

Automatic student modeling in adaptive educational systems through probabilistic learning style combinations: a qualitative comparison between two innovative stochastic approaches

Fabiano A. Dorça · Luciano V. Lima ·
Márcia A. Fernandes · Carlos R. Lopes

Received: 22 December 2011 / Accepted: 10 May 2012 / Published online: 10 July 2012
© The Brazilian Computer Society 2012

Abstract Considering learning and how to improve students' performances, adaptive educational systems must know the way in which an individual student learns best. In this context, this work presents a comparison between two innovative approaches to automatically detect and precisely adjust students' learning styles during an adaptive course. These approaches take into account the nondeterministic and nonstationary aspects of learning styles. They are based upon two stochastic techniques: Markov chains and genetic algorithms. We found that the genetic algorithm (GA) based approach detects learning styles earlier and consequently provides personalized content earlier, making the learning process easier. The Markov based approach produces more fine-tuned results, taking into account strengths of learning styles.

Keywords Automatic learning styles assessment · Student modeling · Stochastic detection · E-learning · Adaptive educational systems

F. A. Dorça (✉) · M. A. Fernandes · C. R. Lopes
Faculty of Computer Science, FACOM, Federal University of Uberlândia (UFU), Campus Santa Mônica, Bloco 1B,
Av. João Naves de Ávila, 2121, Bairro Santa Mônica,
Uberlândia, Minas Gerais, CEP 38400-902, Brazil
e-mail: fabiano@facom.ufu.br

M. A. Fernandes
e-mail: marcia@facom.ufu.br

C. R. Lopes
e-mail: crlopes@facom.ufu.br

L. V. Lima
Faculty of Electrical Engineering, FEELT, Federal University of Uberlândia (UFU), Campus Santa Mônica, Bloco 1B,
Av. João Naves de Ávila, 2121, Bairro Santa Mônica,
Uberlândia, Minas Gerais, CEP 38400-902, Brazil
e-mail: vieira@ufu.br

1 Introduction

Research reveals that students' performances are improved if the learning environment gives support to their specific learning styles (LS). On the other hand, learners whose LS are not supported by the learning environment, may have more difficulties during the learning process [1,28,31,32,38]. LS and their effects on learning processes are carefully exposed by Coffield [13]. Their related instructional strategies have been massively studied in the new learning space introduced by the Internet, where many researchers point out that linking LS to appropriate learning resources is an important stimulus for learning processes. According to [3], the only way to improve performance is to improve education. If the system supports and facilitates high quality of teaching, good results are achieved.

In order to provide adaptivity, students' characteristics have to be known first. However, the traditional approaches for detection of LS in adaptive educational systems (AES) are inefficient. Graf et al. [29] comments on the loss of precision in self-assessment questionnaires, like Index of Learning Styles Questionnaire (ILS) [53], due to the necessity of students' metacognitive knowledge [20]. Furthermore, self-assessment questionnaires take a long time and demand patience to be answered by students. According to [55], this method has shown to be time-consuming and unreliable. In this way, many approaches to assessing students' LS have been proposed. However, in general, they present problems which make them either inefficient or difficult to be implemented, implanted and used, as shown in Sect. 2.

It is well-known that this topic is definitely relevant in the area of AES, since learning styles are of fundamental importance for learning effectiveness. In this context, this paper presents an experimental study on two alternative approaches to assessing students' LS. These approaches are based on

two well-known stochastic techniques, Markov chains and genetic algorithms, and they aim to automatically detect LS, taking into account that, as pointed out by Felder [19], LS can be changed over time in an unexpected and unpredictable way, and even learners with strong preferences for a specific LS can act, sometimes, in a different way.

An advantage of our approach is to constantly revise, correct and adjust the initial information about students' LS, by observing their performance, while they interact with learning resources that fit to a specific LS combination (LSC), selected by a stochastic process, as shown in Sect. 6. As a result, our approaches gradually and constantly update the student model (SM) [54], which effectively converges towards students' real LS, as shown in Sect. 7, which presents a qualitative comparison between the Markov chains based approach and the genetic algorithms based approach. Finally, Sect. 8 presents conclusions and future work.

2 Related works

A diversity of approaches for automatic detection of LS has been proposed, as can be seen in [11,27,30]. In general, these approaches use deterministic inference systems for detecting students' behavioural patterns. These systems infer the LS based on students' actions. One of the problems with these systems is the uncertainty, difficulty and complexity of developing and implementing rules that are able to infer LS effectively through students' actions and to treat students' behaviour as evidences and not as possibilities. Besides, in some systems like AHA! [14], these rules must be defined by the tutor, making the system more difficult to be used.

In the approach proposed by Limongelli et al. [41], the SM is initialized through the Index of Learning Styles (ILS) Questionnaire, and the system updates the students' LS according to the LS associated to a node, considering the time spent for the node and the score obtained with the post-test. Advantages of our approach are that the use of self-assessment questionnaires is not compulsory (as shown in Sect. 7), and the process of updating students' LS is not deterministic but stochastic, considering that not only LS but also many factors exert some influence on students' performance, as stated by [1,28,31,32,38,44].

More complex approaches can be seen in [8–10,22,37,55–57]. These approaches use learning machine techniques, such as Bayesian and neural networks. Some of the problems with these approaches are both high complexity and computational cost, which are thought to be serious concerns when considering a high number of students using the AES simultaneously. Besides, in general, these approaches are highly coupled, either to the system or to the whole teaching process, making them harder to be reused in other systems. In some

of these approaches, once acquired, the students' LS remain the same throughout the entire learning process [11].

Another well-known problem with these approaches is the complication generated by concept drift and concept shift [11]. It is well-explained by Castillo et al. [11] that, as a rule, supervised learning assumes the stability of the target concept. Therefore, in many real world problems, when data is collected over an extended period of time, the learning task can be complicated by changes in the target concept.

Yannibelli et al. [55] present a genetic algorithm approach for automatically identifying and tracking students' LS over time, based on the actions they take during a course. The aim of the algorithm is to detect the combinations of actions that the student usually performs to learn. A student-preferred combination of actions is then mapped to LS preferences. A problem with this approach, as stated before in this section, is the difficulty of mapping actions to LS, treating students' behaviour as evidences and not as possibilities. One advantage of our approach is to directly detect students' LS based on their performance, as depicted later in this paper.

According to [23], the quality of an AES critically depends on the quality of its student modeling. The system might implement a precise adaptation strategy and provide students with personalized learning content, but if its estimations of students' knowledge and preferences are inconsistent, the adaptive interventions it produces are unlikely to be effective. In this case, the concept of consistency indicates whether the SM describes correctly the student's characteristics.

In this scenario, adaptive decision models, which are able to better adapt to students' LS, are desirable. In this context, we believe that our approach brings advantages due to the following specific features:

- it considers that not only LS but also many factors exert some influence on students' performance, making it harder to infer students' LS based only on fixed behavioural pattern rules, because students' behaviour and performance may be influenced by other factors besides LS. Some of these factors are pointed out by [1,28,31,32,38,44];
- it considers that the influence exerted by each LS on students' behavior is unknown [4];
- it considers that LS can change over time in an unpredictable way. These changes may be associated with other factors, such as knowledge domain, as analyzed by Jones et al. [35];
- it considers that it is impossible to know the precision of the results obtained from self-assessment questionnaires (which may have inconsistencies) [11,49,50];
- it eliminates the necessity to discover students' behavioural patterns, considering that it is hard or impossible to obtain such patterns, considering that students with the same learning styles preferences may, sometimes, act

differently, taking into account the concept of tendencies, which means that even a learner with, e.g., a strong active LS can act sometimes in a reflective way [25];

- it is uncoupled from any learning management system (LMS), being independent from any specific students' actions in a specific system, as it always occurs in traditional approaches [22,30];
- it takes into account the dynamic nature of LS, which may change when the knowledge domain changes [37] or naturally evolve over time [44];
- it eliminates the necessity of using complex machine learning techniques, which are difficult to implement and may bring problems such as the complications related to concept drift and concept shift, as exposed by Castilho et al. [11];
- it eliminates the necessity of using drift-detection methods and dealing with concept drift and concept shift, which are automatically handled by the approach described in this paper.

According to [54], building a student model involves defining crucial matters, such as the level of specialization of the students that are at their knowledge and capabilities and also the way of giving assistance, providing feedback and interpreting the behavior of the learner. In this work, we focus on modeling students' learning capabilities, known as LS.

Our approach is based on the Felder and Silverman learning styles model (FS). The next section presents important aspects of FS to our work.

3 Learning styles

According to [3], the term learning style may include more than 70 different models with conflicting assumptions about learning, and with different designs and starting Points. In this work, we consider the Felder and Silverman's definition, where LS are defined as the characteristics, strengths and preferences in the way people receive and process information [19]. It refers to the fact that each student has their own particular method or set of strategies when learning.

Theories of LS simply assume that everyone can learn, but in different ways and levels [3]. There are many different theories and models of learning styles with varying dimensions and variables. They focus on different aspects, cognitive processes, skills, sensory modalities, learning processes, thinking styles, etc. Well-known theories and models of LS have been proposed by Kolb [39]; Honey and Mumford [33]; Entwistle [18]; Pask [48]; Felder and Silverman [19]. Each one of these models describes different aspects in which students prefer to learn.

Graf and Kinshuk [25] point out that the FS is one of the most frequently used in AES [7]. Besides, Kuljis and

Liu [40] claim that FS is the most appropriate model for the implementation of AES. According to Kinshuk et al. [38], FS combines the main models, such as Kolb [39], Pask [48] and the Myers-Briggs indicator [6].

According to Graf and Kinshuk [25,26], the FS uses the concept of dimensions, and, therefore, describes LS more thoroughly. As proposed by Felder and Silverman [19], each learner has a dominant preference in each of the four dimensions: Processing (active/reflective); Perception (sensitive/intuitive); Input (visual/verbal); Understanding (sequential/global). Each preference tells us about how a student learns best and the related pedagogical strategies for effective learning. According to FS, each learner has a preference within the scope of each one of the four dimensions described above, which is measured on a scale from +11 to -11. This feature makes it possible to describe the strength of the learners' preferences [38].

As described by Graf et al. [25], LS are considered to be flexibly stable, which means that they are relatively stable but they can change over time. For instance, when learners train their weak LS. Furthermore, FS is based on the concept of tendencies, which means that even learners with, e.g., a strong active LS can act sometimes in a reflective way [25].

A very important characteristic of FS for our work is that it uses scales to classify students instead of using defined types. In this way, the strength of each LS is finely measured [19]. Another important aspect of FS is that it considers LS as tendencies and students may act differently in specific situations, that is, in a nondeterministic way, as pointed out by Kinshuk et al. [38]. Therefore, we can consider students' LS as probabilities in the four-dimensional FS model, as depicted in the Sect. 6.

In this context, our work introduces the use of stochastic techniques to effectively provide adaptation and diagnose students' LS. Particularly, we analyze and compare the use of Markov chains and genetic algorithms for handling adaptation and automatic detection of students' LS. The next section presents important aspects of MC for this work. The main foundations of GA are briefly presented in Sect. 5.

4 Markov chains

A Markov chain (MC) is a mathematical system that represents changes of state between a finite number of possible states, and they are often described by a directed graph, where the edges are labeled by the probabilities of going from one state to others. The changes of state are called transitions, and the probabilities associated with transitions are called transition probabilities. A MC is a stochastic process characterized as memoryless, that means, the next state depends only on the current state and not on the sequence of events

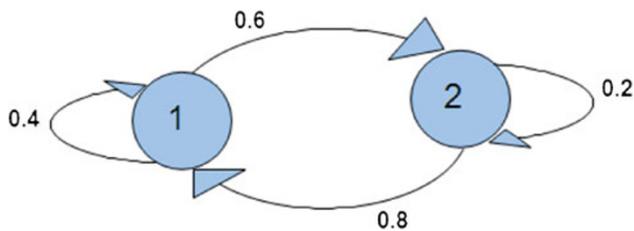


Fig. 1 Markov chain

that preceded it. This property is called a Markov property. Figure 1 presents an example of MC.

The MC shown in Fig. 1 represents a stochastic process in which a random variable X_t defines the state of the system on time t . There are two possible states: 1 and 2. The transition probabilities from one state to another are also described in the picture. The set of all states and transition probabilities completely characterizes a MC.

Formally, a MC is a sequence of random variables X_1, X_2, X_3, \dots with the Markov property, namely that, given the present state, the future and past states are independent. Formally, $P(X_{n+1} = x | X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = P(X_{n+1} = x | X_n = x_n)$, which means that the probability of going from state i to state j in n time steps is $p_{ij}^{(n)} = P(X_n = j | X_0 = i)$, and the single-step transition is $p_{ij} = P(X_1 = j | X_0 = i)$.

The transition probabilities of a MC are described by a transition matrix [45]. Each line represents the transition probabilities from a state to the others. Therefore, the transition probability from state 1 to state 2 is defined in the position given by the intersection of line 1 with column 2 of the transition matrix. For example, the transition matrix T below represents the transition probabilities that appears in the MC shown in Fig. 1.

$$T = \begin{bmatrix} 0.4 & 0.6 \\ 0.8 & 0.2 \end{bmatrix}$$

It is important to notice that the rows of T sum to 1. This is because P is a stochastic matrix. Each line describes the transition probabilities $P(X_{n+1} = x | X_n = x_n)$ from one state to others, in time n . A MC is a discrete-time random process with the Markov property. Also, a MC has a discrete (finite or countable) state-space. A discrete-time random process represents a process that is in a certain state, in a certain time n , with the state changing randomly throughout time. The time n represents a step of the process, and the conditional probability distribution for the process at the next step depends only on the current state. An example is given by Fig. 2, where the probabilities of weather conditions (modeled as either sunny (state 1) or rainy (state 2)) on next day is given by the weather on the current day. The transition matrix T below represents the MC depicted in Fig. 2.

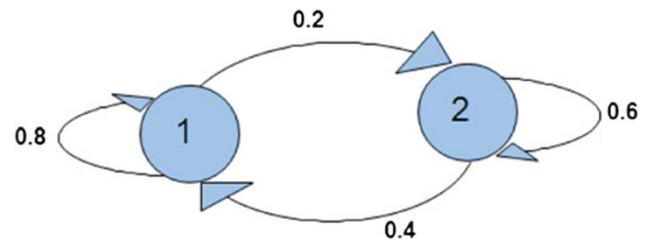


Fig. 2 Probabilities of weather conditions

$$T = \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix}$$

The matrix T represents the weather model in which a sunny day is 80 % likely to be followed by another sunny day, and a rainy day is 60 % likely to be followed by another rainy day. Consequently, p_{ij} is the probability of a day of type i be followed by a day of type j .

The next section presents important aspects of genetic algorithms to our work.

5 Genetic algorithms

A genetic algorithm (GA) is an adaptive search technique based on Darwin's theory of evolution, which is characterized by an iterative process and work in parallel on a number of potential solutions for a problem.

The fitness value is a numerical value that expresses the performance of an individual (a possible solution) for solving the problem. The notion of fitness is fundamental to the application of GA, in which the degree of success depends critically on the definition of a fitness function that ensures that individuals can be differentiated according to their capacity for solving the problem. Individuals evolve through an iterative process [17].

This process leads to the evolution of a population of individuals that are better fitted to their environment than the individuals that they were created from, just as in natural adaptation. GAs often perform well approximating solutions to all types of problems because they do not make any assumption about the underlying fitness function, which is specific to each problem [17].

As stated in [17], the general scheme of a GA is given in Algorithm 1.

Algorithm 1 Scheme of a GA

```

INITIALIZE population with random candidate solutions;
EVALUATE each candidate;
while not TERMINATION CONDITION is satisfied do
  SELECT parents;
  RECOMBINE pairs of parents;
  MUTATE the resulting offspring;
  EVALUATE new candidates;
  SELECT individuals for the next generation;
end while

```

The main features of GAs are [17]:

- they are population based, i.e., they process a whole collection of candidate solutions simultaneously;
- they use recombination to mix information of more candidate solutions into a new one;
- they are stochastic.

As pointed out in [17], the most important components of a GA are:

- representation (definition of individuals);
- evaluation function (or fitness function);
- population;
- parent selection mechanism;
- variation operators (recombination and mutation);
- survivor selection mechanism (replacement).

GAs operate on a population of potential solutions applying the principle of survival of the fittest to produce better and better approximations to a solution [12]. Individuals, or current approximations, are encoded as strings, or chromosomes, composed over some alphabet. The most commonly used representation in GAs is the binary alphabet {0, 1}, although other representations can be used, e.g., ternary, integer, real-valued, etc. [12]. Genetic operators are used in GAs to generate diversity and to combine existing solutions into others. Genetic variation is a necessity for the process of evolution. Genetic operators used in genetic algorithms are analogous to those in the natural world: survival of the fittest (selection), reproduction (crossover or recombination) and mutation [24]. The next section presents our approaches in detail.

6 Two alternative approaches to assessing students' learning styles

In this section, we present in detail our stochastic approaches for automatic detection of students' LS, which use probabilistic LS combinations (LSC), as presented hereafter. Although the idea of automatic detection of learning styles is not new, the techniques used are novel. Our approaches use information from a student's performance for updating the SM frequently while the student is using the system for learning.

In this way, students' LS are dynamically and constantly revised and corrected, leading to fine-tuned SMs, which are consistent with students' real LS (in our work, the concept of consistency indicates whether the SM describes correctly the student's LS). As a consequence, this process allows AES to provide more accurate adaptivity. A conclusion can be drawn that our approach takes into account the development of an advanced student modeling approach concerning LS,

which combines the automatic, dynamic, and global student modeling aspects, as pointed out in [25].

Due to the stochastic nature of LS, our approaches are based on a probabilistic learning styles combination [21]. A learning styles combination (LSC) is a 4-tuple composed by one LS from each FS dimension, as stated by Definition 1.

Definition 1 (*Learning styles combination (LSC)*) $LSC = (a, b, c, d) | a \in D1, b \in D2, c \in D3, d \in D4$ such that:
 $D1 = \text{Active}(A), \text{Reflective}(R)$
 $D2 = \text{Sensitive}(S), \text{Intuitive}(I)$
 $D3 = \text{Visual}(Vi), \text{Verbal}(Ve)$
 $D4 = \text{Sequential}(\text{Seq}), \text{Global}(G)$

Therefore, there are 16 possible learning styles combinations, as stated by Definition 2.

Definition 2 (*Learning styles combinations (LSCs)*) $LSCs = (A, Vi, S, \text{Seq}), (A, Vi, S, G), (R, Vi, S, \text{Seq}), (R, Vi, S, G), (A, Ve, S, \text{Seq}), (A, Ve, S, G), (R, Ve, S, \text{Seq}), (R, Ve, S, G), (A, Vi, I, \text{Seq}), (A, Vi, I, G), (R, Vi, I, \text{Seq}), (R, Vi, I, G), (A, Ve, I, \text{Seq}), (A, Ve, I, G), (R, Ve, I, \text{Seq}), (R, Ve, I, G)$.

We propose that during a learning session the student should interact with a set of learning objects (LO) [34] that satisfies a specific LSC, probabilistically selected according to the student's LS preferences stored in the SM. Which means that, in our approach, a LSC is a specific combination of four random variables [47]. Therefore, in our approach, the student's LS describes the probability of random variables a, b, c, d , considering Definition 1.

In this context, in our approach, student's LS are stored as values in the interval [0, 100] instead of [-11, +11] representing a student's probability of preference for a specific LS in a FS dimension. Therefore, the student's preferences are stored as probabilities. Considering this model, a student's LS are represented according to Definition 3.

Definition 3 (*Learning styles (LS)*) $LS = (Pr_A, Pr_R), (Pr_S, Pr_I), (Pr_{Vi}, Pr_{Ve}), (Pr_{Seq}, Pr_G)$ such that $Pr_A + Pr_R = 100, Pr_S + Pr_I = 100, Pr_{Vi} + Pr_{Ve} = 100, Pr_{Seq} + Pr_G = 100$

Table 1 presents an example of a student model.

Therefore, in Table 1 we can consider that the student probably is Reflective, Intuitive, Visual and Sequential. In

Table 1 Student model

Learning styles							
Processing		Perception		Input		Understanding	
Act.	Ref.	Sen.	Int.	Vis.	Ver.	Seq.	Glo.
35.0	65.0	17.0	83.0	89.0	11.0	84.0	16.0

this context, the greatest advantage of our approach is to stochastically consider all LSCs according to the student's **supposed** LS, that **may** be wrong or **may** change over time, as pointed out in Sect. 2. As shown in Sect. 7, this characteristic allows system to effectively discover and fine-tune the SM.

At this point, it is important to mention that although Table 1 presents a SM exclusively based on LS, there are many other factors that could be important for learning, such as [43]: initial user knowledge; objective and plans; cognitive capacities; preferences; academic profile; age and type of student; cognitive style (affective, impulsive, etc.); personality aspects (introverted, extroverted, etc.); deficiencies (visual or others) and traces of the personality. Additional important factors for learning are discussed in [16]. Furthermore, a SM should include information referring to the specific knowledge that the system judges that the user possesses on the domain. Therefore, in AES, SM has increased relevance: when the student reaches the objectives of the course, the system must be able to readapt, for example, to his knowledge [43]. We consider this characteristic in our approach, as it can be seen in Sect. 7.

Additionally, according to Martins et al. [43], the learning process is more efficient when it is built over previously acquired knowledge. In addition, Martins et al. [43] reports that in AES the emphasis is placed on students' knowledge in the domain application and LS, in order to allow them to reach the learning objectives proposed in their training. A variety of student modeling approaches, techniques and standards are used to implement SM, and some of them are depicted by Thompson [52].

This is just to say that the SM is usually more complicated than the one presented in this paper. But, to reach our goal, we needed to point the attention only to LS, excluding other important characteristics of the SM. As a consequence, we isolated the around complexity, in order to come up with an efficient model to automatically detect students' LS. But, it is important to mention that when using our approach with an existing LMS, such characteristics of the SM should be considered for providing adaptivity. Considering those characteristics in student modeling does not affect the operation of our approach. This is possible because our approach is intrinsically able to deal with uncertainty in the process of detection of LS. The uncertainty appears due to the diversity of factors that exert influence on the learning process, as stated in [1, 28, 31, 32, 38, 44].

Following such, we present two alternative approaches to automatically detecting LS: a MC based approach and a GA based approach. Both are stochastic approaches, which apply different stochastic techniques. The pros and cons of each one of these approaches are opportunely discussed in Sect. 7, where we present a comparative experimental study on them.

6.1 MC based approach

In this approach, we consider a stochastic process modeled by four concurrent MC [45], which are depicted in Fig. 3. In Fig. 3, the MC were modeled by considering the SM presented in Table 1. In this approach, each state represents a preference in a LS dimension. Therefore, in Fig. 3a, state 1 represents the active preference and state 2 represents the reflective preference. In Fig. 3b, state 1 represents the sensitive preference and state 2 represents the intuitive preference. In Fig. 3c, state 1 represents the visual preference and state 2 represents the verbal preference. In Fig. 3d, state 1 represents the sequential preference and state 2 represents the global preference. The probabilities of occurrence of a state are described by transition matrices [45]. Therefore, a preference from each dimension is selected to compose a LSC to be considered in a learning session during the learning process. As a result, the SM is constantly updated, as explained later in this section. The transition matrices $D1$, $D2$, $D3$, $D4$ describe the four random variables represented in Fig. 3.

$$D1 = \begin{bmatrix} 0.35 & 0.65 \\ 0.35 & 0.65 \end{bmatrix} \quad D2 = \begin{bmatrix} 0.17 & 0.83 \\ 0.17 & 0.83 \end{bmatrix}$$

$$D3 = \begin{bmatrix} 0.89 & 0.11 \\ 0.89 & 0.11 \end{bmatrix} \quad D4 = \begin{bmatrix} 0.84 & 0.16 \\ 0.84 & 0.16 \end{bmatrix}$$

$D1$ represents the MC depicted in Fig. 3a, $D2$ represents the MC depicted in Fig. 3b, $D3$ represents the MC depicted in Fig. 3c and $D4$ represents the MC depicted in Fig. 3d. This approach has naturally evolved to a GA based approach, considering that GA are also stochastic process, due to its probabilistic nature. The GA based approach is depicted below.

6.2 GA based approach

Taking into consideration that LS are probabilities, a LSC can be probably considered during a learning session according to the probability distribution shown in Table 2, considering the SM depicted in Table 1. This approach aims to automatically detect students' LS based on GA techniques [12, 17]. The selection of a LSC from a population during a learning session is done by a stochastic selection method [24]. We use the Roulette Wheel Selection due to its adequacy to our approach [24].

Therefore, each LSC is considered as an individual, which has a fitness given by its probability of preference by the student, as shown in Table 2. An individual's probability of selection is given by the proportion between their fitness and the entire population's fitness, as shown in (1).

$$Pr_i = \frac{\text{fitness}_i}{\sum_{x=1}^n \text{fitness}_x} \quad (1)$$

As pointed out in Sect. 5, a GA has some important components. They are:

Fig. 3 Stochastic process, modeled by four concurrent Markov chains. **a** Processing: 1 active, 2 reflective, **b** perception: 1 sensitive, 2 intuitive, **c** input: 1 visual, 2 verbal, **d** understanding: 1 sequential, 2 global

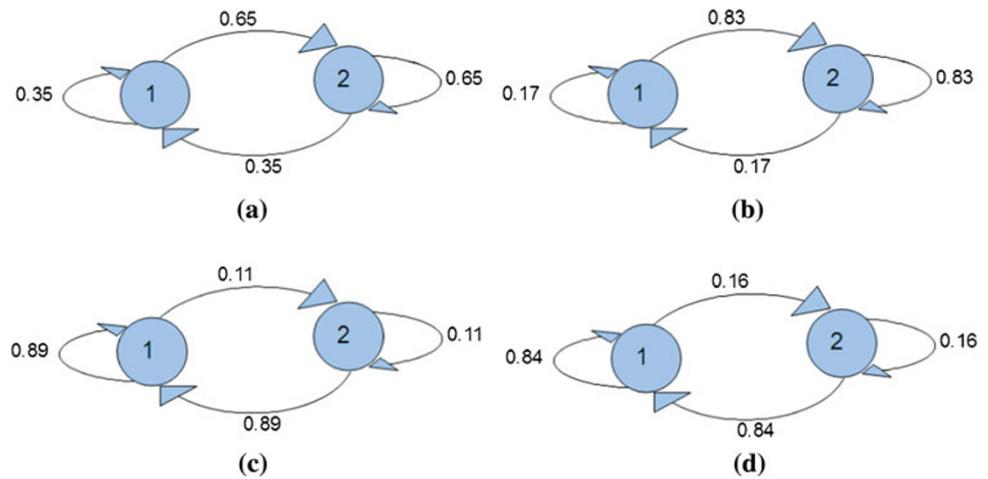


Table 2 LSC probability distribution

LSC	Probabilities
$P(S,Vi,A,Seq)$	$0.17 \times 0.89 \times 0.35 \times 0.84 = 0.045$
$P(S,Vi,A,G)$	$0.17 \times 0.89 \times 0.35 \times 0.16 = 0.008$
$P(S,Vi,R,Seq)$	$0.17 \times 0.89 \times 0.65 \times 0.84 = 0.083$
$P(S,Vi,R,G)$	$0.17 \times 0.89 \times 0.65 \times 0.16 = 0.016$
$P(S,Ve,A,Seq)$	$0.17 \times 0.11 \times 0.35 \times 0.84 = 0.005$
$P(S,Ve,A,G)$	$0.17 \times 0.11 \times 0.35 \times 0.16 = 0.002$
$P(S,Ve,R,Seq)$	$0.17 \times 0.11 \times 0.65 \times 0.84 = 0.010$
$P(S,Ve,R,G)$	$0.17 \times 0.11 \times 0.65 \times 0.16 = 0.003$
$P(I,Vi,A,Seq)$	$0.83 \times 0.89 \times 0.35 \times 0.84 = 0.217$
$P(I,Vi,A,G)$	$0.83 \times 0.89 \times 0.35 \times 0.16 = 0.043$
$P(I,Vi,R,Seq)$	$0.83 \times 0.89 \times 0.65 \times 0.84 = 0.403$
$P(I,Vi,R,G)$	$0.83 \times 0.89 \times 0.65 \times 0.16 = 0.076$
$P(I,Ve,A,Seq)$	$0.83 \times 0.11 \times 0.35 \times 0.84 = 0.026$
$P(I,Ve,A,G)$	$0.83 \times 0.11 \times 0.35 \times 0.16 = 0.005$
$P(I,Ve,R,Seq)$	$0.83 \times 0.11 \times 0.65 \times 0.84 = 0.049$
$P(I,Ve,R,G)$	$0.83 \times 0.11 \times 0.65 \times 0.16 = 0.009$
Sum of probabilities of all LSC	1.000

The role of survivor selection is to deterministically choose which individuals will be allowed in the next generation. This decision is based on their fitness values, favoring those with higher quality. But, we have to consider that if the SM is inconsistent, the best fitted LSC may not be the preferred one by the student. Therefore, we do not use any survivor selection mechanism in this approach.

A binary representation of LSCs was used, where preferences A, S, Vi, Seq are represented by 0 and preferences R, I, Ve, G are represented by 1. Therefore, we have the following $LSCs = \{(0,0,0,0), (0,0,0,1), (1,0,0,0), (1,0,0,1), (0,1,0,0), (0,1,0,1), (1,1,0,0), (1,1,0,1), (0,0,1,0), (0,0,1,1), (1,0,1,0), (1,0,1,1), (0,1,1,0), (0,1,1,1), (1,1,1,0), (1,1,1,1)\}$. After each learning session, we apply recombination and mutation operators [12]. The recombination operator recombines LSCs in order to (probably) produce more fit individuals.

We are using here the single-point crossover [12]. This crossover operation is not necessarily performed on all individuals. Instead, it is applied with a probability P_x when the pairs are chosen for breeding. The mutation operator is then applied to the new LSC, with a probability P_m (mutation rate).

- *Representation*: an individual is a LSC, according to Definition 1.
- *Evaluation function*: fitness function $P(LSC)$, which calculates how much a LSC is preferred by the student, as expounded in Table 2.
- *Population*: binary representations of LSCs, where preferences A, S, Vi, Seq are represented by 0 and preferences R, I, Ve, G are represented by 1.
- *Parent selection mechanism*: Roulette Wheel Selection [24]
- *Variation operators*: recombination and mutation, as depicted below
- *Survivor selection mechanism*: not applied, as explained below

6.3 Updating the student model

When a student shows a learning problem during a learning session (unsatisfactory performance), LS in SM that appear in the currently selected LSC are decremented, considering a probable inconsistency in these preferences. LS in SM that do not appear in the currently selected LSC are incremented (reinforced), making them stronger, considering that the learning difficulties appeared because they were not present in the selected LSC.

This approach for gradually updating student LS is based on Reinforcement Learning [36] techniques and is a critical part of our work, which is currently being adjusted. While these updates are executed, SM becomes more consistent,

providing support to more accurate adaptivity. Therefore, adaptivity becomes more accurate and students' performances are improved. As stated before, in our work, the concept of consistency indicates whether the SM describes correctly the student's LS.

It is well-known that a variety of factors should be taken into account for students' performance evaluation and learning problems detection, as pointed out in [16, 42]. It's a complex problem and a lot of approaches have been proposed to solve it. For testing our approach without this complexity, we considered a simulated learning process, which is a stochastic process that infers students' performances, taking into account some aspects related to the impact of LS on learning processes, as depicted in [1, 28, 31, 32, 38]. The next section presents some experiments and discusses their results.

7 Comparing results

This section aims to present an experimental study and the results obtained through tests with both approaches.

7.1 Methodology

Both approaches have been tested through a set of experiments. Some of them are expounded in this section and their results are compared and discussed. The experiments with the GA based approach were performed considering a population of 100 individuals (LSC), in which the first 16 individuals are copied from LSCs, described in Definition 2, and the other individuals are randomly generated. Additionally, the mutation rate Pm was set with 0.2, and Px was set with 0.1. This configuration has generated encouraging results.

Each experiment was repeated 20 times. Therefore, we could observe the process under different circumstances and identical conditions. It was possible to notice that the resulting sequences during an experiment were different, but the final results were very similar. So, the nondeterministic and convergence aspects intrinsic to the student modeling process were very clear to us.

Four experiments and their results are shown in this section. The execution of an experiment finishes when the student achieves all learning goals. We considered 30 concepts to be learned by students and 6 cumulative cognitive levels to be achieved in each concept, based on the Bloom's Taxonomy for Knowledge [2]. Therefore, the simulated learning process, in these experiments, should have, at least, 180 learning sessions (or iterations) in order to achieve all learning goals ($30 \times 6 = 180$).

When students have good performances during a learning session, their cognitive level in a concept evolves, until they reach the maximum cognitive level to the concept. When students fail, their cognitive level in the concept does not evolve.

Therefore, the easier the learning process, the fewer iterations necessary to achieve all learning goals. The better adapted the content is, the easier the learning process is, as pointed out in Sect. 1.

The simulation process needs to know the student's real LS (SRLS) and the strength of each preference (strong/moderate/weak or balanced). For each experiment, we show, graphically, how student's probable LS (SPLS), stored in SM, is updated during the learning process. In each graph, the x -axis shows the numbers of the iterations of the learning process and the y -axis shows the updating of the SPLS throughout the learning process. The main goal was to observe how SPLS are gradually updated and tuned along the iterations of the learning process. In order to validate the approaches, we considered two variables:

- consistency: the SPLS effectively converged to the SRLS during the learning process?
- efficiency: the SPLS converged to the SRLS in reasonable time? i.e., the SPLS became consistent in the beginning of the learning process?

The results obtained through experiments show that, considering these variables, both approaches are valid. We could observe different levels of consistency and efficiency when comparing the approaches, as it is depicted hereafter.

7.2 Experiment 1

Firstly, we considered a student with the following SRLS: {reflective (strong), sensitive (strong), visual (moderate), global (weak)}. The SPLS initially stored in the SM is shown in Table 3. Therefore, the SM is initially inconsistent (doesn't express the SRLS correctly), specifically in dimensions active/reflective and sensitive/intuitive. Figure 4 presents how SPLS were updated during this experiment, considering the Markov based approach. The Fig. 4a shows how the SPLS, in dimension processing, were updated along the learning process. The Fig. 4b shows how the SPLS, in dimension perception, were updated along the learning process. The Fig. 4c shows how the SPLS, in dimension input, were updated along the learning process. The Fig. 4d shows how the SPLS, in dimension understanding, were updated along the learning process. As it can be seen, the SPLS became

Table 3 SM—experiment 1

Learning styles							
Processing		Perception		Input		Understanding	
Act.	Ref.	Sen.	Int.	Vis.	Ver.	Seq.	Glb.
70.0	30.0	35.0	65.0	60.0	40.0	45.0	55.0

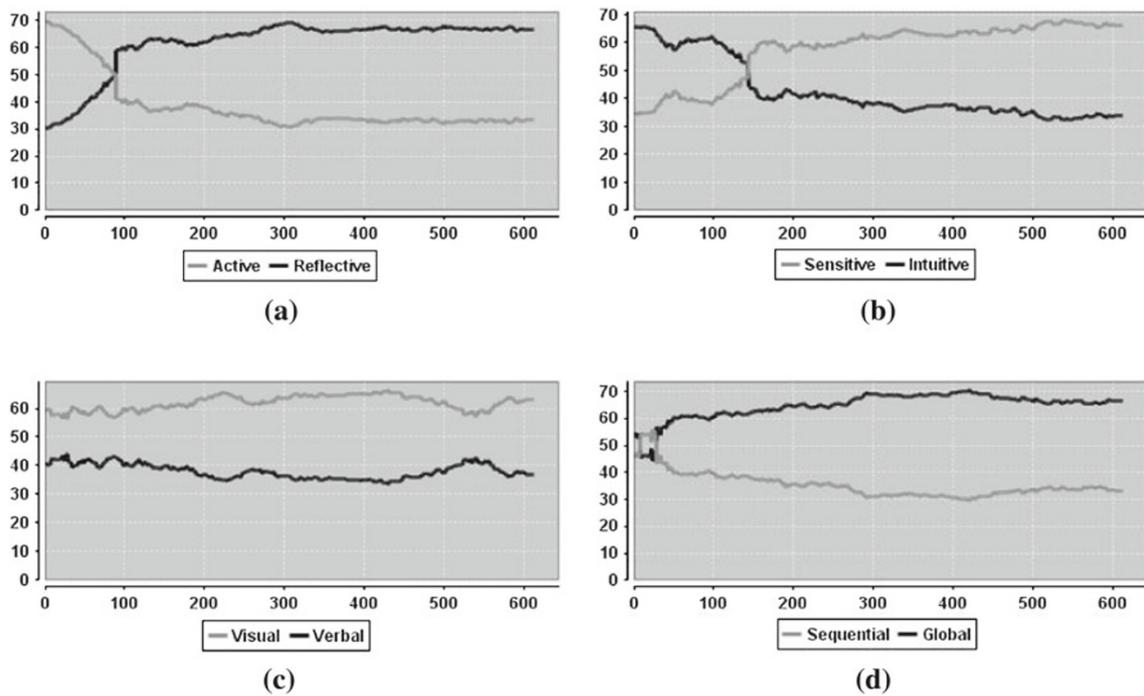


Fig. 4 Experiment 1—results from MC based approach. **a** Processing, **b** perception, **c** input, **d** understanding

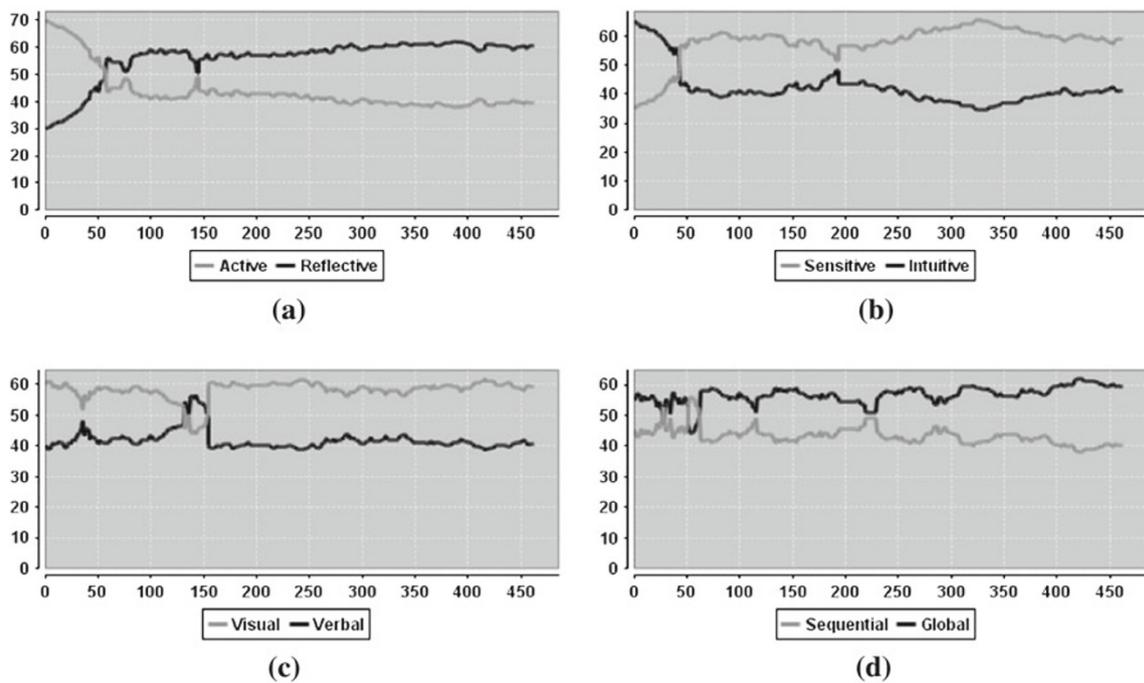


Fig. 5 Experiment 1—results from GA based approach. **a** Processing, **b** perception, **c** input, **d** understanding

consistent with the SRLS during the learning process, considering its four dimensions. Figure 5 presents how SPLS were updated during this experiment, considering the GA based approach. The Fig. 5a shows how the SPLS, in dimension processing, were updated along the learning process. The Fig. 5b shows how the SPLS, in dimension perception, were

updated along the learning process. The Fig. 5c shows how the SPLS, in dimension input, were updated along the learning process. The Fig. 5d shows how the SPLS, in dimension understanding, were updated along the learning process. As it can be seen, during the learning process, the SPLS became consistent with the SRLS considering its four dimensions.

Table 4 SM—experiment 2

Learning styles							
Processing		Perception		Input		Understanding	
Act.	Ref.	Sen.	Int.	Vis.	Ver.	Seq.	Glb.
50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0

In both approaches, all repetitions of this experiment produced a consistent SM during the learning process. As it can be seen, the GA based approach took less iterations than the Markov based approach. It happens because, as it can be noticed in Figs. 4 and 5, the GA based approach detected the student's LS earlier than the Markov based approach. Therefore, the teaching process performed by the GA based approach could provide more adequate content earlier, which provided accurate adaptivity earlier and made the learning process easier, and, as a consequence, improved the student's performance. As a result, less iterations were needed by the student to complete learning process. This fact was observed in every experiment performed.

7.3 Experiment 2

In the following experiment, we consider the case in which there is no initial information available about the SPLS, as shown in Table 4. The SRLS considered in this experiment are: SRLS = {reflective (weak), intuitive (strong), visual (moderate), sequential (weak)} Figure 6 presents how SPLS were updated during this experiment, considering the Markov based approach. The Fig. 6a shows how the SPLS, in dimension processing, were updated along the learning process. The Fig. 6b shows how the SPLS, in dimension perception, were updated along the learning process. The Fig. 6c shows how the SPLS, in dimension input, were updated along the learning process. The Fig. 6d shows how the SPLS, in dimension understanding, were updated along the learning process. As it can be seen, during the learning process, the SPLS became consistent with the SRLS considering its four dimensions. Figure 7 presents how SPLS were updated during this experiment, considering the GA based approach. The Fig. 7a shows how the SPLS, in dimension processing, were updated along the learning process. The Fig. 7b shows how the SPLS, in dimension perception, were updated along the learning process. The Fig. 7c shows how the SPLS, in dimension input, were updated along the learning process. The Fig. 7d shows how the SPLS, in dimension understanding, were updated along the learning process. As it can be seen, during the learning process, the SPLS became consistent with the SRLS considering its four dimensions. Figures 6 and 7 let us notice that, in this experiment, less iterations were necessary to complete the learning process than in experiment 1.

This occurred because inconsistencies in the SM seem to be worse than the lack of information. When the system doesn't have any initial available information about SPLS, it can discover preferences faster and provide accurate adaptivity earlier. In both approaches, all repetitions of this experiment produced a consistent SM.

7.4 Experiment 3

This experiment sets the initial SM with a result obtained from experiment 2. The goal is to observe how the SM is fine-tuned by the system during the learning process, even when it is initially consistent. Table 5 shows the SM used in this experiment. Figure 8 presents how the SPLS were updated during this experiment, considering the MC based approach. The Fig. 8a shows how the SPLS, in dimension processing, were updated along the learning process. The Fig. 8b shows how the SPLS, in dimension perception, were updated along the learning process. The Fig. 8c shows how the SPLS, in dimension input, were updated along the learning process. The Fig. 8d shows how the SPLS, in dimension understanding, were updated along the learning process. Figure 9 presents how the SPLS were updated during this experiment, considering the GA based approach. The Fig. 9a shows how the SPLS, in dimension processing, were updated along the learning process. The Fig. 9b shows how the SPLS, in dimension perception, were updated along the learning process. The Fig. 9c shows how the SPLS, in dimension input, were updated along the learning process. The Fig. 9d shows how the SPLS, in dimension understanding, were updated along the learning process. As it can be seen, during the learning process, the SPLS became consistent with the SRLS considering its four dimensions. As it can be seen, we had a considerable reduction in the number of iterations, due to the initial consistency of the SM. Which means that having precise information about student's preferences and providing support to them during the learning process results in great positive effects on a student's performance, as pointed out in [1,28,31,32,38]. In both approaches, all repetitions of this experiment produced a consistent SM. We can clearly notice the difference between strengths (strong/moderate/weak) in Fig. 8 (MC based approach). But it is not very clear in Fig. 9 (GA based approach). This is an interesting fact observed during the experiments: the MC based approach seems to be more sensitive than the GA based approach with respect to the strengths of the SRLS.

7.5 Experiment 4

This experiment considers the case in which the SM are initially inconsistent in all dimensions, as shown in Table 6. The SRLS are given by:

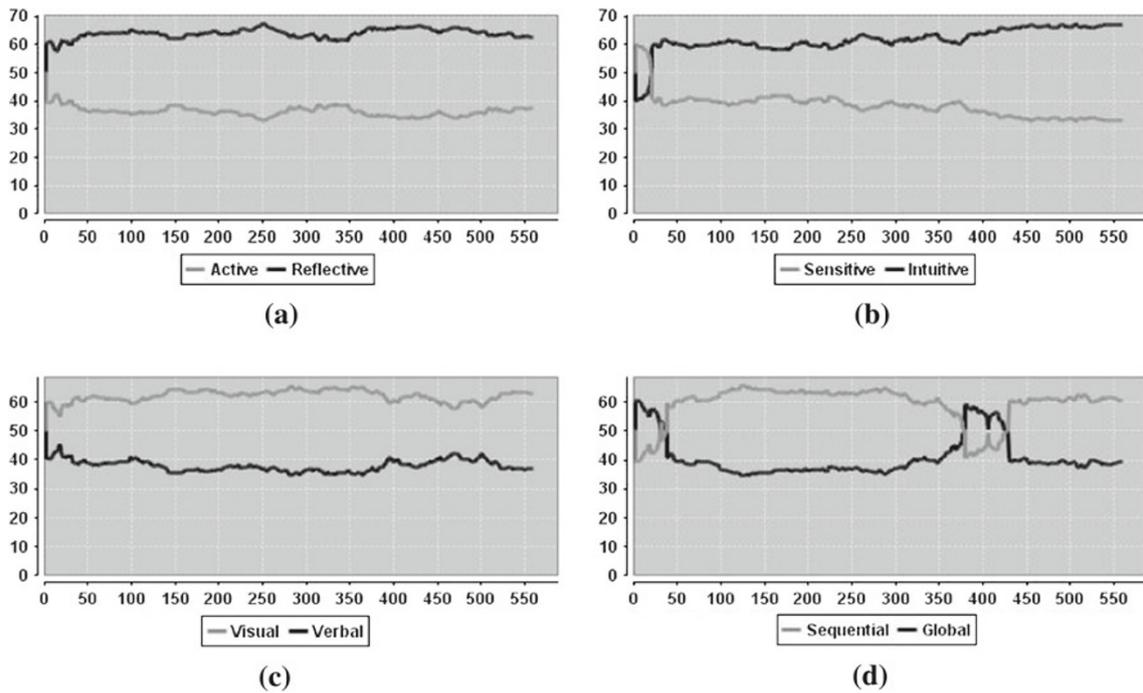


Fig. 6 Experiment 2—results from MC based approach. **a** Processing, **b** perception, **c** input, **d** understanding

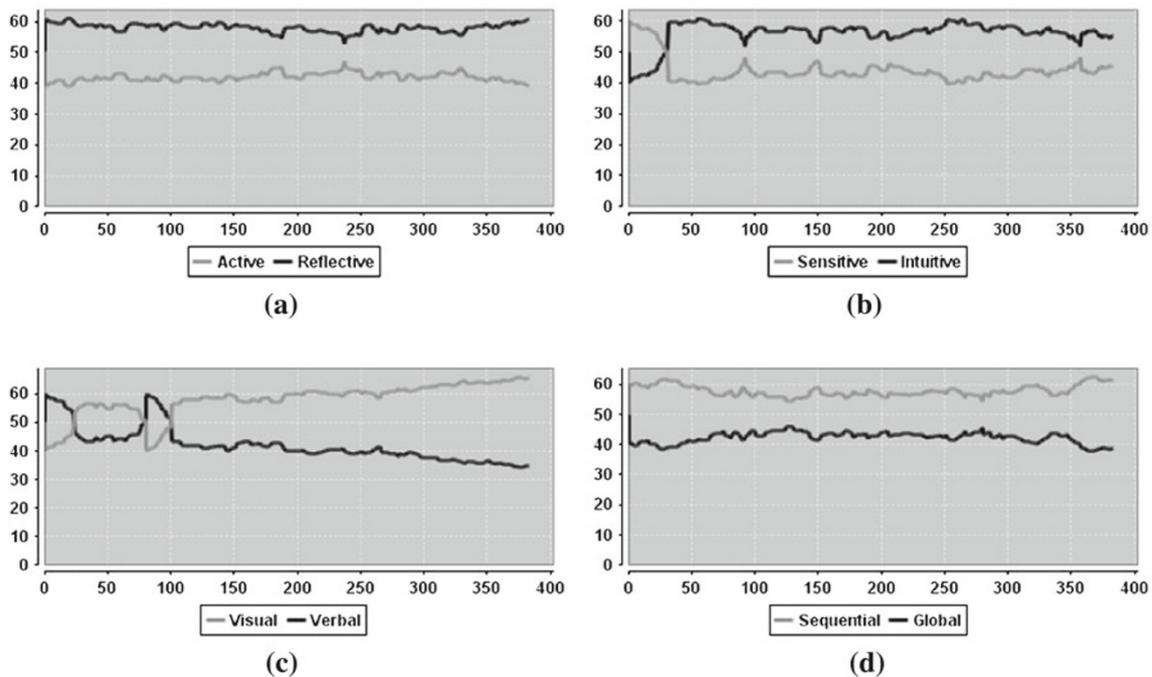


Fig. 7 Experiment 2—results from GA based approach. **a** Processing, **b** perception, **c** input, **d** understanding

SRLS = {Reflective (weak), Intuitive (strong), Verbal (moderate), Global (weak)}

Figure 10 presents how the SPLS were updated during this experiment, considering the MC based approach. The Fig. 10a shows how the SPLS, in dimension processing,

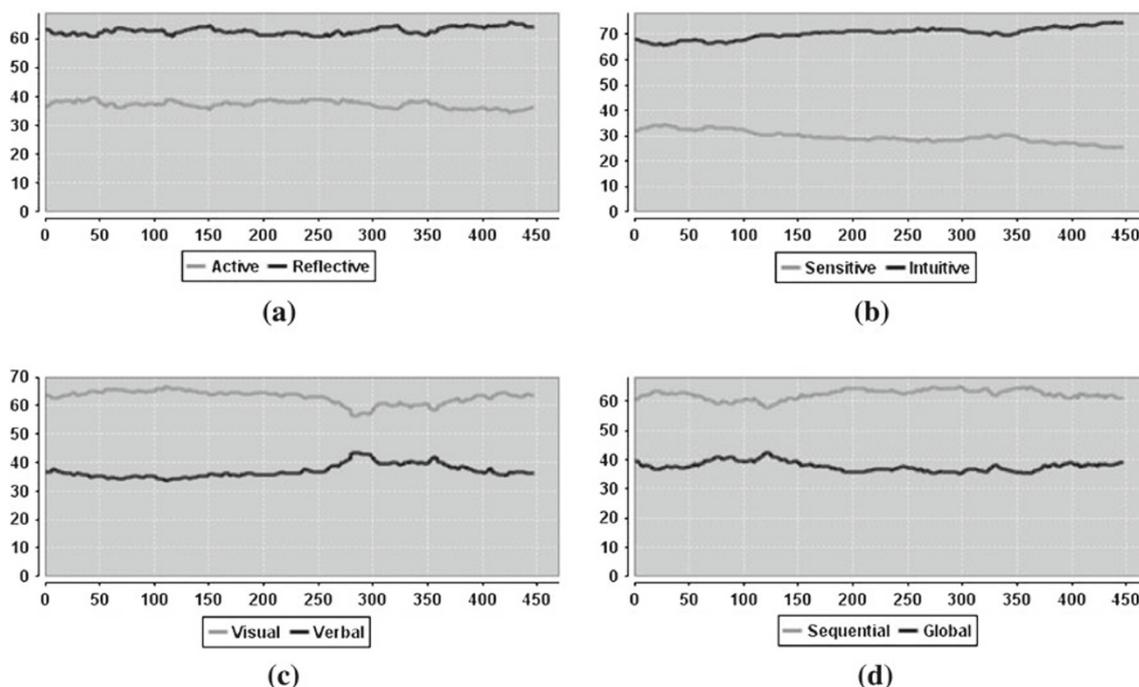
were updated along the learning process. The Fig. 10b shows how the SPLS, in dimension perception, were updated along the learning process. The Fig. 10c shows how the SPLS, in dimension input, were updated along the learning process. The Fig. 10d shows how the SPLS, in dimension

Table 5 SM—experiment 3

Learning styles							
Processing		Perception		Input		Understanding	
Act.	Ref.	Sen.	Int.	Vis.	Ver.	Seq.	Glb.
36.8	63.2	27.9	72.1	68.1	31.9	65.0	35.0

understanding, were updated along the learning process. As it can be seen, during the learning process, the SPLS became consistent with the SRLS considering its four dimensions. Figure 11 presents how the SPLS in the SM was updated during this experiment, considering the GA based approach. The Fig. 11a shows how the SPLS, in dimension processing, were updated along the learning process. The Fig. 11b shows how the SPLS, in dimension perception, were updated along the learning process. The Fig. 11c shows how the SPLS, in dimension input, were updated along the learning process. The Fig. 11d shows how the SPLS, in dimension understanding, were updated along the learning process. As it can be seen, during the learning process, the SPLS became consistent with the SRLS considering its four dimensions. As it can be seen, the SPLS were efficiently corrected. In both approaches, all repetitions of this experiment produced a consistent SM. The amount of iterations was smaller than in experiment 1, due to the fact that in this experiment only one strong preference was inconsistent. As pointed out by Felder *apud* [31], strong preferences produce

stronger negative effects on the students' performances when they are not supported by the learning process. During these experiments, we could notice that both approaches were able to efficiently discover the SRLS early in the learning process. Furthermore, we could notice that, although the GA based approach detects LS earlier than the Markov based approach, and consequently provides personalized content earlier, which makes the learning process easier, the Markov based approach produces more fine-tuned results, taking into account the strengths of the SRLS. Advantages of the Markov based approach are that it doesn't have to keep a population of LSC for each student, and it doesn't have to spend additional computational resources in order to compute next generations during the learning process, as it is done by the GA based approach. Finally, we believe that the results obtained from these experiments validate the proposed approaches, which can be easily implemented in an existing LMS, like Modular Object-Oriented Dynamic Learning Environment (Moodle) [46] or simulation model for education development (SIMEduc) [15], and tested with real students. Simulating the learning process was a very important part of our work, which allowed us to test, adjust and correct our approaches since the very beginning, optimizing the development process. In both approaches, a huge number of tests, adjustments and corrections were done in order to achieve these results. Therefore, without using simulation, it should be impossible to come up with this approach within reasonable time, due to the large amount of time necessary to do experiments with real students and real learning processes.

**Fig. 8** Experiment 3—results from MC based approach. **a** Processing, **b** perception, **c** input, **d** understanding

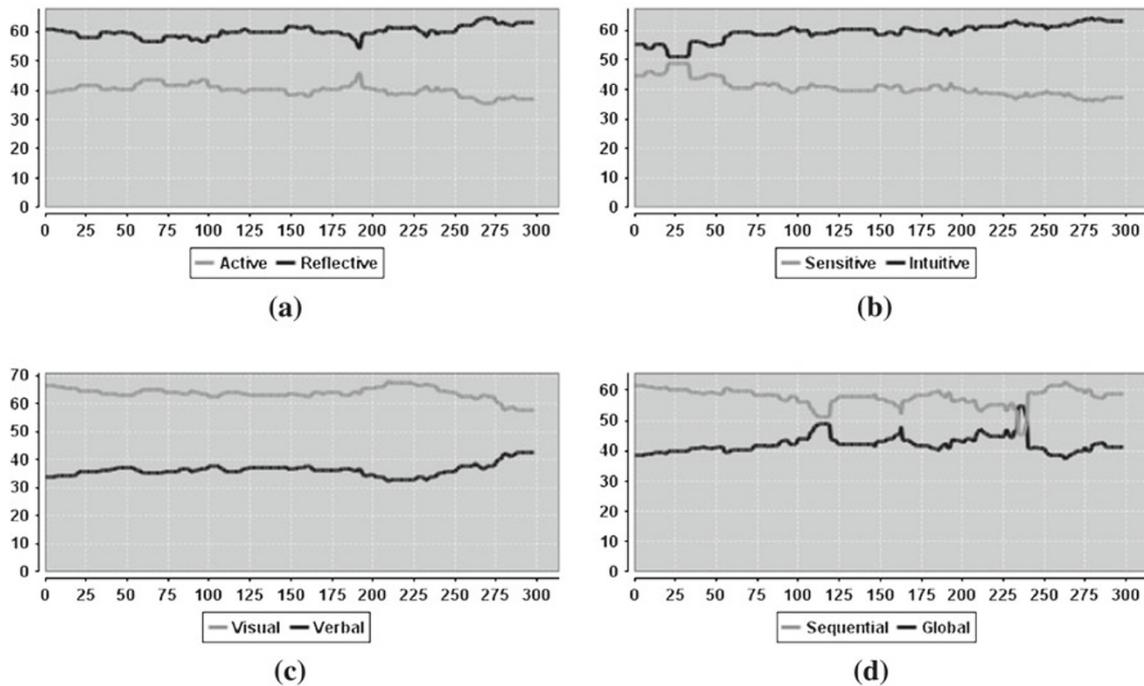


Fig. 9 Experiment 3—results from GA based approach. **a** Processing, **b** perception, **c** input, **d** understanding

Table 6 SM—experiment 4

Learning styles							
Processing		Perception		Input		Understanding	
Act.	Ref.	Sen.	Int.	Vis.	Ver.	Seq.	Glb.
70	30	70	30	70	30	70	30

Performing experiments with real students demands long-term data logs, as it can be seen in [23,51]. Moreover, without using simulation, it would be very difficult to validate the proposed approaches, due to the impossibility to know the SRLS of real students with certainty. Consequently, it would be impossible to compare the SRLS with the SPLS, and, as a result, it would be impossible to measure the consistency and efficiency of our approaches. Therefore, we consider that it is very important to initially test new approaches in simulated environments, and only after an initial study, test it in real environments.

8 Conclusions and future work

AES has been considered a promising approach to increase the efficiency in computer-aided learning. A necessary characteristic in this approach is the precise, dynamic and continuous identification of students’ LS in order to provide well-adapted learning experiences. In this context, one

challenge is the development of systems able to efficiently acquire students’ LS preferences. The information about students’ LS preferences, acquired by psychometric instruments, encloses some degree of uncertainty [49,50]. Furthermore, in most of the existing approaches, the assumptions about students’ LS, once acquired, are no longer updated. In this context, this work presents two alternative approaches to automatically detect and precisely adjust students’ LS preferences considering FS, based on the nondeterministic and nonstationary aspects of LS [25]. Because of the probabilistic and dynamic factors enclosed in students’ LS modeling, our approach gradually and constantly modifies the SM using a set of rules that detect which LS should be adjusted at a specific point of the learning process, considering the student’s performance. In this way, SM converges to the students’ real LS, considering fine-tuned strengths, as showed in Sect. 7. We found out, through experiments, that the GA based approach detects LS earlier than the Markov based approach, and consequently provides personalized content earlier, making the learning process easier. On the other hand, the Markov based approach produces more fine-tuned results than the GA based approach, taking into account the strengths of the LS. Another advantage of the Markov based approach is that it doesn’t need to keep track of a population of LSC for each student, and, consequently, it doesn’t have to spend additional computational resources in order to compute next generations of the populations during the learning process, as it is done by the GA based approach. Finally, the proposed approaches solve some important problems ignored

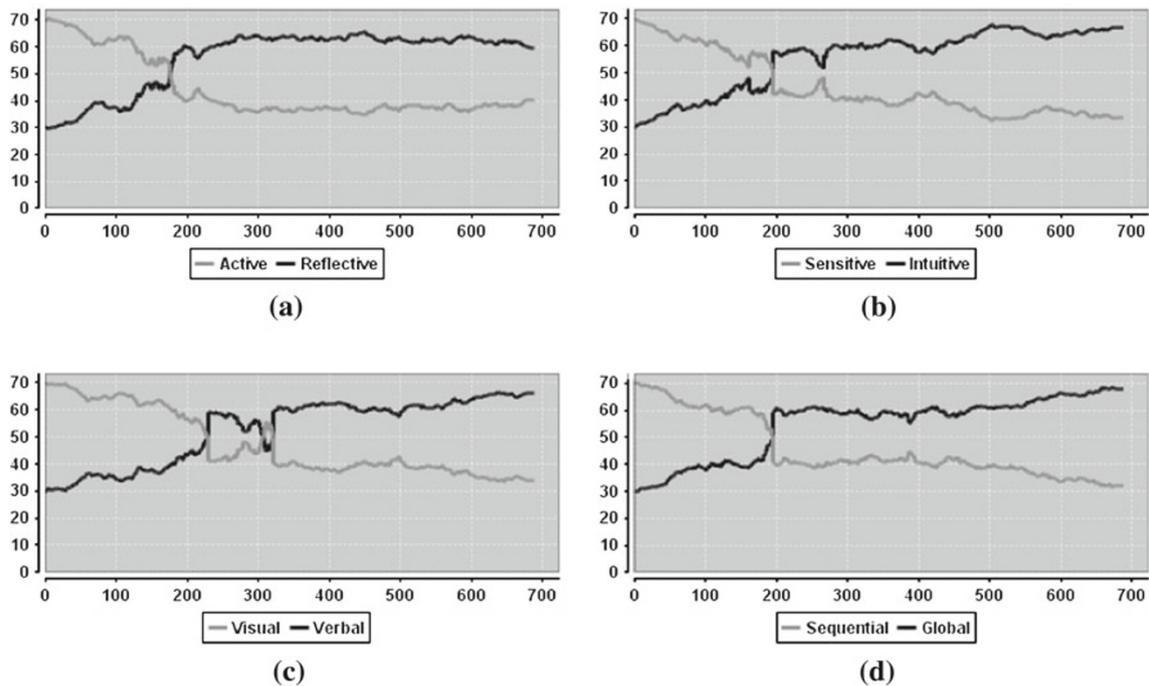


Fig. 10 Experiment 4—results from MC based approach. **a** Processing, **b** perception, **c** input, **d** understanding

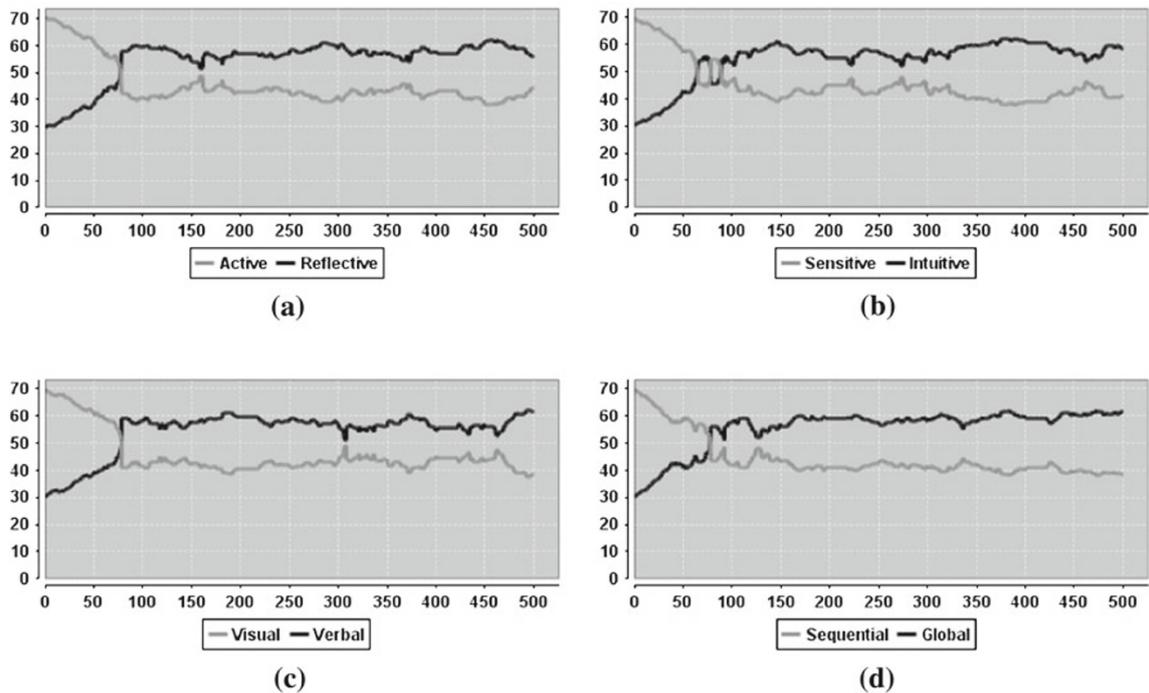


Fig. 11 Experiment 4—results from GA based approach. **a** Processing, **b** perception, **c** input, **d** understanding

in most of the analyzed approaches, and brings advantages, due to specific points, as showed in Sect. 2. The experiments with these approaches were done through computer simulation, which took into account how LS preferences exert influence on students' performances, as described by some

researchers, e.g., [1,28,31,32,38]. The evaluation of AES is a difficult task, as pointed out in [5]. Therefore, testing our approach through simulation was vital, due to the time and human resources needed to test it with real students. Now that we have achieved good results through simulation, we feel

confident to implement our approach in an existing LMS, like SIMEduc [15] and Moodle [46], and test it with real courses and real students, as a near future work. In order to achieve this goal, we are working on the development of a function able to efficiently map LO characteristics to students' LS.

References

- Alfonseca E, Carro R, Martín E, Ortigosa A, Paredes P (2006) The impact of learning styles on student grouping for collaborative learning: a case study. *User Model User-Adapted Interact (UMUAI)* 16(3):377–401
- Bloom B, Krathwohl D (1956) Taxonomy of educational objectives: the classification of educational goals. In: *Handbook I: cognitive domain*. Addison Wesley Publishing Company, New York, pp 64–81
- Bostrom L (2011) Students' learning styles compared with their teachers' learning styles in secondary schools. *Inst Learn Styles J* 1 (2011)
- Botsios S, Georgiou D, Safouris N (2008) Contributions to adaptive educational hypermedia systems via on-line learning style estimation. *Educ Technol Soc* 12(4):322–339
- Bravo J, Ortigosa A (2006) Validating the evaluation of adaptive systems by user profile simulation. In: *Proceedings of workshop held at the fourth international conference on adaptive hypermedia and adaptive web-based systems (AH2006)*, pp 479–483
- Briggs-Myers I (1957) The Myers–Briggs type indicator. Educational Testing Service, Princeton
- Brusilovskiy P (2001) Adaptive educational hypermedia. In: *International PEG conference*. Citeseer, pp 8–12
- Cabada R, Estrada M, Garcia C (2009) A fuzzy-neural network for classifying learning styles in a web 2.0 and mobile learning environment. In: *Web Congress, 2009. LE-WEB'09. Latin American*. IEEE, pp 177–182
- Carmona C, Castillo G, Millán E (2007) Discovering student preferences in E-learning. In: *Proceedings of the international workshop on applying data mining in, e-learning*
- Carmona C, Castillo G et al (2008) Designing a dynamic Bayesian network for modeling students learning styles. In: *Eighth IEEE international conference on advanced learning technologies*. IEEE, pp 346–350
- Castillo G, Gama J, Breda A (2005) An adaptive predictive model for student modeling. In: *Advances in web-based education: personalized learning, environments*. Information Science Publishing, Hershey, pp 70–92
- Chipperfield A, Fleming P, Pohlheim H, Fonseca C (1994) Genetic algorithm toolbox for use with matlab. Department of Computer Science, University of Ilmenau, Ilmenau, Germany
- Coffield F, Moseley D, Hall E, Ecclestone K (2009) Learning styles and pedagogy in post-16 learning: a systematic and critical review. National Centre for Vocational Education Research (NCVER), Berkeley
- De Bra P, Smits D, Stash N (2006) Creating and delivering adaptive courses with AHA! In: *Innovative approaches for learning and knowledge sharing*, pp 21–33
- Dorça F, Lopes C, Fernandes M (2003) A multiagent architecture for distance education systems. In: *Advanced learning technologies, 2003. Proceedings. The 3rd IEEE international conference*. IEEE, pp 368–369
- Dorça F, Lopes C, Fernandes M, Lopes R (2009) Adaptivity supported by neural networks in web-based educational systems. *J Educ Inform Cybern (JEIC)* 1
- Eiben A, Smith J (2003) *Introduction to evolutionary computing*. Springer, Berlin
- Entwistle N (1981) *Styles of learning and teaching*. Wiley, Chichester
- Felder R, Silverman L (1988) Learning and teaching styles in engineering education. *Eng Educ* 78(7):674–681
- Felder R, Spurlin J (2005) Applications, reliability and validity of the index of learning styles. *Int J Eng Educ* 21(1):103–112
- Franzoni AL, Assar S (2009) Student learning styles adaptation method based on teaching strategies and electronic media. *Educ Technol Soc* 12(4):15–29
- García P, Amandi A, Schiaffino S, Campo M (2007) Evaluating Bayesian networks' precision for detecting students' learning styles. *Comput Educ* 49(3):794–808
- Gogvadze G, Sosnovsky S, Isotani S, McLaren B (2011) Evaluating a Bayesian student model of decimal misconceptions. In: *Proceedings of the 4th international conference on educational data mining*
- Goldberg D (1989) *Genetic algorithms in search, optimization, and machine learning*. Addison-wesley, Boston (1989)
- Graf S, Kinshuk K (2009) Advanced adaptivity in learning management systems by considering learning styles. In: *Proceedings of the 2009 IEEE/WIC/ACM international joint conference on web intelligence and intelligent agent technology, vol 03*. IEEE Computer Society, pp 235–238
- Graf S, Kinshuk C (2010) A flexible mechanism for providing adaptivity based on learning styles in learning management systems. In: *2010 10th IEEE international conference on advanced learning technologies*. IEEE, pp 30–34
- Graf S, Kinshuk K (2010) Using cognitive traits for improving the detection of learning styles. In: *Database and expert systems applications (DEXA), 2010 workshop*. IEEE, pp 74–78
- Graf S, Lan C, Liu T et al (2009) Investigations about the effects and effectiveness of adaptivity for students with different learning styles. In: *2009 ninth IEEE international conference on advanced learning technologies*. IEEE, pp 415–419
- Graf S, Lin T (2007) Analysing the relationship between learning styles and cognitive traits. In: *Advanced learning technologies, 2007. ICALT 2007. Seventh IEEE international conference*. IEEE, pp 235–239
- Graf S, Liu T (2008) Identifying learning styles in learning management systems by using indications from students' behaviour. In: *Advanced learning technologies, 2008. ICALT'08. Eighth IEEE international conference*. IEEE, pp 482–486
- Graf S, Liu TC, Kinshuk K (2008) Interactions between students learning styles, achievement and behaviour in mismatched courses. In: *Proceedings of the international conference on cognition and exploratory learning in digital age (CELDA 2008)*. IADIS International Conference, pp 223–230
- Haider M, Sinha A, Chaudhary B (2010) An Investigation of relationship between learning styles and performance of learners. *Int J Eng Sci Technol* 2(7):2813–2819
- Honey P, Mumford A (1992) *The manual of learning styles*
- IEEE: LOM (Learning Object Metadata) (2010) IEEE Learning Technology Standards Committee. <http://ltsc.ieee.org/wg12/index.html>
- Jones C, Reichard C, Mokhtari K (2003) Are students learning styles discipline specific? *Commun Coll J Res Pract* 27(5):363–375
- Kaelbling L, Littman M, Moore A (1996) Reinforcement learning: a survey. *Arxiv, preprint cs/9605103* (1996)
- Kelly D, Tangney B (2005) 'First Aid for You': getting to know your learning style using machine learning. In: *Advanced learning technologies, 2005. ICALT 2005. Fifth IEEE international conference*. IEEE, pp 1–3

38. Kinshuk K, Liu T-C, Graf S (2009) Coping with mismatched courses: students' behaviour and performance in courses mismatched to their learning styles. *Educ Technol Res Develop* 57(6):739–752
39. Kolb D (1984) *Experiential learning: experience as the source of learning and development*. Prentice-Hall, Englewood
40. Kuljis J, Liu F (2005) A comparison of learning style theories on the suitability for elearning. In: Proceedings of the IASTED conference on web technologies, applications, and services, pp 191–197
41. Limongelli C, Sciarrone F, Temperini M, Vaste G (2009) Adaptive learning with the LS-plan system: a field evaluation. *IEEE Trans Learn Technol* 2(3):203–215
42. Lopes R, Dorça F, Fernandes M, Lopes C (2008) Um sistema de avaliação em EAD baseado em lógica Fuzzy. In: Simpósio Brasileiro de Informática na Educação. Brazilian Computer Society (SBC), pp 30–34
43. Martins AC, Faria L, Vaz de Carvalho C, Carrapatoso E (2008) User modeling in adaptive hypermedia educational systems. *Educ Technol Soc* 11(1), 194–207 (2008)
44. Messick S (1976) Personal styles and educational options. In: *Individuality in, learning*. Jossey Bass, San Francisco, pp 327–368
45. Meyn S, Tweedie R, Glynn P (2009) *Markov chains and stochastic stability*, vol 2. Cambridge University Press, Cambridge
46. Moodle (2010) <http://www.moodle.org/>
47. Papoulis A, Pillai S, Unnikrishna S (2002) *Probability, random variables, and stochastic processes*, vol 73660116. McGraw-Hill, New York
48. Pask G (1976) Styles and strategies of learning. *Br J Educ Psychol* 46:128–148
49. Price L (2004) Individual differences in learning: Cognitive control, cognitive style, and learning style. *Educ Psychol* 24(5):681–698
50. Roberts M, Erdos G (1993) Strategy selection and metacognition. *Educ Psychol* 13(3):259–266
51. Shein P, Chiou W (2011) Teachers as role models for students' learning styles. In: *Social behavior and personality*, 2011 39(8), 1097–1104. <http://dx.doi.org/10.2224/sbp>
52. Thompson J (1996) *Student modeling in an intelligent tutoring system*. Phd thesis. Faculty of the Graduate School of Engineering of the Air Force Institute of Technology
53. Van Zwanenberg N, Wilkinson L, Anderson A (2000) Felder and silverman's index of learning styles and honey and mumford's learning styles questionnaire: how do they compare and do they predict academic performance? *Educ Psychol* 20(3):365–380
54. Virvou M, Troussas C (2011) Web-based student modeling for learning multiple languages. In: *Information society (i-Society)*, 2011 international conference. Dept. of Inf., Univ. of Piraeus, Athens
55. Yannibelli V, Godoy D, Amandi A (2006) A genetic algorithm approach to recognize students' learning styles. *Interact Learn Environ* 14(1), 55–78. doi:10.1080/10494820600733565
56. Zatarain R, Barrón-Estrada L, Reyes-García C, Reyes-Galaviz O (2011) Applying intelligent systems for modeling students' learning styles used for mobile and web-based systems. *Studies in computational intelligence*, vol 318. Springer, Berlin, pp 3–22
57. Zatarain-Cabada R, Barrón-Estrada M, Zepeda-Sánchez L, Sandoval G, Osorio-Velazquez J, Urias-Barrientos J (2009) A Kohonen network for modeling students' learning styles in web 2.0 collaborative learning systems. *MICAI (2009) Advances in Artificial Intelligence*, pp 512–520