Using Q-Learning for Recommend Learning Object on e-Learning System

by M Said Hasibuan

Submission date: 12-Jan-2021 07:08PM (UTC+0700)

Submission ID: 1486269651

File name: eecsiok.pdf (279.34K)

Word count: 1535

Character count: 7838

Using Q-Learning for Recommend Learning Object on e-Learning System

Hasibuan S Muhammad
Information System Institute Business and Informatics Darmajaya

ABSTRACT
This paper discusses the software development used for the selection of instructional materials based on an appropriate learning style. Q-Learning is used to estimate state-action learners using Q-Learning. It is conducted by calculating the value of the action done to reward students in selecting teaching materials, so the pattern matches are found. This study uses the Felder Silverman Learning Style Model (FSLSM). Tests showed that the system is able to perform the selection (action) of instructional materials for teaching materials that match learning styles. Q-Learning Lessons can be applied to optimize the selection of instructional materials with minimal effort.

Keywords ; Q-learning; learning style; e-learning

I. INTRODUCTION

There are several theories refer to learning styles, such as Felder-Silverman learning style model (FSLSM) [1], Honey and Mumford [2], and Kolb's Learning style model [3], and VARK learning style model [4].

According to Felder and Silverman , learners with a strong preference for a specific learning style might have difficulties in learning if their learning styles are not considered by the teaching environment. On the other hand, providing courses that fit the learning styles of learners makes learning easier for them and leads to better progress. However, for providing adaptivity, the learning styles of students need to be known first. Therefore, student modelling is a crucial aspect of each adaptive system.

In this paper, we introduce an automatic student modelling approach for detecting learning style preferences according to the Felder-Silverman learning style model (FSLSM). We use Q-Learning [5] as an algorithm of reinforcement learning, which is an asynchronous method of Dynamic Programming (DP). This algorithm provides an agent which is able to act in optimal learning. A similar study was carried out by [6] using Temporal Differential (TD), where the agent tries to perform an action on a particular state and an evaluation of the rewards is received. It is conducted by calculating the value of the action done to reward students in selecting teaching materials, so the pattern matches are found. Tests show that the system is able

to perform the selection (action) of instructional materials for teaching materials that match learning styles. Q-Learning Lessons can be applied to optimize the selection of i instructional materials with minimal effort. The next step repeats all the actions in all states, demonstrating that learning as a whole will get the optimal reward.

The rest of the paper is structured as follows: section 2 describes the theory foundation of the Q-Learning algorithm and the calculation of the value of Q, section 3 presents the methods of the software testing selection of instructional materials with eight styles of learning, and in section 4 we conclude our experiments.

II. LEARNING ALGORITHM

The development of computer science has evolved to intelligent systems. An intelligent system is able to perform actions similar to the human brain functions in solving problems. Reinforcement Learning is one of the developments of the artificial intelligence branch that can be applied to the machine or robot to determine the ideal action automatically in certain circumstances.

Furthermore, Graf [7] used Bayes' method to capture the learning style, however his method has not explained the delivery learning object.

In order to maximize system performance, Q-Learning is one simple way for agents learning the reinforcement learning method to obtain optimal performance. The Q-Learning algorithm is based on MDP (Markov Decision Problem) with discreet state and action. At time t, an agent performs an action at the state st to the next state st +1 and reward r (st, at).

The purpose of this study is to obtain the maximum number of policies, where policy Π is an action that was chosen to obtain the optimal reward. The relationship of each parameter in the reinforcement learning method is as follows:

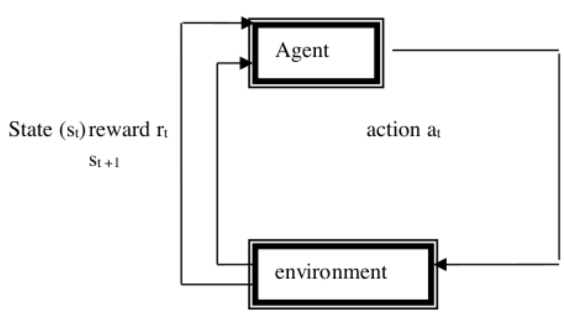

Figure 1.relation on system

The value Q (s, a) is the approximate number of the expected reward, which is obtained by selecting an action in a state which is the optimal policy. The optimal action is from each state, where the value Q(s,a) is the highest. The calculation of the value of Q(s,a) depends on the value of the transition probability if the transition probability P (st, at) and P (st, at) (rt) is known, then the function value can be calculated by:

$$Q(s_t,a_t)= \sum_{t=0}^{\infty} r_t\, P_{(st,at)}\, (r_t) + Y \sum_{t=0}^{\infty} V^{\Pi}{}_{(st+i)}\, P_{\ (st,at)}\, (S_{t+1}) \ldots. \ (1)$$

If the transition probability P (st, at) (St +1) and P (st, at) (rt) is not known, then the function value can be calculated by:

$$Q(s_t,a_t) = (r_{t\,+}\, Y \max Q\ (s_t + i, a_t + 1) \ldots\ldots\ldots\ldots\ldots..(2)$$

III. METHODS

In the selection of teaching materials, in this study there are three basic parameters used for this algorithm, namely student as state, selection of learning style as action, and reward as the result of the use of teaching materials' evaluation. Developing software in this research, we have performed tests to determine the appropriate instructional materials to the eight learning styles and teaching materials objectives. Each learning style features a learning object. The following figure describes the mechanism of the e-learning system.

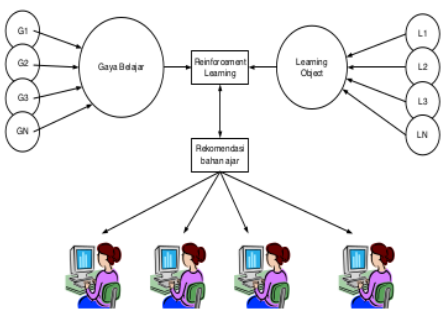

Figure 2 Architecture System

IV. RESULTS

Our work shows there are four states and eight actions in this experiment and it performed the calculation of the value of reward for each action. After a comparison of reward values, there will be the selection of action with a maximum value. From the test results, the reward value of each action can be known, so it can be used for the determination of learning styles and teaching materials in the next lesson. There are eight learning styles namely G1, G2, G3, G4, G5, G6, G7, and G8 as the identification of learning styles possessed. Each state is described as a student and there are several actions that can be done, namely:

- Action in state P1 is [G1, LO1], [G2, LO2], [G3, LO3], [G4, LO4], [G5, LO5], [G6, LO6], [G7, LO7], [G8 , LO8]
- Action in state P2 is [G1, LO1], [G2, LO2], [G3, LO3], [G4, LO4], [G5, LO5], [G6, LO6], [G7, LO7], [G8 , LO8]
- Action in state P3 is [G1, LO1], [G2, LO2], [G3, LO3], [G4, LO4], [G5, LO5], [G6, LO6], [G7, LO7], [G8 , LO8]
- Action on the state of P4 is [G1, LO1], [G2, LO2], [G3, LO3], [G4, LO4], [G5, LO5], [G6, LO6], [G7, LO7], [G8 , LO8]

The algorithms in the making of this software are:
1. Giving value for the parameter Y and the reward matrix R. Y is the future rewards that are worth between 100-80, 79-60 and 59-0
The rewards of each action are:
R [P1 (G1, LO1)] = 100, [P1 (G2, LO2)] = 70
R [P2 (G1, LO1) = 70, R [P2 (G2, LO2) = 100
R [P3 (G4, LO4) = 50, R [P3 (G5, LO5) = 70
R [P4 (G6, LO6) = 40, R [P4 (G7, LO7) = 45

I. Here is a table of results of the learning process of the value of student learning styles:

Table Result

Student	Learning Style	Result
P1	G1,LO1	100
P1	G2,LO2	70
P2	G1,LO1	70
P2	G2,LO2	100
P3	G4,LO4	50
P3	G5,LO5	70
P4	G6,LO6	40
P4	G7,LO7	45

Table Result

CONCLUSION

Reinforcement learning has the ability to be an intelligent technique that is able to optimise students in participating in online learning. RL can help students to obtain appropriate teaching materials from the learning process earlier. This study used RL in an effort to create personalised teaching materials according to the student. From the results, nearly 70% of the selected learning style scored above 60.

REFERENCES

[1] R. Felder and L. Silverman, "Learning and teaching styles in engineering education," *Eng. Educ.*, vol. 78, no. June, pp. 674–681, 1988.

[2] P. Honey et.al, "The Learning Styles Helper's Guide," *Peter Honey Publ.*, vol. 1, no. 1, pp. 1–3, 2006.

[3] A. Y. Kolb and D. a Kolb, "The Kolb Learning Style Inventory — Version 3 . 1 2005 Technical Specifi cations," *LSI Tech. Man.*, pp. 1–72, 2005.

[4] W. L. Leite, M. Svinicki, and Y. Shi, "Attempted Validation of the Scores of the VARK: Learning Styles Inventory With Multitrait-Multimethod Confirmatory Factor Analysis Models," *Educ. Psychol. Meas.*, vol. 70, no. 2, pp. 323–339, 2010.

[5] I. Bagus, M. Mahendra, and M. R. Widyanto, "DESAIN SISTEM REKOMENDASI PADA E-LEARNING BERBASIS REINFORCEMENT LEARNING," pp. 2–5, 2007.

[6] A. Gosavi, "Reinforcement learning: A tutorial survey and recent advances," *INFORMS J. Comput.*, vol. 21, no. 2, pp. 178–192, 2009.

[7] S. Graf and Kinshuk, "An approach for detecting learning styles in learning management systems. Proceedings of the," *Proc. Int. Conf. Adv. Learn. Technol. (ICALT 06)*, pp. 161–163, 2006.

Using Q-Learning for Recommend Learning Object on e-Learning System

ORIGINALITY REPORT

14%	2%	14%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

4%

★ "Advances in Semantic Media Adaptation and Personalization", Springer Nature, 2008

Publication

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		