

---

## Case-based reasoning in computers and human cognition: a mathematical framework

---

Michael Gr. Voskoglou

School of Technological Applications,  
Graduate Technological Educational Institute (T.E.I.),  
263 34 Patras, Greece  
E-mail: voskoglou@teipat.gr  
E-mail: mvosk@hol.gr

**Abstract:** Case-based reasoning (CBR) is a recent approach to problem solving and learning for computers and humans. In this paper, we introduce a finite Markov chain on the main steps of the CBR process. Using this approach we succeed in calculating the probabilities for the CBR process to be at a certain step in a certain phase of the solution of a real-world problem and we obtain a measure of the effectiveness of a CBR system in solving similar new problems. Next, the steps of the CBR process are represented as fuzzy subsets of a set of linguistic labels characterising the success of the CBR process in each of the above steps. Thus, we build a fuzzy model for the representation of a CBR system and we use the total possibilistic uncertainty as a measure of its effectiveness in solving new related problems. Examples are also given to illustrate our results.

**Keywords:** finite Markov chains; case-based reasoning; CBR; cognition; fuzzy sets; uncertainty; machine intelligence.

**Reference** to this paper should be made as follows: Voskoglou, M.Gr. (2013) 'Case-based reasoning in computers and human cognition: a mathematical framework', *Int. J. Machine Intelligence and Sensory Signal Processing*, Vol. 1, No. 1, pp.3–22.

**Biographical notes:** Michael Gr. Voskoglou (BSc, MSc, MPhil, PhD in Mathematics) is currently a Professor of Mathematical Sciences at the Graduate Technological Educational Institute of Patras, Greece. He is the author of eight books (seven in Greek and one in English language) and of about 240 papers published in reputed journals and proceedings of conferences of 22 countries in five continents, with many references from other researchers. He is also a reviewer of the AMS and member of the editorial board or referee in several mathematical journals. His research interests include algebra, fuzzy sets, Markov chains and mathematics education.

---

### 1 Introduction

Case-based reasoning (CBR) is a recent approach to problem-solving and learning for computers and people that has got a lot of attention over the last few years. Broadly construed, CBR is the process of solving new problems based on the solutions of similar past problems. The term *problem-solving* is used here in a wide sense, coherent with

common practice within the area of knowledge-based systems in general. This means that it is not necessarily the finding of a concrete solution to an application problem, it may be any problem put forth by the user. For example, to justify or criticise a solution proposed by the user, to interpret a problem situation, to generate a set of possible solutions, or generate explanations in observable data, are also problem-solving situations.

A lawyer, who advocates a particular outcome in a trial based on legal precedents, or an auto mechanic, who fixes an engine by recalling another car that exhibited similar symptoms, or even a physician, who considers the diagnosis and treatment of a previous patient having similar symptoms to determine the disease and treatment for the patient in front of him, are using CBR; in other words CBR is a prominent kind of analogy making.

In CBR's terminology, a case denotes a problem situation. A previously experienced situation, which has been captured and learned in a way that it can be reused in the solving of future problems, is referred as a *past case*, *previous case*, *stored case*, or *retained case*. Correspondingly, a *new case*, or *unsolved case*, is the description of a new problem to be solved. The CBR systems expertise is embodied in a collection (library) of past cases rather, than being encoded in classical rules. Each case typically contains a description of the problem plus a solution and/or the outcomes. The knowledge and reasoning process used by an expert to solve the problem is not recorded, but is implicit in the solution. A case-library can be a powerful corporate resource allowing everyone in an organisation to tap in the corporate library, when handling a new problem. CBR allows the case-library to be developed incrementally, while its maintenance is relatively easy and can be carried out by domain experts.

CBR's coupling to learning occurs as a natural by-product of problem-solving. When a problem is successfully solved, the experience is retained in order to solve similar problems in future. When an attempt to solve a problem fails, the reason for the failure is identified and remembered in order to avoid the same mistake in future. This process was termed as *failure-driven learning* (Schank, 1981). Thus, CBR is a cyclic and integrated process of solving a problem, learning from this experience, solving a new problem, etc. Effective learning in CBR, sometimes referred as *case-based learning*, requires a well worked out set of methods in order to extract relevant knowledge from the experience, integrate a case into an existing knowledge structure and index the case for later matching with similar cases.

For more details about the history, methodology, applications and development trends of CBR the reader may look at Voskoglou (2008, 2011) and their references.

## **2 CBR in computers and human cognition**

The first trails into the CBR field have come from the study of analogical reasoning (Gentner, 1983) and – further back – from theories of concept formation, problem-solving and learning within philosophy and psychology (e.g., Wittgenstein, 1953; Smith and Medin, 1981; etc). For example, Wittgenstein observed that concepts, which are part of the natural world, like bird, tree, chair, car, etc., are *polymorphic* and therefore it is not possible to come up with a classical definition, but it is better to be defined by their sets of instances, or cases.

Memory is the repository of knowledge and therefore the question is what kind of memory accounts for observed cognitive behaviours. A leading theory has been the

semantic memory model. Psychologists devoted much attention to this theory (Collins and Quillian, 1969; Rumelhart et al., 1972; Kintsch, 1972; etc.), as have artificial intelligence (AI) researchers (Quillian, 1968; Woods, 1975; etc.), who attempted to create computer programmes that model cognitive processes. The *semantic memory* model typically represents static facts about the world and therefore this type of knowledge does not change over time. However, it was observed that this model did not account for all the data; e.g., it does not explain how knowledge is incorporated into memory and where does the information come from.

To address these and other questions Tulvin (1972, 1983) proposed a theory of *episodic memory* as an adjunct to semantic memory. Episodic memory receives and stores information about temporally dated episodes or events. The retrieval of information from the episodic store serves as a special type of input into episodic memory and thus changes the contents of the episodic memory store.

CBR traces its roots in AI to the work of Roger Schank and his students at Yale University, USA in the early 1980s. Schank (1975) proposed a conceptual memory that combined semantic memory with Tulvin's episodic memory. *Scripts* (Schank and Abelson, 1977) were proposed as a knowledge structure for the conceptual memory. The acquisition of scripts, which are analogous to Minsky's (1975) *frames*, is the result of repeated exposure to a given situation. As a psychological theory of memory scripts suggested that people would remember an event in terms of its associated script. However, an experiment by Bower et al. (1979) showed that subjects often confused events that have similar scripts: e.g., one might mix up waiting room scenes from a visit to a doctor with a visit to a dentist. These data required a revision in script theory. Schank (1979, 1980) postulated a more general structure to account for the diverse and heterogeneous nature of episodic memory, called memory organisation packet (MOP). MOP's can be viewed as metascripts; e.g., a professional office visit MOP can be instantiated and specified for both the doctor and the dentist, thus providing the basis for confusion between these two events.

However, more important than the MOP knowledge was the new emphasis on the basic memory processes of reminding and learning. Schank proposed a theory of learning based on reminding, according to which we can classify a new episode in terms of past similar cases. Schank's model of *dynamic memory* (Schank, 1982) was the basis of the earliest CBR systems that might be called case-based reasoners: Kolodner's CYRUS (1983) and Lebowitz's IPP (1983). The basic idea of Schank's model is to organise specific cases, which share similar properties, under a more general structure called a generalised episode (GE). During storing of a new case, when a feature of it matches a feature of an existing past case, a new GE is created. Thus, the organisation and structure of memory is dynamic, i.e., changes over time. Similar parts of two case descriptions are generalised in to a new GE and the cases are indexed under this GE by their different features. Concerning CYRUS, it was basically a question-answering system with knowledge of the various travels and meetings of former US Secretary of State Cyrus Vance and the case memory model developed for this system has later served as basis for several other CBR systems including MEDIATOR, PERSUADER, JULIA, etc.

An alternative approach for the representation of cases in a CBR system is the *category and exemplar model*, produced by the work of Bruce Porter and his group at the University of Texas. In this model the case memory is embedded in a network of categories, cases and index pointers. Each case is associated with a category. Finding a

case in the case library that matches an input description is done by combining the features of the new problem case into a pointer to the category that shares most of these features. A new case is stored in a category by searching for a matching case and by establishing the appropriate feature indices. The above model applied first to the PROTOS system (Porter and Bareiss, 1986; Bareiss, 1989), where emphasis is given to the combination of the general with the specific knowledge obtained through the study of cases.

An alternative case memory model was produced by the work of Edwina Rissland and her group at the University of Massachusetts, interested in the role of precedence reasoning in legal judgements (Rissland, 1983). This work resulted in the HYPO (Ashley, 1991) and CABARET (Skalak and Rissland, 1992) systems, where cases are grouped under a set of domain-specific dimensions. Another early significant contribution to CBR is the memory-based reasoning (MBR) model of Stanfill and Waltz (1988), designed for parallel computation rather than knowledge-based matching.

### 3 The steps of the CBR process

CBR has been formalised for purposes of computer and human reasoning as a four-step process, known as the 'dynamic model of the CBR cycle'. These steps involve:

- 1 Retrieve the most similar to the new problem past case, or cases.
- 2 Reuse the information and knowledge in that case to solve the problem.
- 3 Revise the proposed solution.
- 4 Retain the parts of this experience likely to be useful for future problem-solving.

In more detail, an initial description of a problem defines a new case. This new case is used to *retrieve* the most similar case, or cases, from the library of previous cases. The subtasks of the retrieving procedure involve: Identifying a set of relevant problem descriptors, matching the case and returning a set of sufficiently similar cases, given a similarity threshold of some kind, and selecting the best case from the set of cases returned.

Some systems retrieve cases based largely on superficial syntactic similarities among problem descriptors, while advanced systems use semantic similarities.

The retrieved case (or cases) is combined, through *reuse*, with the new case into a solved case, i.e., a proposed solution of the initial problem. The reusing procedure focuses on identifying the differences between the retrieved and the current case, as well as the part of the retrieved case which can be transferred to the new case. CBR methods are implemented by retrieval methods (to retrieve past cases), a language of preferences (to select the best case) and a form of derivational analogy (to reuse the retrieved method into the current problem).

Through the *revising* procedure this solution is tested for success, e.g., by being applied to the real world environment, or a simulation of it, or evaluated by a teacher, and repaired, if failed. This provides an opportunity to learn from failure.

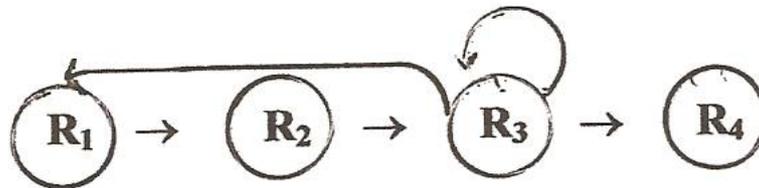
Finally, during the *retaining* procedure useful experience is retained for future reuse, and the case base is updated by a new learned case, or by modification of some existing cases. The above process involves deciding what information to retain and in what form

to retain it, how to index the case for future retrieval, and integrating the new case into the case library.

The ‘general knowledge’ usually plays a part in the CBR cycle by supporting the CBR process. This support however may range from very weak (or none) to very strong, depending on the type of the CBR method. By general knowledge we mean general, domain-dependent knowledge, as opposed to specific knowledge embodied by cases. For example, in the case of the lawyer, mentioned in our introduction, who advocates a particular outcome in a trial based on legal precedents, the general knowledge is expressed through the existing relevant laws and the correlations between them and the case of the trial. A set of rules may have the same role in other CBR cases.

According to the above description one can design the ‘flow-diagram’ of the CBR process shown in Figure 1.

**Figure 1** CBR flow-diagram



Notice that Riesbeck and Bain (1987), Slade (1991), Lei et al. (2001), Aamodt and Plaza (2004), Voskoglou (2008, 2011) and others presented detailed flowcharts illustrating the basic steps of the CBR process. However, the simple flow-diagram of Figure 1 is very useful in helping us to build a stochastic (Markov chain) model that enables a mathematical representation of the CBR process.

#### 4 THE 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 a ‘memory’ of only one state. This means that the probability of entering a certain state at a certain phase, although it is not necessarily independent of previous phases, depends at most on the state occupied in the previous phase (Kemeny et al 1964). This property 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 and Snell (1976).

We shall construct a Markov chain model for the mathematical description of the CBR process. For this, assuming that the CBR process has the Markov property, we introduce a finite Markov chain having as states the four steps of the CBR 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 modelling 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  $R_i$  to  $R_j$ , for  $i, j = 1, 2, 3, 4$ , then the matrix  $A = [p_{ij}]$  is said to be the transition matrix of the chain.

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

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

In fact, starting from  $R_1$  the chain moves at the next phase to  $R_2$ . Therefore,  $p_{11} = p_{13} = p_{14} = 0$  and  $p_{12} = 1$ . Similarly, being at  $R_2$  the chain moves at the next phase to  $R_3$ . Therefore,  $p_{21} = p_{22} = p_{24} = 0$  and  $p_{23} = 1$ . Also, being at  $R_3$  the chain at the next phase either remains there, or it returns to  $R_1$ , or it proceeds to  $R_4$ . Therefore,  $p_{32} = 0$ , while the probabilities  $p_{31}$ ,  $p_{33}$  and  $p_{34}$  are unknown, with  $p_{31} + p_{33} + p_{34} = 1$  (probability of a certain event). Finally, when chain reaches  $R_4$  it is impossible to leave it, because the solution of the new problem by the CBR approach finishes there. In other words,  $R_4$  is an absorbing state of the chain. Therefore,  $p_{41} = p_{42} = p_{43} = 0$  and  $p_{44} = 1$ .

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

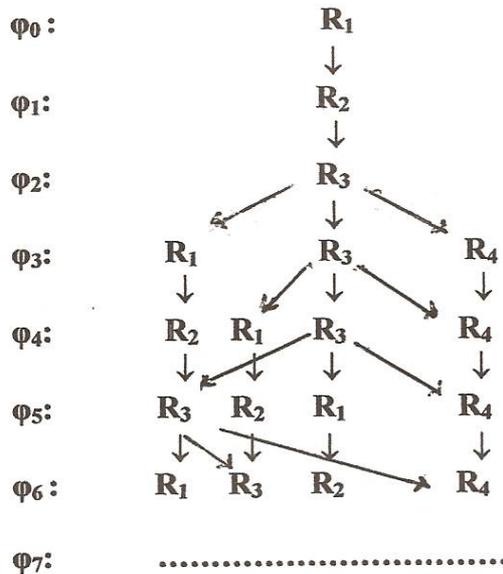
$$P_i = [p_1^{(i)} p_2^{(i)} p_3^{(i)} p_4^{(i)}]$$

the row-matrix giving the probabilities  $p_j^{(i)}$  for the chain to be in each of the states  $R_j, j = 1, 2, 3, 4$ , at phase  $\varphi_i, i = 1, 2, \dots$  of the chain. We obviously have that

$$\sum_{j=1}^4 p_j^{(i)} = 1.$$

The above row-matrix is called the ‘probability vector’ of the chain at phase  $\varphi_i$ .

**Figure 2** Tree of correspondence among states and phases of the Markov model



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.

From this tree becomes evident that

$$P_0 = [1 \ 0 \ 0 \ 0], P_1 = [0 \ 1 \ 0 \ 0], P_2 = [0 \ 0 \ 1 \ 0], \text{ and } P_3 = [p_{31} \ 0 \ p_{33} \ p_{34}].$$

Further it is well known that

$$P_{i+1} = P_i A, \quad i = 0, 1, 2, \dots$$

Therefore, we find that

$$P_4 = P_3 A = [p_{33} p_{31} \ p_{31} \ p_{33}^2 \ p_{34} (p_{33} + 1)]$$

$$P_5 = P_4 A = [p_{33}^2 p_{31} \ p_{33} p_{31} \ p_{31} + p_{33}^3 p_{34} (p_{33}^2 + p_{33} + 1)]$$

and so on.

As we have seen above  $R_4$  is an absorbing state of the chain. Further, from the flow-diagram of Figure 1 it is easy to observe that from any other state of the chain it is possible to reach  $R_4$ , not necessarily in one step. Thus, we have an ‘absorbing Markov chain’ with  $R_4$  being its unique absorbing state. Applying standard techniques from the theory of finite absorbing chains we bring the transition matrix  $A$  to ‘its canonical (or standard) form’  $A^*$  by listing the absorbing state first and then partition it as follows:

$$A^* = \begin{array}{c} R_4 \\ R_1 \\ R_2 \\ R_3 \end{array} \begin{array}{c} R_4 \\ R_1 \\ R_2 \\ R_3 \end{array} \left[ \begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ - & - & - & - \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ p_{34} & p_{31} & 0 & p_{33} \end{array} \right].$$

Symbolically, we can write

$$A^* = \left[ \begin{array}{c|c} I & 0 \\ - & - \\ R & Q \end{array} \right],$$

where  $Q$  stands for the transition matrix of the non-absorbing states.

Let  $I_3$  denote the  $3 \times 3$  unitary matrix. Since we work with an absorbing chain, the matrix  $I_3 - Q$  has an inverse (Kemeny and Snell, 1976; Theorem 3.2.1), say  $N$ , which is called the ‘fundamental matrix’ of the chain. Using the well-known from linear algebra method of determinants in calculating the inverse of a non-singular matrix we can write

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

In the above formula  $adj(I_3 - Q)$  denotes the adjoint matrix of  $I_3 - Q$  (i.e., the matrix of the algebraic complements of the entries of the transpose matrix of  $(I_3 - Q)$  and  $D(I_3 - Q)$  denotes the determinant of  $I_3 - Q$ . 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 that the entry  $n_{ij}$  of  $N$  gives the mean number of times in state  $R_j$  when the chain is started in state  $R_i$  (Kemeny and Snell, 1976; Theorem 3.2.4). Therefore, since the present chain is always starting from  $R_1$ , the sum

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

gives the mean number of phases of the chain before absorption. In other words, 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. The ideal case is when the CBR process is completed straightforwardly, i.e., without ‘backwards’ from  $R_3$  to  $R_1$ , or ‘stays’ to  $R_3$  (see Figure 1). In this case we have that  $p_{31} = p_{33} = 0$  and  $p_{34} = 1$ , therefore  $t = 3$ . Thus, in general we have that  $t \geq 3$ .

The following simple example illustrates our results:

*Example:* Consider the case of a physician, who takes into account the diagnosis and treatment of a previous patient having similar symptoms in order to determine the disease and treatment for the patient in front of him. Obviously the physician is using CBR. If the initial treatment fails to improve the health of the patient, then the physician either revises the treatment (this means stay to  $R_3$  for two successive phases), or, in more difficult cases, gets a reminding of a previous similar failure and uses the failure case to improve its understanding of the present failure and correct it (this means transfer from  $R_3$  to  $R_1$ ). The process is completed, when the physician succeeds to cure his patient.

Assume that the recorded statistical data show that the probabilities of a straightforward cure of the patient and of each of the above two reactions of the physician in case of failure are equal to each other. This means that  $p_{13} = p_{33} = p_{34} = \frac{1}{3}$  and therefore  $t = 7$ , i.e., the mean number of steps for the cure of the patient is 8.

Further, one finds that

$$P_3 = \begin{bmatrix} \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \end{bmatrix}, \quad P_4 = \begin{bmatrix} \frac{1}{9} & \frac{1}{3} & \frac{1}{9} & \frac{4}{9} \end{bmatrix}, \quad P_5 = \begin{bmatrix} \frac{1}{27} & \frac{1}{9} & \frac{4}{9} & \frac{13}{27} \end{bmatrix}$$

and so on. Observing for example the probability vector  $P_5$  one concludes that the probability for the CBR process to be at the step of revision ( $R_3$ ) in the sixth phase after its starting is  $\frac{4}{9}$ , or approximately 44.44%, the corresponding probability to be at the

step of retaining the acquired experience ( $R_4$ ) is  $\frac{13}{27}$ , or approximately 48.15% (in this case it is possible that the CBR process has arrived to the absorbing state  $R_4$  in an earlier phase), etc.

*Note:* Knowing the exact ‘movements’ during the CBR process one can calculate the number of steps needed for the absorption of the chain directly from the flow-diagram of Figure 1. For example, considering the above case of the physician, assume that the initial treatment given to the patient failed to cure him and the physician got a reminding of a similar failure in the past in order to correct it. Assume further that the new treatment did not give the expected results and the physician revised it again succeeding in this way to cure the patient. According to the assumptions mentioned above it is easy, through the flow-diagram, to find that the number of steps needed for the absorption is exactly 8.

## 5 Measuring the effectiveness of a CBR system

The challenge in CBR is to come up with methods that are suited for problem-solving and learning in particular subject domains and for particular application environments. In line with the main steps of the CBR process, core problems addressed by CBR research can be grouped into five areas: Representation of cases and methods for retrieval, reuse, revision and retaining the acquired experience. A CBR system should support the problems appearing in the above five areas. A good CBR system should support a variety of retrieval mechanisms and allow them to be mixed when necessary. In addition, the system should be able to handle large case libraries with the retrieval time increasing linearly (at worst) with the number of cases.

Let us consider a CBR system including a library of  $n$  recorded past cases and let  $t_i$ , as it has been calculated in the previous section, be the mean number of steps for the completion of the CBR process for case  $c_i$ ,  $i = 1, 2, \dots, n$ . Each  $t_i$  could be stored in the system’s library together with the corresponding case  $c_i$ . We define then the system’s effectiveness (in solving new related problems), say  $t$ , to be the mean value of the  $t_i$ ’s of its stored cases, i.e., we have that

$$t = \frac{\sum_{i=1}^n t_i}{n}.$$

The more problems are solved in future applications through the given system, the bigger becomes the number  $n$  of the stored cases in the system’s library and therefore the value of  $t$  is changing. As  $n$  increases it is normally expected that  $t$  will decrease, because the values of the  $t_i$ ’s of the new stored cases would be decreasing. In fact, the bigger is  $n$ , the better would be the chance of a new case to ‘fit’ well (i.e., to have minor differences) with a known past case, and therefore the less would be the difficulty of solving the corresponding problem by the CBR process. Thus, we could say that a CBR system ‘behaves well’ if, when  $n$  tends to infinity, then its effectiveness  $t$  tends to 3.

*Example:* Consider a CBR system that has been designed in terms of Schank’s model of dynamic memory for the representation of cases, as we have briefly described it in the second section of this paper. In order to calculate the effectiveness of a system of this type we need first to calculate the effectiveness of each GE contained in it and then take the mean value of them.

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 third section of this chapter, and that  $c_3$  includes one ‘return’ from  $R_3$  to  $R_1$  and two ‘stays’ to  $R_3$ . Then  $t_1 = 3$  and  $t_2 = 7$ , while for the calculation of  $t_3$  observe that  $p_{31} = p_{34} = \frac{1}{4}$  and  $p_{33} = \frac{1}{2}$ , therefore  $t_3 = 8$ . Thus, the effectiveness of this GE is equal to  $t = \frac{3+7+8}{3} = 6$ .

Notice that a complex GE may contain some more specific GE’s (e.g., see Figure 3 in page 12 of Aamodt and Plaza, 2004). In this case we only need to calculate the efficiency of the complex GE by considering all its cases, regardless if they belong or not to one or more of the specific GE’s contained in it.

*Note:* As we have seen in the second section, an alternative approach for the representation of cases in a CBR system is the category and exemplar model of Porter and Bareiss. The process of calculating the effectiveness of a system of such type is analogous to the process described in the above example, the only difference being that one has to work with categories rather, instead of GE’s. In a similar way one may calculate the effectiveness of systems corresponding to the other case memory models (Rissland’s and MBR) that we have described in the second section.

## 6 A fuzzy model for the representation of a CBR system

They are often situations in everyday life in which definitions have not clear boundaries; e.g., this happens when we speak about the ‘high mountains’ of a country, the ‘good players’ of a football team, etc. The fuzzy sets theory was created (Zadeh, 1965) in response to have a mathematical representation of such kind of situations. For special facts on fuzzy sets and on uncertainty theory we refer freely to Klir and Folger (1988).

Let us consider a CBR system whose library contains  $n$  past cases,  $n \geq 2$ . We denote by  $R_i$ ,  $i = 1, 2, 3$ , the steps of retrieval, reuse and revision and by  $a, b, c, d$ , and  $e$  the linguistic labels of negligible, low, intermediate, high and complete degree of success respectively for each of the  $R_i$ ’s. Set

$$U = \{a, b, c, d, e\}$$

We are going to represent  $R_i$ ’s as fuzzy sets in  $U$ . For this, if  $n_{ia}$ ,  $n_{ib}$ ,  $n_{ic}$ ,  $n_{id}$  and  $n_{ie}$  respectively denote the number of cases where it has been achieved negligible, low, intermediate, high and complete degree of success for the state  $R_i$ ,  $i = 1, 2, 3$ , we define the membership function  $m_{R_i}$  in terms of the frequencies, i.e., by

$$m_{R_i}(x) = \frac{n_{ix}}{n}$$

for each  $x$  in  $U$ . Thus, we can write

$$R_i = \left\{ \left( x, \frac{n_{ix}}{n} \right) : x \in U \right\}, \quad i = 1, 2, 3$$

The reason, for which we did not include the last step  $R_4$  of the CBR process in our fuzzy representation, is that all past cases, either successful, or not, are retained in the system’s

library and therefore there is no fuzziness in this case. In other words keeping the same notation we have that  $n_{4a} = n_{4b} = n_{4c} = n_{4d} = 0$  and  $n_{4e} = 1$ .

In order to represent all possible *profiles* (*overall states*) of a case during the CBR process, we consider a fuzzy relation, say  $R$ , in  $U^3$  of the form

$$R = \{(s, m_R(s)) : s = (x, y, z) \in U^3\}$$

To determine properly the membership function  $m_R$  we give the following definition:

*Definition:* A profile  $s = (x, y, z)$ , with  $x, y, z$  in  $U$ , is said to be well ordered if  $x$  corresponds to a degree of success equal or greater than  $y$ , and  $y$  corresponds to a degree of success equal or greater than  $z$ .

For example, profile  $(c, c, a)$  is well-ordered, while  $(b, a, c)$  is not.

We define now the membership degree of  $s$  to be

$$m_R(s) = m_{R1}(x) m_{R2}(y) m_{R3}(z)$$

if  $s$  is a well-ordered profile, and zero otherwise. In fact, if for example  $(b, a, c)$  possessed a non-zero membership degree, given that the degree of success at the step of reuse is negligible how the proposed solution could be revised?

In order to simplify our notation we shall write  $m_s$  instead of  $m_R(s)$ . Then the *possibility*  $r_s$  of the profile  $s$  is given by

$$r_s = \frac{m_s}{\max\{m_s\}}$$

where  $\max\{m_s\}$  denotes the maximal value of  $m_s$ , for all  $s$  in  $U^3$ . In other words,  $r_s$  is the 'relative membership degree' of  $s$  with respect to the other profiles.

During the CBR process it might be used reasoning that involves amplified inferences, whose content is beyond the available evidence and hence obtain conclusions not entailed in the given premises. The appearance of conflict in the conclusions requires that the conclusions be appropriately adjusted so that the resulting generalisation is free of conflict. The value of total conflict during the CBR process can be measured by the *strife function*  $S(r)$  [Klir, (1995), p.28] on the ordered possibility distribution

$$r : r_1 = 1 \geq r_2 \geq \dots \geq r_n \geq r_{n+1}$$

of the profiles defined by:

$$S(r) = \frac{1}{\log 2} \left[ \sum_{i=2}^n (r_i - r_{i+1}) \log \frac{i}{\sum_{j=1}^i r_j} \right]$$

In general, the amount of information obtained by an action can be measured by the reduction of uncertainty that results from the action. Thus, the *total possibilistic uncertainty*  $T(r)$  during the CBR process can be used as a measure for the system's effectiveness in solving new related problems. The value of  $T(r)$  is measured by the sum of the strife  $S(r)$  and *non-specificity*  $N(r)$  [Klir, (1995), p.28], defined by:

$$N(r) = \frac{1}{\log 2} \left[ \sum_{i=2}^n (r_i - r_{i+1}) \log i \right]$$

In contrast to strife, which, as we have already seen, expresses conflicts among the various sets of alternatives, non specificity is connected with the sizes (cardinalities) of relevant sets of alternatives. The lower is the value of  $T(r)$  (i.e., the greater is the reduction of the initially existing system's uncertainty), the higher is the effectiveness of the CBR system in solving new related problems. In fact, the higher is the reduction of system's uncertainty, the greater is the amount of information obtained by the implemented action (solution of the new given problem).

Assume now that one wants to study the combined results of the behaviour of  $k$  different systems,  $k \geq 2$ , designed for the solution of the same type of problems via the CBR process. Then it becomes necessary to introduce the *fuzzy variables*  $R_1(t)$ ,  $R_2(t)$  and  $R_3(t)$  with  $t = 1, 2, \dots, k$ . The values of the above variables represent the steps of the CBR process for each of the  $k$  CBR systems as fuzzy sets in  $U$ ; e.g.,  $R_1(2)$  represents the step of retrieval for the second system. In order to measure the degree of evidence of the combined results of the  $k$  systems, it becomes necessary to define the possibility  $r(s)$  of each profile  $s$  with respect to the sum of the membership degrees of  $s$  for all systems. For this, we introduce the *pseudo-frequencies*

$$f(s) = \sum_{t=1}^k m_s(t)$$

and we define

$$r(s) = \frac{f(s)}{\max\{f(s)\}}$$

where  $\max\{f(s)\}$  denotes the maximal pseudo-frequency. Obviously, the same method could be applied when one wants to study the behaviour of a system during the CBR process for the solution of  $k$  different related problems.

## 7 An application of the fuzzy model

Let us consider a CBR system with an existing library of 105 past cases, where in no case there was a failure at the step of retrieval of a past case for the solution of the corresponding problem. More explicitly, let us assume that in 51 cases we had an intermediate success in retrieving a suitable past case, in 24 cases high, and in 30 cases we had a complete success respectively. Of course the existence of a certain criterion is necessary in order to be able to characterise the degree of success of retrieval for each of the past cases. Thus, the step of retrieval can be represented as a fuzzy set in  $U$  as

$$R_1 = \left\{ (a, 0), (b, 0), \left( c, \frac{51}{105} \right), \left( d, \frac{24}{105} \right), \left( e, \frac{30}{105} \right) \right\}$$

Assume further that in a similar way we obtained that

$$R_2 = \left\{ \left( a, \frac{18}{105} \right), \left( b, \frac{18}{105} \right), \left( c, \frac{48}{105} \right), \left( d, \frac{21}{105} \right), (e, 0) \right\}$$

and

$$R_3 = \left\{ \left( a, \frac{36}{105} \right), \left( b, \frac{30}{105} \right), \left( c, \frac{39}{105} \right), (d, 0), (e, 0) \right\}$$

It is a straightforward process now to calculate the membership degrees of all the possible profiles [see column of  $m_s(1)$  in Table 1]. For example, if

$$s = (c, b, a)$$

then

$$m_s = m_{R1}(c) \cdot m_{R2}(b) \cdot m_{R3}(a) = \frac{51}{105} \frac{18}{105} \frac{36}{105} \approx 0.029.$$

It turns out that  $(c, c, c)$  is the profile with the maximal membership degree 0.082 and therefore the possibility of each  $s$  in  $U^3$  is given by

$$r_s = \frac{m_s}{0.082}$$

For example, the possibility of  $(c, b, a)$  is

$$\frac{0.029}{0.082} \approx 0.353$$

while the possibility of  $(c, c, c)$  is of course equal to 1.

The total number of the system's profiles is equal to  $5^3 = 125$  (ordered samples with replacement of 3 objects taken from 5). Calculating the possibilities of these profiles [see column of  $r_s(1)$  in Table 1] one finds that their ordered possibility distribution  $r$  is:

$$\begin{aligned} r_1 = 1, r_2 = 0,92, r_3 = 0,768, r_4 = 0,512, r_5 = 0,476, r_6 = 0,415, r_7 = 0,402, \\ r_8 = 0.378, r_9 = r_{10} = 0.341, r_{11} = 0.329, r_{12} = 0.317, r_{13} = 0.305, r_{14} = 0.293, \\ r_{15} = r_{16} = 0.256, r_{17} = 0.20, r_{18} = 0.195, r_{19} = 0.171, r_{20} = r_{21} = r_{22} = 0.159, \\ r_{23} = 0.134, r_{24} = r_{25} = \dots = r_{125} = 0 \end{aligned}$$

Therefore, the total possibilistic uncertainty of the system is

$$T(r) = S(r) + N(r) = 0.565 + 2.405 = 2.97$$

Next we shall study the combined results of the behaviour of the above system and of another system, designed for the solution of the same type of problems via the CBR process, with an existing library of 90 past cases. Assume that working as before we found for the second system that

$$R_1 = \left\{ (a, 0), \left( b, \frac{18}{90} \right), \left( c, \frac{45}{90} \right), \left( d, \frac{27}{90} \right), (e, 0) \right\}$$

$$R_2 = \left\{ \left( a, \frac{18}{90} \right), \left( b, \frac{24}{90} \right), \left( c, \frac{48}{90} \right), (d, 0), (e, 0) \right\}$$

and

$$R_3 = \left\{ \left( a, \frac{36}{90} \right), \left( b, \frac{27}{90} \right), \left( c, \frac{27}{90} \right), (d, 0), (e, 0) \right\}$$

The calculation of all possible profiles gives the results shown in column of  $m_s(2)$  in Table 1. It turns out that  $(c, c, a)$  is the profile possessing the maximal membership degree 0.107 and therefore the possibility of each  $s$  is given by

$$r_s = \frac{m_s}{0.107}$$

[see column of  $r_s(2)$  in Table 1].

**Table 1** Profiles with non zero pseudo-frequencies (the outcomes are of accuracy up to the third decimal point)

$A_1$	$A_2$	$A_3$	$m_s(1)$	$r_s(1)$	$m_s(2)$	$r_s(2)$	$f(s)$	$r(s)$
b	b	b	0	0	0.016	0.150	0.016	0.087
b	b	a	0	0	0.021	0.196	0.021	0.115
b	a	a	0	0	0.016	0.150	0.016	0.087
c	c	c	0.082	1	0.080	0.748	0.162	0.885
c	c	a	0.076	0.927	0.107	1	0.183	1
c	c	b	0.063	0.768	0.008	0.075	0.071	0.388
c	a	a	0.028	0.341	0.040	0.374	0.068	0.372
c	b	a	0.028	0.341	0.053	0.495	0.081	0.443
c	b	b	0.024	0.293	0.040	0.374	0.064	0.350
d	d	a	0.016	0.495	0	0	0.016	0.087
d	d	b	0.013	0.159	0	0	0.013	0.074
d	d	c	0.021	0.256	0	0	0.021	0.115
d	a	a	0.013	0.159	0.024	0.224	0.037	0.202
d	b	a	0.013	0.159	0.032	0.299	0.045	0.246
d	b	b	0.011	0.134	0.024	0.224	0.035	0.191
d	c	a	0.031	0.378	0.064	0.598	0.095	0.519
d	c	b	0.026	0.317	0.048	0.449	0.074	0.404
d	c	c	0.034	0.415	0.048	0.449	0.082	0.448
e	a	a	0.017	0.207	0	0	0.017	0.093
e	b	b	0.014	0.171	0	0	0.014	0.077
e	c	a	0.039	0.476	0	0	0.039	0.213
e	c	b	0.033	0.402	0	0	0.033	0.180
e	c	c	0.042	0.512	0	0	0.042	0.230
e	d	a	0.025	0.305	0	0	0.025	0.137
e	d	b	0.021	0.256	0	0	0.021	0.115
e	d	c	0.027	0.329	0	0	0.027	0.148

Finally, in the same way as above, one finds that

$$T(r) = S(r) + N(r) = 0.452 + 1.87 = 2.322$$

Thus, since  $2.322 < 2.97$ , the effectiveness of the second system in solving new related problems is better than that of the first one. This happens despite to the fact that the profile  $(c, c, c)$  with the maximal possibility of appearance in the first system is a more satisfactory profile than the corresponding profile  $(c, c, a)$  of the second system.

Notice that in general, the more are the stored past cases in the system's library, the greater is expected to be its effectiveness in solving new related problems. In fact, the more are the past cases, the greater is the probability for a new problem to fit satisfactorily to one of them. Therefore, the fact that the second system was found to be more effective than the first one, although not impossible to happen, it is rather unexpected in general.

We introduce now the fuzzy variables  $R_1(t)$ ,  $R_2(t)$ , and  $R_3(t)$ , with  $t = 1, 2$ . Then the pseudo-frequency of each profile  $s$  is given by

$$f(s) = m_s(1) + m_s(2)$$

(see the corresponding column of Table 1). It turns out that  $(c, c, a)$  is the profile with the highest pseudo-frequency 0.183 and therefore the possibility of each profile is given by

$$r(s) = \frac{f(s)}{0.183}$$

The possibilities of all profiles having non-zero pseudo-frequencies are given in the last column of Table 1.

## 8 Discussion and conclusions

CBR is one of the central ideas in the nowadays AI. Its applications are especially efficient in helping information managers to increase efficiency and reduce cost by substantially automating processes. Detailed flowcharts illustrating in the traditional (descriptive) way the basic steps of the CBR process have been designed during the last 25 years by several researchers (see the references given in Section 3).

In the present article, based on a simple flow-diagram of CBR (Figure 1), we developed an innovative stochastic model for a more effective description and mathematical formulation of the CBR process. The model was built by introducing a finite, absorbing Markov chain on the major steps of the CBR process. By applying standard results of the relevant theory we succeeded in calculating the probabilities for the CBR process to be at any of its steps at a certain phase of the solution procedure of a given problem. We also obtained a measure for the effectiveness of the corresponding CBR system in solving this problem. Mathematics does not explain the natural behaviour of an object, it simply describes it. This description however is so much effective, so that an elementary mathematical equation can describe simply and clearly a relation, that in order to be expressed with words could need entire pages. We believe that this is exactly the main advantage of our Markov model compared with the above referred traditional methodologies. Similar Markov models have been also derived by the author in earlier papers for the description of several situations appearing in Mathematics Education and

Management. For example, see Voskoglou and Perdikaris (1991), Voskoglou (1996a, 1996b, 2000, 2007, 2009), etc.

Fuzzy logic is widely used in machine learning and control (e.g., see Mamdani, 1974; Hajek, 1998; Gerla, 2005; etc). In fact, although genetic algorithms and neural networks can perform just as well as fuzzy logic in many cases, the former has the advantage that the solution to the problem can be cast in terms that human operators can understand. This makes it easier to mechanise tasks that are already successfully performed by humans. Our innovative fuzzy model for the description of a CBR system presented in this paper was developed by representing the steps of retrieval, reuse and revision of the CBR process as fuzzy subsets of the set  $U$  of the linguistic labels of negligible, low, intermediate, high and complete success of the CBR system in each of the above steps. In this fuzzy environment we used the total possibilistic uncertainty  $T(r)$  on the ordered possibility distribution  $r$  as a measure of the CBR system's performance in solving new related problems. In fact, the lower is the final value of  $T(r)$  after the solution procedure, the greater the reduction of the initially existing  $T(r)$  and therefore the greater the new information obtained. Analogous fuzzy models have been constructed by the author for representing the processes of learning (Voskoglou, 1999, 2009), of mathematical modelling (Voskoglou, 2010), of problem-solving (Voskoglou, 2012a) and of analogical reasoning (Voskoglou, 2012b).

There are also other fuzzy measures in use for the assessment of a system's performance. One such is the well-known *Shannon's entropy* (Shannon 1948). For use in a fuzzy environment, this measure is expressed in terms of the Dempster-Shafer mathematical theory of evidence in the form

$$H = -\frac{1}{\ln n} \sum_{s=1}^n m_s \ln m_s \pi$$

where  $n$  is the total number of the system's profiles [Klir, (1995), p.20]. The value of  $H$  measures the system's total probabilistic uncertainty and the associated information. Therefore, similarly with  $T(r)$  above, the lower is the final value of  $H$ , the better the system's performance. It should be mentioned here that in evaluating fuzzy data the probability of a system's profile  $s$  is calculated in terms of the membership degrees of all system's profiles by

$$p_s = \frac{m_s}{\sum_{s \in U^3} m_s}.$$

The above fuzzy measure has been used in assessing students' analogical problem-solving skills (Voskoglou, 2012b), to determine the continuity of the van Hiele levels of geometric thinking (Perdikaris, 2011), etc. An advantage of adopting  $H$  as a measure instead of  $T(r)$  is that  $H$  is calculated directly from the membership degrees of all profiles  $s$  without being necessary to calculate their probabilities  $p_s$ . In contrast, the calculation of  $T(r)$  presupposes the calculation of the possibilities  $r_s$  of all profiles first. However, according to Shackle (1961) human reasoning can be formalised more adequately by possibility rather, than by probability theory.

Another popular approach is the 'centroid' method, in which the *centre of mass* of the graph of the membership function involved provides an alternative measure of the

system's performance ( ). As we have seen above, in an earlier paper (Voskoglou, 1999) we have formalised the process of learning a subject matter by the individuals (and especially the process of learning mathematics by students) using a fuzzy logic approach similar to that described in this paper for CBR. Later (Voskoglou, 2009) we expanded this argument by using the  $T(r)$  of a student group as a measure of its learning skills. Meanwhile, Subbotin et al. (2004), based on our fuzzy model for the learning process (Voskoglou, 1999), they used the 'centroid' method for measuring students learning skills. Recently, together with Prof. Subbotin, we have applied this approach as an assessment tool of students' analogical problem-solving abilities (Voskoglou and Subbotin, 2012).

We emphasise that the above, three in total, fuzzy measures 'are treating differently the idea of a system's performance'. In fact, in the first two cases [ $T(r)$  and  $H$ ] the system's uncertainty is connected to its capacity in obtaining relevant information. In contrast, the weighted average plays the main role in the third case, i.e., the result of the performance close to the ideal performance has much more weight than the one close to the lower end. In other words, while the first two cases are looking to the *average performance*, the third one is mostly looking at the *quality of the performance*. Consequently, some differences could appear in boundary cases in evaluating a system's performance by these two different approaches. Therefore, it is argued that a combined use of them could help the user in founding the ideal profile of performance according to his/her personal criteria of goals. This could help him/her in choosing the appropriate system among the existing ones for solving his/her problems.

Analogous attempts to use the fuzzy logic in the area of student modelling and student diagnosis in particular and in education in general have been attempted by other researchers as well, e.g., Perdikaris (1996), Espin and Oliveras (1997), Ma and Zhou (2000), Ajello and Spagnolo (2002), etc. Spagnolo and Gras (2004) have also used analogous methods in studying fuzzy through statistic implication. This is a new approach in Zadeh's (2001) classification of the various approaches of fuzzy logic (possibilistic, probabilistic, veristic) that can possibly have some interesting applications in AI.

In conclusion, our fuzzy model for the CBR is not restricted in providing quantitative information only [value of  $T(r)$ , etc.], but it also gives a qualitative view of a CBR system's behaviour. In fact, through it one studies the possibilities of all system's profiles and gets – in terms of the linguistic labels – a comprehensive idea about the degree of success in each step of the CBR process. Another advantage of this model is that it gives to the user the possibility to study the combined results of the behaviour of two, or more, CBR systems designed for the solution of the same type of problems, or alternatively the behaviour of the CBR system in use during the solution of two, or more, new related problems. On the other hand, the characterisation of the system's performance in terms of a set of linguistic labels which are fuzzy themselves is a disadvantage of the fuzzy model, because this characterisation depends, as we have seen above, on the user's personal criteria.

Concerning our Markov model, although it looks rather easier to be applied in practice by a non-expert, it is self-restricted to provide quantitative information only for the corresponding situations, i.e., a measure of a CBR system's efficiency in solving related problems and short-run forecasts (probabilities) for the evolution of the CBR process. Because of this, this model is helpful only in understanding the 'ideal behaviour'

of a CBR system proceeding linearly from real-world problems to acceptable solutions and reports on them. However, the common experience shows that things are not always like that in practice, especially when dealing with complicated problems depending upon a large number of variables and/or parameters. In these cases a system's qualitative study, provided only by the fuzzy model, becomes necessary in comparing and validating the results obtained. Therefore, a combined use of the fuzzy and stochastic models seems to be the best solution in achieving a worthy of credit mathematical analysis of the CBR process.

Finally, we ought to notice that, although we have presented some examples in both cases (Markov and fuzzy models) to illustrate our results, further study and research are needed in order to expand these models to work over some more complex real world problems (e.g., diagnostic and control systems, help desk applications, decision-making, systems supporting architectural and industrial design, learning and knowledge based systems, etc.) and prove their efficiency over them.

## References

- Aamodt, A. and Plaza, E. (2004) 'Case-based reasoning: foundational issues, methodological variations, and system approaches', *Artificial Intelligence Communications*, Vol. 7, No. 1, pp.39–52.
- Ajello, M. and Spagnolo, F. (2002) 'Some experimental observations on common sense and fuzzy logic', *Proceedings International Conference on Mathematics Education into the 21st Century*, pp.35–39, Palermo, Italy.
- Ashley, K. (1991) *Modeling Legal Arguments: Reasoning with Cases and Hypotheticals*, MIT Press, Bradford Books, Cambridge.
- Bareiss, R. (1989) *Exemplar-based Knowledge Acquisition: A Unified Approach to Concept Representation, Classification, and Learning*, Academic Press, Boston.
- Bower, G., Black, J. and Turner, T. (1979) 'Scripts in memory for text', *Cognitive Psychology*, Vol. 11, pp.177–220.
- Collins, A. and Quilliam, M. (1969) 'Retrieval time from semantic memory', *Journal of Verbal Learning and Verbal Behavior*, Vol. 8, pp.240–247.
- Espin, E.A. and Oliveras, C.M.L. (1997) 'Introduction to the use of fuzzy logic in the assessment of mathematics teachers', in Gagatsis, A. (Ed.): *Proceedings of the 1st Mediterranean Conference on Mathematics Education (MEDCONF 97)*, pp.107–113, Nicosia, Cyprus.
- Gentner, D. (1983) 'Structure mapping – a theoretical framework for analogy', *Cognitive Science*, Vol. 7, pp.155–170.
- Gerla, G. (2005) 'Fuzzy logic programming and fuzzy control', *Studia Logica*, Vol. 79, pp.231–235.
- Hajek, P. (1998) *Meta mathematics of Fuzzy Logic*, Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Kemeny, J. and Snell, J.L. (1976) *Finite Markov Chains*, Springer-Verlag, New York.
- Kemeny, J.G., Schleifer, A., Snell, J.L. and Thompson, G.L. (1964) *Finite Mathematics with Business Applications*, Prentice-Hall, Englewood Cliffs, NJ.
- Kintsch, W. (1972) 'Notes on the structure of semantic memory', in Tulvin, E. and Donaldson, W. (Eds.): *Organization of Memory*, pp.247–308, Academic, New York.
- Klir, G. J. (1995) 'Principles of uncertainty: what are they? Why do we need them?', *Fuzzy Sets and Systems*, Vol. 74, pp.15–31.
- Klir, G.J. and Folger, T.A. (1988) *Fuzzy Sets, Uncertainty and Information*, Prentice Hall, London.
- Kolodner, J. (1983) 'Reconstructive memory: a computer model', *Cognitive Science*, Vol. 7, pp.281–328.

- Lebowitz, M. (1983) 'Memory-based parsing', *Artificial Intelligence*, Vol. 21, pp.363–404.
- Lei, Y., Peng, Y. and Ruan, X. (2001) 'Applying case-based reasoning to cold forcing process planning', *Journal of Materials Processing Technology*, Vol. 112, pp.12–16.
- Ma, J. and Zhou, D. (2000) 'Fuzzy set approach to the assessment of student-centered learning', *IEEE Transactions on Education*, Vol. 43, No. 2, pp.237–241.
- Mamdani, E.H. (1974) 'Application of fuzzy algorithms for the control of a simple dynamic plant', *Proceedings IEEE*, pp.121–159.
- Minsky, M. (1975) 'A framework for representing knowledge', in Wilson, P. (Ed): *The Psychology of Computer Vision*, pp.211–227, McGraw-Hill, New York.
- Perdikaris, S. (1996) 'A system framework for fuzzy sets in the van Hiele level theory of geometric reasoning', *International Journal of Mathematical Education in Science and Technology*, Vol. 27, No. 2, pp.273–278.
- Perdikaris, S. (2011) 'Using fuzzy sets to determine the continuity of the van Hiele levels', *Journal of Mathematical Sciences and Mathematics Education*, Vol. 6, No. 1, pp.39–46.
- Porter, B. and Bareiss, B. (1986) 'PROTOS: an experiment in knowledge acquisition for heuristic classification tasks', in *Proceedings of the First International Meeting on Advances in Learning (IMAL)*, pp.159–174, Les Arcs, France.
- Quilliam, M. (1968) 'Semantic memory', in Minsky, M. (Ed.): *Semantic Information Processing*, pp.227–353, MIT Press, Cambridge, Mass.
- Riesbeck, C. and Bain, W. (1987) 'A methodology for implementing case-based reasoning systems', Technical Report, Lockheed.
- Rissland, E. (1983) 'Examples in legal reasoning: legal hypotheticals', in *Proceedings 8th International Joint Conference on Artificial Intelligence (IJCAI)*, Karlsruhe.
- Rumelhart, D., Lindsay, P. and Norman, D. (1972) 'A process model for long-term memory', in Tulving, E. and Donaldson (Eds.): *Organization of Memory*, pp.197–246, Academic, New York.
- Schank, R. (1975) 'The structure of episodes in memory', in Bobrow, D.G. and Collins, A. (Eds.): *Representation and Understanding*, pp.237–272, Academic, New York.
- Schank, R. (1979) 'Reminding and memory organization: an introduction to MOPS', Technical Report 170, Dept. of Computer Science, Yale University.
- Schank, R. (1980) 'Language and memory', *Cognitive Science*, Vol. 4, No. 3, pp.243–284.
- Schank, R. (1981) 'Failure driven memory', *Cognition and Brain Theory*, Vol. 4, No. 1, pp.41–60.
- Schank, R. (1982) *Dynamic Memory: A Theory of Reminding and Learning in Computers and People*, Cambridge Univ. Press.
- Schank, R. and Abelson, R. (1977) *Scripts, Plans, Goals and Understanding*, Lawrence Erlbaum, Hillsdale, NJ.
- Shackle, G.L.S. (1961) *Decision, Order and Time in Human Affairs*, Cambridge University Press, Cambridge.
- Shannon, C.E. (1948) 'A mathematical theory of communications', *Bell Systems Technical Journal*, Vol. 27, pp.379–423, 623–656.
- Skalak, C.B. and Rissland, E. (1992) 'Arguments and cases: an inevitable twining', *Artificial Intelligence and Law, An International Journal*, Vol. 1, No. 1, pp.3–48.
- Slade, S. (1991) 'Case-based reasoning: a research paradigm', *Artificial Intelligence Magazine*, Vol. 12, No. 1, pp.42–55.
- Smith, E. and Medin, D. (1981) *Categories and Concepts*, Harvard University Press.
- Spagnolo, F. and Gras, R. (2004) 'Fuzzy implication through statistic implication: a new approach in Zadeh's framework', in Scott, D., Kurgan, L., Musilek, P., Pedrycz, W. and Reformat, M. (Eds.): *23rd International Conference of the North American Fuzzy Information Processing Society (NAFIPS 2004)*, IEEE, Vol. 1, pp.425–429.

- Stanfill, C. and Waltz, D. (1988) 'The memory-based reasoning paradigm', in *Case-based Reasoning, Proceedings from a Workshop*, pp.414–424, Morgan Kaufmann, Clearwater Beach, Florida.
- Subbotin, I., Badkoo-behi, H. and Bilotskii, N. (2004) 'Application of fuzzy logic to learning assessment', *Didactics of Mathematics: Problems and Investigations*, Vol. 22, pp.38–41.
- Tulving, E. (1972) 'Episodic and semantic memory', in Tulving, E. and Donaldson, W. (Eds.): *Organization of Memory*, pp.381–403, Academic Press.
- Tulving, E. (1983) *Elements of Episodic Memory*, Oxford University Press, Oxford.
- Voskoglou, M.Gr. (1996a) 'The use of Markov chains to describe the process of learning', *Theta: A Journal of Mathematics*, Vol. 10, No. 1, pp.36–40, Manchester Metropolitan University, UK.
- Voskoglou, M.Gr. (1996b) 'An application of ergodic Markov chains to analogical problem solving', *The Mathematics Education*, Vol. 30, No. 2, pp.95–108, India.
- Voskoglou, M.Gr. (1999) 'An application of fuzzy sets to the process of learning', *Heuristics and Didactics of Exact Sciences*, Vol. 10, pp.9–13.
- Voskoglou, M.Gr. (2000) 'An application of Markov chains to decision making', *Studia KUPIECKIE*, Vol. 6, pp.69–76, University of Lodz, Poland.
- Voskoglou, M.Gr. (2007) 'A stochastic model for the modelling process', in Haines, C., Galbraith, P., Bloom, W. and Khan, S. (Eds.): *Mathematical Modelling: Education, Engineering and Economics (ICTMA12)*, pp.149–157, Horwood, Chichester, UK.
- Voskoglou, M.Gr. (2008) 'Case-based reasoning: a recent theory for problem-solving and learning', *Communications in Computer and Information Science (WSKS 08)*, Vol. 19, pp.314–319, Springer.
- Voskoglou, M.Gr. (2009) 'Fuzziness or probability in the process of learning? A general question illustrated by examples from teaching mathematics', *The Journal of Fuzzy Mathematics*, International Fuzzy Mathematics Institute, Los Angeles, Vol. 17, No. 3, pp.79–686.
- Voskoglou, M.Gr. (2010) 'Use of total possibilistic uncertainty as a measure of students modelling capacities', *International Journal of Mathematical Education in Science and Technology*, Vol. 41, No. 8, pp.1051–1060.
- Voskoglou, M.Gr. (2011) 'Case-based reasoning: history, methodology and development trends', *International Journal of Psychology Research*, Nova Publishers, NY, Vol. 6, No. 5, pp.505–530.
- Voskoglou, M.Gr. (2012a) 'Fuzzy logic and uncertainty in problem solving', *Journal of Mathematical Sciences and Mathematics Education*, Vol. 7, No. 1, pp.37–49.
- Voskoglou, M.Gr. (2012b) 'A fuzzy model for analogical problem solving', *International Journal of Fuzzy Logic Systems*, Vol. 2, No. 1, pp.1–10.
- Voskoglou, M.Gr. and Perdikaris, S. (1991) 'A Markov chain model in problem-solving', *International Journal of Mathematical Education in Science and Technology*, Vol. 24, No. 3, pp.443–447.
- Voskoglou, M.Gr. and Subbotin, I.Ya. (2012) 'Fuzzy models for analogical reasoning', *International Journal of Applications of Fuzzy Sets*, Vol. 2, pp.19–38.
- Wittgenstein, L. (1953) *Philosophical Investigations*, pp.31–34, Blackwell.
- Woods, W. (1975) 'What's a link: foundations for semantic networks', in Bobrow, D. and Collins, A. (Eds.): *Representation and Understanding*, pp.35–82, Academic, New York.
- Zadeh, L.A. (1965) *Fuzzy Sets, Information and Control*, Vol. 8, pp.338–353.
- Zadeh, L.A. (2001) 'From computing with numbers to computing with words, from manipulation of measurements to manipulation of perception', in Cantoni, V., Di Gesu, V., Setti, A. and Tegolo, D. (Eds.): *Human and Machine Perception (Thinking, Deciding and Acting)*, Kluwer Academic, New York.