

On Predicting Learning Styles in Conversational Intelligent Tutoring Systems using Fuzzy Decision Trees

Keeley Crockett^a, Annabel Latham^a, Nicola Whitton^b

^aSchool of Computing, Mathematics and Digital Technology,

^bEducational and Social Research Institute

The Manchester Metropolitan University, Chester Street,

Manchester, M1 5GD, UK

Telephone (+44) (0) 161 247 1497

email K.Crockett@mmu.ac.uk; A.Latham@mmu.ac.uk; N.Whitton@mmu.ac.uk

Corresponding Author is Keeley Crockett

Abstract

Intelligent Tutoring Systems personalise learning for students with different backgrounds, abilities, behaviours and knowledge. One way to personalise learning is through consideration of individual differences in preferred learning style. OSCAR is the name of a Conversational Intelligent Tutoring System that models a person's learning style using natural language dialogue during tutoring in order to dynamically predict, and personalise, their tutoring session. Prediction of learning style is undertaken by capturing independent behaviour variables during the tutoring conversation with the highest value variable determining the student's learning style. A weakness of this approach is that it does not take into consideration the interactions between behaviour variables and, due to the uncertainty inherently present in modelling learning styles, small differences in behaviour can lead to incorrect predictions. Consequently, the learner is presented with tutoring material not suited to their learning style. This paper proposes a new method that uses fuzzy decision trees to build a series of fuzzy predictive models combining these variables for all dimensions of the Felder Silverman Learning Styles model. Results using live data show the fuzzy models have increased the predictive accuracy of OSCAR-CITS across four learning style dimensions and facilitated the discovery of some interesting relationships amongst behaviour variables.

Keywords

Intelligent tutoring systems; conversational agents; architectures for educational technology system; fuzzy decision trees

1. Introduction

Intelligent Tutoring Systems (ITS) are adaptive educational systems that utilise intelligent technologies to provide individualised instruction (Brusilovsky and Peylo, 2003; Mitrovic, 2003; Papanikolaou et al., 2003; Popescu 2008, 2009, 2010; Aparicio et al., 2012; Jia-Jiunn et al., 2012). Over a period of time, ITS build a model of the goals, preferences and knowledge of each student, and use this model to adapt the teaching style and provide a degree of intelligent assistance. Personalisation of a student's learning experience has been achieved in such systems through the detection of emotional and cognitive states (Garcia et al., 2007; Garcia-Valdez et al., 2010); affective states including facial and body language (Annar et al., 2010; D'Mello and Graesser, 2012a, 2012b); personality (Leontidis and Halatsis, 2009); learning progressions (Rus et al., 2013) and learning styles (Latham et al., 2014). Reyes et al., (2013) highlighted the current challenges in the development of ITS, which include the expense of system construction, no or minimal reflection on material by students to feedback given by the ITS, no

evaluation standards and the question of whether the students' focus is on learning how to use the system rather than the actual learning content.

A recent review by Sani and Aris (2014) of 50 ITS-related journal papers published between 2008 and 2013 shows a clear trend in the use of artificial intelligence techniques to improve performance and efficiency of ITS. Of particular relevance to our work is the use of fuzzy and Bayesian logic, which has been used increasingly to handle uncertainty with the ITS student model. A brief review of learning systems inspired by computational intelligence can be found in section 2.2.

Learning styles recognise that individuals have instructional preferences in how they choose to learn. The authors recognise that the concept of learning styles in education is contentious, with multiple models existing, looking at a variety of constructs including learning styles, attitudes, cognitive styles, and approaches to studying, with various levels of grounding in data and robustness of measurement. Some models of learning styles, such as the VAK (visual-auditory-kinaesthetic) model have almost achieved acceptance, particularly at school level, through their very ubiquity despite having no clear provenance (Sharp et al., 2008). However, the notion of learning styles can provide insights into the variety and diversity of learning experiences and preferences, but a critical review by Coffield et al. (2004, p138) suggests that there is theoretical incoherence in the field, with variable quality of models, and that "it matters fundamentally which instrument is chosen".

The principal learning styles model used in the field of engineering education is the Felder and Silverman (FS) model (Felder and Silverman, 1998) which is diagnosed by a formal questionnaire. The FS model defines four independent dimensions of preferred learning style: perception (sensory-intuitive), input (visual-auditory), processing (active-reflective) and understanding (sequential-global) which cover different aspects of the learning process. Research has shown that adapting to a student's learning style increases their learning ability which is often assessed by teachers through their performance (Coffield et al. (2014). Learning styles have been widely used to enhance ITS by presenting learning material matched to individual students' preferred styles (Stash et al., 2004; Wang et al., 2006; Popescu, 2008; Fossati et al., 2009). When adapting to learning styles in ITS there are two main issues: how will learning styles be modelled? And how will the system adapt to individual learning styles?

Existing ITS typically capture learning styles using a self-assessment questionnaire (Spalleck, 2003; Wang et al., 2006), or by analysing a student's behaviour within the ITS (D'Mello and Graesser, 2012a; Latham et al., 2012a). Questionnaires are time-consuming to develop and validate, arduous to administer and often not completed accurately. Predicting learning style using a history of student behaviour means adaptation is delayed until several modules have been completed, and also a student's learning style may change over time or for different topics. WELSA (Popescu, 2010) and EDUCE (Kelly and Tangney, 2006) are two ITS that estimate learning style dynamically during the tutoring session for the purpose of curriculum sequencing. However, they do not include a conversational interface or other intelligent tutoring technologies. Dorca et al.

(2013) proposed three strategies for the modelling of student learning styles through reinforcement learning in adaptive ITS. This work highlighted the importance of ongoing adaptation to learning styles and continuous updates to the student model during the actual tutoring session. However, the strategies were only tested through a simulation experiment and no real students were actually involved.

During the past ten years there has been substantial effort in including both dialogue and affect into ITS which has led to the development of Conversational Intelligent Tutoring Systems (CITS). CITS are complex to develop (Cha et al., 2006; Kumar et al., 2010; Latham et al., 2012a, 2012b, Porayska-Pomsta and Mellish, 2013) yet necessary within a constructivist pedagogy (von Glasersfeld 1995; Piaget, 2011) in order to effectively mimic a human tutor and support the construction of a student's knowledge through discussion. Conversational Agents (CA) can provide flexible and robust natural language interfaces to ITS (Latham et al., 2012a, 2012b, 2014). CA (O'Shea et al., 2011; Crockett et al., 2012) functionality includes the ability to explain, reason, emphasise and pursue a course of action based on interactions with humans and other agents. All CA features can be incorporated into a CITS to generate personalised learning through the understanding and correct interpretation of natural language dialogue. Examples of CITS include: AutoTutor and its variants (Azevedo et al., 2003; Graesser et al., 2005a, 2005b; D'Mello and Graesser, 2012a, 2012b; Nye et al. 2014) helps students construct knowledge about computer literacy and physics; CIRCSIM-tutor (Woo Woo et al., 2006) which engages students in discussion to solve physiology problems; and AVIS (Kumar et al., 2010) a CA tutor with 'social conversational skills' that supports small teams of learners in a collaborative learning scenario. Discussions that take place within these CITS do not consider learning styles during tutoring.

OSCAR is the name of a Conversational Intelligent Tutoring System (referred to as OSCAR-CITS in this paper) that implicitly models a student's learning style during a tutoring conversation and dynamically adapts its style of tutoring to suit his or her preference (Latham, 2011; Latham et al., 2012a, 2012b, 2014). Modelling of learning style occurs by inferring or calculating the values of a number of independent behaviour variables (representing the characteristics an individual exhibits) which are captured during the conversational tutoring. In order to determine an individual's preferred learning style (in accordance with FS model (Felder and Silverman, 1998)) and provide the most suitable learning material, the value of a related variable is used. The current method reported in (Latham et al., 2012a) has been shown to be successful (despite not considering the interactions between different behaviour variables) and adaptation to learning styles during a tutorial has been shown to improve the learning experience (Latham et al., 2014). However, a simple winner-takes-all strategy is employed whereby the single best behaviour variable associated with the learning style determines the adaptation for a particular question. Combining behaviour variables would factor in all behaviour extracted from the dialogue to determine the learning style.

The motivation for the research presented in this paper was based on the following observations:

1) From the earlier experiments (reported in (Latham et al., 2012a)), only one of the 41 independent behaviour variables captured during the conversation is used to predict each learning style. For example, in the experiments reported in (Latham et al., 2012b), the sequential and global learning styles are associated with three different independent behaviour variables but predictions only use logic rules (sequential) and approach to queries Q9 (global) in predictions. By combining the behaviour variables, it may be possible to improve the accuracy of prediction.

2) No proposal has been made for how to deal with conflict resolution, i.e. when the selected attribute for each of the two learning styles of a dimension gives opposing classes of learning styles (e.g. the learner is classed as both sequential and global for the understanding dimension).

3) In earlier work (Latham et al., 2012a) behaviour variables have been mapped to individual learning styles according to the FS model. However, it may be that all aspects of behaviour (even that commonly associated with different learning style dimensions) can be useful in modelling an individual's preference, as learning styles models only describe general traits of different learners. By combining the values of all behaviour variables, regardless of the associated learning styles, further information may be mined that improves the prediction accuracies.

The work presented in this paper aims to overcome the weakness of OSCAR-CITS by modelling the interactions between the behaviour variables from individuals who have participated in a tutoring session. Data collected from these tutorials will allow the creation of fuzzy models which take into account all behaviours which should comprehensively predict the learning style dimension rather than using a winner-takes-all strategy. This will be achieved through the use of fuzzy decision trees to automatically find patterns of learner behaviour (using the independent behaviour variables) and produce a set of rules that can improve learning style dimension prediction accuracy. Better prediction by the fuzzy models will allow for better adaptation of tutoring materials to take place and improve the student learning experience. The resulting experimental study will then consider whether it is possible to accurately estimate a learner's learning style from a two-way tutoring discourse using these fuzzy models. This paper proposes a two-phase methodology for modelling learning style using fuzzy decision trees (Crockett et al., 2006; Gasir et al., 2012; Liu et al., 2013; Marsala, 2013). Fuzzy decision trees are based on the principles of classical decision trees that can automatically extract the most relevant features for a given problem from a set of data without human intervention. Classical decision trees are interpretable by humans and are robust and efficient but the decision making process

is often rigid with only one path (rule) in the tree firing thus limiting amount of information contributing towards the predicting the outcome. Application of fuzzy theory, allows all information in the decision tree to contribute to the prediction phase by a matter of degree and allows for human uncertainty to be modelled. Initial work using fuzzy decision trees by the authors (Crockett et al., 2013) has shown that combining such behaviour variables using machine learning techniques has shown comparable prediction levels, however this work did not explore the predictive accuracy of a full learning styles model nor critically analyse the behaviour variables in the context of the models produced.

In terms of the research described in this paper, to be able to predict an individual's learning style fuzzy trees will first be created in an induction phase and then applied to a set of individual learners data to perform prediction. In the induction phase, fuzzy decision trees will be induced from the independent behaviour variables captured through previous tutoring conversations. During the prediction phase all branches within the fuzzy tree will fire to some degree, and a fuzzy inference strategy will be applied to combine the information to provide an overall prediction of an individual's learning style. In order to improve the predictive accuracy of each individual's learning style, a fuzzy decision tree is learned automatically from a set of data about a domain and provides a set of rules for classifying or predicting an outcome.

Initial work (Crockett et al., 2013) tested out the use of fuzzy decision trees in the development of a Fuzzy Learning Styles Predictor Module which would interface with the predictive version of OSCAR-CITS. This work demonstrated that combining behaviour variables using machine learning algorithms (including support vector machines, the J48 decision tree and fuzzy trees) gave a higher predictive accuracy compared with the accuracy obtained using one variable as reported in (Latham et al., 2012a). The first limitation of this work was that only two learning dimensions of the FS model were considered, and given choice and execution of learning style model is critical to successful learning, the need to investigate the predictive ability across all dimensions is essential to evaluate the overall predictive accuracy of OSCAR-CITS. Secondly, the initial work did not examine the information content of the behaviour variables and their relationships in building predictive models of each dimension and no detailed analysis was conducted.

The novel contribution of this paper can be summarised as follows:

- 1) The enhancement and extension of the Fuzzy Learning Styles Predictor (FLSP) module (first reported in Crockett et al., 2013) to use a simpler fuzzy inference and fuzzification strategy to allow transparency in the decision making process leading to prediction of a learning style dimension. This enables transparency in the model for tracking the importance of particular behaviour variable combinations influencing an individual's learning style.
- 2) An optimised and simplified approach of inducing all four fuzzy learning style models from behavioural variables for all dimensions in the FS model: Sensory – Intuitive, Visual – Verbal, Active – Reflective, Sequential – Global. This

allows a complete picture of how OSCAR-FLSP performs in terms of predictive accuracy across a complete learning style model.

- 3) A comparative evaluation for all FS dimensions between the predictive accuracy of OSCAR-CITS compared with OSCAR-FLSP and the classical decision tree algorithm C4.5 obtained from building fuzzy decision tree models from a Learner Behaviour dataset.

The paper is organized as follows: Section 2 describes the foundations of the work, including the OSCAR-CITS and an overview of fuzzy decision trees. Section 3 presents the new OSCAR-FLSP (Fuzzy Learning Style Predictor) architecture and describes the two phase methodology of inducing fuzzy decision trees from a previous conversational interaction. Section 4 explains the experimental process that was undertaken on the four dimensions of the FS model. Finally, section 5 presents the results and discussion.

2. Foundations

This section starts by providing an overview of the Felder and Silverman Learning Styles Model and why it was selected as a model to personalise learning in the conversational intelligent tutoring system used within this study. Related work in the field of conversational intelligent tutoring systems is provided with a detailed description of the original OSCAR-CITS, Fuzzy Logic and Other Machine Learning Approaches. The current state of the art in learning systems inspired by machine learning, specially fuzzy logic is briefly highlighted. This includes a review of how existing machine learning approaches which incorporate fuzzy decision trees have been used to induce fuzzy models for predicting dimensions of the FS learning styles model.

2.1 Felder and Silverman Learning Styles Model

Learning styles describe the way in which groups of people prefer to learn. There are several contradictory theories about learning, and therefore no single agreed definition. Some learning theories are based on psychological theories such as personality traits and intellect whereas others focus on brain functioning or the learning environment. One established and popular model is known as the Felder-Silverman (FS) learning styles model (Felder and Silverman, 1988) which was developed to describe the learning styles in engineering education and suggest different teaching styles to address learners' needs. The Felder and Silverman model, uses four aspects of learning style to describe the learning style of engineering students. The four dimensions of learning style described in the FS model are associated with different steps in the receipt and processing of information:

- The *perception* dimension describes learners as *sensory* or *intuitive* depending on the way in which they prefer perceive the world, externally through observing and gathering data through the senses (sensory) or internally through speculation or hunches (intuitive).
- The *input* dimension classes learners as *visual* or *verbal* according to their preferred method of receiving external information, e.g. seeing diagrams (visual) or reading textual explanations (verbal).
- The *processing* dimension labels learners as *active* or *reflective* corresponding to the way they convert information into knowledge, e.g. by discussing information (active) or through introspective thought (reflective).
- The *understanding* dimension describes learners as *sequential* or *global* depending on their progression towards understanding, e.g. whether they take continual steps (sequential) or first gain an overview of a topic before exploring the detail (global).

There are 16 (2⁴) learning styles overall (an example being sensory/visual/ active/sequential). Each learning style dimension may be thought of as an axis with the opposite learning styles at each end and learners are placed on each axis according to the strength of their preferred learning style. By defining independent dimensions the FS model describes the detail of learner tendencies, including the strength of preference as well as the nature of learning styles. The Index of Learning Styles (ILS) (Felder and Silverman, 1998) is a 44-question self-assessment diagnostic questionnaire that determines an individual's learning style under the FS model. For each dimension, answers are compared and result in a learning style and a score from 1 to 11 (in steps of 2). Scores of 1 or 3 place the learner at the centre of the axis and indicate a low preference for that learning style, which Felder & Silverman call neutral preferences. The FS model has been selected in this study for a number of reasons: the dimensions are distinct and independent, teaching styles exist that map onto preferred learning styles, the model was designed for engineering students and therefore applicable in on line tutoring systems that deliver computing related courses.

2.2 Conversational Intelligent Tutoring Systems Approaches

Conversational Intelligent Tutoring System (CITS) are designed to model a human tutor by directing a tutoring conversation in natural language with a user in the same style as a human. Their development typically brings together two strands of research: intelligent tutoring systems and conversational agents. Intelligent tutoring systems (ITS) extend the traditional computerised learning systems 'one-size-fits-all' approach by capturing and modelling individual traits used to personalise the instruction (Brusilovsky and Peylo 2003). This involves presenting learning material in a style and order to suit the learner (e.g. by presenting learning material matched to preferred learning styles), and also proactively helping learners, e.g.

by giving intelligent feedback on incomplete or erroneous solutions and guidance to assist learners in constructing solutions to problems. Within the research field of CITS, there are two leading systems: AutoTutor and OSCAR-CITS.

AutoTutor uses an animated conversational agent (CA) to present tutoring questions and engage in mixed initiative dialogue (where dialogue flow is controlled by both the agent and the user) whilst guiding students towards constructing a solution. As well as showing AutoTutor's response textually on screen, the animated CA speaks the text and shows appropriate facial expressions and gestures, although it is the dialogue content rather than the animation and speech that influences learning (Graesser et al., 2003). AutoTutor has since been expanded to adapt to learner emotions as well as their knowledge (D'Mello et al., 2009, Graesser et al., 2008a). Emotions such as boredom and frustration are modelled using sensors to detect facial expressions and body posture as well as dialogue patterns. AutoTutor responds by, for example, giving a hint to a frustrated student or giving a challenging problem to a bored student. Adapting to emotions improved deeper learning for students with low knowledge, however some students were irritated by the empathic AutoTutor (D'Mello et al., 2010b) and the detection of boredom and confusion from dialogue patterns alone was poor (D'Mello et al., 2008). The necessity for sensors to detect emotion in a real-life learning environment has many difficulties and facial recognition is a complex task. As well as requiring cameras with sufficient resolution (which is expensive), successful recognition of emotions requires a consistent environment, for example adequate lighting and learners seated in a fixed position. These barriers constrain the widespread use of AutoTutor in real learning environments but the extensive research on AutoTutor has contributed much to the CITS field.

OSCAR-CITS is a sophisticated multi-agent system that aims to imitate a human tutor by delivering an adaptive natural language tutorial based upon a learner's preferred learning style. Like a human tutor, OSCAR-CITS automatically models an individual's learning styles using cues captured from his or her behaviour during the tutoring conversation, and then dynamically adapts its tutoring style to suit each individual's preferences. By adapting the pedagogic approach to the preferred style (as encouraged in Felder & Silverman 1988), OSCAR-CITS aims to provide a more effective learning experience that results in improved learner motivation and a deeper understanding. Results showed that participants completing a tutorial adapted to their learning styles performed 12% better in answering tutorial questions than those without suitable adaptation (Latham et al. 2014). Users value OSCAR-CITS's personalised tutorials, which enable them to discuss and construct knowledge, give instant feedback and are available 24-hours a day (Latham et al., 2012a, 2012b, 2014) By modelling a classroom tutorial style and engaging in a natural language dialogue, OSCAR-CITS is intuitive to use, which helps learners feel more confident. In addition to implementing an intelligent conversational agent to converse with learners, OSCAR-CITS incorporates several intelligent techniques, including:

- presenting learning material in a style and sequence adapted to the user's learning style (curriculum sequencing);
- providing analysis and detailed feedback of errors and omissions in solutions (intelligent solution analysis);
- modelling the approach used by human tutors to discuss questions and give hints (problem solving support) to encourage learners to construct their own knowledge.

The original OSCAR-CITS architecture can be seen in Figure 1 and interface in Figure 2.

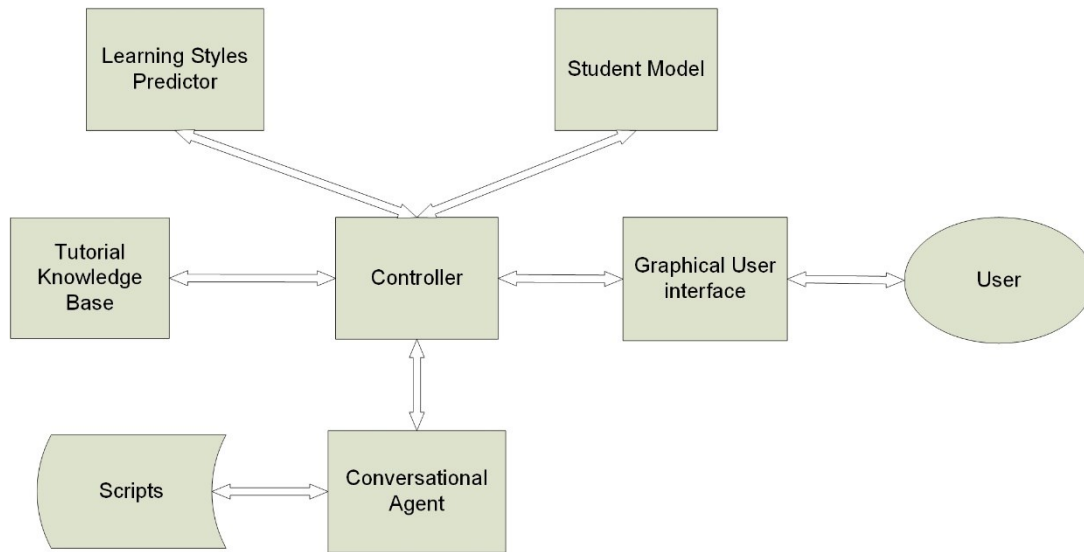


Fig.1. Original OSCAR-CITS architecture (Latham et al. (2012))

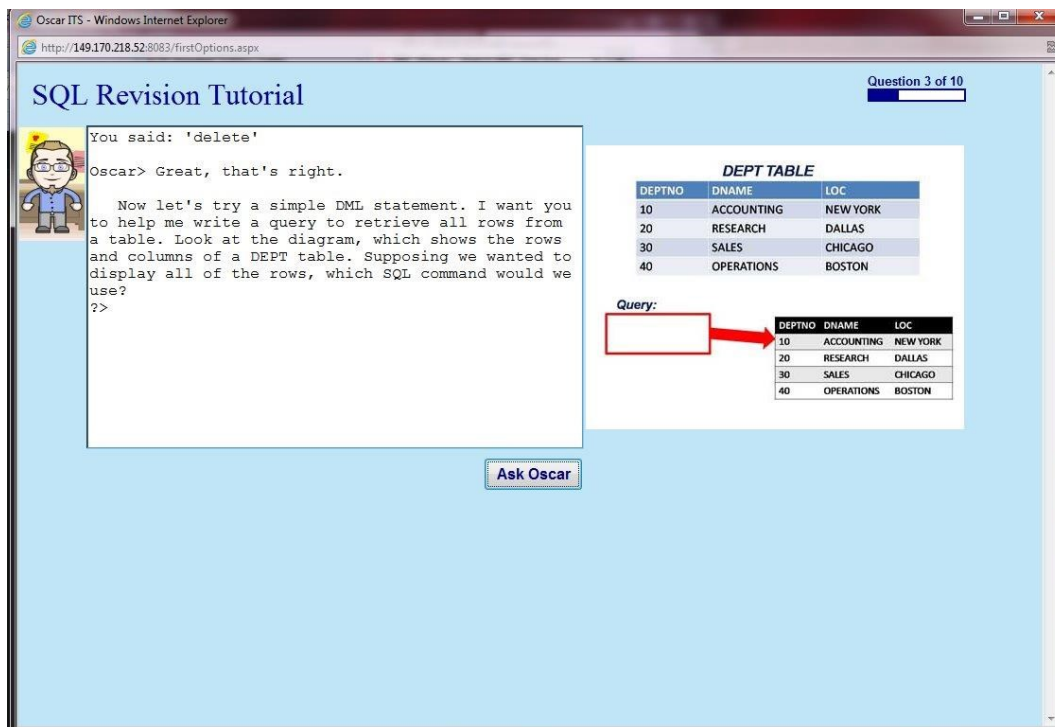


Fig. 2. OSCAR-CITS conducting a SQL tutorial

OSCAR-CITS is a complex system; its modular design facilitates reuse and ease of maintenance through separation of learning styles and domain knowledge bases from its functionality. As such, OSCAR-CITS is independent of the learning styles model and tutoring discipline, allowing it to be extended for different domains by adopting a 3-phase methodology that describes how to construct a Learning Styles Predictor, Tutorial Knowledge Base and Conversational Agent (Latham et al. (2012a)). The methodology features both generic components and tools to aid development (behaviour variables, key words, logic rules, 3-level conversation model, question styles and templates).

The Conversational Agent also includes a logging component that records information about the behaviour of the learner, such as timing of interactions, the number of words used, the number of times FAQs are asked and the type of tutor resource accessed (Latham et al. (2012a)). The tutoring conversation is also recorded, along with information about the student knowledge of the topic being discussed and the adaptation employed. A full description of the architecture and development methodology of OSCAR-CITS with its implementation to deliver an SQL tutorial for university students can be found in Latham (2011). The current live version of OSCAR-CITS is delivering SQL revision tutorials at Manchester Metropolitan University (MMU).

In OSCAR-CITS implementation, the knowledge of typical learner behaviours described in the Felder and Silverman Learning Styles model is expressed as 33 logic rules that were incorporated into the OSCAR-CITS Learning Styles Predictor agent (Latham et al., 2012a). The logic rules follow a typical IF-THEN format. Figure 3 shows five examples of logic rules used by OSCAR-CITS to increment learning style values which are recorded within the associated variables in Table 2 during tutoring.

1. Example rule to test how verbose the learner is:

```
IF (total-student-wordcount-per-interaction (wordcount) >= average-student-wordcount-per-interaction (mean_Wordcount))
```

```
THEN
```

```
    INCREMENT VERBAL;
```

```
ELSE
```

```
    INCREMENT VISUAL;
```

2. Example rule testing whether showing an example visually helps the learner receive and process information:

```

IF (answer IS (wrong OR don't-know) AND show-movie)
THEN
{
    IF (next-answer IS right)
    THEN
        (INCREMENT VISUAL) AND (INCREMENT ACTIVE);
    }

```

3. Example rule testing whether the learner prefers to try to solve the problem all at once:

```

IF (question-template-applied IS choice-of-approach)
THEN
{
    IF (student-chooses-onego OR student-attempts-query)
    THEN
        (INCREMENT GLOBAL) AND (INCREMENT ACTIVE);
    }

```

4. Example rule testing whether the learner prefers to solve problems one step at a time:

```

IF (question-template-applied IS choice-of-approach)
THEN
{
    IF (student-chooses-steps OR student-chooses-don't-know)
    THEN
        (INCREMENT SEQUENTIAL) AND (INCREMENT REFLECTIVE);
    }

```

5. Example rule to test how comfortable the student is with words and with detail:

```

IF (answer IS (wrong OR don't-know) AND answer-given-in-explanation-text)
THEN
    (INCREMENT INTUITOR) AND (INCREMENT VISUAL);

```

Fig. 3. Example of logic rules used to adjust student learning style based on tutoring conversation

The empirical evaluation of OSCAR-CITS modelling of learning styles, conducted in a real teaching/learning environment, is described by Latham et al. (2012a, 2014) and summarised here. A conversational tutorial was developed for undergraduate computing students on structured query language (SQL). The OSCAR-CITS tutorial was integrated into several undergraduate modules at MMU, and learner behaviours were captured throughout the tutorial. All participants were asked to complete the Index of Learning Styles (ILS) questionnaire (Felder and Silverman, 1998) and its results were compared to the OSCAR-CITS predictions to compute the prediction accuracy. The research investigated whether it was possible to predict learning style from

behaviour captured during a two-way conversational tutorial with a CITS (Latham, 2011). The experiments (reported in Latham et al. (2012a)) had 95 participants, and showed it was possible to predict learning style, giving best prediction accuracies for individual learning styles of 61-100%, as shown in Table 1.

Table 1 OSCAR-CITS best prediction accuracy.

	n	Sensory	Intuitive	Visual	Verbal	Active	Reflective	Sequential	Global
OSCAR-CITS	95	70%	80%	94%	71%	100%	73%	82%	61%

To evaluate the usability of OSCAR-CITS all participants were asked to complete a feedback questionnaire (Latham et al. 2014). The questionnaire consisted of eight questions to be rated using a six-point LIKERT scale (to force participants to express an opinion one way or another rather than selecting the centre rating), four questions requiring a Yes/No answer, and three open questions.. The user evaluation showed that OSCAR-CITS was well received by participants, with 87% of participants rating the tutoring highly (with 51% giving the highest rating). 94% of participants found the tutoring helpful (with 72% giving the highest rating) and 85% of participants felt that OSCAR-CITS had helped them to revise. Overall, 89% of participants stated that they would use a resource like OSCAR-CITS if it were available. Slightly more than half of participants (52%) stated that they would use OSCAR-CITS instead of reading a book, and a surprising 35% of participants stated that they would use OSCAR-CITS instead of attending a face-to-face tutorial.

The current version of OSCAR-CITS models learning style using a number of separate behaviour traits exhibited during the conversation as well as the result of a combined set of logic rules. The results reported show good prediction accuracies for each learning style individually; however the mechanisms have not been combined to give an overall classification for each learning style dimension.

2.3 Fuzzy logic and Other Machine Learning Approaches and Applications

Research into the application of machine intelligence to ITS has been undertaken since 1996 when Hawkes and Derry (1996) proposed a fuzzy model that captured the way in which human tutors evaluated borderline tutoring decisions in accordance with student behaviour. Fuzzy models, based on fuzzy set theory resembles human reasoning in its use of approximate information uncertainty to inform decision making. Since then there have been numerous examples (Chenn-Jung et al., 2007; Chen and Liu, 2008; Cabada et al., 2009; Fazlollahtabar and Mahdavi, 2009; Chih-Yueh et al., 2011; Zarandim et al., 2012; Hsieh et al., 2012; Lin et al., 2013; Vassen et al., 2014) where machine intelligence has successfully been applied within the context of ITS, such as:

- a virtual teaching assistant that evaluates student solutions and generate hints in a program tutorial (Chih-Yueh et al., 2011);
- the application of k-means clustering to identify help-seeking strategies in ITS (Vassen et al., 2014);
- an adaptive web-based learning system that is optimized using a multi-layer feed forward neural network to predict a student's cognitive styles (Cabada et al., 2009);
- the creation of a personalised on-line programming tutorial through identification of learning styles by mining student web logs using the AprioriAll algorithm (Klasnja-Milicevic et al., 2011). In Klasnja-Milicevic et al. (2011) clustered learners based on their learning style and then used the AprioriAll pattern mining algorithm to extract behavioural patterns from log files. By comparing learner behaviours to each cluster, learning styles were identified and learning material recommended.

A detailed review of state of the art in the automatic detection of learning styles can be found in (Feldman et al., 2014)

Bayesian networks are probabilistic models that have been used to model the relationships between learning styles and behaviour factors. Garcia et al., (2007) used Bayesian networks to infer student learning styles from a history of their behaviour in using the ITS. Three dimensions of the FS model were modelled, with precisions of 58-77%. Enhancing the Bayesian model improved precisions to 66-80%. EDUCE (Kelly and Tangney, 2006) offers different resources styled using four of the Gardner multiple intelligences (Gardner, 1993) and used Naïve Bayes (a form of simple probabilistic classifier) to predict which resources students prefer, based on past choices.

Artificial neural networks are computational models inspired by the neural structure of the brain, which have been used to classify student learning styles based on behaviour. Villaverde et al. (2006) used a neural network to predict student learning style for three dimensions of the FS model. The neural network analysed student behaviour in an ITS to automatically model learning styles, achieving an accuracy of 69.3%. However, only ten behaviour factors (input neurons) were used. A known weakness with neural networks is their 'black box' nature, which means that no information is learned about which behaviour factors are most significant in predicting a learning style.

The use of fuzzy logic to handle uncertainty and imprecision in learning systems is particularly relevant to the study presented in this article. Work by Nykanen (2006) examined the use of fuzzy logic in educational design in two distinct areas: how to integrate imprecision in predictive systems due to the lack of tutoring information; and how fuzzy descriptions and expressions can be used to provide an intuitive and linguistic interface. Nykanen (2006) developed a fuzzy model that could predict failing students on a university mathematics course by identifying students who required additional assistance. The DEPTHS (Design Pattern Teaching Help System) utilized a knowledge assessment method based on fuzzy rules to assess student understanding of domain concepts (Jovanović and Gašević, 2012).

A genetic fuzzy expert system that automatically estimates the difficulty level of questions in a specific competitive learning system in the virtual learning environment Moodle showed promising initial results (Verdu et al., 2014). This system consisted of a fuzzy model generator that learned characteristics of the questions in the environment for a number of difficulty levels. Initial classifications of questions into levels of difficulty were made by a teacher and are therefore open to subjective interpretation. The authors emphasized that more work is required on how to map fuzzy sets to levels of question difficulty perceived by students. Fuzzy rule-based systems have also been developed in conjunction with collaborative filtering algorithms to recommend the most appropriate learning objects to students based upon their learning style (Arce-Cardenas and Garcia-Valdez, 2012) and fuzzy expert systems architectures are also being considered in the development of ITS (Xu et al., 2002) through the use of fuzzy cognitive maps. More recently, a prototype fuzzy logic based personalized learning system for adaptive English learning was proposed (Hsieh et al., 2012). Whilst each of the above learning systems (inspired by computational intelligence techniques) have shown degrees of success using small experimental groups, none attempt to predict the learning style through obtaining variables captured through natural language dialogue which occurs naturally during a tutoring conversation.

Decision tree methodologies provide a natural set of tools for extracting rules from both symbolic and numeric data. The aim is to induce a set of rules for classifying objects from knowledge of a subset of the domain of objects whose attribute values and classes are given (Quinlan, 1993). The resulting set of rules formulates the decision tree model which is comprised of attributes from the domain that allow the user to comprehend, for a new case passing through the tree, how a decision is made to some degree of accuracy. In other words, all data from the domain is used to build a tree of questions with answers at the leaves. To answer a new query, start from the tree root, and answer the questions until a leaf node is reached and then return the answer specified at that leaf node.

The process of classification can be simply defined as the task of discovering rules or patterns from a set of data. The objective of any classification task is to at least equal and essentially exceed a human decision-maker in a consistent and practical manner. In studies, such as the one presented in this paper, the classification is provided by the results of the formal Index of Learning Styles questionnaire and not by human labels. The benefit of using computational intelligence techniques for classification is that they are independent of other factors that may influence a human's intuitive labelling during tutorials, such as the tutor's mood or preconceptions about the student. The success of any classification algorithm can be measured in terms of accuracy, which is determined by its performance on an unseen set of data (test set) and its speed, the time taken to learn a set of rules or discover patterns from a training set of data. Providing that the training set is randomly sampled from the given data set, the accuracy of the test set can be seen to be an un-biased estimate of performance. For example, an induced

decision tree could be used to determine what kinds of behaviour, exhibited by a student during a conversational agent tutorial, lead to a particular learning style being preferred over another.

Classical decision trees have also been used to predict the value of a variable (e.g. learning style) based on a number of input variables (e.g. behaviour factors). Cha et al. (2006) designed their ITS interface to capture behaviour related to the FS model, using decision trees and Hidden Markov Models to classify learning styles from learner choices and behaviour. Ozpolat and Akar (2009) used the NBTtree algorithm to classify student learning styles from the content of learning objects rather than behaviour (using keywords matched in Internet search terms). In their study involving 30 participants, they achieved accuracies of between 70 and 73.3% across different learning style dimensions. Chen and Liu (2008) used decision trees and K-means clustering to automatically identify cognitive styles from learning patterns

Fuzzy decision trees (Crockett et al., 2006) are an extension of classical decision trees where the main fundamental difference between fuzzy and crisp trees is that with fuzzy trees, gradual transitions exist between attribute values. For example, consider a Bank is trying to use a system to predict whether to offer a person a loan based upon the data they have supplied from an application form. Within the induced crisp tree there exists a node for the attribute age. Values less than or equal to 21 (left branch) strongly suggest a person will not get a loan, whilst values greater than 21 (right branch) suggest that more information about the person needs to be analyzed to make a decision. The decision to split on the value of 21 for age is a strict decision that the tree algorithm has learned from the data and it is not flexible. Previous research (Crockett et al., 2006; Crockett et al., 20012; Crockett et al., 2013) has identified that new cases whose values are close to the associated node value may be misclassified due to this inflexibility. In this case, for example, values between say 18 and 24 may not be clear cut in terms of determining who gets a loan. Applying Fuzzy trees to deal with this type of uncertainty present at node values in crisp trees enables a more flexible approach to decision-making. This is achieved by modelling each attribute within the tree as a fuzzy set, which generates a degree of membership for each related attribute value. In addition, the reasoning process within the tree allows all rules to be fired to some degree, with the final classification being the result of combining all membership grades. A number of methodologies exist for the development of fuzzy decision trees, with the main distinction being whether to induce the fuzzy decision tree from a raw dataset or to first fuzzify the dataset using either experts in the domain or via an automatic method.

Previous work (Crockett et al., 2006) has shown that the generation of a fuzzy decision tree from crisp induced trees produces a generalized and robust rule set without the need for human expert intervention during the fuzzification process. In order to predict learning styles from the independent variables captured during the tutorial conversation a fuzzy tree algorithm known as FIA (Crockett et al., 2006) will be used. One of the strengths of the algorithm is that no pre-fuzzification of the data is required in order to induce the fuzzy tree. FIA provides a mechanism for the direct fuzzification of both discrete and continuous

attributes and the ability to use a range of fuzzy inference techniques. In the context of this work FIA has undergone modification and will be described in section 3.

3. Methodology

3.1 An Architecture for a Fuzzy Learning Styles Predictor within a CITS

This section describes the addition of a new module, known as the Fuzzy Learning Styles Predictor (FLSP) module to the original OSCAR-CITS architecture reported by (Latham et al., 2012a). The complete architecture, referred to as OSCAR-FLSP is shown in Figure 4.

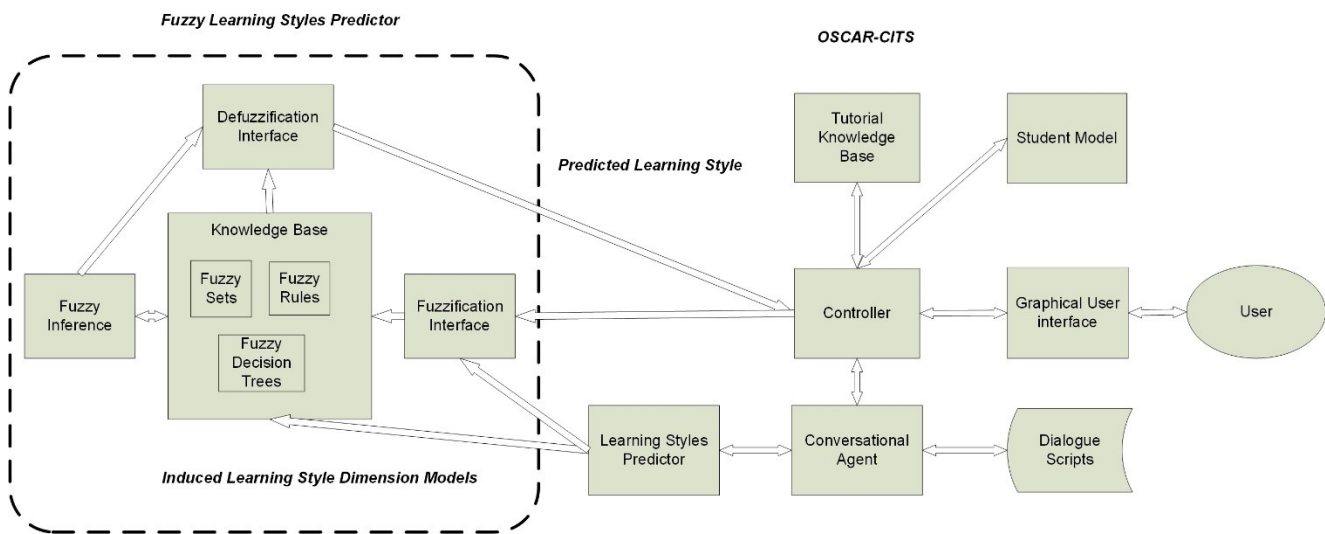


Fig.4. OSCAR-CITS with Fuzzy Learning Styles Predictor Module

When a learner first engages with the tutorial through the textual conversational interface, variables reflected through the learner’s behaviour are detected via the conversational agent (and its dialogue scripts), collated, and analyzed via the controller. The controller is responsible for managing the tutorial, and for communicating with all components. The tutorial knowledge base manages course information, such as topics and their breakdowns, related tests and teaching material. All information within the knowledge base is categorised according to teaching style (related to learning style). At the start of the first tutoring session, no initial learning style values exist for a student. During the tutoring conversation, learning style values will be incremented depending on the learner’s statements and behaviour. Periodically, the value pairs of each learning style dimension will be compared to reveal the student’s overall learning style tendency for that dimension (i.e. the greater value). Learning style values depend on an individual’s unique tutoring session, and if no evidence is gathered relating to a particular learning style dimension, that learning style will remain unclassified. The student model records information such as the student

registration number and password and a predicted numerical value for each learning style. The conversational agent receives natural language text and information about topic and learning style from the interface and stores a log file, a state file and a time log which enable the analysis of behaviour such as timings, word and interaction counts during experimentation. The conversational agent accesses a database of dialogue scripts in order to match the input and generate a response. The graphical user interface displays a webpage showing questionnaires, tests, images, documents and interactive movies and sends communication to and from the user.

The FLSP module consists of four principal components:

- *Fuzzification Interface* that takes raw conversational cues (e.g. number of interactions to complete a question) and maps to an appropriately defined fuzzy set(s).
- *Knowledge Base* consisting of induced fuzzy rules obtained from the fuzzy decision tree models and a series of fuzzy sets representing branches within the tree. The fuzzy decision tree models are induced from the behavioural variable knowledge base.
- *Fuzzy Inference Mechanism* for combining grades of membership in all eight dimensions of the four learning styles.
- *Defuzzification Interface* to translate the aggregated fuzzy sets into a fuzzy singleton to predict the learning style dimension at a point in time.

The benefit of using a fuzzy decision tree in the FLSP is that when a new case (set of data comprising behaviour variables and associated values for a specific learner) passes through the tree, all branches in the tree will contribute towards the decision classification (i.e. the learning style dimension) to some degree. Interaction between behaviour variables captured during the tutoring conversation is possible. This is in comparison to traditional decision tree classifiers where only one route through a tree can be chosen. Each component of the FLSP will now be described in the sub-sections that follow.

3.1.1 Fuzzification Interface

Fuzzification is the process of converting crisp values into their fuzzy representations to allow a degree of uncertainty to be represented. The construction of any fuzzy system requires the identification of all control and solution variables within the system to be modelled. In this application, a dataset comprising 41 independent behaviour variables (described in section 4.2) captured during the tutoring conversation is used in the first instance to induce a classical decision tree using the established decision tree algorithm C4.5 (Quinlan, 1993). The behaviour variables are included in the tree based upon their information content with respect to how much they contribute to the prediction of a learning style dimension and will be subject to fuzzification. Figure 5a shows a portion of a crisp decision tree showing the root node *Rules_GLO* that was induced (learned)

from the Learner dataset (Table 2, Appendix). Global learners (on the understanding dimension) typically make intuitive leaps, have difficulty working with material they have not understood and prefer to jump directly to complex material (Felder and Henriques 1995). *Rules_GLO* is the score resulting from rules that fired during the tutoring conversation and has the range $\{0...10\}$. The decision tree has learned from the dataset that the number of rules fired, captured by the behaviour variable *Rules_GLO*, is the most significant attribute in accurately predicting the sequential-global learning dimension. However, *Rules_GLO* splits on the value of 5 which is a generalised crisp value known as the decision threshold for that attribute (dt) i.e. the most significant value of any given attribute, learnt through the induction process, which contributes to the most number of correct classifications overall. Previous work (Crockett et al., 2006; Gasir et al., 2012) has found that such crisp values create hard decision boundaries that do not give the best decision-making, as they do not capture uncertainty inherent in the data.

A solution is to fuzzify this boundary to create a fuzzy region around each node in the tree using two opposing fuzzy membership functions as defined in equation (1). This will enable a new case passing through the tree to be assigned degrees of membership to each branch. Therefore, if a person's tutorial triggered 5 global logic rules they would have a membership grade of 0.5 in both the *Rules_GLO_Left* and *Rules_GLO_Right* membership functions representing maximum uncertainty. Figure 5b shows a sample pair of membership functions for the behaviour variable *Rules_GLO*. When the number of global rules triggered during the tutoring session is between 3.6 and 6.4 then partial degrees of membership are generated for both *Rules_GLO_Left* and *Rules_GLO_Right*.

This allows for uncertainty of each student's learning styles to be taken into consideration during the tutorial. Once all branches throughout the tree have been assigned grades of membership in a similar manner, the fuzzy decision tree will use fuzzy inference to combine membership grades of all behaviour variables in the tree to predict the learning style dimension.

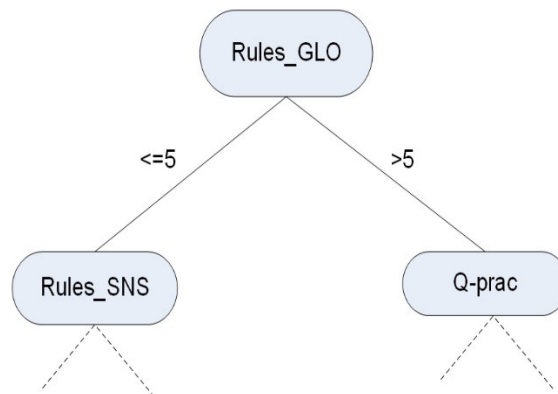


Fig.5a. Root node *Rules_GLO* of induced Classification tree for global-sequential dimension

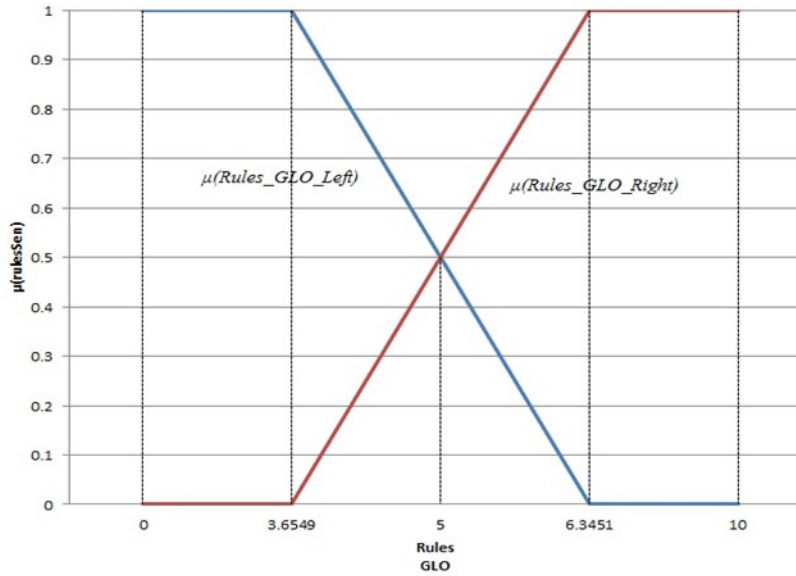


Fig.5b. Example pair of opposing fuzzy membership functions for root node *Rules_GLO*.

The shape of the membership function is very significant in any problem domain. Slight adjustments can have considerable effects on final predictions of a system. The approach adopted for acquiring the shape of a membership function is often dependent on the application. The construction of the membership functions may arise directly from the ‘expert’ (manual) or by automatic generation (automatic). For the purpose of this work, simple linear membership functions have been used.

Each membership function will have an associated domain (dm_i, dn_i) whose scope is determined by the behaviour variable i at each specific node in the tree. To represent a left branch, linear increasing and decreasing membership functions L, may be defined as (Crockett et al., 2006):

$$L \uparrow (dm, dn, x) = \left[\begin{array}{l} 0 \rightarrow x \leq dm \\ \frac{x - dm}{dn - dm} \rightarrow dm \leq x \leq dn \\ 1 \rightarrow x \geq dn \end{array} \right] \quad (1)$$

$$L \downarrow (dm, dn, x) = \left[\begin{array}{l} 1 \rightarrow x \leq dm \\ 1 - \frac{x - dm}{dn - dm} \rightarrow dm \leq x \leq dn \\ 0 \rightarrow x \geq dn \end{array} \right] \quad (2)$$

where dm generates the zero membership value and is the lower bound of L defined as $dm = dt - n\sigma$

where n is a real number $n \rightarrow [0.0, \infty]$ and σ is the standard deviation.

dn generates the complete membership value (maximum membership value) and is the upper bound of L defined as $dn = dt + n\sigma$ where n is a real number $n \rightarrow [0.0, \infty]$ and σ is the standard deviation, and x is the value of variable i . n is empirically determined.

In figure 5b, the domains of Rules_GLO_Left and Rules_GLO_Right are determined using equations (1) and (2). Empirical optimisation using a genetic algorithm is used to determine the most optimal coverage of fuzzification around the decision threshold (dt) of each attribute (Crockett et al., 2006). In this work, optimal equates to obtaining the set of membership function domains which maximise the classification accuracy of the fuzzy decision tree. For Rules_GLO_left and Rules_GLO_right, the value of dt for the Rules_GLO attribute = 5 and through optimisation of n , $dm = 3.6549$ and $dn = 6.3451$. Therefore the degree of uncertainty is only present when a value of Rules_GLO is between $\{3.6549, 6.3451\}$.

3.1.2 Creating the Knowledge Base

In order to create a knowledge base for predicting learning styles of students, a series of fuzzy decision trees were induced from the data described in Table 1 (Section 4.2). This section describes the generalised algorithm used to induce a fuzzy decision tree. The algorithm was inspired from previous work (Crockett et al., 2006) and consists of two phases.

Phase 1: Construction of the knowledge base

1. Generate a crisp classification tree, C , from a set of data, S consisting of i -variables $\{v_1, v_2, \dots, v_i\}$ of domain, D which are used to describe a single classification outcome. In this application, Quinlan's established C4.5 decision tree algorithm was used (Quinlan, 1993).
2. For each variable, v_i , at a node within the classification tree, C , define an opposing membership function pair $L \uparrow_v$ and $L \downarrow_v$ as defined in eq(1) and eq(2).
3. Interpret and implement the classification tree, C as a set of Fuzzy IF-THEN rules to formulate the FLSP knowledge base. Each path from the root of the tree to the leaf node is represented by one rule within the knowledge base and is defined as:

IF x_1 is A_1 AND x_2 is A_2 AND x_n is A_n THEN y is $B_{1..j}$

where x_n are attribute values

A, B are fuzzy sets representing linguistic labels over the universe of discourse, and

y is the consequent real number $y \in \{0,1\}$ (a fuzzy singleton) for outcome j determined from defuzzification of $B_{l,j}$.

Phase 2: Application of fuzzy inference

In order to combine the outcomes of rules within the knowledge base, the antecedents of each individual rule need to be combined together. One technique is the min-max composition developed by Zadeh (1992), which first involves taking the minimum of the membership values of each individual rules by performing a logical AND operation on them. The final stage of the min-max composition is the application of the maximum operator, which takes the consequents from each rule and performs a logical OR by taking the maximum value. Due to the simplicity of the rules, the output is a singleton that will be used to determine the degree to which an individual exhibits traits of a certain learning style dimension. This application of inference can be defined as follows (Crockett et al., 2006):

The process of learning from a set of data, S involves a transformation function F which accepts as input S and produces a defuzzified outcome O , which is a mapping

$$F(S) \rightarrow O \tag{3}$$

Applying an inference technique onto an existing tree consisting of x branches involves the combination of M membership function values $\{\mu_1, \mu_2, \dots, \mu_m\}$ of all root to leaf node paths.

Let T be a set of all possible outcomes $\{t_1, t_2, \dots, t_j\}$ defined from an existing classification tree where j is the total number of outcomes. Fuzzy inference of S will involve an inference mechanism, IM which consists of an intersection function f^\wedge , which takes in M and a t-norm $TMin$, which produces a set of minimum outcomes $\{TMin_1, TMin_2, \dots, TMin_j\}$ where j is the number of leaf nodes, and a union function f^\vee , which combines the output from f^\wedge to produce a maximum membership grade O . Let $f^\wedge, f^\vee, O \in \{0.0,1\}$ consisting of real numbers, \mathfrak{R} .

- i. Apply the fuzzy intersection function, f^\wedge

This involves combining membership grades down all paths, from root to leaves within the tree.

$$f^\wedge (\{\mu_1, \mu_2, \dots, \mu_{b_j}\}) \rightarrow TMin \{TMin_1, TMin_2, \dots, TMin_j\} \tag{4}$$

ii. Apply possibilistic strength of leaf node

The possibilistic strength of a leaf node represents the possibility that an example reaching a leaf node will have the same outcome as the leaf. Let \mathcal{P} be a set of leaf possibilities $\{p_1, p_2, \dots, p_y\}$ where y is the total number of leaf outcomes then

$$f \wedge (\{\mu_1, \mu_2, \dots, \mu_w\}) \rightarrow TMin \{(TMin_1 * p_1), (TMin_2 * p_2), \dots, (TMin_j * p_y)\} \quad (5)$$

Each leaf possibility p_y is applied to the corresponding membership grade at each leaf node $TMin_j$, after the intersection operation has been applied to combine grades of membership down one tree path.

iii. Apply fuzzy union function $f \vee$

The fuzzy union operator is applied in-order to combine the membership grades from all leaf nodes in order to produce a representative final grade of membership.

$$f \vee (\{(TMin_1 * p_1), (TMin_2 * p_2), \dots, (TMin_j * p_y)\}) \rightarrow \mathcal{O} \quad (6)$$

\mathcal{O} is the fuzzy singleton used to determine the success of correct classification having taken place for \mathcal{S} .

iv. Optimise the membership function and inference weightings (if applicable)

A full description of the original version of this algorithm can be found in (Crockett et al., 2006)

3.1.3 Defuzzification Interface

The unique characteristics of a specific classification problem will often determine the defuzzification strategy used. The experiments in this paper have been undertaken using a discrete outcome Learner data set (described in section 4.2) where the outcome is the learning style for a dimension. The output of the application of fuzzy inference is the generation of a fuzzy singleton (section 3.3) which is used to determine this learning style. At any point in time, a student's overall learning style is indicated by comparing the two fuzzy singleton values for each dimension of a given learning style, with the larger (max) value representing their learning style.

3.2 Methodology for generating fuzzy decision trees from conversational tutorials

3.2.1 Study Overview

An empirical study was undertaken in a real teaching/learning environment to evaluate the success of the fuzzy learning styles predictor (FLSP) model when incorporated into the OSCAR-FLSP. The experiment investigated whether the Oscar-FLSP was successful at predicting learning styles by testing the hypothesis:

H1: It is possible to accurately estimate a learner's learning style from a two-way tutoring discourse using a fuzzy logic learning styles predictor model embedded into a conversational intelligent tutoring system.

In order to test this hypothesis, a study was conducted using a conversational tutorial on structured query language (SQL) that was delivered using OSCAR-CITS. Learners were undergraduate computing students within Manchester Metropolitan University who had studied a module on database fundamentals for at least one term, delivered by a human tutor. After anonymous registration, learners completed an online version of the formal ILS questionnaire, recorded in the student model to provide a benchmark classification for comparison of the experimental results. Before starting the conversational tutorial, learners were presented with a pre-tutorial 12 question multiple-choice question (MCQ) test, known as the pre-test, to assess their existing SQL knowledge. The pre-test results were stored in the student model. Next, OSCAR-CITS directed a two-way conversational SQL revision tutorial which took on average 43 minutes, with each participant following an individual learning path depending on their existing knowledge and the dialogue. There were ten main tutorial questions. At the end of the tutorial, participants again completed the same MCQ test (known as the post-test) to assess any learning gain (Fossati et al., 2009), with the results stored in the student model. Next, OSCAR-CITS presented participants with a comparison of their test results (indicating their learning gain) and some feedback on their tutorial performance. Finally, participants were asked to complete a user feedback questionnaire.

3.2.2 Experimental Learner Data Set

75 undergraduate university students who had previous experience of SQL completed the formal ILS questionnaire and the OSCAR-CITS SQL Revision tutorial. During the tutorial their behaviour and test scores were logged. Information relevant to the prediction of learning styles was captured in 42 different behaviour variables during the tutorial. Table 2, available in the appendix, provides a high level description of the variables captured during the conversation. In addition, for each LS dimension the learning style assessed by the ILS questionnaire was stored in the class variable. Each behaviour variable is associated with predicting at least one learning style dimension. For example, one variable known as No_interactions, captured the number of conversational interactions the user had with OSCAR-CITS during the session. Variables Q5_Choice and

Q9_choice, were used to capture a student's approach to writing SQL queries (i.e. whether they try all at once or want to try one step at a time) which is indicative of the sequential-global and sensory-intuitive dimensions. The capturing of each behaviour variable involved individual sets of complex rules, which are defined in detail in (Latham et al., 2012a). Most learning style dimensions had an approximately even split of students, apart from the Visual/Verbal dimension where most students were Visual. This is consistent with the FS model, which suggests that most engineering students are Visual. However, despite this learning style not obtaining a fair representation it has still been included in this study for completeness. This was not unexpected and in accordance with the FS model for students studying science and engineering based subjects. Table 2 includes both the statistical information for all numeric attributes, and quantities for discrete values captured through interactions with all 75 participants.

3.2.3 Experimental Methodology

Using the experimental Learner dataset described in section 4.2, four sets of experiments were undertaken on all four of the dimensions of Felder-Silverman model (Felder and Silverman, 1998). For each experiment set, stratified 10-fold cross validation was used. Following the methodology described in section 3, first crisp decision trees were created from the dataset using the statistical chi-square pruning technique with significance levels of 0.1% in order to produce highly optimised C4.5 trees that would allow good generalisation to take place. Figure 6, shows as an example the highest optimised fuzzy decision tree for the sequential-global learning dimension, generated from 10-fold cross validation. At each leaf the possibilistic strength, p is shown.

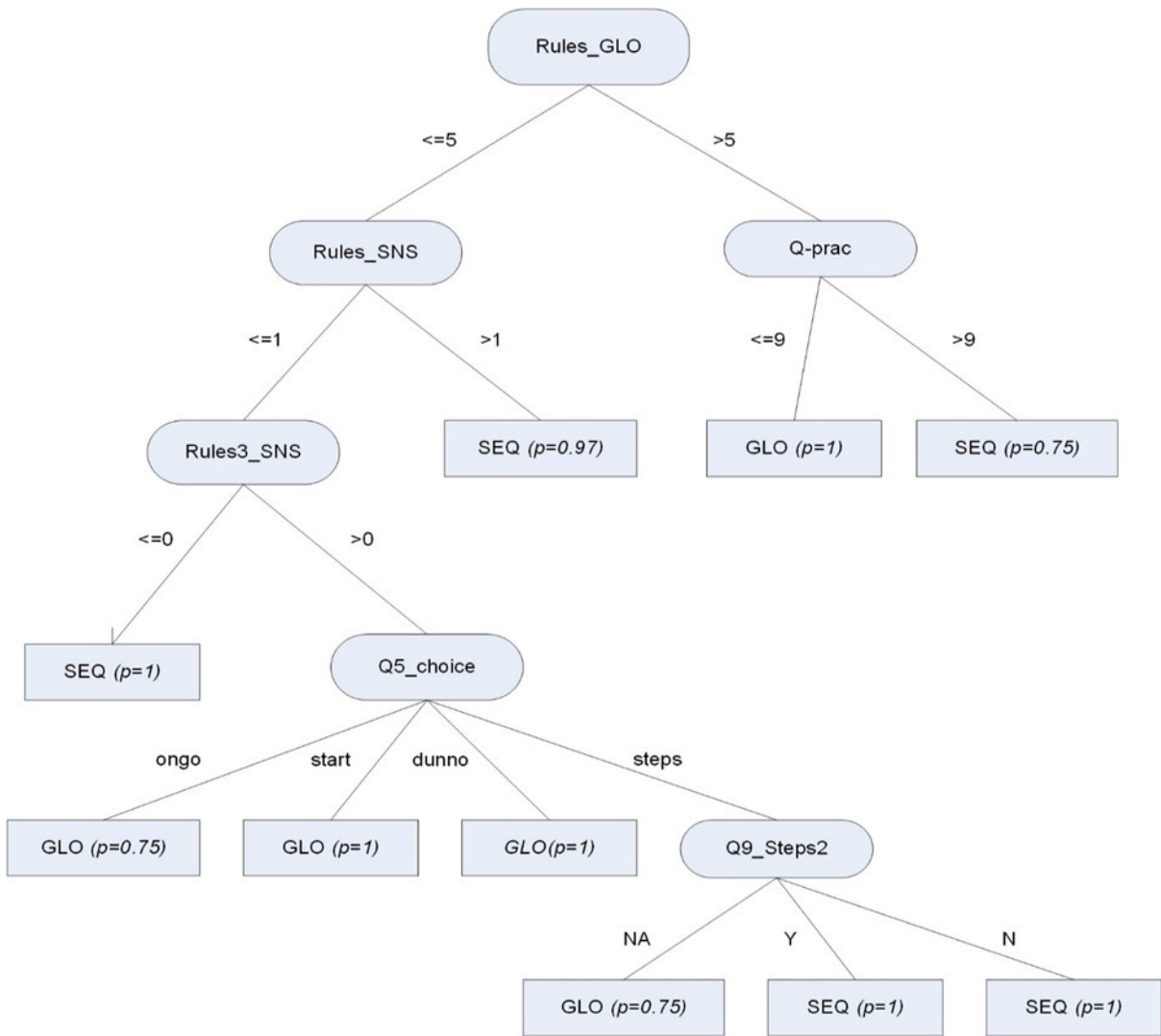


Fig. 6. Fuzzy induced tree representing sequential-global dimension

Fuzzy membership functions were generated for both discrete and continuous attributes. Initial membership functions were selected to be equivalent to crisp sets, with the initial fuzzy classification trees becoming an alternative representation of the C4.5 decision tree. For experiments reported in this paper the amount of fuzzification either side of each tree node was increased linearly from zero fuzzification up to $n = 5$ for each tree i.e. maximum fuzzification per node was 5 times the standard deviation of the attribute as deduced from the training set.

Using a one-to-one mapping, each tree was then transformed into a set of Fuzzy IF-THEN rules where each path from root to leaf node represents one rule. For example, the fuzzy induced tree in Figure 6 would be transformed into 10 fuzzy IF-THEN rules as shown in Figure 7. For numerical behaviour variables, left branches represent values less than or equal to the

partitioning threshold value and are represented by associated fuzzy sets labelled LOW. Subsequently, right branches, greater than the partitioning threshold value, are represented by associated fuzzy sets labelled HIGH. LOW and HIGH for each behaviour variable have unique membership functions. The final fuzzy rule base contained a total of 31 rules, 10 for the sequential-global; 7 for the sensory-intuitive; 8 for active-reflective and 6 for visual-verbal.

1. IF *Rules_GLO* ARE LOW AND *Rules_SNS* ARE LOW AND *Rules3_SNS* ARE LOW THEN learning style is SEQUENTIAL
2. IF *Rules_GLO* ARE LOW AND *Rules_SNS* ARE LOW AND *Rules3_SNS* ARE HIGH AND *Q5_choice* is {ongo} THEN learning style is GLOBAL
3. IF *Rules_GLO* ARE LOW AND *Rules_SNS* ARE LOW AND *Rules3_SNS* ARE HIGH AND *Q5_choice* is {start} THEN learning style is GLOBAL
4. IF *Rules_GLO* ARE LOW AND *Rules_SNS* ARE LOW AND *Rules3_SNS* ARE HIGH AND *Q5_choice* is {dunno} THEN learning style is GLOBAL
5. IF *Rules_GLO* ARE LOW AND *Rules_SNS* ARE LOW AND *Rules3_SNS* ARE HIGH AND *Q5_choice* is {steps} AND *Q9_Steps2* is {NA} THEN learning style is GLOBAL
6. IF *Rules_GLO* ARE LOW AND *Rules_SNS* ARE LOW AND *Rules3_SNS* ARE HIGH AND *Q5_choice* is {steps} AND *Q9_Steps2* is {Y} THEN learning style is SEQUENTIAL
7. IF *Rules_GLO* ARE LOW AND *Rules_SNS* ARE LOW AND *Rules3_SNS* ARE HIGH AND *Q5_choice* is {steps} AND *Q9_Steps2* is {N} THEN learning style is SEQUENTIAL
8. IF *Rules_GLO* ARE LOW AND *Rules_SNS* ARE HIGH THEN learning style is SEQUENTIAL
9. IF *Rules_GLO* ARE HIGH AND *Q-prac* ARE LOW THEN learning style is GLOBAL
10. IF *Rules_GLO* ARE HIGH AND *Q-prac* ARE HIGH THEN learning style is SEQUENTIAL

Fig. 7. Fuzzy IF-THEN rules for the sequential-global Learning Dimension.

To illustrate the application of fuzzy inference (phase 2), consider the fuzzy IF-THEN rules in figure 7 which have been generated from the fuzzy tree modelling the sequential and global dimension in figure 6.

Given the following participant case passing through the tree where *Rules_GLO* = 5, *Rules_SNS* = 1, *Rules3_SNS* = 1, *Q-prac* = 0.4 and *Q5_choice* = start. The classification outcome can be calculated by generating the grades of membership for each fuzzy set for the associated variable values i.e. *Rules_SNS* = 1 has a fuzzy set membership of 0.78 in *Rules_SNS are low* fuzzy set and then applying fuzzy inference as follows:

$$R1: \text{Min} \{0.5, 0.78, 0\} * 1 = 0$$

$$R2: \text{Min} \{0.5, 0.78, 1.2, 0\} * 0.75 = 0$$

$$R3: \text{Min} (0.5, 0.78, 1.2, 1) * 1 = 1.2$$

$$R4: \text{Min} \{0.5, 0.78, 1.2, 0\} * 1 = 0$$

$$R5: \text{Min} \{0.5, 0.78, 1.2, 0, 0\} * 0.75 = 0$$

$$R6: \text{Min} \{0.5, 0.78, 1.2, 0, 1\} * 1 = 0$$

$$R7: \text{Min} \{0.5, 0.78, 1.2, 0, 0\} * 1 = 0$$

$$R8: \text{Min} \{0.5, 0.23\} * 1 = 0.23$$

$$R9: \text{Min} \{0.5, 1\} * 1 = 0.5$$

$$R_{10}: \text{Min} \{0.5, 0\} * 0.75 = 0$$

$$\text{Class Outcome} = \text{Max} \{R_1..R_{10}\} = 0.5 (R_9) = \text{Global}.$$

The participant would be classified in this case as a global learner.

4. Results and Discussion

4.1 Predictive Results per Learning Dimension

Tables 3 to 6 present the prediction accuracies of all learning style dimensions where the number of participants, $m = 75$. In each Table the average results of 10 fold cross validation are shown. For the sequential-global dimension (Table 3), columns %AVG SEQ and %AVG GLO show the percentage accuracy of predicting sequential and global learners respectively, with the overall prediction accuracy for the sequential-global learning dimension shown in the %AVG SEQ/GLO column. Tables 4 to 6 present the results for the other dimensions in a similar way. DT (C4.5) provides the predictive accuracy of a crisp decision tree model where multiple behaviour variables are combined based upon their effect on the outcome, but uncertainty is not taken into consideration. The FLSP rows show the results obtained by using fuzzy classification trees within the FLSP model. The OSCAR-CITS row represents and names the best behaviour variable predictor used by the original OSCAR-CITS to determine the learning style. For both sequential-global and sensory-intuitive dimensions, this variable was Q9_choice, which essentially records whether the learner chooses to approach a complex question in one go or not. For active-reflective, OSCAR-CITS, the key predictive variable was Q-style which recorded whether the participant was more successful at practical or theoretical questions. For visual-verbal, OSCAR-CITS, the key predictive variable was mean_read_time, which records whether the participant's average reading time during the tutorial was longer or shorter than the mean reading time across the group.

Table 3 Sequential - Global

Method	%AVG	% AVG	% AVG
<i>m=75</i>	SEQ	GLO	SEQ/GLO
OSCAR-CITS (Q9_choice)	70	59	65
DT (C4.5)	80	70	76
OSCAR-FLSP	93	83	88

Table 4 Sensory-Intuitive

Method	%AVG	% AVG	% AVG
<i>m=75</i>	SEN	INT	SEN/INT
OSCAR-CITS (Q9_choice)	70	59	65
DT (C4.5)	84	80	82
OSCAR-FLSP	98	73	86

Table 5 Visual-Verbal

Method	%AVG	% AVG	% AVG
<i>m=75</i>	VIS	VER	VIS/VER
OSCAR-CITS (Mean-read_time)	72	50	61
DT (C4.5)	95	50	73
OSCAR-FLSP	100	70	85

Table 6 Active-Reflective

Method	%AVG	% AVG	% AVG
<i>m=75</i>	ACT	REF	ACT-REF
OSCAR-CITS (Q-style)	53	73	63
DT (C4.5)	69	77	73
OSCAR-FLSP	74	94	84

Table 7 Learning styles prediction results summary and comparison of best prediction accuracy of learning styles.

	Sensory	Intuitive	Visual	Verbal	Active	Reflective	Sequential	Global
<i>Prior Probability</i>	60%	40%	87%	13%	57%	43%	60%	40%
<i>OSCAR-FLSP</i>	98%	73%	100%	70%	74%	94%	93%	84%
<i>OSCAR-CITS</i>	70%	59%	72%	50%	53%	73%	70%	59%
+/-	+28%	+14%	+28%	+20%	+21%	+21%	+23%	+25%

4.2 Predictive Ability

In discussing the significance of the results in Tables 3 to 6, the prior probability will be used as a fair comparison instead of chance (50%) due to the uneven spread of learning styles across the sample. Prior probability is based on the distribution of learning styles across the sample. On first observation, the results in Table 7 indicate that the introduction of the FLSP models has predicted learning style better than the prior probability. When a single predictive behaviour variable is used, OSCAR-CITS also predicts learning style with higher accuracy across all dimensions except visual and active. What is noticeable is the increase in the prediction accuracy when the FLSP is used, which ranged from 14% to 28%. A test for the difference of two proportions (Dietterich, 1998) was used to measure the difference between the error rate of OSCAR-CITS and OSCAR-FLSP. Low p-values mean the difference between the two is statistically significant. The results indicated a significant difference in the predictive accuracy between OSCAR-CITS and OSCAR-FLSP dependent on learning style dimension; sensory-intuitive ($p < 0.01$), sequential-global ($p < 0.05$), active-reflective ($p < 0.05$) and visual-verbal ($p < 0.05$). Thus, the results have indicated that H1 is true.

The results in Tables 3 and 4 show that for sequential-global and sensory-intuitive dimensions, when only one variable is used (OSCAR); the same variable (*Q9_choice*) was identified as being the most significant in terms of prediction accuracy. An initial observation of the results shows that a correlation might exist between the two learning dimensions as measured by this particular question as sequential people typically like to learn step by step and sensory people prefer to go slowly and carefully, whereas global people typically like to jump straight in and intuitive people may be more fast and careless. The results in Table 3 for OSCAR show that 70% of students who chose to answer the tutorial Question 9 one step at a time are Sequential learners. Using this variable (*Q9_choice*) OSCAR-CITS would predict that students in this group are Sequential learners and be accurate 70% of the time. However, if the FLSP model had been integrated, where 6 behaviour variables (as shown in the tree in Figure 5) were used in combination to determine Sequential learners, OSCAR-CITS would have predicted correctly 93% of the time – a 33% improvement.

The results of inducing fuzzy trees yielded some interesting results with regards to the models produced (e.g. Figure 5). In modelling the sequential-global dimension, the attribute *Rules_GLO* (representing the number of global related rules that had fired during the conversation) was seen to be the most significant detector of learning style. However, what was surprising was the number of sensory attributes (*Rules_SNS* and *Rules3_SNS*) that were also included in the model. This again supported the correlation between sequential-global and sensory-intuitive dimensions.

The results for the visual-verbal dimension in Table 5 are not conclusive as only 13% of the students captured were classified as verbal learners. The uneven distribution of undergraduate learners on the visual-verbal FS dimension, as

noted in the FS model, leads to consideration of the question whether it is important for a CITS to classify this dimension and therefore if adopting the FS model the importance of predicting visual-verbal learning styles should be considered carefully.

A comparison of predictive results directly with other current CITS is not possible as they do not predict learning styles. Work has been done using menu-based tutoring systems embodied within ITS that predict learning styles at a slight lower level than OSCAR-CITS using the winner takes all strategy (Ozpolat and Akar et al., 2009) and thus the results presented from the FLSP compare favourably. However, it must be noted that in (Ozpolat and Akar et al., 2009) classification is determined through the addition of a neutral class for each dimension which is a different approach to that presented in this paper.

4.3 Implications for CITS Design

Regardless of the debates surrounding learning style theory and practice as a whole there is considerable support in the educational community for the development of such models as they have been shown to increase motivation of students (Ghaedi and Jam, 2014), improve learning gain and even reduce learning time (Coffield et al., 2004). Key challenges lie in the identification of learning styles, avoiding the weaknesses of traditional paper-based approaches, and in making meaningful adaptations to a learning environment based on such classifications. It also raises the pedagogic question of whether dynamic style detection and automatic personalisation is a preferable approach to learner education and choice in relation to learning styles. Automatic detection of learning styles within a learning system such as a CITS can be achieved over a period of time (such as a tutorial) and is more accurate in relation to a specific learner when using live student data. However, employing a computational intelligence technique (such as fuzzy decision trees) that allows models to be built from the behaviour of learners using multiple pieces of information captured during the tutorial substantially increases the predictive accuracy.

The FLSP solves the problem of predicting a user's learning style using only the highest ranking single behaviour variable. Adapting to learning styles within an CITS tutorial has been shown to significantly improve the learning experience (Latham et al., 2014) but this is highly dependent on the accuracy of the student learning style model. The FLSP has produced four models – one for each learning style dimension, which have learned combinations of behaviour variables that result in a more accurate and comprehensive overall learning styles model. The next stage is to embed these models into the adaptive OSCAR-CITS and conduct a further study to assess the impact on learner performance.

The methodology and architecture for developing a CITS with a Fuzzy Learning Styles Predictor Module are independent of the learning styles model and subject domain, and can be adapted. The FLSP creates a rule base that is readable by the teacher and ensures decision transparency in how the system determined the learning style. The introduction of fuzzification for each behaviour variable allows the model to deal with the uncertainty in trying to automatically detect the correct dimension

through natural language conversation. Consequently, this makes the FLSP much less sensitive to noise within natural language.

5. Conclusion and Future Work

The motivation of this work stemmed from the uncertainty in predicting a person's learning style for the purpose of CITS tutorial delivery. The current OSCAR-CITS predicts the learning style through one behaviour variable captured during the tutoring conversation – typically the one that scored the highest. Whilst this prediction was good, it was not based on the combination of behaviour exhibited by each person during the tutorial and the learner was on occasion being presented with material not suited to their learning style.

This paper has proposed a new method that uses fuzzy decision trees to build a series of fuzzy learning style predictive (FLSP) models using behaviour variables captured from natural language within a CITS tutorial for the four dimensions of the Felder and Silverman Learning Styles model. Seventy-five undergraduate students completed a formal ILS questionnaire and a tutorial in SQL using the OSCAR-CITS as part of their revision studies. During the tutorial, their behaviour was logged in 41 different behaviour variables that were used in this study to build the Learner data set. Experiments were carried out for all four learning style dimensions: perception (sensory-intuitive), understanding (sequential-global), input (visual-verbal) and processing (active, reflective) to build a set of four fuzzy decision trees, each with the ability to predict a dimension with a degree of possibility. The results have indicated that the use of fuzzy predictive models has substantially increased the predictive accuracy of the OSCAR-FLSP compared with the one variable predictor currently used. The prediction accuracy improvement across all dimensions ranged from 14% to 28% and a significant difference in the predictive accuracy between OSCAR-CITS and OSCAR-FLSP was obtained for all learning dimensions as indicated by the test for the difference of two proportions: sensory-intuitive ($p < 0.01$), sequential-global ($p < 0.05$), active-reflective ($p < 0.05$) and visual-verbal ($p < 0.05$).

The new Fuzzy Learning Styles Predictor module is a part of the intelligent background processing in the new OSCAR-CITS architecture (Figure 4). Therefore, it is hidden from the user. In modelling the student learning style with more accuracy, the expectation is that the improved model will deliver better adaptation decisions that should impact positively on the usability of the system. However, it is not possible to demonstrate this until the fuzzy predictive models are incorporated into the adaptive OSCAR-CITS in future work.

Fuzzy decision trees are one of a number of computational intelligence techniques that learn from data to build models. However, decision trees are most suited to handling uncertainty and imprecision that is inherent in humans. The models that have been induced in this work have created a transparent set of rules where it is clear to see which behaviour variables impact most on the predictive accuracy for a given learning dimension. A model, however is only as good at the data and therefore

further experiments would be needed using different learners and different CITS tutorials. Further work will incorporate the fuzzy predictive models into the adaptive OSCAR-CITS (Latham et al., 2014) in order to provide an improved and more personalised learning experience.

ACKNOWLEDGEMENTS

The authors thank ConvAgent Limited for the use of their InfoChat Conversational Agent and PatternScript scripting language.

APPENDIX

Table 2. Learner Dataset Description

Attribute No	FIELD	DESCRIPTION	TYPE	VALUES / STATS
1	Rules_VIS	Contains score for Visual learning style dimension resulting from OSCAR-CITS logic rule	Numeric	Min: 1 / Max: 11 Mean: 6.733 StdDev: 2.767
2	Rules_VRB	Contains score for Verbal learning style dimension resulting from OSCAR-CITS logic rule.	Numeric	Min: 0 / Max: 11 Mean: 3.4 StdDev: 2.427
3	Rules3_VIS	Contains score for Visual learning style dimension resulting from OSCAR-CITS extended logic rule - includes whether the participant answered question correctly.	Numeric	Min: 0 / Max: 16 Mean: 4.413 StdDev: 2.881
4	Rules3_VRB	Contains score for Verbal learning style dimension resulting from OSCAR-CITS extended logic rule – includes whether the participant answered question correctly.	Numeric	Min: 0 / Max: 12 Mean: 3.053 StdDev: 2.117
5	Rules_SEQ	Contains score for Sequential learning style dimension resulting from OSCAR-CITS logic rule.	Numeric	Min: 1 / Max: 11 Mean: 6.2 StdDev: 2.307
6	Rules_GLO	Contains score for Global learning style dimension resulting from OSCAR-CITS logic rule.	Numeric	Min: 1 / Max: 10 Mean: 4.933 StdDev: 1.848
7	Rules3_SEQ	Contains score for Sequential learning style dimension resulting from OSCAR-CITS extended logic rule – only applied if participant gave the correct answer.	Numeric	Min: 0 / Max: 10 Mean: 6.067 StdDev: 2.062
8	Rules3_GLO	Contains score for Global learning style dimension resulting from OSCAR-CITS extended logic rule – only applied if participant gave the correct answer.	Numeric	Min: 0 / Max: 8 Mean: 4.347 StdDev: 1.573
9	Q5_choice	The choice the learner makes in response to a question on how to approach a complex question (Question 5 in the tutorial) – in one go (ONEGO) or steps (STEPS), or just starts solving the problem (START).	Discrete	ONEGO: 9 START: 5 STEPS: 16 steps: 19 onego: 9 start: 15 dunno: 2

10	Q5_STEPS2	Extends Q5_choice, by reporting the choice the learner made in how to approach a complex question and whether the question was then answered correctly.	Discrete	Y: 49 NA: 16 N: 10
11	Q9_choice	The choice the learner makes in how to approach a complex question (Question 9 in the tutorial) – in one go or steps.	Discrete	ONEGO: 10 START: 6 STEPS: 13 DUNNO: 1 steps: 17 onego: 7 start: 16 dunno: 5
12	Q9_STEPS2	Extends Q9-choice, by reporting the choice the learner made in how to approach a complex question and whether the question was then answered correctly.	Discrete	NA: 14 Y: 54 N: 7
13	Q4_detail	Records whether a learner got the answer right when the answer was given in the question explanation text.	Discrete	Y:60 N: 15
14	Q_prac	Indicate whether the student performs better during the tutorial on practical questions.	Numeric	Min: 0.2 /Max: 1 Mean: 0.673 StdDev: 0.174
15	Q_theo	Indicates whether the student performs better during the tutorial on theoretical questions.	Numeric	Min: 0.13/ Max: 1 Mean: 0.678 StdDev: 0.183
16	Q_style	Records student performance on Q_prac and Q_theo. (Practical (P), Theory (T), Same on both(S))	Discrete	P: 31 S: 5 T: 39
17	Keyword_example	Records the number of times particular keyword “example” was uttered by the learner during the tutorial.	Numeric	Min: 0 / Max: 3 Mean: 0.04 StdDev:0.346
18	Keyword_show	Records the number of times particular keyword “show” was uttered by the learner during the tutorial.	Numeric	Min: 0 /Max: 3 Mean: 0.627 StdDev:0.749
19	Keyword_see	Records the number of times particular keyword “see” was uttered by the learner during the tutorial.	Numeric	Min: 0 / Max: 1 Mean: 0.013 StdDev:0.115
20	Keyword_tell	Records the number of times particular keyword “tell” was uttered by the learner during the tutorial.	Numeric	Min: 0 / Max: 2 Mean: 0.04 StdDev:0.257
21	Keyword_picture	Records the number of times particular keyword “picture” was uttered by the learner during the tutorial	Numeric	Min: 0 / Max: 1 Mean: 0.013 StdDev: 0.115
22	Keyword_discuss	Records the number of times particular keyword “discuss” was uttered by the learner during the tutorial.	Numeric	Min: 0 / Max: 2 Mean: 0.027 StdDev: 0.231
23	Rules_ACT	Contains score for Active learning style dimension resulting from OSCAR-CITS logic rule.	Numeric	Min: 0 / Max: 11 Mean: 5.52 StdDev: 2.549
24	Rules_REF	Contains score for Reflective learning style dimension resulting from OSCAR-CITS logic rule.	Numeric	Min: 0 / Max: 11 Mean: 3.013 StdDev: 2.883
25	Rules3_ACT	Contains score for Active learning style dimension resulting from OSCAR-CITS extended logic rule – only applied if participant gave the correct answer.	Numeric	Min: 1 / Max: 10 Mean: 3.907 StdDev: 1.612
26	Rules3_REF	Contains score for Reflective learning style dimension resulting from OSCAR-CITS extended logic rule – only applied if participant gave the correct answer.	Numeric	Min: 0 / Max: 3 Mean: 0.267 StdDev: 0.6
27	Rules_SNS	Contains score for Sensory learning style dimension resulting from OSCAR-CITS logic rule.	Numeric	Min: 0 / Max: 11 Mean: 4.72

				StdDev: 3.451
28	Rules_INT	Contains score for Intuitive learning style dimension resulting from OSCAR-CITS logic rule.	Numeric	Min: 0 / Max: 11 Mean: 4.52 StdDev: 2.811
29	Rules3_SNS	Contains score for Sensory learning style dimension resulting from OSCAR-CITS extended logic rule - only applied if participant gave the correct answer.	Numeric	Min: 0 / Max: 4 Mean: 1.307 StdDev: 0.822
30	Rules3_INT	Contains score for Intuitive learning style dimension resulting from OSCAR-CITS extended logic rule – only applied if participant gave the correct answer.	Numeric	Min: 1 / Max: 14 Mean: 4.293 StdDev: 2.358
31	MCQ_theory	Indicates whether the student performs better in the MCQ on questions related to a particular style of tutorial question: theoretical or practical.	Numeric	Min: 0 / Max: 1 Mean: 0.524 StdDev: 0.37
32	MCQ_pract	Indicates whether the student performs better in the MCQ on questions related to a particular style of tutorial question: theoretical or practical.	Numeric	Min: 0 / Max: 1 Mean: 0.532 StdDev: 0.373
33	Duration	Records the duration of the tutorial in seconds.	Numeric	Min: 845 / Max: 14527 Mean: 2445 StdDev: 1638
34	Mean_Duration	Records whether the duration was longer or shorter than the mean durations across the study group.	Discrete	Min: 845 Max: 14527 Mean: 2445 StdDev: 1638
35	Time_per_int	Records the time per interaction.	Numeric	Min: 18.322 Max: 415.057 Mean: 52.339 StdDev: 45.906
36	Mean_time_int	Records whether the duration per interaction was longer or shorter than the mean durations across the group.	Discrete	ABOVE: 25 BELOW: 50
37	Read_time	Contains the mean reading time for 10 words in seconds.	Numeric	Min: 4.305 Max: 102.303 Mean: 12.23 StdDev: 11.295
38	MeanRdTime	Records whether the reading time was longer or shorter than the mean reading time across the group (in seconds).	Discrete	ABOVE: 22 BELOW: 53
39	Wordcount	Contains the learner's word count for the tutorial.	Numeric	Min: 67 / Max: 352 Mean: 164.027 StdDev: 68.666
40	Mean_Wordcount	Records whether the word count was longer or shorter than the mean word count across the group.	Discrete	BELOW: 44 ABOVE: 31
41	No_interactions	Contains the number of interactions (Dialogue exchanges).	Numeric	Min: 24 / Max: 84 Mean: 48.84 StdDev: 9.895
42	Mean_interactions	Records whether the number of interactions was longer or shorter than the mean interactions across the group.	Discrete	BELOW: 39 ABOVE: 36
43	Class	Contains the learner's learning style class given by the ILS results. ACTIVE (ACT), REFLECTIVE (REF), GLOBAL (GLO), SEQUENTIAL (SEQ), VISUAL (VIS), VERBAL (VRB), INTUTIVE (INT), SENSORY (SNS).	Discrete	ACT: 43 REF: 32 GLO: 30 SEQ: 45 VIS: 65 VRB: 10 INT: 30 SNS: 45

REFERENCES

- Arce-C'rdenas, F.; Garcia-Valdez, M. 2012. Learning Objects for Intelligent Environments, 8th International Conference on Intelligent Environments, pp.351-354.
- Ammar, M.B., Neji, M., Alimi, A.M., Gouarderes, G. 2010. The Affective Tutoring System. *Expert Systems with Applications* 37, pp.3013–302.
- Aparicio, F Rodríguez, M. Rubio, M. Hernando, A, 2012. An intelligent information access system assisting a case based learning methodology evaluated in higher education with medical students. *Computers & Education* 58(4), pp.1282-1295.
- Azevedo, R. Witherspoon, A. Chauncey, A. Burkett, C. Fike, A. Brusilovsky, P., Peylo, C. 2003. MetaTutor: A MetaCognitive Tool for Enhancing Self-Regulated Learning, *Adaptive and Intelligent Web-based Educational Systems*. *Int.J. Artificial Intelligence in Education* 13, pp.156–169.
- Brusilovsky, P. and Peylo, C. 2003 Adaptive and intelligent Web-based educational systems. In P. Brusilovsky and C. Peylo (eds.), *International Journal of Artificial Intelligence in Education* 13 (2-4), Special Issue on Adaptive and Intelligent Web-based Educational Systems, pp.159-172.
- Cabada, R.Z.; Estrada, M.L.B.; Garcia, C.A.R. 2009. A Fuzzy-Neural Network for Classifying Learning Styles in a Web 2.0 and Mobile Learning Environment, *Web Congress, 2009. LE-WEB '09. Latin American*, DOI: 10.1109/LA-WEB.2009.18. pp.177-182.
- Cha, H.J., Kim, Y.S., Park, S.H., Yoon, T.B., Jung, Y.M., Lee, J.H. 2006. Learning styles diagnosis based on user interface behaviours for the customization of learning interfaces in an intelligent tutoring system. In: Ikeda, M., Ashley, K.D., Chan, T.-W. (eds.) *ITS 2006*, Springer LNCS, (4053), pp.513–524.
- Chen, S. & Liu, X. 2008. An Integrated Approach for Modeling Learning Patterns of Students in Web-Based Instruction: A Cognitive Style Perspective. *ACM Transactions on Computer Human Interaction*, 15(1).
- Chih-Yueh Chou, Huang, B, H., & Lin, C. J. 2011. Complementary Machine Intelligence and Human Intelligence in Virtual Teaching Assistant for Tutoring Program Tracing. *Computers & Education* 57(4), pp.2303-2312.
- Chenn-Jung Huang, Ming-Chou Liu, San-Shine Chu, Chih-Lun Cheng. 2007. An intelligent learning diagnosis system for Web-based thematic learning platform. *Computers & Education* 48(4), pp.658-679.

Coffield, F. Moseley, D., Hall, E., & Ecclestone, K. 2004. Should We Be Using Learning Styles? What Research Has to Say to Practice. Learning and Skills Research Centre, London. ISBN 1853389145

Crockett, K, O'Shea, J, Bandar, Z. 2011. Goal Orientated Conversational Agents: Applications to Benefit Society. Springer-Verlag, Lecture notes in computer science, Agent and Multi-Agent Systems: Technologies and Applications, pp.16-25.

Crockett, K., Bandar, Z., O'Shea, J. Mclean, D. 2006. On Constructing a Fuzzy Inference Framework using Crisp Decision Trees. Fuzzy Sets and Systems 157(21), pp.2809-2832.

Crockett, K, Latham, A, Mclean, D, O'Shea, J. 2013. A fuzzy model for predicting learning styles using behavioral cues in an conversational intelligent tutoring system, 2013 IEEE International Conference on Fuzzy Systems, DOI: 10.1109/FUZZ-IEEE.2013.6622382, pp.1-8.

Gasir, F. Crockett, K, Bandar, Z. 2012. Inducing Fuzzy Regression Tree Forests Using Artificial Immune Systems. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems (IJUFKS), Print ISSN: 0218-4885, Online ISSN: 1793-6411, 20(133).

D'Mello, S. K. & Graesser, A. C. 2012a. AutoTutor and affective AutoTutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. ACM Transactions on Interactive Intelligent Systems, 2(4), 23.

D'Mello, C., & Graessner, A. 2012b. Dynamics of affective states during complex learning. Learning and Instruction, 22(2), pp.145-157.

D'Mello, S. K., Lehman, B. Pekrun, R., & Graesser, A. C. 2014. Confusion Can be Beneficial For Learning, Learning & Instruction, 29(1), pp.153-170.

G. Dietterich, 1998. Approximate statistical tests for comparing supervised classification learning algorithms, Neural Computing., vol. 10, no. 7, pp.1895–1924.

D'Mello, S. K., & Kory, J. 2015. A Review and Meta-Analysis of Multimodal Affect Detection Systems, ACM Computing Surveys,47(3), 43 pp.1-43:46

Dorça,F. Lima, L. Fernandes, M. Lopes, C. 2013. Comparing strategies for modeling students learning styles through reinforcement learning in adaptive and intelligent educational systems: An experimental analysis, Expert Systems with Applications, 40(6), pp.2092-2101.

Fazlollahtabar H.,Mahdavi, I. 2009. User/tutor optimal learning path in e-learning using comprehensive neuro-fuzzy approach, Educational Research Review, 4(2), pp.142-155.

Felder, R., Silverman, L.K. 1998. Learning and Teaching Styles in Engineering Education. Journal of Engineering Education, 78(7), pp.674–68.

- Felder, R. M. and Henriques, E.R. 1995. Learning and Teaching Styles in Foreign and Second Language Education, *Foreign Language Annals*, vol. 28 (1), pp. 21–31.
- Feldman, J. Monteserin, A. Amandi, A. 2014 Automatic detection of learning styles: state of the art, *Artificial Intelligence Review*, Springer, DOI: 10.1007/s10462-014-9422.
- Fossati, D., Di Eugenio, B., Brown, C.W., Ohlsson, S., Cosejo, D.G. & Chen, L. 2009. Supporting Computer Science Curriculum: Exploring and Learning Linked Lists with iList. *IEEE Trans. Learning Technologies* 2 (2), pp.107-120.
- Gardner H. 1993. *Frames of mind*, 2nd ed. New York: Basic Books.
- Garcia, P., Amandi, A., Schiaffino, S., Campo, M. 2007. Evaluating Bayesian Networks' Precision For Detecting Students' Learning Styles. *Computers & Education*, 49(3), pp.794-808.
- Garcia-Valdez, M. Alanis, A. Parra, B. 2010. Fuzzy inference for Learning Object Recommendation, *IEEE International Conference on Fuzzy Systems*, DOI:10.1109/FUZZY.2010.5584322, pp.1-6.
- Ghaedi, Z. Jam, B. 2014. Relationship between Learning Styles and Motivation for Higher Education in EFL Students. *Theory and Practice in Language Studies*, Vol. 4, No. 6, pp. 1232-1237.
- Graesser, A., Chipman, P., Haynes, B.C., Olney, A. 2005a. AutoTutor: An Intelligent Tutoring System With Mixed-Initiative Dialogue. *IEEE Transactions on Education*, 48(4), pp.612–618.
- Graesser, A.C., Person, N., Lu, Z., Jeon, M.G. McDaniel, B. 2005b. Learning While Holding a Conversation With a Computer. *Technology-Based Education: Bringing Researchers and Practitioners Together*. Information Age Publishing Inc, pp.143-167.
- Hawkes, L. Derry, S. 1996. Advances in local student modeling using informal fuzzy reasoning. *International Journal of Human-Computer Studies*, DOI: doi:10.1006/ijhc.1996.0075, 45(6).
- Hsieh, Tung-Cheng, Tzone-I. Wang, Chien-Yuan Su, and Ming-Che Lee. 2012. A Fuzzy Logic-based Personalized Learning System for Supporting Adaptive English Learning. *Educational Technology & Society*, 15(1), pp.273-288.
- Jovanović, J. Gašević, D. 2012. Student modeling and assessment in intelligent tutoring of software patterns, *Expert Systems with Applications*, 39(1), pp.210-222.
- Klašnja-Milićević, A. Vesin, B. Ivanović, M. Budimac, Z. 2011. E-Learning personalization based on hybrid recommendation strategy and learning style identification, *Computers & Education*, 56(3), pp.885-899.
- Kelly, D. & Tangney, B. 2006. Adapting to intelligence profile in an adaptive educational system. *Interacting with Computers*, 18, pp.385-409.
- Kumar, R. Ai, A. Beuth, J. Rosé, C. 2010. Socially Capable Conversational Tutors Can Be Effective in Collaborative Learning Situations Conference: Intelligent Tutoring Systems, 10th International Conference, ITS 2010, Pittsburgh, PA, USA, June 14-18, 2010, Proceedings, Part I. DOI: 10.1007/978-3-642-13388-6_20, pp.156-164.

- Kelly, D. Tangney, N. 2006. Adapting to intelligence profile in an adaptive educational system, *Interacting with Computers*, 18, pp.385-409.
- Latham, A 2011. PhD Thesis: Towards Socially-Intelligent Agents, Manchester Metropolitan University.
- Latham, A. Crockett, K. Mclean, D. Edmonds, B 2012a, A conversational intelligent tutoring system to automatically predict learning styles
Computers & Education, 59(1), pp.95-109.
- Latham, A., Crockett, K., McLean, D. Edmonds, B., 2012b. Adaptive Tutoring in an Intelligent Conversational Agent System, *Transactions on Computational Collective Intelligence* vol. VIII, Springer-Verlag, 7430, pp.148-167.
- Latham, A. Crockett, K. Mclean, D. 2014, An adaptation algorithm for an intelligent natural language tutoring system, *Computers & Education* 71, pp.97-110.
- Chun Fu Lin, Yu-chu Yeh, Yu Hsin Hung, Ray I Chang, 2013. Data mining for providing a personalized learning path in creativity: An application of decision trees, *Computers & Education*, 68, pp.199-210.
- Liu, X. Feng, X Pedrycz W. 2013. Extraction of fuzzy rules from fuzzy decision trees: An axiomatic fuzzy sets (AFS) approach, *Data & Knowledge Engineering*, 84, pp.1-25.
- Leontidis, M., Halatsis, C. 2009. Integrating Learning Styles and Personality Traits into an Affective Model to Support Learner's Learning. Spaniol, M., et al. (eds.) *Advances in Web Based Learning – ICWL 2009*, Springer LNCS, 5686, pp.225–34.
- Jia-Jiunn Lo, Ya-Chen Chan, Shiou-Wen Yeh. 2012. Designing an adaptive web-based learning system based on students' cognitive styles identified online. *Computers & Education* 58(1), pp.209-222.
- Marsala, C., Fuzzy decision trees for dynamic data. 2013. 2013 IEEE Conference on Evolving and Adaptive Intelligent Systems, pp.17-24.
- Mitrovic, A. An Intelligent SQL Tutor on the Web. 2003. *International Journal Artificial Intelligence in Education*, 13, pp.171–195.
- Nykänen, Ossi. 2006. Inducing fuzzy models for student classification. *Educational Technology & Society*, 9(2), pp.223-234.
- Nye, B.D., Graesser, A.C., Hu, X. 2014. AutoTutor and family: A review of 17 years of natural language tutoring. *International Journal of Artificial Intelligence in Education*, 24(4), 427-469. doi:10.1007/s40593-014-0029-5
- Ozpolat E., Akar G. B., 2009. Automatic detection of learning styles for an e-learning system, *Computers & Education*, Elsevier, 53, pp.355-367.
- O'Shea, J. Bandar, Z. Crockett, K. 2011. *Systems Engineering and Conversational Agents*, A. Tolk and L.C. Jain (Eds.): *Intelligence-Based Systems Engineering*, Intelligent Systems Reference Library 10, ISBN 978-3-642-17930-3.

- Papanikolaou, K.A., Grigoriadou, M., Kornilakis, H., Magoulas, G.D. 2003. Personalizing the Interaction in a Web-based Educational Hypermedia System: the case of INSPIRE. *User Modeling and User-Adapted Interaction*, 13, pp.213–267.
- Paviotti, G., Rossi, P.G., & Zarka, D. 2012. *Intelligent Tutoring Systems: An Overview*. Lecce: Pensa Multimedia, Italy, ISBN 9788867600489.
- Piaget, J. Piaget's Theory of Constructivism. TeAchnology.Teachnology, Inc. 2011. Web. 1 Feb. 2011. Available: <<http://www.teach-nology.com/currenttrends/constructivism/piaget/>> Date Accessed: 1/12/2015.
- Popescu, E.: An Artificial Intelligence Course Used to Investigate Students' Learning Style. 2008 In: Li, F., Zhao, J., Shih, T.K., Lau, R., Li, Q., McLeod, D. (eds.) Springer LNCS, 5145, pp.122–131.
- Popescu, E. 2009. Diagnosing Students' Learning Style in an Educational Hypermedia System, in: Constantinos Mourlas, Nikos Tsianos and Panagiotis Germanakos (Eds): *Cognitive and Emotional Processes in Web-based Education: Integrating Human Factors and Personalization*, Advances in Web-Based Learning Book Series, IGI Global, pp.187-208.
- Popescu, E. 2010. Adaptation provisioning with respect to learning styles in a Web-based educational system: an experimental study. *Journal of Computer Assisted Learning*, 26, pp.243-257
- Porayska-Pomstaa, P. Mellish, C. 2013. Modelling human tutors' feedback to inform natural language interfaces for learning, *International Journal of Human-Computer Studies*, 71 (6), pp.703–724
- Quinlan, J, R. 1993. *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers.
- Reyes, J. Navarro-Almanza, R. Gomez-Tagle, Y. Licea, G. Huertas, C. Quinto, G. 2013. Orchestrating an Adaptive Intelligent Tutoring System: Towards Integrating the User Profile for Learning Improvement, *Procedia - Social and Behavioral Sciences*, ISSN 1877-0428, <http://dx.doi.org/10.1016/j.sbspro.2013.12.227>, 106, pp.1986-1999.
- Rus, V., Baggett, W., Gire, E., Franceschetti, D., Conley, M., Graesser, A.C. 2013. Towards Learner Models based on Learning Progressions in DeepTutor, *Learner Models* (Eds. Sottolare, R.), Army Research Lab.
- Sani, S. Aris, T. 2014. Computational Intelligence Approaches for Student/Tutor Modeling: A Review, 2014th International Conference on Intelligent Systems, and Simulation, DOI 10.1109/ISMS.2014.21, IEEE Computer Society.
- Sharp, J. G., Byrne, J., & Bowker, R. 2008. The trouble with VAK. *Educational Futures*, 1(1), pp.89–97.
- Spallek, H. 2003. Adaptive Hypermedia: A New Paradigm for Educational Software. *Advances in Dental Research*, 17(1), pp.38–42.
- Stash, N. V., Cristea, A. I., & De Bra, P. M. 2004. Authoring of learning styles in adaptive hypermedia: problems and solutions. In *Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters*, ACM, pp.114-123.
- Vaessen, B. Prins, F. Jeuring, J. 2014. University students' achievement goals and help-seeking strategies in an intelligent tutoring system, *Computers & Education*, 72, pp.196-208.

- Verdú, E. Regueras, M. Castro, J. García, R. 2012. A genetic fuzzy expert system for automatic question classification in a competitive learning environment, *Expert Systems with Applications*, 39(8), pp.7471-7478.
- Villaverde,, E. Godoy, D. Amandi, A. 2006. Learning styles' recognition in e-learning environments with feed-forward neural networks, *Journal of Computer-assisted Learning*, 22, pp.197-206.
- von Glasersfeld, E. 1995. A constructivist approach to teaching. In L. Steffe & J. Gale (Eds.). *Constructivism in education*, New Jersey: Lawrence Erlbaum Associates, Inc, pp.3-16.
- Wang, H.C., Li, T.Y., Chang, C.Y. 2006. A web-based tutoring system with styles-matching strategy for spatial geometric transformation. *Interacting with Computers*, 18, pp.331–355.
- Weber, G., Brusilovsky, P. 2001. ELM-ART: An adaptive versatile system for Web-based instruction *International Journal of Artificial Intelligence in Education*, 12(4), pp.351–384.
- Woo Woo, C. Evens, M. Freedman, R. Glass, M, Seop Shim, L. Zhang, Y, Zhou, Y. and Michael, J. 2006. An intelligent tutoring system that generates a natural language dialogue using dynamic multi-level planning, *Artificial Intelligence in Medicine*, 38, pp.25-46.
- Zadeh, L. 1992. Knowledge Representation In Fuzzy Logic. In *An Introduction To Fuzzy Logic Applications In Intelligent systems*, edited by Yager, R and Zadeh, L, Kluwer Academic Publishers.
- Zarandi, M. Khademian, M. Minaei-Bidgoli, B. Türkşen, I. 2012. A Fuzzy Expert System Architecture for Intelligent Tutoring Systems: A Cognitive Mapping Approach, *Journal of Intelligent Learning Systems and Applications*, 4(1), pp.29-40.