps that lead to derivation of the formulas are outlined so a user will be able to understand the theoretical basis of the model and the logic behind any simplifying assumptions that were necessary. Finally, the steps which enable one to calibrate the model (i.e., determine the model's free parameters) for a particular student are included.

It is necessary to note that it is not the goal of this paper to analyze the complicated mechanisms of storing and recalling

of information from human memory. Similarly, this paper does not try to provide a new insight into the structure of the memory, information flow, or interactions occurring in the memory. The aim of this paper is to propose a useful tool that deals with the memory as a whole, by using easily measurable quantities, and is mathematically accessible to an ordinary educated person.

2. Results and Discussion

The model was built upon the premise of the Atkinson-Shiffrin (AS) concept of human memory, which assumes the existence of three memory blocks: sensory registry (SR), short-term memory (STM), and long-term memory (LTM) [1]. It should be noted that the Atkinson-Shiffrin scheme of human memory has some drawbacks that were criticized in a number of works [2–7]. Today, there are many more advanced models of human memory [8, 9] which partly address the issues of the Atkinson-Shiffrin model. However, the Atkinson-Shiffrin model is undoubtedly a suitable base for a simple mathematical model. In the presented model it is assumed that, while SR and STM are essential for storing the information to long-term memory, they do not significantly contribute to the total amount of memorized information.

Moreover, if the rate of assimilation of information and the rate of forgetting in SR and STM are several orders of magnitude greater than in LTM, it is safe to assume that SR and STM are incapable of storing information at the timescale of hours or years. In the following text, when referring to the available volume of information (AVI), the information stored in the long-term memory is to be considered. It is possible that certain memories are inaccessible not because they are no longer present in the memory, but because another information processing makes them hard or impossible to recall [10]. Such inaccessible information is not included in AVI. The symbol ν refers to AVI throughout the text. Furthermore, if it is said that the information is learned, it means that the information was not only stored, but also can be recalled. On the other hand, if it is said that the information has been forgotten, it does not necessarily mean that the information has been erased; it may just not be possible to recall it.

2.1. Forgetting. While it is possible to willingly stop the learning process, it is not an option with the process of forgetting. Since forgetting occurs all the time, it is difficult to directly observe the process of learning. On the other hand, observation of the forgetting process alone is relatively simple. Ebbinghaus was among the first who started to observe forgetting process, and in his seminal work he found, besides other things, that the process of forgetting is relatively fast at the start and gradually slows down over time [11].

A study on a large set of experimental data found that long-term memory retention can be described by several mathematical functions that are in a good correlation with experimental data [12]. Considering the properties of the functions and the analysis carried out in [12–15], the following

functions were chosen for the description of the forgetting process:

$$\nu(\nu_0, t) = \nu_0 e^{-\lambda t} + \omega, \quad (1)$$

$$\nu(\nu_0, t) = \frac{\nu_0}{(1 + \alpha t)^\beta}, \quad (2)$$

$$\nu(\nu_0, t) = \frac{\nu_0}{(1 + \alpha t)^\beta} e^{-\lambda t}. \quad (3)$$

For physical reasons time cannot be negative, so the domain of definition is, regardless of other parameters, the set $D(\nu) = \langle 0, \infty \rangle$. Coefficients $\alpha, \beta, \nu_0, \lambda, \omega \in R^+$ are parameters that define the shapes of the curves, and their values can be determined from experimental data.

Furthermore, at least for the purposes of the presented model, it is interesting to know whether or not the distributive property of forgetting functions (1)–(3) is satisfied. In such a case, the equation $\nu(\nu(\nu_0, t_1), t_2) = \nu(\nu_0, t_1 + t_2)$ must be valid for arbitrary $t_1, t_2 \in R^+$. If functions (1)–(3) do not obey this logical condition, forgetting process *depends on the observation* and the model can run into problems with calibration. For example, if the memory of the same person was tested twice in such a manner that the second test followed just after the first, then one would expect that after the second test the amount of forgotten information would be the same as if a single test (taking the same time as test 1 + test 2) was performed. This does not hold true if the distributive property is not satisfied.

In the next sections, the differential forms of formulas (1)–(3) are derived.

2.1.1. Exponential Model of Forgetting. Time derivative of function (1) is $d\nu/dt = -\lambda\nu_0 e^{-\lambda t}$. If (1) is taken into account and basic algebraic operations are performed, the exponential law of forgetting in differential form is obtained as follows:

$$\frac{d\nu}{dt} = -\lambda(\nu - \omega). \quad (4)$$

It can be seen that function (1) is decreasing on the whole domain if the logical condition $\nu > \omega$ is satisfied. The limit of this function at the upper boundary of the domain approaches ω . Hence the range of the function is the interval $H(\nu) = (\omega, \nu_0)$. Equation (4) further implies that λ (in units s^{-1}) determines the curve steepness.

It is important to note that function (1) does not obey the distributive property because

$$\begin{aligned} \nu(\nu(\nu_0, t_1), t_2) &= (\nu_0 e^{-\lambda t_1} + \omega) e^{-\lambda t_2} + \omega \\ &= \nu_0 e^{-\lambda(t_1+t_2)} + \omega(1 + e^{-\lambda t_2}) \\ &\neq \nu_0 e^{-\lambda(t_1+t_2)} + \omega \\ &= \nu(\nu_0, t_1 + t_2). \end{aligned} \quad (5)$$

In this case, the distributive property failure is caused by the presence of permastore asymptotic term ω . Although

Bahricks permastore [16] is established and well-accepted phenomenon which comes into play mainly on a long-term time scale, for the purposes of the presented short-term model the permastore asymptotic term should be considered to be zero. Further argumentation supporting this decision can be found in Section 2.2.2.

2.1.2. Power-Law Model of Forgetting. Time derivative of function (2) leads to $dv/dt = -\nu_0\alpha\beta/(1+\alpha t)^{\beta+1}$. If (2) is taken into account and basic algebraic operations are performed, the power law of forgetting in differential form can be obtained as follows:

$$\frac{dv}{dt} = -\frac{\alpha\beta}{1+\alpha t} \nu. \quad (6)$$

The function ν is decreasing on the whole domain and the limit of this function at the upper boundary of the domain approaches zero. Hence the range of the function is the interval $H(\nu) = (0, \nu_0)$. Equation (6) further implies that coefficients α and β determine the curve steepness. Coefficient β is dimensionless; the dimension of α is s^{-1} .

Note that function (2) does not obey the distributive property because

$$\begin{aligned} \nu(\nu(\nu_0, t_1), t_2) &= \frac{\nu_0/(1+\alpha t_1)^\beta}{(1+\alpha t_2)^\beta} \\ &= \frac{\nu_0}{[(1+\alpha(t_1+t_2) + \alpha^2 t_1 t_2)]^\beta} \quad (7) \\ &\neq \frac{\nu_0}{[(1+\alpha(t_1+t_2))]^\beta} \\ &= \nu(\nu_0, t_1+t_2). \end{aligned}$$

On the other hand, good arguments in favor of the power law model of forgetting exist. These arguments are supported by experiments as well as by theoretical works [14, 17–19]. In general, this model is widely used in the available literature; therefore, the model of learning with power-like type of forgetting is also derived in this paper (see Section 2.2.3).

2.1.3. Combined Power-Exponential Model of Forgetting. This model is described by function (3). Because this function is a combination of functions (1) and (2), all the benefits of the mentioned functions are to be expected. In the special case of $\alpha = 0$ or $\beta = 0$, (3) is reduced to (1); in case of $\lambda = 0$ (3) leads to (2). The model is based on the assumption that the decay rate of a memory trace is slowed down by the interference with other memory traces and by its *fragility*. Detailed derivation of formula (3) can be found in [15].

Time derivative of function (3) is $dv/dt = \nu_0 e^{-\lambda t} ((-\lambda(1+\alpha t)^\beta - \alpha\beta(1+\alpha t)^{\beta-1})/(1+\alpha t)^{2\beta})$. Taking (3) into account, after basic algebraic operations the combined power-exponential law of forgetting in differential form becomes

$$\frac{dv}{dt} = -\left(\lambda + \frac{\alpha\beta}{1+\alpha t}\right) \nu. \quad (8)$$

The function ν is decreasing on the whole domain and the limit of this function at the upper boundary of the domain approaches zero. Hence the range of the function is the interval $H(\nu) = (0, \nu_0)$.

Function (3) does not satisfy the distributive property because

$$\begin{aligned} \nu(\nu(\nu_0, t_1), t_2) &= \frac{\nu_0 e^{-\lambda t_1}}{(1+\alpha t_1)^\beta} \frac{e^{-\lambda t_2}}{(1+\alpha t_2)^\beta} \\ &= \frac{\nu_0 e^{-\lambda(t_1+t_2)}}{[(1+\alpha(t_1+t_2) + \alpha^2 t_1 t_2)]^\beta} \quad (9) \\ &\neq \frac{\nu_0 e^{-\lambda(t_1+t_2)}}{[(1+\alpha(t_1+t_2))]^\beta} \\ &= \nu(\nu_0, t_1+t_2). \end{aligned}$$

In spite of this fact, the combined power-exponential retention function has the potential to be one of the most accurate models when compared to experimental data. Taking this fact into account, the model of learning with combined power-exponential type of forgetting is also derived in this paper (see Section 2.2.4).

2.2. Learning

2.2.1. Idealized Model of Learning. Learning is a process that occurs concurrently with forgetting. While learning is the result of our decision, we are unable to control the process of forgetting. Before a realistic model of learning is synthesized, it will be necessary to describe an idealized case in which forgetting, figuratively speaking, is *turned off*. One must assume that “pure” learning can be described by the following equation:

$$\frac{dv}{dt} = u. \quad (10)$$

In the following text the symbol u is reserved for the *rate of learning*. Its dimension is s^{-1} . In reality, the rate of learning is affected by many factors, including but not limited to fatigue, amount of sleep, and mental wellbeing. However, in order to obtain analytical solutions of differential equations, throughout this study it is assumed that the rate of learning is *constant in time*. This is also one of the model’s free parameters that can be determined experimentally for a particular user. In such a case (10) is a separable first order ordinary differential equation. Assuming that at the start of learning, $t = 0$, the memory contains $\nu(0) = \nu_0$ information, by integration of (10) one can arrive at the solution

$$\nu = \nu_0 + ut. \quad (11)$$

In the initial phase of learning, if one ignores forgetting and considers learning to be ideal, the time necessary for memorizing the volume of information ν becomes

$$t = \frac{\nu - \nu_0}{u}. \quad (12)$$

2.2.2. Model of Learning with Exponential Type of Forgetting.

For the sake of simplicity, this model will be later referred to as the exponential model of learning. The basic idea of the model is that the process of learning consists of two concurrent processes, ideal learning and forgetting. Hence the overall increase of the AVI in time dt is the difference between acquired volume and forgotten volume of information. In the case of exponential forgetting, this concept can be formulated mathematically in the following form:

$$\frac{dv}{dt} = u - \lambda(v - \omega). \quad (13)$$

This differential equation is separable and for the initial condition $v(0) = v_0$ the solution is

$$v(\nu_0, t) = \left(\nu_0 - \frac{u}{\lambda} - \omega \right) e^{-\lambda t} + \frac{u}{\lambda} + \omega. \quad (14)$$

Function (14) has some interesting features and implications. In time $t = 0$ it is equal to the initial AVI. If the condition $1 < (u/\lambda + \omega)/\nu_0$ is met for $\lambda, u, \nu_0 \in R^+$, the function is increasing over the whole domain. In the limit case $t \rightarrow \infty$ it approaches the value $\nu_\infty = u/\lambda + \omega$, $\nu_\infty \in R^+$; therefore, function (14) is bounded. Evidently, the exponential model supports the hypothesis that an individual is incapable of learning an unlimited amount of information. If one learned over an infinite amount of time, this individual would have approached their upper AVI limit ν_∞ . In the following text this number will be referred to as *the capacity of a student*.

It should be mentioned at this point that according to solution (14) the amount of learned information increases up to ω even if the rate of learning is zero (i.e., the student does not learn at all). Clearly, this result is far from reality and another reason why the asymptotic term ω should be disregarded in presented model.

It will now be shown that solution (14) meets the distributive property $v(\nu(\nu_0, t_1), t_2) = \nu(\nu_0, t_1 + t_2)$ for any choice of variables $\nu_0, t_1, t_2 \in R^+$. Consider

$$\begin{aligned} & \left[\left(\nu_0 - \frac{u}{\lambda} - \omega \right) e^{-\lambda t_1} + \frac{u}{\lambda} + \omega \right] - \frac{u}{\lambda} - \omega \Big) e^{-\lambda t_2} + \frac{u}{\lambda} + \omega \\ &= \left(\nu_0 - \frac{u}{\lambda} - \omega \right) e^{-\lambda(t_1+t_2)} + \frac{u}{\lambda} + \omega. \end{aligned} \quad (15)$$

Furthermore, the amount of time needed for acquiring the volume of information ν can be derived from (14) as follows:

$$t = \frac{1}{\lambda} \ln \frac{\nu_0 - u/\lambda - \omega}{\nu - u/\lambda - \omega}. \quad (16)$$

The idea that an individual is capable of receiving new information continuously is utopian. After a specific amount of time (let us call it τ) it is necessary to interrupt the learning process since the rate of learning cannot be considered constant, and assimilation of information becomes less effective. At this point it is advisable to allow one's memory to regenerate. Some studies even suggest that relaxation and

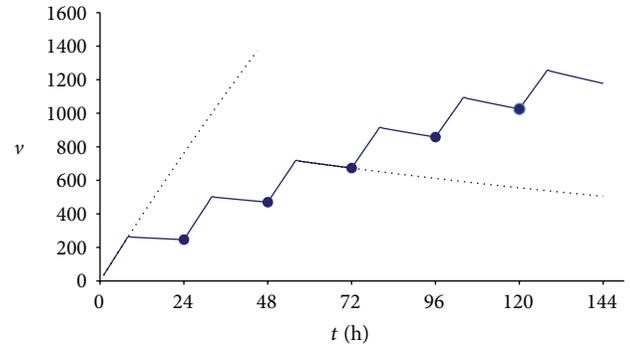


FIGURE 1: Periodic alternation of learning and forgetting. Vertical axis represents the number of memorized words of a foreign language; horizontal axis represents time.

sleep are vital for the consolidation of memory traces of declarative type [20]. An attempt to mathematically describe the process of learning interrupted with periodic breaks is needed. It must be assumed that during the time τ the student is able to learn at a constant rate u . During this phase, learning and forgetting occur concurrently as described in (13). After this phase, the student needs to relax for the period $T - \tau$ to be able to study effectively again. At this point the student is merely forgetting, and (1) is valid. Symbol T , introduced previously, refers to the period of human circadian rhythm and is equal to 24 h. The alternation between learning and forgetting phases is depicted in Figure 1.

It is evident that the AVI remaining in memory after the first closed cycle of learning and forgetting will be

$$\nu_1 = \left[\left(\nu_0 - \frac{u}{\lambda} - \omega \right) e^{-\lambda \tau} + \frac{u}{\lambda} + \omega \right] e^{-\lambda(T-\tau)} + \omega. \quad (17)$$

Consequently, after n closed cycles, the AVI will be

$$\nu_n = \left[\left(\nu_{n-1} - \frac{u}{\lambda} - \omega \right) e^{-\lambda \tau} + \frac{u}{\lambda} + \omega \right] e^{-\lambda(T-\tau)} + \omega. \quad (18)$$

Learning is progressive only until the AVI is increasing with the number of cycles; that is, $\nu_{n+1} > \nu_n$ or $[(\nu_n - u/\lambda - \omega)e^{-\lambda \tau} + u/\lambda + \omega]e^{-\lambda(T-\tau)} + \omega > \nu_n$. After a series of algebraic operations one will find that in the limit case $n \rightarrow \infty$ the progression ν_n is approaching the value

$$\tilde{\nu}_\infty = \lim_{n \rightarrow \infty} \nu_n = \frac{(u/\lambda + \omega)(e^{\lambda \tau} - 1) + \omega e^{\lambda T}}{e^{\lambda T} - 1}, \quad (19)$$

which will be referred to as *the real capacity of a student*.

It will now be shown that the presence of the asymptotic term ω leads to further illogical implications. For example, assume a scenario in which the student decides not to study at all. In other words, the period τ during which the student is supposed to be assimilating information is equal to zero. Logically, one would expect that after an arbitrary amount of time the student will still know nothing. However, (19) states that the amount of his or her knowledge will be growing until it reaches the boundary $\omega(e^{\lambda T}/(e^{\lambda T} - 1)) \geq 0$. This logical discrepancy can only be removed if $\omega = 0$. In the remainder

of the text, the equations that do not contain the asymptotic term ω will therefore be used.

It can be shown by means of mathematical analysis that the longer the student is able to assimilate information at a constant rate u , the greater their *real capacity* is. The following holds true:

$$\frac{\partial \tilde{v}_\infty}{\partial \tau} = u \frac{e^{\lambda \tau}}{e^{\lambda T} - 1} \geq 0, \quad \forall \tau \in \langle 0, T \rangle. \quad (20)$$

If the student knows his or her capacity \tilde{v}_∞ , before the beginning of the study they can calculate if it is feasible to memorize a desired amount of information. If the answer is positive, then according to recurrent relation (18) they can calculate how many closed cycles (in other words, how many days) will be needed to achieve their goal.

Coefficient \tilde{v}_∞ could also be a tool to compare one student to another or to determine the student's ability to complete their study at a given college or university. \tilde{v}_∞ could also be a reference point in a selection process for jobs requiring good memory and so forth.

2.2.3. Model of Learning with Power-Like Type of Forgetting. This model will be later referred to as the power-law model of learning. Once again, the increase of the AVI in time dt is the difference between acquired and forgotten information. The member responsible for forgetting is represented by the right side of (2) as follows:

$$\frac{dv}{dt} = u - \frac{\alpha \beta}{1 + \alpha t} v. \quad (21)$$

$$v_n = \frac{u(1 + \alpha \tau) / [\alpha(\beta + 1)] + (v_{n-1} - u / [\alpha(\beta + 1)]) / (1 + \alpha \tau)^\beta}{[1 + \alpha(T - \tau)]^\beta}. \quad (23)$$

The progression of numbers $\{v_n\}_{n=1}^\infty$ is increasing and bounded. In the limit case $t \rightarrow \infty$ we will obtain

$$\tilde{v}_\infty = \lim_{n \rightarrow \infty} v_n = \frac{u}{\alpha(\beta + 1)} \frac{(1 + \alpha \tau)^{\beta+1} - 1}{(1 + \alpha \tau)^\beta [1 + \alpha(T - \tau)]^\beta - 1}. \quad (24)$$

Therefore, the real capacity of the memory is bounded even in the power-law model.

2.2.4. Model of Learning with Combined Power-Exponential Type of Forgetting. This model will be later referred to as the combined model of learning. To obtain the increase of the volume of information in time one can write

$$\frac{dv}{dt} = u - \left(\lambda + \frac{\alpha \beta}{1 + \alpha t} \right) v. \quad (25)$$

Equation (25) is a first order linear differential equation which can be solved by method of variation of the constant,

Formula (21) is a first order linear differential equation which can be solved by method of variation of the constant. Assuming a reasonable initial condition $v(0) = v_0$, one will obtain the solution in the following form:

$$v(v_0, t) = \frac{u}{\alpha(\beta + 1)} (1 + \alpha t) + \frac{v_0 - u / [\alpha(\beta + 1)]}{(1 + \alpha t)^\beta}. \quad (22)$$

It can be easily verified that the solution of (22) does not satisfy the distributive property.

At the beginning (in time $t = 0$) the volume of information is v_0 . If the parameters were unrestricted ($\alpha, \beta, u, v_0 \in R^+$), function (21) might not necessarily be monotonous and it might have a local minimum, the existence of which is not justifiable. To rule out this possibility, one will require the time derivative to be nonnegative on the entire domain; that is, $dv/dt \geq 0$ for $t \in R^+$. It can be easily shown that this condition will be satisfied if $v_0 \alpha \beta / u < 1$. An even stronger condition is the inequality $v_0 \alpha (\beta + 1) / u < 1$, which rules out the existence of local extremes on the domain of real numbers. When the aforementioned conditions are met, function (22) is nondecreasing. In the limit case $t \rightarrow \infty$ the AVI diverges to infinity, which seemingly supports the idea of limitless capacity of human memory. However, it is unrealistic to assume that a student will learn without breaks, so it is necessary to examine the case of cyclic periods of learning and relaxation. The meaning of symbols τ and T is the same as in the case of exponential model. After the n th cycle, the AVI will be

assuming a reasonable initial condition $v(0) = v_0$. The solution of (25) leads to an integral, the value of which is impossible to find unless we introduce a restriction that $\beta \in N_0$. In such a case, after a series of recurrent *per partes* integrations, we will obtain the solution of (25) in the following form:

$$v(v_0, \beta, t) = \left[v_0 - \frac{u}{\lambda} \sum_{i=0}^{\beta} (-1)^i \frac{\beta!}{(\beta - i)!} \left(\frac{\alpha}{\lambda} \right)^i \right] \frac{e^{-\lambda t}}{(1 + \alpha t)^\beta} + \frac{u}{\lambda} \sum_{i=0}^{\beta} (-1)^i \frac{\beta!}{(\beta - i)!} \left[\frac{\alpha}{\lambda(1 + \alpha t)} \right]^i. \quad (26)$$

If β is a noninteger number, the approximate solution of (25) can be obtained by linear interpolation method. Let $m \leq \beta \leq m + 1$, $m \in N_0$, $\beta \in (m, m + 1)$, and let us further assume that $v(v_0, m, t)$ and $v(v_0, m + 1, t)$ are the solutions of (25) for a whole-number parameter β . The solution for β belonging

to the real numbers, specifically to the interval $(m, m + 1)$, $m \in N_0$, can be approximated by the function

$$\begin{aligned} v(v_0, \beta, t) = & (m + 1 - \beta) v(v_0, m, t) \\ & + (\beta - m) v(v_0, m + 1, t). \end{aligned} \quad (27)$$

It can be easily verified that the solution of (25) is again not compliant with the distributive property.

Finally, one can derive the formulas describing cyclic alternation of learning and forgetting. After the first phase of learning that takes time τ and after the following phase of memory regeneration that takes time $T - \tau$, the volume of information stored in the memory will be

$$\begin{aligned} v_1 = & \left\{ \left[v_0 - v_\infty \sum_{i=0}^{\beta} (-1)^i \frac{\beta!}{(\beta-i)!} \left(\frac{\alpha}{\lambda}\right)^i \right] \frac{e^{-\lambda\tau}}{(1+\alpha\tau)^\beta} \right. \\ & \left. + v_\infty \sum_{i=0}^{\beta} (-1)^i \frac{\beta!}{(\beta-i)!} \left[\frac{\alpha}{\lambda(1+\alpha\tau)} \right]^i \right\} \\ & \times \frac{e^{-\lambda(T-\tau)}}{[1+\alpha(T-\tau)]^\beta}. \end{aligned} \quad (28)$$

Analogically, after the n th cycle the volume of information stored in the memory will be

$$\begin{aligned} v_n = & \left\{ \left[v_{n-1} - v_\infty \sum_{i=0}^{\beta} (-1)^i \frac{\beta!}{(\beta-i)!} \left(\frac{\alpha}{\lambda}\right)^i \right] \frac{e^{-\lambda\tau}}{(1+\alpha\tau)^\beta} \right. \\ & \left. + v_\infty \sum_{i=0}^{\beta} (-1)^i \frac{\beta!}{(\beta-i)!} \left[\frac{\alpha}{\lambda(1+\alpha\tau)} \right]^i \right\} \\ & \times \frac{e^{-\lambda(T-\tau)}}{[1+\alpha(T-\tau)]^\beta}. \end{aligned} \quad (29)$$

If one investigates the limit case $t \rightarrow \infty$, after numerous algebraic operations, one will obtain the real capacity of the memory as follows:

$$\begin{aligned} \tilde{v}_\infty = & \frac{u}{\lambda} \left(\left(\frac{e^{\lambda\tau} \sum_{i=0}^{\beta} (-1)^i (\beta! / (\beta-i)!) (\alpha/\lambda)^i (1+\alpha\tau)^{\beta-i}}{\sum_{i=0}^{\beta} (-1)^i (\beta! / (\beta-i)!) (\alpha/\lambda)^i} - 1 \right) \right. \\ & \left. \times \left(\frac{e^{\lambda T} [1+\alpha(T-\tau)]^\beta (1+\alpha\tau)^\beta}{\sum_{i=0}^{\beta} (-1)^i (\beta! / (\beta-i)!) (\alpha/\lambda)^i} - 1 \right)^{-1} \right). \end{aligned} \quad (30)$$

Note that formula (30) is valid only for the whole-number parameter β . Transition to the real β can be done by linear interpolation; see (27).

2.2.5. Usability of the Models. The presented models are expected to describe the assimilation of information satisfactorily, under the assumption that retention functions (1)–(3) give a true picture about forgetting process. The models are limited to scenarios where logical bonds are not created during the process of learning and mnemonic devices or other

memory aids and associations cannot be utilized. In the case of memorizing meaningful logical material, the models are limited to situations when the subject is pressed for time or working under pressure. This essentially prevents them from realizing the logical bonds (connections to prior knowledge) between memorized information. Some examples in which these models are ineffective are scientific disciplines such as mathematics, physics, or informatics. Some examples of suitable candidates are the study of foreign languages, history, law, chosen parts of medicine, pharmacology, chemistry, biology, botanic, and so forth.

2.3. Calibration of the Models for a Particular Student. In order to make the formulas derived earlier usable in practice, the student must first find the constants defining their ability to learn and remember the memorized information. Regardless of the chosen theoretical model, it is necessary to know the rate of learning u . In the case of the exponential model one must further know the constant λ , in the power-law model the constants α and β , and in combined exponential-power model all of the mentioned constants. In the case of periodic alternation of learning and forgetting one needs to know the parameter τ as well.

First, it is necessary to describe an experiment in which one observes the process of forgetting, which is in the exponential model represented by function (1). Due to the aforementioned reasons, the permastore term ω will be considered to be zero in the remainder of the text.

The student will memorize an initial volume of information v_0 . After that, he or she will wait for at least 30 seconds. The student will then be tested to verify the accuracy of what they really remembered. The 30-second minimum delay between memorizing and evaluation will ensure that the recalled information is not retrieved from the short-term or sensory memory. For the next 60 minutes the student does not learn and avoids conscious or unconscious recalling or repeating of the memorized information. After this time, he or she is to be examined in order to find out how much of the initially remembered information he or she is still able to recall.

This process is repeated at least two more times. Each time a new set of information is memorized and the interval between memorizing and examining is prolonged for an appropriate time step (e.g., one hour). This will lead to at least four ordered pairs $\{[0, v_0]; [t_1, v_1]; [t_2, v_2]; [t_3, v_3]\}$. The constant λ can be determined by the least squares method in order to minimize the deviation of experimental data from function (1).

From an experimental point of view, the calibration procedure for the power-law and the combined models is the same. When one considers that these models contain more free parameters, it is advisable to increase the number of ordered pairs and use one of the methods of nonlinear regression.

It is now necessary to determine the rate of learning u . It is important to note that forgetting cannot be suppressed by will power. Any experiment dealing with learning must include the premise that learning and forgetting are expected

to occur concurrently; therefore, both of these phenomena will be observed.

In the case of the exponential model, this complex process is described by the equation $v = (\nu_0 - u/\lambda)e^{-\lambda t} + u/\lambda$. The student is to pick an unknown set of information; hence, $\nu_0 = 0$ and (14) reduces to $v = (u/\lambda)(1 - e^{-\lambda t})$. The student must memorize the given set of information v and will write down the time needed to accomplish this task. The volume of information should be memorizable in a relatively short time; otherwise, the nonlinearity coupled with ongoing forgetting would start to be apparent. Therefore the condition $\lambda t \approx 0$ should be met. In such a case one can approximate (14) with sufficient accuracy by second order Taylor series around the origin $t = 0$: $v \doteq ut - u\lambda t^2$. When one uses this formula, the rate of learning becomes

$$u = \frac{\nu}{t - \lambda t^2}. \tag{31}$$

In the case of the power-law model, assuming that $\nu_0 = 0$ and $\alpha t \approx 0$, one can approximate (22) with sufficient accuracy by second order Taylor series around the origin $t = 0$: $v \doteq ut - \beta t^2$. This formula generates the rate of learning to be

$$u = \frac{\nu}{t - \beta t^2}. \tag{32}$$

In the case of the combined model, under the conditions, mentioned previously, one can approximate (26) with sufficient accuracy by second order Taylor series around the origin $t = 0$: $v \doteq ut - u(\alpha\beta + \lambda)t^2$. This formula results in the rate of learning to be

$$u = \frac{\nu}{t - (\alpha\beta + \lambda)t^2}. \tag{33}$$

It is very likely that if a student repeats the calibration process under the same conditions several times, he or she will obtain different values of parameters α , β , λ , and u . This can be caused by the fact that the forgetting term of used model does not obey the distributive property or memory characteristics changed over time, or because the described calibration methods provide only an estimate of real values of the parameters.

Imagine if one repeated the calibration process n -times. One would get n -values $\{\lambda_1, \dots, \lambda_n\}$, $\{\alpha_1, \dots, \alpha_n\}$, $\{\beta_1, \dots, \beta_n\}$, and $\{u_1, \dots, u_n\}$ that would be distributed normally around the mean values $\bar{\lambda}$, $\bar{\alpha}$, $\bar{\beta}$, and \bar{u} , with the variance σ_λ^2 , σ_α^2 , σ_β^2 , and σ_u^2 . The mean value of function v in the exponential model will then be

$$\nu(\nu_0, t) = \frac{\iint_0^\infty [(\nu_0 - u/\lambda)e^{-\lambda t} + (u/\lambda)] e^{-(\bar{\lambda}-\lambda)^2/2\sigma_\lambda^2} e^{-(\bar{u}-u)^2/2\sigma_u^2} d\lambda du}{\int_0^\infty e^{-(\bar{\lambda}-\lambda)^2/2\sigma_\lambda^2} d\lambda \int_0^\infty e^{-(\bar{u}-u)^2/2\sigma_u^2} du}. \tag{34}$$

Analogically, for the power-law model one can write

$$\nu(\nu_0, t) = \frac{\iiint_0^\infty (\nu_0/(1 + \alpha t)^\beta + u/[\alpha(\beta + 1)] [1 + \alpha t - (1 + \alpha t)^{-\beta}]) e^{-[(\bar{\alpha}-\alpha)^2/2\sigma_\alpha^2 + (\bar{\beta}-\beta)^2/2\sigma_\beta^2 + (\bar{u}-u)^2/2\sigma_u^2]} d\alpha d\beta du}{\iiint_0^\infty e^{-[(\bar{\alpha}-\alpha)^2/2\sigma_\alpha^2 + (\bar{\beta}-\beta)^2/2\sigma_\beta^2 + (\bar{u}-u)^2/2\sigma_u^2]} d\alpha d\beta du} \tag{35}$$

and for the combined power-exponential model we will obtain (in shortened form)

$$\nu(\nu_0, t) = \frac{\iiint_0^\infty \nu(\nu_0, \alpha, \beta, \lambda, u) e^{-(\bar{\lambda}-\lambda)^2/2\sigma_\lambda^2 - (\bar{\alpha}-\alpha)^2/2\sigma_\alpha^2 - (\bar{\beta}-\beta)^2/2\sigma_\beta^2 - (\bar{u}-u)^2/2\sigma_u^2} d\lambda d\alpha d\beta du}{\iiint_0^\infty e^{-(\bar{\lambda}-\lambda)^2/2\sigma_\lambda^2 - (\bar{\alpha}-\alpha)^2/2\sigma_\alpha^2 - (\bar{\beta}-\beta)^2/2\sigma_\beta^2 - (\bar{u}-u)^2/2\sigma_u^2} d\lambda d\alpha d\beta du}. \tag{36}$$

Considering the complexity of these integrals, the process of integration needs to be realized numerically.

Finally, in the formulas describing the real capacity of the memory it is necessary to find the last of the parameters, the time τ . This parameter represents the part of the day during which one is able to effectively assimilate information at a constant rate. This rate can be determined as follows. Using

a small sample of study material, the student will determine his or her rate of learning by one of the formulas (31)–(33). The rate of learning should be determined several times during the learning process (e.g., once per hour). In this manner, the ordered pairs $\{[t_0, u_0]; [t_1, u_1]; \dots; [t_n, u_n]\}$ will be obtained. Here t_i represents the time elapsed from the start of the learning process and u_i are the measured rates

of learning. This process should be repeated until the rate of learning decreases under a certain threshold due to fatigue and loss of concentration ($u_{\text{threshold}} \cong 90\%u_0 = u_n$). At this point the rate of learning can no longer be considered constant and the following estimation can be made: $\tau \cong t_n$.

It is known for a long time and experimentally verified [11, 12, 21] that a memory trace of information that is repeatedly remembered lasts longer than a trace that is remembered for the first time. Similarly, repeated memorizing of information that one is unable to recall takes shorter amount of time than information remembered for the first time. Furthermore, it must be noted that none of the presented models takes these factors into account. The authors are planning to deal with these issues in the future.

3. Conclusion

In the presented work the relations describing the concurrent processes of learning and forgetting were derived mathematically. The relations are based on experimentally verified retention functions. It has been shown that if these functions realistically depict the process of forgetting on the time scale of hours to days, then one of the implications is that the capacity of human memory is limited and a person cannot, even in theory, learn a limitless amount of information. Presented models have a predictive potential to estimate the time needed to learn a given set of information in a circadian rhythm. This feature could be especially useful for students facing examinations under time pressure. An outline which describes the mechanism of calibration of the models for a particular student is also included in the study. This tool could also prove to be useful for students that are in the process of choosing the right scientific discipline or for people looking for the right job. It is assumed that the model gives realistic results only if the nature of the memorized material does not allow for the creation of logical bonds (connections to prior learning) and does not provide options for utilization of mnemonic devices or other memory aids.

At the beginning, three key criteria were defined that a model should meet to be usable in practice. It must be *simple enough* not to deter potential users. The solutions must exist in *analytical form* and the *process of calibration for a particular user must be simple*. These requirements are primarily met by the exponential and power-law model. The combined model lacks simplicity and contains many free parameters, making the calibration process difficult. Moreover, the power-law and combined models do not satisfy the distributive property. When one takes the aforementioned facts into account, it becomes apparent that the most practical and usable model is the model of learning with exponential type of forgetting. It should be emphasized that the equations in Section 2.2.2, for the sake of better compatibility with some psychological articles do contain the asymptotic permastore term ω . However, it was shown that the presence of this term creates a discrepancy between the model's predictions and logical expectations at the short-time scale. Therefore, it is suggested that this term should be considered to be zero at least for the purposes of the presented model.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors gratefully acknowledge the Scientific Grant Agency (VEGA) of the Ministry of Education of Slovak Republic and the Slovak Academy of Sciences for supporting this work under Grant no. 1/1245/12.

References

- [1] R. C. Atkinson and R. M. Shiffrin, "Human memory: a proposed system and its control processes," *Psychology of Learning and Motivation*, vol. 2, pp. 89–195, 1968.
- [2] A. W. Melton, "Implications of short-term memory for a general theory of memory," *Journal of Verbal Learning and Verbal Behavior*, vol. 2, no. 1, pp. 1–21, 1963.
- [3] B. B. Murdock Jr., "Short-term memory," *Psychology of Learning and Motivation*, vol. 5, pp. 67–127, 1972.
- [4] E. Tulving and R. D. Patterson, "Functional units and retrieval processes in free recall," *Journal of Experimental Psychology*, vol. 77, no. 2, pp. 239–248, 1968.
- [5] T. Shallice and E. K. Warrington, "Independent functioning of verbal memory stores: a neuropsychological study," *The Quarterly Journal of Experimental Psychology*, vol. 22, no. 2, pp. 261–273, 1970.
- [6] E. Tarnow, "Why the Atkinson-Shiffrin model was wrong from the beginning," *WebmedCentral Neurology*, vol. 1, no. 10, pp. 1–13, 2010.
- [7] E. Tarnow, "There is no capacity limited buffer in the Murdock (1962) free recall data," *Cognitive Neurodynamics*, vol. 4, no. 4, pp. 395–397, 2010.
- [8] A. D. Baddeley and G. Hitch, "Working Memory," *Psychology of learning and motivation: advances in research and theory*, vol. 8, pp. 47–89, 1974.
- [9] A. Baddeley, "Working memory: theories, models, and controversies," *Annual Review of Psychology*, vol. 63, pp. 1–29, 2012.
- [10] D. Laming, "Failure to recall," *Psychological Review*, vol. 116, no. 1, pp. 157–186, 2009.
- [11] H. Ebbinghaus, *Über Das Gedächtnis. Untersuchungen Zur Experimentellen Psychologie*, Duncker and Humblot, Leipzig, Germany, 1885.
- [12] D. C. Rubin and A. E. Wenzel, "One hundred years of forgetting: a quantitative description of retention," *Psychological Review*, vol. 103, no. 4, pp. 734–760, 1996.
- [13] D. C. Rubin, S. Hinton, and A. Wenzel, "The precise time course of retention," *Journal of Experimental Psychology*, vol. 25, no. 5, pp. 1161–1176, 1999.
- [14] J. T. Wixted and S. K. Carpenter, "The Wickelgren power law and the Ebbinghaus savings function," *Psychological Science*, vol. 18, no. 2, pp. 133–134, 2007.
- [15] W. A. Wickelgren, "Single-trace fragility theory of memory dynamics," *Memory and Cognition*, vol. 2, no. 4, pp. 775–780, 1974.
- [16] H. P. Bahrick, "Semantic memory content in permastore: fifty years of memory for Spanish learned in school," *Journal of Experimental Psychology*, vol. 113, no. 1, pp. 1–29, 1984.

- [17] 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, pp. 731–739, 1997.
- [18] J. T. Wixted and E. B. Ebbesen, "On the form of forgetting," *Psychological Science*, vol. 2, no. 6, pp. 409–415, 1991.
- [19] R. B. Anderson, "The power law as an emergent property," *Memory and Cognition*, vol. 29, no. 7, pp. 1061–1068, 2001.
- [20] S. Ruch, O. Marques, S. B. Duss et al., "Sleep stage II contributes to the consolidation of declarative memories," *Neuropsychologia*, vol. 50, no. 10, pp. 2389–2396, 2012.
- [21] C. W. Luh, "The conditions of retention," *Psychological Monographs*, vol. 31, article 142, 1922.