Applying the Absorbing and the Ergodic Markov Chain Theory to CBR

Michael Gr. Voskoglou

***Abstract*—**In the present work it is shown that in cases where is possible to apply either the absorbing or the ergodig Markov chain theory for modeling purposes - by making the proper modifications to the corresponding real situation – the results obtained are equivalent. For this, the Case-Based Reasoning paradigm is used, which is the process of solving problems (frequently with the help of computer systems) based on the solutions of previously solved analogous problems

***Keywords*—**Absorbing Markov Chain (AMC), Case-Based Reasoning (CBR), Ergodic Markov Chain (EMC), Problem Solving (PS).

I. Introduction

IN an earlier work we have modeled the Case-Based Reasoning (CBR) process by introducing an Absorbing Markov Chain (AMC) on its steps and through it we have obtained a method for evaluating the effectiveness of a CBR system [1].

Here, by modifying properly our hypothesis about the CBR process, we manage to do the same thing with an Ergodic Markov Chain (EMC) on the steps of the CBR process instead of the AMC. For reasons of comparison the same examples are used with [1]. It is finally shown that the results obtained in both cases are the same.

The rest of the paper is organized as follows: In Section II a brief account is given of the earlier AMC model for the CBR process. In Section III the EMC model is developed in detail and through it a measure is obtained in Section IV for the effectiveness of a CBR system. The paper closes with the general conclusions stated in Section V.

II. THE AMC MODEL FOR CBR

A. Absorbing Markov Chains

It is recalled that a Markov Chain (MC) is a stochastic process moving in a sequence of phases (steps) through a set of states and having only a *one step memory* (Markov property). In other words, the probability of entering a certain state at a certain phase depends only on the state occupied in the previous phase and not in older phases. However, in practical applications the Markov property is usually weakened by accepting that the above probability, although it mainly depends on the state occupied in the previous phase, it may not be co0mpletely independent from older phases [2].

When its set of its states is a finite set, the corresponding MC is called a Finite MC (FMC). For general facts on FMC's we refer to Chapter 2 of the book [3].

A state of a MC is called absorbing if once entered cannot be left and a MC is called 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..

B. The CBR Process

Roughly speaking CBR is the process of solving problems (frequently with the help of computers) based on the solutions of previously solved analogous problems. For more details and examples about the CBR process we refer to [4] and to the relevant references contained in it.

CBR has been formalized for purposes of computer and human reasoning as a four steps process involving the following actions:

- R_1: *Retrieve* the most similar to the new problem past case.
- R_2: *Reuse* the information and knowledge of the retrieved case for designing the solution of the new problem.
- R_3: *Revise* the proposed solution for use with the new problem. .
- R_4: *Retain* the part of this experience likely to be useful for future problem-solving.

Through the revision the solution is tested for success. If successful, the revised solution is directly retained in the CBR system's library; otherwise it is repaired and evaluated again. When the final result is a failure, the system tries to compare it to a previous analogous failure (transfer from R_3 back to R_1) and uses it in order to understand the present failure, which is finally retained in the library. The CBR process is completed in R_4. According to the above description the flow diagram of the CBR process is that shown in Figure 1 that has been taken from [1]

M, Gr. Voskoglou is with the Graduate Technological Educational Institute (TEI) pf Western Greece, 26334 Patras, Greece.

Phone - Fax: 0002610328631, Mobile: 00306978600391, e-mail: voskoglou@teiwest.gr ; mvoskoglou@gmail.com,

URL: http://eclass.teipat.gr/courses/523102

.

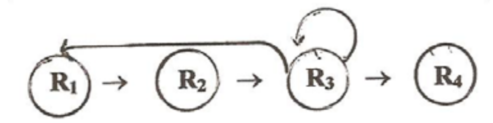

Fig. 1: The flow diagram of the CBR process

C. The model

We introduce a finite MC having as states the four steps of the CBR process. Denote by p_{ij} the **transition probability** from state R_i to R_j, for i, j=1, 2, 3, 4. Then, with the help of Fig. 1, one finds that the **transition matrix** of the MC is

$$A = \begin{array}{c} \\ R_1 \\ R_2 \\ R_3 \\ R_4 \end{array} \begin{array}{c} \begin{array}{cccc} R_1 & R_2 & R_3 & R_4 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ p_{31} & 0 & p_{33} & p_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array},$$

with $p_{31} + p_{33} + p_{34} = 1$ (probability of the certain event).

We obviously have an AMC with R_4 being its unique absorbing state. Applying standard techniques of the theory of AMC's we bring A to its **canonical form** A* by listing its absorbing state first and then we make a partition of A* to sub-matrices as follows:

\

$$A^* = \begin{array}{c} \\ R_4 \\ - \\ R_1 \\ R_2 \\ R_3 \end{array} \begin{array}{c} \begin{array}{ccccc} R_4 & & R_1 & R_2 & R_3 \end{array} \\ \begin{bmatrix} 1 & | & 0 & 0 & 0 \\ - & - & - & - & - \\ 0 & | & 0 & 1 & 0 \\ 0 & | & 0 & 0 & 1 \\ p_{34} & | & p_{31} & 0 & p_{33} \end{bmatrix} \end{array}.$$

$$= \begin{bmatrix} I & | & 0 \\ - & | & - \\ R & | & Q \end{bmatrix}.$$

Here Q stands for the transition matrix of the non absorbing states. Then the **fundamental matrix** of the chain is given by

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

In the above equation I_3 denotes the 3X3 unitary matrix, adj $(I_3 - Q)$ denotes the adjoin matrix and $D(I_3 - Q)$ denotes the determinant of
I_3-Q. Therefore, a straightforward calculation gives that

$$N = \frac{1}{1 - p_{31} - p_{33}} \begin{bmatrix} 1 - p_{33} & 1 - p_{33} & 1 \\ -p_{31} & 1 - p_{33} & 1 \\ p_{31} & p_{31} & 1 \end{bmatrix}$$

$$= [n_{ij}]$$

It is well known ([5], Theorem 3.2.4) that the entry n_{ij} of N gives the mean number of times at state R_j when the chain is started in state R_i. Therefore, since the present chain is always starting from R_1, the mean number of its phases before the absorption is given by the sum

$$t = n_{11} + n_{12} + n_{13} = \frac{3 - 2p_{33}}{1 - p_{31} - p_{33}}$$

Therefore, the mean number of steps for the completion of the CBR process is t+1. It becomes evident that the bigger is the value of t, the greater is the difficulty encountered for the solution of the given problem via the CBR process. Another indication of that difficulty is the total time spent for the completion of the CBR process, which however is negligible in practice when using computers.

III. THE EMC MODEL FOR CBR

A. The general MC Model

Given a FMC, the row-matrix
$P_k = [p_1^{(k)}\ p_2^{(k)} \ldots p_n^{(k)}]$, known as 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,…. Obviously we have that

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

Using conditional probabilities on can show ([3], Chapter 2, Proposition 1) that for all non negative integers k we have

$$P_{k+1} = P_k A \qquad (1)$$

Therefore a straightforward induction on *k* gives that

$$P_k = P_0 A^k \qquad (2)$$

Equation (2) enables one to make **short run** forecasts for the evolution of the various situations that can be represented by a finite MC.

B. Ergodic Markov Chains

A MC is said to be an EMC, if it is possible to go between any two states, not necessarily in one step.

It is well known ([7], Theorem 5.1.1) that, as the number of its steps tends to infinity (long run), an EMC tends to an equilibrium situation, in which the probability vector P_k takes a constant value $P = [p_1\ p_2\ \ldots.\ p_n]$, called the limiting probability vector of the EMC. Therefore, as a direct consequence of equation (1), the equilibrium situation is characterized by the equation

$$P = P\text{A} \quad (3)$$

The entries of P express the probabilities of the EMC to be in each of its states in the long run, or in other words the importance (gravity) of each state of the EMC.

Let us now demote with m_{ij} the mean number of times in state S_i between two successive occurrences of the state S_j, i, j = 1, 2, …., n. It is well known then ([5], Theorem 6.2.3) that

$$m_{ij} = \frac{p_i}{p_j} \quad (4),$$

where p_i and p_j are the corresponding limiting probabilities.

C. The revised model for CBR

Let us now assume that, when the CBR process is completed in R_4, a new analogous problem is forwarded to the CBR system for solution. Therefore the process is transferred back to R_1 and a new circle is repeated. According to the above assumption the flow diagram of Fig. 1 for the CBR process can be revised as shown in Fig. 2.

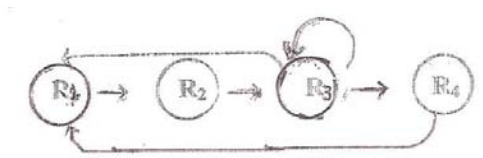

Fig. 2: The revised flow diagram of the CBR process

Accordingly, the transition matrix of the MC introduced on the steps of the CBR process, which is obviously now an EMC, takes the form

$$\text{A} = \begin{array}{c} \\ R_1 \\ R_2 \\ R_3 \\ R_4 \end{array}\begin{array}{c} \begin{array}{cccc} R_1 & R_2 & R_3 & R_4 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ p_{31} & 0 & p_{33} & p_{34} \\ 1 & 0 & 0 & 0 \end{bmatrix} \end{array}.$$

By equation (3) one finds that in the long run we have for the equilibrium situation of the EMC that

$$[p_1 \quad p_2 \quad p_3 \quad p_4] = [p_1 \quad p_2 \quad p_3 \quad p_4]\text{A} \text{ or}$$
$$[p_1 \quad p_2 \quad p_3 \quad p_4] =$$
$$[p_3p_{31} + p_4 \quad p_1 \quad p_2 + p_3p_{33} \quad p_3p_{34}]$$

Consequently it turns out that

$$p_1 = p_3p_{31} + p_4, \quad p_2 = p_1, \quad p_3 = p_2 + p_3p_{33}, \quad p_4 = p_3p_{34} \quad (5)$$

Adding by members the first three of the equations (5) one

$$p_1 + p_2 + p_3 = p_3p_{31} + p_4 + p_1 + p_2 + p_3p_{33}$$

finds that $\Leftrightarrow p_3 = p_4 + p_3(p_{31} + p_{33})$

$$\Leftrightarrow p_3 = p_4 + p_3(1 - p_{34}) \Leftrightarrow p_4 = p_3p_{34}$$

Therefore, the fourth of the equations (5) is equivalent to the rest of them. Consider now the linear system L of the first three of the equations (5) and of the equation $p_1+p_2+p_3+p_4 = 1$. It is straightforward to check that the determinant of L is equal to

$$D = \begin{vmatrix} 1 & 0 & -p_{31} & -1 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & p_{33}-1 & 0 \\ 1 & 1 & 1 & 1 \end{vmatrix} = 4 - 3p_{33} - p_{31}$$

Also $$D_{p_1} = \begin{vmatrix} 0 & 0 & -p_{31} & -1 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & p_{33}-1 & 0 \\ 1 & 1 & 1 & 1 \end{vmatrix} = 1 - p_{33}$$

Therefore, by the Cramer's rule one finds that

$$p_1 = \frac{D_{p_1}}{D} = \frac{1 - p_{33}}{4 - 3p_{33} - p_{31}} = p_2 \quad (6).$$

In the same way one finds that

$$P_3 = \frac{1}{4 - 3p_{33} - p_{31}} \text{ and } P_4 = \frac{1 - p_{33} - p_{31}}{4 - 3p_{33} - p_{31}}$$
$$= \frac{p_{34}}{4 - 3p_{33} - p_{31}} \quad (7).$$

The values of the p_i's give the probabilities of the CBR process to be in step Ri in the long run, i = 1, 2, 3, 4. Furthermore, since R_1 is the starting state of the EMC it becomes evident that the sum $m = m_{14} + m_{24} + m_{34}$ calculates the mean number of steps of the EMC between two successive occurrences of the state R_4. Therefore, the mean number of steps for the completion of the CBR process will be m+1, since it includes also the step R_4. With the help of equation (4) one finds that

$$m = \frac{p_1 + p_2 + p_3}{p_4} = \frac{1 - p_4}{p_4} \quad (8)$$

It becomes evident that the greater is the value of *m*, the more are the difficulties during the CBR process. The other factor indicating those difficulties, i.e. the total time spent for the CBR process, is practically negligible when using computers.

D) Example

For reasons of comparison the present example has been taken from Section 3 of [1]

A physician, in order to determine the disease and suggest the analogous treatment to a patient, takes into account the diagnosis and treatment of a previous patient having similar symptoms. If the initial treatment fails to improve the health of the patient, then the physician either revises the treatment (stay to R_3 for two successive phases), or gets a reminding of a previous similar failure and uses the failure case to improve the understanding of the present failure (transfer from R_3 to R_1).

Assume that the recorded statistical data show that the probabilities of a straightforward cure of the patient as well as of each of the above two reactions of the physician in case of failure of the initial treatment are equal to each other. Therefore $p_{31} = p_{33} = p_{34} = \frac{1}{3}$. Then equations (6) and (7) give that $p_1 = p_2 = \frac{1}{4}$,

$p_3 = \frac{3}{8}$ and $p_4 = \frac{1}{8}$. That means that in this case the step of revision (R_3) has the greatest gravity among the steps of the CBR process.

Also equation (8) gives that $m = 7$. Consequently the mean number of steps for the completion of the CBR process is 8. The same outcome was found in [1] by using the AMC model for the CBR process that has been presented in Section II.

IV. ASSESSING THE EFFECTIVENESS OF A CBR SYSTEM

A) Calculating the effectiveness of a CBR system

Let us consider a CBR system including a library of *n* recorded past cases and let m_i be the outcome of equation (8) for the case c_i, i=1,2,…, *n*. Each m_i can be stored in the system's library together with the corresponding case. Then we define the system's **effectiveness,** say *E*, to be the mean value of the m_i's of its stored cases, i.e. we have that

$$E = \frac{\sum_{i=1}^{n} m_i}{n} \quad (9).$$

The more problems are solved through the given CBR system, the bigger becomes the number *n* of the stored cases in its library and therefore the value of *E* is changing. As *n* is increasing it is normally expected that *E* will decrease, because the values of the m_i*'s* of the new stored cases will be normally decreasing. In fact, the bigger is *n*, the greater would be the probability for a new case to have minor differences with a past case, and therefore the less would be the difficulty of solving the corresponding problem via the CBR process. Thus we could say that a CBR system behaves well if, when *n* tends to infinity, then its effectiveness tends to 3, which, according to the flow-diagram of Fig. 2, is equal to the minimum number of steps between two successive occurrences of R_4.

B) Example

For reasons of comparison the present example has been taken from Section 4 of [1]

Consider a CBR system that has been designed in terms of Schank's model of dynamic memory for the representation of cases [6]. The basic idea of this model is to organize specific cases, which share similar properties, under a more general structure called a generalized episode (GE). During the storing of a new case, when a feature of it matches a feature of an existing past case, a new GE is created. Hence the memory structure of the system is in fact dynamic, in the sense that similar parts of two case descriptions are dynamically generalized to a new GE and the cases are indexed under this GE by their different features.

In order to calculate the effectiveness of a system of this type we need first to calculate the effectiveness of each of the GE's contained in it. For example, assume that the given system contains a GE including three cases, say c_1, c_2 and c_3. Assume further that c_1 corresponds to a straightforward successful application of the CBR process, that c_2 is the case described in the example of the previous section, and that c_3 includes one "return" from R_3 to R_1 and two "stays" to R_3. Then $m_1 = 3$ and $m_2 = 7$. For calculating m_3 observe first that $p_{31} = p_{34} = \frac{1}{4}$ and $p_{33} = \frac{1}{2}$. Therefore, the second of equations (7) gives that $p_4 = \frac{1}{9}$ and equation (8) gives that $m_3 = 8$. Therefore, the effectiveness of this GE is equal to $E = \frac{3+7+8}{3} = 6$.

The same outcome was found in [1] by defining a CBR system's effectiveness in terms of the AMC model for the CBR process that has been presented in Section II.

Next we calculate the effectiveness of all the other GE's of the CBR system the mean value of which gives the system's total effectiveness.

Notice that a complex GE may contain some more specific GE's including some common cases (see Figure 3 in page 12 of [7]). Then we calculate the effectiveness of the complex GE by considering all its cases only once, regardless if they belong or not to one or more of the specific GE's contained in it.

There are also alternative models for the representation of cases in a CBR system like the category and exemplar model of Porter and Bareiss [8], the Rissland's and Ashley's HYPO system for legal reasoning [9], the MBR model of Stanfill & Waltz [10], etc. The process of calculating the effectiveness of a CBR system cfunctioning by one of those alternative models is analogous to the process described in the previous example for the Schank's model.

V. CONCLUSION

The theory of MC's, that is based on probability theory and uses Linear Algebra as its basic tool, can be applied for making short and long run forecasts for the evolution of certain situations characterized by randomness.

In the paper at hands we have modeled the CBR process by introducing an EMC on its steps. This model provides the same outcomes with the AMC for CBR that we have developed in an earlier work [1]. For reasons of comparison we have properly modified here the same examples that have been used in [1].

Our earlier research contains also applications of MC's to Mathematics Education (learning mathematics, problem solving, mathematical modeling, etc.; e.g. see Chapter 3 of [3]), while the application of the MC theory for modeling other human activities is also among the priorities of our future research.

References

[1] M.Gr. Voskoglou, "An Absorbing Markov Chain Model for Case-Based Reasoning", International Journal of Computers (IARAS), 2, 52-60, 2017.

[2] J.G. Kemeny, A.J. Schleifer, & J.L. Snell, Finite Mathematics with Business Applications, Prentice Hall, London, 1962.

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

[4] M. Gr. Voskoglou & A-B. M Salem, "Analogy-Based and Case-Based Reasoning: Two Sides of the Same Coin", International Journal of Applications of Fuzzy Sets and Artificial Intelligence, 4, 5-51, 2014.

[5] J.G. Kemeny & J.L. Snell, Finite Markov Chains, Springer-Verlag, New York, 1976.

[6] R. Schank, Dynamic memory: A theory of reminding and learning in computers and people, Cambridge University Press, 1982.

[7] A. Aamodt & E. Plaza, "Case- Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches", Artificial Intelligence Communications, 7(1), 39- 52, 1994.

[8] B. Porter & B. Bareiss, "PROTOS: An experiment in knowledge acquisition for heuristic classification tasks", Proceedings of the 1st International Meeting on Advances in Learning (IMAL), pp. 159-174, Les Arcs, France, 1986.

[9] E. Rissland,, "Examples in legal reasoning: Legal hypotheticals", Proceedings of the 8th International Joint Conference on Artificial Intelligence (IJCAI), Karlsruhe, 1983

[10] C. Stanfill & D. Waltz, "The memory-based reasoning paradigm", Case-based reasoning Proceedings from a workshop, pp. 414-424, Morgan Kaufmann Publications, Clearwater Beach, Florida, 1988.

Michael Gr. Voskoglou (B.Sc., M.Sc., M.Phil. , Ph.D. in Mathematics) is currently an Emeritus Professor of Mathematical Sciences at the Graduate Technological Educational Institute of Western Greece.

He used to be a Visiting Researcher at the Bulgarian Academy of Sciences (1997-2000), a Visiting Professor at the Universities of Warsaw (2009), of Applied Scinces of Berlin (2010) and at the National Institute of Technology of Durgapur, India (2016).

Prof. Voskoglou is the author of 16 books (9 in Greek and 7 in English language) and of about 500 papers published in reputed mathematical journals and proceedings of international conferences of more than 30 countries of the five continents (2018 Google Scholar Citations: 2401, H-idex 24).

He is also the Editor in Chief of the "International Journal of Applications of Fuzzy Sets and Artificial Intelligence" (ISSN 2241-1240), reviewer of the AMS and member of the Editorial Board or referee in many mathematical journals. His research interests include Algebra, Fuzzy Sets, Markov Chains, Artificial Intelligence and Mathematics Education.