A Comparative Study Between The Application Of The Existing MDP And The Enhanced MDP Model In Modeling Students Comprehension In A Multimedia Learning Environment

Achi Ifeanyi Isaiah, Agwu Chukwuemeka Odi, Alo Uzoma Rita, Ogbu Nwani Henry

Abstract: The analysis of an existing process or model in order to discover the weakness and then developing an improved version of that same process or model has become one of the best practices of software engineering in recent times. The (Markov Decision Process) MDP is an effective mathematical modeling tool used for Sequential Decision Making (SDM) under uncertainty and has been used in many other fields such as manufacturing, industrial, medicine and so on but has been underutilized in the Decision Making Process (DMP) of the educational sector. Therefore, this research paper provides a comparative study on the application of the existing (standard) MDP and the improved MDP models in modeling the student’s comprehension in a multimedia learning environment under uncertainty. The improved Markov decision processes integrate the decision making process in the model to ensure that multiple decisions are made in a shorter time. Hence, the significant advantages over the standard MDP. This paper compares the improved MDPs model to the standard MDP models by solving the problem of modeling students in a multimedia modeling system in order to assign the optimal learning style to the learning student using both models. Both models result, when compared to each other, is the same in providing the optimal policy for obtaining and assigning the optimal learning style to the learning students but different in the time of arriving at a decision. The result obtained from the experiment performed shows that the computation time for solving the improved MDP model is significantly smaller than that of solving with the standard MDP model.

Keywords: Markov decision processes, decision analysis, Markov processes

1. INTRODUCTION
The MDP model has been an integral part of solving critical problems of decision-making process in the field of artificial intelligence especially in the aspect of machine learning [1]. This complex problem requires the use of a more advanced MDP model. Previously, problems were solved using the existing, standard MDP model. This model has some setbacks especially in their inability to model and solve very complex outcomes or events (problems) in a shorter time, especially when outcomes or events occur (or may reoccur) over a range of time. As a result of these setbacks, this paper prefers a solution by proposing the enhanced MDP as a better mathematical modeling tool than the existing standard existing MDP model. This enhanced MDP model could model recurrent problems faced within the learning environment states and future events. Since the introduction of the existing MDP for solving real-life problems by Berkenkamp F, Turchetta M, Schoellig A, and Krause A, [2] the model has become more popular and their usage has grown so much especially in modeling states in a learning environment. Hence, the need for the improvement of this standard existing MDP model by proposing the enhanced MDP that is more effective than the existing MDP in problem solving not only in the learning domain but to the world at large.

However, the existing, standard MDP model cannot be used in a very complex sequential decision-making process[3] whereby problems are represented in steps of the iteration process. This implies that a large number of embedded decision nodes will be generated along the path of the decision making process[4]. Again, computation done by the existing MDP model is impractical most especially when there are large numbers of possible embedded decisions. This is because each iteration in a standard existing MDP model can evaluate only one set of decision rules at a time [4] as against the improved or enhanced MDP model. For example, consider a learning system whereby each learning student is assigned a learning style based on one time modeling process of a sample student as against the enhanced model where each individual student is sequentially or dynamically modeled so as to assign the optimal learning style to each student from the available three learning styles as presented in this paper. The purpose of this paper is to provide a comparative study between the existing MDP model and the enhanced MDP model and to demonstrate the use of both the existing and the enhanced MDP model to solve a decision based problem with sequential decisions that must be made under uncertainty. MDP is a very vital mathematical modeling tool which has been used in various areas such as industries, financial institutions, manufacturing and inventory control system[5] just to mention but a few. It has been used in the field of education and in a learning environment but not to the extent of dynamically modeling students in learning environment which is the focuses of this paper. This paper presents three learning styles to each learning students and we did apply both existing and the enhanced MDP in modeling students’ comprehension state. The three learning styles presented in this paper are as follows audio, visual/textual and hybrid learning style. For each learning style, one out of the three is expected to be the optimal learning style for the learning student. In this paper, both the existing and the enhanced models were subjected to the same modeling process and then the outcome from both

• Achi I. I. is currently pursuing a doctorate degree program in Computer Science, University of Nigeria, Nsukka, Nigeria, PH-2347039147549, Email: bishopchai@yahoo.com
• AGWU C.O., is a PhD holder in Computer Science Department, Ebonyi State University, Abakaliki, Nigeria, PH-2348033954892, Email: emeka2010@yahoo.com
• Abi R. U., is a PhD holder in Computer Science, Alex Ekwueme Federal University, Nduru-Alike, Nigeria, PH-2347034477250, Email: uzomarita@yahoo.com
• Ogbu N. H., is a PhD holder in Computer Science, Ebonyi State University, Abakaliki, Nigeria. PH-2348037954462, Email: greatrio2001@yahoo.com
processes was evaluated. We conclude from our result that the enhanced MDP model is a better model tool than that the existing MDP model.

2. Review of Existing MDP Application in Modeling

This research paper explores MDP applications in modeling generally prior to its application in modeling student comprehension in a multimedia learning environment as described in this paper. Among these, is the work done by Mota et al [5] given a set of available learning materials, this approach enables the ITS to track the students’ difficulties and provide the right material at the right time. They modeled the learning process using the existing MDP, the Partially Observable Markov Decision Process (POMDP), where the hidden information corresponds to the student’s familiarity with each of the topics to be learned. After teaching, quiz and tests are used to monitor the comprehension state of the learning students. Their performance in those test exercises now forms the basis for which the ITS uses to determine the optimal learning material for the student. John W. Coffey [6] in his work came up with intelligent or adaptive tutoring systems which focused on how to determine what resources to present to students as they make their way through a course of study. This decision on how to determine the best resource to present to the student was achieved using the existing MDP model. Zhu X, et al [7] presented in their work a theoretical study of machine teaching in the setting where the teacher must use the same training set to teach multiple learners. This problem is a theoretical abstraction of the real-world classroom setting in which the teacher delivers the same lecture to academically diverse students. They defined a minimax teaching criterion to guarantee the performance of the worst learner in the class. They also proved that the teaching dimension increases with class diversity. They concluded that effort and time required to ramp-up a system is largely dependent on the effectiveness of the human decision-making process to select the most promising sequence of steps that will provide the optimum performance required by the system. Doltsinis et al [8] in their work used the existing MDP based model to choose the best sequence of action that improves the system by producing the desired output irrespective of the time involved. Hasibuan et al [9] in their work used the existing MDP to model a system that enables the user to choose the best learning approach from the two options given-data driven and literature base. It depends on the learner’s prior knowledge to attain accuracy in choosing the learning style for each learning students. Both approaches focus on retrieving data from the interaction of learners from the system and use it as the basis for selecting the optimum learning style. However, according to researchers, the learner interaction approach to this system only has an accuracy value below 80 %. Hamtini in his work [10] describes a dynamic technique that used MDP model for identifying learners best learning styles based on their behavior in the learning environment influenced by the literature approach. Allen, L et al [11] based their research on an existing system called the iSTART which was used to model student reading comprehension ability. In their work, they determined student reading comprehension based on evaluating the student on how the student can explain the content read to another person. This was achieved using the self-explanation performance algorithm and does not adapt content based on reading ability. This decision was arrived at using the MDP modeling tool which helped to determine the best way the student can achieve comprehension. The self-explanatory algorithm on its own is based on the natural language processing technique of artificial intelligence which was used to build models of students’ comprehension ability from the linguistic properties of their self-explanations. Yahya et al [12] in their work, adopted a decision support system based on MDP in e-learning in order to model the visual-auditory-kinesthetic learning style focusing on learning disabilities children. Hence, proposed an e-learning decision support system architecture to estimate students’ learning styles automatically using the literature-based method. The calculation to estimate each of the student’s learning styles is based on number of visits and the time spent on learning objects with respect to the visual-auditory-kinesthetic learning style ok.

3. THE EXISTING MDP MODEL

A Markov decision process (MDP) is a discrete-time stochastic decision process that provides a mathematical model for modeling both simple and complex decision making in situations where outcomes could be partly random (Partially observable MDP) and partly under the control (Constrained MDP) of a decision maker. MDP is a four-tuple (S,A,Pa,Ra) [13] where

- S is a finite set of states
- A is a finite set of actions available from state s
- Pa(s,S') = Pr(S=t | s, a=a) is the probability that action a in state s in time t will lead to s' in time t+1
- Ra(s,S') is the immediate reward (or expected immediate reward) received after transitioning from state s to s' due to action a

The aim of MDPs is to [14]

i. Obtain the value function for a state s which is defined in terms of the well-known Bellman equations and
ii. To find an optimal policy that maximizes the value function in each state of the MDP, typically the expected discounted sum over a potentially infinite horizon:

$$\sum_{t=0}^{\infty} \gamma^t R(s_t,a_t)$$

(Where we choose the action a_t given by the policy)

Where $\gamma$ the discount is a factor and satisfies $0 \leq \gamma \leq 1$ (for example $\gamma = 1/(1+r)$ when the discount rate is r). $\gamma$ is typically close to 1 [14]. The essence of the discount factor is to ensure that the decision maker takes action early and does not postpone them indefinitely. This MDP aim could be achieved via

a. Dynamic programming
b. Value iteration Process and
c. Policy iteration Process

This research adopted the Policy iteration Process to achieve that goal. At every point in time, the MDP is in a state s, and the decision maker is often time confronted with knowing the best action that is most rewarding $Ra(s',S')$. The probability that the process moves into its new state s' is influenced by the chosen action of the decision maker, which is given by the state transition function $Pa(s',S')$. Therefore, the transition to the next state s' is dependent on the current state s' and the decision maker action a[15].

John W. Coffey [6] in his work

Doltsinis et al [8] in their work

Yahya et al [12] in their work

Hamtini in his work

Allen, L et al [11] based their research

WUSTR©2020
www.ijstr.org

2365
4. Solving for MDP

In solving MDP, the solution is a policy that describes the best action for each state in the MDP, known as the optimal policy [16]. The best action is the one that has the highest reward value. This optimal policy can be found through a variety of methods, but this research is focused on using the Policy iteration method.

The algorithm for solving the Policy Iteration Process is summarized in two steps[16]:

i. Value Update which is achieved by solving the equation below

\[ V(s) := \sum_{s'} P_{\pi(s)}(s,s') \left( R_{\pi(s)}(s,s') + \gamma V(s') \right) \]

ii. Policy Update is also obtained using the equation below

\[ \pi(s) := \arg\max_a \left\{ \sum_{s'} P(s'|s,a) \left( R(s'|s,a) + \gamma V(s') \right) \right\} \]

The elaborate version of the algorithm is as follows[16]:

a. Initialization
   Choose an arbitrary policy \( \pi' \) and compute using the equation

b. Policy Evaluation using the equation

\[ V(s) := \sum_{s'} P_{\pi(s)}(s,s') \left( R_{\pi(s)}(s,s') + \gamma V(s') \right) \]

Repeat Until \( v = V' \)

c. Policy Improvement

\[ \pi(s) := \arg\max_a \left\{ \sum_{s'} P(s'|s,a) \left( R(s'|s,a) + \gamma V(s') \right) \right\} \]

Repeat Until \( \pi = \pi' \)
Else Go to b

d. End

5. THE ENHANCED MDP MODEL

The enhanced MDP shares several attributes with the standard, existing MDP like the same framework (component), same goal. The difference is on how each model achieves the set goal. The goal of an MDP is to obtain the value function via value iteration process for a state s which is defined in terms of the well known Bellman equations and to find an optimal policy that maximizes the value function in each state of the MDP via policy iteration process, typically the expected discounted sum over a potentially infinite horizon [17]:

6. COMPARING BOTH MODELS

The difference between the enhanced MDP and the standard, existing MDP is on how both models perform the MDP goal that is in obtaining the value function for each state which is achieved by value iteration, dynamic programming and policy iteration. In this paper, the policy iteration process algorithm was used to calculate the value function

i. Value Function

In order to get the value function in the enhanced MDP modeling, the \( \pi \) in the equation for solving for value function in the existing MDP model is not used in the enhanced MDP model instead the value of \( \pi(s) \) is calculated within \( V(s) \) whenever it is needed. Substituting \( \pi(s) \) in the calculation of \( V(s) \) ensures that there is no continuous repetition of steps until convergence which is really time-consuming. Bellow is the enhanced MDP model

\[ V_{\pi_t}(s) := \max_a \left\{ \sum_{s'} P_a(s,s') \left( R_a(s,s') + \gamma V_{\pi_t}(s') \right) \right\}, \]

Where \( i \) is the iteration number. Value iteration starts at \( i = 0 \) and \( V_0 \) as a guess of the value function. It then iterates, repeatedly computing \( V_{i+1} \) for all states s, until \( V \) converges with the left-hand side equal to the right-hand side which was included as a special case of the value iteration method for MDPs.

ii. Policy Iteration
starts at \( i = 0 \) and \( V_0 \) and compute using the equation

\[
V_{i+1}(s) := \max_a \left\{ \sum_s P_a(s, s')(R_a(s, s') + \gamma V_i(s')) \right\},
\]

ii. Policy Evaluation using the equation

\[
\begin{align*}
&V_i(s) := R_a(s) \quad \text{for all states } s, \quad \text{if } s \in S^f, \\
&V_i(s) := \max_a \left\{ \sum_s P_a(s, s')(R_a(s, s') + \gamma V_{i-1}(s')) \right\}, \\
&\text{for all } s \in S - S^f.
\end{align*}
\]

iii. Policy Execution

iv. End

7. Evaluating the modeling Processes

To further validate our findings on both modeling processes, an experiment was performed which exposed both modeling process to two lecturers, from two universities in Nigeria. One from Ebonyi State University (EBSU) and the other from Alex Ekwueme Federal University Ndufu Alike, Ikwo(AE-FUNAI), that is each lecturer used both the enhanced and the standard existing MDP model in modeling learning students from both institutions of higher learning. This paper allowed each lecturer to select five students each at random from both institutions, irrespective of their age, level, and department. Both models presented a multimedia system with three learning styles. The idea is to find out the optimum learning style, that is the learning style that best suit each learning student. The learning styles presented by these models are audio, video (textual) and hybrid (audio and video) learning style. The focus of this paper is to compare both models and find out the advantage(s) of the enhanced MDP model over the standard existing MDP model. As well stated in the previous section, the advantage is time-based. The enhanced MDP model is faster in solving complex problems than the existing MDP model. The result obtained in the experiment has also validated that fact. The experiment is based on modeling the various students by teaching them an introductory course in Biology using the proposed three learning styles - audio, visual (textual) and hybrid. The result of the data collected from the experiment is shown below:

i. Lecturer (From EBSU) has the following data

<table>
<thead>
<tr>
<th>Student</th>
<th>LS1</th>
<th>LS2</th>
<th>LS3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student A</td>
<td>30</td>
<td>29</td>
<td>21</td>
<td>80</td>
</tr>
<tr>
<td>Student B</td>
<td>23</td>
<td>22</td>
<td>25</td>
<td>70</td>
</tr>
<tr>
<td>Student C</td>
<td>19</td>
<td>18</td>
<td>19</td>
<td>56</td>
</tr>
<tr>
<td>Student D</td>
<td>44</td>
<td>33</td>
<td>22</td>
<td>89</td>
</tr>
<tr>
<td>Student E</td>
<td>46</td>
<td>37</td>
<td>22</td>
<td>105</td>
</tr>
</tbody>
</table>

Below is the graph that compares the total minutes spent on the modeling process against each student for both model for EBSU:

Table 2: Showing the data obtained from AE-FUNAI

<table>
<thead>
<tr>
<th>Students</th>
<th>Modeling Time(minutes) (Existing Model)</th>
<th>Modeling Time(minutes) (Enhanced Model)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LS1</td>
<td>LS2</td>
</tr>
<tr>
<td>Student A</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>Student B</td>
<td>08</td>
<td>09</td>
</tr>
<tr>
<td>Student C</td>
<td>07</td>
<td>06</td>
</tr>
<tr>
<td>Student D</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>Student E</td>
<td>12</td>
<td>11</td>
</tr>
</tbody>
</table>

Below is the graph that compares the total minutes spent on the modeling process against each student for both model for AE-FUNAI:


