SaGe Framework - Mapping of SARSA to Adaptive e-Learning using Learning Styles

Balasubramanian Velusamy[1], S. Margret Anouncia[2], George Abraham[3]
[1,2,3]School of Computing Science and Engineering
VIT University, Vellore, India – 632 014.
[1] vbalasubramanian@vit.ac.in
[2] smargretanouncia@vit.ac.in
[3] george.abraham1912@gmail.com

***Abstract* - This paper proposes a mathematical framework – the SaGe framework, which maps the adaptive nature of the e-learning environment to the reinforcement learning approach. An adaptive e-learning system works in a similar fashion to an intelligent agent that assesses the various interactions of the user with the system, learns from these interactions and uses this knowledge base to select the best possible action to earn the maximum reward. This is the same methodology of reinforcement learning. The adaptive nature of the e-learning environment is provided by the assessment of the individual differences in the learning styles of the students. In our approach we chose the learning styles provided by the dimensions of the Felder-Silverman Learning Style Model (FSLSM) and the SARSA algorithm for the reinforcement learning. Mapping of the reinforcement SARSA algorithm to an adaptive e-learning system is asserted by the time-based update of the Q-values of the SARSA algorithm.**

Keywords - e-Learning, Reinforcement Learning, SARSA, Learning Styles, FSLSM.

I. INTRODUCTION

With the increase in the popularity and importance of the Internet, the face of the education system has changed due to the introduction of e-Learning. The demands of e-learning systems that cater to all needs in various fields of educations has been ever increasing [9][13]. There are different types of learners present. Some may be fast learners while others may be slow, some may need more graphical images or representations to understand a concept better while other may understand by just reading some text. As a result of all these variations, personalization is one of the most important characteristics of e-learning. The various preferences and requirements of an individual can be captured in a learner model that can be extracted from personality factors like learning styles, behavioural factors like user's browsing history and knowledge factors like user's prior knowledge [4]. Majority of the research work carried out are based on the learning styles as these are the most dynamic and give the best results if catered to properly [5].

The main challenge is the detection of the learning styles. Researchers have described various learning styles models like Dunn & Dunn [18], Myers-Briggs [10], Kolb [3], Felder-Silverman [21] and Honey & Mumford [16]. Research has proved that the Felder-Silverman Learning Style Model (FSLSM) is the most suited for the engineering students' environment as it also considers the psychological aspects of a person [20]. The Index of Learning Styles [19] is a questionnaire-based approach for detection of learning styles based on the FSLSM. The problem with questionnaire-based approach is that it suffers from the "inaccurate self-conceptions of students" [17][23][24] at a specific time. Moreover these questionnaires are incapable of tracking the changes in a learner's learning style.

As a result of these problems, various researches have been conducted to come out with alternate automated solutions for learning style detection. These works can be broadly classified into two groups: data-driven approach and literature-based approach. Some of the noticeable works in the data-driven approach are by using Bayesian Networks [14][15], NBTree classifiers [4] and Genetic Algorithms [26]. Literature-based approach is a relatively new method with some of the noticeable works being done by Graf *et al.* [23][24], Dung and Florea [17] and Simsek *et al.* [12]. Abraham *et al.* [7] have made a detailed study of almost all the prominent works done in this domain.

Personalization of the e-Learning system means to provide a system that adapts according to the learners' learning process. In the next section, we will talk about adaptive web-based education.

II. ADAPTIVE WEB-BASED EDUCATION

The idea of an adaptive system was initially stressed by Bursilovsky and Peylo [13]. They talk about improving the system of web-based education by providing an Adaptive and Intelligent Web-Based Educational System (AIWBES) as an alternative to the traditional systems. AIWBES adapts to the learners' needs, knowledge and behaviour like a human teacher would do. An adaptive system modifies its solutions to a problem based on various factors, for instance the learners' previous experience with the system whereas an intelligent system provides the same solution irrespective of the different needs of the learners. AIWBES is a mixture of adaptive hypermedia technologies and intelligent tutoring technologies. It also contains adaptive information filtering, intelligent monitoring and intelligent collaborative learning. Adaptive hypermedia mainly consists of adaptive presentation and adaptive navigation support while intelligent tutoring mainly consists of curriculum sequencing, problem solving support and intelligent solution analysis.

III. REINFORCEMENT LEARNING

In Computer Science, reinforcement learning [22] is an area of machine learning concerned with what actions an agent, i.e. an intelligent program, should take in an environment so as to maximize the cumulative reward. The agent learns by sensing various parameters from the environment, exploring various possibilities for a better reward and when performing the same task again, exploiting the already learned best path.

The e-learning system that we are talking about is similar to an intelligent agent. It senses the various user interactions and has to decide on the best possible responses to the user so as to enhance the learning experience of the user. If the same user uses the system again, then the previously learned options that were best suited to the person should be given again, if not for a better option.

Reinforcement learning scenarios are described by states, actions and rewards. There exist two main reinforcement learning algorithms – Q-Learning Algorithm [1][2] and SARSA Algorithm [6].

IV. SARSA ALGORITHM

SARSA [6] algorithm is an improvement on the Q-learning [2] algorithm which is a form of model-free reinforcement learning [1]. The problem domain consists of an agent, its various states S, and a set of actions per state A. The agent can move from one state to another by performing some action a A. The transition, i.e. the next state gives a reward to the agent. The goal of the agent is to maximize the total reward. This is achieved by optimizing the actions for each state. Hence, there exists a function Q that calculates the quality of each state-action combination. Initially Q returns a fixed value that the designer has set. Then during each step when the agent is rewarded, new values are calculated and updated [25].

In SARSA algorithm, the rule that updates the Q-value depends on the current state s_t, the action the agent choses a_t, the reward r, the next state that the agent will be in after taking the action s_{t+1} and the action that the agent will take in the new state a_{t+1}. The name SARSA stands for the same – state, action, reward, state (next state), action (next action).

The following represents the SARSA algorithm [6][8][25].

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[R(s_t, a_t) + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)] \tag{1}$$

where,

$\leftarrow$ - Updating the old value
t - Current interaction
t+1 - Next interaction
$Q(s_t, a_t)$ - The Q-values of the current interaction.
$R(s_t, a_t)$ - Reward obtained for performing action a_t in s_t
α - Learning rate ($0 \leq \alpha \leq 1$)
γ - Discount factor that decides the importance of future rewards ($0 \leq \gamma < 1$)

SARSA algorithm is also called as an “On-Policy” of “Policy Dependent” algorithm as it depends on the decision process used to select an action given a certain state. It does not have a greedy approach like the Q-learning algorithm.

The policy selection process is not always selecting the action that results in the maximum Q-value as this will lead to a phenomenon of “local maxima”. Instead, it is determined on a factor epsilon ε which determines the extent to which the actions are randomized.

There are three types of action selection policies:

(1) ε – greedy
Action with the highest estimated reward is chosen independent of the Q-value estimates.

(2) ε – soft
Best action is selected with a probability of 1 - ε and rest of the time the actions are selected uniformly.

(3) Softmax
The problem with both the above methods is the uniform selection that may result in the worst possible action being selected as second best. A solution to this is the softmax policy where a rank or a weight is assigned to each action according to its action-value estimate and actions are selected based on these weights, as a result of which the worst actions are unlikely chosen.

In the SARSA equation, if the discount factor (γ value) is set to 0, i.e. $\gamma = 0$ then from Eq. (1) we get,

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[R(s_t, a_t) - Q(s_t, a_t)] \quad (2)$$

This means that the updating can happen then and there itself without the consideration of a new state.

Now, in Eq. (2) if the learning rate (α value) is set to 0, i.e. α=0 then

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) \quad (3)$$

This means that no learning takes place and the value remains as it is.

Else, if in Eq. (2) the learning rate (α value) is set to 1, i.e. α=1 then

$$\begin{aligned} Q(s_t, a_t) &\leftarrow Q(s_t, a_t) + R(s_t, a_t) - Q(s_t, a_t) \\ &\leftarrow R(s_t, a_t) \end{aligned} \quad (4)$$

This means that the agent will consider only the most recent information, i.e. the reward only. This is too simplistic.

If α=0.5 (for example), the old and the new Q-values meet half way, given the reward.

Hence, when the discount factor γ is set to 0, the agent will consider only the current reward. It has only short term greedy goals.

In the SARSA algorithm, if the discount factor (γ value) is set to 1, i.e. γ=1 and also α=1 then from Eq. (1) we get,

$$\begin{aligned} Q(s_t, a_t) &\leftarrow Q(s_t, a_t) + R(s_t, a_t) + Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t) \\ &\leftarrow R(s_t, a_t) + Q(s_{t+1}, a_{t+1}) \end{aligned} \quad (5)$$

This means that the updated Q-value for a state is equal to the sum of reward and the Q-value of the next state. As γ tends to 1 or more, the more it optimizes for long-term goals. It is ideally set in between 0 and 1 [25].

V. SaGe FRAMEWORK – A GENERIC MAPPING TO E-LEARNING USING LEARNING STYLES

As of now, we are restricting the learning styles based on the FSLSM model to only the domain of VISUAL or VERBAL learners. In this section, we propose the SaGe Framework that models the adaptive e-learning environment using the SARSA algorithm. The SaGe framework arises from the necessity to take into consideration the dynamic nature of the learning style preference of the students, a factor that is not considered in the standard evaluation procedure using the FSLSM model, i.e. the Index of Learning Styles (ILS) questionnaire. The ILS questionnaire is a self-calculating online available tool that contains a set of 44 questions that is used to assess the learning styles of the students based on the responses gathered. As we have discussed earlier, the adaptive learning process in an e-learning environment is similar to the reinforcement learning process, and hence SARSA algorithm is best suited for the same.

Fig. 1 explains the generalized structure of how the mapping can be done between SARSA algorithm and learning on the basis of learning styles, while Table I gives the tabular representation of the same. Consider the following nomenclature for the figure:

- "VI" and "VE" denotes the states of learners, i.e. Visual and Verbal, respectively.
- "VI_LQ" denotes the action of providing a Visual Lesson followed by a Quiz.
- "VE_LQ" denotes the action of providing a Verbal Lesson followed by a Quiz.
- "+1" denotes a positive reward
- "-1" denotes a negative reward.
- The rewards are calculated as follows:

  If (Quiz_Marks >= 5) {Positive Reward, i.e. +1}

  Else {Negative Reward, i.e. -1}

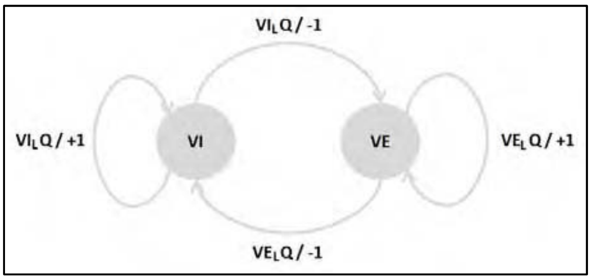

Fig. 1. SaGe Framework – State Transition Diagram

TABLE I
Tabular Representation of State Transition Diagram

State (s_t)	Action (a_t)	Reward (R)	Next State (s_{t+1})
VI	VI_LQ	+1	VI
VI	VI_LQ	-1	VE
VE	VE_LQ	+1	VE
VE	VE_LQ	-1	VI

For mapping this learning process to the SARSA algorithm, we need a few assumptions:

- The initial state of the student is "VI"
- We use softmax policy for selecting the next action in which weights are assigned to the actions based on their action-value estimate; hence an action with a +1 reward gets the maximum weight. It has also been observed that use of softmax selection policy for action selection yields in better results [11].
- The discount factor γ is set to 1 and the learning rate α is assumed to be 0.8, based on the proof mentioned in Section IV.
- Initial Q-value of all states is 0.

When we observe both Fig. 1 as well as Table I, we can see that the student moves from state VI back to the same state on selecting the best action. Once it moves to VI, it will again select the same action, for best case scenario.

From Eq. (1), we get:

$$\begin{aligned} \mathrm{Q}(\mathrm{VI},\mathrm{VI_LQ}) &\leftarrow \mathrm{Q}(\mathrm{VI},\mathrm{VI_LQ}) + \alpha[\mathrm{R}(\mathrm{VI},\mathrm{VI_LQ}) + \gamma \mathrm{Q}_{t+1}(\mathrm{VI},\mathrm{VI_LQ}) - \mathrm{Q}(\mathrm{VI},\mathrm{VI_LQ})] \quad (6)\\ &\leftarrow 0 + 0.8[1 + 1(0) - 0] \\ &\leftarrow 0.8 \end{aligned}$$

In the same way, we can continue to calculate the Q-values of all the state-action combination and populate the Q-table which is used for further analysis. When the student uses the system again, the system is now aware of the quality of each of the possible state-action pairs and hence by default can suggest the student the best action that it needs to take. This suggestion will be as per its assessed learning style.

However there is also a possibility that the student can choose some other action apart from the one that is suggested. This may be due to the fact that for that specific topic, he/she finds it easier to understand in the changed manner. In this scenario, the Q-values of the student are again calculated and the table updated. This indicates the property of personalization and adaptation of the system as per the user's needs.

VI. ANALYSIS

In this analysis, we try to showcase both the scenario when the learning style of the student remains the same for the repeated evaluation as well as when it changes. The actual analysis was done on a class of 35 students of Masters of Technology in Computer Science and Engineering from VIT University, Vellore, India and the subject under consideration was Web Services. For better portrayal of the essence of the results and discussion, we have assumed a narrowed down space set of a class of 5 students.

For providing a proper analysis and support that the SaGe framework is a better approach towards modelling the adaptive learning process, we first asses the learning styles of the students based on the existing approach, i.e. the Index of Learning Style (ILS) questionnaire. Table II gives the learning style preferences of the students, as assessed from the ILS tool.

TABLE II
Learning Styles of Students based on ILS

Student	Learning Style
S1	VI (Visual)
S2	VE (Verbal)
S3	VI (Visual)
S4	VI (Visual)
S5	VE (Verbal)

Initially no information about the learning style preference (visual or verbal) of the student is available. In the first round of evaluation, all the students are given both types of lessons one at a time followed by a quiz on that respective lesson to analyse their understanding about what they have understood. There are two visual and two verbal lessons, conducted one at a time followed by their respective quiz. Table III gives the complete summary of the performance of all the students in Round 1.

TABLE III
Summary of Round 1 Assessment

Student	Lesson	Quiz	Reward
S1	VI_1	Q1	-1
S1	VE_1	Q2	-1
S1	VI_2	Q3	+1
S1	VE_2	Q4	-1
S2	VI_1	Q1	-1
S2	VE_1	Q2	+1
S2	VI_2	Q3	-1
S2	VE_2	Q4	+1
S3	VI_1	Q1	+1
S3	VE_1	Q2	-1
S3	VI_2	Q3	+1
S3	VE_2	Q4	-1
S4	VI_1	Q1	+1
S4	VE_1	Q2	-1
S4	VI_2	Q3	+1
S4	VE_2	Q4	+1
S5	VI_1	Q1	-1
S5	VE_1	Q2	+1
S5	VI_2	Q3	+1
S5	VE_2	Q4	+1

From Table III, we can assess the learning style inclination of the students.

Consider for student S1, (VI1 + VI2) = 0 while (VE1 + VE2) = -2. Therefore, as (VI1 + VI2) > (VE1 + VE2), hence S1 has an inclination towards visual style of learning.

Computing in the same way, Table IV gives the learning style preference of all the students after the 1st round of assessment.

TABLE IV
Learning Style Preference after Round 1

Student	Learning Style
S1	VI (Visual)
S2	VE (Verbal)
S3	VI (Visual)
S4	VI (Visual)
S5	VE (Verbal)

After analysing the learning styles of the students, they are then given lessons and quizzes based only on their learning styles and are evaluated again. Hence, S1, S3 and S4 gets only visual contents while S2 and S5 gets only verbal contents. This is similar to the fact that in SARSA algorithm, once the Q-values are known of the students, then when the students interact with the system again, the best actions as per the Q-values are selected. Assuming that the students are again subjected to a series of four lessons, as shown in Table V:

TABLE V
Summary of Round 2 Analysis

Student	Lesson	Quiz	Reward
S1	VI_1	Q1	+1
S1	VI_2	Q2	+1
S1	VI_3	Q3	+1
S1	VI_4	Q4	-1
S2	VE_1	Q1	+1
S2	VE_2	Q2	+1
S2	VE_3	Q3	+1
S2	VE_4	Q4	+1
S3	VI_1	Q1	-1
S3	VI_2	Q2	-1
S3	VI_3	Q3	+1
S3	VI_4	Q4	-1
S4	VI_1	Q1	+1
S4	VI_2	Q2	-1
S4	VI_3	Q3	+1
S4	VI_4	Q4	+1
S5	VE_1	Q1	-1
S5	VE_2	Q2	-1
S5	VE_3	Q3	-1
S5	VE_4	Q4	+1

From Table V, the learning styles of the students are again assessed. If the sum of all the rewards is non-zero, then the learning style of the student stays as it is, else it is an indication in the change of learning style.

Considering for student S1, $(VI_1 + VI_2 + VI_3 + VI_4) = 2 > 0$, i.e. the student S1 still has a visual learning style whereas for student S3, $(VI_1 + VI_2 + VI_3 + VI_4) = -2 < 0$, i.e. the student S3 has a change in the learning style preference from visual to verbal style.

Computing in the same manner, Table VI gives the learning style preference of the students after Round 2 evaluation.

TABLE VI
Learning Style Preference after Round 2

Student	Learning Style
S1	VI (Visual)
S2	VE (Verbal)
S3	VE (Verbal)
S4	VI (Visual)
S5	VI (Visual)

It is observed that after Round 2 evaluation, the learning style preferences of S3 and S5 changed from their earlier analysed preference. Hence the content delivered to them again needs to be adapted as per their learning style. This portrays the essence of adaptation in the e-learning system on the basis of learning styles and also maps to the process of computing new Q-values as explained in the SARSA mapping.

When the same set of students come for learning again, the type of content needs to be changed as per their preferences that has been learned after the second evaluation.

TABLE VII
Summary of Round 3 Analysis

Student	Lesson	Quiz	Reward
S1	VI_1	Q1	+1
S1	VI_2	Q2	-1
S1	VI_3	Q3	+1
S1	VI_4	Q4	-1
S2	VE_1	Q1	+1
S2	VE_2	Q2	+1
S2	VE_3	Q3	+1
S2	VE_4	Q4	+1
S3	VE_1	Q1	-1
S3	VE_2	Q2	+1
S3	VE_3	Q3	+1
S3	VE_4	Q4	+1
S4	VI_1	Q1	+1
S4	VI_2	Q2	+1
S4	VI_3	Q3	-1
S4	VI_4	Q4	+1
S5	VI_1	Q1	+1
S5	VI_2	Q2	-1
S5	VI_3	Q3	+1
S5	VI_4	Q4	+1

Table VII gives the summary of the rewards of all the students after Round 3 of evaluation. It is again observed that the learning style preference of student S1 changes from visual to verbal, even though it was constant in the first two rounds while S2 and S4's preferences remain same throughout, as shown in Table VIII.

TABLE VIII
Learning Style Preference after Round 3

Student	Learning Style
S1	VE (Verbal)
S2	VE (Verbal)
S3	VE (Verbal)
S4	VI (Visual)
S5	VI (Visual)

On analysing Table III, Table V and Table VII, we can see that there is a possibility of the learning styles of the students to remain constant as well as vary. This variation is portrayed in Fig. 2.

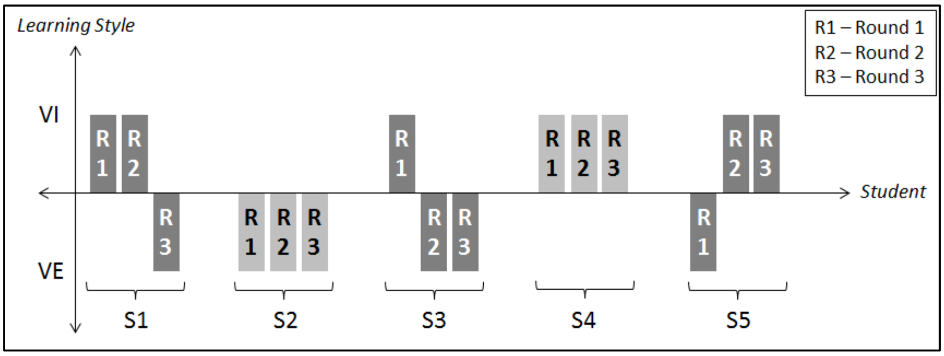

Fig. 2. Learning Style Variation of the Students

From this variation, it can be observed that the change is learning style will definitely occur, as it can be affected due to the complexity of the topic being learnt and other external factors even if we restrict the learning to the same subject. The intention of the system is to adapt to the change in learning style based on what is being learnt after each round of evaluation, hence the name reinforcement learning.

Considering the majority preference of the students, Table IX gives their final learning style preference.

TABLE IX
Final Learning Style Preference

Student	Learning Style
S1	VI (Visual)
S2	VE (Verbal)
S3	VE (Verbal)
S4	VI (Visual)
S5	VI (Visual)

For finding the precision of the approach, we use the formula mentioned in [15]:

$$\text{Precision} = \frac{\sum_{i=1}^{n} Sim\ (LS_{SaGe}, LS_{ILS})}{n} \quad (7)$$

In Eq. (7), *Sim* stands for similarity and is equal to 1 if the values, obtained using the SaGe framework and the ILS, are same and 0 if they are opposite; *n* stands for the number of students studied upon.

Comparing Table II and Table IX, we obtain from Eq. (7),

$$\begin{aligned}\text{Precision} &= (1+1+0+1+0)/5 \\ &= 3/5 \\ &= 0.6 \equiv 60\%\end{aligned} \quad (8)$$

This precision obtained in Eq. (8) is for the normalized sample set consisting of just 5 students. On evaluating the actual sample set of 34 students, the precision obtained was 79.41% for the input dimension of the FSLSM, where there was a similarity for 27 out of the 34 students studied upon.

VII. CONCLUSION

The SaGe Framework, i.e. a reinforcement approach toward adaptive e-learning using SARSA algorithm has been adapted and analysed in engineering education. A time-based update of Q-values leads to the identification of the individual differences in learning styles dynamically. Here, the learners' visual and verbal preference based on the FSLSM dimensions has been analysed and the framework gives a precision of 79.41%. This framework is a mathematical evaluation model and can be further expanded using technologies like semantic agents and can also be expanded to incorporate the other dimensions of FSLSM.

ACKNOWLEDGEMENT

The authors would like to thank the Cognitive Learning Research Laboratory (CLRL) at the School of Computing Science and Engineering (SCSE), VIT University, Vellore, India for providing the support and the opportunity to carry out this research work.

REFERENCES

[1] C. J. C. H. Watkins and P. Dayan, "Q-Learning," *Machine Learning*, Vol. 8, pp. 279 – 2992, 1992.
[2] C. J. C. H. Watkins, "Learning from delayed rewards," PhD Thesis, University of Cambridge, England, 1989.
[3] D. A. Kolb, *Experiential Learning: Experience as the Source of Learning and Development*, Prentice-Hall, Englewood Cliffs, New Jersey, 1984.
[4] E. Ozpolat and G. B. Akbar, "Automatic detection of learning styles for an e-learning system," *International Journal of Computer and Education*, Elsevier, Vol. 53, pp. 355-367. DOI: 10.1016/j.compedu.2009.02.018, 2009.
[5] F. Dag and A. Geycer, "Relations between online learning and learning styles," *Procedia Social and Behavioral Sciences - I*, Elsevier Publications, pp. 862–871, 2009.
[6] G. A. Rummery and M. Niranjan, "On-Line Q-Learning Using Connectionist Systems," Cambridge University, CUED/F-INFENG/TR166, 1994.
[7] G. Abraham, V. Balasubramanian and RA. K. Saravanaguru, "Semantic Agents for Learning Style Recognition in an e-Learning Environment," *International Journal of Evaluation and Research in Education (IJERE)*, Vol. *2*(1), 2013.
[8] H. A. Van Hasselt (2013) Short Introduction To Some Reinforcement Learning Algorithms. [Online]. Available: http://homepages.cwi.nl/~hasselt/rl_algs/Sarsa.html
[9] J. Nath, S. Ghosh, S. Agarwal and A. Nath, "E-Learning Methodologies and its Trends in Modern Information Technology," *Journal of Global Research in Computer Science*, Vol. 3(4), pp. 48–52, 2012.
[10] M. I. Briggs and P. B. Myer, *Gifts Differing: Understanding Personality Type*, Mountain View, CA: Davies-Black Publishing. ISBN 0-89106-074-X, 1980.
[11] M. Tokic and G. Palm, "Value-Difference Based Exploration: Adaptive Control between Epsilon-Greedy and Softmax," *KI 2011: Advances in Artificial Intelligence*, Vol. 7006, pp. 335 – 346, 2011
[12] O. Simsek, N. Atman, M. M. Inveigle and Y. D. Arian, "Diagnosis of Learning Styles Based on Active or Reflective Dimension of Felder and Silverman's Learning Style Model in a Learning Management System," in *Proc. ICCSA'10*, 2010, Part II, LCNS 6017, pp. 544 – 555.
[13] P. Bursilovsky and C. Peylo, "Adaptive and Intelligent Web-Based Educational Systems," *International Journal of Artificial Intelligence in Education*, Vol. 13, pp. 56–169, 2003.
[14] P. Garcia, S. Schiaffino and A. Amandi, "eTeacher: Providing personalized assistance to e-learning students," *International Journal of Computer and Education*, Elsevier Publications, Vol. 51 pp. 1744–1754, 2008.

[15] P. Garcia, S. Schiaffino, A. Amandi and M. Campo, "Evaluating Bayesian networks' precision for detecting students' learning styles," *International Journal of Computer and Education*, Elsevier Publications, Vol. 49, pp. 794–808. 2007.
[16] P. Honey and A. Mumford, *The Manual of Learning Styles*, Maidenhead, 1992.
[17] P. Q. Dung and A. M. Florea, "An approach for detecting learning styles in learning management systems based on learners' behaviors," in *Proc. IPEDR'12*, 2012, IACSIT Press, Singapore, pp:171-177.
[18] R. Dunn and K. Dunn, *Teaching students through their individual learning styles: A practical approach*, Reston, VA: Reston Publishing Company, 1978.
[19] R. M. Felder and B. A. Soloman (2012) Index of Learning Styles Questionnaire. [Online]. Available: http://www.engr.ncsu.edu/learningstyles/ilsweb.html
[20] R. M. Felder and J. Spurlin, "Applications, Reliability and Validity of the Index of Learning Styles," *International Journal of Engineering Education*, Vol. 21(1), pp. 103-112, 2005.
[21] R. M. Felder and L. K. Silverman, "Learning and Teaching Styles in Engineering Education," *Engineering Education*, Vol. 78(7), pp.674-681, 1988.
[22] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, MIT Press, Cambridge, MA. ISBN: 0-262-19398-1, 1998.
[23] S. Graf, "Adaptivity in Learning Management Systems Focussing on Learning Styles," PhD Thesis, Vienna University of Technology, 2007. [Online]. Available: sgraf.athabascau.ca/publications/PhDthesis_SabineGraf.pdf
[24] S. Graf, Kinshuk and T. C. Liu, "Identifying Learning Styles in Learning Management Systems by Using Indications from Students' Behaviour," in *Proc. ICALT'08*, 2008, pp. 482-486. DOI: 10.1109/ICALT.2008.84
[25] T. Eden, A. Knittel and R. Van Uffelen (2013) Tutorial on Reinforcement Learning. [Online]. Available: http://www.cse.unsw.edu.au/~cs9417ml/RL1/index.html
[26] Y. C. Chang, W. Y. Kao, C. P. Chu and C. H. Chiu, "A learning style classification mechanism for e-learning," *International Journal of Computer and Education*, Elsevier Publications, Vol. 53, pp. 273–285, 2009.