Monitoring students’ actions and using teachers’ expertise in implementing and evaluating the neural network-based fuzzy diagnostic model

Regina Stathacopoulou a,*, Maria Grigoriadou a, Maria Samarakou b,1, Denis Mitropoulos a

a Department of Informatics and Telecommunications, University of Athens, Panepistimiopolis, GR-15784 Athens, Greece
b Department of Energy Technology, Technological Education Institute of Athens, Ag. Spyridonos Str., GR-12210 Egaleo, Athens, Greece

Abstract

In this paper, the implementation of a neural network-based fuzzy modeling approach to assess aspects of students’ learning style in the discovery learning environment “Vectors in Physics and Mathematics” is presented. Fuzzy logic is used to provide a linguistic description of students’ behavior and learning characteristics, as they have been elicited from teachers, and to handle the inherent uncertainty associated with teachers’ subjective assessments. Neural networks are used to add learning and generalization abilities to the fuzzy model by encoding teachers’ experience through supervised neural-network learning. The neural network-based fuzzy diagnostic model is a general diagnostic model which is implemented in an Intelligent Learning Environment by eliciting teachers’ expertise regarding students’ characteristics based on real students’ observation and on data being collected from students’ interaction. The model has been successfully implemented, trained and tested in the learning environment “Vectors in Physics and Mathematics” by using the recommendations of a group of five experienced teachers. The performance of our model in real classroom conditions has been evaluated during an experiment with an experienced Physics teacher and 49 students of secondary school attending Physics lessons.

Keywords: Student model; Intelligent Learning Environments; Fuzzy logic; Neural networks; Learning styles

1. Introduction

An Intelligent Learning Environment (ILE) is a relatively new kind of Artificial Intelligence (AI) computer-based educational system, which is able to support student-driven learning and knowledge acquisition (Brusilovski, 1994). ILEs are considered to be generalizations of traditional Intelligent Tutoring systems (ITS), which are based on objectivist epistemology and embrace instructional environments which make use of theories of constructivism and situated cognition (Akhras & Self, 2002). Student models are distinguishing features of both Intelligent Tutoring Systems (ITS) (VanLehn, 1988; Wenger, 1987) and Intelligent Learning Environments (ILE) (Akhras & Self, 2002; Brusilovski, 1994). A student model enables the system to adapt its behavior and pedagogical decisions to the individual student who uses it (Brusilovski, 1994; Wenger, 1987).

Ideally, a student model should include all aspects of students’ behavior and knowledge which have repercussions on their performance and learning (Wenger, 1987). In practice, the contents of a student model depend on the application. It normally includes learner goals and plans, capabilities, attitudes and/or knowledge or beliefs, and is used as a tool for adapting ILE behavior to the individual student (Brusilovski, 1994; Holt, Dubs, Jones, & Greer, 1991; Self, 1999; VanLehn, 1988). Inferring a student model is called diagnosis, because it is much like the medical task of inferring a hidden physiological state.
from observable signs (VanLehn, 1988), i.e. the ILE uncovers a hidden cognitive state (student characteristics) from observable behavior.

The term student behavior can be used to refer to a student’s observable response to a particular stimulus in a given domain which, together with the stimulus, serves as the primary input to the student modeling system (Sison & Shimura, 1998). The input can be an action or the result of that action, and can also include intermediate results (Sison & Shimura, 1998). From this input, the diagnosis unit must infer a student’s unobservable behavior (VanLehn, 1988). Clearly, the less information the unit has the harder its task is (VanLehn, 1988). This makes student modeling a difficult process, given that the evidence about a student’s behavior provided by the student’s inputs to an ILE is usually scanty (Self, 1991), and contains a good deal of uncertainty (Jameson, 1996). A variety of AI techniques have been proposed for this purpose.

Bayesian networks have been proposed in ANDES (Conati, Gertner, & VanLehn, 2002; VanLehn & Niu, 2001), in order to relate—in a probabilistic way—a student’s observable behavior to a particular piece of his/her knowledge and in the Prime Climb educational game (Conati, 2002), in order to relate students’ observable behavior to their emotional state. Unsupervised machine learning techniques have also been proposed (Sison, Masayuki, & Shimura, 1998), in order to discover classes of errors, which represent misconceptions and other knowledge errors, from discrepancies in students’ behavior.

Another approach to handle the inherent uncertainty in a student’s behavior and to achieve a human description of knowledge is to use fuzzy logic. Fuzzy logic techniques have been proposed in a variety of user and student modeling approaches (Jameson, 1996). If a user modeling system adopts this approach, its reasoning may be particularly easy for designers and users to understand and/or to modify (Jameson, 1996).

One of the first attempts in using fuzzy student modeling, which was revised some years later (Hawkes & Derry, 1996), has been proposed in TAPS by Hawkes, Derry, and Rundensteiner (1990). In this context, fuzzy logic has been proposed as a flexible and realistic method of easily capturing the way human tutors might evaluate a student and handle tutoring decisions which are not clear-cut. Towards this direction, several other attempts have been proposed in the literature to model student knowledge, mental states and progress as well as student cognitive abilities and personal characteristics. A comprehensive review can be found in Jameson (1996).

Neural networks have also been proposed in student modeling due to their abilities to learn from noisy or incomplete patterns of students’ behavior and generalize over similar examples (Beck & Woolf, 1998; Posey & Hawkes, 1996). This generalized knowledge can then be used to recognize unknown sequences. A problem which arises when trying to apply a neural network to modeling human behavior is knowledge representation. The black-box characteristics of neural networks cannot offer much help, since the weights learned are often difficult for humans to interpret. To alleviate this situation, a neural network approach in which each node and connection has symbolic meaning has been proposed in TAPS (Posey & Hawkes, 1996). The back-propagation algorithm has been used to modify weights which represent importance measures of attributes associated with student performance, in order to refine and expand incomplete expert knowledge. Hybrid rule-based approaches integrating symbolic rules with neurocomputing have been proposed for knowledge representation in Intelligent Tutoring Systems, in order to create improved representations (Hatzilygeroudis & Prentzas, 2004).

Along these lines, this paper presents a neuro-fuzzy synergism for student diagnosis, by monitoring students’ actions in an Intelligent Learning Environment and using teachers’ expertise in implementing, training, testing and evaluating the neural-network based fuzzy diagnostic model in assessing aspects of students’ learning style. The Intelligent Learning Environment consists of the learning environment “Vectors in Physics and Mathematics” (Grigoriadou, Mitropoulos, Samarakou, Solomonidou, & Stavridou, 1999a), and the neural network-based fuzzy diagnostic model (Stathacopoulou, Magoulas, Grigoriadou, & Samarakou, 2005), which is a general diagnostic model which can be used in any learning environment according to designers’ and teachers’ suggestions. The proposed approach allows handling uncertainty of student behavior, by expressing teachers’ qualitative knowledge in a clearly interpretable way with the use of fuzzy logic, while offering the possibility of adaptation to the learning environment and to teachers’ personal constructs in classifying and discriminating among students by employing a neural network implementation of the fuzzy diagnostic model.

The paper is organized as follows. Section 2 presents the Intelligent Learning Environment, giving details on student interaction with the learning environment and providing a brief description of fuzzy knowledge representation and neural network implementation of the neural network-based fuzzy diagnostic model. Section 3 presents how the close monitoring of students’ actions and the use of data being collected from students’ interaction, allow us to generate hypotheses and to classify students regarding their learning characteristics. Section 4 presents the aspects of student learning style diagnosed by our model, the implementation, training and testing of our model using teachers’ expertise, students’ observation and logfiles as well as the evaluation of our model in real classroom conditions. Lastly, conclusions are drawn and directions for future work are presented.

2. The Intelligent Learning Environment

The Intelligent Learning Environment consists of the learning environment “Vectors in Physics and Mathematics” (Grigoriadou et al., 1999a), and the neural network-based fuzzy diagnostic model (Stathacopoulou et al., 2005).
2.1. The learning environment

The learning environment “Vectors in Physics and Mathematics” (Grigoriadou et al., 1999a, 1999b) is a discovery learning environment which has been designed and developed according to the constructivist theory of learning. The learning environment encourages students to take active control of their learning process, to express and support their ideas, to make predictions and hypotheses and test them by conducting experiments (Vosniadou, Ioannides, Dimitrakopoulou, & Papademetriou, 2001). Moreover, the design is based on the situated approach to cognition, which emphasizes students’ engagement in meaningful and purposeful activities (Vosniadou, 1996). Within this framework, the learning environment supports students’ active involvement in authentic activities, which correspond to real world processes, allowing them to control their own learning procedure, while providing them with help and guidance, when this is necessary.

The learning environment “Vectors in Physics and Mathematics” (see Fig. 1) aims at helping teachers to instruct and students to construct the concepts of vectors in physics and mathematics in secondary schools. In order to design the software and to choose its thematic units, we have taken into account the conceptual difficulties (Solomonidou, Stavridou, & Christidis, 1997) that secondary education students develop while learning mathematical and physical entities represented by vectors. We have also taken into account their difficulties in conceptualizing various phenomena which correspond to physical entities and which could be the source of misconceptions and inert knowledge (Arons, 1990; Driver, Guesne, & Tiberghien, 1985).

The thematic units of the learning environment are: Position and Displacement; Motion; Forces and Equilibrium; Forces and Motion; Forces and Momentum. Each one of these units contains several scenarios, which refer to real-life situations. Examples of such scenarios are: “Going fishing”, “planning a journey”, “which ship moves faster?”, “travelling in the islands”, “playing golf”, “bodies in equilibrium” (see Fig. 2), “imaginary climbing”, “falling objects”, “away from the earth”, etc. Students carry out selected activities within these scenarios. The activities simulate phenomena and actions happening in the real world. Students can control and observe the progress of these phenomena, take measurements, change various parameters, examine “what if” scenarios and compare them to reality, express their own model by drawing the vectors which compose it and, finally, observe the behavior of their model—where this is attainable—by comparing it to the scientific model. The environment also includes a dictionary of useful terms and concepts.

In order to implement the neural network-based fuzzy diagnostic model, we use the scenario “bodies in equilibrium” (Fig. 2) of the unit “Forces and Equilibrium”. The environment resembles a simple mechanics laboratory. A table appears on the screen and several objects such as boxes, cords, a spring and a pulley are available for use by the students. Students can drag and drop these objects in order to compose an equilibrium experiment. Several tools are also available, which help students draw and manipulate vectors representing forces, carry out measurements, etc.

Within this scenario students have the opportunity to carry out different equilibrium experiments by selecting one or two from the available objects from the object box (see Fig. 2) and using them to compose a system in equilibrium. Students have the opportunity to carry out equilibrium experiments in two different ways: either (a) by using the 5 different worksheets, compose the proposed equilibrium experiments in each work sheet and carry out the equilibrium experiments following the suggestions of the worksheet or (b) by using the available objects and tools of the learning environment in all possible ways in order to compose as many equilibrium experiments as possible and to draw forces acting on each object in every
equilibrium experiment. For example, they can place a single box of 40 N weight or 20 N weight on the table or they can select a box and the spring or a rope and hang the box from the ceiling through them, or they can place a box of 40 N or 20 N on the table and then place another box on top of the first. They can also select different surfaces for the table (different static friction) in case of experiments with the pulley and a box. In this way, students can compose 16 different equilibrium experiments. When composing each equilibrium experiment, they have to decide on the kind (gravitational/contact) and the properties (magnitude and direction) of the forces acting upon each object and draw them according to their conceptions.

In Fig. 3 an example activity with an equilibrium experiment with a spring and a box on the ceiling is shown. Students draw the forces acting on the box according to their opinions. In this way they are allowed to give their own Newtonian model for the equilibrium of the box. They can then use the “test” button to observe the behavior of their model. For example, if the resultant force is not equal to zero, the box will move towards the direction of this force. Students can also click the “reality” button, in order to observe the scientific model, i.e. the representation of the correct forces acting on the box.

2.2. The neural network-based fuzzy diagnostic model

As mentioned in Section 1, we use the neural network-based fuzzy diagnostic model, analytically presented in Stathacopoulou et al. (2005), to encode teachers’ knowledge in assessing students’ learning characteristics in order to implement the diagnostic process updating the student
model. The neural network-based fuzzy diagnostic model is a general diagnostic model which can be used in any learning environment according to designers' and teachers' suggestions. Fuzzy logic is used to provide a linguistic description of students' behavior and learning characteristics, as they have been elicited from teachers, and to handle the inherent uncertainty associated with teachers' subjective assessments. A decision is made with a set of fuzzy systems, with an approach which is closer to the human decision making process, since decisions are made by combining fuzzy evidences, each one contributing to the final decision to some degree. In addition, through the mode of qualitative reasoning teachers' knowledge is represented in a way that can be interpreted by designers of Intelligent Learning Environments. Neural networks are used to add learning and generalization abilities to the fuzzy model, in order to encode teachers' intuitive assessments—available by means of examples—into the system.

2.2.1. Knowledge representation scheme

Depending on the type of learning environment, domain and instructional design, teachers provide the learning characteristics they use to discriminate among students for the purpose of adapting their educational strategies to students' individual differences. Teachers may provide several student characteristics (de Vincente & Pain, 2002; Derry & Potts, 1998) related to student knowledge, learning abilities, motivation, learning strategies and learning styles. The output of the neural network-based fuzzy diagnostic model updates the student model regarding L different learning characteristics C_1, C_2, ..., C_L, such as aspects of a student's learning style, learning abilities and motivation.

Depending on the learning characteristic teachers provide the types of evidences they use to discriminate among students (de Vincente & Pain, 2002; Derry & Potts, 1998). Teachers may provide several types of evidences B_1, B_2, ..., B_k such as, students' total time working on the scenario, students' not random mouse moves, students' mouse moves near the selected objects, the number of students' conceptual errors, the total time spent on task, the number of idle intervals etc. The set of names of the types of evidences B = {B_1, B_2, ..., B_k} describes linguistically the k aspects of a student's observable behavior which will serve as inputs to the diagnostic process. For each type of evidence B_i (i = 1, 2, ..., k), we define with the help of expert teachers, students' actions related to the type of evidence B_i (i = 1, 2, ..., k), from students' actions during the interaction with the learning environment. For each type of evidence B_i (i = 1, 2, ..., k), a measured numeric value x_i (i = 1, 2, ..., k) is calculated for a student, from student's actions related to the type of evidence B_i (i = 1, 2, ..., k), (for example, by calculating the total time working on the scenario or by counting idle intervals). Each measured numeric value x_i (i = 1, 2, ..., k) takes its values in a set of positive numbers U_i (i = 1, 2, ..., k).

Thus, we define the numeric input X = {x_1, ..., x_k} to the neural network-based fuzzy diagnostic model, where x_i ∈ U_i, U_i (i = 1, 2, ..., k) is the universe of discourse of the ith input; each x_i ∈ U_i, U_i ⊆ R^n (i = 1, 2, ..., k) represents the measured values of B_i.

The process consists of three stages: fuzzification, inference, and defuzzification (see Fig. 4).

2.2.1.1. Fuzzification stage (stage 1). This stage represents teachers' subjective linguistic description of a student's behavior B = {B_1, B_2, ..., B_k} (e.g. s/he draws a vector after a long time, s/he has answered enough questions in the pre-test). We use linguistic variables (Zadeh, 1989) to describe the types of evidence B_1, ..., B_k, of student behavior B={B_1, B_2, ..., B_n, ..., B_k}, which enable teachers to discriminate among different students' learning characteristics. Each variable B_i (i = 1, 2, ..., k) can take a different number of linguistic values f_i. The number f_i of the linguistic values of each linguistic variable B_i (i = 1, 2, ..., k) and their names V_1h, V_2h, ..., V_fh are defined by the developer with the help from teachers, and depend on the variable B_i (i = 1, 2, ..., k). The set

Fig. 4. Schematic of the diagnostic model.
$T(B_i) = \{ V_{i1}, V_{i2}, \ldots, V_{ir} \}$ is the term set of $B_i$ ($i = 1, 2, \ldots, k$). For example, let us consider the linguistic variable $B_1$ = “total time on the scenario”. The corresponding term set could be $T(\text{total time on the scenario}) = \{ V_{i1}, V_{i2}, V_{i3} \} = \{ \text{Short, Normal, Long} \}$ or $T(\text{total time on the scenario}) = \{ V_{i1}, V_{i2}, V_{i3}, V_{i4}, V_{i5} \} = \{ \text{Very Short, Short, Normal, Long, Very Long} \}$ including three ($f_i = 3$), or five ($f_i = 5$) linguistic values respectively, depending on the required resolution. The overall observable behavior is described by the set $T = \{ T(B_1), \ldots, T(B_i), \ldots, T(B_k) \}$ of all term sets and by the set $B$. Thus, the set $T = \{ T(B_1), \ldots, T(B_i), \ldots, T(B_k) \} = \{ V_{i1}, V_{i2}, \ldots, V_{ij}, \ldots, V_i, V_{i5}, \ldots, V_{ik}, V_{k2}, \ldots, V_{k8} \}$ is a set of $j_1 + \cdots + j_i + \cdots + j_k$ words or sentences which describes the ranges of the types of evidence $B_i$ ($i = 1, 2, \ldots, k$) of a student’s behavior $B$. At the fuzzification stage, the numeric input $X = \{ x_1, x_2, \ldots, x_k \}$, where $x_1 \in U_1, x_2 \in U_2, \ldots, x_k \in U_k$ and $U_i$ is the universe of discourse of the $i$th input element, $U_1, U_2, \ldots, U_k \subset \mathbb{R}^+$ is fuzzified and transformed into membership degrees to the linguistic values $V_{i1}, V_{i2}, \ldots, V_{ij}$ which describe a student’s behavior $B = \{ B_1, \ldots, B_i, \ldots, B_k \}$. Therefore, the overall student behavior can be considered as a matrix $Y = \{ Y_1, \ldots, Y_j, \ldots, Y_k \} = \{ (y_{i1}, y_{i2}, \ldots, y_{ij}), \ldots, (y_{i1}, y_{i2}, \ldots, y_{ij}), \ldots, (y_{i1}, y_{i2}, \ldots, y_{ij}) \}$ of numeric values in $[0, 1]$, which represent the degree of membership $y_{ij}$ of each element $x_i$ ($i = 1, 2, \ldots, k$) of an input pattern $X = \{ x_1, \ldots, x_m, \ldots, x_k \}$, into the term sets of $B_i$ ($i = 1, 2, \ldots, k$) which use linguistic values $V_{i1}, V_{i2}, \ldots, V_{ij}$.

2.2.1.2. Inference stage (stage 2). This stage represents teachers’ reasoning in categorizing students qualitatively according to their abilities and personal characteristics, such as attentive, rather slow, good, etc. In particular, an approximation of fuzzy IF–THEN rules is performed, which represents teachers’ reasoning in the qualitative assessment of students’ characteristics. For example, if a student’s total time on the scenario is large and the number of attempts to find the correct forces is large and the number of random mouse moves and clicking buttons is small, then the student is very interested in the scenario.

In our model, a qualitative description of a student’s characteristics $C_1, C_2, \ldots, C_L$ is performed by treating student characteristics as linguistic variables. Each linguistic variable $C_j$ ($j = 1, 2, \ldots, L$) can take a different number of linguistic values $m_j$. The set $T(C_j) = \{ C_{ji}, C_{j2}, \ldots, C_{jm} \}$ is the term set of $C_j$ ($j = 1, 2, \ldots, L$). The expert-teachers set the number $m_j$ of the linguistic values and their names $C_{ji}, C_{j2}, \ldots, C_{jm}$ for each characteristic $C_j$ ($j = 1, 2, \ldots, L$) according to their personal judgement. For example, if we treat the linguistic variable $C_1$ = “student interest” using five linguistic values ($m_1 = 5$), then the term set could be: $T(C_1) = T(\text{student interest}) = \{ C_{i1}, C_{i2}, C_{i3}, C_{i4}, C_{i5} \} = \{ \text{very bored, bored, neither interested neither bored, interested, very interested} \}$. In this way, a mode of qualitative reasoning, in which the preconditions and the consequents of the IF–THEN rules involve fuzzy variables (Zadeh, 1989), is used to provide an imprecise description of teachers’ reasoning:

IF $B_j$ is $V_{ij1}$ AND $B_2$ is $V_{i2j}$ AND $B_3$ is $V_{i3k}$ THEN $C_1$ is $C_{i1j}$ AND $C_2$ is $C_{i2j}$ AND $C_3$ is $C_{i3k}$

where $I_1 = 1, 2, \ldots, j_1$; $I_2 = 1, 2, \ldots, j_2$; $I_3 = 1, 2, \ldots, j_3$; $J_1 = 1, 2, \ldots, m_1$; $J_2 = 1, 2, \ldots, m_2$; $J_3 = 1, 2, \ldots, m_3$.

The inference stage, provides a fuzzy assessment $C_j = [c_{j1}, c_{j2}, \ldots, c_{jm}]$ of the characteristic $C_j$, $C_{1j}, \ldots, C_{Lj}$ by assessing membership degrees $c_{j1}, c_{j2}, \ldots, c_{jm}$ to the linguistic values $C_{ji}, C_{j2}, \ldots, C_{jm}$ of the linguistic variable $C_j$ ($j = 1, 2, \ldots, L$) that describe the characteristic $C_j$ ($j = 1, 2, \ldots, L$). We use fuzzy relations (Pedrycz, 1991) operated with the max-min composition operator in order to infer a fuzzy assessment $C_j = [c_{j1}, c_{j2}, \ldots, c_{jm}]$ from a fuzzy preconditional. These relations represent a human teacher’s estimation of the degree of association between a fuzzy preconditional and a fuzzy assessment $c_{j1}, c_{j2}, \ldots, c_{jm}$ of a particular learning characteristic $C_j$ ($j = 1, 2, \ldots, L$).

2.2.1.3. Defuzzification stage (stage 3). This stage represents teachers’ final decision in classifying a student in one of the predefined linguistic values $C_{ji}, C_{j2}, \ldots, C_{jm}$ of the characteristic $C_j$ ($j = 1, 2, \ldots, L$). This process is performed by weighing the fuzzy assessment $c_j$ ($j = 1, 2, \ldots, L$). The fuzzy assessments $c_j = [c_{j1}, c_{j2}, \ldots, c_{jm}]$ ($j = 1, 2, \ldots, L$) are defuzzified to non-fuzzy values, that is to say, to decisions on one of the linguistic values $C_{ji}, C_{j2}, \ldots, C_{jm}$ ($j = 1, 2, \ldots, L$) of the learning characteristic $C_j$ ($j = 1, 2, \ldots, L$).

2.2.2. Neural network implementation of the fuzzy model

The fuzzy model is implemented with a set of neural networks as described in Stathacopoulou et al. (2005). A student’s evaluation regarding each learning characteristic, $C_1, C_2, \ldots, C_L$ is assessed by processing the numerical input $X = \{ x_1, x_2, \ldots, x_k \}$, of a student’s behavior through a set of neural networks. The process consists of three stages: fuzzification, inference, and defuzzification (see Fig. 4).

2.2.2.1. Fuzzification stage (stage 1). In the first stage the numeric input $X = \{ x_1, x_2, \ldots, x_k \}$, is fuzzified with a set of $k$ fuzzifiers and transformed to membership degrees $Y = \{ Y_1, \ldots, Y_i, \ldots, Y_k \} = \{ (y_{i1}, y_{i2}, \ldots, y_{ij}), \ldots, (y_{i1}, y_{i2}, \ldots, y_{ij}), \ldots, (y_{i1}, y_{i2}, \ldots, y_{ij}) \}$ of the linguistic values $V_{i1}, V_{i2}, \ldots, V_{ij}$ which describe each type of evidence $B_i$ ($i = 1, 2, \ldots, k$). As it usually happens in practice, the form of a membership function depends on sample points describing experts’, in this case teachers’, opinions (Kasabov, 1996). We have adopted an approach which simplifies the implementation by approximating the membership functions $y_{ij}(x_i), y_{ij2}(x_i), \ldots, y_{ij}(x_i)\ x_i \in U_i$ of the linguistic
values $V_{i1}, V_{i2}, \ldots, V_{iM}$ using a library of regular shapes and therefore implement the fuzzifiers with a set of $k$ fixed weight neural networks which calculate such regular shapes (Lin & Lee, 1995). We have used sigmoid functions as membership functions $y_{if}(x) = \frac{1}{1 + e^{-x}}$, for the extreme linguistic values $V_{i1}$ and $V_{iM}$, and the pseudotrapezoidal function (composed of two sigmoid functions) as membership functions $y_{ij}(x)$ for the intermediate values $V_{i2}, \ldots, V_{iM-1}$. Since membership functions are subjective and generally context-dependent, i.e. teacher— and subject matter—dependent (Zadeh, 1972), a set $M = \{m_1, m_2, \ldots, m_i, \ldots, m_k\}$ of parameters which adjust the membership functions (Stathacopoulou et al., 2005) is defined, to allow a range of adaptations to teachers’ subjective linguistic description.

2.2.2.2. Inference stage (stage 2). In the second stage, the inference process provides a fuzzy assessment $c_1, c_2, \ldots, c_j, \ldots, c_L$, of a student’s learning characteristics $C_1, C_2, \ldots, C_j, \ldots, C_L$, by assessing membership degrees $c_{j1}, c_{j2}, \ldots, c_{jM}$ to the linguistic values $C_{j1}, C_{j2}, \ldots, C_{jM}$ which describe each $C_j$ ($j = 1, 2, \ldots, L$). To this end, a set of $L$ specialized fuzzy systems, where each system $j$ ($j = 1, 2, \ldots, L$) infers about a particular learning characteristic $C_j$ ($j = 1, 2, \ldots, L$) is used. A fuzzy system of this type is a neural network containing a precondition layer which combines linguistic values in order to form fuzzy preconditions and a neural network which implements the fuzzy relation operated with the max-min composition. The neural network which implements the fuzzy relation is adjusted to teachers’ reasoning, in case where teachers’ reasoning is available in the form of fuzzy IF–THEN, or is trained with a Hebbian-style learning approach (Dae-Sik & Hyung-Il, 1997), in case where teachers’ reasoning is available in the form of examples.

2.2.2.3. Defuzzification stage (stage 3). In the third stage, the fuzzy assessments $c_j = [c_{j1}, c_{j2}, \ldots, c_{jM}]$ ($j = 1, 2, \ldots, L$) are defuzzified to non-fuzzy values $C_{j1}, C_{j2}, \ldots, C_{jM}$ ($j = 1, 2, \ldots, L$) of the learning characteristic $C_j$ ($j = 1, 2, \ldots, L$), by using a defuzzifier from the ensemble of the $M$ defuzzifiers. Each defuzzifier has a different number of inputs, therefore it defuzzifies a different number of linguistic values $m_j$. Therefore, depending on the number of linguistic values $m_j$ ($m_j > 2$) of each learning characteristic $C_j$ ($j = 1, 2, \ldots, L$) an appropriate ($m_j - 1$) defuzzifier $M$ is used in order to assess a student’s learning characteristic. At the defuzzification stage a set of $M$ neural networks is used, which are trained with a modified backpropagation algorithm that uses variable stepsize (called BPVS) (Magoulas, Vrahatis, & Androulakis, 1997).

3. Monitoring students’ actions

Human tutors obtain diagnostic information from observing what students say and do, and how something is said and done, i.e. tone of voice, inflection, hesitancy, timing of students’ responses, etc. (Derry & Potts, 1998), whereas ILEs are handicapped in this regard, since the communication channel between the student and computer is very restricted (usually a keyboard and a mouse) (Wenger, 1987). Researchers are beginning to develop technology to accurately assess actions in less restricted domains, as it is reported in Kort and Reilly (2002), but such types of devices are not available in the Intelligent Learning Environments which are currently in use. However, some indirect information, which approximates students’ unobservable behavior, can be obtained (VanLehn, 1988). An appropriately designed interface (where it is possible, for example, to time each keystroke) can facilitate the process of collecting the best available information on what the student is doing, in order to make diagnosis computationally tractable (Conati et al., 2002). The amount of time between a student’s actions is one type of information, which can be used as input for diagnosis, since chronometric data have been used in psychology for years as a basis for deciding between potential models of human cognition (VanLehn, 1988). For example, ANDES (Conati et al., 2002) uses a mechanism to track a student’s attention, which employs a masking interface that allows to measure how long the student focuses on a particular solution step. The latency data help the ITS determine probabilistically when the student has constructed a self-explanation mentally. Furthermore and more recently, researchers have tried to extract useful information from teachers about the detection of students’ motivational state, by showing them only students’ pre-recorded screen interactions with the instructional system (de Vincente & Pain, 2002). Despite teachers’ original doubts, most of the participants made a considerable number of reasoned inferences about students’ motivational state, and at the end of the study they commented that the task was actually not so difficult and that there was quite a lot of information available to them in order to perform these inferences.

In the learning environment “Vectors in Physics and Mathematics” we allow for a close monitoring of students’ actions over time, where each response such as a keystroke, mouse move or drag is timed and recorded. In this way, the learning environment stores all available information on what a student is doing in a logfile, recording each student action with a time stamp. Typical examples of student actions include: selection of objects for experimentation, selection of available tools, mouse moves, mouse drags or clicks on tools or objects, mouse drags when he/she tries to draw a vector, details about the vectors (forces) that the student draws, i.e. magnitude, direction and kind, as well as the time the action has been performed.

Figs. 5 and 6 present a part of two different students’ logfiles, named “Neocleous” and “Vas”, in tabular form. Each entry of the record corresponds to a student’s action, with a time-stamp showing minutes and seconds having elapsed from the start of the scenario “Bodies in equilibrium”. Words in quotes refer to tools/buttons, and pairs of unquoted numbers refer to mouse cursor positions.
Entries in standard font refer to mouse moves or idle mouse states e.g. the entry (“test” 12 min, 3 s) denotes that the student moves the mouse over the “Test” button but does not click it; the entries (5610 4950 11 min, 48 s) and (5460 2685 11 min, 50 s) denote two subsequent mouse cursor positions on the screen. Entries in bold refer to particular mouse events, i.e. selecting/clicking buttons (e.g. the entry (“test” 12 min, 4 s) denotes that the student clicks the “Test” button), dragging objects (as for example when the student moves an object or he/she draws a vector). A typical example of a drawing action is shown in the first column of the logfile of the student named “Neo- cleous” illustrated in Fig. 5 under the entry (“create vector” 11 min, 41 s), which denotes that the student clicks the “Create Vector” button. The process involves mouse drag, starting from the mouse position 5850 5610 denoted in row (5850 5610 11 min, 42 s) and ending at the mouse position denoted in row (5910 6750 11 min, 45 s). Entries in brackets provide a short description of the student’s actions denoted in the entries before in brackets (e.g. [Creates a gravitational force on box 40 magnitude 40 and direction -90 (force A)]) is the result of the drawing action starting with the action (“create vector” 11 min, 41 s)).

This close monitoring of a student’s behavior allows us not only to observe students’ misconceptions regarding the forces acting on each object, but also—by comparing logs—to identify differences in students’ interaction with the learning environment such as the presence or absence of idle intervals, unnecessary selections of objects or tools, purposeless navigations or clicking buttons. For example, the student of Fig. 5 named “Neo-cleous”, by using the appropriate tools, draws the correct forces acting on the box, a gravitational and a contact with correct direction.

Afterwards, he uses the “Test” button in order to observe the behavior of his model, and corrects only the magnitude of the second force. The student of Fig. 6 named “Vas” does not draw the correct forces; instead he draws two horizontal gravitational forces opposite in direction and magnitude 25N and 27N. Then, he uses the “Test” button to observe the behavior of his model, and corrects only the magnitude of the second force to 25 N. In addition, we observe six idle intervals in the logfile of the student named “Neo-cleous” illustrated in Fig. 5, and we suppose that the student temporarily stops to think, whereas in the logfile of the student named “Vas” illustrated in Fig. 6 we observe a lot of use of the “test” and “reality” buttons, opening and closing the “attribute list” as well as unnecessary navigations over objects or tools, facts which make us suppose that the student does not understand how to compose the equilibrium experiment and searches a solution in the available tools and selections of the learning environment.

This close monitoring of a student’s behavior allows us to generate hypotheses, to draw inferences about students’ unobservable behavior or beliefs and to classify students regarding their learning characteristics by observing how a student interacts with the learning environment from the actions recorded in the logfiles. In addition, this close monitoring allows us to generate simulated students’ logfiles by duplicating and changing real students’ logfiles (Stathacopoulou, Grigoriadou, Samarakou, & Magoulas, 2004), which are resembling to real students’ logfiles and can be used for training and testing, as will be described in Section 4.
4. Using teachers' expertise and students' observation and logfiles in implementing the neural network-based fuzzy diagnostic model

In order to elicit teachers' knowledge mentioned in Section 2.2 and implement the neural network-based fuzzy diagnostic model in evaluating students' learning characteristics according to designers' and teachers' suggestions, a group of five expert teachers has been used: three of them were experienced in teaching physics in secondary education, one of them was an expert in didactics of physics, and the last one was an expert in the design of educational software. In collaboration with the group of expert teachers we have chosen an aspect of the surface/deep model. In collaboration with the group of expert teachers we have chosen an aspect of the surface/deep model.

Fig. 6. Part of a student's logfile named “Vas” composing an equilibrium experiment with a spring and a box of 20 N.
been used to implement, train, test and evaluate the neural network-based fuzzy diagnostic model in order to assess an aspect of the surface/deep approach of a student’s learning style.

4.1. The deep/surface approach to learning

There is a growing interest in using learning styles as a source for adaptation in educational hypermedia systems. A comprehensive review can be found in Papanikolaou and Grigoriadou (2004). In these systems, diagnosis of student learning styles is usually based on questionnaires submitted by the students and/or observed behavior choices. Diagnosing a student’s learning style by monitoring his/her interaction with the system is an open research question (Papanikolaou & Grigoriadou, 2004).

A variety of learning style categorizations have been proposed in the literature (Riding & Rayner, 1998; Schmeck, 1988). In collaboration with the group of expert teachers we have chosen, as mentioned before, an aspect of the surface/deep approach (Biggs, 1987a; Marton & Saljo, 1997) of a student’s learning style, to be assessed with the neuro-fuzzy diagnostic model. The deep approach to learning is associated with constructivist teaching, which emphasises that learners actively construct knowledge for themselves, whereas the surface approach is linked to the emphasis that learners actively construct knowledge for themselves, whereas the surface approach is linked to the content (as opposed to their intention to complete task requirement and treating the task as an external imposition) were suggested by our group of experts as fundamental characteristics of deep/surface learning style to be assessed. The two characteristics were labelled as “student tendency to learn by discovery in a deep or surface way” and assessed as the one and only characteristic $C$ ($L = 1$, therefore for simplicity $C_1 = C$) by the neural network-based fuzzy diagnostic model. A student is classified as surface or deep with respect to his/her interaction with the learning environment. The output of the diagnostic process was described with five linguistic values ($n = 5$) in the term set $T(C) = \{ C_1, C_2, \ldots, C_5 \}$ and the following names $T(C) = \{ Deep, Rather Deep, Average, Rather surface, surface \}$.

4.2. Implementing the neural network-based fuzzy diagnostic model

Investigating and selecting the appropriate measures of students’ observable behavior to serve as indicators of students’ learning style regarding each categorization is a critical issue in designing the diagnostic process of students’ learning style (Papanikolaou & Grigoriadou, 2004). We use teachers’ expertise, which is based on real students’ observations, in order to specify the types of evidence $B_i$ ($i = 1, 2, \ldots, k$) which enable them to discriminate between the most positive and the most negative characterization (Derry & Potts, 1998), and which will serve as inputs to the diagnostic process. In order to elicit teachers’ knowledge regarding the types of evidence $B_i$ ($i = 1, 2, \ldots, k$) they use to discriminate among students interacting with the learning environment in respect of their “tendency to learn by discovery in a deep or surface way”, we have conducted an experiment with 18 students with the assistance of the group of expert teachers. During the experiment the group of teachers was asked to observe 18 students interacting with the scenario “Bodies in equilibrium”. At the same time, students’ interactions were recorded in the logfiles.

In order to record as many students’ strategies as possible (Kordaki, 2003; Kordaki & Potari, 2002), students were asked to use the available objects and tools of the learning environment in all possible ways, in order to compose as many equilibrium experiments as possible, and to draw forces acting on each object in every equilibrium experiment.

The group of expert teachers was asked to reach consensus and to specify the types of evidences $k$ which enable them to discriminate among students classified with the most positive characterization ($C_1 = “deep”$) and the most negative characterization ($C_5 = “surface”$) regarding “student tendency to learn by discovery in a deep or surface way”. They were asked in other words to provide us with the names $B_i$ ($i = 1, 2, \ldots, k$) of the linguistic variables of a student’s observable behavior $B = \{ B_1, \ldots, B_k \}$ which will serve as inputs for diagnosis, as well as with the number of the linguistic values $f_i$ of each linguistic
variable $B_i$ ($i = 1, 2, \ldots, k$) and their names $V_{i1}, V_{i2}, \ldots, V_{it}$ which assign the term set $T(B_i) = \{V_{i1}, V_{i2}, \ldots, V_{it}\}$ and the set of term sets $T = \{T(B_1), \ldots, T(B_i), \ldots, T(B_k)\}$ that describes a student’s behavior. In addition, the group of expert teachers was asked to identify the actions related to each type of evidence $B_i$, from the actions in the logfiles mentioned in Section 3, which are used to calculate the numeric value $x_i$ ($i = 1, 2, \ldots, k$) of each type of evidence $B_i$.

The group of expert teachers specified the types of evidences by using studies in cognitive psychology and their personal experience derived from observing real students interacting with the learning environment, as well as the logfiles of students’ interaction during the experiment. For example, the number of times a student tested his/her ideas or compared his/her ideas with reality was taken into account as an indication of whether the student was using trial and error strategies. In addition, during the experiment, students exhibited different behavior regarding the use of the “Test” and “Reality” buttons before and after drawing a vector in order to compose an equilibrium experiment. Some students always used the “Test” and “Reality” buttons before drawing the first vector, after drawing each of the vectors which compose the equilibrium experiment, and after changing the properties (magnitude and direction), while some other students rarely used the “Test” and “Reality” buttons. Fig. 7 illustrates, in tabular form, the use of the “Test” and “Reality” buttons before and after drawing a vector and some more actions, selected from three students’ logfiles, named “Giger”, “Dim” and “Poulos”, who use the “Test” and “Reality” buttons differently before and after drawing a vector. The three students in Fig. 7 are composing an equilibrium experiment with a spring on the ceiling and a box (see Fig. 3). Fig. 7 illustrates vectors drawn on the box by the students, the use of the “Test” and “Reality” buttons before and after drawing a vector, deleting or changing the properties of a vector, and some other students’ actions. Symbols $A'$ and $B'$ denote the first and the second vector respectively the student draws on the box. Symbols G and C denote the kind of force represented by the vector (Gravitational or Contact). Symbols 1t, 2t, \ldots, nt denote the number of subsequent uses of the “Test” and “Reality” button. Symbol + between symbols 1t, 2t, \ldots, nt denotes that the student uses the “Test” or “Reality” button after a series of other actions or after an idle interval.

<table>
<thead>
<tr>
<th>Duration</th>
<th>Objects</th>
<th>before drawing vectors</th>
<th>Drawing vectors</th>
<th>after drawing vectors</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 min &amp; 25 sec</td>
<td>Giger</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B20</td>
<td>2t+1t</td>
<td>1t+1t=1t+1t</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>B', C, 91, 20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>change direction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1t+1t</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>B', G, -91</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>change direction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dim</td>
<td>Dim</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B40</td>
<td>1t</td>
<td>1t+1t=1t+1t</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>B', G, 90, 40</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>delete vector</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1t+1t</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>B', G, -91, 37</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>delete vector</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1t+1t</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>B', G, 90, 40</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poulos</td>
<td>Poulos</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B20</td>
<td>4t</td>
<td>1t</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>B', G, -180, 19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>delete vector</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2t+1t</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>B', C, -180, 27</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>change magnitude</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1t+1t</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1t+1t</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 7. Use of the “Test” and “Reality” buttons before and after drawing a vector recorded in three different students’ logfiles.
The group of expert teachers suggested three linguistic variables \( B_1, B_2, B_3 \) associated with students’ actions within the 16 different equilibrium experiments of the scenario “Bodies in equilibrium”, which describes a student’s behavior \( B = \{ B_1, B_2, B_3 \} \) to be used in the diagnosis of “student tendency to learn by discovery in a deep or surface way”. Students’ actions before trying to draw a vector or after making an incorrect attempt are taken into account in \( B_1 = \text{the number of times a student tests his/her ideas or compares his/her ideas with reality} \), described with three linguistic values \( (f_1 = 3) \) and by the term set \( T(B_1) = \{ V_{11}, V_{12}, \ldots, V_{1i} \} = \{ \text{Seldom, Sometimes, Frequently} \} \). A student’s actions, such as idle intervals recorded in the logfiles (described in Section 3, and illustrated in the logfile of the student named “Neoclaneous” in Fig. 5), during drawing vectors or after testing an incorrect idea, are taken into account in \( B_2 = \text{the number of times a student consults the dictionary or temporarily stops to think} \), described with three linguistic values \( (f_2 = 3) \) and by the term set \( T(B_2) = \{ V_{21}, V_{22}, \ldots, V_{2j} \} = \{ \text{Sometimes, Frequently, Always} \} \). In addition, the time needed to draw the correct forces is taken into account in the linguistic variable \( B_3 = \text{experiment carry out speed} \), described with three linguistic values \( (f_3 = 3) \) and by the term set \( T(B_3) = \{ V_{31}, V_{32}, \ldots, V_{3j} \} = \{ \text{Fast, Medium, Slow} \} \).

Using the real students’ logfiles obtained during the experiment we define, with the assistance from the group of expert teachers, the universe of discourses \( U_i (i = 1, 2, 3) \) for each measured input \( x_i (i = 1, 2, 3) \), representing the measured type of evidence \( B_i (i = 1, 2, 3) \), as well as the membership functions \( y_{1i}(x_i, m_i), y_{2i}(x_i, m_i), y_{3i}(x_i, m_i), x_i \in U_i, m_i \in U_i, U_i \subset \mathbb{R}^+ \) which associate the linguistic values \( V_{1i}, V_{2i}, V_{3i} (i = 1, 2, 3) \) of each linguistic variable \( B_i (i = 1, 2, 3) \) with the universe of discourse \( U_i (i = 1, 2, 3) \).

The universe of discourse \( U_1 \) of the linguistic variable \( B_1 = \text{the number of times a student tests his/her ideas or compares his/her ideas with reality} \), which is calculated from the number of times in each equilibrium experiment recorded in the logfiles, was set to \([0, 35]\). In the 18 students’ logfiles, a threshold value of two times was found, since all students used the “Test” and “Reality” buttons once, regardless of whether they composed a successful or unsuccessful equilibrium experiment. Thus, three and four times have a large membership degree in the linguistic value \( V_{11} = \text{“seldom”} \), and the interval \((\text{Biggs, 1987a, 1994})\) is the overlapping area of the membership functions \( y_{11}(x_1), x_1 \in [0, 35] \), and \( y_{12}(x_1), x_1 \in [0, 35] \), of the linguistic values \( V_{11} = \text{“seldom”} \) and \( V_{12} = \text{“sometimes”} \) respectively. Fig. 8 illustrates the membership functions \( y_{11}(x_1), y_{12}(x_1), y_{13}(x_1), x_1 \in [0, 35] \), of the three linguistic values \( V_{11} = \text{“seldom”} \), \( V_{12} = \text{“sometimes”} \) and \( V_{13} = \text{“frequently”} \) in their universe of discourse \( U_1 = [0, 35] \).

Students’ action temporarily stops in order to think, which is used in the calculations of \( x_2 \), is measured from a student’s idle intervals when composing an equilibrium experiment, by counting therefore idle intervals before and after: drawing a vector, the use of “Test” or “Reality” button, the use of tools that manipulate vectors, the use of dictionary of useful terms and concepts and the use of the button “Clear Screen” in order to compose the same equi-

![Fig. 8. Membership functions for the three linguistic values of the linguistic variable \( B_1 = \text{“the number of times a student tests his/her ideas or compares his/her ideas with reality”} \).](image-url)
librium experiment again. In calculating the measured value \( x_2 \), idle intervals greater than 7 s have been used. In the universe of discourse \( U_2 \) of \( B_2 = \text{the number of times a student consults the dictionary or temporarily stops to think} \), the number of times per time interval has been used, since the number of times depends on the total time a student has used the learning environment. We have used a time interval of 300 s (5 min), since from students’ logfiles it was found that this is approximately the minimum amount of time needed by a student to compose an equilibrium experiment.

The universe of discourse is defined, by processing the logfiles, in the interval \([0, 35]\). Fig. 9 illustrates the membership functions in their universe of discourse \( U_2 = [0, 35] \). Fig. 9 illustrates the membership functions \( y_21(x_2) \), \( y_22(x_2) \), \( y_23(x_2) \), \( x_2 \in [0, 20] \), of the three linguistic values \( V_21 = \text{sometimes} \), \( V_22 = \text{frequently} \) and \( V_23 = \text{always} \) in their universe of discourse \( U_2 = [0, 20] \).

The linguistic variable \( B_1 = \text{experiment carry out speed} \) is determined by computing the average percentage of time needed to draw the correct forces of each equilibrium experiment (Grigoriadou et al., 1999a). The time needed to draw the forces applied to an object was compared against the time that the group of experts teachers defined as the average time multiplied by two; thus, the universe of discourse was set to \([0, 100]\). In case a student has composed more than one equilibrium experiments, the measured value \( x_3 \) is calculated as the mean value of all equilibrium experiments. Fig. 10 illustrates the membership functions in their universe of discourse \( U_3 = [0, 100] \).

In addition, the group of experts also suggested taking into account students’ prior experience with the interface of the learning environment, given that after observing students interact with the learning environment, it was realised that the time a student needs to draw the correct forces may also include the time needed to use the available tools that manipulate vectors and draw forces. Thus, the association between the universe of discourse \( U_3 \) of the linguistic variable \( B_3 \) and the membership degrees \( y_31, y_32, y_33 \) of the linguistic values \( V_31, V_32, V_33 \) can be adjusted for students who have prior experience in using the software, by using the adjusting parameter \( m_3 \), \( m_3 \in [0, 100] \), which represents the expected mean value of \( x_3 \), which moves the center of membership functions \( y_31(x_3, m_3), y_32(x_3, m_3), y_33(x_3, m_3) \) (continuous lines) and the adjusted membership functions for students with prior experience in using the software \( y_31(x_3, m'_3), y_32(x_3, m'_3), y_33(x_3, m'_3) \) (dotted lines) of the three linguistic values \( \{V_31\} = \text{fast} \), \( V_32 = \text{medium} \), \( V_33 = \text{slow} \) of the linguistic variable \( B_3 = \text{experiment carry out speed} \).

4.3. Training and testing

The neural network-based fuzzy diagnostic model has been trained and tested off-line using a set of simulated students’ logfiles, which is generated using the set of real students’ logfiles (Stathacopoulou et al., 2004) provided during the experiment which was carried out with the assistance of the group of experts teachers. In order to train the neural network-based fuzzy diagnostic model, a set of simulated students with predefined membership degrees \( \{y_{11}, y_{12}, y_{13}\}, \{y_{21}, y_{22}, y_{23}\}, \{y_{31}, y_{32}, y_{33}\} \) to the linguistic values \( \{V_{11}, V_{12}, V_{13}\}, \{V_{21}, V_{22}, V_{23}\}, \{V_{31}, V_{32}, V_{33}\} \) of the linguistic variables \( B_1, B_2, B_3 \) of their observable behavior.

![Fig. 9. Membership functions for the three linguistic values of the linguistic variable B2 = the number of times a student consults the dictionary or temporarily stops to think.](image-url)
B was generated. In this way the set included all combina-
tions of the linguistic values \{V_1, V_{12}, V_{13}\}, \{V_{21}, V_{22}, V_{23}\}, \{V_{31}, V_{32}, V_{33}\}, of the linguistic variables \(B_1, B_2, B_3\), representing the preconditions at the inference stage. The group of expert teachers classified the set of simulated students with respect to their “tendency to learn in a deep or surface way” in one of the linguistic values of the term set \(T(C) = \{C_1, C_2, C_3, C_4, C_5\} = \{\text{Deep}, \text{Rather Deep}, \text{Average}, \text{Rather Surface}, \text{Surface}\}\), to be used in training the neural network-based fuzzy diagnostic model described in Section 2.2.

In order to test the performance of the neural network-
based fuzzy diagnostic model, three test sets have been gen-
erated with predefined linguistic values \{V_1, V_{12}, V_{13}\}, \{V_{21}, V_{22}, V_{23}\}, \{V_{31}, V_{32}, V_{33}\}, in the linguistic variables \(B_1, B_2, B_3\) of their observable behavior \(B\), and predefined membership degrees \{y_{11}, y_{12}, y_{13}\}, \{y_{21}, y_{22}, y_{23}\}, \{y_{31}, y_{32}, y_{33}\}, to these values as well. The first set contains patterns with clear-cut descriptions of students’ observable behavior \(B = \{B_1, B_2, B_3\}\), i.e. their membership degrees \{y_{11}, y_{12}, y_{13}\} \((i = 1, 2, 3)\) in the linguistic values \(V_{1i}, V_{2i}, V_{3i}\) \((i = 1, 2, 3)\) of each linguistic variable \(B_i\) \((i = 1, 2, 3)\) are close to 1. The second set involves a lot of uncertainty; there are no clear-cut cases due to lack of well-defined boundaries in evaluating students’ observable behavior. This set includes marginal cases, i.e. patterns that contain membership degrees \{y_{1i}, y_{2i}, y_{3i}\} \((i = 1, 2, 3)\) close to 0.5 in two linguistic variables \(V_{1i}, V_{2i}, V_{3i}\) \((i = 1, 2, 3)\) of one or more than one linguistic variable \(B_i\) \((i = 1, 2, 3)\). This data set was used to test the capability of the model to handle the uncertainty incorporated in the marginal cases of students’ observable behavior, produced by lack of well-defined boundaries in associating the universe of discourse \(U_i\) of the measured value \(x_i (i = 1, 2, 3) x_i \in U_i\), \(U_i \subset \mathbb{R}\) of the linguistic variable \(B_i (i = 1, 2, 3)\) with the linguistic values \(V_{1i}, V_{2i}, V_{3i}\) \((i = 1, 2, 3)\). This capability is usually not supported in a non-fuzzy rule-based environment.

The third set consists of special marginal cases, which are possible to cause conflicting judgments, if they are processed by means of classic IF–THEN rules. A typical example is when two IF–THEN rules with close preconditions categorize the student into two different non-adjointing categories, as will be described below. Fig. 11 illustrates the special marginal cases in assessing student tendency to learn by discovery in a deep or surface way from a student’s observable behavior.

At this point it is useful to illustrate the behavior of our model in the special marginal cases with an example. Let us consider the two close preconditions of student behavior, presented in Fig. 11 as marginal case number 7. This is the case of two students, a student who “frequently” (V_{13}) tests his/her ideas or compares his/her ideas with reality (linguistic variable \(B_3\)), who “sometimes” (V_{21}) consults the dictionary or temporarily stops after testing an incorrect idea in order to think (\(B_2\)), and has “slow” (V_{33}) experiment carry out speed (\(B_3\)), and another student who “sometimes” (V_{12}) tests or compares ideas with reality (\(B_1\)), and “frequently” (V_{22}) consults or thinks (\(B_2\)), and has “medium” (V_{32}) experiment carry out speed (\(B_3\)). The first student’s tendency to learn by discovery in a deep or surface way is classified as “Surface”, if evaluated with IF–THEN rules, and the second’s as “Average” respectively. In our model, when the membership degrees \(y_{13}, y_{21}, y_{33}\), and \(y_{12}, y_{22}, y_{32}\), to the above linguistic values \(V_{13}, V_{21}, V_{33}\), and

![Fig. 10. Membership functions for the three linguistic values of the linguistic variable \(B = \text{"experiment carry out speed"}\).](image)
mulate the input values
fuzzy diagnostic model. The patterns of these data sets for-
face $(V_{12}, V_{22}, V_{32})$ of a student’s observable behavior
$B = \{B_1, B_2, B_3\}$ are equal to 1, these two evaluation
decisions provide the following fuzzy assessments vectors at the
output of the inference stage: $[c_1, c_2, c_3, c_4, c_5] = [0, 0, 0.05,
0.05, 0.9]$ and $[c_1, c_2, c_3, c_4, c_5] = [0, 0.05, 0.9, 0.05, 0]$ for
the first and the second case respectively. Finally, after
defuzzification, the students are classified into two quite
different non-adjoining categories $C_5 = “surface”$ and
$C_3 = “average”$.

Let us now consider a special marginal case where student
observable behavior $B = \{B_1, B_2, B_3\}$ causes two rules
to fire. This may be the case of a student who tests his/her ideas or compares his/her ideas with reality ($B_1$) “frequently”
($V_{13}$) with a membership degree $y_{13} = 0.4$, and “sometimes”
($V_{12}$) with a membership degree $y_{12} = 0.59$. The stu-
dent also “sometimes” ($V_{21}$), with a membership degree
$y_{21} = 0.4$ and “frequently” ($V_{22}$) with a degree $y_{22} = 0.59$
consults the dictionary or temporality stops to think ($B_2$).
The same student also has experiment carry out speed
($B_3$) that is “slow” ($V_{33}$) with a membership degree
$y_{33} = 0.4$ and “medium” ($V_{32}$) with a degree $y_{32} = 0.59$.
This complicated case will provide the following fuzzy
assessment vector at the end of the inference stage: $[c_1,
c_2, c_3, c_4, c_5] = [0, 0.05, 0.59, 0.4, 0.4]$. The final decision
at the output of the defuzifier is that this student’s tendency
to learn by discovery in a deep or surface way is
“Rather Surface”, a decision between the two categories
which is indeed consistent with the group of expert teachers’
evaluations when classifying similar marginal cases
of real or simulated students. Fig. 11 illustrates all reported
cases.

The three test data sets of simulated students have been
classified by the group of expert teachers according to the
term set \{Deep, Rather Deep, Average, Rather Surface, Surface\}.
The classifications of the three data sets of simulated stu-
dents by the group of experts were compared against clas-
sifications by the neuro-fuzzy diagnostic model. As
analytically presented in Stathacopoulou et al. (2005),
the overall average success in diagnosis reached 94%, i.e.
100%, 96%, 86% for each of the three data sets respectively.
The average success in diagnosis for the first test set
reached 100%. The model also provided an excellent aver-
age performance, 96%, in evaluating marginal cases (sec-
test set), in accordance to the group of expert
teachers’ assessments. In the third test set (special marginal
cases) an average performance of 86% was achieved,
by fine-tuning the rules.

\subsection*{4.4. Evaluating the neural network-based fuzzy diagnostic model in real classroom conditions}

The neural network-based fuzzy diagnostic model, has
been implemented trained and tested off-line in accordance
to the group of expert teachers’ expertise as described in
Sections 4.2 and 4.3, and attached to the learning environ-
ment “Vectors in Physics and Mathematics”. In order to
evaluate the performance of our model in classifying stu-
dents in real classroom conditions we have conducted an
experiment with an experienced Physics teacher and 49 stu-
dents of secondary school attending Physics lessons. Stu-
dents’ classification by the teacher regarding “student tendency to learn by discovery in a deep or surface way” in
the term set \{Deep, Rather Deep, Average, Rather Surface, Surface\} was compared with students’ classification by the
neural network-based fuzzy diagnostic model. The overall
success in diagnosis has been 55%. We have also analyzed
percentage of success of our model for each linguistic
value. The percentage of success for the linguistic values
Surface, Rather Surface, Average, Rather Deep, was 44 %,
67%, 63%, 50%, respectively. Percentage of success for
the linguistic value deep was not calculated, since the tea-
cher did not classify any students to the linguistic value

<table>
<thead>
<tr>
<th>Special marginal cases</th>
<th>Linguistic description of student's behavior in the first case</th>
<th>Classification of the first case</th>
<th>Linguistic description of student's behavior in the second case</th>
<th>Classification of the second case</th>
<th>Classification (C) of the special marginal case</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$V_{11}$, $V_{21}$, $V_{31}$</td>
<td>$C_2$</td>
<td>$V_{22}$, $V_{21}$, $V_{31}$</td>
<td>$C_4$</td>
<td>$C_3$</td>
</tr>
<tr>
<td>2</td>
<td>$V_{11}$, $V_{22}$, $V_{31}$</td>
<td>$C_2$</td>
<td>$V_{22}$, $V_{21}$, $V_{32}$</td>
<td>$C_4$</td>
<td>$C_3$</td>
</tr>
<tr>
<td>3</td>
<td>$V_{11}$, $V_{21}$, $V_{33}$</td>
<td>$C_2$</td>
<td>$V_{22}$, $V_{21}$, $V_{33}$</td>
<td>$C_4$</td>
<td>$C_3$</td>
</tr>
<tr>
<td>4</td>
<td>$V_{11}$, $V_{22}$, $V_{31}$</td>
<td>$C_1$</td>
<td>$V_{12}$, $V_{21}$, $V_{31}$</td>
<td>$C_5$</td>
<td>$C_2$</td>
</tr>
<tr>
<td>5</td>
<td>$V_{11}$, $V_{22}$, $V_{31}$</td>
<td>$C_1$</td>
<td>$V_{12}$, $V_{22}$, $V_{31}$</td>
<td>$C_5$</td>
<td>$C_2$</td>
</tr>
<tr>
<td>6</td>
<td>$V_{11}$, $V_{22}$, $V_{31}$</td>
<td>$C_1$</td>
<td>$V_{12}$, $V_{22}$, $V_{32}$</td>
<td>$C_5$</td>
<td>$C_2$</td>
</tr>
<tr>
<td>7</td>
<td>$V_{11}$, $V_{22}$, $V_{31}$</td>
<td>$C_1$</td>
<td>$V_{12}$, $V_{22}$, $V_{33}$</td>
<td>$C_5$</td>
<td>$C_2$</td>
</tr>
<tr>
<td>8</td>
<td>$V_{11}$, $V_{22}$, $V_{31}$</td>
<td>$C_1$</td>
<td>$V_{12}$, $V_{22}$, $V_{33}$</td>
<td>$C_5$</td>
<td>$C_2$</td>
</tr>
<tr>
<td>9</td>
<td>$V_{11}$, $V_{22}$, $V_{31}$</td>
<td>$C_1$</td>
<td>$V_{12}$, $V_{22}$, $V_{33}$</td>
<td>$C_5$</td>
<td>$C_2$</td>
</tr>
</tbody>
</table>

Fig. 11. Special marginal cases reported during training and testing.
deep. Fig. 12 illustrates the classification success of the neural network-based fuzzy diagnostic model for each linguistic value. As we can see in Fig. 12 the best classification success (67%) has been performed in the linguistic value Rather Surface.

As described in Section 4.3 the overall success in diagnosis of our model when training and testing with a set of simulated students classified by the group of expert teachers has been 94%, while as already mentioned the overall success in diagnosis in real classroom conditions has been 55%. In order to explain the above deviation we have further analyzed the 49 students’ logfiles recorded during interaction. We divided students’ logfiles in two subsets: subset 1 that includes students, which at the end of their interaction with the learning environment succeeded to draw the correct forces in most equilibrium experiments they performed (34 students) and subset 2 that includes students which at the end of their interaction with the learning environment did not succeed to draw the correct forces in most equilibrium experiment they performed (15 students). The overall success in diagnosis has been 68% and 27% for subset 1 and subset 2 respectively. We have also analyzed percentage of success of our model for each linguistic value for the two subsets. Fig. 13 illustrates comparative results of classification success in each linguistic value for the two subsets of students’ logfiles. As we can see in Fig. 13 our model has a better performance 58%, 78%, 75%, 60% for the linguistic values Surface, Rather Surface, Average, Rather Deep, respectively when classifying students which at the end of their interaction with the learning environment succeeded to draw the correct forces (subset 1) in comparison with the performance 17%, 33%, 33%, 33% when classifying students which at the end of their interaction with the learning environment did not succeed to draw the correct forces (subset 2). The above deviation in classification success of our model for the two subsets of students’ logfiles is explained from the fact that our model was trained with simulated students’ logfiles which at the end of their interaction with the learning environment succeeded to draw the correct forces in most equilibrium experiments they performed.

Furthermore, we have also analyzed the types of classification errors performed by the neural network-based fuzzy diagnostic model for all students (49 students) as well as for the two subsets of students, subset 1 (34 students) and subset 2 (15 students). The type of error is particularly important, as the outcome of the diagnosis has an impact on the pedagogical strategy adopted for each student. In classifications performed by the neural network-based fuzzy diagnostic model in real classroom conditions regarding “student tendency to learn by discovery in a deep or surface way” we have identified three types of errors. Type 1 error happens, when a student has been incorrectly classified in an adjoining category, i.e. with rank difference of one in the linguistic values \{C_1, C_2, C_3, C_4, C_5\}. For example, this type of error occurs, when a student is classified by the teacher as \(C_4 = \text{“rather surface”}\) (regarding his/her tendency to learn by discovery in a deep or surface way), but our model classifies him/her in the category \(C_5 = \text{“surface”}\) or \(C_3 = \text{“average”}\). On the other hand, when the same student is classified by our model as \(C_3 = \text{“rather deep”}\) (rank difference of three), a type 3 error occurs. When the same student is classified by our model as \(C_1 = \text{“deep”}\) (rank difference of two), a type 2 error occurs. The above error occurs, when the student is classified by the teacher as \(C_4 = \text{“rather surface”}\) (regarding his/her tendency to learn by discovery in a deep or surface way), but our model classifies him/her in the category \(C_5 = \text{“surface”}\) or \(C_3 = \text{“average”}\). On the other hand, when the same student is classified by our model as \(C_3 = \text{“rather deep”}\) (rank difference of three), a type 3 error occurs. Fig. 14 illustrates the overall classification success and types of errors for all students and the two subsets. As shown in Fig. 14 our model has a better performance in overall classification success (68%) for subset 1 in comparison with overall classification success 55% and 27% for all students and for subset 2 respectively. In addition our model has a better performance in types of errors when classifying subset 1 with type 1 error 24%, type 2 error 9% and type 3 error 0%, in comparison with 20%, 12%, 12% and 13%, 20%, 40% when classifying all students and subset 2 respectively.

The performance of our model has been further analyzed for subset 1 (34 students which at the end of their interaction with the learning environment succeeded to draw the correct forces in most equilibrium experiments...
they performed). We compared the number of classifications for each linguistic value provided by the teacher and our model respectively. Fig. 15 illustrates the percentage of students from subset 1 classified in each linguistic value by the teacher and our model respectively. As shown in Fig. 15, the teacher classified 35%, 26%, 24%, 15%, 0% of students of subset 1 in the linguistic values surface, rather surface, average, rather deep and deep respectively, while the neural network-based fuzzy diagnostic model classified 21%, 29%, 26%, 18%, 6% of students of subset 1 in the linguistic values surface, rather surface, average, rather deep and deep respectively.

As described in Sections 4.2 and 4.3 the neural network-based fuzzy diagnostic model has been implemented, trained and tested in order to be used to assess “student tendency to learn by discovery in a deep or surface way” in the learning environment “Vectors in Physics and Mathematics” following suggestions and experience of the group of five expert teachers. As we can see in Fig. 15 the specific teacher when classifying students of subset 1 disagrees with the group of expert teachers and provides a more strict classification for students regarding “student tendency to learn by discovery in a deep or surface way”. As we can see in Fig. 15 classification provided by our model, follows an almost bell shaped function, while classification provided by the specific teacher does not have the same shape and the higher percentage of students (35%) is classified in the linguistic value “surface”. As described in Stathacopoulou et al. (2005), one of the goals when designing the neural network-based fuzzy diagnostic model is to propose a model that can be tailored to individual teacher’s experience in evaluating students. The neuro-fuzzy implementation of the model can easily handle subjectivity of teachers’ suggestions and reasoning, creating that way a model tailored to the needs of a particular teacher in case of disagreement. In order to tailor the model to the teacher’s subjective assessments following his suggestions, by using the same linguistic variables and the same linguistic values for student observable behavior as suggested by the group of experts, adjustments can be made in two ways: in the association between the linguistic values and the universes of discourse by changing the adjusting parameters $m_1$, $m_2$, $m_3$ which represent the expected mean value of the numerical input $X = \{x_1, x_2, x_3\}$ (thus, slightly altering the shape of the membership functions, i.e. the degree of membership to each linguistic value), and at the inference stage by changing the weights in the fuzzy relations network.

In order to tune our model to teacher’s subjectivity, we have analyzed the misclassifications of our model for each linguistic value when classifying students of subset 1 (34 students). Fig. 16 illustrates the percentage of students classified by our model in a linguistic value different from the linguistic value classified by the teacher. As shown in Fig. 13 our model misclassifies a 42%, 22%, 25%, 40% of students of subset 1, for the linguistic values surface, rather surface, average, rather deep, respectively. As we can see in Fig. 16 type of error 2 is present in the linguistic values surface and rather surface, since our model misclassifies students classified by the teacher as surface (42%), 25% of them in the linguistic value rather surface and 17% of them in the linguistic value average, misclassifies students classified by the teacher as rather surface (22%), classifying a 11% of them as average and 11% as rather deep, misclassifies students classified by the teacher as average (25%) as rather deep and misclassifies students classified by the teacher as rather deep (40%) as deep.

In order to tune our model to teacher’s personal view, we have reconsidered with the assistance of the specific teacher the association of linguistic values $V_{i1}, V_{i2}, V_{i3}$ ($i = 1, 2, 3$) and the universe of discourse $U_i$ ($i = 1, 2, 3$) therefore the membership functions $y_{i1}(x_i, m_i)$, $y_{i2}(x_i, m_i)$, $y_{i3}(x_i, m_i)$, $x_i \in U_i$, $m_i \in U_i$, $U_i \subset \mathbb{R}^+$ ($i = 1, 2, 3$) which associate the linguistic values $V_{i1}, V_{i2}, V_{i3}$ ($i = 1, 2, 3$) of each linguistic variable $B_i$ ($i = 1, 2, 3$) with the universe of discourse $U_i$ ($i = 1, 2, 3$). Analysing the linguistic
description of the specific teacher in sample points of the universe of discourse \( U_i \) \((i = 1, 2, 3)\), we found that the specific teacher when associating the universe of discourse \( U_1 \) of the linguistic variable \( B_1 = \text{"the number of times a student tests his/her ideas or compares his/her ideas with reality"} \) with the linguistic values \( V_{11} = \text{"seldom"} \) \( V_{12} = \text{"sometimes"} \) \( V_{13} = \text{"frequently"} \) provided different interval

for the maximum membership degrees as well as for the overlapping areas of the membership functions \( y_{11}(x_1) \), \( y_{12}(x_1) \), \( y_{13}(x_1) \), \( x_1 \in [0, 35] \).

As already mentioned and described in Section 2.2.2, the association between the universe of discourse \( U_1 \) of the linguistic variable \( B_1 \) and the membership degrees \( y_{11}, y_{12}, y_{13} \) of the linguistic values \( V_{11}, V_{12}, V_{13} \) can be adjusted on-line

---

**Fig. 16.** Percentage of misclassifications provided by the neural network-based fuzzy diagnostic model in each linguistic value for subset 1.

**Fig. 17.** Adjusted membership functions for the linguistic variable \( B_1 = \text{"the number of times a student tests his/her ideas or compares his/her ideas with reality"} \).
to the specific teacher’s linguistic description, using the adjusting parameter $m_1$, $m_1 \in [0, 35]$, which represents the expected mean value of $x_1$, which moves the center of membership functions $y_{11}(x_1, m_1), \ y_{12}(x_1, m_1), \ y_{13}(x_1, m_1)$, $x_1 \in [0, 35], m_1 \in [0, 35]$. Fig. 17 illustrates the membership functions $y_{11}(x_1, m_1), \ y_{12}(x_1, m_1), \ y_{13}(x_1, m_1)$ (continuous lines) and the adjusted membership functions to teacher’s personal view $y_{11}(x_1, m_0), \ y_{12}(x_1, m_0), \ y_{13}(x_1, m_0)$ (dotted lines) of the three linguistic values $V_{11} = \"seldom\", \ V_{12} = \"sometimes\", \ V_{13} = \"frequently\”$ of the linguistic variable $B_1 = \"the number of times a student tests his/her ideas or compares his/her ideas with reality\”$.

After adjusting the membership functions to teacher’s personal view, we again compared classification of students of subset 1 by the teacher regarding “student tendency to learn by discovery in a deep or surface way” in the term set {Deep, Rather Deep, Average, Rather Surface, Surface} with classification by the neural network-based fuzzy diagnostic model. The overall success in diagnosis has been 88%, which an improved performance of our model in comparison with 68% without adjustments. Fig. 18 illustrates the percentage of students classified after adjustments by our model in a linguistic value different from the linguistic value classified by the teacher. As shown in Fig. 18 our model misclassifies a 25%, 11%, 12.5%, 20% of students of subset 1, for the linguistic values surface, rather surface, average, rather deep, respectively. As we can see in Fig. 18, the type of error 2 is not present in the linguistic values surface and rather surface. Our model after adjustments misclassifies students classified by the teacher as surface (25%) in the linguistic value rather surface, misclassifies students classified by the teacher as rather surface (11%) as average, misclassifies students classified by the teacher as average (12.5%) as rather deep and misclassifies students classified by the teacher as rather deep (20%) as deep.

5. Conclusions and future work

This paper presents how teachers’ expertise and students’ logfiles have been used in implementing, training, testing and evaluating the neural network-based fuzzy diagnostic model, which is a general diagnostic model, in diagnosing aspects of students’ learning style in the Intelligent Learning Environment “Vectors in Physics and Mathematics”. Experimental results of implementing, training and testing using simulated students and the recommendations and expertise of a group of five experienced teachers show that the neural network-based fuzzy diagnostic model manages the inherent uncertainty associated with human tutors’ expertise in diagnosing aspects of students’ learning style successfully, especially for marginal cases where our model accurately evaluates students by synthesizing conflicting assessments. Evaluation of our model in real classroom conditions, with an experienced Physics teacher and 49 students, has shown that our model provides better results, with overall classification success 68%, as concern a type 1 error 24%, type 2 error 9% and type 3 error 0% when classifying students which at the end of their interaction with the learning environment succeeded to draw the correct forces in most equilibrium experiments they performed (subset 1) in comparison with 27%, 13%, 20%, 40% when classifying students which did not succeed to draw the correct forces (subset 2). The performance of our model has been further analyzed for subset 1 in order to be tailored to individual teacher’s experience in

![Fig. 18. Percentage of misclassifications provided by the neural network-based fuzzy diagnostic model in each linguistic value for subset 1 after adjustments.](image-url)
evaluating students. Results of our model after adjusting the membership functions to teacher’s personal view, shown an improved performance with overall classification success 88%, and there are no type of errors 2 and 3, in comparison with overall classification success 68%, while type of error 1 is present with 24%, and type of error 2 with 9%.

Results of implementing, training, testing and evaluating, have shown the potential of the neural network-based fuzzy diagnostic model, but further work needs to be undertaken, in order to fully explore the benefits and limitations of our approach. Our current work targets in implementing our model in diagnosing aspects of students’ motivation and knowledge level, as well as in implementing our model in distance learning in an adaptive educational hypermedia. In addition, we are currently implementing an interface with which, the teacher can interact and modify the adjusting parameters of the membership functions, as well as in the fuzzy relations. We are also implementing intelligent help for students during learning interaction, based on student classification regarding their leaning style provided by our model. Since deep learners often prefer self-regulated learning, while surface learners often prefer externally regulated learning (Vernunt, 1992) (Vernmetten, Lodewijks, & Vernunt, 2001), diagnosis of an aspect of students’ deep or surface learning style in the learning environment “Vectors in Physics and Mathematics” will be used in providing students with the appropriate work sheets intended to guide their learning.

References


