



Title	Applying an evolutionary approach for learning path optimization in the next-generation e-learning systems
Author(s)	Tam, VWL; Fung, ST; Yi, J; Lam, EY
Citation	The IEEE 13th International Conference on Advanced Learning Technologies (ICALT 2013), Beijing, China, 15-18 July 2013. In Conference Proceedings, 2013, p. 120-122
Issued Date	2013
URL	http://hdl.handle.net/10722/189846
Rights	IEEE International Conference on Advanced Learning Technologies Proceedings. Copyright © IEEE Computer Society.

Applying An Evolutionary Approach for Learning Path Optimization in the Next-Generation E-Learning Systems

Vincent Tam, S.T. Fung, Alex Yi and Edmund Y. Lam
Department of Electrical and Electronic Engineering
The University of Hong Kong, Pokfulam Road
Hong Kong
Email: vtam@eee.hku.hk

Abstract—Learning analytics is targeted to better understand and optimize the process of learning and its environments through the measurement, collection and analysis of learners' data and contexts. To advise people's learning in a specific subject, most intelligent e-learning systems would require course instructors to explicitly input some prior knowledge about the subject such as all the pre-requisite requirements between course modules. Yet human experts may sometimes have conflicting views leading to less desirable learning outcomes. In a previous study, we proposed a complete system framework of learning analytics to perform an explicit semantic analysis on the course materials, followed by a heuristic-based concept clustering algorithm to group relevant concepts before finding their relationship measures, and lastly employing a simple yet efficient evolutionary approach to return the optimal learning sequence. In this paper, we carefully consider to enhance the original evolutionary optimizer with the hill-climbing heuristic, and also critically evaluate the impacts of various experts' recommended learning sequences possibly with conflicting views to optimize the learning paths for the next-generation e-learning systems. More importantly, the integration of heuristics can make our proposed framework more self-adaptive to less structured knowledge domains with conflicting views. To demonstrate the feasibility of our prototype, we implemented a prototype of the proposed e-learning system framework for learning analytics. Our empirical evaluation clearly revealed many possible advantages of our proposal with interesting directions for future investigation.

Keywords—concept clustering; evolutionary optimizers; hill climbing; learning path optimization.

I. INTRODUCTION

Learning analytics, a growing area of interest to educators, administrators and researchers, is aimed to better understand and optimize the process of learning and its environments through the measurement, collection and analysis of learners' data and contexts. Most online e-learning systems allow the learners or students to specify their own learners' profiles and learning styles. To advise people's learning in a specific subject, many intelligent e-learning systems [1] would also require course instructors to explicitly input some prior knowledge about the subject such as all the pre-requisite requirements between course modules and/or some relationship measures between involved concepts to be arbitrarily and explicitly specified by the course instructors such that

an optimizer will be ultimately employed to find an optimal learning sequence of involved concepts or modules for each individual learner after considering his/her past performance, learner's profile, learning style, and relevant learners' profiles [1]. However, there are several pitfalls in relying solely on the course instructor's input on the relationship among the involved concepts possibly due to the individual biases by human experts. Furthermore, the decision will become more complicated when various instructors hold conflicting views on the relationship among the involved concepts that may hinder any plausible logical deduction.

There were some previous works focused on using statistical or machine learning approaches for the semantic analysis [1] of the relevant course materials and/or the ultimate optimization [2] of learning paths or sequences of involved course concepts or modules. For instance, Wong and Looi [3] proposed an adaptive learning pathway generation approach using the ant colony optimization method. Furthermore, we proposed in a previous work a complete e-learning system framework for learning analytics that can perform an explicit semantic analysis (ESA) [2] on the course materials, possibly aided by the relevant Wiki articles for any missing information about the involved concepts, to formulate the set of individual concepts. This was followed by using a heuristic-based concept clustering algorithm to group relevant concepts before finding their relationship measures. Lastly, a simple yet efficient evolutionary optimizer was used to return the optimal learning sequence. In this paper, we carefully consider to enhance the original evolutionary optimizer with the hill-climbing heuristic, and also critically evaluate the impacts of various experts' recommended learning sequences possibly with conflicting views to optimize the learning paths for the next-generation e-learning systems. To demonstrate the feasibility of our prototype, we implemented a prototype of the proposed e-learning system framework for learning analytics. Our empirical evaluation clearly revealed the possible strengths of our proposal. There are many interesting directions for future investigation including the possible integration of our proposed complete system framework for learning analytics with the cloud computing technologies and/or other existing

e-learning systems. Definitely, this will have significant impacts for developing more adaptive, personalized and next-generation e-learning systems through enhanced ontology analysis on any cloud computing platform. In addition, the integration of hill-climbing and/or other nature-inspired heuristics is worth exploring as it will help to make our proposed framework more self-adaptive and thus potentially extensible to less structured knowledge domains with conflicting views.

This paper is organised as follows. Section II describes the preliminaries that are important for our subsequent discussion, and our previous proposal of learning path optimization for learning analytics. Section III discusses about our enhanced evolutionary optimizer integrated with the hill-climbing heuristic to intelligently select the refined rules as extracted from the ontology analysis to optimize the process of learning for the next-generation e-learning systems. Section IV gives a thorough comparison of our implemented prototype against that of the benchmarking shortest-path optimizer on real engineering courses offered in the University of Hong Kong, and the uses of the hill-climbing heuristic to critically evaluate the impacts of various experts' for the learning path optimisation. Lastly, Section V summarises this work and shed light on various possible directions for future investigation.

II. PREVIOUS WORK

The goal of most previous work to find a good learning path is essentially to search for a fixed sequence of all the relevant concepts while satisfying the knowledge or prerequisite requirement of such concepts behind each course module. A systematic way is to perform ontology analysis [4] or statistical algorithms to extract keywords that may possibly denote key concepts in relevant course modules as based on its frequency of occurrence in the course materials and/or other reasonable factors.

Beside ontology analysis, there is another alternative approach to link up relevant modules with edges in the underlying graph by applying statistical methods. Among these methods, Chen [1], [4] proposed to use the students' answers in a quiz to deduce the implicit knowledge structure of the concerned course. The motive of using students' quiz answers for inferencing is based on the assumption that if most student simultaneously gave wrong answers to any two questions covering concept *A* and concept *B* respectively, we may then deduce that concept *A* and concept *B* may have some association. This method is simple and efficient since it only requires students' answers to construct the underlying concept graph and its linking edges, and students answers could be easily collected in most e-learning system. However, to extract meaningful correlation of concepts/modules from student's answers through this approach, all the students should sensibly give their answers in the quiz, which is an extremely difficult task and there is no vigorous

way to detect whether the students are behaving sensibly or not during the quiz.

In our previous proposal of learning path optimization for learning analytics, we consider to perform an explicit semantic analysis, followed by enhancing the ontology analysis through concept clustering, that is essentially systematic grouping of closely related concepts, and lastly applying an evolutionary optimizer to find an optimal learning path of involved concepts or modules. Given a set of n course modules and their corresponding course materials, we can obtain an optimal learning path through our proposed framework. For more detail on the complete framework, refer to [2].

III. A HEURISTIC-BASED EVOLUTIONARY OPTIMIZATION APPROACH

The original evolutionary algorithm is simple yet efficient. Each chromosome string is defined as a sequence of n integers ranged from 1 to n in which each integer denotes the corresponding course module. Thus, the whole chromosome string represents the learning path/sequence of the course modules covered in the whole course.

Our enhanced evolutionary algorithm is detailed as follows.

- **Fitness function** : the fitness function is a performance indicator used to determine the quality of the generated learning path as measured by the number of precedence rules being violated by the learning path itself. Basically, the more rules the generated learning path is violated, the worse the quality of the generated learning path.
- **Reproduction**: reproduction is the operation to generate new chromosomes by manipulating the "parent chromosomes". Essentially, it includes the crossover, mutation and random generation operation. In our enhanced evolutionary algorithm to optimize for the learning paths, the size of the reproduction pool is 100 chromosomes. After each iteration, the best 5 chromosomes will be carried to the next iteration with 80 new chromosomes generated by the crossover operator, 10 chromosomes generated by the mutation operator and the last 5 generated randomly.
- **Crossover operation**: in each crossover operation, two chromosomes (**X** and **Y**) will be selected by roulette-wheel selection to perform their crossover. In order to avoid illogical learning path, that is having multiple occurrence of the same integers inside the chromosome string, and also retaining the basic sequential order of both chromosomes, a special segment-based crossover scheme is used in which the randomly selected segment ranging from $i \dots j$, where $i < j$, will be swapped between the two chromosome **X** and **Y**.
- **Mutation operation**: the mutation operation randomly selects two indice and swap the serial numbers in the involved chromosome.

Our enhanced evolutionray algorithm will iterate to apply both crossover and mutation operators to evolve new offsprings of learning paths until the solution is converged, or the maximum number of generations preset as 10,000 is exceeded. In each iteration, after the new offsprings are generated, our enhanced evolutionary optimizer will invoke the hill-climbing heuristic function to perform the hill climbing on the fitness function to check whether the quality of the new learning paths is better than that of the parent chromosome. In case not, the hill-climbing function will call the crossover and mutation operator continuously until the resulting quality of the generated learning path is improved.

IV. AN EMPIRICAL EVALUATION

To demonstrate the effectiveness of our proposal, a prototype of our ontology based analyser integrated with the enhanced evolutionary algorithm was implemented and evaluated on 4 undergraduate Engineering courses.

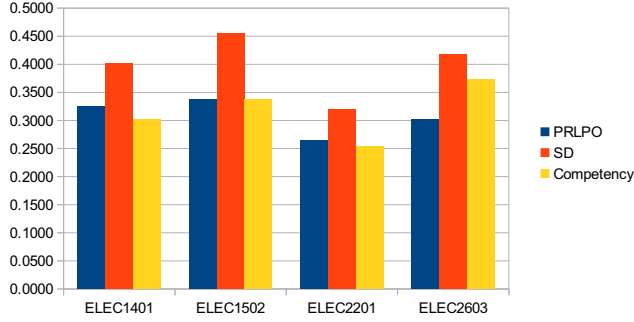


Figure 1. Performance of learning paths generated by our proposal (Ours) against those generated by the shortest distance (SD) and another competency based approach.

Figure 1 gives the total violated distance of learning paths generated by our proposal against those generated by using the SD and competency based approach. The results clearly show that our proposal consistently outperforms the SD approach by returning better learning paths in all cases. It is worth noting that the performance difference in ELEC1502 and ELEC2603 is significantly larger. This is probably due to the fact that the amount of knowledge encapsulated in a course concept is larger in ELEC1502 and ELEC2603 as we regard each set of lecture notes as one single concept. In such case, most of the concepts are loosely correlated and having more course modules requiring course modules as knowledge requirement, thus making it difficult for the SD approach to search for a reasonably good learning path.

When conflicting views as opposite rules occurring in different experts' recommended learning sequences, our evolutionary optimizer may apply some filtering mechanism such as formulating a common sets of rules among all recommended sequences or simply using the hill-climber (HC)

Cases	With HC	Without HC	% of Impr.
fitness	0.9874	0.8721	13.22%
common rules	43.1	44	-2.05%
recomm. path 1	2078.1	2151.9	-3.43%
recomm. path 2	2087.2	2184.8	-4.47%
recomm. path 3	2147.8	1847.7	16.24%
recomm. path 4	930.7	914.4	1.78%

Table I
THE AVERAGED QUALITIES OF LEARNING PATHS FOR ELEC1401 WITH/WITHOUT HC.

to determine the quality of each individual recommendation to the overall search before biasing toward some recommended sequences. Table I shows the averaged qualities of the learning paths generated by the HC over 10 runs on the fitness values, common rules or any of the 4 expert's recommended paths for ELEC1401. It is obvious that the largest gain can be obtained by the HC on the path 3 and also the fitness values.

V. CONCLUDING REMARKS

In this paper, we consider a heuristic-based evolutionary algorithm integrated with the hill climbing heuristic for a more thorough e-learning system framework that carefully integrates semantic semantic analysis, concept clustering and learning path optimization. Besides promoting the fitness values of newly generated learning paths during the evolutionary search, the hill climber can also be used to carefully select the more valuable reference path out of experts' recommendation. To demonstrate the feasibility of our proposal, a prototype of our enhanced evolutionary optimizer integrated with concept clustering and rule-based optimizer was implemented. Its performance was compared favorably against the benchmarking shortest-distance optimizer on various actual courses. More importantly, our proposal clearly demonstrates the importance of heuristics and rule selection for the overall performance of learning path optimization for learning analytics.

REFERENCES

- [1] C.-M. Chen, C.-J. Peng, and J.-Y. Shiue, "Ontology-based concept map for planning personalized learning path," in *Proceedings of the IEEE Cybernetics and Intelligent Systems*, Chengdu, November 2008, pp. 1337–1342, ISBN: 978-1-4244-1673-8.
- [2] S. Fung, V. Tam, and E. Y. Lam, "Enhancing learning paths with concept clustering and rule-based optimization," in *Proceedings of the International Conference on Advanced Learning Technologies (ICALT)*. IEEE, July 2011, pp. 167 – 171.
- [3] L. Wong and C. Looi, "Adaptable Learning Pathway Generation with Ant Colony Optimization," *Educational Technology and Society*, vol. 12, no. 3, pp. 309–326, 2009.
- [4] C. Chen, "Ontology-based concept map for planning a personalised learning path," *British Journal of Educational Technology*, vol. 40, no. 6, pp. 1028–1058, 2009.