A Learning Design Recommendation System Based on Markov Decision Processes

G. DURAND, F. LAPLANTE AND R. KOP
National Research Council of Canada

As learning environments are gaining in features and in complexity, the e-learning industry is more and more interested in features easing teachers' work. Learning design being a critical and time consuming task could be facilitated by intelligent components helping teachers build their learning activities. The Intelligent Learning Design Recommendation System (ILD-RS) is such a software component, designed to recommend learning paths during the learning design phase in a Learning Management System (LMS). Although ILD-RS exploits several parameters which are sometimes subject to controversy, such as learning styles and teaching styles, the main interest of the component lies on its algorithm based on Markov decision processes that takes into account the teacher's use to refine its accuracy.

Key Words and Phrases: Learning design, recommendation system, learning style, Markov decision processes.

1. INTRODUCTION
The design of learning activities (usually called learning design [Durand and Downes 2009]) is a critical step for all teachers, even experienced ones. A design mistake can easily lead to ill-adapted learning activities preventing learners and teachers from meeting their learning and teaching objectives. Many initiatives have been undertaken during the last ten years in the e-learning research community to develop learning design tools for teachers. Some of those initiatives, such as the proposal of learning design formalisms and the implementation of learning design environments, have drawn a lot of interest.

Learning design formalisms are usually XML based specifications that allow the definition of shareable learning activities. The most popular learning design formalism is IMS-LD\(^1\). IMS-LD was designed to be both human readable and machine computable. This duality was exploited in several learning design environments proposing that teachers or instructional designers design and play with learning activities.

Those learning design environments have been built on the idea that the end-user would have a good knowledge of the IMS-LD; however, experience has shown that the use of such tools was not easy and in the end did not meet the essential requirement: assist teachers efficiently.

Cognitive overload had already been identified as a big challenge for Learning Management Systems' editors since the recent integration of numerous business features that significantly raised the amount of information processing required to use these tools. Considering this state of affairs, the main actors of the e-learning industry, though initially interested, did not take up the proposed learning design approaches as they preferred to find other solutions suitable for helping teachers, but at the same time minimizing the cognitive overload. A good example is the Desire2Learn\(^2\) Instructional Design Wizard which guides teachers in choosing a course structure according to a defined instructional design approach (ADDIE). As the end-user does not need to understand the inner learning design formalism of the platform, the specification can be optimized for platform interoperability. SLD 2.0 [Durand et al. 2010], a recent learning design formalism, was indeed designed for that purpose, getting rid of not needed features of IMS-LD and adding missing computable data. Instructional Design methods have been used for more than half a century

Authors' addresses: G. Durand, LCT-IIT, National Research Council of Canada, Moncton, N.-B., Canada. E-mail: Guillaume.Durand@NRC-CNRC.gc.ca; F. Laplante, LCT-IIT, National Research Council of Canada, Moncton, N.-B., Canada. E-mail: Francois.Laplante@NRC-CNRC.gc.ca; R. Kop, LCT-IIT, National Research Council of Canada, Moncton, N.-B., Canada. E-mail: Frederika.Kop@NRC-CNRC.gc.ca.

\(^1\) IMS-LD: http://www.imsglobal.org/learningdesign/
\(^2\) Desire2Learn: http://www.desire2learn.com
bringing a great improvement in the training and education industry; however, the application of those methods does not guarantee success.

Moreover, Instructional Design is more an industrial approach to designing the structure of courses, while learning design would rather help in the choice of activities and content so as to meet the needs of specific learners and teachers.

The Intelligent Learning Design (ILD) initiative aims at developing a seamless component called ILD-RS (Recommendation System) that aims to help teachers with the design of learning activities. ILD-RS works as a recommendation system proposing learning activities to the teacher during the learning design phase. ILD-RS makes recommendations based on criteria related to the learners, teachers and available learning objects. The recommendation mechanism is reinforced by the usage of the learning design. This means that ILD-RS can produce recommendations with few prior learning design phases and improve its recommendations as the number of learning designs increases. Based on Markov decision processes, ILD-RS recommendation criteria include teaching styles and learning styles even though they are still subject to controversy.

2. LEARNING AND TEACHING STYLE IN THE LEARNING DESIGN PERSPECTIVE

2.1 Controversy around Learning Styles

Active research into learning and teaching styles has taken place over at least the past 40 years [Coffield et al. 2004] and is still being carried out up to this date [Coffield et al. 2004; Paschler et al. 2008; Grasha 1996]. There has been a surge in interest in learning styles and teaching styles since the 1990s as the development of diagnostic tools was followed by commercialization and a market in learning style measurement tools was established. The interest stemmed from a curiosity of educationalists and psychologists in human learning and in the measurement of learning. Coffield et al. argue that:

"There is a strong intuitive appeal in the idea that teachers and course designers should pay closer attention to students’ learning styles – by diagnosing them, by encouraging students to reflect on them and by designing teaching and learning interventions around them. Further evidence for the idea that we have individual learning styles appears to be offered when teachers notice that students vary enormously in the speed and manner with which they pick up new information and ideas, and the confidence with which they process and use them.” [Coffield et al. 2004]

A wide variety of learning styles has been identified over the years and Pashler et al. [2008] highlight that there is evidence showing that children and adults express preferences for different types of information and in how they would like information to be presented to them, in addition to differences in predisposition to different kinds of thinking.

Recently, the emerging market in learning styles measurement instruments has led to research in the validation of these tools. Two major studies were carried out in the past six years that researched the validity of learning styles measurement tests and their use to determine teacher interventions. [Coffield et al. 2004; Paschler et al. 2008]. A multitude of learning styles measuring instruments were developed, but not much reliable evidence to show that they actually measured what they said they would measure, or a justification to use them as a basis for teacher interventions. Of course at first sight that is a little worrying when you have the idea for a technology driven application or tool that will connect learning and teaching styles.

However, the same authors highlighted other issues related to this approach: research methodologies used in testing the tools were not very reliable, or at least were varied in their application, so it was not necessarily the use of learning styles to advance teaching and education that was problematic, but the evaluation of the learning style measurement instruments [Paschler et al. 2008].

Another issue that was raised is the complexity of the learning process. It is not just the learning or teaching style that influences people’s learning, but also the topic of study, the context, mode and environment in which learning takes place. These factors might also influence each other to add to the complexity. The concept of ‘learning strategies’ was highlighted for instance by
Entwistle as a more valuable way to determine people’s learning preferences as they also take into consideration the circumstances in which people learn [Coffield et al. 2004].

A possible positive influence stressed in the learning styles debate was the meta-cognitive effect they might have on learning and that reflection on one’s own learning and development will help one’s learning [Schon 2002; Carroll et al. 2008].

A closer look was given to the learning preferences that Grasha [1996] identified after his research in undergraduate classes. He highlighted six preferences:
1. Competitive,
2. Collaborative,
3. Avoidant,
4. Participant,
5. Dependent,
6. Independent.

These do not only reveal something about the learners themselves, but also about their interactions with others and of their actions while learning, and thus take several contextual factors into consideration.

2.2 Matching of Learning Styles
In the past century it has been made clear by numerous educationalists that something special happens in the interaction between teacher and learner. Bonnett [2002] for instance suggests that teaching and learning should include a genuine dialogue and that teacher and learner are both required to ‘invest something of themselves’ in learning, “which results in personal fulfillment and genuine receptivity’. Biesta [2006] when discussing communication in the teaching and learning process argues that ‘it is because people share in a common activity, that their ideas and emotions are transformed as a result of and in function of the activity in which they participate’. Montgomery & Groat [1998] also highlight the importance of the interaction between teacher and learner in a ‘collective dialogue’, while Kop [2009] speaks of ‘the dance between tutor and learner’. This implies that communication between them is important, but also their personalities, or their styles. Enough has been written about learning styles, but to define teaching styles is not as easy. Grasha discusses what teaching styles might be.

He says:

“Our teaching style represents those enduring personal qualities and behaviors that appear in how we conduct our classes. Thus, it is both something that defines us, that guides and defines our instructional processes, and that has effects on students and their ability to learn.” [Grasha 1996]

He stresses, however, that looking at individual traits in the interaction between learner and teacher would be ineffective for making any systematic change. He researched teaching styles and emphasized a number of influences that were seen to be important in the literature: ‘General modes of classroom behavior, characteristics associated with a popular instructor, the teaching methods employed, behaviors common to all college faculty, the roles teachers play, personality traits, archetypal forms, metaphors for teaching’ [Grasha 1996]; the latter ones of course related to theories of mind, knowledge and learning that have been developed and changed over the past decades [Kop 2010]. From these he derived five teaching styles and found through his research that most people don’t only have one teaching style, but have at least a secondary one, depending on context and circumstances in which they teach. The five teaching styles Grasha used:
1. The teacher as expert, who has all the knowledge and expertise that learners need;
2. The teacher as formal authority who doesn’t only have all knowledge and expertise required, but also has status that is related to this expertise and role in faculty.
3. The teacher as personal model. This style would entail the teacher leading by example, showing good behavior and positive examples of how to think, encouraging students to follow suit.
4. The teacher as facilitator involves a teacher who is the ‘guide on the side’ rather than the ‘sage on the stage’ and is more the helpful knowledgeable other, than the authority.
5. The teacher as delegator, who will support learners in their quest for autonomy and self-direction.
It is clear that to even start identifying the category in which an instructor were to fit, would require a certain level of reflection on his or her own practice. Reflection has been seen as an important activity at the heart of professional teaching, in order for teachers to rise above the role of a technician and move towards the role of a professional [Schon 2002] and to ensure an awareness of classroom and other contextual influences on their practice [Boud and Walker 2002].

2.3 Challenges of Developing a Technology Based Tool to Enhance the Learning Process and the Learning Experience
Grasha [1996] was well aware of situational influences on learning. He advocated the inclusion of teaching goals in the learning mix and saw three options for teachers to use the Grasha-Reichmann Learning Styles Scales (GRLSS): by either designing instructional processes to accommodate particular styles, or by designing them in such a way that creative mismatches would be created, or a combination of the two. This led him to a clustering of teaching styles with particular learning styles for particular circumstances [Grasha 1996]. These clusters form the basis for our recommendation system as they include some of the most important influences on the learning in a formal education environment, teacher and learner interactions, but in addition to that, the clustering offers the possibility for the adaptation of their use to particular contexts. Moreover, to add to ILD-RS’s applicability over a wide range of contexts, ‘field of study’ and ‘time duration’ were included in our algorithm.

3. INTELLIGENT LEARNING DESIGN RECOMMENDATION SYSTEM

3.2 Overview
ILD-RS is a software component able to recommend to the learning designer appropriate learning paths during the learning design phase: ILD-RS does not substitute but assist the learning designer. To do so, ILD-RS takes as input all available learning objects, the relevant field of study, the maximum duration of the learning path (in minutes), the learner’s grade, current competencies, desired competencies and learning style as represented by his or her scores on the GRLSS [Grasha 1996]. These inputs are fed into a prediction algorithm to produce appropriate recommendations. The prediction method uses a probabilistic approach based on Markov decision processes [Bellman 1957] to decide which learning path to recommend. For this purpose the prediction method exploits a transition matrix. This transition matrix is regularly updated according to the learning designs created by the users.

3.2 Markov Decision Process and ILD-RS
Markov decision processes (MDP) are widely used in operations research to study a wide range of optimization problems. MDP provides a formal framework to model decision making, such as the decisions a teacher would make in choosing and sequencing learning objects to constitute a learning path. By using MDP, the objective is to simulate some of the features of the teacher’s decision process in order to offer learning path recommendations which teachers are free to accept or decline while they are building their own learning activities.

A Markov decision process is a 4-tuple $< S, A, T, R >$ where:

- $S$ is a finite set of states,
- $A$ is a finite set of actions,
- $T: S \times A \rightarrow \text{Prob}(S)$ specifies a probability distribution for each state and action over next states,
- $R(s, a)$ is a reward function.

A MDP is assumed to respect the Markov property; the probability distribution $T$ is constituted of independent probabilities. In other words the probability of going from state $s_1$ to $s_2$ does not depend on previous states but only on the current state $s_1$. When a learner completes learning objects in a learning path, he/she goes from one learning object to another. By analogy with MDP, we have chosen to represent learning objects as states ($S$), the transition from one learning object to another as an action ($A$). $T$ represents the probability distribution of transitioning from one learning object to another. $T$ is represented by a transition matrix used in our prediction method. $R$ is the immediate reward obtained by transitioning from one learning object to another. Contrasting from
T that represents an average probability of transitioning from one learning object to another, R adds learner’s and teacher’s specificities required to take an accurate decision.

Our reward function $R$ is defined as:

$$R(s, a) = \frac{P(s, a)}{||TS(Teacher), TS(s')|| + ||LS(Learner), LS(s')||}$$

The learning object $s'$ is reached from $s$ after the transition $a$ $||TS(Teacher), TS(s')||$ is a distance factor between the teacher’s teaching styles and the learning object $s'$ teaching styles. Consequently, $||LS(User), LS(s')||$ represents a distance factor between the learning styles of a learner or a group of learners and the learning styles associated to the learning object $s'$. Usually in MDP, a policy $\pi$ is defined based on the reward function to help the decision maker, in our case our prediction method, to make the right decision. ILD-RS policy, $\pi$, is defined as follows:

$$\pi = \max \frac{\sum_{s=0}^{s_n} R(s, a)}{n}$$

As a first policy, we have chosen to implement a stationary policy. Our stationary policy does not change over time and aims to systematically choose the transition between two learning objects $s_0$ and $s_n$ offering the biggest reward while minimizing the number of visited states between initial and final states of the learning path.

According to our policy the chosen path between initial state A and final state D will consist of the learning objects ABCD (light grey path in Fig. 1) with a global reward of $7 \left( \frac{5+12+7}{4} \right)$.

### 3.3 Implementation

**Model objects implementation:**

Learning objects are represented with a custom XML format where learning objects are composed of a title, field of study, grade level, duration in minutes and teaching style of its corresponding teacher in the form of his or her results on the Grasha-Riechmann Teaching Styles Inventory [Grasha 1996]. Learning objects contain competencies, composed of a short textual description and a grade level. Competencies may also contain prerequisites, represented by other competencies. Fig. 2 shows an example of a simple learning object.

```xml
<learningObject title="FRACTIONS_MODULE_3" grade="6" duration="30" field="MATH" teacherStyles="4.8,4.3,2.5,2.2,1.9">
    <competencies>
        <competency description="IMPROPER FRACTIONS" grade="6">
            <prerequisites>
                <competency description="SIMPLE FRACTIONS" grade="6"/>
            </prerequisites>
        </competency>
    </competencies>
</learningObject>
```

Fig. 2. Example of learning object format
ILD-RS relies on a prediction algorithm supplemented with a training method in order to produce recommendations.

**Prediction method implementation**

ILD-RS begins by taking as input the relevant field of study, the maximum duration of the learning path (in minutes) and the learner’s grade, current competencies, desired competencies and student learning styles. Using the available learning objects, ILD-RS generates all possible learning paths (in the form of sequences of learning objects) leading from a set of initial competencies, corresponding to the students current competencies, to a point where the desired competencies have all been targeted. This is accomplished by first selecting only the learning objects of the correct field of study and grade level. Learning paths are generated by recursively examining these learning objects to verify that their prerequisites are met by the current competencies (the union of the student’s current competencies and those given by the learning objects in the current path), if so, they are added to the current learning path. This continues until the path contains all of the required competencies; however, if the path exceeds the maximum allowed duration, it is dropped.

The generated possible learning paths are examined to determine their probability of being of interest to the teacher. According to our MDP policy, the learning path offering the biggest cumulative reward and learning object number ratio will be selected. For each learning path, the potential reward is determined by examining every pair of consecutive learning objects within the learning path and determining the probability of successful learning for each transition between learning objects as well as the learning styles and teaching styles distance. The probability of successful learning for each transition is determined using a transition matrix \( T \) of dimension \( m \), where \( m \) is the number of known relevant learning objects and \( P(s_n|s_{n-1}) \) corresponds to the probability of the transition between learning object \( s_{n-1} \) and learning object \( s_n \) leading to successful learning. However, in this implementation, the values in our transition matrix do not actually represent probabilities but rather a score based on relative probability. The Markov property is still respected since the value associated to a transition between two states does not depend on the previous visited states as expressed in the following formula:

\[
P(s_n|s_{n-1}, ..., s_0) = P(s_n|s_{n-1})
\]

For each learning object, the student’s learning styles and the teachers teaching styles are compared in order to calculate the learning style and teaching style distance factor. For every object where the learning style and teaching style did not fall in compatible clusters [Grasha 1996], a “distance” value is increased by an adjustable value. The probability score of the learning path is divided by the learning and teaching styles “distance” and this resulting value is normalized by being divided by the number of learning object transitions within the learning path. This final value represents a global reward for the learning path as a whole. The learning paths with the highest rewards are then recommended. If the teacher decides to follow the recommendation, the recommendation is fed to the training method as a successful recommendation; otherwise, if the teacher does not follow the recommendation, his/her learning design is fed to the training method as an unsuccessful.

**Training method**

In the training method, the transition matrix is updated by adjusting the values of the confirmed paths’ learning object transitions, incrementing the values by an adjustable amount if the path was successful, otherwise decrementing them by a separate adjustable amount. Thus, the prediction method is continuously being adjusted based on results, implementing a form of reinforcement learning that is however far from the theory around reinforcement learning in MDP [Sutton and Barto 1998].

Once confirmed, learning paths can be exported in the proprietary XML format. ILD-RS can also output the learning paths in SLD 2.0 format [Durand et al. 2010].

So as to minimise impact from stale data, the transition matrix resets values that have not been used for a given number of recommendations (the “lifespan” of the values). If, after resetting
expired values, there is a learning object which has all of its transition values (both transition to and transition from) equal to the “neutral” value (that is to say, all values are equal to the initial or reset value), it is removed from the matrix entirely.

Many parameters, such as the initial, increment, decrement and lifespan values for the transition matrix, “distance” values and the number of learning paths to be recommended are all adjustable. As of yet, optimal values for these parameters have not been determined and they remain fully adjustable with arbitrary default values.

3.4 Related Work
Many approaches have emerged in the past few years to ease learners’ task of finding learning materials. Among those approaches, the application of data mining and recommendation technologies to provide learners with adequate learning materials is the most prominent one. Thus, by recommending educational papers [Tang and McCalla 2010], internet links [Godoy and Armandi 2010] or learning objects [Idris et al 2009] authors use collaborative and or content-based filtering approaches for their recommendation systems. There are two ways to build such a recommendation system. Usually, some profile clusters are discovered among a learner or content population (clustering) and new learners or content are classified (via classifiers) among those clusters. Depending on the classification, learners receive recommendations based on the resources completed by other learners sharing the same cluster and/or with contents similar to those they have already completed. Other approaches, especially in the course generation research community address the need for recommending not only learning objects but sequences of learning objects. Course generation has been studied for a while especially in the context of intelligent tutoring systems, where the ability to deliver individualized courses is a must have. Though numerous solutions have been proposed, using statistical methods [Karampiperis & Sampson 2005], decision rules [Vassileva & Deters 1998], and Hierarchical Task Network Planning [Ulrich 2005], it seems that very few have investigated the capacities of MDP, especially to exploit teachers’ knowledge to reinforce the accuracy of the predictions.

4. CONCLUSION AND FUTURE WORKS
MDP offers an interesting theoretical framework to think and design learning design recommendation systems. The rich MDP theory on policies and policy evaluation is definitely an asset in such applications where the choice of criteria and their use is still subject of controversy. In a further study, learning styles and teaching styles could be used directly at the policy level allowing policies using learning styles and teaching styles and policies not using those criteria to be evaluated. Learning styles and teaching styles are criteria among others whose use an MDP approach could help to discuss. The dense MDP policy evaluation research has provided the community with promising policy evaluation algorithms we are looking forward to explore.

From a learning design perspective, we are aware that ILD-RS only predicts sequences of learning objects and by doing so addresses a narrow part of learning design. This means that some complicated learning activities involving parallel and/or group sequences of learning object are not managed yet. As is, ILD-RS can only predict a sequence of learning objects to be performed by learners. This limit can have a pedagogical impact by limiting the use of specific instructional strategies. For example in the case of cooperative learning, where learners cooperate within an activity, the strategy has to be used within the learning object but cannot be explicitly designed in the learning design yet.

In learning design, learning activities are usually designed before being undertaken by learners. As such, the learning activity is static and if the activity has to change (or may benefit a change), the teacher has to redesign the activity. In order to help the teacher to “redefine it ‘on-the-fly’” [Jullien et al. 2009], another component is being designed to give teachers or autonomous learners the ability to offer learners a continuously calculated learning path. This new component would be comparable to a navigation system calculating an itinerary among available learning objects.

ACKNOWLEDGEMENT
We would like to thank our colleagues from National Research Council of Canada for their precious help with a special mention to Dr. Dan Tulpan.
REFERENCES
Carroll, F., Kop, R., AND Woodward, C., 2008. Sowing the Seeds of Learner Autonomy: Transforming the VLE into a Third Place through the use of Web 2.0 Tools, In proceedings 7th European Conference on e-Learning, University of Cyprus, Cyprus, pp. 152-160.
Kop, R., 2010. Networked Connectivity and Adult Learning: Social Media, the Knowledgeable Other and Distance Education, Ph.D. Thesis, Swansea University, UK.