International Education and Culture Studies
E-ISSN 2766-6778
Vol. 1, No. 1, 2021
Published by STSL Press

Original Paper

Modelling the Teaching Process: A Markov Chain Approach

Michael Gr. Voskoglou[1]

[1] University of Peloponnese, Ex Graduate T.E.I. of Western Greece, Greece

Correspondence: Michael Gr. Voskoglou, mvoskoglou@gmail.com

Abstract

The topic of Markov chains is a synthesis of Probability and matrix theory enabling one to tackle mathematically a variety of problems characterized by randomness and connected to many sectors of the human activity. In this work a finite Markov chain is introduced representing mathematically the teaching process which is based on the ideas of constructivism for learning. Interesting conclusions are derived and a measure is obtained for the teaching effectiveness. An example on teaching the derivative to fresher university students is also presented illustrating our results.

Keywords: Markov chain, fundamental matrix, social constructivism

1. Introduction

Learning is the most important component of the human cognition. Many theories and models have been developed by the experts on the field for analyzing the process of learning, explaining its mechanisms and make it accessible to everyone. *Constructivism* is a relatively recent theory for learning, which is based on Piaget's ideas on the subject. This theory was formally introduced by von Clasersfeld during the 1970's. According to it, knowledge is not passively acquired from the environment, but is actively developed by the learner through a process of adaptation based on and constantly modified by the learner's experience of the world (Taber, 2011). As a result of the failure of the introduction of "new mathematics" in basic education, the application of the ideas of constructivism in the teaching process has become very popular during the last decades, especially in school education. The steps of a typical method of teaching which is based on the principles of constructivism are the following:

- *Engagement* (S_1): This is the first step of the learning process connecting the past with the present learning situations and orienting student reasoning on the learning conclusions of the current activities.

- *Investigation* (S_2): Here students study and investigate the environment on the purpose of creating a common set of relative to the learning subject experiences by determining and constructing notions, procedures and the necessary skills.

- *Explanation* (S_3): In this step students explain, analyze and formalize the notions that have been constructed in the previous step, while the teacher introduces formal terms, definitions and explanations for the new concepts and processes and demonstrates new skills.

- *Deeper Understanding* (S_4): In that step students develop a deeper and broader conceptual understanding and obtain more knowledge about the new acquired information through elaboration and practice on the new topic.

- *Evaluation* (S_5): In the last step of the learning process, students are encouraged to evaluate with the help of the given by the teacher homework their understanding and abilities on the new topic, while the instructor assess with oral and written tests and/or questions the student performance on the new topic.

Depending on the student reactions in the classroom, there are forward or backward transitions between the three intermediate steps (S_2, S_3 and S_4) of this framework during the teaching process, the flow-diagram of which is shown in Figure 1.

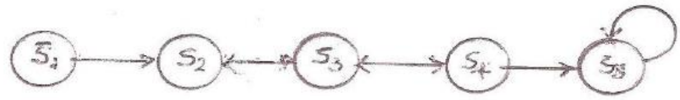

Figure 1. The Flow-diagram of the Teaching Process

Note that the three intermediate steps S_2, S_3 and S_4 of the teaching model developed above appear analogous to the phases of *exploration, formalization* and *assimilation* respectively for the process of learning Mathematics presented by G. Polya in his famous two volume book “Mathematical Discovery” (Wiley: NY, 1962/65).

In this article a Markov chain is introduced on the steps of the teaching process and through it interesting conclusions are derived for the teaching effectiveness. The rest of the article is formulated as follows: The next section provides a brief account of the basic principles of the theory of finite Markov chains, which is necessary for the better understanding of the contents of this work. In section 3 an absorbing Markov chain model is developed for the teaching process based on the principles of constructivism. The model is illustrated by an example on teaching the derivative to a group of fresher university students. Section 3 closes with an important remark connected to a different technique used in an author’s earlier work for modelling the teaching process by using an ergodic Markov chain. The general conclusions of this study are stated in section 4, the last section of the article.

2. Markov Chains

A simple way to define a *Markov chain (MC)* is the following: A MC is a process depending on a random variable that takes values from a set of states which move in a sequence of steps (phases), and having memory of one step only. *This means that the probability of the random variable to enter a certain state in a certain step depends on* the state occupied in the previous step and not in earlier steps. This is called the *Markov property*. In practice, however, in order to be able to model as many cases as possible by using MCs, one may accept that the probability of entering a certain state in a certain step, although it could not be completely independent of previous steps, it mainly depends on the state occupied in the previous step (Kemeny, 1962).

The basic concepts of MCs were introduced by Andrey Markov (Figure 2) in 1907 on coding literal texts. Since then the MC theory was developed by a number of leading mathematicians, such as A. Kolmogorov, W. Feller, etc. However, only from the 1960’s the importance of this theory to the natural, social and applied sciences has been recognized (Suppes & Atkinson, 1960; Bartholomew, 1976; Kemeny & Snell, 1976; Domingos & Lowd, 2009; Zucchini, MacDonald, & Langrock, 2009; Davis, 2017).

Figure 2. A. Markov (1856-1922)

2.1 Finite Markov Chains

A MC with a finite set of states is called a *finite MC.* For general facts on finite MCs we refer to Chapter 2 of the book (Voskoglou, 2017). Let us consider a finite MC having n states, say S_1, S_2, …, S_n, with n a nonnegative integer, $n \geq 2$. Denote by p_{ij} the *transition probability* from state S_i to state S_j, i, j = 1, 2,…, n ; then the matrix $A = [p_{ij}]$ is called the *transition matrix* of the MC. The transition from a state to any other state (including its self) is the certain event, therefore it turns out that, for all i =1,…,n, is

$$p_{i1} + p_{i2} + \ldots + p_{in} = 1 \tag{1}$$

The row-matrix $P_k = [p_1^{(k)}\ p_2^{(k)} \ldots p_n^{(k)}]$, called the *probability vector* of the MC, gives the probabilities $p_i^{(k)}$ for the MC to be in state i at step k , for i = 1, 2,…, n and k = 0, 1, 2,… It becomes obvious again that

$$p_1^{(k)} + p_2^{(k)} + \ldots + p_n^{(k)} = 1 \tag{2}$$

Using conditional probabilities on can show (Voskoglou, 2017; Chapter 2, Proposition 1) that for all nonnegative integers k is

$$P_{k+1} = P_k A \tag{3}$$

Therefore an induction on k gives that, for all integers $k \geq 1$ is

$$P_k = P_0 A^k \tag{4}$$

Equations (3) and (4) enable one to make *short run* predictions for the evolution of the corresponding phenomenon which is represented by the MC. In practice most problems involving the use of MCs can be solved by distinguishing between two types of finite MCs, the *absorbing MCs (AMCs)* and the *ergodic MCs (EMCs).*

2.2 Absorbing Markov Chains

A state of a MC is called *absorbing* if, once entered, it cannot be left. Further a MC is said to be an AMC if it has at least one absorbing state and if from every state it is possible to reach an absorbing state, not necessarily in one step.

Working with an AMC with k absorbing states, $1 \leq k < n$, one brings its transition matrix A to its *canonical* (or *standard*) *form* A^* by listing the absorbing states first and then makes a partition of A^* to sub-matrices as follows

$$A^* = \begin{bmatrix} I_k & \vert & O \\ - & \vert & - \\ R & \vert & Q \end{bmatrix} \tag{5}$$

In the above partition of A^*, I_k denotes the unitary k X k matrix, O is a zero matrix, R is the $(n-k)$ X k transition matrix from the non-absorbing to the absorbing states and Q is the $(n-k)$ X $(n-k)$ transition matrix between the non-absorbing states.

It can be shown (Voskoglou & Perdikaris, 1991; Section 2) that the square matrix $I_{n-k} - Q$, where I_{n-k} denotes the unitary $n\text{-}k$ X $n\text{-}k$ matrix is always an invertible matrix. The *fundamental matrix N* of the AMC is defined to be the inverse matrix of $I_{n-k} - Q$. Therefore (Morris, 1978; Section 2.4)

$$N = [n_{ij}] = (I_{n-k} - Q)^{-1} = \frac{1}{D\ (I_{n-k} - Q)} adj\ (I_{n-k} - Q) \tag{6}$$

In equation (6) $D\ (I_{n-k} - Q)$ and $adj\ (I_{n-k} - Q)$ denote the determinant and the *adjoin* of the matrix $I_{n-k} - Q$ respectively It is recalled that the adjoin of a matrix M is the matrix of the *algebraic complements* of the *transpose* matrix M^t of M, which is obtained by turning the rows of M to columns and vice versa. It is also recalled that the algebraic complement m_{ij}' of an element m_{ij} of M is calculated by the equation

$$m_{ij}' = (-1)^{i+j} D_{ij} \tag{7}$$

In equation (7) D_{ij} denotes the determinant of the matrix obtained by deleting the *i*-th row and the *j*-th column of M.

It is well known (Kemeny & Snell, 1976; Chapter 3) that the element n_{ij} of the fundamental matrix *N* gives the mean number of times in state S_i before the absorption, when the starting state of the AMC is S_j, where S_i and S_j are non-absorbing states.

2.3 Ergodic Markov Chains

A MC is said to be an EMC, if it is possible to move, in one or more steps, between any two states. As the number of its steps tends to infinity (long run), an EMC tends to an equilibrium situation, where the probability vector P_k takes a constant price P = [p_1 p_2 … p_n], called the *limiting probability vector* of the EMC (Kemeny & Snell, 1976; Theorem 5.1.1). Therefore, in the equilibrium situation we have, by equation (3), that

$$P = PA \quad \text{, where } p_1 + p_2 + \ldots + p_n = 1 \tag{8}$$

The limiting probabilities of the row-matrix P are the probabilities for the EMC to be in each of its states in the long run, thus expressing the importance (gravity) of each state of the EMC.

Further, if m_{ij} is the mean number of times in state S_i between two successive occurrences of the state S_j, i, j = 1, 2, …, n, we have (Kemeny & Snell, 1976; Theorem 6.2.3) that

$$m_{ij} = \frac{p_i}{p_j} \tag{9}$$

In equation (9) p_i and p_j are the corresponding limiting probabilities.

3. A Markov Chain Model for the Teaching Process

3.1 The Model

We define a finite MC having as states S_i, i = 1, 2,…, 5, the corresponding steps of the teaching process as they have been defined in section 1. From the flow-diagram of Figure 1 it becomes evident that this chain is an AMC with S_1 being its starting state and S_5 being its unique absorbing state. The minimum number of steps before the absorption is 4 and this happens when we have no backward transitions between the three middle states S_2, S_3 and S_4 of the chain. Denote by p_{ij} the transition probability of the MC from state Si to state Sj, for i, j =1, 2,…,5. Then the transition matrix *A* of the chain is equal to

$$A = \begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & p_{32} & 0 & p_{34} & 0 \\ 0 & 0 & p_{43} & 0 & p_{45} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}, \text{ with } p_{32} + p_{34} = p_{43} + p_{45} = 1$$

The transition matrix can be written in its canonical form *A** as follows:

$$A^*=\begin{array}{c} \\ S_5 \\ - \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{array}\begin{array}{c} \begin{array}{cccccc} S_5 & & S_1 & S_2 & S_3 & S_4 \end{array} \\ \left[\begin{array}{cccccc} 1 & | & 0 & 0 & 0 & 0 \\ - & - & - & - & - & - \\ 0 & | & 0 & 1 & 0 & 0 \\ 0 & | & 0 & 0 & 1 & 0 \\ 0 & | & 0 & p_{32} & 0 & p_{34} \\ p_{45} & | & 0 & 0 & p_{43} & 0 \end{array}\right] \end{array} = \left[\begin{array}{ccc} I_1 & | & O \\ - & | & - \\ R & | & Q \end{array}\right].$$

Denote by I_4 the 4X4 unitary matrix. Then the fundamental matrix of the AMC is

$$N = [n_{ij}] = (I_4 - Q)^{-1} = \frac{1}{1-p_{34}p_{43}-p_{32}}\begin{bmatrix} 1-p_{34}p_{43}-p_{32} & 1-p_{34}p_{32} & 1 & p_{34} \\ 0 & -1+p_{34}p_{34} & -1 & -p_{34} \\ 0 & p_{32} & 1 & p_{34} \\ 0 & -p_{32}p_{43} & -p_{43} & -1+p_{32} \end{bmatrix}$$

Therefore, since S_1 is the starting state of the above AMC, it becomes evident that the mean number of steps before the absorption is given by the sum

$$T = n_{11} + n_{12} + n_{13} + n_{14} = \frac{3-2p_{43}p_{34}-p_{32}+p_{34}}{1-p_{34}p_{43}-p_{32}} \tag{10}$$

It becomes also evident that the bigger is T, the more are the student difficulties during the teaching process. Another indication of the student difficulties is the total time spent for the completion of the teaching process. However, the time is usually fixed in a formal teaching procedure in the classroom, which means that in this case T is a measure of the student difficulties.

3.2 A Classroom Application

The application presented here took place at the Graduate Technological Educational Institute of Western Greece during the teaching of the derivative to a group of fresher students of engineering. The instructor used the teaching framework that has been described in our Introduction as follows:

Engagement: The student attention was turned to the fact that the definition of the tangent of a circle as a straight line having a unique common point with its circumference does not hold for other curves (e.g., for the parabola). Therefore, there is a need to search for a definition of the tangent covering all cases and in particular of the tangent at a point of the graph of a given function.

Investigation: The discussion in the class led to the conclusion that the tangent at a point A of the graph of a given function $y=f(x)$ can be considered as the limit position of the secant line of the graph through the points A(a, $f(a)$) and B(b, $f(b)$), when the point B is moving approaching to A either from the left, or from the right (Figure 3). But the slope of the secant line AB is equal to $\frac{f(b)-f(a)}{b-a}$, therefore the slope of the tangent of the graph at A is equal to the limit of the above ratio when b tends to a.

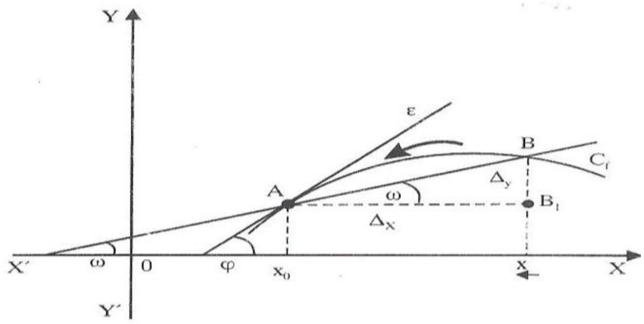

Figure 3. Tangent at a Point of the Graph of a Given Function

Explanation: Based on what it has been discussed at the step of investigation, the instructor presented the formal definition of the derivative number $f'(a)$ at a point A(a, $f(a)$) of a given function $y=f(x)$ as the limit (if there exists) of $\frac{f(b)-f(a)}{b-a}$ when $a \to b$, and of the tangent of the graph of $y=f(x)$ at A as the straight line through A with slope $f'(a)$. Some examples followed of calculating the derivative at a given point of a function and the tangent of its graph at this point. Then the definition of the derivative function $y' = f'(x)$ of the function $y=f(x)$ was given and suitable examples were presented to show that its domain is a subset of the domain of $y=f(x)$.

Deeper Understanding: Here the fact that the derivative $y' = f'(x)$ expresses the rate of change of the function $y=f(x)$ with respect to x was emphasized and its physical meaning was also presented connected to the speed and the acceleration at a moment of time of a moving object under the action of a steady force. The fundamental properties of the derivatives followed (sum, product, composite function, etc.) as well as a list of formulas calculating the derivatives of the basic functions and applications of them.

Evaluation: At the end of the teaching process a number of exercises and problems analogous to those solved in the classroom were given to students on the purpose of checking at home their understanding of the subject. A week later a written test was performed in the classroom enabling the instructor to assess the student progress.

It has been observed that the student reactions during the teaching process led to 2 transitions of the discussion from state S_3 (formalization) back to state S_2 (exploration). Therefore, since from state S_2 the chain moves always to S_3 (Figure 1), we had 3 in total transitions from S_2 to S_3. The instructor also observed 3 transitions from S_4 (assimilation) back to S_3. Therefore, since from state S_3 the chain moves always to state S_4 (Figure 1), we had 4 in total transitions from S_3 to S_4. In other words we had 3+3 = 6 in total “arrivals” to S_3, 2 “departures” from S_3 to S_2 and 4 “departures” from S_3 to S_4. Therefore $p_{32} = \frac{2}{6}$ and $p_{34} = \frac{4}{6}$. In the same way one finds that $p_{43} = \frac{3}{4}$ and $p_{45} = \frac{1}{4}$. Replacing the above values of the transition probabilities to equation (10) one finds that the mean number T of steps before the absorption of the MC is equal to 14. Consequently, since the minimum number of steps before the absorption is 4, the students faced significant difficulties during the teaching process. This means that the instructor should find ways to improve his teaching procedure of the same subject in future.

3.3 An Important Remark

In certain cases it is possible to develop either an AMC or an EMC model for representing the same situation. In case of the teaching process, for example, the flow diagram of Figure 1 could be revised by assuming that, when the teaching process of a subject matter is integrated, then a new process starts by the instructor for teaching the next subject of the course. That means that the teaching process is transferred back from step S_5 to S_1 to be repeated from the beginning again. The revised flow diagram of the teaching process, therefore, takes the form of Figure 4.

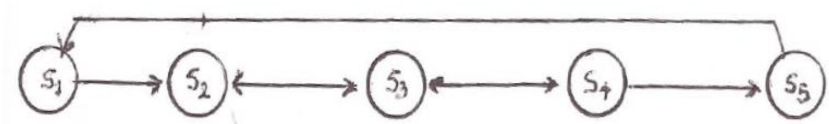

Figure 4. Revised Flow Diagram of the Teaching Process

In this case the resulting MC on the steps of the teaching process is obviously an EMC. Since S_1 is the starting state of the EMC it becomes evident that the sum $m = m_{15} + m_{25} + m_{35} + m_{45}$ calculates the mean number of steps of the EMC between two successive occurrences of the state S_5. Therefore, the mean number of steps for the completion of the teaching process will be m+1, since it includes also the step S_5. With the help of equation (9) one finds that

$$m = \frac{p_1 + p_2 + p_3 + p_4}{p_5} = \frac{1 - p_5}{p_5} \tag{11}$$

The value of the limiting probability p_5 is calculated with the help of the matrix equation (8). In this equation the transition matrix of the EMC differs from the corresponding matrix A of the AMC of section 3.1 only with respect to the last row, where 1 is transferred from the fifth to the first column and its other entries are 0. Performing the necessary calculations, equation (8) leads to a linear system of five equations with respect to the p_i's, i = 1, 2, 3, 4, 5. Adding by members the first four of those equations one finds the fifth one. Thus, replacing the fifth equation with the equation $p_1+p_2+p_3+p_4+p_5 = 1$ and solving the resulting 5X5 linear system one finds the required value of p_5 and with the help of equation (11) calculates m (for more details see Voskoglou, 2019). It becomes evident that, the greater is the value of m the more the student difficulties during the teaching process. Concerning the classroom application of section 3.2, after performing all the necessary calculations one finds that m+1 = T.

4. Conclusion

The application of the ideas of constructivism has become very popular during the last decades for designing teaching models. In this work the teaching process was modelled by introducing an AMC on its consecutive steps. This approach helps the teacher to evaluate the student difficulties and based on it to reorganize his/her plans for teaching the same subject in future. Although the theoretical development of the MC model is quite laborious, its final application is easy and straightforward. It requires only to count the backward movements among the three middle steps S_2, S_3 and S_4 of the teaching process. MCs appear as a very promising tool for use in several other applications of Education (e.g., see Chapter 3 of Voskoglou, 2017) and especially for modelling the smart learning systems of Artificial Intelligence (Voskoglou, 2020). This could be proved as a very interesting suggestion for future research on the subject.

References

Bartholomew, R. C. (1973). *Stochastic Models for Social Processes*. J. Wiley and Sons, London.

Davis, M. H. A. (2017). *Markov Models and Optimization*. Create Space Independent Publishing Platform, Columbia, SC., USA.

Domingos, P., & Lowd, D. (2009). *Markov Logic: An Interface Layer for Artificial Intelligence*. Morgan & Claypool Publishers, Williston, VT, USA.

Kemeny, J. G., Schleifer, A. J., & Snell, J. L. (1962). *Finite Mathematics with Business Applications*. Prentice Hall, London.

Kemeny, J., & Snell, J. l. (1976). *Finite Markov Chains*. Springer-Verlag, New York.

Morris, A. O. (1978). *An Introduction to Linear Algebra*. Van Nostrand-Beinhold Company Ltd., Berkshire, England.

Suppes, P., & Atkinson, R. C. (1960). *Markov Learning Models for Multiperson Interactions*. Stanford University Press, Stanford-California, USA.

Taber, K. S. (2011). Constructivism as educational theory: Contingency in learning, and optimally guided instruction. In J. Hassaskhah (Ed.), *Educational Theory* (Chapter 2, pp. 39-61). Nova Science Publishers, NY.

Voskoglou, M. Gr. (2017). *Finite Markov Chain and Fuzzy Logic Assessment Models: Emerging Research and Opportunities*. Create Space Independent Publishing Platform (Amazon), Columbia, SC, USA.

Voskoglou, M. Gr. (2019). An Application of Ergodic Markov Chain to the Process of Teaching Mathematics. *American Journal of Applied Mathematics and Statistics*, *7*(5), 187-190.

Voskoglou, M. Gr., & Perdikaris, S. C. (1991). A Markov chain model in problem solving. *International Journal of Mathematical Education in Science and. Technology*, *22*, 909-914. https://doi.org/10.1080/0020739910220607

Voskoglou, M. Gr., & Salem, A.-B. M. (2020). Benefits and Limitations of the Artificial with Respect to the Traditional Learning of Mathematics. *Mathematics*, *8*, 611. https://doi.org/10.3390/math8040611

Zucchini, I. L., MacDonald, I. L., & Langrock, R. (2009). *Hidden Markov Models for Time Series*. CRC Press, NY.