A Stochastic Model for the Process of Learning

Michael Gr. Voskoglou, MSc., PhD
Professor of Mathematical Sciences, Faculty of Technological Applications , Graduate Technological Educational Institute (T.E.I.) 263 34 Patras – Greece voskoglou@teipat.gr

Abstract

A Markov chain is introduced to the major steps of the process of learning a subject matter by a group of students in the classroom, in order to obtain a mathematical representation of the above process. A classroom experiment for learning mathematics is also presented illustrating the applicability of our results in practice.

Introduction

There are very many theories and models developed from psychologists and education researchers for the description of the mechanisms of learning. Nowadays it is widely accepted that any instance of learning involves the use of already existing knowledge. Voss (1987) developed an argument that learning consists of successive problem – solving activities, in which the input information is represented of existing knowledge, with the solution occurring when the input is appropriately interpreted.

The whole process involves the following steps: ***Representation*** of the stimulus input, which is relied upon the individual's ability to use contents of his (her) memory to find information, which will facilitate a solution development; ***interpretation*** of the input data, through which the new knowledge is obtained; ***generalization*** of the new knowledge to a variety of situations, and ***categorization*** of the generalized knowledge, so that the individual becomes able to relate the new information to his (her) knowledge structures known as schemata, or scripts, or frames.

The process of learning a subject matter in the classroom

In order to describe the process of learning a subject matter in the classroom one must keep in mind that, as it frequently happens, a learner may not be able to pass successfully through all the steps of the learning process in the time available into the classroom. Therefore it is convenient, for purely technical reasons, to include in this case one more step in the sketch of the process described in the previous section, the step of ***failure to reach categorization.***

We are going to construct a 'flow-diagram' representing the whole process. For this, let us denote by S_i, i=1,2,….,5, the steps of representation, interpretation, generalization, categorization, and failure to reach categorization respectively. The starting state is always S_1. From S_1 the learner proceeds to S_2. Facing difficulties there he (she) may return to S_1 to search for more information that will facilitate the interpretation procedure. Then he (she) must go back to S_2 to continue the process. From S_2 the learner is expected to proceed to S_3, unless if he (she) is unable to interpret the input data during the learning process in the classroom. In this case he (she) proceeds directly to S_5, and the process finishes there for him (her). From S_3 the learner, if he (she) has difficulties during the generalization procedure, may return to S_2 for a better understanding of the subject. Then he (she) comes back to S_3, wherefrom he (she) proceeds either to S_4 or to S_5 and in both cases the process finishes there.

According to the above description the flow-diagram of the process of learning a subject matter in the classroom by a group of students is that shown in Figure 1.

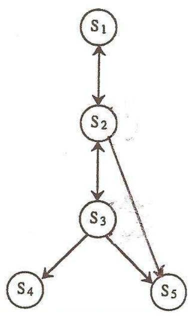

Figure 1: Flow-diagram of the learning process in the classroom

The stochastic (Markov) model

Roughly speaking a ***Markov chain*** is a stochastic process that moves in a sequence of phases through a set of states and has "no memory". This means that the probability of entering a certain state in a certain phase, although it is not necessarily independent of previous phases, depends at most on the state occupied in the previous phase. This is known as the ***Markov property.***

When its set of states is a finite set, then we speak about a ***finite Markov chain.*** For special facts on such type of chains we refer freely to Kemeny & Snell, (1976).

Here we are going to build a Markov chain model for the mathematical description of the process of learning a subject matter in the classroom. For this, assuming that the learning process has the Markov property, we introduce a finite Markov chain having as states the five steps of the learning process described in the previous section. The above assumption is a simplification (not far away from the truth) made to the real system in order to transfer from it to the ***"assumed real system"***. This is a standard technique applied during the mathematical modeling process of a real world problem, which enables the formulation of the problem in a form ready for mathematical treatment (Voskoglou, 2007; section 1).

Denote by p_{ij} the transition probability from state S_i to S_j, for i,j=1,2,3,4,5, then the matrix A=[p_{ij}] is said to be the ***transition matrix*** of the chain.

According to the flow-diagram of the learning process shown in Figure 1 we find that

$$\begin{matrix} & S_1 & S_2 & S_3 & S_4 & S_5 \end{matrix}$$

$$A = \begin{matrix} S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{matrix} \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ p_{21} & 0 & p_{23} & 0 & p_{25} \\ 0 & p_{32} & 0 & p_{34} & p_{35} \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

where we obviously have that $p_{21}+p_{23}+p_{25}=p_{32}+p_{34}+p_{35}=1$

Further let us denote by $\varphi_0,\varphi_1,\varphi_2,\ldots$ .. the successive phases of the above chain , and also denote by

$$P_i=[p_1^{(i)}p_2^{(i)}p_3^{(i)}p_4^{(i)}p_5^{(i)}]$$

the row - matrix giving the probabilities $p_j^{(i)}$ for the chain to be in each of the states S_j, j=1,2,3,4,5 in the phase φi, i=1,2,.... , where we obviously have that

$$\sum_{j=1}^{5} p_j^{(i)} = 1.$$

The above row-matrix is called the ***probability vector*** of the chain at phase φ_i. From the transition matrix A and the flow-diagram of Figure 1 we obtain the tree of correspondence among the several phases of the chain and its states shown in Figure 2.

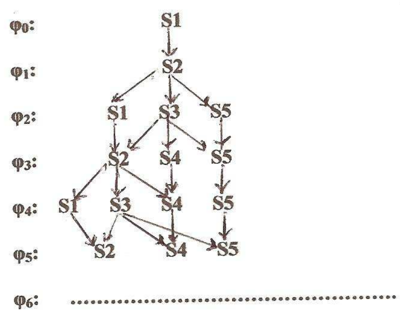

Figure 2: Tree of correspondence among states and phases of the Markov chain

From the above tree becomes evident that P_0 = [1 0 0 0 0], P_1 = [0 1 0 0 0], and P_2=[p_{21} 0 p_{23} 0 p_{25}]. Further it is well known that

$$P_{i+1} = P_iA, \quad i=0,1,2,\ldots .$$

Therefore we find that

$$P_3 = P_2A = [0 \;\; p_{21}+p_{23}p_{32} \;\; 0 \;\; p_{23}p_{34} \;\; p_{23}p_{35}+p_{25}] \quad (1),$$

$P_4=P_3A=$……. , and so on.

Observe now that, when the chain reaches either state S_4, or S_5, it is impossible to leave them, because the learning process finishes there. In other words S_4 and S_5 are absorbing states of the chain. Further, from Figure 1 it becomes evident that from every state it is possible to go to an absorbing state (not necessarily in one step). Thus we have an ***absorbing Markov chain***. Applying standard techniques from theory of absorbing chains we bring the transition matrix A to its ***canonical (or standard) form*** A* by listing the absorbing states first and then we make a partition of A* as follows:

$$A^* = \begin{array}{c} \\ S_4 \\ S_5 \\ \\ S_1 \\ S_2 \\ S_3 \end{array} \begin{array}{c} \begin{array}{cccccc} S_4 & S_5 & & S_1 & S_2 & S_3 \end{array} \\ \begin{bmatrix} 1 & 0 & | & 0 & 0 & 0 \\ 0 & 1 & | & 0 & 0 & 0 \\ - & - & | & - & - & - \\ 0 & 0 & | & 0 & 1 & 0 \\ 0 & p_{25} & | & p_{21} & 0 & p_{23} \\ p_{34} & p_{35} & | & 0 & p_{32} & 0 \end{bmatrix} \end{array}.$$

Symbolically we can write

$$A^* = \begin{bmatrix} I & | & 0 \\ - & | & - \\ R & | & Q \end{bmatrix},$$

where Q is the transition matrix of the non absorbing states and R the transition matrix from the non absorbing to the absorbing states.

Next we consider the ***fundamental matrix*** N of the chain, which is given by

$$N = (I_3-Q)^{-1} = \frac{adj(I_3-Q)}{D(I_3-Q)} \;,$$

where I_3 denotes the 3X3 unitary matrix, $adj(I_3\text{-}Q)$ denotes the adjoin matrix of I_3-Q, and $D(I_3\text{-}Q)$ denotes the determinant of I_3-Q. A straightforward calculation gives that

$$N=\frac{1}{1-p_{23}p_{32}-p_{21}}=\begin{bmatrix} 1-p_{32}p_{23} & 1 & p_{23} \\ p_{21} & 1 & p_{23} \\ p_{21}p_{32} & p_{32} & 1-p_{21} \end{bmatrix}.$$

Finally we consider the 3X2 matrix

$$B=NR=\frac{1}{1-p_{23}p_{32}-p_{21}}\begin{bmatrix} p_{23}p_{34} & p_{25}+p_{23}p_{35} \\ p_{23}p_{34} & p_{25}+p_{23}p_{35} \\ (1-p_{21})p_{34} & p_{32}p_{25}+p_{35}(1-p_{21}) \end{bmatrix}.$$

We write symbolically $B=[b_{ij}]$, with i=1,2,3 and j=4,5. It is well known then that b_{ij} gives the probability that, starting at state S_i, the process is absorbed at state S_j. Thus the probability for a learner to pass successfully through all the states of the learning process in the classroom is given by

$$b_{14}=\frac{p_{23}p_{34}}{1-p_{23}p_{32}-p_{21}} \quad (2).$$

The calculation of b_{14} enables the teacher to check the efficiency of his (her) lectures. It also could be used either as a measure of comparison of the efficiencies of the lectures of different teachers, or as a measure of the learning abilities of different groups of students. The following classroom experiment for learning mathematics illustrates the applicability of our model in practice.

A classroom experiment for learning mathematics

The present experiment took place recently at the Graduate Technological Educational Institute of Patras (Greece), when I was teaching to a group of 30 students of the School of Technological Applications (i.e. to future engineers) the use of the derivative for the maximization and minimization of a function. During my 2 hours lecture I used the method of rediscovery (Voskoglou, 1997). Thus, after a short introduction to the subject, I left my students to work alone on their papers. I was inspecting their works, and from time to time I was giving them some instructions, or hints. After the basic theoretical conclusions I gave them some exercises to solve first, and at the final step some problems including applications to constructions and economics.

During the experiment I found that 4 students were completely unable to understand the subject. Also 10 students faced difficulties before understanding the basic ideas (they looked back to their notes of my previous lectures and/or asked for help). Furthermore 5 students, although it seemed that they understood the basic theoretical ideas, were unable to apply them in order to solve the given exercises and problems. The other 21 students solved the exercises, but 8 of them faced difficulties before they came through. At the last step 10 students solved the problems and 11 they didn't (or solved a small part of them). Interpreting these data with respect to the flow-diagram of Figure 1 I was led to the following conclusions, which are represented in Figure 3.

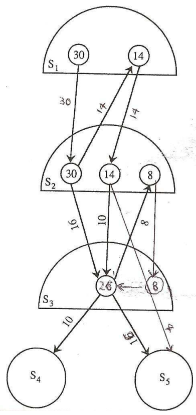

Figure 3: Representation of the experiment's data

- Initially the 30 students proceeded from S_1 to S_2, but 14 of them faced difficulties to interpret the input data. Therefore they returned to S_1 to search for more information that will facilitate the interpretation procedure, wherefrom they came back to S_2. Finally 4 of them reached directly the absorbing state S_5, because they didn't manage to interpret the new knowledge.
- The remaining 26 students proceeded to S_3, but 8 of them faced difficulties to generalize the new knowledge to a variety of situations, and they returned to S_2 for a better understanding of the new information. Then they came back to S_3.
- At the last step 10 students, who solved the exercises and problems, completed successfully the learning process in the classroom and therefore they reached the absorbing state S_4. The other 16 students, i.e. 5 students who didn't manage to solve the exercises and problems, and 11 who solved the exercises, but not the problems, reached the absorbing state S_5.

Therefore, since we had a total of 52 'arrivals' to S_2, 14 'departures' from S_2 to S_1, 34 'departures' from S_2 to S_3, and 4 'departures' from S_2 to S_5, it follows that $p_{21}=\frac{14}{52}$, $p_{23}=\frac{34}{52}$, and $p_{25}=\frac{4}{52}$. In the same way one finds that $p_{32}=\frac{8}{34}$, $p_{34}=\frac{10}{34}$, and $p_{35}=\frac{16}{34}$.

Replacing the values of the p_{ij}'s in equalities (1) and (2) of the previous section we get that $P_3=[0 \ \frac{22}{52} \ 0 \ \frac{10}{52} \ \frac{20}{52}]$ and $b_{14}=\frac{1}{3}$. Interpreting these data with respect to our model we find that the probabilities for a student to be in phase φ_3 of the process of learning in the classroom (i.e. 3 phases after its start) at the steps of representation, interpretation, generalization, categorization, or failure to reach categorization are approximately 0, 42,31%, 0, 19,23%

and 38,46% respectively, while the probability to pass successfully through all the steps of the process is approximately 33,33%.

Remarks and further examples

Most real world problems concerning applications of finite Markov chains can be solved by distinguishing between two types of such chains, the absorbing (e.g. the case of our model in the present paper) and the ergodic ones (Voskoglou, 2006; section 3). We recall that a Markov chain is said to be an ***ergodic chain***, if it is possible to go between any two states, not necessarily in one step.

In Voskoglou (1996) an ergodic chain is introduced for the study of the analogical problem-solving process in the classroom, while in Voskoglou and Perdikaris (1991) the problem-solving process (in general) is described through the introduction of an absorbing Markov chain to the main steps of the process.

In Voskoglou (1994) an absorbing Markov chain is introduced to the major steps through which one would proceed in order to effect the study of a real system (modelling process). An alternative form of the above model is introduced in Voskoglou (2007) for the description of the mathematical modelling process in the classroom. In this case it is assumed that after the completion of the solution process of each problem a new problem is given from the teacher to the class and therefore the process is repeated again. Thus the resulting Markov chain is an ergodic one.

In Voskoglou (2000) an absorbing Markov chain is introduced to the main steps of the decision making process performed in order to choose the best among the existing solutions of a given problem, and examples are presented to illustrate the applicability of the model to "real" decision making situations.

We could mention very many other known applications of Markov chains for the solution of real world problems in almost every sector of the human activity, but this is rather out of the scope of the present paper.

Final conclusions

The theory of Markov chains is a successful combination of Linear Algebra and Probability, which enables one to make forecasts for the evolution of various phenomena of the real world.

In the present paper we built a Markov model for the description of the process of learning a subject matter by a group of students in the classroom. In this way we succeeded to calculate the probabilities for a student to be at any of the major steps of the learning process in each of its phases in the classroom, as well as the probability to pass successfully through all the steps of the learning process in the classroom. Our results are illustrated by a classroom experiment for learning mathematics.

References

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

Voss, J. F. (1987), Learning and transfer in subject learning: A problem solving model, Int.J. Educ. Research, 11, 607-622.

Voskoglou, M. Gr. and Perdikaris, S. C. (1991), A Markov chain model in problem-solving, Int. J. Math. Educ. Sci. Technol., 22, 909-914.

Voskoglou, M. Gr. (1994), An application of Markov chains to the process of modelling, Int. J. Math. Educ. Sci. Technol., 25, 475-480.

Voskoglou, M. Gr. (1996), An application of ergodic Markov chains to analogical problem solving, The Mathematics Education (India), Vol.XXX, No 2, 96-108.

Voskoglou, M. Gr. (1997), Some remarks of the use of rediscovery in the teaching of mathematics, Proceedings of the 1st Mediterranean Conference on Mathematical Education, 124-128, Cyprus.

Voskoglou, M. Gr. (2000), An application of Markov chains to Decision Making, Studia Kupieckie (Univ. of Lodz), 6, 69-76.

Voskoglou, M. Gr. (2006), Applications of Markov chains to Management and Economics, Proceedings of MASSEE International Congress (MICOM 2006). Dodunekov, S., et al (Eds), 140-141, Cyprus.

Voskoglou, M. Gr. (2007), A stochastic model for the modelling process. In: Chaines, Cr. , et al (Eds), Mathematical Modelling: Education, Engineering and Economics (ICTMA 12), 149-157, Horwood Publishing, Chichester.