ilmedia

th.
TECHNISCHE UNIVERSITÄT
ILMENAU

Knauf, Rainer; Sakurai, Yoshitaka; Tsuruta, Setsuo ; Takada, Kouhei; Dohi, Shinichi

Personalized curriculum composition by learner profile driven data mining

Personalized Curriculum Composition by Learner Profile Driven Data Mining

Rainer Knauf

Faculty of Computer Science and Automation
Ilmenau University of Technology
Ilmenau, Germany

Yoshitaka Sakurai, Setsuo Tsuruta, Kouhei Takada,
Shinichi Dohi

School of Information Environment
Tokyo Denki University
Chiba New Town, Japan

Abstract—The paper is focused on modeling, processing, evaluating and refining processes with humans involved like (not only, but also e-) learning. A formerly developed concept called storyboarding has been applied at Tokyo Denki University (TDU) to model the various ways to study at this university. Along with this storyboard, we developed a Data Mining Technology to estimate success chances of curricula. Here, we introduce a learner profiling concept that represents the students' individual properties, talents and preferences personalized data mining.[1]

Keywords—modeling learning processes, storyboarding, personalized data mining, learner modeling

I. INTRODUCTION

Generally, learning systems suffer from a lack of an explicit and adaptive didactic design. University education is especially effected by this lack, because university professors are not necessarily educational experts. One way of didactic support is providing a modeling concept for didactic design, which allows the anticipation of the learning processes.

Since e-learning systems are digital by their very nature, their introduction raises the issue of modeling didactic design in a way that offers the possibility to apply Knowledge Engineering (KE) techniques such as Machine Learning (ML) and Data Mining (DM). For this purpose, the modeling concept needs to be an at least semi-formal knowledge representation.

A modeling concept called storyboarding [1] has been developed formerly as a means of modeling learning processes. Besides providing didactic support, this semi-formal model is setting the stage to apply KE technologies to verify and validate the didactics behind a learning process. The verification may include both logical consistency issues and formally to check didactic issues.

A storyboard provides a road map for a lesson, a course, a subject to teach, or a complete study. According to different learning and teaching preferences, it includes alternative paths and possible detours if certain concepts to be learned need reinforcement. Using modern media technology, a storyboard

also plays the role of a server that provides the appropriate content material when deemed required.

By storyboarding, didactics can be refined according to revealed weaknesses and proven excellence. Successful didactic patterns can be explored by applying DM techniques to the various ways students went through a storyboard and their related success. As a result, future instructors and students may utilize these results by preferring those ways through a storyboard, which turned out to be the most promising ones.

In [2], a knowledge mining technology, which allows students to utilize mined "experience" of former students to compose curricula with an optimal success chance, is introduced.

However, individual learning plans should not only be based on the success of former students who went similar ways. Additionally, individual properties, talents and preferences should be considered. For example, some students are more talented for analytical challenges, some are more successful in creative or composing tasks, and others may have an extraordinary talent to memorize a lot of factual knowledge. Consequently, we need to include individual learner profiles to avoid lavishing the students with suggestions that don't match their individual preferences and talents.

The paper is organized as follows. Section 2 introduces the storyboard concept including the present state of the current development. Section 3 provides an overview on our Knowledge Mining technique to compose optimal curricula for university studies. In section 4, current approaches of adaptation trough user modeling are discussed and summarized. Section 5 introduces and discusses the learner model that we derived for our application. In section 6, we introduce a technology to use the derived learner model for a personalized curriculum composition. Finally, section 7 provides a summary and outlook.

II. STORYBOARDING

Our storyboard concept was introduced in [1] und later refined (see [2] for the latest version). A storyboard is a nested hierarchy of directed graphs with annotated nodes and annotated edges. Nodes are *scenes* or *episodes*. *Scenes* are not further structured, *episodes* have a sub-graph as its implementation. Also, there is exactly one *start node* and one *end node* in each (sub-) graph. *Edges* specify transitions

between nodes and may be single-color or bi-color. Nodes and edges can carry attributes.

A storyboard may be seen as a model of an anticipated reception process that is interpreted as follows:

Scenes denote a non-decomposable learning activity that can be implemented in any way, e.g. by the presentation of a (media) document, opening a tool that supports learning (URL or e-learning system) or an informal activity description. *Episodes* are defined by their sub-graph.

Graphs are interpreted by the paths, on which they can be *traversed*. A *Start Node* of a (sub-) graph defines the starting point of a legal graph traversing. An *End Node* of a (sub-) graph defines the final target point of a legal graph traversing.

Edges denote transitions between nodes. There are rules to leave a node by an outgoing edge: (1) The outgoing edge must have the same color as the incoming edge by which the node was reached. (2) If there is a condition specified as the edge's key attribute, this condition has to be met for leaving the node by this edge. So the colors express the dependence of ways leaving a node from the way of arriving there.

Key attributes of nodes specify application driven information, which is necessary for all nodes of the same type, e.g. actors and locations. *Key attributes of edges* specify conditions, which have to be true for traversing on this edge. *Free attributes* specify whatever the storyboard author wants the user to know: didactic intentions, useful methods, necessary equipment, e.g.

For further information, the reader may see [2] or [3], e.g.

III. Curriculum Composition by Data Mining

A basic objective of storyboarding is to use Knowledge Engineering technologies on the (semi-) formal process models [3]. In particular, we aim at inductively "learning" successful storyboard patterns and recommendable paths. This is some sort of meta-learning, i.e. the learning of learning knowledge. It is performed by an analysis of the paths where former students went through the storyboard [2].

To show the feasibility and benefit of high level storyboarding for its qualified assistance of students suffering from the "jungle of opportunities and constraints" in university education, a simple prototype was developed to evaluate curricula created or modified by the students in advance of their study [3][2]. For this purpose, we introduced a concept to estimate success chances of curricula, which are composed by students at the School of Information Environment (SIE) of the Tokyo Denki University (TDU) in their curriculum planning class in the first semester.

For such curricula we developed a DM technique, which is applied to storyboard paths that (former) students went. Based on these examples, the success chance of intended paths can be estimated [2]. The technique consists in two steps, (1) constructing a decision from the examples of former students and (2) applying this decision tree to the planned curricula. It is described more detailed in [2].

However, individual learning plans should not only be based on individual quantitative capability (like the Grade Point Average at TDU) or the success of former students who went similar ways. Additionally, individual properties, talents and preferences should be considered. For example, some students are more talented for analytical challenges, some are more successful in creative or composing tasks, and others may have an extraordinary talent to memorize a lot of factual knowledge. Therefore, we included individual learner profiles to avoid lavishing the students with suggestions that don't match their individual preferences and talents.

IV. Lerner Modeling Approaches

There are several general approaches in the literature, which can be classified according to the following subsections.

A. Modeling a Competence State

In *Overlay models*, the learner's knowledge L is considered as a subset of the teacher's knowledge T. The difference $T - L$ is considered as the learner's lack of knowledge. The learning objective is T. There is no mechanism to differentiate between the knowledge the learner has not yet grasped and the knowledge the student has not yet been exposed. It does not employ strategies to help in case of misconceptions.

Differential models are a more structured variant of the latter. They separate the entire domain knowledge into learned by the student $(T \cap L)$ and not-learned by the student $(T - L)$.

Perturbation models are refined derivate of both, which includes misconceptions ("known incorrectly") or errors.

Constrained-based models represent both the domain and the learner's knowledge as a set of constraints. Constraints represent basic concepts or rules of the domain. Constraints violated by the learner are recorded.

However, all *competence state models* request an explicit representation of the topical subject knowledge. Therefore, they are not appropriate for universal models for managing learning activities.

B. Modeling the Learning Process

Models of this class represent students' beliefs in conceptual graphs. Graph reasoning algorithms identify a belief b out of a pre-defined set $B = \{b_1, b_2, ..., b_n\}$. These models can be used to predict the kind of mistakes the student will/is likely to make. These models also need a representation of domain knowledge and thus, are not appropriate for our purpose.

C. Modeling Performance Traits

A step away from representing a concrete domain towards universality is a classification of domains into classes of skills, which are challenged, when learning the particular domain knowledge. This still requires some domain analysis, but no explicit representation of the knowledge to be learnt. However, we will show later that this "analyses" can be mined, too.

The most famous model of this class is Gardner's theory of Multiple Intelligences [5][2]. According to Gardner, a learner's intelligence cannot be considered as a whole (and rated by a single IQ), but in various dimensions, namely (1) *Linguistic intelligence*, (2) *Logical-mathematical intelligence*, (3) *Musical intelligence*, (4) *Bodily-kinesthetic intelligence*, (5) *Spatial*

---

[2] The reader may enjoy visiting (accessed on March 17, 2009) *http://www.businessballs.com/howardgardnermultipleintelligences.htm*

(visual, creative-artistic) intelligence, (6) *Interpersonal (social) intelligence*, and (7) *Intrapersonal (emotional) intelligence*.

D.   Modeling Learning Styles

Modeling learning styles has the advantage not to depend on the domain and is, therefore, promising for our purpose. The most researched theory within this category is the Felder-Silverman Learning Style model [6][3] as illustrated in **Table 1**.

Table 1: Felder-Silverman Learning Styles

Active	Reflective
• *Trying something out* • *Social oriented*	• *Think about material* • *Impersonal oriented*
Sensing	**Intuitive**
• *Existing ways* • *Concrete material* • *Careful with details*	• *New ways* • *Abstract material* • *Not careful with details*
Visual	**Verbal**
• *pictures*	• *spoken words* • *written words* • *difficulty with visual style*
Sequential	**Global**
• *detail oriented* • *sequential progress* • *from parts to the whole*	• *overall picture* • *non-sequential progress* • *relations/connections*

Each learner is characterized by a specific preference for each of these four dimensions:

- Abstraction: *Sensory* learners prefer concrete, practical, and procedural information. They look for the facts. *Intuitive* learners prefer conceptual, innovative, and theoretical information. They look for the meaning.

- Perception: *Visual* learners prefer graphs, pictures, and diagrams. They look for visual representations of information. *Verbal* learners prefer to hear or read information. They look for explanations with words.

- Inference: *Active* learners prefer to manipulate objects, do physical experiments, and learn by trying. They enjoy working in groups to figure out problems. This is a more inductive way to learn. *Reflective* learners prefer to think things through, to evaluate options, and learn by analysis. They enjoy figuring out a problem on their own. This is a more deductive way to learn.

- Perspective: *Sequential* learners prefer to have information presented linearly and in an orderly manner. They put together the details in order to understand the big picture emerges. *Global* learners prefer a holistic and systematic approach. They see the big picture first and then fill in the details.

E.   Modeling Cognitive Traits

Recent modeling approaches go one step closer to the genesis of learning behavior and model cognitive traits such as

---

[3] One may check his learning style at (accessed on March 17, 2009) http://www.engr.ncsu.edu/learningstyles/ilsweb.html

done in [7]. This approach also enjoys the advantage to be domain independent. [7] introduces a Cognitive Trait Model (CTM), which includes three items, namely (1) *Working Memory Capacity* (WMC), (2) *Inductive Reasoning Ability* (IRA), and (3) *Divergent Associative Learning Ability* (DALA). From a more technical point of view,

- WMC is comparable with a *Random Access Memory* (RAM), but human working memory is not just a memory, but has some processing abilities as well,

- IRA refers to *Inductive Inference* respectively *Data Mining*, if one refrains from 100% hypotheses consistency with the data, and

- DALA refers to *Analog Inference* ("associative") respectively *Case Based Reasoning* (CBR) with taking into account differences ("divergent").

For these traits, the domain-independence and persistence over time is shown in [7]. Furthermore, behavior patterns are described, which indicate the particular trait, and they are gathered in a learning system as well as by web-based psychometric tools. Also, [7] derives so called Manifestations of Traits (MOT), which are typical (not only learning) behavior patterns. From these MOTs, five typical behavior patterns (implementation patterns) are derived, which can be identified by an analysis of the learners interaction with the machine. As an example, **Table 2** shows the implementation patterns for WMC.

Table 2: Implementation Patterns for WMC

	Working Memory Capacity	
	low	**high**
1	non-linear navigational pattern	linear navigational pattern
2	constant reverse navigation	rare (or none) reverse navigation
3	unable to learn from excursions	able to learn from excursions
4	unable to perform simultaneous tasks	able to perform simultaneous tasks
5	inefficient retrieval from long term memory	efficient retrieval from long term memory

F.   Mixed Modeling Approaches

There are also commonly used approaches which are of a mixed type such as the IMS Learner Information (LIP)[4] and the model of Mödritscher and Brusilovski[5], which suffer from both the need of domain representation and containing individual learner information that is difficult to appraise in our application.

V.   DERIVED LERNER MODEL

Performance based models bag the risk, that the one or other performance is a result of hard training, but not a natural talent. Vice versa, a bad performance in a particular task does not necessarily mean, the performer is generally not able. It

---

[4] See http://www.imsglobal.org/profiles/lipinfo01.html, accessed on February 21, 2009
[5] See http://www.moedritscher.com/papers/paper_moedritscher_et_al_adaptivestandard_2004a.pdf, accessed on February 19, 2009

may also have its reasons in being too tired, being mentally occupied by anything else, or just being not interested.

However, the integration into our application (which is described below) heavily depends on a profile acquisition by a questionnaire test by most successful students in each subject. In [7] the development of online tests is reported. However, we did not have the opportunity so far to inspect these tests and to check, whether they are appropriate for our application scenario. Also, we are not sure, whether these web-based tools are available over the long time of experimentation.

With respect to the requirement to be domain independent for use in universal learning systems, learning style models also seem appropriate for our application. Learning styles are exactly the point we aim at, when modeling learning processes.

On the other hand, our application (curriculum planning) seems to be more driven by competence traits, which are domain-related, but don't need an explicit representation of the domain knowledge.

We have a tendency to assume that

- learning styles need to be focused for "fine grained" learning processes (a lecture, a course) and

- competence trait models are more appropriate if higher level (coarse grained) learning processes (such as a complete study) modeling.

So far, we consider both in our model, which is defined as an array of 11 attribute-value pairs that contains 7 intelligence attributes and 4 learning style attributes. Both can be appraised by questionnaires that are available to the public in the web (see footnotes 2 and 3).

To make the dimensions of both sources comparable to each other and see the quantitative relations, we normalized them in a way that they all have the same range of values. The intelligence dimensions rage from 10 to 40. The learning style dimensions range from -11 to +11 (opposite algebraic sign for opposite styles). The normalization can be done by

- $v = result/40$ for the intelligence dimensions and

- $v=(result+11)/22$ for the learning style dimensions.

Finally, our learner model looks as shown in **Table 3**.

Table 3: Derived learner profile

d_1	Linguistic intelligence	$0 \leq v_1 \leq 1$
d_2	Logical-mathematical intelligence	$0 \leq v_2 \leq 1$
d_3	Musical intelligence	$0 \leq v_3 \leq 1$
d_4	Bodily-kinesthetic intelligence	$0 \leq v_4 \leq 1$
d_5	Spatial intelligence	$0 \leq v_5 \leq 1$
d_6	Interpersonal intelligence	$0 \leq v_6 \leq 1$
d_7	Intrapersonal intelligence	$0 \leq v_7 \leq 1$
d_8	Active vs. Reflective style	$0 \leq v_8 \leq 1$
d_9	Sensing vs. Intuitive style	$0 \leq v_9 \leq 1$
d_{10}	Visual vs. Verbal style	$0 \leq v_{10} \leq 1$
d_{11}	Sequential vs. Global style	$0 \leq v_{11} \leq 1$

One could argue the use of explicit questioning techniques is not appropriate. They bag the risk of cultural bias, acquiescent responses, and inter-group bias.

In fact, a trace of the learners' interaction is more reliable than questionnaire results, because it is driven by only the learners' interest in learning and it is limited to exactly what we want to uncover, the learners' cognitive capacities and preferences.

In our application, we just have the trace of students through the storyboard, i.e. the students' paths along with their performance level in each subject. Indeed, performance levels in subject don't say anything about learning preferences. The only way to derive learning related properties from it is attaching "guessed" properties of successful learners to each subject.

There are subject-related successful students, which have a high ability in understanding the domain of this particular subject. This indicates that their traits are appropriate for understanding especially this subject in especially the way this subject is taught by its particular teacher.

The assumption is that there is a link between

- typical "competence traits" (according to Gardener) and subjects that typically challenge the one or other "kind of intelligence" more than others and

- typical teaching methods (according to Felder and Silverman) and subjects that are typically taught with these methods.

Learners with a high "logic-mathematical intelligence" may typically be top in Mathematics and Physics, learners with a high "spatial (visual, creative-artistic) intelligence" may be good in Geometry and Civil Engineering, learners with a high "interpersonal intelligence" may be good in Social Science and so on. Accordingly, learners with a more visual learning style may be top in subjects, for which picture-like material is predominantly, learners with a more active learning style may be top in subjects, in which teamwork is predominant and so on. We do not presume to appraise these links, because this was quite arbitrary. Therefore, we developed the idea to "mine" this link by considering the successful students in each subject.

How to "guess" profiles of students and link it to subjects? There is no way to ask all students to reveal individual data by questionnaires. In particular, students, who have something to "hide" (a bad performance), will not cooperate. However, students with a very excellent performance in a certain subject, who are available (i.e. still at the university) may be cooperative.

So we could ask only the very successful students of each subject with maximum grade points to fill in the questionnaires on the multiple intelligences distribution and the learning styles as mentioned above on a voluntary basis.

With their data, we could look, whether there is a correlation between learning success in a particular subject and the profile of its successful learners. If so, a typical "success profile" $P(s)$ for each subject s can be estimated by the average profile of its most successful learners $L(s):=\{l:$ learner l received maximum Grade Points in subject s and provided

questionnaire answers*} with individual successful learners' $l^i \in L(s)$, $i=1...|L(s)|$ profiles according to the above profile definition as an 11-dimensional array $p(l^i)=[d_1^i, d_2^i, ..., d_{11}^i]$:

$$p(s) = \frac{1}{|L(s)|} \begin{bmatrix} \sum_{i=1}^{|L(s)|} d_1^i \\ \sum_{i=1}^{|L(s)|} d_2^i \\ \vdots \\ \sum_{i=1}^{|L(s)|} d_{11}^i \end{bmatrix}$$

This calculation has to be done for each subject separately and the set of "most successful students" differs from subject to subject, of course. The idea behind is to mine a "typical success profile" for each subject separately.

All (other) learners' $l^i \notin L(s)$ profiles $p(l^i)=[d_1^i, d_2^i, ..., d_{11}^i]$ are estimated as a (success-) weighted average value of each profile dimension over all subjects the student took so far. The weight factor is the success in the related subject; it should be 1 for subjects with best marks and 0 in subjects in which the student failed. Let

- $S^i = \{s^1, s^2, ..., s^m\}$ be the set of subjects the learner l^i took so far,

- $succ_j^i$ be the success degree of learner l^i in subject s^j with

$$succ_j^i = \begin{cases} 1.00, \text{if learner } i \text{ received in subject } j \text{ mark S} \\ 0.80, \text{if learner } i \text{ received in subject } j \text{ mark A} \\ 0.60, \text{if learner } i \text{ received in subject } j \text{ mark B} \\ 0.40, \text{if learner } i \text{ received in subject } j \text{ mark C} \\ 0.20, \text{if learner } i \text{ received in subject } j \text{ mark D} \\ 0.00, \text{if learner } i \text{ received in subject } j \text{ mark E} \end{cases}$$

- $p(s^j)=[d_1^j, d_2^j, ..., d_{11}^j]$ be the profile of the subject s^j mined from its most successful learners as described above.

Then, a learner's profile $p(l^i)=[d_1^i, d_2^i, ..., d_{11}^i]$ can be estimated (mined) from the subject profiles $p(s^j)=[d_1^j, d_2^j, ..., d_{11}^j]$ of subjects that the learner took ($s^j \in S^i$) by

$$p(l^i) := \frac{1}{\sum_{j=1}^{|S^i|} succ_j^i} \begin{bmatrix} \sum_{j=1}^{|S^i|} (succ_j^i * d_1^j) \\ \sum_{j=1}^{|S^i|} (succ_j^i * d_2^j) \\ \vdots \\ \sum_{j=1}^{|S^i|} (succ_j^i * d_{11}^j) \end{bmatrix}$$

Finally, we have a profile $p(l^i)=[d_1^i, d_2^i, ..., d_{11}^i]$ for each student and a profile for each subject $p(s^j)=[d_1^j, d_2^j, ..., d_{11}^j]$. The latter one is related to its most successful learners.

The entire approach is a mixed technology, i.e. it uses both questionnaires for the students with very significant data (most successful students) students traces for all other students. The very best students of each subject are very significant for our technology, because their profile is obviously very appropriate for a particular subject taught and examined in a particular way by its particular teacher. Therefore, they form the typical "success" profile of this subject.

VI. INTEGRATION OF THE MODEL

First of all, we have to establish a cognitive profile for each involved learner by partly the questionnaires for a few of them and the computation of their analysis result for all other learners as explained above.

The general way to evaluate a submitted curriculum plan as briefly outlined in section 3 is unchanged. However, the evaluation of a submitted curriculum and the suggestion to optimize it is done only with those (former) students (whose paths are represented in the decision tree), which have a *similar* learner profile.

What is "similarity", in our application setting? Students have a "similar" profile, if they have similar talents and similar weaknesses. In other words, their profiles are similar, if the quantitative relations in-between the profile attributes are about the same.

Technically, two profiles are similar, if their (11-dimensional) vectors point in almost the same direction within the 11-dimensional Euclidean Space. For easy comparison of similarities, it should be just a scalar value.

There are several approaches to determine a scalar similarity between two vectors $X = [x_1, x_2, ..., x_n]$ and $Y = [y_1, y_2, ..., y_n]$:

(1) Cosine Coefficient $\dfrac{\sum_{k=1}^{n} x_k y_k}{\sqrt{\sum_{k=1}^{n} x_k^2} * \sqrt{\sum_{k=1}^{n} y_k^2}}$

(2) Dice-Coefficient $\dfrac{2 * \sum_{k=1}^{n} w_{xk} * w_{yk}}{\sum_{k=1}^{n} w_{xk} * \sum_{k=1}^{n} w_{yk}}$

(3) Jaccard-Coefficient $\dfrac{\sum_{k=1}^{n} w_{xk} * w_{yk}}{\sum_{k=1}^{n} w_{xk} + \sum_{k=1}^{n} w_{yk} - \sum_{k=1}^{n} w_{xk} * w_{yk}}$

(4) Overlap-Coefficient $\dfrac{\sum_{k=1}^{n} \min(w_{xk}, w_{yk})}{\min(\sum_{k=1}^{n} w_{xk}, \sum_{k=1}^{n} w_{yk})}$

For (2), (3), and (4) $w_{zk} = \begin{cases} z_k, \text{if } z_k \neq 0 \\ 0, \text{otherwise} \end{cases}$.

The Dice – Coefficient and Jaccard – Coefficient measure are not appropriate for our application for two reasons. (1) They consider non-zero values only. In our application, zero values can occur and have a meaning, too. (2) Similarities are not comparable, because the value range of the similarity depends on the values in the vectors. In our application, we want to compute a number of "most similar" profiles. Thus, similarities must be comparable with each other.

The Overlap-Coefficient is more or less a counting of dimensions, in which one vector has a lower value than the other one. Also, it does not consider the degree of the

2072

difference. This Coefficient may be good for statements about the overall Intelligence, but not for comparing the more detailed complete profiles.

Finally, the Cosine - Coefficient meets our requirements. Similarity of profiles should mean similar relations in-between its components. This is exactly, what the Cosine Coefficient measures, the cosine of the angel between the vectors, which is 1 for identical vectors and zero for orthogonal ones.

To include the profiles and their similarities in the curriculum evaluation (see section 3), we modify the curriculum evaluation procedure as follows.

The procedures to construct a decision tree and to use this tree for curriculum evaluation stay unchanged. The only difference is that we construct the decision tree exclusively from learners with profiles that are most similar to the one under evaluation.

How to compose the subset $S^{sim} \subseteq S$ most similar students?

- One way is to state a trigger value of similarity sim^{min}, at and below which all paths belong to S^{sim}:

$$S^{sim} := \{ p: sim(p, p^{eval}) \leq sim^{min} \}$$

From our point of view, this is a little arbitrary. Such a value is difficult to determine. Maybe, after gaining sufficient data from practice, this may be appropriate. So far it is not.

- A way around the above problem is to use the *k-nearest neighbor method* (k-NN method) and state a number *k* of students, who's paths are most similar to the submitted path:

$$|S^{sim}| = k$$

$$\forall p^{sim} \in S^{sim} \ \forall p \in (S \setminus S^{sim}): sim(p^{sim}, p^{eval}) \leq sim(p, p^{eval})$$

This is quite arbitrary, too. Also, determining a reasonable value for such number is at best possible after having a sufficient amount of data, too. Additionally, this approach does not work, it there are too few students, who's paths form the decision tree.

- A solution that avoids all above mentioned drawbacks, is to state a portion (a percentage *prc*) who's paths are most similar to the submitted one:

$$|S^{sim}| = prc/100 * |S|$$

$$\forall p^{sim} \in S^{sim} \ \forall p \in (S \setminus S^{sim}): sim(p^{sim}, p^{eval}) \leq sim(p, p^{eval})$$

By doing the latter, both of the following performance features are guaranteed:

(1) The estimation of success chances is based on individual preferences, talents, and weaknesses.

(2) The suggestion of a remaining learning path (subjects recommended to optimize the success of study) is adapted to individual properties, because it is calculated on the base of examples with a similar profile.

VII. SUMMARY AND OUTLOOK

The research reported here is focused on modeling, processing, evaluating and refining processes with humans involved like (not only, but also e-) learning. A formerly developed concept called storyboarding is briefly introduced.

Along with a storyboard application, we developed a technology to estimate success chances of curricula, which are composed by students.

However, individual learning plans should not only be based on individual quantitative capability or the success of former students who went similar ways. Additionally, individual properties, talents and preferences should be considered. Here, we addressed this point by introducing into this technology.

Extensions of this technology by individual contexts like (1) the educational history before entering the university, (2) career plans, i.e. a desired kind of position after the university study, and (3) social issues such as family status, residential area, and so on, are under development Context Based Reasoning (CxBR) Technologies are considered to derive optimal curricula.

REFERENCES

[1] K.P. Jantke and R. Knauf, "Didactic design though storyboarding: Standard concepts for standard tools", 4th International Symposium on Information and Communication Technologies, Workshop on Dissemination of e-Learning Technologies and Applications, Cape Town, South Africa, 2005, New York: ACM Press, ISBN 0-9544145-6-X, pp. 20-25.

[2] R. Knauf, R. Böck, Y. Sakurai, S. Dohi, and S. Tsuruta, "Knowledge mining for supporting learning processes", The 2008 IEEE International Conference on Systems, Man, and Cybernetics (SMC 2008), Singapore, 2008, IEEE Catalog number CFP08SMC-USB, ISBN 978-1-4244-2384-2, Library of Congress: 2008903109, IEEE, Piscataway, NJ, USA, pp. 2615-2621.

[3] Y. Sakurai, S. Dohi, S. Tsuruta, R. Knauf, R., "Modeling Academic Education Processes by Dynamic Storyboarding", in press, Journal of Educational Technology & Society, vol. 12, ISSN 1436-4522 (online) and 1176-3647 (print). International Forum of Educational Technology & Society (IFETS), 2009.

[4] R. Knauf, Y. Sakurai, S. Tsuruta, "Applying Knowledge Engineering Methods to Didactic Knowledge - First Steps towards an Ultimate Goal", Proc. of 2008 IEEE World Congress on Computational Intelligence (WCCI 2008), International Joint Conference on Neuronal Networks (IJCNN 2008), Special Session on Computational Intelligence Approaches for E-Learning, Hong Kong, IEEE Catalog Number CFP08IJS-CDR, ISBN 978-1-4244-1821-1, ISSN 1098-7576, pp. 38-45.

[5] H. Gardner, "Frames of Mind: The Theory of Multiple Intelligences", Basic Books, 1993.

[6] R.M. Felder, L.K. Silverman, "Learning and teaching styles in engineering education", Engineering Education, 78(7), 1988, pp. 674–681.

[7] T.-Y. Lin, "Cognitive Trait Model for Adaptive Learning Environments", Ph.D. Tesis, Massey University, Palmerston North, New Zealand, 2007.