# Optimization of Personalized Learning Pathways Based on Competencies and Outcome

Jianhua Lin
Department of Computer Science
Eastern Connecticut State University
Willimantic, CT 06226
linj@easternct.edu

***Abstract*—Personalized learning that is tailored to individual needs, preference, and interests may improve student learning experience and outcome. With the aid of computing technology, it is becoming possible to deliver personalized learning to a large and diverse student population. One of the key problems involved is the determination of the pathway which each learner follows to complete a learning program. Existing methods generally rely on *a priori* knowledge of subject content prerequisite relationship or constraint to determine the sequencing of instructional materials without any consideration of student learning outcome. In this paper, we formulate the selection of learning pathways as an optimization problem based on competencies and student learning outcome. We show that the resulting pathway selection problem can be modeled as a Markov Decision Process (MDP). Decision rules can thus be defined and applied to select personalized learning pathways to optimize student learning outcome according to desired performance criteria.**

***Keywords-personalized learning; adaptive learning; learning pathway selection***

## I. Introduction

Personalized learning delivers significant benefits by providing learners with an educational experience that is paced to individual needs and tailored to different learning preferences and interests [1]. In order to enable personalized learning at scale for a large and diverse student population, a technology-based solution is necessary. A growing number of technology products has emerged in recent years to support personalized learning [2, 3]. More advanced adaptive techniques have also been developed to further enhance the student learning experience using data analytics to direct the personalization [4].

One of the key problems in personalized learning is the selection of a learning pathway which each learner follows to complete a learning program. In a traditional classroom, it is generally the teacher's responsibility to develop a plan of instruction that facilitates student learning and progress. Current solutions for the sequencing of learning materials typically rely on some *a priori* knowledge of the prerequisite relationship among the subject content without any consideration of student learning outcome [5]. Various methods have been developed using techniques such as workflow management [6, 7] or swarm intelligence optimization [8].

In this paper, we formulate the learning pathway selection as an optimization problem based competencies and assessment of student learning. The objective is to find the optimal personalized learning pathway that enables each learner to achieve the best learning outcome possible. We show that selecting an optimal learning pathway can be modeled as a Markov Decision Process (MDP) for which various solution techniques are known [9]. We show that it is possible to develop solutions for selecting personalized learning pathways to optimize student learning outcome according to desired performance criteria.

## II. Optimizing Learning Pathway Selection

In a typical personalized learning environment, instructional content is provided in a repository of learning objects, reusable learning units which can be assembled together to create lessons or courses [10]. Each learning object may contain a variety of learning resources that a learner can use to gain some specific competency which measures the mastery of skill or knowledge [11]. While competencies can be defined in various forms [12, 13, 14], for the purpose of this paper, it is sufficient to consider a competency as a specific learning outcome that can be explicitly assessed. The overall performance from the assessment of the competencies provides an indicator of student learning. We use competencies to define the possible pathways which a learner follows to complete a learning program. Different learning pathways may lead to different performance result for learners so the goal is to find the optimal pathway that enables each learner to achieve the best learning outcome possible.

Let $C = \{c_1, c_2, \ldots, c_n\}$ represent a finite set of competencies to be completed in a learning program. Each of the competencies is associated with a set of Learning Objects (LO), used by learners to acquire the corresponding knowledge or skills. Each of the competencies is also associated with a set of Assessment Modules (AM), which can measure a learner's mastery of the competency. When a learner completes an assessment module, a performance value between *0* and *1* is received, with *0* representing the lowest performance and *1* the highest.

**A learning pathway** is defined to be a permutation of $C$, $P_\pi = (c_{\pi_1}, c_{\pi_2}, \cdots, c_{\pi_n})$, prescribing the order in which a learner follows to complete the competencies. Any part of a learning pathway constitutes a pathway for a competency

2161-377X/16 $31.00 © 2016 IEEE
DOI 10.1109/ICALT.2016.99

subset of $C$. Specifically, for each $t = 1, 2, \ldots, n$, the sequence of the first $t$ competencies in $P_\pi$, $P_{\pi_t} = (c_{\pi_1}, c_{\pi_2}, \cdots, c_{\pi_t})$, represents a learning pathway of the competency subset $\{c_{\pi_1}, c_{\pi_2}, \cdots, c_{\pi_t}\}$.

Let $r(P_{\pi_t}) \in [0, 1]$ denote the performance assessment result received by a learner for competency $c_{\pi_t}$ after completing the learning pathway $P_{\pi_t}$. The cumulative performance of a learning pathway $P_\pi$ for a learner is defined as follows:

$$R(P_\pi, W) = \sum_{t=1}^{n} w_t r(P_{\pi_t})$$

where $W = \{w_1, w_2, \ldots, w_n\}$, is a set of weights assigned to reflect the relative importance of each competency in the set:

$$\sum_{i=1}^{n} w_i = 1 \text{ and } w_i \in [0, 1], 1 \le i \le n$$

**The learning pathway optimization problem**: For a given competency set $C = \{c_1, c_2, \ldots, c_n\}$, a learning pathway $P_\pi$ is *optimal* for a learner if $R(P_\pi, W) \ge R(P_{\pi'}, W)$ for all possible learning pathways $P_{\pi'}$.

As the total number of permutations for $n$ competencies is $n!$, finding an optimal learning pathway through exhaustive search is not feasible. To find a solution, we show how the pathway optimization problem can be modeled as a well-known Markov decision process (MDP) for which various algorithmic techniques have been developed.

## III. Learning Pathway Selection as a Markov Decision Process

Markov decision processes are general models that have been widely used to solve sequential decision making problems in many application domains [9]. Consider a finite horizon, discrete time Markov decision process in which decisions are made at points of time referred to as decision epochs, $T = \{1, 2, \ldots, n\}$. At each decision epoch, the process occupies a state $s \in S$, where $S$ is the set of all possible states. Each state is associated with a set of allowable actions, $A_s$. When an action $a \in A_s$ is taken in state $s$ at decision epoch $t$, a reward, $r_t(s, a)$ is received by the decision maker and the process will transition to another state $s'$ with a probability $p_t(s'|s, a)$. The collection of all these objects together is referred to as a Markov decision process (MDP):

$$\{T, S, A_s, p_t(s'|s, a), r_t(s, a)\}$$

Given a Markov decision process, a *decision rule* prescribes the method for action selection in each state at a specified decision epoch. Since decision rules lead directly to the reward amount received by the decision maker, they are the main focus of the decision process. The complexity of decision rules varies widely, depending on how they incorporate past information in history and expected outcome of future decisions to select actions. The general problem in a Markov decision process is to discover decision rules that are optimal for a given optimality criterion. For example, if the reward $r_t(s, a)$ is nonnegative for all the states and actions, one optimization criterion would be maximize the expected total reward. That is, let $\pi = (d_1, d_2, \ldots, d_{n-1})$ be the sequence of decision rules applied at all decision epochs, called a *policy*. Note that, in a finite horizon MDP, no action is taken at the last decision epoch $n$. If the sequence of rewards,

$$(r_1(s_1, a_1), r_2(s_2, a_2), \cdots, r_{n-1}(s_{n-1}, a_{n-1}), r_n(s_n))$$

is realized under the decision policy $\pi$, the total reward is defined as

$$R(\pi) = \sum_{t=1}^{n-1} r_t(s_t, a_t) + r_n(s_n)$$

For a given Markov decision process, a decision policy $\pi$ is said to be *optimal* if $R(\pi) \ge R(\pi')$ for all possible decision policies $\pi'$.

The learning pathway selection problem defined in this paper clearly demonstrates the characteristics of a sequential decision process. A learner goes through a series of learning competencies and receives a performance assessment score for each, much like a reward in an MDP. An important requirement of the Markov model is that the transition probability and reward functions depend on only the current state of the process and the action taken by the decision in that state. In a learning pathway, however, the learner's performance result at each step depends on the history of all the past competencies completed, not just the one immediately prior, as required by the Markov model. Decision making problems involving such non-Markovian rewards or actions are not uncommon and various techniques have been developed to deal with them [15, 16, 17]. One common method is to convert the original decision process into an equivalent Markov one by formulating additional states that capture all the relevant information contained in the history of the decision process [15]. Existing, well-developed solutions for MDPs can then be applied.

It is indeed possible to translate the learning pathway selection problem as a finite horizon, discrete time Markov decision problem. For a given finite set of competencies, $C = \{c_1, c_2, \ldots, c_n\}$, let $P_\pi(C)$ denote a learning pathway for $C$ and $P_\pi(B)$ represent a learning pathway for any subset $B$ of $C$. In this context, the learning pathway of an empty set, $P_\pi(\{\})$ is an empty pathway of length $0$. If $P_{\pi_t} = (c_{\pi_1}, c_{\pi_2}, \cdots, c_{\pi_t})$ is a pathway of length $t$, we use the expression, $P_{\pi_t} + c_i$, to form a new pathway of length $t+1$, $(c_{\pi_1}, c_{\pi_2}, \cdots, c_{\pi_t}, c_i)$, assuming $c_i$ is not in the original pathway. For each pathway $P_\pi(B)$, we assign a state, $s(P_\pi(B))$, for it and the set of all possible states is

$$S = \bigcup_{B \subseteq C} \{s(P_\pi(B)), P_\pi(B) \epsilon \Pi(B)\}$$

where $\Pi(B)$ is the set of all possible learning pathways for a set of competencies $B$. We use $s_0$ to represent the state associated with the empty pathway. For each state, the set of actions is defined as $A_s = \{m(c_k), c_k \in C - B\}$, where $B$ is the set of competencies contained in the pathway associated with the state $s$. The reward and transition probability functions are defined respectively as follows:

$$r_t(s, m(c_k)) = r(P_\pi(B)+c_k)$$
$$p_t(s(P_\pi(B)+c_k) \mid s(P_\pi(B)), m(c_k)) = 1$$

For a given competency set $C = \{c_1, c_2, \ldots, c_n\}$ with weights, $W = \{w_1, w_2, \cdots, w_n\}$, the learning pathway optimization problem can be defined as a finite horizon Markov decision problem,

$$\{T, S, A_s, p_t(s'|s,a), r_t(s,a)\}$$

with the optimization criterion starting from state $s_0$,

$$R(\pi) = \sum_{t=0}^{n-1} w_{t+1} r_t(s_t, a_t) + r_n(s_n)$$

where $r_n(s_n)$ is zero as no action is taken at the final decision epoch.

The formulation of the learning pathway selection problem as a Markov decision problem allows us to apply the many known results and solutions from MDP theory. Specifically, *the principle of optimality* [9] indicates that, for a learning pathway to be optimal, any part of the pathway has to be optimal. That is, if the cumulative performance of a learning pathway $P_\pi = (c_{\pi_1}, c_{\pi_2}, \cdots, c_{\pi_n})$

$$R(P_\pi, W) = \sum_{t=1}^{n} w_t r(P_{\pi_t})$$

is the highest among all the pathways on $C$, then for any $i \epsilon \{1, 2, \cdots, n-1\}$,

$$R_i(P_\pi, W) = \sum_{t=1}^{i} w_t r(P_{\pi_t})$$

must also be the highest among all the pathways on $\{c_{\pi_1}, c_{\pi_2}, \cdots, c_{\pi_i}\}$. This principle is the basis for the use of dynamic programming to finding an optimal solution.

Dynamic programming works by searching through the state space and eliminating any non-optimal partial solutions based on the principle of optimality. For the pathway selection problem, unfortunately, the state space is too big for such search to be realistic and a heuristic algorithm will be necessary. We have developed a practical solution to limit the state space and select the best possible personalized learning pathway in polynomial time. Due to the page limit, we are not able to present the detail of our algorithm here.

## IV. Conclusion

The selection of a learning pathway in personalized adaptive learning can be formulated as an optimization problem based on competencies and student learning outcome and modeled as a Markov decision process. It is thus possible to apply dynamic programming for the selection of an optimal learning pathway for each individual learner according to a desired outcome measure. We have developed a practical heuristic solution to address the complexity involved in the optimization. We are currently implementing our solution in college-level classes and will gather and report the results. Our solution can be used as a basis for applying different learning analytics to improve student learning based on a wide variety of different performance and outcome measures.

## References


[1] National Education Technology Plan, “Transforming American Education: Learning Powered by Technology”, http://tech.ed.gov/wp-content/uploads/2013/10/netp2010.pdf, 2010.

[2] A. Newman, “Learning to Adapt: Understanding the Adaptive Learning Supplier Landscape”, Tyton Partners white paper, 2013.

[3] M. Vandewaetere and G. Clarebout, “Advanced Technologies for Personalized Learning, Instruction, and Performance”, Handbook of Research on Educational Communications and Technology, 2014, pp. 425-437.

[4] A. Newman, “Learning to Adapt: A Case for Accelerating Adaptive Learning in Higher Education”, Tyton Partners white paper, 2013.

[5] C. Howlin and D. Lynch, “A framework for the delivery of personalized adaptive content”, 2014 International Conference on Web and Open Access to Learning (ICWOAL), pp. 1-5.

[6] M. Cesarini, M. Monga, R. Tedesco, “Carrying on the e-Learning Process with a Workflow Management Engine”, Proceedings of the 2004 ACM Symposium on Applied Computing, pp. 940-945.

[7] H. Indzhov, P. Zhelyazkova, and G. Totkov, “Supporting adaptive e-Learning: an approach based on open-source learning management systems”, CompSysTech '13: Proceedings of the 14th International Conference on Computer Systems and Technologies, 2013, pp. 321-328.

[8] L. de Marcos, J. Martínez, and J. Gutierrez, “Swarm intelligence in e-learning: a learning object sequencing agent based on competencies”, GECCO '08: Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation, 2008, pp. 17-24.

[9] M. L. Puterman, Markov Decision Processes: Discrete Stochastic Dynamic Programming, New York: Wiley, 2005.

[10] D.A. Wiley, “Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy, in the Instructional Use of Learning Objects”, Wiley, D.A. ed., 2000.

[11] IMS, “Reusable Definition of Competency or Educational Objective - Best Practice and Implementation Guide”, IMS Global Learning Consortium, 2002.

[12] IMS, “Reusable Definition of Competency or Educational Objective - Information Model”, IMS Global Learning Consortium, 2002.

[13] HR-XML, “Competencies (Measurable Characteristics) Recommendation”, HR-XML Consortium, 2006.

[14] IEEE, “Learning Technology Standards Committee (LTSC), Draft Standard for Learning Technology - Data Model for Reusable Competency Definitions”, 2007.

[15] F. Bacchus, C. Boutilier, and A. Grove, “Rewarding behaviors”, Proceedings of the Thirteenth American National Conference on Artificial Intelligence (AAAI), 1996, pp. 1160–1167.

[16] F. Bacchus, C. Boutilier, and A. Grove, “Structured solution methods for non-Markovian decision processes”, Proceedings of the Fourteenth American National Conference on Artificial Intelligence (AAAI), 1997, pp. 112–117.

[17] S. Thiébaux, C. Gretton, J. Slaney, D. Price, and F. Kabanza, “Decision-theoretic planning with non-Markovian rewards”, Journal of Artificial Intelligence Research (JAIR) 25, 2006, pp. 17–74.