An Extension of Fuzzy Deformable Prototypes for predicting student performance on Web-based Tutoring Systems

M. Rosario Vázquez[1] Francisco P. Romero[2] Jorge Ruiz-Vanoye[3] Jose A. Olivas[2] Jesus Serrano-Guerrero[2]

[1]Autonomous University of Carmen, Mexico, marosariovazquez@gmail.com
[2]University of Castilla La Mancha, Spain, {franciscoP.romero, JoseAngel.Olivas, Jesus.Serrano}@uclm.es
[3]Autonomous University of Hidalgo State, Mexico, jorge@ruizvanoye.com

Abstract

This paper presents an extension of Fuzzy Deformable Prototypes (FDPs) based on the use of interval type-2 fuzzy sets. The aim is to improve FDPs' capabilities for managing uncertainty and imprecision. This extension is applied to predict the academic performance of the students who make use of Web-based tutoring systems. The prediction model contains patterns of behavior that are used to determine the future academic performance of new students based on their affinity with the prototypes previously discovered. Interval Type 2 Fuzzy Sets (IT2FS) were used to handle the imprecision of the academic data caused by the overlapping between the fuzzy representations of prototypes.

Keywords: Fuzzy Deformable Prototypes, Interval Type 2 Fuzzy Sets, Web-based Tutoring Systems

1. Introduction

Nowadays, studentsâ performance is a top priority for academic institutions. This is why educators are specilly focused in the use of tools different tools improve academic achievement. Web tutoring systems have proved useful to improve the academic success. On the other hand,predictions models enable educational institutions to identify students with low academic performance. Based on these models, it is possible to plan strategies that contribute to increase the academic level.

Researchers have focused their attention on predicting academic performance of students who make use of the tutoring systems by analyzing large volumes of data [1]. The tutoring systems decompose a given task in a series of activities to help the student to reach the solution. The works published around the 2010 KDD Cup competition [2] are good examples of the analysis of large volumes of data that generate the tutoring systems. The contest aims to predict the probability of a student answering a particular activity that belongs to a specific problem correctly. On the other hand, in the filed of computing with words and perceptions, perceptual calculation [3] has been applied to evaluate Learning Outcomes in an outcome-based education system[4].

Some research is based on the assumption that the student receives feedback and this feedback influences in their learning process, so they established a relation between hints requested by students who use tutoring systems and the amount of learning acquired. Koedinger et al. [5] proposed a method to estimate the knowledge acquired by students in these contexts. They analyzed the data contained in the log file that generates a tutoring system. Also, they analyzed the skills required in each issue, and the number of attempts until a success is obtained. Beck [6] proposed a model to obtain the academic performance of the students from the time they use to solve each problem in a tutoring system. In this work, a logistic curve represents the evolution of the students' academic performance. Beck changed the theory IRT (Item Response Theory), which provides a framework for predicting the probability that a student with a certain level of ability to successfully answer a question, by introducing new criteria such as response time and the difficulty of each question. On the other hand, [7] proposed a method to estimate the student's academic performance based on the difficulty of the problem computed across three factors: correctness, time and the requested help.

In this paper, fuzzy logic techniques are used to predict the academic performance of a student in the context of Web-based tutoring systems. The use of Fuzzy Deformable Prototypes [8] permit to adapt the behavior patterns (prototypes) discovered to completely describe a new situation. The objective of this research is to extend the capabiliteies of Fuzzy Deformable Prototypes by adding the use of type-2 fuzzy sets to manage the imprecision of existing data. With these new capabilities, it is possible to obtain a model and predict the academic performance of the students who make use of Web-based tutoring systems.

The remainder of the paper is organized as follows: section 2 presents the background of the research, section 3 briefly describes the proposal to predict student performance on Web-based Tutoring Systems, section 4 describes the case of study and in section 5 the conclusions and future work are outlined.

2. Background

2.1. Fuzzy Prototypes and Fuzzy Deformable Prototypes

This work uses the concept of Fuzzy Deformable Prototypes (FDPs) [8]. The definition of FDPs inherits some features of the Zadeh's fuzzy prototype approach [9], but adding some extensions in order to manage the complexity of the real world problems. The principle to obtain a fuzzy prototype of a population is to stratify ς in grouping objects sharing the same membership degree (see Eq.1).

$$\varsigma = H/\varsigma_{good} + M/\varsigma_{border} + L/\varsigma_{poor} \qquad (1)$$

where ς_{good}, ς_{border} and ς_{poor} are multi sets of good, borderline and poor elements respectively and H, M and L are fuzzy numbers which represent the corresponding *high*, *medium* and *low* membership degrees respectively.

For each level of stratification of ς, this fuzzy prototype is obtained using an iterative process of compactification. During the iterative process, an object maximally summarized from each level of stratification is obtained which can be viewed as a fuzzy prototype.

Given a number of prototypes for a category, it may be meaningful to compute the collective property of the prototypes and to consider that as the reference for the category. Such use of fuzzy prototypes has been suggested for this purpose. In this case, the aim must be to generate conceptual prototypes (Zadehâs approach: fuzzy schemes) that allow us to evaluate new situations from these patterns. The definition of FDPs [8] includes the following extensions of Zadeh's fuzzy prototype approach: the number of fuzzy prototypes depends on the problem, categories are structured using typicality degrees and the shapes of the categories have not been defined. Moreover, FDPs can also be represented as fuzzy sets. It means that it is possible to calculate a membership degree between an element and the fuzzy set. The use of FDP's allows to deform [10] the most similar prototypes to a new situation $(w_1, w_2 \dots w_n)$, and define it using a linear combination with the membership degrees (μp_i) as coefficients (Eq. 2).

$$C_{real}\left(w_1 \dots w_n\right) = \left|\sum \mu_{p_i}\left(v_1 \dots v_n\right)\right| \qquad (2)$$

In other words, a fact or a set of facts is associated with a paradigm so that the paradigm interprets the situation.Thus, it is possible to make predictions according to this interpretation. To generalize, many of the predictions depend on the way the most similar paradigm or prototype for the circumstances of the problem is found.

2.2. Type-2 Fuzzy Sets

In the mid 60's, Zadeh [11] proposed the type-1 fuzzy sets as an alternative to handle the existing imprecision in human reasoning. However, it is insufficient to handle the very large, imprecise data sets with inherent uncertainties. Therefore, in 1975, Zadeh [12] proposed type-2 fuzzy sets as an extension of traditional type-1 fuzzy sets. According to [13], a type-2 fuzzy set is defined by a fuzzy membership function. Each degree of membership is a type-1 fuzzy set in the range $[0, 1]$ unlike type-1 fuzzy sets whose membership is a real value in the range [0,1]. In type-2 fuzzy sets there is a 3D membership function by a primary and secondary membership degree. Thus, a type-2 fuzzy set is defined by this membership function; where x is the primary variable, u is the secondary variable and J_x is the primary membership grade (see Eq. 3).

$$\tilde{A} = ((x, u), \mu(x, u)), \forall x \in X, \forall u \in J_x \subseteq [0.1] \qquad (3)$$

To make the secondary membership equal to the unit function, it gets a two dimensional membership function (see Fig. 1), which is limited by two type-1 membership functions: the so-called upper UMF (Upper Membership Function) and the lower called LMF (Lower Membership Function) (see Fig. 1). The area between the UMF and the LMF of a $\tilde{A}$ fuzzy set is called FOU (FootPrint Of Uncertain). The UMF$(\tilde{A})$ is defined by the upper limit of the FOU and is defined by $\bar{\mu}_A(x)$, $\forall x \in X$, and the LMF$(\tilde{A})$ is associated with the lower limit of the FOU and it is defined by $\underline{\mu}_A(x)$, $\forall x \in X$.

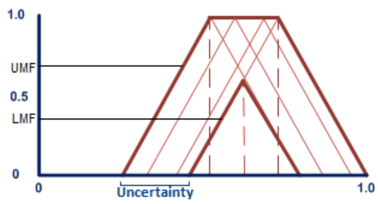

Figure 1: FootPrint Of Uncertain (FOU)

The FOU represents the uncertainty that exists in a type-2 fuzzy set, and it is defined as the union of all grades of primary membership (see Eq. 4).

$$FOU(\tilde{A}) = \bigcup_{\forall x \in X} J_x \qquad (4)$$

In [15] methodology for obtaining the UMF and the LMF of an FOU through the representation theorem proposed by [14] is defined.

(see Eq. 5).

$$\tilde{A} = \bigcup_{i=1}^{n} A^{(i)} \qquad (5)$$

where is the i-th T1FS embedded; $\bigcup$ it is the symbol of union.

3. Modeling Prototypes through Type-2 Fuzzy Sets

The behavior of a student in a Web-based Tutoring System can have different interpretations from different points of view. Therefore, it is not enough to apply a type-1 fuzzy model to make accurate predictions about the student performance. These uncertainty can be seen in the overlapped fuzzy representation of the prototypes. Therefore, T2FS can represent better the uncertainty involved in the learning process than conventional T1FS due the extra degree of freedom. In this proposal, first, the prototypes are discovered by the Fuzzy Prototypical Knowledge Discovery process and formally represented through type-1 fuzzy sets as explained in [8] (see Fig. 2) and then, the fuzzy partition of the prototypes to obtain type-2 fuzzy prototypes is carried out. This last process take place in the following manner:

- The overlapping level of the type-1 fuzzy set representation is computed, i.e, it is estimated the level of overlap of the prototypes represented through fuzzy numbers.
- The type-1 embedded prototype are established, i.e., type-1 fuzzy sets embedded within the fuzzy representation of type-2 are found.
- The UMF and the LMF of the IT2FS are calculated.

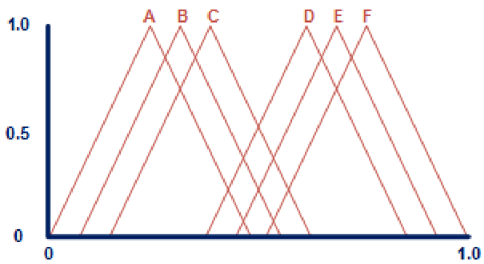

Figure 2: Fuzzy representation of the prototypes. Each fuzzy prototype is defined by a letter.

3.1. Level of overlap of type-1 fuzzy sets

The overlapping-level of the type-1 fuzzy prototypes were computed through Eq. 6:

$$S(A,B) = 1 - \frac{\sum_{i=1}^{p} |a_i - b_i|}{p} \quad (6)$$

where, A and B are triangular fuzzy numbers; a_i and b_i are the fuzzy prototypes parameters and p is the number of parameters of each fuzzy number. Table 1 shows the similarity that exists between each pair of prototypes.

Table 1: Reduced Similarity Matrix

	A	B	C	D	E	F
A	-	0.9	0.8	0.5	0.46	0.37
B	-	-	0.9	0.6	0.54	0.47
C	-	-	-	0.7	0.64	0.57
D	-	-	-	-	**0.94**	0.87
E	-	-	-	-	-	**0.94**
F	-	-	-	-	-	-

3.2. Determination of the type-1 embedded prototypes

A prototype represented through an IT2FS consists of the union of all its embedded prototypes. The following process is performed to determine the number of type-2 fuzzy prototypes, as well as the type-1 fuzzy prototypes embedded in them.

From the reduced similarity matrix (Table 1), the prototypes of greater similarity are determined. In the example, the prototypes (D,E), and (E,F) have a similarity of 0.94 in the range of 0 to 1. On the other hand, the prototypes E and F are selected and they are grouped to form the composite object (E,F) (Table 2). To calculate the similarity of the composite object (E,F) he maximum of the similarities with its components is chosen. For example, the similarity of the composite object (E,F) with D will be:

$$max[(E,D),(F,D)] = max(0.94, 0.87) = 0.94 \quad (7)$$

In the same way, each fuzzy prototype is computed. Table 3 shows that the greatest similarity is between (D) and (E,F) with 0.94, so they are grouped to form the composite object $(D,(E,F))$. In the same way, the maximum similarity of the object with each one of the prototypes in Table 3 is estimated (see Table 4).

Table 2: Composite Matrix I

	A	B	C	D	(E,F)
A	-	0.9	0.8	0.5	0.46
B	-	-	0.9	0.6	0.54
C	-	-	-	0.7	0.64
D	-	-	-	-	**0.94**
(E,F)	-	-	-	-	-

Table 3: Composite Matrix II

	A	B	C	(D,(E,F))
A	-	**0.9**	0.8	0.5
B	-	-	**0.9**	0.6
C	-	-	-	0.7
(D,(E,F))	-	-	-	-

Table 3 shows that the greater similarity between A,B and B,C is 0.9. Then, B and C are grouped forming the composite object (B,C) (Table 4. Then, the similarity of the compound object with the each of the elements of the array is computed.

Table 4: Composite Matrix III

	A	(B,C)	(D,(E,F))
A	-	**0.9**	0.5
(B,C)	-	-	0.6
(D,(E,F))	-	-	-

Table 5 shows the composite objects which represent the type-2 fuzzy prototypes. Each element of the composite object represents a type-1 fuzzy prototype embedded in it. For example, a type-2 fuzzy prototype results from the fusion of the prototypes A, B and C and the second prototype from the prototypes C, D and F. Finally, there is a similarity of 0.5 between both type-2 fuzzy prototypes, that is below the threshold set by the user.

Table 5: Final Matrix

	(A,B,C)	(D,E,F)
(A,B,C)	-	0.5
(D,E,F)	-	-

3.2.1. Definition of the UMF and the LMF of the IT2FS

As it is mentioned below, an IT2FS is delimited by two type-1 fuzzy sets; the UMF (Upper Membership Function) and the LMF (Lower Membership Function), which represent the high and low membership with the IT2FS. The area between UMF and LMF is the FOU (Footprint Of Uncertainty), which allows to model the existing imprecision. The Representation Theorem (RT) is used to represent and manipulate the overlap of the prototypes to the IT2FS.

The following steps, based on the methodology defined by [15], are used to obtain the UMF and the LMF of the IT2FS for representing the fuzzy prototypes:

- The support of the UMF_j (α_{UMF} is obtained through the minimum of the a_i ($\underline{a}$), and the maximum of the b_i ($\bar{b}$; where a_i and b_i represent the left and right ends of the bracket of the i-th T1FS embedded in the FOU_j; c_i is the centre of the i-th T1FS.

$$\alpha_{UMF_j} = (min(a_i), max(b_i)) \tag{8}$$

- The bottom ($\underline{c}$) and upper ($\bar{c}$) centroids of the UMF_j are computed. The bottom centroid is the minimum of c_i, and the upper centroid the maximum of the c_i.

$$\underline{c} = min(c_i)\bar{c} = max(c_i) \tag{9}$$

In this way, you get a UMF_j with trapezoidal-shaped and defined by the points $(\underline{a}, 0), (\underline{c}, 1), (\bar{c}, 1), (\bar{b}, 0)$.

- The support of the LMF_j (α_{LMF}) is computed through the maximum of the a_i ($\bar{a}$) and the minimum of the b_i ($\underline{b}$) (see Eq.10). The LMF is triangular in shape so that equations 11 and 12 are used to calculate its centre and maximum height.

$$\begin{aligned} \alpha_{LMF} &= (max(a_i), min(b_i)) \\ \alpha_{LMF} &= \bar{a}, \underline{b} \end{aligned} \tag{10}$$

$$p = \frac{\underline{b}(\bar{c} - \bar{a}) + \bar{a}(\underline{b} - \underline{c})}{(\bar{c} - \bar{a}) + (\underline{b} - \underline{c})} \tag{11}$$

$$\mu_p = \frac{\underline{b} - p}{\underline{b} - \underline{c}} \tag{12}$$

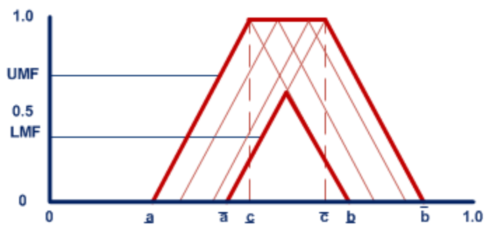

Figure 3: Fuzzy prototype represented through IT2FS

In order to evaluate a new situation which matches with the prototypes represented by IT2FSs, it is necessary to compute the affinity degree with each of the FOUs that define each prototype. Eq. 13 is used to calculate the actual membership of the situation with each FOU. The UMF and the LMF are type-1 membership functions. To obtain the LMF is necessary to calculate its height and its center according to the equations defined by [15] (see Eqs. 11 and 12).

$$\mu_{FOUj} = \begin{cases} UMF_i & \text{si } LMF_i = 0 \\ UMF_i - \frac{UMF_i - LMF_i}{2} & \text{si } LMF_i > 0 \end{cases} \tag{13}$$

where UMF_i and LMF_i correspond to the upper and lower membership with the i-th IT2FS; and μ_{FOU_j} is the actual membership with the j-th prototype. Finally, to describe the new situation using the prototypes of data (which were obtained by the FPKD), a deformation process based on the affinity degree of the new situation with the fuzzy prototypes is done. For this purpose, it is possible to use the technique previously defined in Eq. 2.

4. Application for predicting student performance

In this work, a cognitive model for predicting the academic performance of students is built using the FPKD Process. In the following subsections is shown a brief description of how this model is constructed based on data generated by a Web-based tutoring system.

4.1. Data Selection and Preparation

A collection of data generated by a web-based tutoring system containing all the interactions of the student with the system. The algebra 2008-2009 train collection provided by the 2010 KDD Cup competition[2] has been selected as example. It was generated by a tutoring system for the teaching of algebra called Carnegie Learning Algebra System, containing more than 8 million of records belonging to 3310 students from different educational institutions that used the tutoring system during the 2008-2009. The collection contains 23 characteristics such as: student ID, the name of the problem solved, the unit to which it belongs, etc. The characteristics that affect the academic performance of the student have been selected: the number of attempts to resolve each step, the response of the resolute step (1 if it is correct, 0 wrong), the hints requested as well as the skills used to solve a step.

In the preparation of the data, the arithmetic mean has been used to deal with the missing values. In addition, some attributes has been divided into others, such is the case of the attribute hierarchy unit.

4.1.1. Transformation

A set of tasks have been carried out to detect the patterns of students academic perfomance solving algebra problems:

- *Task 1 - Calculate the performance:* The central axis of all tutoring systems are the tasks (exercises, problems) that the student have to solve. Tasks are split up into a set of activities called steps. For the academic development of each student, it is necessary to analyse the way the student addresses the problem with regard to the sequence of steps followed. The performance is compared with students who solved the same problems to compute the difficulty, the speed, and whether the problem solved it correctly or not.
  - *Difficulty*: According to the evolution of the student, the increase of steps in a problem represents the difficulty adressed by the student to solve the problem. In this way, the difficulty is defined by Eq. 14:

$$difficulty = \frac{StepsCarriedOut - AverageSteps}{AverageSteps} \quad (14)$$

    where *StepsCarriedOut* refers to the number of activities in which the tutor system broken down the problem, *AverageSteps* refers to the average number of steps taken by the students who adresssed the same problem.
  - *Rate of correctness*: The rate of correctness of the problem is calculated through the equation 15:

$$correctnessRate = \frac{RC}{PR} - correctnessAverage \quad (15)$$

    where RC refers to the correctly answered steps n a problem, PR refers to the number of steps in which the tutoring system decomposed the problem for each student.
  - *Speed Factor*: The speed to solve a problem is defined by Eq. 16.

$$speedFactor = \frac{TE - TP}{TP} \quad (16)$$

    where TE referred to the time spent by the student to solve the problem, TP refers to the average time spent by students who solve the same problem.
  - *Help Ratio*: The hints requested by the student were used to compute the help ratio defined by Eq. 17:

$$Rateofhelp = \frac{PS - PP}{PP} \quad (17)$$

    where PP refers to the average number of hints requested by students who solve the same problem, PS refers to the hints requested by the student.

- *Task 2 - Compute the academic performance of students per unit.* For each student, is/her performance is computed through the obtained average performance in each unit. The evolution of the student can be represented as a sigmoidal function (see Fig. 4). This representation allows to trace the academic evolution of students along of their academic course. This representation is divided into 3 parts: Low, Average and High. Each part is divided into 3 sectors: S1,S2 and S3. The sector S1 means students resolved problems correctly are minimized, in a sector S2 there is an increment in the number of problems solved correctly. Finally, in sector S3 there is a stabilization of problems correctly solved.

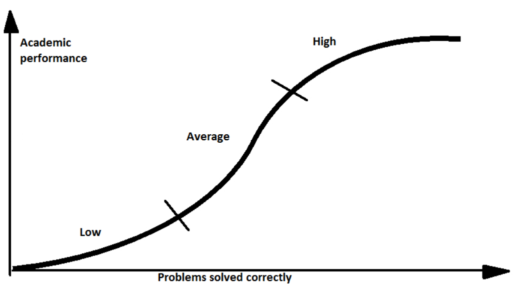

Figure 4: Representation of Academic Evolution

- *Task 3 - Calculate the academic students' performance*: The performance obtained in the completed units is computed through an aggregation function. In this way, each student is represented by a single record. A heuristic measure is defined to estimate the students' academic performance (see Eq. 18).

$$r_i = \frac{difficulty * correctnessrate}{speedFactor} \tag{18}$$

where *difficulty* is the difficulty that the student i faced to solve problems; *rateOfCorrectez* is the percent of correct answers obtained for him/her. *speedFactor* is the average of time used to solve the problems. This heuristic measure is used to classify students based on their expected academic performance levels. In this way, five performance levels are obtained (low, average, almost high, high and very high).

4.2. Fuzzy representation of prototypes

Table 6 shows the obtained prototypes of students' academic performace. Students with low academic performance solves with at least 2 attempts and time of 50 seconds, they often ask for hints. Their percentage of problems solved correctly is 35%. Students with high academic achievement solve problems correctly on the first attempt, in a time of 25 seconds with non-assistance of the tutoring system.

Table 6: Prototypes obtained from the Knowledge Discovery Process

	Prototypes				
Domain	*L*	*A*	A_H	*H*	V_H
Av. per. correct/ problem	35	62	75	84	96
Min. per. of correct./problem	19	56	72	82	92
Max. per. of correct/problem	50	67	78	87	100
attempts/problem	2	1	1	1	1
Time (seconds) /problems	50	41	33	30	25
hints / problem	1	0	0	0	0

The fuzzy representation of the prototypes was carried out through fuzzy numbers, which allow to obtain the affinity degree of a new student with each of them. For simplicity, triangular and trapezoidal membership functions are used. The following steps are carried out to represent the prototypes:

1. The centers of the prototypes are calculated by computing the arithmetic mean as centers of the prototypes that fall under each of them (Eq. 19).

$$C_{Pi} = \frac{\sum r_j}{n} \tag{19}$$

where is r_j the performance of the j-th sample belonging to the i-th prototype, n is the number of samples that fall under the i-th prototype.
2. The base of the fuzzy numbers is calculated. The fuzzy numbers that represent the prototypes are symmetrical, i.e., the base is divided into two equal segments. The standard distance from zero to the center of the first prototype is considered to calculate the size of the segment [8]. Fig. 5 shows the fuzzy representation of the prototypes of the Table 6.

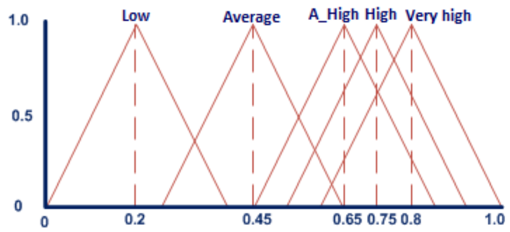

Figure 5: Fuzzy prototypes represented by type-1 fuzzy sets

As can be seen in Fig. 5, there is overlapping between the fuzzy representations of the prototypes. The overlap represents the similarity between the fuzzy prototypes. Therefore, it is necessary to represent the prototypes as type-2 fuzzy sets as it is mentioned below.

Table 7 shows the reduced similarity matrix which it is obtained from fuzzy representations of the prototypes. *L*, *A*, A_H, *H*, and V_H represent, low, average, almost high, high and very high, respectively.

Table 7: Reduced Similarity Matrix

	L	A	A_H	H	V_H
L	-	0.75	0.55	0.47	0.4
A	-	-	0.8	0.72	0.65
A_H	-	-	-	**0.92**	0.85
H	-	-	-	-	**0.92**
V_H	-	-	-	-	-

The type-1 fuzzy sets fusion process stops when the user-defined threshold is greater than the similarity values ($\alpha < 0.85$). The prototypes associated with fuzzy prototypes embedded in the IT2FS are aggregated in order to get only one that represent them. The methodology defined by [15] is used to establish the type-1 fuzzy prototypes embedded within the IT2FS. Table 8 and Fig. 6 show the resulting fuzzy prototypes, two represented through T1FS (*Low* and *average*), and one represented through IT2FS.The IT2FS fuzzy prototype embeds two type-1 prototypes (A_H, H, V_H).

Table 8: Final Similarity Matrix

	L	A	(A_H,(H,V_H))
L	-	0.75	0.55
A	-	-	0.8
(A_H,(H,V_H))	-	-	-

Table 9 shows the parametric definition of the three prototypes. For each prototype, a panel of experts defined three evolution sectors (S1, S2 y S3).

Table 9: Parametric Definition of the Prototypes

Domain	Prototypes								
	Low			Average			High		
	S1	S2	S3	S1	S2	S3	S1	S2	S3
Per. of correctness problem	8	25	42	55	63	71	79	87	96
Min. per. correct/problem	0	17	33	51	60	68	75	84	93
Max. per. correct/problem	16	32	50	59	67	74	83	92	100
attempts/problem	3	2	1	2	1	1	2	1	1
Time (seconds) /problems	106	50	22	100	41	20	100	30	18
hint / problem	2	1	0	1	0	0	1	0	0

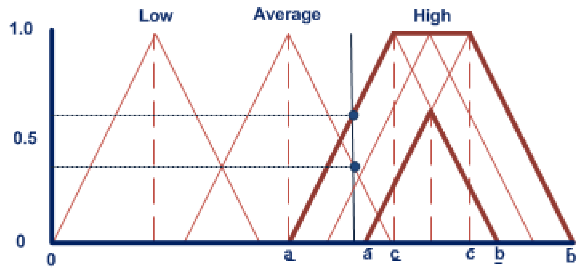

Figure 6: Fuzzy prototypes represented through IT2FS

4.3. Estimation of student academic performance

The following set of characteristics has been selected by a panel of experts as relevant to predict the students' academic performance:

- *Cognitive ability*: the ability of the student to solve problems (high/0.8, average/0.5, low/0.2)).
- *Interpretation*: the ability to understand the problem, it refers to the capacity of the student to interpret the problem (high/0.75, average/0.45, low/0.15).
- *Difficulty Level of the problem*: high/0.85, Media/0.45, low/0.15.

For example, when the student registers to the system for the first time, the teacher introduces the following values: cognitive ability: average/0.5, interpretation: high/0.75, complexity of the problem: average/0.45. In this way, a unique value to define this new student is obtained by applying an aggregation function over these values ($S = 0.56$). The student (S) has a positive affinity with the prototype *average* and the prototype *high* (see Fig. 6). The affinity degree with prototype *average* is obtained using the classic procedure defined in the literature, for this example the obtained value is 0.45 ($\mu_{average} = 0.45$). The affinity degree with prototype high (FOU_{high}) is computed through the affinities of the new case with the UMF (Upper Membership Function) and the LMF (Lower Membership Function). The UMF and the LMF are type-1 membership functions. To obtain the LMF is necessary to calculate its height and its center according to the equations defined by Liu-Mendel [15] (Eqs. 11 and 12). For the above example, $\mu_{UMF_{high}} = 0.55$ and $\mu_{LMF_{high}} = 0$.

The parametric definition of the new prototype (Table 10), i.e., the estimation of student's academic performance, is obtained by the linear combination (Eq. 2) of the parametric definition of the previously defined prototypes (see Table 9).

Table 10: New student estimated performance

Domain	S1	S2	S3
Per. of correctness average/ problem	68	76	85
Min. percentage of correctness/ problem	64	73	82
Max. percentage of correctness/ problem	72	81	88
Num. attempts/problem	2	1	1
Time (seconds) /problem	100	35	19
hints / problem	1	0	0

Table 11 shows the obtained prototypes after applying different aggregation operators on embedded prototypes in IT2FS. However, only resulting prototype obtained through *max* operator covers the complete domain so that it was used.

These results obtained using fuzzy prototypes represented by IT2FS are better than the obtained by the method on type-1 FDPs. Therefore, the IT2FS method represents the students' academic performance with more flexibility than the system based on T1FS.

5. Conclusions and Future Works

This paper introduces an extension of the Fuzzy Deformable Prototypes to improve the capacity to manage the imprecision of existing data. The proposed method aims to model and predict the academic performance of the students who use a Web-based tutoring systems. Interval Type 2 Fuzzy Sets were used to handle the imprecision of the academic data caused by the overlapping between the fuzzy representations of prototypes.

As future work, it will be developed a predictive model that allows predicting the evolution of the students' academic performance along its academic course.

Acknowledgments

This work has been funded by the Mexican Government through the PROgrama de MEjora al Profesorado (PROMEP/103.5/09/4362), and the Autonomous University of Carmen (UNACAR).

Table 11: Estimated mark with prototypes type 1 and prototypes type 2

	Type 1 prototype			Type 2 prototype			Real performance
x	$\mu_{average}$	μ_{high}	Estimated mark	$\mu_{average}$	μ_{high}	Estimated mark	
0.47	0.9	0	57	0.9	0.1	66	63
0.52	0.65	0	41	0.65	0.35	71	69
0.56	0.45	0	28	0.45	0.55	76	72
0.68	0	0.68	59	0	0.71	62	60
0.72	0	0.4	35	0	1	96	93
0.77	0	0.15	13	0	0.63	55	60
0.83	0	0.35	30	0	0.45	39	35
0.88	0	0.1	9	0	0.6	52	50
0.92	0	0.4	35	0	0.4	35	30
0.96	0	0.2	17	0	0.2	17	15

References

[1] C. Romero, P.. Espejo, A. Zafra, J.R. Romero, and S. Ventura. Web Usage Mining for Predicting Final Marks of Students That Use Moodle Courses. *Computer Applications in Engineering Education*, 21:135–146. 2013

[2] J. Stamper,A. Niculescu-Mizil, S. Ritter, G.J. Gordon, K.R. Koedinger. Algebra I 2008-2009. Challenge data set from KDD Cup 2010 Educational Data Mining Challenge. 2010. Find it at http://pslcdatashop.web.cmu.edu/KDDCup/downloads.jsp

[3] J. M. Mendel. An architecture for making judgments using computing with words. *International Journal of Applied Mathematics and Computer Science*, 12(3):325–335, 2002.

[4] M. Hao and J.M. Mendel. Perceptual computer application in learning outcome evaluation. In *Fuzzy Information Processing Society (NAFIPS), 2012 Annual Meeting of the North American*, pages 1–6, 2012.

[5] K. R. Koedinger and S. Mathan. Distinguishing qualitatively different kinds of learning using log files and learning curves. In *ITS 2004 Log Analysis Workshop. Maceio*, pages 39–46, 2004.

[6] J. E. Beck. Engagement tracing: Using response times to model student disengagement. In *International Conference on Artificial Intelligence and*, pages 88–95. IOS Press, 2005.

[7] I. Arroyo, H. Meheranian, and B.P. Woolf. Effort-based tutoring: An empirical approach to intelligent tutoring. In Ryan Shaun Joazeiro de Baker, Agathe Merceron, and Philip I. Pavlik Jr., editors, *EDM*, pages 1–10, 2010.

[8] J.A. Olivas. *Contribución a un estudio experimental de predicción basada en Categorías Deformables Borrosas*. Tesis de doctorado, Universidad de Castilla-La Mancha, 2000.

[9] L. A. Zadeh. A note on prototype set theory and fuzzy sets. *Cognition*, 12:291–297, 1982.

[10] H. Bremermann. Pattern recognition by deformable prototypes. In *Structural Stability, the Theory of Catastrophes, and Applications in the Sciences*, volume 525 of *Lecture Notes in Mathematics*, pp. 15–57. Springer, 1976.

[11] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8(3):338 – 353, 1965.

[12] L.A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning. *Information Sciences*, 8(3):199 – 249, 1975.

[13] N.N. Karnik and J.M. Mendel. Introduction to type-2 fuzzy logic systems. In *1998 Int. Fuzzy Systems Conference* , volume 2, pages 915–920, May 1998.

[14] J. M. Mendel and Robert I. Bob John. Type-2 fuzzy sets made simple. *Fuzzy Systems, IEEE Transactions on*, 10(2):117–127, Apr 2002.

[15] F. Liu and J. M. Mendel. An interval approach to fuzzistics for interval type-2 fuzzy sets. In *2007 IEEE Int. Fuzzy Systems Conference*, pages 1–6, 2007.