Abstract. The essence of learning is for the learner to attain a significant level of comprehension after the learning process is completed. The quest to achieve this singular purpose has led to the introduction of several learning techniques in the conventional learning environment, such as asking questions and conducting test after class, just to mention a few. Additionally, technology has been introduced in learning. Even with technological advancements, the learning experience still faces the challenge of learners not attaining the optimum comprehension state after the learning process. This is due to the present systems' inability to model the learner to determine the best methods for achieving maximum comprehension. Hence, this research paper focuses on deriving an improved mathematical model for predicting the learning path to a learner’s optimum comprehension. The
paper presented three learning instructional media (learning paths); textual, audio and a hybrid of audio and video, which this research uses in modelling the learner. This is to enable the improved system predict the best learning path to optimum comprehension for learners. This research paper adopted Reinforcement Learning and the Markov decision process, specifically the Markov Chain approach, in developing an improved model for prediction. The evaluation of this research involved brainstorming on the Bellman’s equation with the aid of the Markov Chain transition state framework, resulting in an improved mean value function of 71.7. This indicates an enhanced comprehension state for the learning students compared to the existing mean value function of 46.0. The results obtained from this research clearly demonstrate that the improved model was able to predict and assign the best learning path to achieve optimum comprehension state for learners.

**Keywords:** Machine Learning; Markov Chain; Markov Transition State Diagram; Intelligent Tutoring System; Computer-based Learning; Markov Decision Process.

1. **INTRODUCTION**

Learning is an integral part of human existence, and consequently, one must determine the most effective medium among the various forms of learning to achieve optimal comprehension after the learning process. This is where decision-making becomes crucial in this research. Historically, little attention has been given to identifying the best medium because options were not provided initially, depriving students of the opportunity to select the learning style that best suits their needs for maximum comprehension. Typically, learners assemble in classrooms or lecture halls where the teacher instructs using a standardized medium. There was no means of assessing individual students to determine the most suitable learning medium for each student’s needs.

In an attempt to address this issue, students are allowed to ask questions where they need clarification, or sometimes the teacher poses questions to assess understanding, or administers tests after teaching a subject matter. The teacher then uses the responses as a basis for evaluating the students’ comprehension state. Recently, teaching has been supplemented by computers, yet these systems lack the capability to assess students’ learning preferences before instruction begins. To tackle this student modeling challenge within the computer-based learning (CBL) environment, Hasibuan et al.[1] developed a learning system that could anticipate learners’ needs prior to instruction. While this system customizes lessons for users, its approach differs from that of the research discussed here. In this research, a survey is conducted for each user to understand their learning behavior or methods before they engage with the computer-based learning system. This implies that the modelling process is achieved outside the system, manually, unlike the improved model where all aspects of the modelling process are integrated within the proposed system. The authors here in the work still use the existing Markov Decision Process (MDP) in the decision-making process based on the data acquired from the user outside the system. This differs from the improved model, which handles all aspects within the system itself.

This research paper focused on improving the teaching and learning process in both traditional and computer-based learning environments by introducing three teaching medium designed to address the limitations of existing models. This model begins by studying students through student modeling to determine the most suitable learning style or medium of instruction for each individual learner. In this research paper, the learning styles incorporated into this improved model include visual (textual), audio, and hybrid (a combination of video and audio), recognizing that different students may comprehend better through different mediums. The challenge lies in developing a model capable of addressing this diversity. To accomplish this, this model employs Reinforcement learning, Markov Decision Process and the Markov Chain transition framework to predict the best learning style for each learning student and assign it to the learner before the learning commences. By implementing this approach, students can learn
using the optimal learning path that maximizes comprehension of the subject matter. The implications of this innovation are significant, benefiting students, teachers, and society at large.

**The Problem Statement**

The problem addressed in this research paper, as identified through a review of past literature, is the inadequacy of existing systems to effectively model learners by providing diverse learning paths to knowledge. These systems struggle to predict the most effective path for achieving optimum comprehension and assigning it to the learner for the learning process. Recent studies have also highlighted that a major cause of poor comprehension during and after learning is the reliance on the same, potentially incorrect, medium of instruction. Despite efforts to vary instructional materials within a medium, such as using different contents, this issue remains unresolved. Hence, the focus of this research is to address this challenge and develop a system capable of providing diverse learning paths tailored to individual learners to enhance comprehension.

**Significance of the research**

The following stakeholders have been identified by the researchers and the impacts of this research to them have been identified as follows:

i. **Learners (students):** this research will help to achieve the following: reduction of learning curve (process), improved understanding of the subject matter, reduction of learning time and improved output.

ii. **Teachers:** this research is expected to achieve: reduction of teaching time, convenience of teaching and improved performance and quality

iii. **Management/Administration:** Management is expected to benefit from this research by the reduction of daily running cost and low cost of infrastructural investment and maintenance since students study at the comfort of their home.

iv. **Government:** this research will help government to: improve the educational sector and achieve educational goal, solve the problem of unemployment and hence achieve improved governance.

v. **Parents** on the other hand will also benefit from the model’s cost effectiveness.

**Analysis of recent studies and publications**

This research paper explored the use of Reinforcement Learning, Markov Decision Process (MDP) and the Markov Chain Transition framework in modeling student comprehension level in a learning environment by predicting the best path to the optimum comprehension as described in this research paper. Recent related research works adopted the standard MDP approach in attempt to resolve poor comprehension state for learners. These systems discovered that a learner might have difficulty learning with just a single path to knowledge acquisition. Though these systems already identified their weaknesses, however they had not been able to proffer an optimal solution to the existing challenges as a result of not having multiple paths to knowledge acquisition. Amongst these recent studies are the research done by Mota *et al* [2] which gave a set of available learning equipment to learners that was able to track the students’ comprehension state with different They modelled the learning process using the standard MDP and the Partially Observable Markov Decision Process (POMDP), where the hidden information corresponds to the student’s familiarity with each of the topics to be learnt. The student’s progress was monitored by their performance in different test exercises and the right equipment was recommended at the right time depending on this performance.

Intelligent or adaptive tutoring systems was developed by John [3] uses standard MDP model to make a decision on which learning resource to offer to learner in order to attain optimal comprehension state. The research did not yield the desired result because the various instructional materials are within a medium.
Zhu et al. [4] conducted a theoretical investigation on machine learning in a situation where teachers are required to instruct several students using the same training set. This scenario is a theoretical abstraction of a real conventional classroom, where a teacher employs a singular learning path to teach learners of various academic backgrounds and interests. The problem still persists on how a teacher will be able to know how best to communicate to the students of various academic background/interest to attain maximum comprehension state. Consequent upon the persisting problem, a minimal learning curriculum was developed, but the success rate was insignificant because the high proficient learners could not benefit much from the system. Hence, it can be deduced that the use of a set of instructional material and minimal learning curriculum to teach learners from different academic background/interest did not provide the desired result of optimal comprehension state.

Doltsinis et al. [5] employed the standard MDP-based model to determine the optimal course of action that enhances comprehension of a learner. According to the research it proposes that the use of the following teaching activities: repeating every sentence, asking questions after each lecture module and conducting test could help to improve comprehension state of learners. However the system had its own odds because the use of those conventional teaching activities led to much time consumption, long learning curve and as such did not achieve the desired result.

Hasibuan et al. [6] modelled a system that allows the user to select between two learning approaches: data-driven and literature-based using the existing standard MDP. Data-driven and literature-based approaches work based on the interactions of the system with the learner which can easily be circumvented by the insincerity of human user. These two approaches may not be efficient because the system will always rely on the responses of the learner to make choice which often times may not be accurate. This is a major setback to the research.

Hamtini [7], outlined a dynamic technique that leverages on the standard MDP model to identify the best way to teach a learner based on learner’s behavior in a learning environment. However, using the learner’s behavior to determine the instructional techniques to be adopted is not the best approach because behavior is intuitive and should not be used to define a learner’s state of being.

Allen et al [8], created a learners model based on the learners reading comprehension ability within an intelligent tutoring system called iSTART. This system delivers content to the learners based on the learner’s ability to read and understand. It attained some measure of success in the sense that lectures are prepared knowing the reading and comprehension state of the learner. With the system, the learners could comprehend to some extent, but there is still more room for improvement because this system is really time consuming; knowing the content the learner will attain high comprehension. Yahya et al [9] used the standard MDP-based decision support system in e-learning to model the visual-auditory-kinesthetic learning style based on the time spent on these learning paths by the learners. This system proposed to teach a particular learner based on the estimate of the time spent on the learning style. This system may teach with audio or video or kinesthetic means which is teaching by practicalizing. This system was actually proposed for children as such cannot work effectively for adults.

Research Gap

This research became necessary due to persistent problems that existing systems have been unable to resolve, particularly the challenge of learners experiencing poor comprehension despite efforts to address it. Existing systems have struggled to effectively model learners using various instructional mediums, predict the most suitable medium for individual learners (i.e., the learning path conducive to optimal comprehension), and subsequently assign it for learning purposes. As a result, this research aims to address these limitations and develop a system capable of effectively addressing the diverse needs of learners to enhance comprehension outcomes.
The research goal
This research paper aims to develop an improved system for predicting the learning path to learners’ optimal comprehension. The specific objective is to derive the improved model used in the development and evaluation of this system.

2. THEORETICAL BACKGROUND

This section helped to x-ray the conceptual terms as well as the technologies employed in this research.

- **Markov Chain**
  Henry, et al.,[10] defines a Markov Chain as a mathematical system that undergoes transitions between states in accordance with specific probabilistic rules. The key feature of a Markov Chain is that its potential future states are fixed, regardless of how the process got to its current state. A Markov Chain can be compared to a basic random walk [11]. A straightforward idea as demonstrated in [12] known as the Markov Chain explains the majority of complex real-time processes. This straightforward idea known as the Markov Chain is used in many artificial intelligence applications, such as speech recognition, text identifiers, path recognition, and many more.

  According to Devim [13], Markov Chain is a popular and reasonably easy method for statistically modelling random processes. They have been applied to a wide range of fields, such as: financial modelling and text generation. Subreddit Simulator, a well-known demonstration in [14], automates the creation of content for a whole subreddit using Markov Chain. Markov Chain is highly accessible and conceptually straightforward because it does not require the application of complex statistical or mathematical ideas. It is an excellent starting point for studying data science and probabilistic modelling methods.

  Markov Chain as mentioned earlier [15] has numerous applications in statistical real-world modelling processes, including the study of cruise control systems in automobiles, passenger lines at airports, currency exchange rates, storage systems like dams, and population increases of specific animal species, amongst other things.

- **Markov Chain Model**
  According to Martin and Arroyo [17], a Markov model is a stochastic technique for randomly altering systems in which it is presumed that past states will not affect future states. Markov models are frequently used to simulate the likelihood of various states as well as the rates at which they transit. Typically, the technique is applied to model systems.

- **Application of Markov Chain Model**
  Many academic disciplines, including biology, economics, the sciences, engineering and others, use Markov Chain as mentioned earlier [17],18. Markov Chain can be used to model the randomness in asset value predictions. The value of the various options could be used by the Markov Chain to model the randomness when making management decisions. This is also relevant to the research project, since the system must select the learner’s preferred learning style from a list of three, in order to achieve the best possible comprehension state. Here, the randomness is modelled by the Markov Chain model using the learning style value. The subsequent sections of this research paper provides a detailed explanation of this model.

Relevance of Markov Chain model to this research
Markov Chain is relevant to this research paper in the following ways:

- Local; overlapping sets of data that serve as clues to our comprehension of regions can be effectively modelled using Markov models.
- Markov Chain offers a stochastic model for determining a learner's unique learning style, which serves as the basis for the learning environment under study. This assignment on stochastic learning styles, which was completed using the Markov Chain
model, may also serve as a helpful model for learning and knowledge dissemination in post-secondary institutions and the general public.

3. RESEARCH METHODOLOGY

Various methods were used to achieve the objectives of this research. This research adopted the quantitative research methodology for data gathering while Object-Oriented Analysis and Design Methodology (OOADM) was used in software design. By brainstorming on the Bellman’s Mathematical Model using Reinforcement Learning and Markov Chain Transition Framework; the improved model for Predicting the Learners’ Path to Optimum Comprehension State was derived and was used to evaluate the system and the result discussed. The system was evaluated by solving the non-linear equation derived for the improved model and the standard classical model for the existing. The value function which connotes comprehension state obtained from both models were compared which was used to validate success in this research

4. DEMONSTRATING MARKOV CHAIN

This research paper presents the Markov Chain to demonstrate the following MDP components:
- States
- Transition
- Rewards
- No action

This research paper presents the above component using the Markov Chain to obtain the value function for each state, the reward, and the transition.

The Markov Chain Modeling

This shows the modelling process with the Markov Chain align with this research paper. This research model has three states; audio, textual/visual, and hybrid learners which is represented in this research paper as states 1, 2 and 3?

- Value of a state, using infinite discount factor is

\[ V(s) = R(s) + \gamma \sum_{s'} P(S'|S) V(S') \] as mentioned earlier [17]

Markov Chain Rules

- The present state must be an improvement to the previous state. To ensure that this happens, the discount factor (\(\gamma\)) is introduced as seen in the above equation for \(V(s)\). It must be between 0 and 1
  \(0 < \gamma < 1\) I discount factor). In this research paper, \(\gamma=0.9\) was used
- Sum of forward probabilities actions must be = 1
- Markov Chain does not always have reward values in all the states at the same time as shown in the example of this research paper, 0 is assigned to state 1 and 3 and 10 is assigned to state 2. Zero(0) assigned to state 1 and 3 symbolises no reward and then 10 assigned to state 2 symbolises that state 2 has a reward, justifying the rule number 3 as stated above. Furthermore, in a case where there are two states, one will be rewarded, and the other will not be rewarded. Where there are three states, one will be rewarded, and the other two will not be rewarded as in the case of this research paper. Again, in a case where there are four cases, three will not be rewarded and one will be rewarded and so on.
**Solving Markov chain**

Assigning the following values using the above rules in an attempt to solve a three-state problem such as the one presented in this research paper, modelling a student within a learning environment with three teaching styles to choose from the student’s needs to be modelled to find out the best teaching style he or she prefers. From the modelling process, the state or teaching style that the student has the highest value (value function) becomes the basis for assigning the best teaching style to the learning student. Below is how this can be achieved using the Markov Chain:

- Assumption based on Markov chain rules number 1: ϒ = 0.9
- Assumption based on Markov chain rule number 3: state 1=0; state 2= 10 and state 3 =0
- Assumptions based on Markov chain rule number 2; for the probable forward action was the basis which those values between stages were chosen on the Markov chain diagram below:

![Example Markov Chain transition diagram](image)

From the value state formula as mentioned earlier [17], which is stated below:

\[ V(s) = R(s) + \gamma \sum_{S'} P(S'|S) V(S') \]

Substituting all our assumptions from the above formula, we obtained the following equations

\[ V(1) = 0 + 0.9(0.5V(1) + 0.5V(2)) \]
\[ V(2) = 10 + 0.9(0.2V(1) + 0.1V(2) + 0.7V(3)) \]
\[ V(3) = 0 + 0.9(0.2V(1) + 0.7V(3)) \]

If we solve for the value of each state from the above equation, we obtain

\[ V(1) = 40.5, \quad V(2) = 49.5, \quad \text{and} \quad V(3) = 49.1 \]

- The next state and final is choosing the best value (optimal value).
- This is possible with the followings:
  - Dynamic programming
  - Value iteration and
  - Policy iteration

This research paper uses policy iteration to select the optimal value which according to this research, is the highest score after the test during the modelling. In this example, the optimal value is 49.5, which is state 2.
Illustrating the Improved MDP Markov chain modelling

The illustration of the modelling process with the improved MDP (Markov Chain) aligns with this research work. This research model has three states; audio, textual/visual, and hybrid learners. This is demonstrated in this research using the Markov Chain as states 1, 2, and 3.

- Value of a state, using an infinite discount factor rate is

\[ V(s) = R(s) + \frac{1}{1+r} \sum_{s'} P(S'|S) V(S') \]

Improved Markov Chain Rules

- The present state must be an improvement to the previous state. To ensure this happens, this research paper introduced an expression with the discount factor rate \(r\). It must be between -0.5 and 0 \((0 < r < -0.5\) discount factor rate). This research paper used; \(r = -0.5\)
- Sum of forward probabilities actions must be = 1
- Markov Chain does not always have reward values as demonstrated in this research paper by assigning 0 to state 1 and 3 and 10 to state 2.

Solving the improved Markov chain

Here, the expected value should be higher than the existing Markov Chain value, which connotes better comprehension. Assigning the following value with regard to the above Markov Chain rules in an attempt to solve the three-state problems as proposed in this research paper; modelling a student within a learning environment which has three teaching styles to choose from. The student needs to be modelled to find the best teaching style he/she prefers. From the modelling process, the state or teaching style that the student has the highest value becomes the basis for assigning the best teaching style to the learning student. Below is how this can be achieved using the Markov Chain:

- Assumptions from rule 1 and 3 respectively; \(r = -0.5\); state 1=0; state 2= 10 and state 3 =0

Assumptions for the probable forward action rule in rule 2 as stated above are the basis for choosing the values seen on the Markov chain diagram below:

![Fig 2. Example Improved Markov Chain transition diagram](image)

From the value state function formula stated below:

\[ V(s) = R(s) + \frac{1}{1+r} \sum_{s'} P(S'|S) V(S') \]

Substituting all assumptions from the above formula, the following equations were derived:

\[ V(1) = 0 + 2(0.5V(1) + 0.5V(2)) \]

\[ V(1) = 0 + 2(0.5V(1) + 0.5V(2)) \]

\[ V(1) = 0 + 2(0.5V(1) + 0.5V(2)) \]
V(2) = 10 + 2(0.2V(1) + 0.1V(2) + 0.7V(3))..................(ii)
V(3) = 0 + 2( + 0.9V(2) + 0.1V(3))..................(iii)

If we solve for the value of each state from the above equation, we obtain
V(1) = 61.1, V(2) = 79.5, and V(3) = 74.5

- The next state and final is choosing the best value (optimal value).
- This possible with the followings:
  - Dynamic programming
  - Value iteration and
  - Policy iteration

We used the policy iteration in this research to select the optimal value which, in this research, was the highest score after the test during the modelling. In this example, the optimal value is 79.5 which is state 2.

6. RESULTS AND DISCUSSION

The results obtained from the mathematical non linear equation for both the existing and the proposed system are shown in table 1.

<table>
<thead>
<tr>
<th>States</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Existing system</td>
<td>40.5</td>
<td>49.5</td>
<td>49.1</td>
<td>46.4</td>
</tr>
<tr>
<td>Improved System</td>
<td>61.1</td>
<td>79.5</td>
<td>74.5</td>
<td>71.7</td>
</tr>
</tbody>
</table>

From the data shown in Table 6.1, the mean value obtained, that is the value function obtained from the non linear equation for the existing system is much less than the mean values function obtained from the improved system. This helped to validate the fact that improved system has a better comprehension state than the existing system. The contributions by the co-authors are specified as follows:

**Contribution by Authors**

First Author is the Lead Author; Second Author is the Data Analyst; Third Author is the Software Designer; Fourth Author contributed in the area of Data Gathering; Fifth Author is the Software Developer; Sixth Author is the System Analyst and content proofreading; Seventh Author created the Database and the Eighth Author did the system architectural design.

7. CONCLUSIONS AND PROSPECT FOR FURTHER RESEARCH

This research paper demonstrates that the derived improved model which was used to develop and evaluate the improved system for predicting the learner’s path to optimum comprehension, provided the best solution to the problem of poor comprehension state that persisted with learners over the years. The result obtained from the performance evaluation of the improved model when compared to the existing models shows that the current system is better as it has helped to improved comprehension state of learners than the existing systems. Therefore, this research has proven that providing various learning path to knowledge and predicting the path to optimum comprehension state have brought about tremendous improvements in comprehension states for learners unlike the previous systems which could neither provide multiple paths to knowledge nor predict the learners path to optimum comprehension. However, this current system still has its own limitation as it only focused on
learners in tertiary institution. This implies that other categories of learners were not considered in this research, especially learners in primary and post primary level of education. Therefore, for further research, this paper recommends that more learning paths to knowledge that will be beneficial to other categories of learners be incorporated into the current system. This will expand the usage of the system; incorporating other categories of learners and thereby helping to improve the comprehension state of other categories of learners as well in the society. This will guarantee effectiveness and much better comprehension state among all categories of learners.

Acknowledgements
We wish to express our profound gratitude to those who contributed to this research work in one way or the other and to see that this research work was completed. An exceptional gratitude goes to Prof F.J. Ogwu and Prof F. S Bakpo of Computer Science Department, University of Ibadan and University of Nigeria, both in Nigeria respectively; they for their assistance and advice that made this research a success. We also extended our profound gratitude to Dr C.N. Udanor a Reader in the department of Computer Science, University of Nigeria, for his contribution to the completion of this study.

Finally, we extend our best regards to our friends and immediate family members for their immense support during this research. Also, we extend our best wishes to everyone whose names could not be mentioned here for their support and contribution to the success of this study.

REFERENCE (TRANSLATED AND TRANSLITERATED)


Text of the article was accepted by Editorial Team 04.12.2023
Godson Kenechukwu Ezeh  
magistr z informačných technológii, vikладač a doslédňník  
kafeďa kom'joterinných nauk  
Inštitút šiľskéhošpoddárskovei nauki i technológii, Enugu, Nígeria  
ezehgodsonkenechukw@gmail.com

John Otosí Uga  
doktor z kom'joterinných nauk, docent i doslédňník  
kafeďa kom'joterinných nauk/inkomfmatiki  
Federalnûj universiteit imenî Alekša Euëmu Nduifu Alíke Ikvo, štat Eboni, Nígeria  
ugahjohn@gmail.com

Anotáció. Osnovnoj metoj nавчання є досягнення учнем розуміння матеріалу після завершення навчального процесу. Прагнення досягти цієї мети призвело до впровадження у звичайному навчальному середовищі нових методів дослідження, зокрема опитування та проведення тестів після занять. Впровадження технологій у навчальний процес змінює його. Але, незважаючи на технологічний прогрес, залишається проблема відсутності у деях учнів прогресу в досягненні оптимального рівня засвоєння матеріалу після завершення навчання. На погляд авторів, це пов'язано з нездатністю сучасних систем передбачити найкращі методи для досягнення максимального розуміння матеріалу кожним учнем. Представлена наукова робота зосереджена на отриманні покращеної математичної моделі для прогнозування шляху навчання до оптимального розуміння матеріалу учнем. У статті представлено три навчальні засоби навчання (шляхи навчання): текстовий, аудіо та гібрид аудіо та відео, які використовуються в цьому дослідженні для моделювання засвоєння знань учнем. Це дає змогу вдосконаленій системі прогнозувати найкращий шлях навчання для оптимального засвоєння матеріалу учнями. У цій дослідницькій роботі було застосовано навчання з підкріпленням і процес прийняття рішень Маркова, зокрема підхід ланцюга Маркова для розробки вдосконаленої моделі прогнозування. Моцний штурм рівняння Беллмана за допомогою структурної схеми переходних станів ланцюга Маркова, в результаті чого функція середнього значення була покращена до 71,7, був складовою оцінки цього дослідження. Це вказує на покращений стан розуміння матеріалу студентами у порівнянні з існуючою функцією середнього значення 46,0. Результати, отримані в цьому дослідженні, демонструють, що вдосконаленій модель змогла передбачити і признати найкращий шлях навчання для досягнення оптимального стану розуміння учнями матеріалу, що вивчається.

Ключове слова: машинне навчання; ланцюг Маркова; діаграма переходних станів Маркова; інтелектуальна система навчання; комп'ютерне навчання; процес прийняття рішень Маркова.