A Personalised E-Learning System Using Reinforcement Learning technique.

Shobana B.T,
Department of Information Technology,
Sri Venkateswara College Of Engineering, Chennai.
shobana@svce.ac.in

Dr.G.A.Sathish Kumar
Department of Electronics and communication Engineering,
Sri Venkateswara College Of Engineering, Chennai.
sathish@svce.ac.in

ABSTRACT

Advancement in Internet and multimedia technologies has positively influenced the efficient use of e-learning systems for education and training. Removing the time and space limitations, various e-learning systems have been developed for diverse audience. Learners have different goals, knowledge levels and learning capabilities, thus development of adaptive and personalised e-learning systems that address all these issues is quite difficult. To incorporate adaptivity into e-learning systems we have proposed a machine learning algorithm called “Reinforcement Learning”. Personalisation is achieved by our system through interaction with the learner, retrieval of current and past information about the learning activities and establishment of a customised and personalized learning path according to the learner’s need. Feedback of the learner on the learning content is obtained and recorded by the system.

KEYWORDS: Personalization, E learning, Reinforcement Learning.

INTRODUCTION

The life of knowledge and human skills today is shorter than ever, mounting the pressure to remain up to date with education and training throughout a career. Lifelong learning is quickly becoming an imperative in today’s world. E-learning is the catalyst that is changing the whole model of learning in this century. It is the computer and network-enabled transfer of skills and knowledge enabling people to learn anytime and anywhere. There are now no longer geographical constraints to learning; E-learning brings learning to people, not people to learning.

People have unique ways of learning (i.e., some people prefer visual content and some may prefer textual). Incorporating adaptivity into E-learning systems in order to make them more effective and personalized according to the individual characteristics of students is a great challenge. To address this challenge Artificial Intelligence was introduced in the late 1980’s into e-learning systems to make them intelligent and...
versatile. Various innovations have been made and a lot more are yet to be made to achieve the same.

The adoption of e-learning in all spheres—corporate, schools, universities, etc—is increasing day by day. This increasing use of e-learning systems provides the motivation to incorporate intelligence in them. An adaptive e-learning system is mainly a system that aims at adapting some of its key functional characteristics (for example, content presentation and/or Navigation support) to the learner needs and preferences. It operates differently for different learners, taking into account information accumulated in the individual or group learner models. Thus any enhancement in this adaptive e-learning domain is always welcoming.

The objective of our research is to incorporate intelligence into e-learning systems through the introduction of Reinforcement Learning technique. The proposed system aims to achieve a personalised and adaptive tutoring system by interacting with the learner, obtaining current and past information about the learning activities and establishing a customised and personalized learning path according to the learner’s need.

• RELATED WORK

We have surveyed various e-learning systems, their contributions and drawbacks. An intelligent tutoring module controlled by BDI agents for an e-learning platform [4] uses the BDI agents (Belief, Desires, Intentions) for adaptive course delivery. But BDI agent is a static approach and does not involve machine learning. Learning styles’ recognition in e-learning environments with feed-forward neural networks [10] uses artificial neural network to analyse the learner’s learning style and deliver course adaptively. It uses supervised learning to train the system. But this type of machine learning is tedious and not effective in a long run.

A personalized e-learning system based on item response theory and artificial neural network approach [1] uses the item response theory to calculate the learner’s capability and artificial neural network with back propagation to suggest relevant learning materials based on the capability of users.

Reinforcement learning is an intelligent technique for learning where agent interacts with environment mapping states to appropriate actions that yield maximum reward. Reinforcement learning is used for personalised search and filtering in various domains. Reinforcement Learning for Personalizing Image Search [7] introduces a personalised reinforcement learning tool to present images that are desirable to learner. Personalized Web-Document Filtering Using Reinforcement Learning [2] performs online information filtering using reinforcement learning. It learns user profiles, retrieves relevant learning objects, calculates the expected value of the learning objects and uses a policy to present filtered learning objects of user’s interest. The same could be adopted in e-learning applications, evidence of which is provided in [5].

• REINFORCEMENT LEARNING
Reinforcement learning is learning what to do-how to map situations to actions-so as to maximize a numerical reward signal. The learner is not told which actions to take, as in most forms of machine learning, but instead must discover which actions yield the most reward by trying them. In the most interesting and challenging cases, actions may affect not only the immediate reward but also the next situation and through that, all subsequent rewards. These two characteristics - trial-and-error search and delayed reward are the most important distinguishing features of reinforcement learning.

Reinforcement learning is different from supervised learning. Supervised learning is learning from examples provided by a knowledgeable external training data. This is an important kind of learning, but alone it is not adequate for learning from interaction. In interactive problems it is often impractical to obtain examples of desired behaviour that are both correct and representative of all the situations in which the agent has to act. In uncharted territory, where one would expect learning to be most beneficial, an agent must be able to learn from its own experience.

One of the challenges that arise in reinforcement learning and not in other kinds of learning is the trade-off between exploration and exploitation. To obtain a lot of reward, a reinforcement learning agent must prefer actions that it has tried in the past and found to be effective in producing reward. But to discover such actions, it has to try actions that it has not selected before. The agent has to exploit what it already knows in order to obtain reward, but it also has to explore in order to make better action selections in the future. The dilemma is that neither exploration nor exploitation can be pursued exclusively without failing at the task. The agent must try a variety of actions and progressively favour those that appear to be best. On a stochastic task, each action must be tried many times to gain a reliable estimate its expected reward. Another key feature of reinforcement learning is that it explicitly considers the whole problem of a goal-directed agent interacting with an uncertain environment.

This is in contrast with many approaches that consider sub problems without addressing how they might fit into a larger picture. Reinforcement learning takes the opposite tack, starting with a complete, interactive, goal-seeking agent. All reinforcement learning agents have explicit goals, can sense aspects of their environments, and can choose actions to influence their environments.

- **SYSTEM ARCHITECTURE**

  ![SYSTEM ARCHITECTURE](wopE758.tmpMicrosoft_PowerPoint_Slide1.sldc)

**Proposed Algorithm**

1. a) Construct an initial learner profile
   
   \[ P = \{ wpvi , wpve , wpse , wpin , wpse , wpg \} \]
   
   b) Get the search query from the learner. Set \( t = 0 \).

2. (Retrieval) Retrieve N learning objects (LOs) matching the search query.

3. (Filtering) a) Estimate relevance value of learning objects.
b) Rank the $N$ learning objects and present $M$ of them to the user where $M \subseteq N$.

4. (Interface) Get the feedback by observing learner behaviour.
5. (Learning) Update the user profile.

- LEARNING STYLES

FELDER-SILVERMAN LEARNING STYLE MODEL (FSLSM)

A learning style is defined, as “the unique collection of individual skills and preferences that affect how a student perceives, gathers, and process learning materials.”

There are several different learning style models including Kolb (1984), Honey and Mumford (1982) and Felder and Silverman (1988). Each proposes different descriptions and classifications of learning styles. In our work we use Felder and Silverman learning model that describes the learning style of a learner in detail, distinguishing preferences on four dimensions. FSLSM is based on tendencies, indicating that learners with a high preference for certain behaviour can also act sometimes differently. There are four dimensions in FSLSM. Each learner is characterized by a specific preference for each of these dimensions.

**Sensing - Intuitive learners**

Sensing learners tend to like learning facts, solving problems by well-established methods and dislike complications and surprises whereas intuitive learners often prefer discovering possibilities and relationships, innovation and dislike repetition. Sensors tend to be patient with details and good at memorizing facts and doing hands-on (laboratory) work; Intuitors may be better at grasping new concepts and are often more comfortable than sensors with abstractions and mathematical formulations. Sensors tend to be more practical and careful than intuitors and dislike courses that have no apparent connection to the real world; intuitors tend to work faster and to be more innovative than sensors and dislike courses that involve a lot of memorization and routine calculations.

**Visual - Verbal Learners**

Visual learners remember best what they see—pictures, diagrams, flow charts, time lines, films, and demonstrations. Verbal learners get more out of words—written and spoken explanations.

**Sequential – Global Learners**

Sequential learners tend to gain understanding in linear steps, with each step following logically from the previous one.
Global learners may be able to solve complex problems quickly or put things together in novel ways once they have grasped the big picture, but they may have difficulty explaining how they did it.

**• LEARNER PROFILE MANAGEMENT**

Student models should be used for tailoring the teaching strategy and dynamically adapting it according to the student’s abilities and previous knowledge. Student Models are often based on various different dimensions. The focus is on learning style dimension.

The learner style is initially obtained by using a modification of “The Index of Learning Styles (ILS)”, developed by Felder and Soloman. We construct an initial learner profile using the FSLSM model that describes the learning style of a learner distinguishing between preferences on four dimensions. Each learner is characterized by a specific preference for each of these dimensions. Learner’s profile is represented as a vector of these dimensions.

\[ P = \{ wpvi , wpve , wpse , wpin , wpseq , wpog \} \]

• RETREIVAL AGENT

Efficient and effective document retrieval techniques are critical in managing the increasing amount of information available in electronic form. Retrieval techniques mainly rely on indexing keywords. Identifying the index terms and selecting an appropriate document is a challenging task. The paper “Document Ranking and the Vector-Space Model” [3] uses several simplifications of the vector-space model and seeks the optimal balance between processing efficiency and retrieval effectiveness.

The task of the retrieval agent is to get a collection of N relevant learning objects to be filtered. We use standard term-indexing techniques [Frakes and Baeza-Yates, 1992].

Formally, a learning object is represented as a term vector \( x_i \)

\[ x_i = (x_{i1}, x_{i2}, \ldots, wx_{ik}, \ldots, wx_{id}) \]

where \( x_{ik} \) is the numeric value that term \( k \) takes on for learning object \( i \), \( d \) is the number of terms used for learning object representation. In our work, we assume that \( x_{ik} \) represents the normalized term frequency, i.e. \( wx_{i,k} \) is proportional to the number of term \( k \) appearing in learning object \( i \).
INFORMATION RETREIVAL

Information Retrieval (IR) is done with the objective of providing, in response to a user query, references to learning objects that would contain the information desired by the user. The system is intended to identify which learning objects the user should read in order to satisfy his/her information requirements. In order to identify which learning objects the user should read with respect to his information requirements some representation of what the learning object is about is needed. Representation in the system as to the contents of the learning objects cannot be expected to be satisfactory. Hence the IR researchers take the view that the system should adopt fairly simple methods of representation and seek approaches that facilitate ranking of learning objects in the order of their estimated usefulness to the query.

VECTOR SPACE MODEL

In the vector space model, a corpus of texts (documents) is transformed into a term-document matrix, displaying for each term its occurrence frequency in each document. Documents are given an extensional, vector representation, in which dimensions of the vector representing a document are the terms occurring in. The aim is to identify and extract words or terms that best characterize the contents of each document using a statistical weighting process. The more frequently a term k occurs in a document xi, the more important for document it is (term frequency \( wx_{i,k} \)).

Hence weight of each term describes the importance of the term in the document.

\[
\begin{align*}
  x_i &= (x_{i1}, x_{i2}, \ldots, w_{xi,k}, \ldots, w_{xi,d}) \\
  \text{where, } x_{i;k} &\text{ is the numeric value that term } k \text{ takes on for document } i, d \text{ is the number of terms used for document representation. }
\end{align*}
\]

Similarly the user query is also represented as a vector of terms. The user query is broken to series of terms that builds a query vector.

\[ Q = \{\text{index, in, dbms}\} \]

As the document and the users-profile are represented by two vectors, the correlation between an individual and a document is done by selecting documents containing terms matching the terms in the user query vector. The total weight of a document is calculated by taking the summation of the weights of the terms in the document matching the terms in the query vector. The documents are finally ranked based on their weights and given as input to the RL Agent.

**DOCUMENT FILTERING**

In addition to the representation as term vector, the learning objects in the repository are also represented as vector of weights representing the specific preference for each of the FSLSM learning style dimensions.

\[ D = \{w_{vi}, w_{ve}, w_{se}, w_{in}, w_{seq}, w_{g}\} \]
This is done by learning the profiles of users. The retrieval agent then selects N learning objects matching the search terms and estimates the value of every selected learning object that is computed as the inner product:

\[ V(D_i) = \sum_{x=1}^{d} P_x(D_i)_x \]

Where,

- \( V \) - value of learning object \( i \),
- \( P_x \) - profile of preference of learner,
- \( (D_i)_x \) - learning object \( i \) represented in terms of its type and kind \( x \),
- \( d \) - no. of learning style dimensions

These learning objects are then ranked based on the value. The ranks of the learning object \( i \) for profile \( p \) is based on its similarity (or relevance) to the learner’s search query. The candidate learning objects are sorted in descending order of value, and the top M of N learning objects are presented to the user. The RL agent loads the first of the M learning objects for learning and suggests the remaining M-1 learning objects to the learner to read.

**IMPLICIT/EXPLICIT FEEDBACK MANAGEMENT**

Once M learning objects are filtered and presented, the user reads (or ignores) the learning objects which the system considers as learner’s feedback. Feedback obtained may be implicit or explicit.

The explicit feedback is provided by the user while or after he reads the learning object. This feedback type is used in an early stage of interaction between the user and system. After some interactions with the user, system transfers to an implicit feedback mode in which the user does not need to give explicit feedback for the presented learning objects.

The implicit feedback is measured automatically by the system without explicit help from the user. This can be done by analyzing user’s behaviours on the learning objects filtered. Several factors can be measured. In this work, we distinguish three factors: reading time \( (rt) \), scrolling \( (sc) \), and closing \( (cl) \) of learning objects. Implicit feedback are therefore measured as vector of these three factors.

\[ F_i = \{rt,sc,cl\} \]

When a user scrolls or takes considerable time to read a learning object or has worked on the system for considerable time before closing then the respective dimensions of the vector is set. Similarly when a user dislikes the type of learning object presented, he can express his dislike explicitly through the use of ‘Dislike’ button, thereby setting the explicit feedback value to negative. Finally reward is obtained by normalizing the implicit and explicit feedback values. The reward calculated is then added to the learner’s profile of preference. Negative reward is a punishment indicating the system’s bad selection of learner relevant learning object.

Say, for a learning object \( x_i \) presented to the user, RL Agent measures a scalar-valued feedback by observing user behaviours as:
\[ r = (\gamma) R_E + (1- \gamma) R_I \]

where \( R_E \) is an explicit feedback and \( R_I \) is an implicit feedback for learning object \( x_i \).

\( \gamma \) is a regulating factor that adjusts the ratio of implicit and explicit feedback. If \( \gamma \) is zero, implicit feedback is only used. The values are normalized to \( 0 < R_E < 1 \) and \( 0 < R_I < 1 \). The parameter \( \gamma \) controls the relative contribution of each feedback.

• **ASSESSMENT AGENT**

The system provides two choices:

i) Presenting a specific learning object of learner’s need

ii) Presenting an entire course

The latter choice includes providing an assessment by the end of every chapter in the course. After studying the material, learners are given a post-test for assessing their learning outcomes and performance. To assess the efficiency of the learning process, no time limit was imposed on the learners to finish studying the provided material of medium level of difficulty. Once the learner finishes studying he/she is presented with a questionnaire assessment test. Based on their performance the system makes a decision from the choice of proceeding to the next chapter or presenting previous/current chapter with primary level of difficulty.

• **EXPERIMENTAL RESULTS**

The performance of the proposed technique was evaluated. We made two different learners who have different learning styles and preferences to use our system. By interacting with the learners our system perceived that learner 1 preferred visual, sensory and sequential contents while learner 2 preferred verbal, intuitive and global contents. Both the learners were then made to learn the same course for which they were taught with different learning objects.

Table below shows the adaptive course delivery for two different learners. We could see that the learners have learnt the same course with different learning objects. According to learner 1’s learning style and preference learning objects 2, 5, 8, 11, 12, 17 and 20 were provided by the system where as for learner 2, learning objects 1, 4, 8, 9, 12, 16 and 18 were provided by the system.
Table 1. Adaptive course delivery for learners.

<table>
<thead>
<tr>
<th>Time (in mins)</th>
<th>Chapter</th>
<th>User 1</th>
<th>User 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-5</td>
<td>Ch 1</td>
<td>LO 2</td>
<td>LO 1</td>
</tr>
<tr>
<td>5-10</td>
<td>Ch 2</td>
<td>LO 5</td>
<td>LO 4</td>
</tr>
<tr>
<td>10-15</td>
<td>Ch 3</td>
<td>LO 8</td>
<td>LO 8</td>
</tr>
<tr>
<td>15-20</td>
<td>Ch 4</td>
<td>LO 11</td>
<td>LO 9</td>
</tr>
<tr>
<td>20-25</td>
<td>Ch 5</td>
<td>LO 12</td>
<td>LO 12</td>
</tr>
<tr>
<td>25-30</td>
<td>Ch 6</td>
<td>LO 17</td>
<td>LO 16</td>
</tr>
<tr>
<td>30-35</td>
<td>Ch 7</td>
<td>LO 20</td>
<td>LO 18</td>
</tr>
</tbody>
</table>

**CONCLUSION**

Reinforcement learning is a goal-directed intelligent technique and is an efficient tool for E-learning applications. The characteristic of the learning by trial and error makes it a flexible method for employing the user’s feedback in the learning process. This helps to generate an efficient personalization tool to assist the user and prevent the tiredness and dissatisfaction. The objective of this research was to implement the RL for designing of the personalized e-learning tool. Our experimental results show that the RL agent can successfully learn the user’s preferences among the various available types of information thereby establishing adaptivity in e-learning.

In this research, we formulated the problem of document filtering as a reinforcement learning problem, and presented a personalized document filtering system that learns to follow user preferences from observations of his behaviors on the presented learning objects. A practical method was described that estimates the user’s relevance feedback from user behaviors such as reading time, scrolling and closing of documents. Our experimental evidence from a field test on a group of users supports that the proposed method effectively adapts to the user’s specific interests. The online nature of reinforcement learning makes it possible to approximate optimal action policies in ways that put more effort into learning. This is the key property that distinguishes reinforcement learning from other relevance feedback methods based on supervised learning. The reinforcement learning formulation gave more emphasis on decision making as to filtering the learning objects rather than just to learn the mappings or profiles. This resulted in better performance on a long run than simple supervised learning methods in the dynamic environments. Our work suggests that reinforcement learning can provide a better framework for personalization of information in the e-learning environments than conventional supervised learning formulation.

**FUTURE WORK**

Greater system scalability can be achieved in a multiple user environment. Implementation can be done on a distributed and web based architecture to enhance collaboration and interaction between different learners.
Mixed Initiative techniques can be used for interacting with learners in natural language.

REFERENCES


[9] Shehab A. Gamalel-Din, Smart E-Learning: A greater perspective: from the fourth to the fifth generation e-learning, 2010
