II. ITS Theoretical Background

To provide individualized tutoring the ITSs are using a set of information from the problem domain, the pedagogical knowledge and the knowledge about the learner. Problem domain and pedagogical knowledge is more static and is changing only when a new tutoring strategy is implemented or a new methodical material is imported/made available in the system. ITSs functional aspects will not be covered in more details as they are out of the scope of this paper. The knowledge about the learner is a subject of rapid change as learner constantly evolves, changes his expectations, skills, experience and requirements. The knowledge about learner can be divided into two blocks: static and dynamic information. Both kinds of information may have relationships and dependences. These relationships enable context reasoning, which is based on a primary context (sensed form sensors) and relationship information, and enable user/system to calculate/gather secondary (calculated) context [6]. For example, the location data provided by ITS mobile application, like GPS coordinates, is meaningless for the tutoring system. However, these coordinates can be transformed into more meaningful data, like location name with its corresponding data. The ITS instead of latitude and longitude can get location name, its application purpose, number of seats in the room, light conditions, and any other relevant information. In any of the context information gathering cases, it is important to evaluate the number of required resources to get the intended result with an acceptable quality and a level of trust. Sometimes information gathering is so expensive and long, that at the end, the gathered information might be outdated and unusable.

For the classroom tutoring learner profiling/modelling is done by a teacher, is based on the teacher's experience, knowledge and intuition, and only for a small groups of learners. For large groups it is not physically possible to evaluate each individual performance, expectations and preferences without any technological support. Therefore, in real life situations these adaptions are implemented at a certain level (usually group) of individualized tutoring material. For example, in the classroom teacher covers a theoretical material, but in the online course management or in the ITS adapts the tutoring material by providing examples and tasks with different learning styles and difficulty level.

The idea of ITSs is to adapt the teaching level and delivery according to learners' skills, knowledge and other needs, and to provide a just-in-time feedback. Traditionally, ITSs observe learners' behaviour and react by providing feedback, hints, suggesting learning topics, searching for tutoring material, etc. [7]. Typical ITSs contain a predefined learning path, which a learner precedes by studying a concept by exploring learning materials [8].

In the both adaptation scenarios there is a need to collect, process, analyse and use the data about the learning process and the environment to facilitate learning process. Data mining algorithms and tools can be used for this purpose.

Data mining applications in ITS domain is a subject of research for number of years. Kotsiantis et al. [9] have done pioneer work in predicting student dropouts by machine learning techniques. They were taking data from educational institution databases, like personal data, course activity in the beginning of course, and used six modelling paradigms – Naive Bayes model, decision trees, back-propagation neutral networks, support vector machines, three-nearest neighbours' algorithm and logistic regression. However, there is still a need for further result evaluation and application in tutoring process. Hamalainen et al. [1] says that unwillingness to use data mining and machine learning techniques in an educational context is partly due to domain-specific problems and other related problems. Nowadays, technology development overcomes this problem through tool simplification and availability improvements, for example, through cloud based [10] [11] machine learning services [12] where all backend complexity is managed by a service provider. This new situation enables researchers to focus on what matters and ignore related indirect activity or complexity.

Hamalainen et al. [1] clarified general modelling paradigms for developing adaptive ITSs and tested them with real course data. They adapted a dual principle of descriptive and predictive modelling. This principle combines classical paradigms of data mining and machine learning (see Fig. 1).

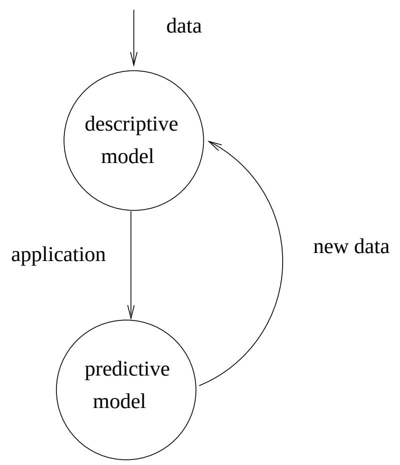

Fig. 1. The iterative process of descriptive and predictive modelling [1].

Descriptive phase analyses learners' observation data and searches for correlations and association rules, and predictive phase generates the learner model. The same process can be used for context modelling. The descriptive model analyses data from the past or surroundings and pattern searching, like correlations and association rules, whereas the predictive model defines structure (variables and their relations). Hamalainen et al. [1] propose to apply the resulting models to the next course. The result serves as a validation of the model and guides the construction process of the next stage. However, this approach of only applying a new model to the next course, but not to the currently ongoing, is limiting as the next course might have different goals, teaching methods, etc. Therefore, for the given example, the risk of inaccurate models increases and can drive all adaptivity to failure. To decrease the above-mentioned risks, it is proposed to apply to the ongoing course the iterative process introduced by Hamalainen et al. [1]. The process used within the same year by building a model on a group of students and then apply the model to the students within the group to recognize student patterns and correlations within the context of an independent environment.

The same challenges apply to the context-aware systems designed for learning support. The survey of context-aware recommender systems presented by Vebert et. al. [3] shows that several context-aware systems for learning have been elaborated, but important challenges still remain. For example, validation of the developed prototypes, their impact assessment on the learning and context-aware system performance in real life learning scenarios. The context acquisition, evaluation, interaction and required data infrastructure are examples of many important challenges in the classroom and the intelligent tutoring systems domains. The main focus in this paper is on context acquisition, evaluation and application. The interaction and the data infrastructure challenges are relevant for large scale purely online based learning systems that are not the subject of this research paper.

The most common context elements that are used for context-aware learning systems are [3]: computing (software, hardware and network), location (qualitative, quantitative, proximity), physical (noise level, weather, lighting), time (timestamps and time intervals), activity (topic, action, task), resource (general, technical, annotation, educational, relation), user (basic info, knowledge, interest/preference, learning style) and social relations. The considered context elements in the scope of this research are: time, activity, resource and the user context. In more details they are discussed in Section IV of this paper.

III. Analysed Hybrid Learning Scenario

This research paper analyses a hybrid learning scenario, where traditional classroom training is provided with additional learning materials available via online ITS prototype (see Fig. 2). Learning materials contain video lectures, theoretical material and related examples. At the end of the course learners' feedback is collected via survey, that gathers information about the teacher performance, used presentation techniques, number of hours per week spent on learning and course topic relation to previously learned courses. Each video lecture is marked with keywords and has detailed covered topics. The learners are identified by their names, surnames, student IDs and e-mail addresses.

Each learner is uniquely identified in the educational institution systems and in ITS prototype (for UI example see Fig. 2) by user ID and username. These identifiers are used to join together the information from various data sources and assign them to one physical person – a learner:

- The Information systems of the educational institution contain data about learners' previous education, performance scores, chosen study program, study start date, list of selected courses, personal information (name, surname, place of living, bank account information, identity number, phone number, e-mail address), information about used university facilities and resources, like, WiFi connection, browsed internet resources, accessed scientific databases and many other information.
- The learning management system (Moodle) of the educational institution stores information about the available learning material, access log information (which resource has been downloaded by whom), number of assigned students, tests, scorecards and messages sent to the learner.
- The ITS prototype contains more detailed information about the learning material which is stored in Learning Object Metadata (LOM) [13] metadata format, as well as detailed information about learner activities and used resources.

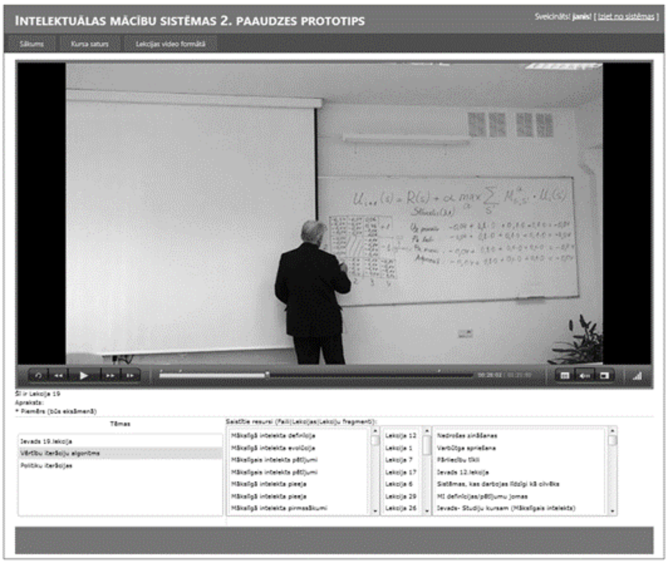

Fig. 2. ITS Lecture video player screenshot.

All these data sources have their own storage format, technologies used, data structures, system owners and application purposes. In order to extract valuable information from the data, there is a need for specific mechanisms and approaches. Context models can serve as models that connect these data sources and enable information extraction in reusable way. Therefore, a simplified learning context model is proposed in the next section.

IV. Learning Context Model and Data References

For the analysed hybrid learning scenario computing, social relation and physical context elements are not relevant. In particular, computing element nowadays is not that important. As an example, mobile smartphones have more capabilities than PCs that are more than 10 years old. These devices can play videos with the same level of performance as modern computers. Industry standards, like HTML5, enable the same user experience across devices (smartphones, tablets, computers). The mobile 3G and 4G networks have the same throughput as LAN connections. The location element can serve as a valuable information source about learning content consumption habits and environmental surroundings and at some point improve the learning content selection. Location awareness and mobile applications have been a subject of interest for context aware system authors for years. There are several prototypes developed where, based on location information, learning material is adaptively presented and/or learning material is broadcasted to mobile device [14], [15], [5]. However, even if mobile device application can be beneficial for some scenarios, still most of the higher education learning is happening in university classrooms where whiteboards at some point are replaced by interactive PowerPoint presentations, where the number of learners can be from few up to several hundreds, where teachers are using one kind of learning material and one kind of approach to deliver it to the learners in the classrooms. In other words, researchers are focusing on searching for new ways to learn instead of gaining full potential out of the existing learning setup at universities.

To meet the above mentioned objective hybrid learning scenario is taken as a basis for context awareness implementation. The learning context model is created from the analysed context elements (time, activity, resource, user). Elements are selected based on the set of data that is available from various data sources in the problem domain and additionally can provide value for the analysed hybrid learning scenario. The time factor is common for all context elements and therefore is not defined as a separate context element.

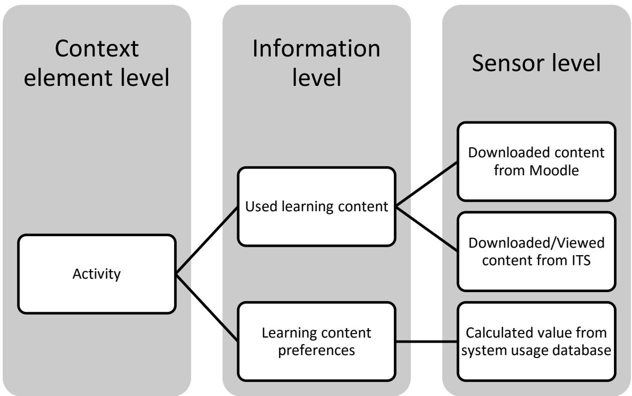

Fig. 3. Context element model structure example.

Fig. 3 presents a part of context element model structure. The context element model consists of tree sublevels. At root level all context element values are presented for further application in ITSs and/or classroom training. At the information level the data from single or multiple physical, logical and/or virtual sensors is aggregated, processed and prepared for further application. At the sensor level raw sensor data is captured and/or read and prepared for analysis.

To remind, the physical sensors might be: light, visual context (video and infra-red cameras), audio, motion, location, gesture, temperature and biometric sensors. In the ITS context these sensors usually are web cameras or sensors that are measuring learner's heart rate, or blood pressure to identify learner's response on learning content [16]. Virtual sensors are software based and are used to observe user actions, like the mouse movement, clicked links in the website, etc. The sensors from this group are among the most widely used in ITS context, because they do not require any physical components and have less impact on learning process. The logical sensors use virtual and/or hardware sensor data and based on the built in mechanisms implement reasoning or the sensor data fusion to generate the data output.

Table I presents examples of structure element, sensor data relationships and data extraction methods demonstrated in Fig. 3.

TABLE I
INFORMATION ELEMENT AND DATA SOURCE RELATIONSHIPS

Information element	**Sensor type**	**Data source**	**Data extraction method**
Used learning content	Virtual sensor	Moodle log	SQL Select statement
	Virtual sensor	ITS log	SQL Select statement
Learning content preferences	Logical sensor	Moodle Log, Resource DB, ITS student model, Survey data	SQL statements and reasoning techniques

V. PROPOSED CONTEXTUAL INFORMATION VERIFICATION APPROACH

Many of the surveyed systems [3] rely on a manual input from the users, moreover, the authors claim that several systems indicated that they use contextual data, but do not describe the methods that are used to capture this data. In the analysed learning scenario (see Section III) manual user input data is available only at the end of the course via learners' survey. Therefore, the automatic/semi-automatic context acquisition techniques should be used. Two step approach is proposed to validate the context acquisition results (see. Fig. 4). The approach consists of two steps to split fully automated and semi-automated activities.

In the first step all available data about the particular context element from the various data sources is aggregated. In the first step's second sub step data values are compared to each other in order to identify differences. If all data items/sources provide the same data/information, then the step two is skipped as the human input will not add any additional value to the context element. The context element value is marked as fully trusted. Then this context element can be used in the ITS or can be evaluated by the teacher to choose/adjust the tutoring strategy.

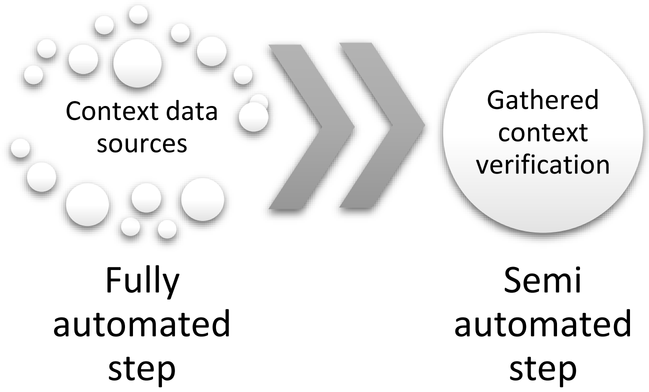

Fig. 4. Contextual information verification approach steps.

In case when values contain different data/information in a step two, based on the level of the trust of the data sources defined in the context model, the element's level of trust is calculated. Finally, if the system is not able to assign the level of trust for the context element, the manual input either from the learner or from teacher is requested. In this case the contextual information management component is storing the calculated as well as manually assigned value of the context element.

This semi-automated contextual verification process can be made fully-automated by implementing various mechanisms of uncertain knowledge management and reasoning techniques, like probabilistic reasoning and probabilistic reasoning over time, supported by simple and complex decision making techniques[17].

VI. Contextual Information Application Scenario

The hybrid learning scenario discussed in Section III has been tested at the Faculty of Computer Science and Information Technology of Riga Technical University for the last 5 years in Master level study course Artificial Intelligence. The ITS prototype collects the data about each learner's interaction with the system. This interaction data provides information about learners' behaviour during the tutoring session, indicated learner's preferences, learning date, time and possibly not well understood topics from the classroom session and/or particular subjects of interest.

The information from the university systems describes learner's performance in other courses studied at Riga Technical University. Annual learner's survey provides subjective information about the number of attended lectures, level of course topic understanding, teacher's performance and teaching approach evaluation and number of hours spent per week on subject learning and downloaded tutoring content from the learning management system. Unfortunately, detailed final examination information is not available at this stage.

This information is used to create the context based descriptive student model that is applied in the ITS and at some degree in the classroom. The observed and collected data about the learning process is used to generate predictive model that includes recommendations to the teacher on possible learning content/teaching strategy adjustments and changes in the way related/recommended learning objects are selected in the ITS. The descriptive model serves as a foundation to generate individual and/or group based learner profiles. That can be used in the context reasoning to draw conclusions and/or describe/plan impact of the chosen approach to the learner's performance.

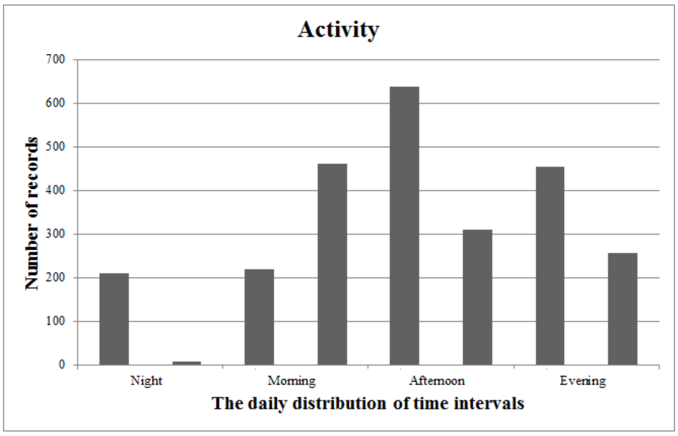

Fig. 5. Activity level across day.

The ITS usage can be reviewed in more detail. In particular, from this data the teacher is able to identify the learning habits of the student group (see Fig. 5). These results show that afternoons and evenings are almost equally popular time to learn for this group of students. The left bar of each day period shows the level of the activity for the first half of the month, consequently the right bar shows the data from the second half of the month. The data demonstrates that afternoons and evenings are the most active periods of system use and this might be due to students' employment and not due to the learning style.

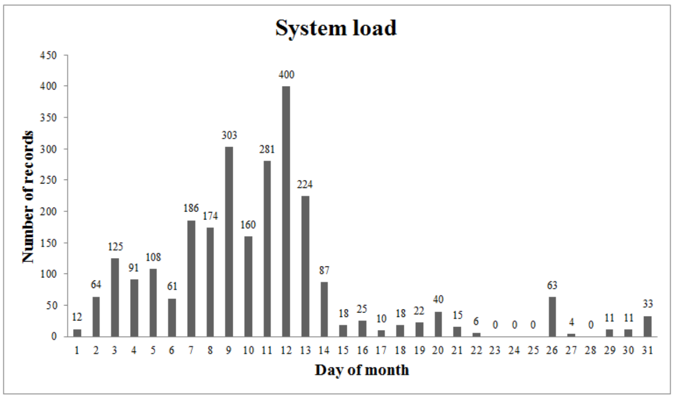

Fig. 6. ITS system load.

This assumption can be verified by checking students' context if it contains the information about employment status. This is only one of the possible perspectives of the data analysis on the activity and students' learning habits.

Additionally, the study of the above mentioned information shows that students are willing to get access to additional learning material and are using it to prepare for the course exam (see Fig. 6). The level of interest significantly drops after the exam day.

The developed ITS prototype provides more detailed information about which lecture, which topic had been of interest most. In the studied example data set Lecture 14 and Lecture 12 with corresponding course topics are among the most used ones. Lecture 1 is not analysed in more details because it was the first lecture that was viewed by all students out of curiosity.

The studied data indicates that the topics covered in both mentioned lectures might not be well understood during the classroom session and students used the opportunity to study these topics once again. That raises more questions, like, was there enough time spent to explain the topics that were covered in these lectures, maybe there is a need to spend more time on the topic and/or related topics.

To validate this assumption more detailed analysis is required, as there might be several reasons for this extra attention from students. For example, if the questions of the final exam have clear and obvious connections between the topics covered in the lecture. This information can be used to identify possible/required changes either to the teacher's performance or the way how this topic has been explained.

The provided example data demonstrates the potential of the contextual data application. The information on individual learning habits joint with information about individual student's employment, hobbies, previous experience, interests, performance in other courses may provide valuable information to the course author and to better understand how students of the particular group study, how they prefer to access information and many other aspects. Most of the required information is already present in the information systems of the educational institution.

VII. Conclusions and Future Work

The context-awareness has the potential to let applications provide more personalized learning experience to learners. The analysed scenario clearly demonstrates that current information systems of the educational institution store valuable information that can be used in context-aware educational systems and further research can be done to improve the information gathering and processing capabilities. Furthermore, it is important for the learning course authors to use this information to improve both virtual and classroom learning experiences.

The availability of lecture video that reflects classroom lecture content is a subject of interest for students. At the moment there is no research done to identify whether the analysed hybrid scenario has positive and/or negative impact on the overall learning process and students' performance.

The lack of frequent feedback from learners limits the ability of the system and course authors to respond to new types and kinds of learners' demands and evaluate ITSs prototype performance. Therefore, further research will focus on automated/semi-automated contextual information gathering and evaluation mechanism development.

References

[1] Hämäläinen, W., Laine, T., and Sutinen, E., "Data mining in personalizing distance education courses", in Editor (Ed.)^(Eds.): "Book Data mining in personalizing distance education courses" (Citeseer, 2004, edn.), pp.

[2] Dey, A.K., "Understanding and Using Context", *Personal Ubiquitous Comput.*, 2001, vol. 5, (1), pp. 4–7. http://dx.doi.org/10.1007/s007790170019

[3] Verbert, K., Manouselis, N., Ochoa, X., Wolpers, M., Drachsler, H., Bosnic, I., and Duval, E. "Context-aware recommender systems for learning: a survey and future challenges", *IEEE Transactions on Learning Technologies*, 2012, vol. 5, (4), pp. 318–335. http://dx.doi.org/10.1109/TLT.2012.11

[4] Anohina, A., Lavendelis, E., and Grundspenkis, J. "The Concept Map-Based Knowledge Assessment System with Reduction of Task Difficulty," *Information Systems Development*, 2009, pp. 1–13.

[5] Lonsdale, P., Baber, C., Sharples, M., Byrne, W., Arvanitis, T.N., Brundell, P., and Beale, R. "Context awareness for MOBIlearn: creating an engaging learning experience in an art museum", *Mobile learning anytime everywhere*, 2005, 115.

[6] Perera, C., Zaslavsky, A., Christen, P., and Georgakopoulos, D., "Context Aware Computing for The Internet of Things: A Survey", 2013

[7] Lavendelis, E. "Open multi-agent architecture and methodology for intelligent tutoring system development," Summary of Doctoral Thesis, 2009, pp. 49.

[8] Devedzic, V., Debenham, J., and Popovic, D. "Teaching formal languages by an intelligent tutoring system," *Educ Tech Soc*, 2000, vol. 3, (2), pp. 36–49.

[9] Kotsiantis, S., Pierrakeas, C., and Pintelas, P. "Preventing student dropout in distance learning using machine learning techniques", in Editor (Ed.)^(Eds.): "Book Preventing student dropout in distance learning using machine learning techniques" (Springer, 2003, edn.), pp. 267–274.

[10] Zhang, Q., Cheng, L., and Boutaba, R.: "Cloud computing: state-of-the-art and research challenges," *Journal of internet services and applications*, 2010, vol. 1, (1), pp. 7–18. http://dx.doi.org/10.1007/s13174-010-0007-6

[11] Kranjc, J., Podpečan, V., and Lavrač, N. "Clowdflows: a cloud based scientific workflow platform": "Machine Learning and Knowledge Discovery in Databases" (Springer, 2012), pp. 816–819.

[12] Low, Y., Bickson, D., Gonzalez, J., Guestrin, C., Kyrola, A., and Hellerstein, J.M. "Distributed GraphLab: a framework for machine learning and data mining in the cloud," in *Proceedings of the VLDB Endowment*, 2012, vol. 5, (8), pp. 716–727. http://dx.doi.org/10.14778/2212351.2212354

[13] Neven, F., and Duval, E., "Reusable learning objects: a survey of LOM-based repositories," in Editor (Ed.)^(Eds.): "Book Reusable learning objects: a survey of LOM-based repositories" (ACM, 2002, edn.), pp. 291–294.

[14] Bellavista, P., Corradi, A., Fanelli, M., and Foschini, L., "A survey of context data distribution for mobile ubiquitous systems," *ACM Computing Surveys (CSUR)*, 2012, vol. 44, (4), pp. 24. http://dx.doi.org/10.1145/2333112.2333119

[15] Gómez, S., Zervas, P., Sampson, D.G., and Fabregat, R. "Context-aware adaptive and personalized mobile learning delivery supported by UoLmP," *Journal of King Saud University-Computer and Information Sciences*, 2014, vol. 26, no. 1, pp. 47–61. http://dx.doi.org/10.1016/j.jksuci.2013.10.008

[16] Rikure, T., "Development of learner psycho-physiological model based adaptive tutoring systems," Summary of doctoral thesis, 2009

[17] Russell, S.J., and Norvig, P., *Artificial intelligence: a modern approach*, Prentice hall Englewood Cliffs, 2010.

Janis Bicans is a doctoral student with the Faculty of Computer Science and Information Technology of Riga Technical University, Riga, Latvia. He started his studies in 2005 and received a Master degree in 2010. The topic of the Master thesis was "Development of learning objects repository and intelligent tutoring system for intellectual learning support".
Since 2010, he has been a Researcher with the Department of Systems Theory and Design of Riga Technical University. His main research interests are sensor data fusion, cortex-aware and intelligent tutoring systems.
Address: 1 Setas Street, Riga, Latvia, LV-1007.
E-mail: Janis.Bicans@rtu.lv