

# Semantic Relation Analysis and Its Application in Cognitive Profiling

Taiyu Lin, Kinshuk<sup>1</sup> and Sabine Graf<sup>2</sup>

Massey University, New Zealand, taiyu.lin@gmail.com

<sup>1</sup>Athabasca University, Canada, kinshuk@ieee.org

<sup>2</sup>Vienna University of Technology, Austria, sabine.graf@ieee.org

## Abstract

*Semantic web is an emerging paradigm that has great potential for the management of web content in a meaningful manner. With more and more semantic information appended to web, the web logs can be used to find valuable information about users. Their preferences, characteristics, cognitive capacity, or goals can be interpreted from the patterns of relations between traversed web nodes. A novel approach called semantic relation analysis (SRA) is proposed in this paper to harvest the new opportunities created by semantic web. In this paper, literature about navigational pattern analysis is presented. Example of using SRA in a learning system is also given. Empirical study about SRA showed promising results.*

## 1. Introduction

With wider adoption of semantic web, a great opportunity is arising: user profiling by the semantic relationships. Semantic relationships are the relationships of classes in the conception of ontology. Activities of users on the web have been recognised as a source to provide information about the users [1]. Traversal of the web pages can therefore be represented as a series of activations of semantic relationships. If a certain pattern of occurrences of relationships can yield information about users (or software agents), be it preferences, characteristics, cognitive capacity or goals, interpretations can then be made to obtain useful information.

This paper introduces a novel approach called *semantic relation analysis (SRA)* for analysing learner's navigational patterns. Based on the learner's action patterns, this approach can perform domain-independent analysis. It is used to create Cognitive Trait Model (CTM) of the learner, which profiles the learners according to their cognitive capacities such as inductive reasoning ability and working memory capacity [2]. This paper starts with an overview of existing approaches of navigational pattern analysis

and compares them with SRA. An example of SRA in a learning system is then demonstrated. Empirical study about the proposed approach of SRA to analyse student behaviour information in CTM is then presented. Finally, a discussion on the benefits of SRA and its potential concludes this paper.

## 2. Content-less and Content-based Analysis

In content-less navigational pattern analysis, every web page is treated as a node in the hyperspace, and every node is treated equally regardless of its content [3]. Relationships between the nodes are not defined. The focus is on the navigational behaviours, and certain navigational pattern may indicate certain type of navigational approach, which can reveal what the users are actually doing. The navigational approaches discussed in [3] include activities such as scanning, browsing, searching, exploring, and wandering. The most important advantage of content-less navigational analysis is its domain-independence, which makes it reusable across different domains.

However, its domain-independence makes it inaccurate in specific situations; following is an example to illustrate this point. While frequently revisiting a particular node may give impression that the learner's working memory capacity is not sufficient enough to allow the learner to proceed on the course smoothly, but if the frequently revisited node is a reference node (e.g. a periodic table in Chemistry) then it is perfectly sensible that every (or nearly all) learners need to revisit the node frequently. Content-less analysis would not be able to differentiate a reference node from an ordinary node.

The other approach is called content-based navigational analysis [3]. It is the primary analysis method for most of the performance-based student models [4]. Performance-based models record learners' progresses and grades in learning domains. The contextualised semantic information of every node is recorded by the learning system based on the semantic information in the concept map or ontology of the

domain. The learning system then makes inferences about the learners (their interests, skill levels, and so on) from their domain performances, and stores these inferences in a student model. It was pointed out that the most serious drawback of the content-based approach is that it is totally domain-dependant, which means that similar effort of analysis has to be carried out for every new domain or curriculum [3]. If domain ontology is modified, inference rules may need to be laboriously re-written.

Given that web pages in the learning system are treated as learning objects, information about the relationship of learning objects is becoming more widely available with the popularity of semantic web. The semantic relationships can therefore become a great source of navigational pattern analysis.

### 3. Semantic Relation Analysis

In the context of learning systems, ontology or concept map of the domain are likely to use one of the learning object metadata standard, such as Learning Object Metadata (LOM) [5]. There are already definitions of relationships in LOM that can readily be adopted to be the relationships of the classes in ontology. Following section discusses the learning object relations.

A learning object in a learning context may relate to other learning objects in the same context or outside the current learning context. In hypermedia learning environments, relationships of learning objects could be instantiated as links. They provide means of navigation in the learning environments. If all learning objects are thought to be nodes in the hyperspace of learning, then their relations are the connections from node to node.

In every learning environment, learning objects relate to learning objects differently. For example, two learning objects (say B1 and B2) may be related to another learning object (say B) by an "IsPartOf" relation, whereas learning object B may be related to each learning object B1 and B2 by a "HasPart" relation. There are many types of relations among learning objects, and they can be put into categories such as IsPartOf and HasPart. Twelve different types of categories have been defined in LOM based on Dublin Core [5]. The number of categories was further extended in [6] to include new members such as HasExample, ExcursionTo, Evaluates, and References.

Categorisation of relations enables SRA to be domain-independent, which is the most important advantage of content-less navigational pattern analysis over content-based analysis. In addition, SRA still

contains some semantic information about the relations of the learning objects, which makes it possible to perform analysis that combines the strengths of content-base and content-less analysis.

A learner's interactions with the learning environment can be represented by a sequence of activations of learning object relations. SRA is used to discover information about the learners. In its application in CTM, SRA is used to discover information about the learner's cognitive traits (e.g. inductive reasoning ability). In order to show how SRA is used in CTM, an example of SRA is presented next.

## 4. Empirical Study

There were a total of 29 subjects participated in this empirical study. The participants were students of Massey University in New Zealand studying in Information System course.

### 4.1 Materials

Subjects in this study used the learning system, in conjunction with a web-based inductive reasoning test called Web-IRA. Web-IRA consists of thirty questions. The questions in Web-IRA include three types of tasks in inductive reasoning tests, namely series extrapolation, analogical reasoning and exclusion, discussed in [6]. Web-IRA has a web interface. What a subject needs is a standard web browser to access Web-IRA. The materials are text and images only. No additional software packages or plug-ins are required to access Web-IRA.

The questions are presented to subjects in a sequential order. A question must be solved before a subject can move to the next one. And once solved, the subject cannot return to it anymore. This is to ensure that the subject cannot go back to the same question to cheat. The Web-IRA system can be either deployed onto a local computer network or the World-Wide-Web (WWW), and the working conditions of both deployments are basically identical – Web-IRA is taken individually by using web browsers. While deployed in a local computer network, experiments can be conducted in a supervised group in computer laboratories.

The ability to deploy Web-IRA on WWW increases its availability especially for those experiments where supervision is too costly or impossible. Web-IRA is also ideal when the experiment needs to bring back participants who are no longer accessible (e.g. graduated) for follow-up study. But due to the higher possibility of interference while

subject is taking the Web-IRA in an unsupervised environment, timing mechanism is built into Web-IRA so that detection of abnormality can be possible. Also the timing mechanism is designed in a way that it starts after the web page is loaded, so that subjects with slow computers or connection speed are not disadvantaged.

In addition, Web-IRA can also be used in the traditional one-supervisor-to-one-subject experimental design. In this case, userIDs and passwords are only given to the supervisor who can then login on behalf of the subject at the beginning of the experiment.

## 4.2 Procedure

Subjects in this study were each given a set of userID and password to login to the learning system and the Web-IRA. In the learning system, a subject could just login and start using it. They could read the descriptions of the concepts in the domain or take quizzes. Quizzes were multiple-choice questions.

For the Web-IRA task, the subject was first shown an instruction page that explained the nature of the task after login. After reading the instruction, the subject could proceed to the first question. After answering the first question, the second question was displayed and so on. After 30 questions, the subject was given a message signalling the end of the task and to close the browser window.

## 4.3 Results and Discussion

The index of performance in the Web-IRA task is the total number of correct answers among the 30 questions. Our data analysis found that the performance of the Web-IRA task is significantly correlated ( $\rho=0.382$ ,  $\text{Sig.}=0.02$ ) to the approximation of inductive reasoning ability in cognitive trait model. This correlation empirically supports the proposed approach of using SRA in cognitive trait model to create cognitive profiles of learners.

## 5. Summary

Semantic relation analysis was made possible by the endeavours to append more semantics in the web. Existing standardisation of learning objects, such as

LOM, provides a well-structured basis for applying semantic relation analysis (SRA) in learning context.

SRA has the advantage of domain-independence that is similar to that in content-less navigational pattern analysis. However, it has more semantic information than what the content-less navigational pattern analysis could offer and thus is free of the inaccuracy problem embedded in the content-less navigational pattern analysis. Because of its domain independence, once an analysis task is done, it is possible to reuse the result in new domains without having to carry out the same analysis task again.

The discussion in this paper demonstrated an exemplary application of the SRA in a learning system to create the cognitive trait models of learners. The strong correlation between the measure from the psychometric tool Web-IRA and the approximation of inductive reasoning ability obtained in cognitive trait model proves the usability and potential of SRA.

In the future, investigation will be directed to investigate more new potential application areas of SRA, and to explore how it can be integrated into the tracking service of SCORM 2004 to provide easier integration mechanism for learning system developers.

## 6. References

- [1] S. Loeber, & A. Cristea, "A WWW Information Seeking Process Model". *Educational Technology & Society*, 6-3 2003, 43-52.
- [2] Kinshuk, T. Lin, A. Patel,. "User Exploration Based Adaptation in Adaptive Learning Systems". *International Journal of Information Systems in Education*, 1-1, 2003, 22-31 (ISSN 1348-236X).
- [3] D. J. Mullier, The Application of Neural Network and Fuzzy Logic Techniques to Educational Hypermedia. *PhD Thesis*, 1999.
- [4] A. Mitrovic, & S. Ohlsson, "Evaluation of a Constraint-Based Tutor for a Database Language". *International Journal on Artificial Intelligence in Education*, 10 3-4, 1999, 238-256.
- [5] LOM LTSC, Draft standard for learning object metadata, (2002), Retrieved 05 May, 2002, from [http://ltsc.ieee.org/doc/wg12/LOM\\_WD6\\_4.pdf](http://ltsc.ieee.org/doc/wg12/LOM_WD6_4.pdf)
- [6] F. J. R. van de Vijver, "Inductive Reasoning in Zambia, Turkey, and The Netherlands: Establishing Cross-Cultural Equivalence". *Intelligence* v30 (4), 2002, pp. 313-351