Automated Acquisition of Domain Model for Adaptive Collaborative Web-Based Learning

Marián Šimko*
Institute of Informatics and Software Engineering
Faculty of Informatics and Information Technologies
Slovak University of Technology in Bratislava
Ilkovičova 3, 842 16 Bratislava, Slovakia
simko@fiit.stuba.sk

Abstract
Adaptive features of web-based educational systems bring substantial advantages for learners. Educational system instruction is adopted to their individual needs resulting in more efficient learning. Adaptation engine responsible for advanced functionality in the educational system relies on the domain model semantically describing subject domain. Quantity of domain descriptions and amount of user-generated data make manual creation and maintenance of semantic descriptions demanding and almost impossible task. In our work we tackled a problem of domain model automated acquisition. We proposed lightweight domain model with respect to acquisition facilitation, together with methods for acquisition of domain model respective parts. We advocate lightweight semantics as suitable form of domain conceptualization for adaptive educational system. In order to evaluate the proposed approach we performed several experiments in the domain of learning programming. We also present case study of lightweight domain modeling in Adaptive LEarning Framework ALEF, which was developed at Slovak University of Technology in Bratislava to support learning for technology oriented courses.

Categories and Subject Descriptors
D.2.13 [**Software Engineering**]: Reusable Software—*domain engineering, reuse models*; H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing—*linguistic processing*; I.2.6 [**Artificial Intelligence**]: Learning—*knowledge acquisition, concept learning*; K.3.1 [**Computers and Education**]: Computer Uses in Education—*collaborative learning, computer-assisted instruction (CAI)*

*Recommended by thesis supervisor: Prof. Mária Bieliková. Defended at Faculty of Informatics and Information Technologies, Slovak University of Technology in Bratislava on March 30, 2012.

Šimko, M. Automated Acquisition of Domain Model for Adaptive Collaborative Web-Based Learning. Information Sciences and Technologies Bulletin of the ACM Slovakia, Vol. 4, No. 2 (2012) 1-9

Keywords
domain model, web-based learning, adaptive learning, collaborative learning, knowledge representation, knowledge acquisition, text mining

1. Introduction
We witness tremendous increase of information overload with wide spread of the Web. It negatively affects our cognitive abilities and effectiveness when solving various tasks. An important approach to overcome this issue is adaptation, which reduces amount of delivered information and tailors it to fit the needs of a particular user. This applies particularly for web-based learning, where adaptive behavior of an educational system results in more effective learning [1, 24].

Web-based learning is also affected by social- and collaboration-oriented Web 2.0 features such as tagging, rating or creating own content, which provide a learner with more competences. Learning experience in such system differs and the learner is active participant in the learning process rather than passive consumer of delivered instructional units [8].

Adaptivity in an educational system is realized by an adaptive engine, which utilizes semantic descriptions of educational content in order to make adaptation decisions. Adaptation adequacy depends on the quality of semantic descriptions (referred to as metadata; often in a form of concepts and relationships), residing in a domain model of a course. Due to the amount of necessary descriptions, metadata initial creation and maintenance is an extremely difficult task to accomplish manually. The problem is even worse when considering so-called adaptive web-based learning 2.0 [19]. Here user-generated content to be adapted has to be assigned necessary semantics as well, hence making the task of metadata creation and maintenance impossible for a human.

In our work we dealt with the problem of automated creation of domain model for adaptive web-based collaborative educational systems. The state-of-the-art approaches to domain modeling for adaptive web-based educational systems suffer from obsolete domain model design . There is no clear distinction between domain conceptualization and the content of resources being presented in the adaptive web-based system. Resources and domain model are often tightly coupled that results into the inflexible and limited adaptation. There also is no explicit support for

"2.0" features in learning; the increasing importance of a user (as one of central concepts of Web 2.0 paradigm) is still not sufficiently reflected into the formal description of adaptive web-based system's domain model. In addition, in the field of adaptive web-based systems there is a limited number of approaches and methods to automatic creation of domain model or its respective parts.

We aimed to address these problems by supporting automated domain model creation for web-based adaptive social educational systems. In particular, thesis goals were:

- to design and evaluate a domain model addressing the drawbacks of the state-of-the-art approaches particularly focusing on:
  - clear separation between domain conceptualization and content,
  - explicit support for modeling "2.0"-based activities such as tagging, annotating, rating, commenting or different forms of collaboration,
  - the possibility to automate certain domain model parts creation (while preserving flexible and adequate adaptation).
- to design and evaluate a method for automated domain model creation:
  - leveraging selected methods and techniques of data mining as well as specifics of educational content,
  - by considering heterogeneous sources of information.

Particularly challenging is development of a proper methodology for domain model evaluation. A "direct" quantitative evaluation is difficult as it is hard to define exact measures that are suitable to assess quality of a domain model for adaptive web-based learning. On the other hand, evaluation involving real-world users and observing impact of the proposed domain model-based adaptation on learning performance is expensive in terms of effort, time and even finance.

2. Related Work

Domain modeling in adaptive web-based systems can be conceptually described in a form of reference models or their parts. The aim of a reference model is to describe an adaptive web-based system at an abstract level by identifying logical components and relations within. Adaptive Hypermedia Application Model (AHAM) aims to separate three concerns of adaptive hypermedia: domain, user and adaptation [25]. When defining the domain model, it builds on a layer viewing "concepts and concept relationship as a generalization of nodes and links". AHAM treats concepts and their content as hypertext nodes and content of such nodes. The notion of concept is interwoven with the content, which is created to represent the concept.

Munich Reference Model is a reference model for adaptive hypermedia applications specified by UML [13]. It is similar to AHAM. Although being described more formally, it still does not completely handle the separation of conceptual and presentational aspect of adaptive application. The LAOS reference model defines two models related to the subject domain: domain model itself and goal and constraint model [5]. The domain model is based on a conceptual map with corresponding operations designed to keep the semantics of course authoring standardized. LAOS's biggest contribution is introducing the goal and constraint model. Although goals were separated from the domain conceptualization, the domain conceptualization itself remained coupled with the content.

State-of-the-art approaches to domain modeling of adaptive web-based systems can be characterized by no genuine separation of concerns and limited explicit support for the social aspects [7].

There is only a small number of approaches to automatic metadata acquisition in adaptive web-based systems we are aware of. The authors of My Online Teacher (MOT), an adaptive web-based system based on reference model LAOS, developed a method for computing similarity between concepts by calculating correspondence weights between concepts attributes [6]. The idea is based on cooccurrence comparison of keywords representing one concept and overall attribute contents of the other concept.

Sosnovsky et al. aim at automated prerequisite and outcome relationships identification [17]. Based on predefined concept pattern detection, they extract concepts from learning objects on C programming language. Considering assignments of concepts to learning objects (programming examples) and ordering of learning objects, they employ simple heuristics to extract prerequisite/outcome relationship between concepts. An interesting example of automated metadata acquisition was performed in the case of adaptive vocabulary acquisition system ELDIT [2], where methods and techniques of natural language processing were employment in order to create relationships between examples of vocabulary entries and vocabulary entries. Relationships are created by performing lexical analysis of underlying text content – a core content of vocabulary already containing examples with single links to vocabulary entries (word definitions) – and comparing it with all vocabulary entries.

None of the described approaches to metadata acquisition is automatic. Although performing machine-supported intelligent processing of the content, many actions must be done by a content author manually (creating attributes, sequencing learning objects, providing initial set of relationship).

There is the overall lack of approaches to knowledge discovery in the domain of adaptive web, however, there are are works from other areas, which relate here.

The accuracy of state-of-the-art approaches to metadata acquisition based on content processing decreases along with a complexity of a task. While accuracy of concept extraction ranges around 80 to 90 % (e.g., [9, 12]), accuracy of hierarchical relationship discovery decreases to 40-60 % (e.g., [3, 10]). Accuracy of non-hierarchical relationships discovery is even lower: 10-20 % (e.g., [15]). It is important to note that there is no uniform methodology involved. Thus, an accuracy of metadata acquisition methods cannot be quantitatively compared. In addition, many approaches use different data sets and different measures, making comparison even more complicated. The presented numbers represent rather an approximated overview of methods' accuracy and reveal to what extent results differ from expectations defined by domain experts.

3. Lightweight Domain Modeling for Adaptive Collaborative Web-Based Learning

We can view an abstraction of the core concept of the adaptive educational system resulting from Web 2.0-induced paradigm shift in learning as depicted on Figure 1. Users (both teachers and students) do not only access the resources (e.g., learning content, annotations), but they are also able to modify them or even create new content. Resources are described with (conceptual) metadata, which are utilized to deliver advanced features supporting learning, interaction and collaboration in the educational system.

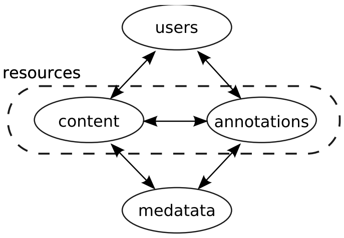

Figure 1: An abstract view of the core concept of adaptive web-based learning 2.0: users create and access resources that are described by metadata.

The core responsibilities of a domain model in an adaptive web-based system reside in forming a base for user model and serving adaptation engine for making decisions. This applies for adaptive educational web-based systems as well.

User model in majority of state-of-the-art overlays a domain model and, basically, it estimates knowledge of a user related to particular domain model knowledge elements (metadata). Based on the user model, an adaptation engine reviews current user knowledge and based on the domain model it derives a set of domain knowledge elements that suits users needs. An adapted form of learning experience is delivered to a user. The core responsibilities point to the fact that it is important to have a domain described in an adequate detail as the accuracy of users' knowledge estimation reflects in quality of adaptation.

Besides providing fundamental functionality, our aim is to cope with the current challenges in adaptive social web-based learning 2.0 by designing a domain model for adaptive educational systems that meets the requirements of emerging social collaborative learning related to both learning flow and collaborating/creating flow of activities, while still addressing "legacy" issues.

3.1 Lightweight Domain Model Overview

We leverage so-called lightweight semantics and proposed lightweight domain model for adaptive web-based educational system. The basic idea is that resources are described by terms that are relevant for a domain. Relationships between terms and their associations to resources constitute core domain conceptualization that forms basis for user modeling and is utilized by adaptation engine.

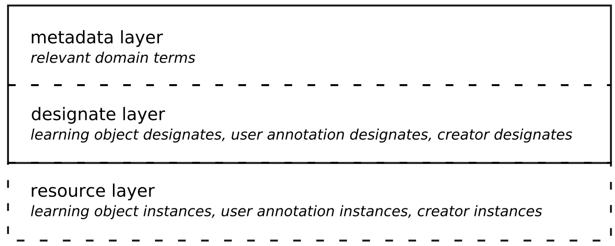

Figure 2: Domain model layers: metadata layer over designate layer. Resource instances are not a part of domain model (solid line).

We take advantage from multilayer design that explicitly differentiates between resources, their abstractions and semantic descriptions.

Domain model consists of designate layer and metadata layer (see Figure 2). It represents a conceptual abstraction over resource instances that are created and modified by content creators. We distinguish dual representation of resources: resource instances and resource designates:

- resource instance – a low level representation (e.g., XML, DocBook) of a resource that is detached from a conceptualization. It allows for easy authoring, maintenance and supports re-usability of the content. In fact, resource instances are not a part of domain model.
- resource designate – an abstraction of a resource residing in a domain model. It contains metadata about resources and relationships to other domain elements. It contains all information, which is relevant for adaptation decisions. It is independent of the actual content representation.

Our approach supports the notion of reusability and extendability in terms of content resource's lower level representation.

Designate layer covers learning objects, user-generated annotations and creators abstractions.

A basic component for education delivery is a learning object. For learning object we adopt a broader definition by IEEE, which defines a learning object as "any entity, digital or non-digital, that may be used for learning, education or training" [11]. The proposed domain model allows to define different types of learning objects to handle different learning activities. For instance, the student can learn a necessary topic from explanations and practice the acquired knowledge using exercises. Learning content extendability results in different forms of interaction supported by the model. Different forms of interaction within one system are aimed to improve student's learning experience and learning outcome.

In the age of ubiquitous Web 2.0, a user – learner – gains a central role and participates in content creation or modification. In a collaborative learning environment a new type of resource element emerges: a user-generated annotation. User-generated annotation we see as fundamental mean of collaboration that facilitates learning by allowing to create and share content (e.g., posting a comment on learning object), organize knowledge (e.g., using tags), or interact with the system and communicate with peers (discussing particular topic). Annotations are created by different users and it is important to track authorship of annotations as their quality or usefulness for learning may differ.

Creator designates represent authors of resources. Besides teachers and domain experts that typically create learning content we also consider learners, who contribute annotations to the content. Learner creator designates are not an integral part of learners' users models. In fact, the overlap of creators and learners may be theoretically very low as learners are not forced to contribute content. However, learners associations with annotations can be useful for user model update and maintenance. In addition to human content creators we also consider artificial ones. Basically, they are methods for automatic domain model creation that can create metadata for learning resources and annotations as well as relationships between metadata entities. The creator designate entities are important part of domain model as they enable to recognize quantitative characteristics of resources produced by different authors (e.g., by differently performing learners or a generation method used with specified confidence of correctness).

Metadata layer is formed by relevant domain terms – easily creatable descriptions that are connected with particular domain topics. It is important to note that relevant domain terms do not represent concepts in strict ontological definition, cf. [4]. They rather represent lexical reference to concepts, which form domain model (unlike relevant domain terms, concepts are not explicit).

Elements in domain model are interconnected via various types of relationships that represent various forms of relatedness between resources, between relevant domain terms and between resources and relevant domain terms. Similarly to domain elements, also domain relationships are associated with a creator – either a human or artificial author.

The proposed domain model meets the following requirements of the adaptive collaborative web-based learning 2.0:

- it clearly separates between domain conceptualization and content,
- it is extendable and allows to define new types of content,
- it comprises explicit support for modeling 2.0-based activities such as tagging, annotating, rating, commenting or different forms of collaboration,
- by design it facilitates own parts creation automation (while preserving adequate adaptation).

Proposed domain model is not restricted neither to predefined types of domain elements nor relationship between them. The goal of multilayer design is to separate metadata and resources, which reflects into ability to easily define new types on entities that makes domain model extendable and suitable for dynamic changes in an educational system triggered by the need to continually improve user experience by supporting learning with new forms of interaction and collaboration.

3.2 Adaptive LEarning Framework – ALEF

We have developed Adaptive LEarning Framework ALEF as a platform for general-purpose adaptive social web-based learning [19]. The proposed lightweight domain model forms one of the three fundamental pillars of the framework, which comprise:

- extensible personalization and course adaptation,
- student active participation in learning process, and
- lightweight domain modeling.

By developing Adaptive Learning Framework ALEF (see Figure 3) we showed that the proposed domain model is capable to describe and conceptualize underlying domain to be involved in adaptive collaborative learning. In ALEF we

- adopted the proposed domain model in order to support learning of technology oriented courses, which are delivered with personalized course navigation support via learning object recommendation,
- incorporated user-generated annotations into education process to support collaborative learning.

To support learning of technology oriented educational courses, we extended the proposed domain model with specific types of resources, metadata and relationships:

- we employed three different types of learning objects: *explanation*, *excercise* and *question* to cover different types of learning activities each providing a different level of interactivity for a learner
- we defined relationship *subsumed-by* allowing us to arrange learning objects to hierarchical book-like structure
- we defined three types of relationships between relevant domain terms: *relatedness* relationship to interconnect paradigmatically similar relevant domain terms, *is-a* relationship to define hierarchical structure between relevant domain terms, and *prerequisite* relationship to define partial ordering of concepts to be learned represented by relevant domain terms.

We used such domain model and created four adaptive courses: Functional programming, Logic programming, Procedural programming and Software engineering. The courses contain 1,352 learning objects, 1,180 relevant domain terms and 8,844 relationships (either between relevant domain terms or between learning objects and relevant domain terms) to date.

On the top of the defined domain model, an adaptive engine was devised by Michlík and Bieliková focusing on recommendation of exercises for limited time learning [16]. The method for exercise recommendation considered various criteria, three of them particularly rooted directly or

Figure 3: Screenshot of adaptive learning framework ALEF (In Slovak). It is divided to three vertical parts (left to right): (i) *navigational part* containing learning objects recommendations (1) and learning objects hierarchical menu (2); (ii) *educational content* containing selected learning object (3) eventually with user-generated annotations; and (iii) *learning and collaboration supporting widgets*: system activity score (4), error reporter (5), tagger (6) and external resource inserter (7).

indirectly in the domain model: conceptual appropriateness of an exercise for a learner, exercise difficulty, last access to an exercise by the learner.

In order to support collaboration during learning, which facilitates learning by creating content and sharing knowledge and we believe it also positively affects motivation, we devised annotation framework in ALEF. We see a user-generated annotation as fundamental mean of collaboration in an educational system. We devised the annotation framework to allow ALEF developers to define new annotation types by extending the domain model, creating necessary user interface and providing the logic for annotation processing.

We take advantage of the proposed domain model and we defined new annotation types as extension of annotation designates together with relationship between resource designates:

- *tag* – a term assigned to a selected part of the content,
- *(learner) question* – a question provided by a learner aimed to assess other learners on particular topic,
- *error report* – a short message describing a content error or inconsistency discovered during learning,
- *comment* – an arbitrary textual remark or note related to a content,
- *external resource* – an external resource identified by URI relevant to a content.
- *annotates* relationship – it associates an annotation with a resource's content.

Together with various annotation types we designed and implemented collaborative adaptive content creator components – annotation widgets – representing a user interface for creating, accessing and processing annotations. The created annotation widgets include tagger, question inserter and presenter, commentator, error reporter, external resource inserter. An example of the widgets are depicted on Figure 3.

We evaluated the proposed approach in a domain of learning programming. Due to the extensivity and complexity of the framework, we have not evaluated all aspects related to collaboration and learning content quality improvement based on annotations. However, we conducted two experiments and showed that [20]:

- learner question creation stimulates collaboration and course content enrichment;
- error reporter improves the quality of the learning content by revealing relevant errors in the content.

4. Automated Domain Model Acquisition

Automated domain modeling is aimed to reduce teacher's effort in providing (a huge number) of semantic description necessary for adaptation. It particularly is very important in dynamic learning environments, where user-generated content emerges often and there is a need to provide it with semantic descriptions.

When devising method for automated acquisition of domain model for adaptive learning, we considered specifics of educational domain such as small size of underlying text corpora; specificity of content, which is educational; specificity of language, which is typically explanatory; and presence of user-generated content.

We propose a method for automated domain acquisition coping with described specifics. In general, the method is based on statistical and linguistic processing of underlying resources created by teachers and eventually also by learners. The goal of the proposed method is automated acquisition of domain model, i.e., identification of relevant domain terms for underlying resources and creation of relationships between them.

The method consists of the following major steps (see Figure 4):

1. resource preprocessing,
2. relevant domain term identification,
3. relationship discovery,
4. domain model finalization.

In the first step, resource representation for further processing is prepared. The learning objects are analyzed, natural language pipeline is applied and extended vector representation is composed. In the second step we identify the most relevant terms for each learning object representation and create resource designate-term associations. Relationship discovery step is crucial for the whole method success. We perform graph, linguistic and statistical analysis to discover two types of relationships: relatedness and is-a relationships. The fourth (optional) step of the method is domain model finalization, where a teacher eventually modifies a created course according to his needs.

The method greatly benefits from lightweight semantic modeling and term-based skeleton of proposed domain model.

4.1 Relevant Domain Terms Identification

In this step we identify relevant domain terms that are suitable descriptors representing lightweight semantics. We build on preceding preprocessing of learning objects and user-generated annotations (if available). The process of relevant domain terms identification for learning objects and user-generated annotations slightly differs.

When processing learning objects, we extract frequently occurring terms. We adjust weights of terms by analyzing learning object formatting and increase relevance of terms according to their importance induced from formatting. When processing user-generated annotations we apply different techniques specific for different types of annotations. For example, we process user comments in a similar manner as learning objects and follow the whole natural language pipeline, whereas tags we only lemmatize. An important step in social annotations processing is filtering content that is irrelevant from domain conceptualization perspective.

4.2 Relationship Discovery

In this step we discover relationships between relevant domain terms already associated with learning objects. Our aim is to create relationships both quantitatively and qualitatively enough to enable adaptation in an educational system. We consider two types of relationships:

- *relatedness* relationship – paradigmatic similarity between terms, the most elemental form of a relationship between relevant domain terms representing lightweight semantics,
- *is-a* relationship – hypo-/hypernymical relationship between relevant domain terms representing that one term is semantically subsumed by another term.

We consider relatedness relationship as the most elemental form of relationship between relevant domain terms. Relatedness-based conceptualization of domain elements draws a basic structure of a domain. Based on such structure, a fundamental reasoning can be performed, which can result e.g., into recommendation of similar learning objects (described by "close" relevant domain terms).

In order to create relatedness relationships we proposed three different methods [21]:

- vector-based approach,
- spreading activation, and
- node centrality analysis.

Each method employs a specific approach to knowledge discovery. Vector-based approach is based on comparison of relevant domain term vector representations derived from vectors of all learning objects it is associated with. Spreading activation leverages actual graph-based representation of learning objects, relevant domain terms and relationships between them and leverages activation spreading metaphor to determine the degree of relatedness between relevant domain terms. By node centrality analysis we propagate actual domain model characteristics into the explicit links between relevant domain terms.

Is-a relationship is more specific type of relationship. It represents basic organizational relationship allowing to arrange relevant domain terms to hierarchy-like structure. Such structure is very important from the domain knowledge modeling perspective. It makes links between relevant domain terms, which represent faster knowledge spread in terms of connectionistic modeling of learner cognitive processes (typical for adaptive educational systems).

Our is-a relationship discovery is based on comprehensive statistic and linguistic analysis of learning object content. We employ tree different techniques:

- explanation phrase processing,
- determination phrase processing, and
- lexical analysis of relevant domain terms.

We proposed a set of rules for matching lexico-syntactic patterns, which indicate is-a relationship between terms in sentences. The proposed rules as well as patterns are created with respect to specifics of educational texts. We identified so-called *explanation phrases* and *determination phrases*, which often incorporate a form of is-a relationship. Explanation phrases aim to explain and expound new topic. They are often bound with a specific verb,

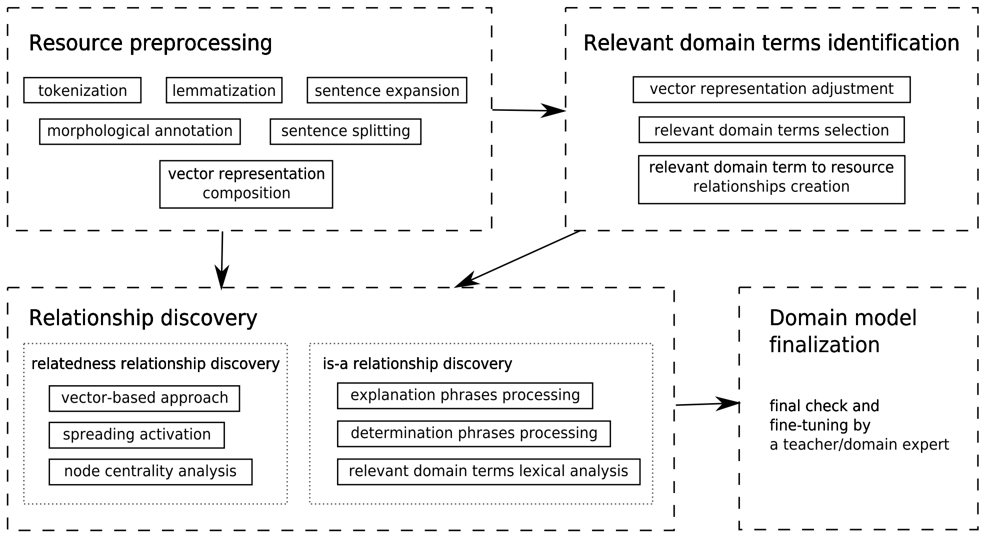

Figure 4: An overview of the method for automated domain model creation.

which we look for in the text. For example, from the sentence "We understand a list in Lisp as finite sequence of symbolic expressions." we derive that is-a(list, sequence of symbolic expressions). On the other hand, determination phrases often introduce new technical terms that are strongly relevant for particular domain and are often accompanied with more general domain relevant terms. For example, from the sentence "Lisp interprets the internal list as a function call to the function FIRST." we derive is-a(FIRST, function).

We complement explanation phrase and determination phrase processing with *lexical analysis of relevant domain terms* themselves. Each of the proposed approaches is aimed at different linguistic aspect of the text. By incorporating all approaches and merging their partial outputs we increase the chance of discovering a number of relationships that is, as a result, sufficient for making proper adaptation decisions.

5. Evaluation

In our work we evaluated relevant domain term relationships discovery. We employed three approaches to evaluation, each considering different criteria on acquired domain model relationships quality: a posteriori evaluation, evaluation against the gold standard and live experiment.

We performed the evaluation in domain of learning programming utilizing two courses: Functional programming and Logic programming. The official learning material for functional programming consists of 79 explanatory learning objects on the functional programming paradigm and programming techniques in the Lisp language. Logic programming part consists of 42 explanatory learning objects related to logic programming and Prolog language. All learning materials are in Slovak.

5.1 Relatedness relationship discovery

The evaluation of relatedness relationship discovery incorporated two steps. We:

- evaluated the automatically acquired domain model for Functional programming against the gold standard – a reference model created manually by the domain experts, and
- used the generated domain model during 2009/10 summer term in Logic programming course in live experiments.

In the first experiment we showed that node centrality analysis variant yields the most precise results with regards to the gold standard. The precision and extended recall[1] weighted harmonic mean measure (extended F-measure) of 0.652 [21]. We considered the results promising but wanted to perform more complex evaluation reducing risks resulting from the very nature of comparing against the gold standard.

Therefore we conducted a series of live experiments in the summer term of academic year 2009/10, where 66 students of Logic programming course were involved. The students were supported during learning by ALEF and learning object recommendation it featured in particular [16]. They were divided into groups based on the domain model used by the adaptive engine. There were two domain models employed: one created manually by a group of domain experts and another created automatically by the proposed method.

The experiments showed that students supported by recommendations based on automatically created domain model perform better indicating that automatically created relationships are more suitable for adaptation that relationships coming from the experts. However, due to the repeating problems with motivation of a small group of students, statistical characteristics of the performed experiments were negatively affected (also as a result of the limited number of participants). The results underline the complexity of live experiments and emphasize risks associated with live user evaluation, which is very expensive in terms of effort and time.

[1] we introduced extended recall measure to consider discrete weights of relationships in the gold standard as defined by experts

5.2 Is-a relationship discovery

The evaluation of is-a relationships involved:

- a posteri evaluation, and
- comparison with gold standard.

Experiments with live users were omitted due to the risks discussed above. In the first part of evaluation four domain experts on lisp language were asked to evaluate the correctness of the automatically acquired is-a relationships. The correctness of generated relationships as judged by experts was reasonable high when considering top 30 % from all generated is-a relationship candidates. The results are even more promising when we consider the fact that the domain experts did not agreed upon correctness of all generated relationships.

In the second part of evaluation we compared the results against the gold standard. We introduced strict measures to assess precision and recall of the generated relationships. We achieved very reasonable topological F-measure ranging from 0.513 to 0.653 (varying with a subset of the gold standard that was used).

6. Conclusions

Nowadays, adaptive web-based learning faces challenges thrown down by Web 2.0-induced paradigm shift in learning. The role of a learner has changed. He has greater autonomy and he became not only passive consumer of information, but he often actively participate in content enrichment or even creation. Improved learning experience facilitates interaction with and collaboration in learning environment. Students together manage and share content and organize the knowledge. Adaptive educational systems themselves have to adapt to new requirements imposed by collaborative web-base learning.

In this work we pursued two major goals related to adaptive collaborative web-based learning in order to overcome drawbacks of the state-of-the-art approaches: the rigidity of domain modeling with respect to social-oriented collaborative learning, and the lack of methods supporting automatic or automated domain model creation. As a result, our core contributions are our proposals of:

- *lightweight domain model for adaptive collaborative web-based learning* – we proposed and evaluated new generation domain model for adaptive collaborative educational systems which clearly separates between domain conceptualization and content, provides explicit support for collaborative interactive learning and it is proposed with regard to automated creation and enrichment facilitated by leveraging user-generated content, while preserving flexible and accurate adaptation [19].
- *method for automated domain model acquisition* – we proposed and evaluated method for automated domain model acquisition that is based on statistical and linguistic processing of underlying resource content. The method consists of several steps, which cover the whole knowledge acquisition process ranging from resource preprocessing to domain model final fine-tuning performed by a teacher, with particular method steps representing partial contributions to the field of adaptive educational systems authoring [22, 21, 14, 18].
- *adaptive learning framework ALEF and its learning content* – the author of this work actively participated in and co-led the development of adaptive learning framework ALEF [19], which is a practical result of a research conducted in the area of adaptive web-based learning presented in our work. ALEF is an ultimate education supporting software solution merging adaptive learning, collaborative learning and interactivity built on the concepts of Web 2.0. Adaptive learning framework is currently actively utilized at Slovak University of Technology in Bratislava. It supports learning in the four courses: Functional programming, Logic programming, Procedural programming, and Principles of software engineering.

Acknowledgements. This work was partially supported by Cultural and Educational Grant Agency of the Ministry of Education of Slovak Republic, grants No. KEGA 3/5187/07 and KEGA 028-025STU-4/2010, Scientific Grant Agency of the Ministry of Education of Slovak Republic and the Slovak Academy of Sciences, grant No. VG1/0675/11, Slovak Research and Development Agency, grant No. APVV-0208-10 and it is the partial result of the Research & Development Operational Programme for the project Support of Center of Excellence for Smart Technologies, Systems and Services II, ITMS 26240120029, co-funded by the ERDF.

References

[1] I. Beaumont and P. Brusilovsky. Adaptive educational hypermedia: From ideas to real systems. In H. Maurer (Ed.), *Proc. of ED-MEDIA'95 – World conf. on educational multimedia and hypermedia*. AACE, pp. 93–98, 1995.

[2] P. Brusilovsky, J. Knapp and J. Gamper. Supporting teachers as content authors in intelligent educational systems. In *Int. Journal of Knowledge and Learning*, vol. 2, no. 3/4, pp. 191–215, 2006.

[3] P. Cimiano, A. Hotho and S. Staab. Comparing conceptual, divisive and agglomerative clustering for learning taxonomies from text. In *Proc. of the European Conf. on Artificial Intelligence (ECAI)*, IOS Press, pp. 435–439, 2004.

[4] P. Cimiano. *Ontology Learning and Population from Text: Algorithms, Evaluation and Applications*. Springer, 347p, 2006.

[5] A.I. Cristea and A. de Mooij. LAOS: Layered WWW AHS Authoring Model and their corresponding Algebraic Operators. In *Proc. The Twelfth Int. World Wide Web Conference. WWW'03. Alternate Track on Education, Budapest*, Hungary, 2003.

[6] A.I. Cristea and A. de Mooij. Designer Adaptation in Adaptive Hypermedia Authoring. In *Proc. of the Int. Conf. on Information Technology: Computers and Communications ITCC'03*. IEEE, pp. 444–448, 2003.

[7] A.I. Cristea, F. Ghali and M. Joy. Social, Personalized Lifelong Learning. In *E-Infrastructures and Technologies for Lifelong Learning (ETLL)*, IGI Global, pp. 90–125, 2011.

[8] S. Downes. E-learning 2.0. *eLearn magazine*. ACM, No. 10, 2005.

[9] E. Drymonas, K. Zervanou and E. Petrakis. Unsupervised ontology acquisition from plain texts: the Ontogain system. In *Proc. of the 15th Int. Conf. on Applications of Natural Language to Information Systems*, NLDB 2010, Springer, pp. 277–287, 2010.

[10] M. Hearst. Automatic acquisition of hyponyms from large text corpora. In *Proc. of the 14th Int. Conf. on Computational Linguistics*, COLING, ACL, pp. 539–545, 1992.

[11] IEEE LTSC. *Draft Standard for Learning Object Metadata. IEEE Standard 1484.12.1*. New York: IEEE, 2002.

[12] X. Jiang and A. Tan. Crctol: A semantic-based domain ontology learning system. *Journal of the American Society for Information Science and Technology* 61, 1, pp. 150–168, 2009.

[13] N. Koch and M. Wirsing. The Munich Reference Model for Adaptive Hypermedia Applications. In *Adaptive Hypermedia and Adaptive Web-Based Systems*, LNCS 2347, Springer, pp. 213–222.

[14] M. Lučanský, M. Šimko and M. Bieliková. Enhancing Automatic Term Recognition Algorithms with HTML Tags Processing. In *Proc. of Int. Conf. on Computer Systems and Technologies, ComSysTech'11*. ACM New York, pp. 173–178, 2011.

[15] A. Mädche and S. Staab. Discovering conceptual relations from text. In *Proc. of the 14th European Conf. on Artificial Intelligence (ECAI)*, IOS Press, pp. 321–325, 2000.

[16] P. Michlík and M., Bieliková. Exercises Recommending for Limited Time Learning. *Procedia Computer Science*. Vol. 1, Issue 2, Elsevier, ISSN 1877-0509, pp. 2821–2828, 2010.

[17] S. Sosnovsky, P. Brusilovsky and M. Yudelson. Supporting Adaptive Hypermedia Authors with Automated Content Indexing. In *Proc. of 2nd Int. Workshop on Authoring of Adaptive and Adaptable Educational Hypermedia at the 3rd Int. Conf. on Adaptive Hypermedia and Adaptive Web-Based Systems (AH'2004)*, pp. 380–389, 2004.

[18] M. Šimko. Automated Domain Model Creation for Adaptive Social Learning System. In *Information Sciences and Technologies Bulletin of the ACM Slovakia*, Vol. 3, No. 2, pp. 119–121, 2011.

[19] M. Šimko, M. Barla and M. Bieliková. ALEF: A Framework for Adaptive Web-based Learning 2.0. In Reynolds, N., Turcsányi Szabó, M. (Eds.): *KCKS 2010, IFIP Advances in Information and Communication Technology*, Vol. 324. Held as Part of World Computer Congress 2010. Springer, pp. 367–378, 2010.

[20] M. Šimko, M. Barla, V. Mihál, M. Unčík and M. Bieliková. Supporting Collaborative Web-Based Education via Annotations. In *Proc. of World Conference on Educational Multimedia, Hypermedia & Telecommunications, ED-MEDIA 2011*. Chesapeake, VA: AACE, pp. 2576–2585, 2011

[21] M. Šimko and M. Bieliková. Automatic Concept Relationships Discovery for an Adaptive E-course. In Barnes, T., Desmarais, M., Romero, C., Ventura, S. (Eds.). *Proc. of Educational Data Mining 2009: 2nd Int. Conf. on Educational Data Mining*. Cordoba, Spain, pp. 171–179, 2009.

[22] M. Šimko and M. Bieliková. Automated Educational Course Metadata Generation Based on Semantics Discovery. In *LNCS 5794, Proc. of European Conf. on Technology Enhanced Learning: Learning in the Synergy of Multiple Disciplines, EC TEL 2009*, Nice, France. Springer, pp. 99–105, 2009.

[23] M. Unčík and M., Bieliková. Annotating Educational Content by Questions Created by Learners. In *Proc. of 5th Int. Workshop on Semantic Media Adaptation and Personalization, SMAP 2010*, IEEE, pp. 13–18, 2010.

[24] S. Weibelzahl and G. Weber. Adapting to prior knowledge of learners. In *Proc. of the 2nd Int. Conf. on Adaptive Hypermedia*, AH'02, Springer, pp. 448–451, 2002

[25] H. Wu, E. de Kort and P. de Bra. Design Issues for General-Purpose Adaptive Hypermedia Systems. In *Proc. of the 12th ACM conference on Hypertext and Hypermedia*, HT '01. ACM, pp. 141–150, 2001.

Selected Papers by the Author

Šimko, M., Barla, M., Mihál, V., Unčík, M., Bieliková, M. Supporting Collaborative Web-Based Education via Annotations. In *Proc. of World Conference on Educational Multimedia, Hypermedia & Telecommunications, ED-MEDIA 2011*. Chesapeake, VA: AACE, pp. 2576–2585, 2011

Šimko, M., Barla, M., Bieliková, M. ALEF: A Framework for Adaptive Web-based Learning 2.0. In Reynolds, N., Turcsányi Szabó, M. (Eds.): *KCKS 2010, IFIP Advances in Information and Communication Technology*, Vol. 324. Held as Part of World Computer Congress 2010. Springer, pp. 367–378, 2010

Barla, M., Bieliková, M., Bou-Ezzedine, A., Kramár, T., Šimko, M, Vozár, O. On the Impact of Adaptive Test Question Selection for Learning Efficiency. *Computers & Education*, Vol. 55, Issue 2, pp. 846–857, 2010.

Majer, T., Šimko, M. Leveraging Microblogs for Resource Ranking. In *Proc. of 38th Conference on Current Trends in Theory and Practice of Computer Science*, SOFSEM 2012. Springer, pp. 518–529, 2012.

Móro, R., Srba, I., Unčík, M., Bieliková, M., Šimko, M. Towards Collaborative Metadata Enrichment for Adaptive Web-Based Learning. In *Proc. of Int. Workshop on Computational Social Networks, Web Intelligence 2011*. IEEE, pp. 106–109, 2011.

Lučanský, M., Šimko, M., Bieliková, M. Enhancing Automatic Term Recognition Algorithms with HTML Tags Processing. In *Proc. of Int. Conf. on Computer Systems and Technologies, ComSysTech'11*. ACM New York, pp. 173–178, 2011.

Šimko, M. Bieliková, M. User Modeling Based on Emergent Domain Semantics. In *LNCS 6075, Proc. of User Modeling, Adaptation and Personalization, UMAP 2010*, Doctoral Consortium. Big Island, Hawaii, Springer, pp. 411–414, 2010.

Šimko, M. Bieliková, M. Automated Educational Course Metadata Generation Based on Semantics Discovery. In *LNCS 5794, Proc. of European Conf. on Technology Enhanced Learning: Learning in the Synergy of Multiple Disciplines, EC TEL 2009*, Nice, France. Springer, pp. 99–105, 2009.

Bieliková, M., Barla, M., Šimko, M. Lightweight Semantics for the "Wild Web". In B. White, P. Isaías and F. M. Santoro (eds.): *Proc. of the IADIS Int. Conf. on WWW/Internet, ICWI 2011*, IADIS Press, pp. xxv–xxxii, 2011.

Šimko, M. Automated Domain Model Creation for Adaptive Social Learning System. In *Information Sciences and Technologies Bulletin of the ACM Slovakia, Special Section on Student Research in Informatics and Information Technologies*, Vol. 3, No. 2, pp. 119–121, 2011.

Bieliková, M., Šimko, M., Barla, M. Personalized Web-Based Learning 2.0. In *Proc. of 8th Int. Conf. on Emerging eLearning Technologies and Applications, ICETA 2010*. Stará Lesná, Slovakia, Košice: Elfa, 2010.

Bieliková, M., Šimko, M., Barla, M., Chudá, D., Michlík, P., Labaj, M., Mihál, V., Unčík, M. ALEF: Web 2.0 Principles in Learning and Collaboration. In *Proc. of the 6th E-learning Conference: E-learning and the Knowledge Society*. Riga: Riga Technical University, pp. 54–59, 2010.

Šimko, M., Bieliková, M. Automatic Concept Relationships Discovery for an Adaptive E-course. In Barnes, T., Desmarais, M., Romero, C., Ventura, S. (Eds.). *Proc. of Educational Data Mining 2009: 2nd Int. Conf. on Educational Data Mining*. Cordoba, Spain, pp. 171–179, 2009.

Šimko, M. Bieliková, M. Leveraging Lightweight Semantics for Search Improvement. In *ACM Student Research Competition ACM.SRC 2009 at Hypertext 2009, 20th ACM Conference on Hypertext and Hypermedia*, Torino, Italy, 2009.

Šimko, M., Bieliková, M. Improving Search Results with Lightweight Semantic Search. In Grobelnik, M., Mika, P., Douc, T. T. and Wang, H. (Eds.). *CEUR, Vol. 491, Proc. of the Workshop on Semantic Search, SemSearch 2009 at the 18th Int. World Wide Web Conference, WWW 2009*. Madrid, Spain, pp. 53–54, 2009.

Šimko, M., Bieliková, M. Automatic Discovery of Relationships for Learning Content Semantic Description. (In Slovak) In *Proc. of 6th Workshop on Intelligent and Knowledge-oriented Technologies*, WIKT 2011, Herľany, Slovakia, pp. 33–38, 2011.

Lučanský, M., Šimko, M. Do HTML tags improve results of Automatic Term Recognition Algorithms? (In Slovak) In *Proc. of Znalosti (Knowledge) 2011*, Stará Lesná, Slovakia, pp. 279–282, 2011.

Šimko, M., Bieliková M. (Semi)Automatic E-Course Metadata Generation. (In Slovak) In *Proc. of Znalosti (Knowledge) 2009*, Brno, Czech republic, pp. 246–257, 2009.

Šimko, M., Bieliková M. Semantic Search on the Web Using Heterogeneous Information Sources. (In Slovak) In *Proc. of 4th Workshop on Intelligent and Knowledge-oriented Technologies*, WIKT 2009, Herľany, Slovakia, pp. 29–33, 2009.

Šimko, M., Bieliková M. Domain Concepts Relationships Discovery. (In Slovak) In *Proc. of 3rd Workshop on Intelligent and Knowledge-oriented Technologies*, WIKT 2008, Smolenice, Slovakia, pp. 17–20, 2008.