



Causal competencies and learning styles: A framework for adaptive instruction

Vive Kumar, Sabine Graf, Kinshuk

School of Computing and Information Systems, Athabasca
University, Canada
vive@athabascau.ca, sabineg@athabascau.ca, kinshuk@
athabascau.ca

Learning in online environments has the potential to better classroom instruction in many avenues. In this context, this article presents two novel technologies, first, a method to causally model learner competencies, both conceptual and metacognitive, and second, a method to identify learning styles of individual learners. We contend that in both cases it would be extremely difficult for human instructors to thoroughly understand the competencies and competency developments of individual learners as well as the individual learning styles and changes to learning styles. We further contend that these two technologies, as part of a singular framework, will assist classroom instructors to complement their understanding of competencies and learning styles of their classes, respectively, and facilitate instructions to be adapted at various levels of granularity.

for citations:

Kumar V., Graf S., Kinshuk (2011), *Causal competencies and learning styles: A framework for adaptive instruction*, Journal of e-Learning and Knowledge Society, English Edition, v.7, n.3, 13-32. ISSN: 1826-6223, e-ISSN: 1971-8829

1 Introduction

In spite of the recent advances, in our opinion, online learning is still being viewed, at best, as supplemental to conventional classroom-based instruction. Not many believe that contemporary eLearning techniques have the capacity to substitute classroom instruction, particularly, the ability of an instructor to monitor a student's progress, to offer context-specific feedback, and to adapt instruction remain as challenges for the advocates of eLearning. In an attempt to mimic, if not better, these capabilities, researchers of eLearning have developed techniques to model a learner's competencies and to identify learning styles adopted by individual learners.

This article, first, offers a method to model competencies of a learner in causal models. Causal reasoning is ubiquitous in our everyday lives. With the help of computer models, researchers offer and demand a better understanding of causal influences. For example, physicians employ diagnostic computer models, engineers employ architectural computer models, financial analysts employ economic models, and volcanologists employ seismic computer models, to accurately predict cause-and-effect relationships among variables in their respective domains. How, then, about a computer model that establishes cause-and-effect relationships between students' learning interactions and their competencies? Here, we review one particular theory called Self-Regulated Learning that attempts to relate learners' performances to their self-regulatory abilities, and present a causal model of the said theory, created using an exploratory framework.

Secondly, this article presents a method to trace specific learning styles adopted by an individual learner. The information about learners' learning styles can help teachers and learners to better understand the learning processes of learners as well as why learning is sometimes difficult for particular learners while easy for others.

We will individually elaborate these two features and their capacity to be integral components in an online learning environment. We will then discuss how these two features come together in a framework to evolve and adapt online learner competency, content, and interactions.

2 Causal Models of Competency

Competency is a proven skill, where the proof normally resides with factors such as years of experience honing that skill, successful coursework corresponding to that skill, positive feedback from referees about that skill, the number of projects executed involving that skill, and a self-evaluation of one's competency of that skill.

Competency can be associated with task level conceptual skills or meta-

cognitive skills. For instance, in the domain of Java programming, skills at the task level include ‘building a flexible architecture’, ‘leveraging proven architectural methodologies’, ‘integrating patterns and frameworks’, ‘optimising programmer developmental tools’, and so on. In this domain, task level skill set not only addresses conceptual target knowledge a programmer is expected to possess (see <http://www.axmor.com/j2ee-development.aspx> for an example set of target knowledge) but also includes the processes that the programmer employs to achieve the conceptual knowledge. For instance, programmers are expected to use software patterns in their code, which is classified as conceptual task-level knowledge. The development environment (such as the Eclipse or BlueJ Java IDE) could be customized to encourage programmers to use software patterns in their code, which is a task level process. Further, programmers can be engaged in an activity to review their code to ensure optimal use of software patterns. Such skills that are typically aimed at improving performances of task level activities belong at the meta level. Self-regulation, co-regulation, and self-reflection are some of the metacognitive skills one would want to see and promote among.

2.1 Causality

Causality is inherent in human life. We recognise a variety of causes. We understand mechanical causes, for example pressing on a gas pedal causes a car to accelerate and move. We understand psychological causes so that being mean to someone causes them to become angry and defensive. Scientists also make use of causal relationships. In many cases, discovering causality is difficult. For instance, ethical reasons prevent us from conducting experiments to see if smoking really causes cancer. Still, we seek to discover such causal relationships from what we observe. We formally represent these relationships, and provide for mathematical and algorithmic manipulations of the relationships. Such discovery, representation, and manipulation are made possible by computer models such as diagnostic models, economic models, and molecular structural models, all to more accurately predict cause-and-effect relationships.

How then, about a computer model that causally predicts students’ future performances based on their past and current observable interactions and study skills in online learning environments? Can a life-long learning model guide a student to reach their full potential with the help of a set of underlying causal relationships? Can a causal model help teachers to predict why students fail to understand certain concepts? Can a causal model simulate how students evolve an understanding of concepts?

Research has shown that high achieving learners exhibit discernible self-regulatory abilities such as goal-setting, self-monitoring, seeking help, and self-

efficacy (Boekaerts & Corno, 2005; Winne & Hadwin, 1998). Would it, then, be possible to employ such causal models to promote these self-regulatory abilities among online learners? If possible, then, can we causally measure how well students apply and transfer these abilities in online learning environments such as Moodle? Answers to these questions start with the creation of a causal model that identifies variables and relationships corresponding to these self-regulatory abilities. In this article, we present one such causal model. Further, we show a computational mechanism to validate the model so that one can believe that the proposed causal model coincides with Self-Regulated Learning (SRL) literature (Boekaerts *et al.*, 2005; Azevedo *et al.*, 2006; Biswas *et al.*, 2009).

Research into theories of education seeks to identify causes of learning performance and difficulty. By discovering such causes we gain the ability to intervene and improve performance. However research is hampered by several difficulties, including difficulties of measuring variables of interest, subjectivity of measurements, limited quantities of data, and difficulty maintaining experimental control outside of artificial situations (Rao & Kumar, 2008).

To some degree, these difficulties can be mitigated by the application of learning software to perform measurement efficiently and consistently (Brokenshire & Kumar, 2009). Causal discovery algorithms provide a means of discovering some causal relationships from observational data. While the use of these causal discoveries to directly influence the learning and the instructional processes remains feasible, it requires further studies to validate the scope and depth of their applicability. For instance, in domains such as Programming Languages, the amount of observed online interactions that can be mapped to specific programming skills is quite high. Further, the reliability and atomicity of an observed individual interaction (e.g., a learner annotation corresponding to a single type of 'bug' in a particular coding assignment) as to its intended purpose, extent of its use, its success, its failure, and its variations over time play a critical part in considering that single interaction as a potential data element for the causal model. On the other hand, domains such as History or Theory of Computing tend to yield observed online interactions pertaining to activities related to reading. While the amount of observed online reading could yield higher numbers, the reliability with which one could map online reading activities to reading comprehension still remains a research challenge. Similarly, the atomicity of reading activities requires further exploration before one could define an explicit mapping between observed reading interactions and a reading skill. Thus, the nature of the online domain and the types of online interaction determine how well one could reliably observe a significant number of learning related interactions within the scope of a particular skill. It is our contention that the context of a learning activity could be more accurately defined (atomicity) and more reliably mapped onto specific skills as long as the

granularity of the learning activity is considerably small. That is, the creation of a context that contains what the learner did in debugging a specific piece of code, when encountering a particular type of error, with the help of a regulated set of resources, while possessing a reasonably defined coding skill set is much more granular and defined than the creation of a context that contains what the learner did in general while reading a paragraph of text.

2.2 Computational Models for Causality

Graphical Causal Models (GCM) are a graph-based technique for representing causal relationships between variables of interest. They began as an outgrowth of Bayesian Belief Networks (BBN) and retain similar semantics, though there are several varieties (Brokenshire & Kumar, 2009). In general, a node represents either a continuous or discrete variable, which may have an associated probability distribution or conditional probability table. Directed edges between nodes indicate a directed causal relationship and undirected edges indicate a possible causal relationship (dependency). Various specialized types of graphs have additional semantics.

GCM can be used to calculate the likelihood of an event given evidence as with a BBN, but can also be used to calculate the likelihood of an event given our intervention to cause another event. These two calculations are not equivalent (see e.g. Pearl, 2000, for detailed discussion). This is of particular interest in education and educational technology where we wish to take steps to improve learning outcomes, not just be abstractly aware of them. Given an accurate GCM, such determinations can be made algorithmically.

Construction of a GCM can be performed manually as a knowledge engineering effort including domain experts, or it can be accomplished automatically using causal discovery algorithms, possibly including some domain knowledge as an aid. There are currently two types of causal discovery algorithms. The first is constraint based algorithms such as the Fast Causal Inference (FCI) algorithm presented in (Spirtes *et al.*, 2000) which provide an ‘equivalence class’ of models which are all consistent with the statistics. The second is Bayesian discovery algorithms which compute likelihoods for complete models by Bayesian updating given data.

Each class has advantages. The equivalence class of the constraint algorithms provides a complete set of consistent models, which can be differentiated using experimental means. However, there is no formal analysis of the reliability of the results of the algorithm when data is noisy. The Bayesian methods provide a fully oriented model with a likelihood of being correct, but do not provide the entire class of models.

To evaluate the potential of graphical causal models for education we con-

structured several models of one particular educational theory, Self-Regulated Learning. Simulation studies have been performed on these models to evaluate the reliability of the algorithm in discovering the model (Brokenshire & Kumar, *op. cit.*). In a non-traditional use of simulation techniques, these studies have indicated the accuracy with which the proposed models could be discovered from raw data. We also investigate a means of using meta-analysis to gather sufficient statistics from existing publications to act as input to the discovery algorithm.

The following discussion uses both BBN and GCM to exemplify the importance of causal models in online learning. Shown below (Figure 1) is an abstract depiction of a Bayesian Causal Network that estimates programming competency in terms of activities related to code design, coding, documenting the code, debugging the code, and testing of the code.

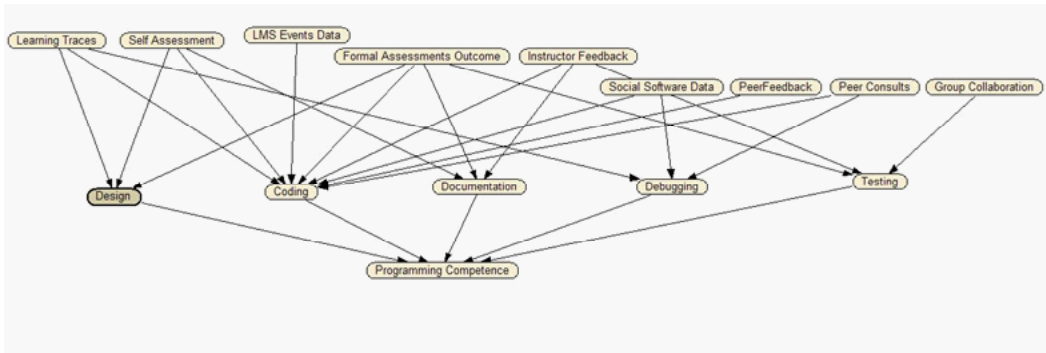


Fig. 1 - A Bayesian Belief Network of Java Competency

The Bayesian network shown in Figure 1 corresponds to 9 competency inputs as shown in the top tier – learning traces, self-assessment, LMS Events data, formal assessments outcome, instructor feedback, social software data, peer feedback, peer consults, and group collaboration. Activities of programmers corresponding to these 9 inputs are then classified into activities contributing to the 5 Java skills – design, coding, documentation, debugging, and testing. In turn, these 5 skills contribute to an estimation of the overall Java programming competency of a learner.

This network has been abstracted based on what could potentially be observed when learners are engaged in online programming activities. For instance, it is quite possible for a peer to offer feedback on the design activities of another learner, as long as these design activities remain observable. At this time, the system that we are currently developing (see <http://kevinhaghighat.com/MILE> for details) only allows coding related activities for the observation of

peers, thus limiting the scope of the Bayesian network shown above. Thus, the scope of the network is limited only by observable and shareable learner interactions. Privacy issues dictate that not all observable interactions are readily shareable. Only the interactions that are explicitly allowed by that learner could be shared.

The node titled “Learning Traces” could employ metacognitive skills such as “Self-Regulation” and supply data pertaining to metacognitive skills. Self-regulation characterizes learners and their learning habits that are typically proactive in nature. Such proactive students are called self-regulated learners and the theory that models and predicts such metacognitive traits is called the theory of Self-Regulated Learning.

2.3 Self-Regulated Learning

Self-Regulated Learning (SRL) theory attempts to explain academic learning and achievement of learners in terms of metacognitive characteristics and processes individuals use to regulate their own behaviour. It emphasizes the student as an active participant in the learning process, as opposed to a passive recipient of information provided by a teacher. It concerns how learners develop learning skills and how they develop expertise in using learning skills effectively (Winne & Hadwin, 1998; Winne, 2001). SRL comprises a set of strategies and tactics employed by learners to regulate their own learning processes. It arises from two key observations. First, learners’ goals for learning take precedence over goals set by teachers, authors of curricula, and developers of learning objects. Second, learners are in charge of how they learn. They choose which study tactics and learning/problem-solving strategies to use as they strive to achieve their goals.

SRL theory covers a large number of variables and situations that interact in a complex and difficult to control environment. The complexity of the environment and large number of variables, many of which are not directly observable, makes it difficult to conduct studies that provide causal relationships. Two types of computational SRL models have been observed in the literature – informal and formal. In informal models (Rao & Kumar, 2008; Shakya *et al.*, 2005), learner interactions in targeted learning activities are mapped onto components of a theoretical framework of SRL (e.g., Winne & Hadwin, *op. cit.*; Zimmerman, 2002), where the mapping between learner activities and SRL variables is informal and is mostly determined by the experimenters. Formal models (e.g., Samsonovich, 2008; Brokenshire & Kumar, *op. cit.*) attempt to formally map learner interactions to specific variables and allow relationships among the variables to evolve into a theoretical framework. In this article, we

focus on formal SRL models where one can ‘discover’ alternative SRL models from existing trace data.

2.4 SRL Model from Popular Theories

In order to create the theory-specific model a literature review was conducted on SRL papers that described the two theories Winne & Hadwin, *op. cit.*; Zimmerman, *op. cit.*) and the relationships these two theories predict, as well as on review articles that summarized the body of empirical work in terms of the theories. This resulted in a collection of papers that were then read closely for variables’ definitions and any correlational or causal relationships proposed between the variables. These relationships were then composed to form a complete model (Brokenshire & Kumar, *op. cit.*).

The models are intended as proofs of concept, demonstrating that the relationships in SRL can be represented in the form of graphical causal models. If we are to use causal models of SRL we must first demonstrate that SRL variables and relationships can in fact be represented in this formalism. The creation of a causal model of SRL from the literature acts as a kind of existence proof, demonstrating that the causal structure of SRL can be captured in this way.

When creating a causal graph the appropriate identification of variables is necessary. There is almost always some choice in how to define a variable, in terms of the discretization of continuous variables or aggregating lower level variables into more abstract variables, and these choices change the structure. Omission of relevant variables can conceal causal effects and aggregation of variables which have different causal structures can result in an unfaithful distribution (Spirtes *et. al.*, *op. cit.*). Identifying the relationships between variables also depends on the presence of other related variables because we determine causal structure by considering how two variables relate in the presence of additional variables. Thus, when using only observational data, excluding some variables from the model can limit the ability to discover causal relationships between modeled variables. Given this, in the case of the theoretical model, it is not expected that the model created corresponds perfectly to either the theory of SRL, or to the correct underlying structure. Approaching such precision in the theoretical model would require enlisting multiple experts in SRL in a knowledge engineering effort.

2.5 Simulation Studies

Algorithms for discovering causal structure from data are only useful if the amount of data required can reasonably be obtained. Theoretical results and simulation studies have shown that simple structures can often be discovered up

to the point of observational equivalence with sample sizes between 1000 and 10000 (Spirtes *et. al.*, *op. cit.*). To evaluate the possibility of learning the causal structure of SRL theory from observational data, we conducted simulation studies using the theory-specific model to establish approximately the quantity of observational data required to correctly recover the equivalence class. If the model accurately reflects the theory, or has a similar structure and sparseness, this should provide an idea of the quantity of data required to learn the model from real data (Brokenshire & Kumar, *op. cit.*). The simulations are run using the TETRAD IV software package (Scheines). Note that the data used in the simulation have been generated by the TETRAD system based on a sample, engineered, theory-specific, reference model. Reference models can be built based entirely from readings of literature corresponding to the theory or from a single meta-analysis that yield correlation matrices. For example, Robbins *et al.* present the results of a meta-analysis of 108 articles relating psychosocial and study skill factors to college outcomes (Robbins *et. al.*, 2004). The studies discussed here only used simulation data and did not use observational real-world data reported in (Rao & Kumar, *op. cit.*; Shakya *et al.*, *op. cit.*).

Since the theory-specific model is not parameterised, and to avoid any biasing effects from a particular parameterisation of the variables, each simulation run used a different randomly generated parameterisation. Each variable was assumed to be discrete, and to take between two and four values. 30 or more simulation runs each were done with samples of 1000, 2000, 5000, 10000, 20000, and 50000 complete data instances. A complete instance of data is a vector with one element for each variable in the sample model. It can be considered a simultaneous measurement of all of the variables in the model. The Partial Ancestral Graphs (PAG) produced by the FCI algorithm at each sample size were compared with the PAG produced directly from the conditional independence relationships, as well as being compared with the theoretical model itself.

The simplified theory-specific reference model is presented in Figure 2, along with the equivalence class found by FCI given perfect data. The reference model uses causal links between variables based on readings in the literature. The equivalence model consists of causal links (e.g., effort --> performance), confounding links (e.g., level of goal challenge o--o interest), and partial links (e.g., strategy knowledge o--> strategy use).

Using TETRAD, we then compared the equivalence class models with the reference graph. One of the key criteria for comparison is the adjacency among nodes. With sample sizes of up to 5000, the majority of the adjacencies are correctly identified but there are large numbers of false negatives where adjacencies were not correctly identified. False positives were rare, with no false positives being the most common case at all sample sizes. This lends credence to the result from the equivalence class that the presence of an adjacency in the

discovered graph is strong evidence for its existence. As sample sizes increased the results for adjacency detection begin to converge to correctly identifying the complete set of adjacencies with very low rates of both false negatives and false positives. Thus, one can conclude that FCI is reliable in generating equivalence models with respect to adjacency when compared with a reference model, for higher sample sizes.

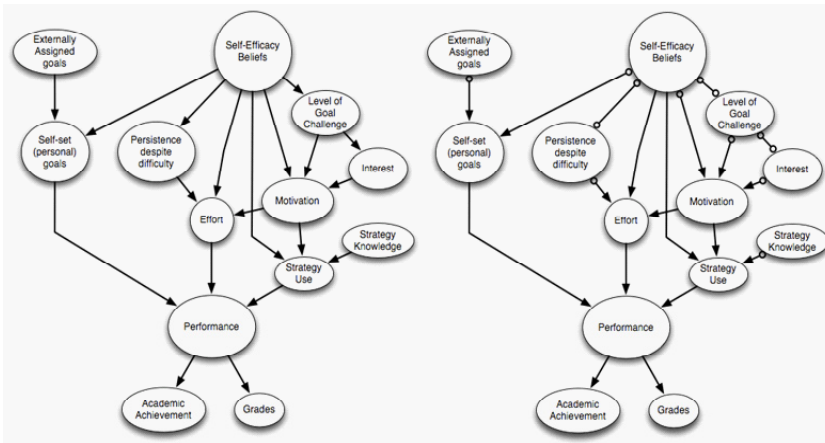


Fig. 2 - the simplified theoretical SRL model and its equivalence class SRL model (Brokenshire & Kumar, *op. cit.*)

Another criteria for comparison is the accuracy of the orientations of the links between nodes. Brokenshire and Kumar (Brokenshire & Kumar, *op. cit.*) report that the orientation results are less consistent. At very low sample sizes of 1000 and 2000, only 9 to 12 of the 16 possible arrow points are correctly identified. As sample sizes increase to 5000 and above, the algorithm begins recovering all of the arrow points it can correctly recover. However, the algorithm produces a large number of false positive arrow points at low sample sizes, and false positives continue to occur, even at sample sizes of 50,000. This limits the confidence one can have in the orientations produced by the algorithm, particularly at low sample sizes. One possible technique to reduce the number of false positive orientations is to include background information about links which we are certain are not allowable due to temporal relationships or other constraints between the variables. Further study should investigate the effects of including such temporal information on simulation results.

One of the key findings of the proposed approach of generating equivalence models using FCI is that the number of experiments required to confirm undetermined endpoints (endpoints labeled with a o). It is quite possible to use less-

than-ideal equivalence models generated by the FCI solely using observed data from an individual learner or a group of learners. These equivalence models could postulate variables that have been observed from the data and potential causality between observed as well as unobserved variables. If a researcher is interested in a particular relation between any two nodes in the model, it is quite possible to verify that such a relation is non-existent in all possible equivalent models produced by the FCI, thus allowing the researcher to proceed with a real-world experiment to observe and validate the said relation. Brokenshire and Kumar (Brokenshire & Kumar, *op. cit.*) show that an estimated 73% reduction in real-world experiments to confirm undetermined endpoints.

3 Identifying Learning Styles

This section focuses on the identification and consideration of learning styles in learning systems. The field of learning styles is complex and many different learning style models exist (e.g., Felder & Silverman, 1988; Kolb, 1984; Honey & Mumford, 1982). While there are still several open issues with respect to learning styles (Coffield *et al.*, 2004), all learning style models agree that learners have different ways in which they prefer to learn. Furthermore, many educational theorists and researchers consider learning styles as an important factor in the learning process and agree that considering them in education has high potential to facilitate learning. For example, Felder pointed out that learners with a strong preference for a specific learning style may have difficulties in learning if the teaching style does not match with their learning style (Felder & Silverman, *op. cit.*; Felder & Soloman, 1997). From theoretical point of view, conclusion can be drawn that incorporating learners' learning styles in a learning environment makes learning easier for learners and increases their learning efficiency. On the other hand, learners whose learning styles are not supported by the learning environment may experience problems in the learning process.

Based on these theoretical arguments, several adaptive learning systems have been developed over the last years. Examples of such systems include CS383 (Carver *et al.*, 1999), WELSA (Popescu, 2008), and TSAL (Tseng *et al.*, 2008). Evaluations of these systems demonstrated the possible benefits of considering learning styles in learning systems, showing that the required time for learning can be decreased and the overall learner satisfaction can be increased.

The first step towards incorporating learning styles in technology enhanced learning is to identify learners' learning styles. Brusilovsky (1996) distinguished between two different ways of student modelling: collaborative and automatic. In the collaborative approach, the learners provide explicit feedback which can

be used to build and update a student model, such as filling out a learning style questionnaire. In the automatic approach, the process of building and updating the student model is done automatically based on the behaviour and actions of learners while they are using the system for learning. The automatic approach is direct and free from the problem of inaccurate self-conceptions of learners. Moreover, it allows learners to focus only on learning rather than additionally providing explicit feedback about their preferences. In contrast to learning style questionnaires, an automatic approach can also be more accurate and less error-prone since it analyses data from a time span rather than data which are gathered at one specific point of time.

Due to the advantages of identifying learning styles through automatic student modelling, we investigated the potential of automatically identifying learning styles with respect to the Felder-Silverman learning style model (FSL-SM) (Felder & Silverman, *op. cit.*). We considered different sources for this identification process, including students' behaviour patterns in online courses, their navigation patterns and their cognitive traits. In this paper, we propose an architecture that considers all three sources.

3.1 Using Behaviour Patterns for Detecting Learning Styles

An automatic approach has been designed, implemented and evaluated Graf *et al.*, 2009), which uses the behaviour and actions of learners, gathered while they are learning, for inferring their learning styles. An important aim of this approach was that it should be applicable for different learning systems. Only few research works exist about automatic identification of learning styles in learning systems (e.g., García *et al.*, 2007; Cha *et al.*, 2006). These works aim at identifying learning styles in particular learning systems and therefore are tailored exactly to these systems by using only those behaviour patterns which are incorporated in the respective systems. Moreover, the investigated courses are created in consideration of learning styles by using particular types of learning objects for detecting learning styles.

When aiming at developing a generic approach for automatic student modelling which can be used for different learning systems, several additional issues have to be considered. First, behaviour patterns have to be selected in a way that most learning systems are able to gather data with respect to these patterns. Furthermore, it needs to be noted that most courses in existing learning systems are not created in consideration of learning styles. Therefore, it is not sufficient that the system can technically track the required information about patterns but teachers also have to use the respective types of learning objects in their courses. Hence, only commonly used types of learning objects were selected as basis for patterns in the automatic student modelling approach.

Moreover, the approach has to consider that nevertheless some data might not be available and therefore has to be able to deal with missing data. Thus, the proposed approach considers a high number of patterns, which is beneficial for identifying learning style accurately.

For each of the four learning style dimensions of FSLSM, relevant behaviour patterns were selected, which were based on commonly used types of learning objects in learning systems. These patterns mainly consider how often students visits particular types of learning objects, how much time they spend on these types of learning objects, how well they do on particular types of questions in quizzes, etc.

For inferring learning styles from these behaviour patterns, a data-driven approach using Bayesian networks and a literature-based approach using a simple rule-based method were implemented. In a study with 127 students, who participated in a university course about object oriented modelling within the learning management system Moodle, both approaches were evaluated. The learning styles calculated from both approaches were compared with the results of the ILS questionnaire (Felder & Soloman, 1997), a 44-item instrument developed by Felder and Soloman for identifying learning styles based on the FSLSM. The evaluation showed that the literature-based approach achieved better results (precision values of the four dimensions ranged from 73.33% to 79.33%) than the data-driven approach (precision values ranged from 62.5% to 68.75%) and identified learning styles with high precision. Hence, the proposed concept including the literature-based approach can be seen as a suitable instrument for automatic detection of learning styles.

The concept for identifying learning styles through the literature-based approach was implemented in a standalone tool called DeLeS (Graf *et al.*, *op. cit.*). DeLeS automatically extracts relevant data from a learning system's database and calculates learning styles by using the literature-based approach. In our research, this tool is used to fill the student model with information about the students' learning styles.

3.2 Improving the Identification Process of learning styles through using additional sources

Besides using behaviour patterns that are based on visits of certain types of learning objects, we also did research on using other sources for improving the identification process of learning styles. We investigated the usage of navigation pattern, indicating how student navigate through a course, as well as students' cognitive traits, in particular their working memory capacity.

Navigational behaviour refers to how learners navigate through the course and in which order they visit certain types of learning objects. The order in

which learners prefer to take in and learn from specific types of learning material and activities, as well as in which order and priority these different types of learning material and activities should be presented for supporting learners with different learning styles is a key aspect of most learning style theories. In our study (Graf *et al.*, 2010), we investigated the navigation patterns of students with different learning styles, using the same data as for the previous study. Assumptions about students' navigation patterns were made based on the learning style theory and then these assumptions were statistically evaluated, using lag sequential analysis (Bakeman & Gottman, 1997). Some examples for such assumptions are that re-submitting exercises is a significant navigational behaviour for active learners and going from one self-assessment quiz to another without solving it is a significant navigational behaviour for intuitive learners. The results of this study showed that most of the investigated assumptions could be confirmed, indicating differences in the students' navigational behaviour depending on their learning styles. These resulting differences in navigational behaviour can contribute in student modeling as additional patterns in the identification process of learning styles.

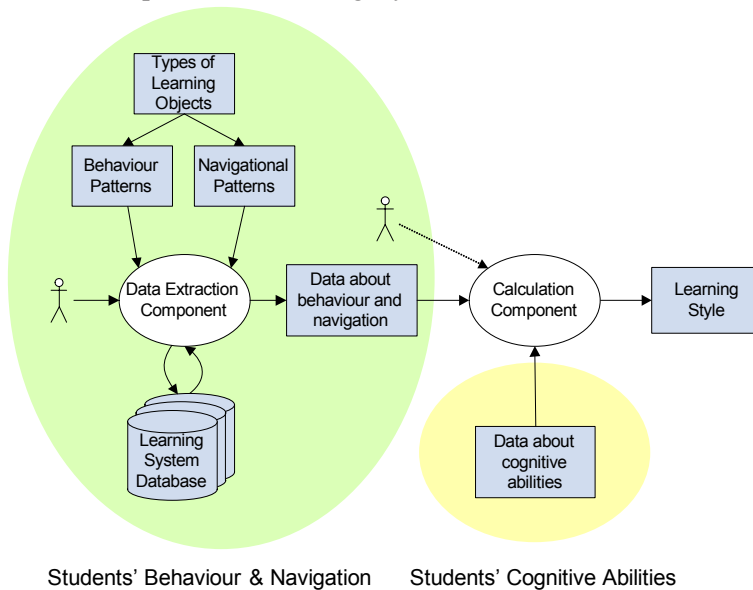


Fig. 3 - Architecture for Identifying Learning Styles

Furthermore, we looked into cognitive abilities, in particular working memory capacity, and investigated the relationships between the four dimensions of FSLSM and working memory capacity. First, a comprehensive literature review was conducted, followed by an experimental study with 39 students,

and then, since both results were promising, a main study with 297 students was conducted (Graf *et al.*, 2009; Web-OSPA, 2010). In both studies, students were asked to fill out the ILS questionnaire (Felder & Soloman, *op. cit.*) in order to get information about their learning styles and perform the Web-OSPA task (Web-OSPA, *op. cit.*) in order to get information about their working memory capacity. The results of these experiments and detailed analysis showed that relationships between working memory capacity and three of the four dimensions of the learning style model exist. These identified relationships have high potential to improve the student modelling process of cognitive abilities and learning styles, as has been demonstrated in an experiment with data from 63 students (Graf & Kinshuk, 2010). In this experiment, results showed that the incorporation of data from the students' working memory capacity can improve the accuracy of the identification process.

Both, navigation patterns and cognitive traits, showed high potential to improve the identification process of learning styles and can therefore be seen as useful extensions to DeLeS. Figure 3 shows the extended architecture of DeLeS, including sources from behaviour patterns, navigation patterns and cognitive abilities. The architecture consists of two main components: the extraction component is responsible for extracting the required data from the learning system's database and the calculation component is responsible for calculating learning styles based on information about students' behaviour patterns, navigation patterns and cognitive abilities.

Conclusions

We have argued for the use of graphical causal models (using Self-Regulated Learning theory as an example) and demonstrated the viability and usefulness of such a course. We contend that graphical causal models provide a useful means of representing the causal claims, in this case, of the underlying SRL theory in a formal and computable form, but they are limited by the availability of sufficient quantities of accurate data.

The exploratory framework taken by the causal discovery algorithms stands in contrast to the confirmatory framework to Structural Equation Modelling. The use of a confirmatory approach in which a model is proposed a priori has the considerable limitation of ignoring the equivalence class of models which can equally account for data. The confirmatory approach is appropriate for disconfirming proposed models, but cannot confirm one model over another equivalent model.

The exploratory framework has the benefit of discovering the complete equivalence class for the available data. A standard challenge of data based methods in machine learning and in science is over-fitting of a model to idiosyncrasies

of the data. The FCI algorithm and related algorithms partially overcome this difficulty by incorporating the faithfulness assumption, but may fail to correctly evaluate relationships when this assumption is violated. The models must of course be tested repeatedly in the same fashion as any proposed theory in order to be considered valid.

The creation of graphical causal models representing educational theories offers multiple benefits. They require a clear and precise specification of the claims of a theory and the definitions of the variables, and represent those claims in an understandable form. This formal, understandable representation should allow for clearer specifications of causal claims in the theoretical literature. More importantly, graphical causal models can answer questions specific to a particular set of learner competencies and how each of these competencies can potentially be improved with respect to dependent variables in the model. Thus, instructors can use GCM to have an overall competency of the entire class of students as well as the competency growth of an individual student over time. This paves way for highly personalised instruction based exclusively on an individual student's GCM.

While the advantages of using GCM are becoming evident, one should also understand the inherent limitations. As discussed earlier in Section 2.1, the quality (atomicity and reliability) and quantity of the data limits model generation. When using causal discovery algorithms (such as FCI), limitations such as the sample size and the difficulty of evaluating high order conditional independence relationships from a reasonable amount of data need to be taken into consideration before employing the algorithms in real-world applications. The solution to such problems is the same as in any observational study: collect more data, and collect better data. Another limitation concerns the Causal Markov Condition and the Causal Faithfulness Condition assumed under the generation of GCM (J Brokenshire & Kumar, *op. cit.*). Thirdly, for a model with a large number of variables, running the tests for conditional independence at conventional significance levels may result in multiple incorrect results given the large number of such tests required. Increasing the thresholds for significance of the statistical decisions changes the type of mistake likely to be made, as correct results may not meet significant thresholds. Given the reliance of the algorithms on patterns of such results, changing the significance of the decisions can produce very different results from FCI and similar algorithms. Fourthly, the FCI algorithm is exponential in the in-degree (number of parents) of the model. For a sparse graph, FCI runs in a reasonable time, but then becomes quickly infeasible for graphs with many parents. This is directly related to the issue of Bayesian networks face with large conditional probability tables with graphs have high average in-degree. Finally, how truly a GCM can represent a theory in light of the fact that multiple viewpoints of the said theory could be

candidates. For instance, from the literature it has been found that the theory of self-regulation has multiple models that are active. Under the same theory, different models could propose different variables and relations to be measured. While GCM provide another way of analysing the correctness of the current models to the said theory, one should note that models do evolve continuously with new research. There are difficulties in establishing causality between two variables in the theory when multiple models address such a causal relation differently at different levels of abstraction.

The proposed GCM is a generic formalism to understand the evolution of competence. One of the key variables of competency is the set of learning styles adopted and exhibited by a single learner. GCM has the capacity to analyse successful learning styles in specific learning situations in a large population of learners. With this data, one can identify the particular learning style that would be beneficial to a learner in given learning situation. Once learners' learning styles are identified, this information can be used for providing adaptive support such as adaptive courses that match learners' learning styles. In order to provide learners with such adaptive courses, a framework of learning objects is suggested, including commonly used types of learning objects such as content objects, outlines, reflection quizzes, self-assessment quizzes, discussion forum activities, additional reading material, animations, exercises, examples, real-life applications, conclusions and assignments. Based on this framework, adaptivity to fit students' learning styles is provided by changing the sequence in which these types of learning objects are presented as well as annotating types of learning objects which fit well to students' learning styles. For example, for an active learner, self-assessment tests, exercises, animations, and forum activities are annotated as particularly important since they support an active learning style where students prefer to learn by trying things out and discussing with others about the learned material. Furthermore, self-assessment tests, exercises and animations are promoted to be presented in the very beginning of a section in order to spark students' interest in the content of the section. The degree of adaptivity can be enhanced with refinements based on information gleaned from GCM. Once the current learning style of the learner has been identified, the sequence of individual instructional activities can be further adapted based on the generic GCM models for a given population of learners.

It is also conceivable to discover learning styles corresponding to specific theories from observed learner interactions and build models based on these observations. Studies could also treat learning styles as causal data and observe their influences on performance. In this article, we promote the notion that, to be effective, adaptive instruction should employ causal competency and learning styles in a complementary fashion, with the possibility of learning styles being embedded in the competency model as a cause or as an effect.

The exploratory causal framework and the learning style framework could feed on each other's data, and strengthen each other's capacity to adapt as well as contribute to adaptive instruction, over time. Together, we believe, they form an integral component of any online learning environment in adapting online content and online interactions.

Acknowledgements

The authors acknowledge the support of NSERC, iCORE, Xerox, and the research-related gift funding by Mr. A. Markin.

REFERENCES

- Pearl J. (2000), *Causality: Models, Reasoning, and Inference*, Cambridge University Press.
- Spirtes P., Glymour C. & Scheines R.. (2000), *Causation, Prediction, and Search*, The MIT Press, 2nd edition.
- Samsonovich A. V., de Jong K. A., Kitsantas A., Peters E. E., Dabbagh N. & Kalbfleisch M. L. (2008), *Cognitive Constructor: An Intelligent Tutoring System based on a Biologically Inspired Cognitive Architecture (BICA)*, in: Wang P., Goertzel B., Franklin S. (eds), *Frontiers in Artificial Intelligence and Applications*, Proceedings of the first AGI Conference, ISBN 978-1-58603-833-5, 311-325.
- Scheines R.. *Tetrad IV*, Computer Software, URL: <http://www.phil.cmu.edu/projects/tetrad/>.
- Boekaerts M. & Corno L.(2005), *Self-regulation in the classroom: A perspective on assessment and intervention*, *Applied Psychology An International Review*, 54 (2),199-231.
- Boekaerts M., Maes S. & Karoly P. (2005), *Self-regulation across domains of applied psychology*, *Applied Psychology: An International Review*, 54 (2), 149-154.
- Azevedo R., Cromley J. G., Winters F. I., Moos D. C. & Greene J.A. (2006), *Using computers as metacognitive tools to foster students' self-regulated learning*, *Technology, Instruction, Cognition, and Learning Journal*, 3, 97-104.
- Biswas G., Roscoe R., Jeong H. & Sulcer B. (2009), *Promoting self-regulated learning skills in agent-based learning environments*, in *Proceedings of the 17th International Conference on Computers in Education*, 67-74, Hong Kong, Asia-Pacific Society for Computers in Education.
- Cook T. (2001), *Randomized experiments in education: A critical examination of the reasons the educational evaluation community has offered for not doing them*, *Educational Evaluation and Policy Analysis*, 24, 175–200.
- Winne, P. H. & Hadwin, A. F. (1998), *Studying as self-regulated engagement in learning*, in: Hacker D., Dunlosky, J., Graesser A. (eds.), *Metacognition in Educational Theory and Practice*, 277-304, Hillsdale, NJ, Lawrence Erlbaum.

- Winne P. H. (2001), *Self-regulated learning viewed from models of information processing*, in: Zimmerman B. J., Schunk D. H. (eds.), *Self-regulated learning and academic achievement: Theoretical perspectives*, 2nd ed., 153-189, Mahwah, NJ, Lawrence Erlbaum Associates.
- Zimmerman B. J. (2002), *Becoming a self-regulated learner: An overview*, *Theory Into Practice*, 41, 64-71.
- Rao S. & Kumar V. (2008), *A theory-centric realtime assessment of programming*, in: *Proceedings of IEEE's International Conference on Advanced Learning Technologies (ICALT.08)*, 139-143.
- Shakya J., Menon S., Doherty L., Jordanov M. & Kumar V.S. (2005), *Recognizing Opportunities for Mixed-Initiative Interactions based on the Principles of Self-Regulated Learning*, in: *Workshop on Mixed-initiative interactions for Problem-solving*. Arlington, VA, AAAI Fall Symposia.
- Brokenshire D. & Kumar V. (2009), *Learning models of self-regulated learning*, in: *Proceedings of the International Conference on Artificial Intelligence in Education (AIED 09)*, 257-264.
- Felder R. M. & Silverman L.K. (1988), *Learning and teaching styles in engineering education*, *Engineering Education*, 78 (7), 674-681.
- Kolb D. A. (1984), *Experiential learning: Experience as the source of learning and development*, Englewood Cliffs, New Jersey, Prentice-Hall.
- Honey P. & Mumford A. (1982), *The manual of learning styles*, Maidenhead, Peter Honey.
- Coffield F., Moseley D., Hall E. & Ecclestone K. (2004), *Should we be using learning styles? What research has to say to practice*, London, Learning and Skills Research Centre / University of Newcastle upon Tyne.
- Felder R. M. & Soloman B. A. (1997), *Index of Learning Styles questionnaire*. URL: <http://www.engr.ncsu.edu/learningstyles/ilsweb.html> (accessed on 24 August, 2011).
- Carver C. A., Howard R. A. & Lane W. D. (1999), *Addressing different learning styles through course hypermedia*, *IEEE Transactions on Education*, 42 (1), 33-38.
- Popescu E. (2008), *Dynamic adaptive hypermedia systems for e-learning*, Ph.D. thesis, Romania, Université de Craiova.
- Tseng J. C. R., Chu H.-C., Hwang G.-J. & Tsai C.-C. (2008), *Development of an adaptive learning system with two sources of personalization information*, *Computers & Education*, 51 (2), 776-786.
- Brusilovsky P. (1996), *Methods and techniques of adaptive hypermedia*, *User Modeling and User-Adapted Interaction*, 6 (2-3), 87-129.
- Graf S., Kinshuk & Liu T.-C. (2009), *Supporting teachers in identifying students' learning styles in learning management systems: An automatic student modelling approach*, *Educational Technology & Society*, 12 (4), 3-14.
- García P., Amandi A., Schiaffino S. & Campo M. (2007), *Evaluating Bayesian networks' precision for detecting students' learning styles*, *Computers & Education*, 49 (3), 794-808.

- Cha H. J., Kim Y. S., Park S. H., Yoon T. B., Jung Y. M. & Lee J.-H. (2006), *Learning style diagnosis based on user interface behavior for the customization of learning interfaces in an intelligent tutoring system*, in: Proceedings of the International Conference on Intelligent Tutoring Systems, LNCS vol. 4053, 513-524, Heidelberg, Springer.
- Graf S., Liu T.-C. & Kinshuk (2010), *Analysis of Learners' Navigational Behaviour and their Learning Styles in an Online Course*, Journal of Computer Assisted Learning, 26 (2), 116-131.
- Bakeman R. & Gottman J.M. (1997), *Observing Interaction: An Introduction to Sequential Analysis*, 2nd edition, Cambridge, University Press, Cambridge.
- Graf S. (2007), *Adaptivity in learning management systems focussing on learning styles*, PhD thesis, Vienna University of Technology.
- Graf S., Liu T.-C., Kinshuk, Chen N.-S. & Yang S. J. H. (2009), *Learning styles and cognitive traits – Their relationship and its benefits in web-based educational systems*, Computers in Human Behavior, 25 (6), 1280-1289.
- Web-OSPAN (2010), URL: <http://kinshuk.athabascau.ca/webospan> (accessed on 24 August, 2011).
- Graf S. & Kinshuk (2010), *Using Cognitive Traits for Improving the Detection of Learning Styles*, in: Workshop Proceedings of the International Conference on Database and Expert Systems Applications (DEXA 2010), 74-78, IEEE Press.
- Robbins S. B., Lauver K., Le H., Davis D., Langley R. & Carlstrom A (2004), *Do psychosocial and study skill factors predict college outcomes? a metaanalysis*, Psychological Bulletin, 130, 261-288.