

Probability and Fuzzy Logic in Analogical Reasoning

M. Gr. Voskoglou

School of Technological Applications
Graduate Technological Educational Institute (T.E.I.), 263 34 Patras, Greece

Received 1 October 2013; accepted 12 November 2013

ABSTRACT

Analogical Reasoning (AR) is a method of processing information that compares the similarities between new and past understood concepts, then using these similarities to gain understanding of the new concept. In this work we develop two mathematical models for the description of the process of AR: A stochastic model by introducing a finite ergodic Markov chain on the steps of the AR process and a fuzzy model by representing the main steps of the AR process as fuzzy subsets of a set of linguistic labels characterizing the individuals' performance in each of these steps. The two models are compared to each other by listing their advantages and disadvantages. Classroom experiments are also performed to illustrate their use in practice.

Keywords: Analogical Reasoning, Problem Solving, Markov chains, Fuzzy Sets, Defuzzification Techniques

1. Introduction

One of the most frequently used strategies (heuristics) in problem-solving (PS) is the strategy of the *analogous problem*: When the solver is not sure of the appropriate procedure to solve a given problem (*target problem*), a good hint would be to look for a similar problem solved in the past (*source problem*) and then try to adapt the solution procedure of this problem for use with the target problem. However this strategy can be difficult to implement in PS, because it requires the solver to attend to information other than the problem to be solved. Thus the solver may come up empty-handed, either because he/she has not solved any similar problems in past, or because he/she fails to realize the relevance of previous problems. But, even if an analogue is retrieved, the solver must know how to use it to determine the solution procedure for the target problem.

Several studies ([2], [3], [4], [10], [18] , etc) have provided detailed models for the process of analogical PS (APS), in which factors associated with instances of successful transfer – that is, use of already existing knowledge to produce new knowledge - are identified. According to these studies the main steps involved in APS include:

- *Representation* of the target problem.
- *Search-retrieval* of a source problem
- *Mapping* of the representations of the target and the source problem.

- *Adaptation* of the solution of the source problem for use with the target problem. More specifically, before solvers working on a problem they usually construct a representation of it. A good representation must include both the surface and the structural (abstract, solution relevant) features of the problem. The former are mainly determined by what are the quantities involved in the problem and the latter by how these quantities are related to each other. The features included in solvers' representations of the target problem are used as retrieval cues for a source problem in memory. When the two problems share structural but not surface structures the source is called a *remote analogue* of the target problem. Analogical mapping requires aligning the two situations—that is, finding the correspondences between the representations of the target and the source problem – and projecting inferences from the source to the target. Once the common alignment and the candidate inferences have been discovered the analogy is evaluated. The last step involves the adaptation of the solution of the analogous problem for use with the target problem, where the correspondences between objects and relations of the two problems must be used.

The successful completion of the above process is referred as *positive analogical transfer* (AR). But the search may also yield *distractor problems* having surface but not structural (solution relevant) common features with the target problem and therefore being only superficially similar to it. Usually the reason for this is a non satisfactory representation of the target problem, containing only its salient surface features, and the resulting consequences on the retrieval cues available for the search process. When a distractor problem is considered as an analogue of the target, we speak about *negative analogical transfer*. This happens if a distractor problem is retrieved as a source problem and the solver fails, through the mapping of the representations of the source and target problem, to realize that the source cannot be considered as an analogue to the target. Therefore the process of mapping is very important in APS playing the role of a "*control system*" for the fitness of the source problem.

In this work we develop two mathematical models for the description of the process of AR: A stochastic model by introducing a finite ergodic Markov chain on the steps of the AR process and a fuzzy model by representing the main steps of the AR process as fuzzy subsets of a set of linguistic labels characterizing the individuals' performance in each of these steps. The two models are compared to each other by listing their advantages and disadvantages. Classroom experiments are also performed to illustrate their use in practice.

2. The stochastic model

For the development of our stochastic model we assume that the APS process has the *Markov property*. This means that the probability of entering a certain step at a certain phase of the process, although it is not necessarily independent of previous phases, it depends at most on the step occupied in the previous phase. Our assumption is a simplification (not far away from the truth) made to the real system that enables the formulation of it to a form ready for mathematical treatment (*assumed real system*, e.g. see [19]; section 1).

We introduce a *finite Markov chain* on the steps of the APS process described above. The states of this chain are: s_1 = representation, s_2 =search-retrieval, s_3 = mapping,

Probability and Fuzzy Logic in Analogical Reasoning

s_4 =adaptation and s_5 =solution of the target problem. For general facts on Markov chains we refer freely to the book [6].

The starting state is always s_1 . When the APS process is completed at s_5 , it is assumed that a new problem is given for solution and therefore the process restarts from s_1 . After the completion of the target problem's representation the solvers proceed from s_1 to s_2 . Being at s_2 and facing difficulties in finding a source problem they may return to s_1 asking for more information from problem's representation. Then they proceed again to s_2 to continue the APS process.

After the retrieval of a source problem the solvers proceed from s_2 to s_3 . If the source is considered to be analogous to the target problem, then they transfer from s_3 to s_4 . Otherwise they return to s_2 searching for a new source problem. Notice that solvers who finally fail to retrieve an analogue through the mapping process cannot proceed further. Therefore they return to s_1 waiting for a new problem to be given for solution.

After the adaptation of the solution of the source for use with the target problem the solvers proceed to the final state s_5 of the solution of the target problem. On the contrary, if during the adaptation process they realize that the source is in fact a distractor problem, they return to s_2 searching for a new source. Solvers who finally fail to adapt the solution of the source for use with the target problem they return from s_4 to s_1 waiting for a new problem to be given for solution. According to the above description the *flow-diagram* of the APS process is that shown in Figure 1.

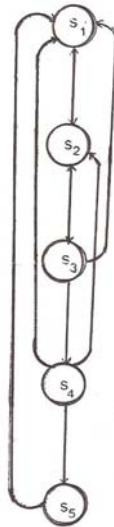


Figure 1. The “flow-diagram” of the process of APS

Denote by p_{ij} the *transition probability* from state s_i to s_j , for $i,j=1,2,3,4,5$. According to the above diagram the *transition matrix* of the chain is:

M. Gr. Voskoglou

$$A = \begin{matrix} & s_1 & s_2 & s_3 & s_4 & s_5 \\ \begin{matrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{matrix} & \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ p_{21} & 0 & p_{23} & 0 & 0 \\ p_{31} & p_{32} & 0 & p_{34} & 0 \\ p_{41} & p_{42} & 0 & 0 & p_{45} \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \end{matrix}$$

Obviously we have that $p_{21}+p_{23}=p_{31}+p_{32}+p_{34}=p_{41}+p_{42}+p_{45}=1$ (1)

It becomes also evident that in our Markov chain it is possible to go between any to states, not necessarily in one step, i.e. it is an *ergodic chain*. For an ergodic chain it is well known that, as the number of its phases tends to infinity (*long run*), the chain tends to an *equilibrium situation* characterized by the equality $P=PA$ (2), where $P=[p_1 \ p_2 \ p_3 \ p_4 \ p_5]$ is the *limiting probability vector* of the chain. The entries of P give the probabilities for the chain to be in each of its states in the long run. Obviously we have that $p_1+ p_2+ p_3+ p_4+ p_5=1$ (3).

From relation (2) one gets easily the following equations:

$$p_1=p_2p_{21}+p_3p_{31}+p_4p_{41}+p_5$$

$$p_2=p_1+p_3p_{32}+p_4p_{42}$$

$$p_3=p_2p_{23}$$

$$p_4=p_3p_{34}$$

$$p_5=p_4p_{45}$$

Adding the first four of the above equations and using relation (1) one finds the fifth equation, which therefore is equivalent with the others. Solving the linear 5X5 system of the first four equations and of equation (3) by the Cramer's rule (it turns out that this system has always a unique solution) one finds that

$$p_1 = \frac{(p_{31}-1)(p_{41}-1) + p_{23}p_{32}(p_{41}-1) - p_{23}p_{42}p_{34}}{D} \quad p_2 = \frac{(p_{31}-1)(p_{41}-1)}{D}$$

$$p_3 = \frac{-p_{23}(p_{41}-1)}{D}, \quad p_4 = \frac{p_{23}p_{34}}{D}, \quad p_5 = \frac{-p_{23}p_{34}(p_{41}-1) - p_{23}p_{34}p_{42}}{D}$$

where $D=(2+p_{23})(p_{31}-1)(p_{41}-1)+p_{23}(2p_{32}-1)(p_{41}-1)+p_{23}p_{34}(1-2p_{42})$ is the determinant of the system.

Further, it is well known that in an ergodic chain the mean number of times in state s_i between two successive occurrences of s_j is given by $\frac{p_i}{p_j}$ ([6]; Theorem 6.2.3).

Therefore, since the process starts again after state s_5 (as a new problem is given for solution), the mean number of times between two successive occurrences of s_5 is given

$$\text{by } m = \sum_{i=1}^4 \frac{p_i}{p_5} = \frac{1 - p_5}{p_5}. \text{ The value of } m \text{ is an indicator of the solvers' difficulties}$$

during the APS process. Another indicator is the time spent for the solution of each problem. However, assuming that the time available for the solution of each problem is prefixed, it becomes evident that m becomes a measure for solvers' difficulties during the APS process. The bigger is m the more the solvers' difficulties during the APS process.

Probability and Fuzzy Logic in Analogical Reasoning

A classroom experiment: In order to illustrate the use of the above model in practice we performed the following experiment, where the subjects were students of the Graduate Technological Educational Institute of Patras, Greece, being at their second term of studies. We formed two groups, with 20 students of the School of Management and Economics in the first and 20 students of the School of Technological Applications (prospective engineers) in the second group.

Three mathematical problems were given for solution to both groups (see Appendix) on topics of the students' first term course of mathematics. In each case and before receiving the target problem students received two other problems together with their solution procedures. They read each problem and its solution procedure and then solved the problem themselves using the given procedure. Subjects were allowed 10 minutes for each problem and they were not given the other problem until after 10 minutes had elapsed. The first of these problems was a remote analogue to the target problem, while the other was a distractor problem. Next the target problem was given and was asked from the students to try to solve it by adapting the solution of one of the previous problems (time allowed 20 minutes). Our instructions stressed the importance of showing all of one's work on paper and emphasized that we were interested in both correct and incorrect solution attempts.

Examining students' papers after the end of the experiment we calculated the following means:

4,2 students from the first group faced difficulties in retrieving a source problem, but they came through after looking back to their representations of the target problem (5,1 students from the second group).

15,1 students from the first group considered through the mapping the collected source as an analogue to the target problem, while the rest of them (4,9 students) searched for a new source. Finally 3,7 from the 4,9 students considered the new source as an analogue to the target problem, while the rest of them (1,2 students) failed to retrieve an analogue through the mapping process (14,8 and 1,6 students from the second group respectively).

Thus $15,1+3,7=18,8$ students from the first group proceeded finally to the step of adaptation ($14,8+1,6=16,4$ students from the second group). From these students 11,1 adapted successfully the solution of the analogue for use with the target problem, while 1,5 students (who had previously considered both the remote analogue and the distractor as source problems) failed to do so. (12,1 and 1,5 students from the second group respectively). The rest (6,2 students from the first and 2,8 from the second group) returned to s_2 to retrieve a new source and through s_3 they came back to s_4 .

Finally 3,3 from these 6, 2 students of the first group adapted successfully the solution of the analogue for use with the target problem and 2,9 failed to do so (1,6 and 1,2 students from the second group respectively). Thus $11,1+3,3=14,4$ students from the first and $12,1+1,6=13,7$ students from the second group solved finally the target problems.

The "movements" of the students of the first group are shown in Figure 2. We observe that we have 35,3 in total "arrivals" to s_2 and 31,1 "departures" from s_2 to s_3 , therefore $p_{23} = \frac{31,1}{35,3} \approx 0,881$. In the same way one finds that $p_{21} = \frac{4,9}{35,3} \approx 0,119$, $p_{31} = \frac{1,2}{31,1} \approx 0,038$

M. Gr. Voskoglou

$$p_{32} = \frac{4,9}{31,1} \approx 0,158, \quad p_{34} = \frac{25}{31,1} \approx 0,804, \quad p_{41} = \frac{4,4}{25} \approx 0,176, \quad p_{42} = \frac{6,2}{25} \approx 0,248 \quad \text{and}$$

$$p_{45} = \frac{14,4}{25} \approx 0,576.$$

Replacing the values of the transition probabilities in the formulae of the model we find that the limiting probability vector for the first group is

$P \approx [0,157 \ 0,259 \ 0,231 \ 0,232 \ 0,121]$ and that $m \approx 7,264$ times.

Operating the analogous calculations for the second group we find that

$P \approx [0,154 \ 0,26 \ 0,237 \ 0,23 \ 0,119]$ and $m \approx 7,404$ times.

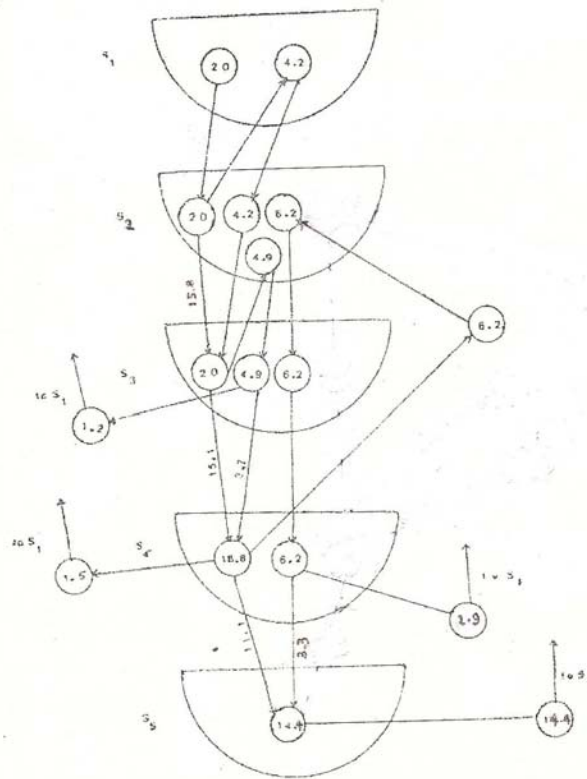


Figure 2. The “movements” of the students of the first group

The elements of P give the several probabilities about the “behavior” of each group during the AR process. Also, since $7,264 < 7,404$, the performance of the first group was slightly better.

According to the design of our experiment students had to choose the source problem between two given problems: A remote analogue to the target and a distractor problem. However, often things are not so simple. In fact, the individuals have usually to search in their memories to retrieve the source among several past problems sharing common surface and/or structural characteristics with the target. We could of course add in our experiment one or more problems among the candidate source problems.

Probability and Fuzzy Logic in Analogical Reasoning

Nevertheless, this manipulation would make the calculation of the transition probabilities between states of the chain more complicated, because the students' movements would be extended to several directions.

3. The fuzzy model

Human reasoning in general is characterized by a degree of uncertainty. In fact, the individual's cognition utilizes concepts that are inherently graded and therefore fuzzy. On the other hand, from the observer's point of view (e.g. teacher) there usually exists vagueness about the degree of the individual's success in each of the steps of the reasoning process.

Fuzzy logic, based on fuzzy sets theory introduced in 1965 by Zadeh (18) provides a rich and meaningful addition to standard logic in general and to probability theory in particular. The applications which may be generated from or adapted to fuzzy logic are wide-ranging (e.g. see [9], [11], [12], [20], etc) and provide the opportunity for modelling under conditions which are inherently imprecisely defined, despite the concerns of classical logicians. Some important applications of fuzzy logic were also attempted in the field of Education (e.g. see [1], [5], [13], [14], [15], [17], [20], [21], etc). All these gave us the impulsion to introduce principles of fuzzy logic to describe in a more effective way the process of AR. For general facts on fuzzy sets we refer freely to the book [7].

Let us consider a group of n analogical problem solvers, $n \geq 2$. Denote by A_i , $i=1,2,3$ the steps of search-retrieval, mapping and adaptation respectively. Denote also by a , b , c , d , and e the linguistic labels of negligible, low, intermediate, high and complete success respectively of the analogical problem solvers in each of the A_i 's. Set $U = \{a, b, c, d, e\}$ and let n_{ia} , n_{ib} , n_{ic} , n_{id} and n_{ie} denote the numbers of analogical problem solvers who faced negligible, low, intermediate, high and complete success at step A_i , $i=1,2,3$. We define the *membership function* m_{A_i} for each x in U , as follows:

$$m_{A_i}(x) = \begin{cases} 1, & \text{if } \frac{4n}{5} < n_{ix} \leq n \\ 0,75, & \text{if } \frac{3n}{5} < n_{ix} \leq \frac{4n}{5} \\ 0,5, & \text{if } \frac{2n}{5} < n_{ix} \leq \frac{3n}{5} \\ 0,25, & \text{if } \frac{n}{5} < n_{ix} \leq \frac{2n}{5} \\ 0, & \text{if } 0 \leq n_{ix} \leq \frac{n}{5} \end{cases}$$

Then A_i is represented as a fuzzy subset of U by

$$A_i = \{(x, m_{A_i}(x)): x \in U\}, i=1, 2, 3.$$

M. Gr. Voskoglou

In order to represent all possible solvers' profiles (*overall states*) during the AR process we consider a *fuzzy relation*, say R, in U^3 (i.e. a fuzzy subset of U^3) of the form $R = \{(s, m_R(s)) : s = (x, y, z) \in U^3\}$.

Since the degree of solvers' success at a certain step depends upon the degree of their success in the previous step and in order to determine properly the membership function m_R we give the following 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, (c, c, a) is a well ordered profile, while (b, a, c) is not.

We define now the *membership degree* of a profile s to be $m_R(s) = m_{A_1}(x)m_{A_2}(y)m_{A_3}(z)$, if s is well ordered, and 0 otherwise.

In fact, if for example profile (b, a, c) possessed a nonzero membership degree, how it could be possible for a solver, who failed at the step of mapping, to perform satisfactorily at the step of adaptation?

Next, for reasons of brevity, we shall write m_s instead of $m_R(s)$. Then the *possibility* r_s of the profile s is defined 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 the possibility of s expresses the "relative membership degree" of s with respect to $\max\{m_s\}$. Calculating the possibilities of all profiles one obtains a qualitative view of the group's performance during the AR process.

Further, the amount of information obtained by an action can be measured by the reduction of uncertainty resulting from this action. Accordingly the individuals' uncertainty during the AR process is connected to their capacity in obtaining relevant information. Therefore a measure of uncertainty could be adopted as a measure of the group's abilities in AR. For example, such a measure that we have used in an earlier paper, when developing an analogous fuzzy model for the problem solving process [21] is the *total possibilistic uncertainty* T of the group.

Here we shall use a classical measure of uncertainty and the associated information (known as the *Shannon's entropy* or the *Shannon- Wiener diversity index*), expressed in terms of the Dempster-Shafer mathematical theory of evidence, for use in a fuzzy environment in the form

$$H = -\frac{1}{\ln n} \sum_{s=1}^n m_s \ln m_s, \text{ where } n \text{ is the total number of elements of the corresponding}$$

fuzzy set ([8]; p.20). In the above formula the sum is divided by $\ln n$ in order to normalize H, so that its maximal value is 1 regardless the value of n .

Adopting H as a measure of the group's abilities in AR it becomes evident that the lower is the value of H (i.e. the higher is the reduction of the corresponding uncertainty), the better the group's abilities. An advantage of adopting H as a measure instead of T is that H is calculated directly from the membership degrees of all profiles s , in contrast to T that presupposes the calculation of the possibilities of all profiles first.

Probability and Fuzzy Logic in Analogical Reasoning

Assume now that one wants to study the *combined results of behaviour* of k different groups, $k \geq 2$, during the same process. For this we introduce the *fuzzy variables* $A_1(t)$, $A_2(t)$ and $A_3(t)$ with $t=1, 2, \dots, k$. The values of these variables represent fuzzy subsets of U corresponding to the steps of the AR process for each of the k groups; e.g. $A_1(2)$ represents the fuzzy subset of U corresponding to the step of search-retrieval for the second group ($t=2$). It becomes evident that, in order to measure the degree of evidence of combined results of the k groups, it is necessary to define the possibility $r(s)$ of each profile s with respect to its membership degrees for all groups. 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 combined results of behaviour of a group during different analogical problem solving processes.

Application: A similar classroom experiment with it described in section 2 was repeated a few days later with two different groups of 20 students of the Technological Educational Institute of Patras being at their second term of studies. The difference was that this time we added one more problem among the candidate source problems in each of the three cases, which was unrelated to the target problem in terms of both their surface and structure features (see Appendix). For the unrelated problem we followed the same process with the other problems, i.e. we gave a solution procedure to students and we allowed 10 minutes to solve it themselves by using the given procedure.

We also added one more case (see Appendix), i.e. we had 4 cases in total. Our characterization of students' performance at each step of the AR process involved:

- Negligible success, if they didn't obtain (at the particular step) positive results for the given problems.
- Low success, if they obtained positive results for 1 only of the given problems.
- Intermediate success, if they obtained positive results for 2 problems.
- High success, if they obtained positive results for 3 problems.
- Complete success, if they obtained positive results for all the given (4 in total) problems.

Examining students' papers of the first group we found that 9, 6 and 5 students achieved intermediate, high and complete success respectively at the step of search-retrieval in terms of choosing the correct problem (i.e. the remote analogue to the target) as the source problem. This means that $n_{1a}=n_{1b}=0$, $n_{1c}=9$, $n_{1d}=6$ and $n_{1e}=5$. Thus, according to the definition of $m_{A_i}(x)$, the step of search-retrieval corresponds to a fuzzy subset of U of the form:

$$A_1 = \{(a,0),(b,0),(c, 0.5),(d, 0.25),(e, 0.25)\}.$$

In the same way we represented the steps of mapping and adaptation as fuzzy subsets of U by

M. Gr. Voskoglou

$A_2 = \{(a,0),(b,0),(c, 0.5),(d, 0.25),(e,0)\}$ and

$A_3 = \{(a, 0.25),(b, 0.25),(c, 0.25),(d,0),(e,0)\}$ respectively.

Next, we calculated the membership degrees of the 5^3 (ordered samples with replacement of 3 objects taken from 5) in total possible students' profiles (see column of $m_s(1)$ in Table 1). For example, for $s=(c, c, a)$ one finds that $m_s = m_{A_1}(c) \cdot m_{A_2}(c) \cdot m_{A_3}(a) = (0.5) \cdot (0.5) \cdot (0.25) = 0.06225$.

It turned out that (c, c, a) was one of the profiles possessing the maximal membership degree and therefore the possibility of each s in U^3 is given by $r_s = \frac{m_s}{0,06225}$.

Using this formula we calculated the possibilities of all profiles (see column of $r_s(1)$ in Table 1).

Finally we calculated the Shannon's entropy in terms of the values of column $m_s(1)$ in Table 1, where $n=125$ and we found that $H=0,289$.

Working as above for the second group we found that

$A_1 = \{(a,0),(b, 0.25),(c, 0.5),(d, 0.25),(e,0)\}$,

$A_2 = \{(a, 0.25),(b, 0.25),(c, 0.5),(d, 0),(e,0)\}$,

$A_3 = \{(a, 0.25),(b, 0.25),(c,0.25),(d,0),(e,0)\}$.

The membership degrees of all possible profiles of the second group are shown in column of $m_s(2)$ of Table 1. It turned out that the maximal membership degree was again 0.06225, therefore the possibility of each s is calculated by the same formula as for the first group. The possibilities of all profiles are shown in column of $r_s(2)$ of Table 1, while for the Shannon's entropy we found that $H=0,312$. Thus, since $0,289 < 0,312$, the general performance of the first group was slightly better.

Next, in order to study the combined results of behaviour of the two groups, we introduced the fuzzy variables $A_i(t)$, $i=1, 2, 3$ and $t=1, 2$, as we have described them in the model. Then the pseudo-frequency of each student profile s is given by $f(s) = m_s(1) + m_s(2)$ (see the corresponding column in Table 1). It turns out that the highest pseudo-frequency is 0,124 and therefore the possibility of each student's profile is given by $r(s) = \frac{f(s)}{0,124}$. The possibilities of all profiles having non-zero pseudo-frequencies are

presented in the last column of Table 1.

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	0.258	0.016	0.129
B	b	a	0	0	0.016	0.258	0.016	0.129
B	a	a	0	0	0.016	0.258	0.016	0.129
C	c	c	0.062	1	0.062	1	0.124	1
C	c	a	0.062	1	0.062	1	0.124	1
C	c	b	0	0	0.031	0.5	0.031	0.25
C	a	a	0	0	0.031	0.5	0.031	0.25
C	b	a	0	0	0.031	0.5	0.031	0.25

Probability and Fuzzy Logic in Analogical Reasoning

C	b	b	0	0	0.031	0.5	0.031	0.25
D	d	a	0.016	0.258	0	0	0.016	0.129
D	d	b	0.016	0.258	0	0	0.016	0.129
D	d	c	0.016	0.258	0	0	0.016	0.129
D	a	a	0	0	0.016	0.258	0.016	0.129
D	b	a	0	0	0.016	0.258	0.016	0.129
D	b	b	0	0	0.016	0.258	0.016	0.129
D	c	a	0.031	0.5	0.031	0.5	0.062	0.5
D	c	b	0.031	0.5	0.031	0.5	0.062	0.5
D	c	c	0.031	0.5	0.031	0.5	0.062	0.5
E	c	a	0.031	0.5	0	0	0.031	0.25
E	c	b	0.031	0.5	0	0	0.031	0.25
E	c	c	0.031	0.5	0	0	0.031	0.25
E	d	a	0.016	0.258	0	0	0.016	0.129
E	d	b	0.016	0.258	0	0	0.016	0.129
E	D	c	0.016	0.258	0	0	0.016	0.129

Table 1. Profiles with non zero membership degrees

4. The centroid defuzzification technique

Defuzzification is the process of producing a quantifiable result in fuzzy logic given fuzzy sets and corresponding membership degrees. A common and useful defuzzification technique is the method of the *centre of gravity*, usually referred as the *centroid method* (e.g. see [16]). According to this method, given a fuzzy subset $A = \{(x, m(x)): x \in U\}$ of the universal set U of the discourse with membership function $m: U \rightarrow [0, 1]$, we correspond to each $x \in U$ an interval of values from a prefixed numerical distribution, which actually means that we replace U with a set of real intervals. Then, we construct the graph F of the membership function $y=m(x)$. There is a commonly used in fuzzy logic approach to measure performance with the pair of numbers (x_c, y_c) as the coordinates of the centre of gravity, say F_c , of the graph F , which we can calculate using the following well-known from mechanics formulas:

$$x_c = \frac{\iint_F x dx dy}{\iint_F dx dy}, y_c = \frac{\iint_F y dx dy}{\iint_F dx dy} \quad (1)$$

Subbotin et al. adapted the centroid method for use with our fuzzy model for the process of learning [17] and they have applied it on comparing students' mathematical learning abilities [13] and in other cases (e.g. see [14], [15], etc).

In this paper we shall apply this method as a defuzzification technique of the fuzzy outputs of our model for AR developed in the previous section. For this, we characterize a student's performance as very low if $x \in [0, 1)$, as low if $x \in [1, 2)$, as intermediate if $x \in [2, 3)$, as high if $x \in [3, 4)$ and as very high if $x \in [4, 5]$. This characterization is obtained through the evaluation of the students' papers. Consequently if $x \in [0, 1)$ then $y=m(x)=m(a)$, if $x \in [1, 2)$ then $y=m(x)=b$, etc.

Now the graph F of the corresponding fuzzy subset of U takes the form of the bar graph of Figure 3 consisting of 5 rectangles, say F_i , $i=1,2,3,4,5$, having the lengths of their sides on the x axis equal to 1.

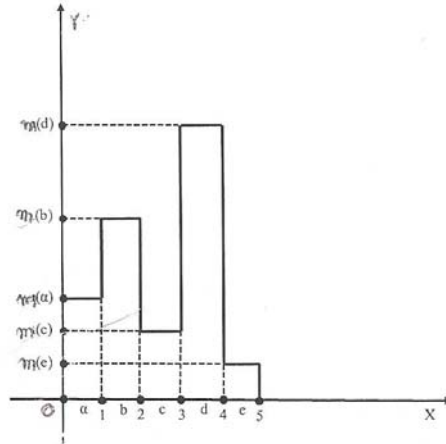


Figure 3. Bar graphical data representation

In this case $\iint_F dx dy$, which is the total area of F , is equal to $\sum_{i=1}^5 y_i$. We also have that

$$\iint_F x dx dy = \sum_{i=1}^5 \iint_{F_i} x dx dy = \sum_{i=1}^5 \int_{i-1}^i \int_0^{y_i} x dy dx = \sum_{i=1}^5 y_i \int_{i-1}^i x dx = \frac{1}{2} \sum_{i=1}^5 (2i-1) y_i, \quad \text{and} \quad \iint_F y dx dy =$$

$$\sum_{i=1}^5 \iint_{F_i} y dx dy = \sum_{i=1}^5 \int_{i-1}^i \int_0^{y_i} y dy dx = \frac{1}{2} \sum_{i=1}^5 y_i^2. \quad \text{Therefore formulas (1) are transformed into the}$$

following form:

$$x_c = \frac{1}{2} \left(\frac{y_1 + 3y_2 + 5y_3 + 7y_4 + 9y_5}{y_1 + y_2 + y_3 + y_4 + y_5} \right),$$

$$y_c = \frac{1}{2} \left(\frac{y_1^2 + y_2^2 + y_3^2 + y_4^2 + y_5^2}{y_1 + y_2 + y_3 + y_4 + y_5} \right).$$

Normalizing our fuzzy data by dividing each $m(x)$, $x \in U$, with the sum of all membership degrees we can assume without loss of the generality that $y_1 + y_2 + y_3 + y_4 + y_5 = 1$. Therefore we can write:

$$x_c = \frac{1}{2} (y_1 + 3y_2 + 5y_3 + 7y_4 + 9y_5),$$

$$y_c = \frac{1}{2} (y_1^2 + y_2^2 + y_3^2 + y_4^2 + y_5^2) \quad (2)$$

with $y_i = \frac{m(x_i)}{\sum_{x \in U} m(x)}$, where $x_1 = a$, $x_2 = b$, $x_3 = c$, $x_4 = d$ and $x_5 = e$.

But $0 \leq (y_1 - y_2)^2 = y_1^2 + y_2^2 - 2y_1 y_2$, therefore $y_1^2 + y_2^2 \geq 2y_1 y_2$, with the equality holding if, and only if, $y_1 = y_2$. In the same way one finds that $y_1^2 + y_3^2 \geq 2y_1 y_3$, and so on. Hence it is easy

Probability and Fuzzy Logic in Analogical Reasoning

to check that $(y_1+y_2+y_3+y_4+y_5)^2 \leq 5(y_1^2+y_2^2+y_3^2+y_4^2+y_5^2)$, with the equality holding if, and only if $y_1=y_2=y_3=y_4=y_5$. But $y_1+y_2+y_3+y_4+y_5 = 1$, therefore $1 \leq 5(y_1^2+y_2^2+y_3^2+y_4^2+y_5^2)$ (3), with the equality holding if, and only if $y_1=y_2=y_3=y_4=y_5=\frac{1}{5}$.

. In this case the first of formulas (2) gives that $x_c = \frac{5}{2}$. Further, combining the inequality (3) with the second of formulas (2) one finds that $1 \leq 10y_c$, or $y_c \geq \frac{1}{10}$. Therefore the unique minimum for y_c corresponds to the centre of gravity $F_m(\frac{5}{2}, \frac{1}{10})$.

The ideal case is when $y_1=y_2=y_3=y_4=0$ and $y_5=1$. Then from formulas (2) we get that $x_c = \frac{9}{2}$ and $y_c = \frac{1}{2}$. Therefore the centre of gravity in the ideal case is the point $F_i(\frac{9}{2}, \frac{1}{2})$. On the other hand the worst case is when $y_1=1$ and $y_2=y_3=y_4=y_5=0$. Then for formulas (2) we find that the centre of gravity is the point $F_w(\frac{1}{2}, \frac{1}{2})$. Thus, the “area”

where the centre of gravity F_c lies is represented by the triangle $F_w F_m F_i$ of Figure 4. Then from elementary geometric considerations it follows that for two groups of students with the same $x_c \geq 2.5$ the group having the centre of gravity which is situated closer to F_i is the group with the higher y_c ; and for two groups with the same $x_c < 2.5$ the group having the centre of gravity which is situated farther to F_w is the group with the lower y_c . Based on the above considerations we formulate our criterion for comparing the groups’ performances as follows:

- Among two or more groups the group with the biggest x_c performs better.
- If two or more groups have the same $x_c \geq 2.5$, then the group with the higher y_c performs better.
- If two or more groups have the same $x_c < 2.5$, then the group with the lower y_c performs better.

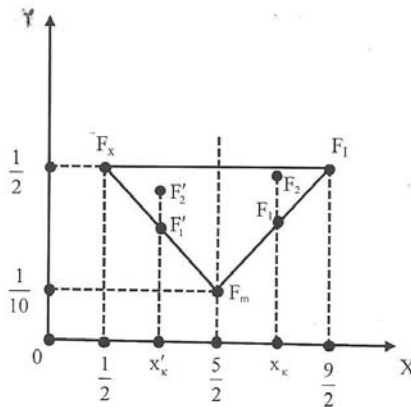


Figure 4. Graphical representation of the “area” of the centre of gravity

Next we give an example illustrating our results in practice.

Example: Let us reconsider the example with the two groups of students of the Graduate Technological Educational Institute of Patras presented in section 3. The fuzzy data obtained was the following:

First group

$$A_{11} = \{(a,0),(b,0),(c, 0.5),(d, 0.25),(e,0.25)\}, A_{12} = \{(a,0),(b,0),(c, 0.5),(d, 0.25),(e,0)\},$$

$$A_{13} = \{(a, 0.25),(b, 0.25),(c, 0.25),(d,0),(e,0)\}.$$

Second group

$$A_{21} = \{(a, 0),(b, 0.25),(c, 0.5),(d, 0.25),(e, 0)\}, A_{22} = \{(a, 0.25),(b, 0.25),(c, 0.5),(d, 0),(e, 0)\},$$

$$A_{23} = \{(a, 0.25),(b, 0.25),(c, 0.25),(d, 0),(e, 0)\}.$$

According to the above notation the first index of A_{ij} denotes the group ($i=1, 2$) and the second index denotes the corresponding step A_j of the AR process. We recall that the steps of the AR process as they have been considered in our model developed in [8] are A_1 : search-retrieval, A_2 : mapping and A_3 : adaptation respectively.

We compare now the two groups' performance with the centroid method. Applying the first of formulas (2) for the first step of search-retrieval we find

$$x_{c11} = \frac{1}{2}(5 \times 0.5 + 7 \times 0.25 + 9 \times 0.25) = 3.25, x_{c21} = \frac{1}{2}(3 \times 0.25 + 5 \times 0.5 + 7 \times 0.25) = 2.25$$

Thus, by our criterion the first group demonstrates better performance.

For the second step of mapping normalizing the membership degrees of A_{12} ($0.5 : 0.75 \approx 0.67$ and $0.25 : 0.75 \approx 0.33$) we get

$$A_{12} = \{(a, 0),(b, 0),(c, 0.67),(d, 0.33),(e, 0)\}.$$

$$\text{Therefore we have } x_{c12} = \frac{1}{2}(5 \times 0.67 + 7 \times 0.33) = 2.83, x_{c22} = \frac{1}{2}(0.25 + 3 \times 0.25 + 5 \times 0.25) = 1.125$$

By our criterion, the first group again demonstrates better performance.

Finally, for the third step of adaptation we have $A_{13} = A_{23} = \{(a, 0.25),(b, 0.25),(c, 0.25),(d, 0),(e, 0)\}$, which obviously means that in this case the performances of both groups are identical.

Based on our calculations we can conclude that the first group demonstrated better performance at the first two steps of the AR process, while at the third step the performances of both groups were identical.

Among the advantages of the centroid method is that no complicated calculations are needed in its final step (application of the formulas obtained). Further it enables one to compare the performances of the student groups' at each stage of the AR process.

Apart from the centroid method there are also other defuzzification techniques in use. For example, in section 3 we have used the total student group's uncertainty during the AR process as a measure of its performance. The above two assessment approaches treat differently the idea of the students' performance and therefore the results obtained may

Probability and Fuzzy Logic in Analogical Reasoning

differ to each other. In fact, in the first case the student group's uncertainty during the AR process is connected to its capacity in obtaining the relevant information. In other words, in this case we are looking for the *average group's performance*. On the other hand, in the case of the centroid technique the *weighted average* plays the main role, i.e. the results of the performance close to the ideal performance have much more weight than those close to the lower end. In other words, in this case we are mostly looking at the *quality* of the performance. It is argued that the combined application of these two approaches helps in finding the ideal profile of performance according to the user's personal criteria of goals and therefore to finally choosing the appropriate approach for measuring the results of his/her experiments.

5. Students' individual assessment

The outputs of our fuzzy model for AR can be used not only for assessing the performance of student groups', but also for the students' individual assessment. In fact, if $n=1$ (we recall that n denotes the number of students' of the group under study), then from the definition of the membership function m_{A_i} given in section 3 it becomes evident that in each A_i , $i = 1, 2, 3$, there exists a unique element x of U with membership degree 1 , while all the others have membership degree 0 . In this case it is straightforward to compare to each other the student's performance at each step of the AR process. For example, if $A_{1I} = \{(a, 0), (b, 0), (c, 0), (d, 1), (e, 0)\}$ and $A_{2I} = \{(a, 0), (b, 0), (c, 1), (d, 0), (e, 0)\}$, then obviously the first student demonstrates a better performance than the second one at the step of search/retrieval. This is trivially crossed by the centroid method, since $x_{c1I} = \frac{7}{2}$ and $x_{c2I} = \frac{5}{2}$.

As a consequence of the above situation ($n=1$), there exists a unique student profile s with $m_s = 1$, while all the others have membership degree 0 . In other words, each student is characterized in this case by a unique profile, which gives us the requested information about his/her total performance. For example, if (c, b, a) and (c, b, b) are the characteristic profiles for students x and y respectively, then clearly y demonstrates a better performance than x . On the contrary, if (d, b, b) and (c, c, b) are the corresponding profiles, then x demonstrates a better performance than y at the step of search/retrieval, but y demonstrates a better performance than x at the step of mapping. Mathematically speaking this means that the students' characteristic profiles define a relationship of partial order among students' with respect to their performance.

Jones developed a fuzzy model to the field of Education involving several theoretical constructs related to assessment, amongst which is a technique for assessing the deviation of a student's knowledge with respect to the teacher's knowledge, which is taken as a reference ([1] and [5]). Here we shall present this technique, properly adapted with respect to our fuzzy model, as an alternative fuzzy method for the students' individual assessment.

Let $X = \{A_1, A_2, A_3\}$ be the set of the steps of the AR process mentioned in the example of section 2. Then a fuzzy subset of X of the form $\{(A_1, m(A_1)), (A_2, m(A_2)), (A_3, m(A_3))\}$ can be assigned to each student, where the membership function m takes the values $0, 0.25, 0.5, 0.75, 1$ according to the level of the student's performance at the

corresponding step. The teacher's fuzzy measurement is always equal to I , which means that the fuzzy subset of X corresponding to the teacher is $\{(A_1, I), (A_2, I), (A_3, I)\}$.

Then the *fuzzy deviation* of the student i with respect to the teacher is defined to be the fuzzy subset

$$D_i = \{(A_1, I - m(A_1)), (A_2, I - m(A_2)), (A_3, I - m(A_3))\} \text{ of } X.$$

This assessment by reference to the teacher provides us with the ideal student as the one with nil deviation in all his/her components and it defines a relationship of partial order among students'. The following example illustrates this theoretical framework in practice.

Example: The same experiment with it described in section 3 was repeated with another group of 35 students of the Technological Educational Institute of Patras, Greece. This time in assessing the students' individual performance by applying the A. Jones technique we found the following types of deviations with respect to the teacher:

$$D_1 = \{(A_1, 0.75), (A_2, 0.75), (A_3, I)\} \text{ (this type of deviation was related with 2 students)}$$

$$D_2 = \{(A_1, 0.5), (A_2, I), (A_3, I)\} \text{ (related with 7 students)}$$

$$D_3 = \{(A_1, 0.5), (A_2, 0.75), (A_3, I)\} \text{ (related with 5 students)}$$

$$D_4 = \{(A_1, 0.5), (A_2, 0.75), (A_3, 0.75)\} \text{ (related with 4 students)}$$

$$D_5 = \{(A_1, 0.25), (A_2, 0.5), (A_3, 0.75)\} \text{ (related with 3 students)}$$

$$D_6 = \{(A_1, 0.25), (A_2, 0.25), (A_3, 0.5)\} \text{ (related with 6 students)}$$

$$D_7 = \{(A_1, 0), (A_2, 0.5), (A_3, 0.75)\} \text{ (related with 1 student)}$$

$$D_8 = \{(A_1, 0), (A_2, 0.5), (A_3, 0.5)\} \text{ (related with 2 students)}$$

$$D_9 = \{(A_1, 0), (A_2, 0.25), (A_3, 0.5)\} \text{ (related with 1 student)}$$

$$D_{10} = \{(A_1, 0), (A_2, 0.25), (A_3, 0.25)\} \text{ (related with 3 students)}$$

$$D_{11} = \{(A_1, 0), (A_2, 0), (A_3, 0.25)\} \text{ (related with 1 student)}$$

On comparing the above types of students' deviations it becomes evident that the students possessing the type D_3 of deviation demonstrate a better performance than those possessing the type D_1 , the students possessing the type D_4 demonstrate a better performance than those possessing the type D_3 and so on. However, the students possessing the type D_1 demonstrate a better performance at the step of mapping than those possessing the type D_2 , who demonstrate a better performance at the step of search/retrieval. Similarly, the students possessing the type D_6 demonstrate a better performance at the steps of mapping and adaptation than those possessing the type D_7 , who demonstrate a better performance at the step of search/retrieval.

Notice that each deviation D_i corresponds to a student's profile s_i , $i = 1, 2, \dots, 11$. For example, the deviation D_1 corresponds to the student $\{(A_1, 0.25), (A_2, 0.25), (A_3, 0)\}$, whose profile is $s_1 = (b, b, a)$. Applying the same argument one finally finds the following profiles characterizing the students' performance in our experiment:

$$s_1 = (b, b, a) \text{ (this profile is related with 2 students)}$$

$$s_2 = (c, a, a) \text{ (related with 7 students)}$$

$$s_3 = (c, b, a) \text{ (related with 5 students)}$$

$$s_4 = (c, b, b) \text{ (related with 4 students)}$$

$$s_5 = (d, c, b) \text{ (related with 3 students)}$$

$$s_6 = (d, d, c) \text{ (related with 6 students)}$$

Probability and Fuzzy Logic in Analogical Reasoning

$s_7 = (e, c, b)$ (related with 1 student)

$s_8 = (e, c, c)$ (related with 2 students)

$s_9 = (e, d, c)$ (related with 1 student)

$s_{10} = (e, d, d)$ (related with 3 students)

$s_{11} = (e, e, d)$ (related with 1 student)

In other words, the A. Jones technique is actually equivalent to our method for the students' individual assessment. The only difference is that the former expresses the fuzzy data with numerical values, while the latter expresses it qualitatively in terms of the fuzzy linguistic labels of U .

Notice also that the teacher may put a target for his/her class and may establish didactic strategies in order to achieve it. For example he/she may ask for the deviation, say D , with respect to the teacher to be $0.25 \leq D \leq 0.5$, for all students and in all steps. In this case the application of the A. Jones technique could help the teacher to determine the divergences with respect to this target and hence to readapt his/her didactic plans in order to diminish these divergences.

6. Conclusions and discussion

In this paper we developed two mathematical models for the description of the process of AR: A stochastic model by introducing a finite ergodic Markov chain on the steps of the AR process and a fuzzy model by representing the main steps of the AR process as fuzzy subsets of a set of linguistic labels characterizing the individuals' performance in each of these steps. The Shannon's entropy (system's uncertainty) and the centroid method were used as defuzzification techniques in the latter case.

Both models give important quantitative information about the abilities of a group of analogical problem solvers'. In the stochastic model the calculation of the transition probabilities between states of the chain is getting more complicated when the source must be chosen among more than two given problems, because the individuals' "movements" in this case are extended to more directions. On the contrary, there is not any particular difficulty in this case with the fuzzy model. Moreover the fuzzy model gives a qualitative view of the group's performance through the calculation of the possibilities of all individuals' profiles during the AR process and it can be also implemented for the students' individual assessment. Finally, an additional advantage of the fuzzy model is that it gives to the researcher the opportunity to study the combined results of the behaviour of two or more groups during the AR process or alternatively to study the combined results of the behaviour of the same group during different analogical problem solving processes.

On the other hand the characterization of the analogical problem solvers' performance in terms of a set of linguistic labels which are fuzzy by themselves is a disadvantage of the fuzzy model, because this characterization depends on the researcher's personal criteria (see for example in section 3 the criteria used in our experiments for characterizing the students' performance). Therefore the combined use of the two models seems to be the best solution in achieving a worthy of credit mathematical analysis of the AR process.

APPENDIX: Problems given for solution in the classroom experiments

CASE 1

M. Gr. Voskoglou

Target problem: A box contains 8 balls numbered from 1 to 8. One makes three successive drawings, putting back the corresponding ball to the box before the next drawing. Find the probability of getting all the balls drawing out of the box different to each other. -

The probability is equal to the quotient of the total number of the ordered samples of 3 objects from 8 (favourable outcomes) to the total number of the corresponding samples with replacement (possible outcomes).

Remote analogue: How many numbers of 2 digits can be formed by using the digits from 1 to 6 and how many of them have their digits different?

Solution procedure given to the students: Find the total number of the ordered samples of 2 objects from 6 with and without replacement respectively.

Distractor problem: A box contains 3 white, 4 blue and 6 black balls. If we draw out 2 balls, what is the probability to be of the same colour?

Solution procedure given to the students: The number of all favourable outcomes is equal to the sum of the total number of combinations of 3, 4 and 6 objects taken 2 at each time respectively, while the number of all possible outcomes is equal to the total number of combinations of 13 objects taken 2 at each time.

Unrelated problem (Used only with the fuzzy model): Find the number of all possible anagrammatisms of the word "SIMPLE". How many of them start with S and how many of them start with S and end with E?

Solution procedure given to the students: The number of all possible anagrammatisms is equal to the total number 6! of permutations of 6 objects. The anagrammatisms starting with S are 5! And the anagrammatisms starting with S and ending with E are 4!

CASE 2

Target problem: Consider the matrices:

$$A = \begin{bmatrix} 1 & -a & -a \\ 0 & 1 & -a \\ 0 & 0 & 1 \end{bmatrix} \quad \text{καα} \quad B = \begin{bmatrix} 0 & -a & -a \\ 0 & 0 & -a \\ 0 & 0 & 0 \end{bmatrix} .$$

Prove that $A^n = A + (n-1)(B + \frac{n}{2}B)$, for every positive integer n . -

Since $A=I+B$, where I stands for the unitary 3X3 matrix, and $B^3=0$, is $A^n=(I+B)^n=I+nB+\frac{n(n-1)}{2}B^2=A+(n-1)B+\frac{n(n-1)}{2}B=A+(n-1)(B+\frac{n}{2}B)$.

Remote analogue: Let α be a nonzero real number. Prove that $\alpha^n = \sum_{i=0}^n \binom{n}{i} (\alpha-1)^i$, for

all positive integers n .

Solution procedure given to the students: Write $\alpha = 1+(\alpha-1)$ and apply the Newton's

formula $(x+b)^n = \sum_{i=0}^n \binom{n}{i} x^{n-i} b^i$, setting $x=1$ and $b=\alpha-1$.

Distractor problem: If A and B are as in the target problem, calculate $(A+B)^2$. -

The students were asked to operate the corresponding calculations.

Probability and Fuzzy Logic in Analogical Reasoning

Unrelated problem (Used only with the fuzzy model): Prove that $1+2+\dots+n=\frac{n(n+1)}{2}$,

for all positive integers n .

The students were asked to apply induction on n .

CASE 3

Target problem: The price of sale of a good depends upon its total demand Q and it is given by $P(Q) = \frac{1}{2}Q - 50$, while the cost of production of the good is given by $C(Q) = \frac{1}{4}Q^2 + 35Q + 25$. Find the quantity Q of the good's total demand maximizing the profit from sale.

The revenue from sale is equal to $P(Q)Q$ and therefore the profit from sale is given by $K(Q) = P(Q)Q - C(Q)$. The maximum of function $K(Q)$ is calculated by using the well known theorem of derivatives.

Remote analogue: A car is entering in a road having initial speed 50 Km/h, which is changed according to the relation $U(t) = 3t^2 - 12t + 50$, where t represents the time (in minutes) during which the car is moving on this road. Find the minimal speed of the car on this road.

The students were asked to apply the well known theorem of derivatives in order to calculate the minimum of the function $U(t)$.

Distractor problem: The price of sale of a good depends upon its total demand Q and it is given by $P(Q) = 25 - Q^2$. The price is finally fixed to 9 monetary units and therefore the consumers who would be willing to pay more than this price benefit. Find the total benefit to consumers (Dowling 1980, paragraph 17.7: *Consumer's surplus*).

Solution procedure given to the students: For $P=9$ and since $Q \geq 0$, it turns out that $Q=4$ metric units. Drawing the graph of the function $P(Q)$ (parabola) it is easy to observe

that the total benefit to consumers is equal to $\int_0^4 P(Q)dQ - 4 \cdot 9$ monetary units

Unrelated problem (Used only with the fuzzy model): Find the area under the curve $y=4x^2+2$.

Solution procedure given to the students: The area is given by $\int (4x^2 + 2)dx$.

CASE 4 (Used only with the fuzzy model)

Target problem: A producer has a stock of wine greater than 500 and less than 750 kilos. He has estimated that, if he had the double quantity of wine and transfused it to bottles of 12, or 25, or 40 kilos, it would be left over 6 kilos at each time. Find the quantity of the stock.

If Q is the quantity of stock, then, since the lowest common multiple of 12, 25 and 40 is 600, $2Q-6$ is a multiple of 600, therefore $2Q=606$, or $2Q=1212$, or $2Q=1818$, etc. But $500 < Q < 750$, therefore $Q=603$ kilos.

Remote analogue: An employer occupies less than 50 workers. If he occupied the triple number of workers and 3 more, then he could distribute them in bands of 8 or 12 or 15 workers. How many workers he occupies?

M. Gr. Voskoglou

Solution procedure given to the students: If x is the number of workers, then, since the lowest common multiple of 8, 12 and 15 is 120, $3x+3$ is a multiple of 120, or $x+1$ is a multiple of 40.

Distractor problem: A producer has a stock of 3400 and 5025 kilos respectively of two different kinds of wine and he decides to distribute these quantities to the maximal possible number of customers. After this distribution, they remained 25 kilos from each kind of wine in his barrels. How many of his customers he succeeded to satisfy with this manipulation?

Solution procedure given to the students: The number of customers is equal to the greatest common divisor of $3400-25$ and $5025-25$.

Unrelated problem: The number of students of a school is between 300 and 400. When they tried marching in rows of 10 the last row had 9 students, while when they tried marching in rows of 9 the last row had 7 students. How many are the students?

Solution procedure given to the students: Let $a=100x+10y+z$ be the number of students of the school. Then $a=10t-1$ for some positive integer t . Therefore $z=9$ and $x=3$. Further $a=9s+7$ for some positive integer s , or $a-7=9s$. But, since 9 divides $a-7$, 9 divides also the sum of the digits of $a-7$, i.e. $(3+y+9)-7=9k$ for some positive integer k , or $y+5=9k$. But $0 < y \leq 9$, therefore $y=4$.

REFERENCES

1. E. A.Espin and C. M. L.Oliveras, Introduction to the Use of Fuzzy Logic in the Assessment of Mathematics Teachers' Professional Knowledge, Proceedings of the First Mediterranean Conference on Mathematics, 107-113, Cyprus, 1997.
2. D.Gentner and C.Toupin, Systematicity and surface similarity in development of analogy, *Cognitive Science*, 10 (1986) 277-300.
3. D.Gentner and A.B.Markman, Structure mapping in analogy and similarity, *American Psychologist* 52 (1997) 45-56.
4. K.J.Holyoak, The pragmatics of analogical transfer, in G. H. Bower (Ed.), *The psychology of learning and motivation*, Vol. 19, 59-87, New York: Academic Press, 1985.
5. A.Jones, A.Kaufman and H.J.Zimmerman (Eds.), *Fuzzy Sets. Theory and Applications*, NATO ASI Series, Series C: Mathematical and Physical Sciences, Vol. 177, Reidel Publishing Company, Dordrecht, Holland, 1986.
6. J.Kemeny and J.Snell, *Finite Markov Chains*, Springer-Verlag, New York, 1976.
7. G.J.Klir and T.A.Folger, *Fuzzy Sets, Uncertainty and Information*, Prentice-Hall, London, 1988.
8. G.J.Klir, Principles of uncertainty: What are they? Why do we need them?, *Fuzzy Sets and Systems*, 74 (1995) 15-31.
9. R.Muthuraj and M.Sridharan, Bipolar Anti Fuzzy HX Group and its Lower Level Sub HX Groups, *Journal of Physical Sciences*, 16 (2012) 157-169.
10. L.R.Novick, Analogical transfer, problem similarity and expertise, *Journal of Educational Psychology: Learning, Memory and Cognition*, 14 (1988) 510-520.

Probability and Fuzzy Logic in Analogical Reasoning

11. R.Pradhan and M.Pal, Generalized Inverse of Block Intuitionistic Fuzzy Matrices, *International Journal of Applications of Fuzzy Sets and Artificial Intelligence*, 3 (2013) 23-38.
12. S.Rezvani, A New Approach Ranking of Exponential Trapezoid Fuzzy Numbers, *Journal of Physical Sciences*, 16 (2012) 45-57.
13. I.Subbotin, H.Badkoobehi and N.Bilotskii, Application of Fuzzy Logic to Learning Assessment, *Didactics of Mathematics: Problems and Investigations*, 22 (2004) 38-41.
14. I.Subbotin, F.Mossovar-Rahmani and N.Bilotskii, Fuzzy logic and the concept of the Zone of Proximate Development, *Didactics of Mathematics: Problems and Investigations*, 36 (2011) 101- 108.
15. I.Subbotin and M. Gr.Voskoglou, Applications of Fuzzy Logic to Case-Based Reasoning, *International Journal of Applications of Fuzzy Sets and Artificial Intelligence*, 1 (2011) 7-18.
16. E.Van Broekhoven and B.D.Baets, Fast and accurate centre of gravity defuzzification of fuzzy system outputs defined on trapezoidal fuzzy partitions, *Fuzzy Sets and Systems*, 157(7) (2006) 904-918.
17. M. Gr.Voskoglou, The process of learning mathematics: A fuzzy set approach, *Heuristics and Didactics of Exact Sciences*, 10 (1999) 9-13.
18. M. Gr. Voskoglou, Analogical Problem Solving and Transfer, *Proceedings 3d Mediterranean Conference on Mathematical Education*, 295-303, Athens, 2003.
19. M. Gr. Voskoglou, A stochastic model for the modelling process, in: C. Chaines, P. Galbraith, W. Blum and S. Khan (Eds.) *Mathematical Modelling: Education, Engineering and Economics (ICTMA 12)* 149-157, Horwood Publishing, Chichester, 2007.
20. M.Gr.Voskoglou, *Stochastic and fuzzy models in Mathematics Education*, Artificial Intelligence and Management, Lambert Academic Publishing, Saarbrucken, Germany, 2011 (for more details look at <http://amzn.com/3846528218>).
21. M.Gr.Voskoglou, Fuzzy Logic and Uncertainty in Problem Solving, *Journal of Mathematical Sciences and Mathematics Education*, 7(1) (2012) 37-19.
22. L.A.Zadeh, Fuzzy Sets, *Information and Control*, 8 (1965) 338-353.