Learning Analytics to Support the Use of Virtual Worlds in the Classroom

Michael D. Kickmeier-Rust and Dietrich Albert

Cognitive Science Section, Knowledge Management Institute
Graz University of Technology
Inffeldgasse 13, 8010 Graz, Austria
{michael.kickmeier-rust,dietrich.albert}@tugraz.at

Abstract. Virtual worlds in education, intelligent tutorial systems, and learning analytics – all these are current buzz words in recent educational research. In this paper we introduce ProNIFA, a tool to support theory-grounded learning analytics developed in the context of the European project Next-Tell. Concretely we describe a log file analysis and presentation module to enable teachers making effectively use of educational scenarios in virtual worlds such as OpenSimulator or Second Life.

Keywords: Learning analytics, CbKST, virtual worlds, OpernSimulator, data visualization.

1 Introduction

Although the hype over open, freely accessible virtual worlds is abating a little bit, there is still a significant amount of interest from teachers, particularly technology-minded ones, to adopt the rich possibilities of virtual worlds for their classroom work. In particular a strong beneficial aspect is seen in using existing communities and worlds to illustrate, to demonstrate, and to experience historical facts and events or other cultures. There is also a strong community for foreign language learning. Examples of how to find such "educationally meaningful" places offers the *SecondLife* destination guide (http://secondlife.com/destinations) ordered by topics or for *OpenSimulator* the *3D Learning Experience Services* (http://3dles.com/en).

But what are virtual worlds or virtual environments? Virtual worlds are persistent, computer-simulated, graphical environments in which individuals can appear through avatars, i.e., artificial representations of themselves. Within these worlds, people can interact with the virtual environment and with others, regardless of their physical locations, through a variety of integrated communication channels ranging from text-based chat to video communications. Virtual worlds, which were first developed in the late 1970s, are viewed as a subset of virtual reality applications that have moved through several stages of development [1].

While virtual worlds are common in (multiplayer) online games and role-playing games (such as *World of Warcraft*), virtual environments are rapidly emerging as a

A. Holzinger and G. Pasi (Eds.): HCI-KDD 2013, LNCS 7947, pp. 358–365, 2013.

complementary means to the physical world for communicating, collaborating, and organizing activities in a variety of fields, ranging from education to management. Of special relevance to the educational sector is that virtual worlds offer a rich potential for collaboration, exploration, and creativity due to characteristics such as immersion, avatar-based interaction, multi-modal collaboration, or the feeling of presence. For example, virtual environments have been used for activities such as brainstorming or iteratively and interactively creating (from objects such as clothing or buildings to interactive art and music shows to simulations of natural disasters), while simultaneously or asynchronously sharing the act of creation with other users. Many of the objects in the virtual world contain scripts that run animations of the object, play media files (such as sound or video) or otherwise enable the user to do or experience something new or perhaps even impossible. Another aspect is the fact that virtual worlds are not constrained by the real world physics (Figures 1). While this may appear to be self-evident, it is worth consideration when imagining and planning what one can do in these virtual environments; it can be just emptiness populated with objects or physically impossible representations such as being depicted as a cancer cell avatar within a human body. Undoubtedly, the *holodeck* of the 1970s is becoming reality. Therefore, virtual worlds and avatar-supported interaction seem to be a convincingly natural playground with an unbelievable diversity of tools to learn.

During the past few years, virtual worlds and virtual teams have received an increasing amount of attention by educational researchers. The results, however, are still unclear. The work of [2], for example, yielded that distinct characteristics of teaming in virtual worlds such as physical distance, device dependence, structural dynamism, or natural and cultural diversity may reduce (learning) performance. On the other hand, in a larger scale study [3] found that aspects like virtual proximity, communication modalities, and task coordination can significantly support performance. Further enabling factors coming from the research are trust, support, encouragement, freedom, challenge, goal clarity, motivation, commitment, sufficient resources and time. Further research addressed also several distinct aspects of virtual environments and teams, for example, the effects of avatar reference frames and realities, multimodality, simulation fidelity, immersion, etc. (e.g., [4, 5]). Summarizing such findings, virtual worlds may (and likely will) serve as a promising basis for educational solutions supporting and facilitating new forms of learning. This holds true for conventional schooling but also for distance education measures.

Fig. 1. Examples for virtual worlds (OpenSimulator and Second Life)

In the past years, *OpenSimulator* (http://opensimulator.org) has been increasing in popularity, particularly for serious purposes (such as education). *OpenSimulator* is an open source multi-platform, multi-user 3D application server that enables individuals and firms across the globe to customize their virtual worlds based on their technology preferences (Figure 1). The project is powered by the efforts of the community members, who devote their time and energy to the development processes. The project has a global reach and the community hosts a very diverse group of actors: independent users, freelance developers, non-profit organizations (e.g., universities), and commercial players.

2 Bringing Virtual Worlds into the Classroom

A learning scenario for *OpenSimulator,* we developed recently, is an English learning adventure. The idea is that an entire class logs into the virtual world and forms small teams. The teams then are supposed to solve a mysterious riddle: *Why is the town they find abandoned? Where are all the people?* There are only a view characters left (e.g., a drunkard or a priest) who can provide the teams with foretelling and throughout the world various hints are hidden. Accomplishing this scavenger hunt, the teams must read and listen to English texts and must understand the (often complex) meaning and must identify the main points. A highly motivational, competitive element is a reward for the team who solves the riddle first. From a pedagogical perspective, the scenario is designed around the so-called *CEFR* skills, a common specification of second language competencies (cf. http://www.cambridgeesol.org/about/standards/cefr.html).

In educational settings, usually activities and test results are stored with scores or qualitative descriptors in an overview and are included separately in a series of results or outcomes alongside other activities. This is a strongly behavior/activity-oriented approach, which most often cannot life up to the demands of 21st century education. Also educational policies in Europe presently are moving from a focus on knowledge to a focus on competency, which is reflected in revisions on curricula in the various countries. For example, in Norway, the learning goals catalogue now covers five broad areas: communication; language learning; culture, society and literature - each of which comprises sets of competencies (e.g., “the ability to read and understand the main content of texts on familiar topics”). Equally, in Austria there is a revision of the curricula in progress heading towards competence-based schooling.

In alignment with this increasing competency and ability focus of modern teaching, the virtual learning scenarios are a perfect teaching context, because they combine experiential, active, constructivist learning, with the need to directly apply the competencies in a meaningful setting, not least receiving direct feedback (e.g., by progressing with the scavenger hunt or by feedback of virtual educational entities). In that sense, a playful use of virtual worlds enables designing instructionally brilliant lesson, grounding for example on the important “First Principles of Education”, as stated by famous M. David Merrill. The principles are (1) demonstration, (2) application, (3) activation of prior knowledge, (4) integration of new information into the mental and physical world of the learner, and (4) an orientation to meaningful tasks.

The problem with a broad virtual scenario (as the adventure described above) is the massive amount of relevant educational data being produced and the inability of teachers to monitor, record, aggregate the information in a formative sense in or to generate a fair and correct model of a learner's activities and competencies – without the support of smart software solution enabling such level of "learning analytics". In other words, imagine an entire school class with 25 students; all are entering a large virtual world and disperse quickly all over this world. The teacher has almost no chance to monitor what is going on in the world, who is active or inactive, who is communicating to whom in which manner, etc.

One option is log file data. Usually, virtual environments such as *OpenSimulator* provide detailed log files for specific sessions. Unfortunately, the amount of information, stored in such log files, is massive and it takes software to analyze the log files – in an educationally meaningful way. This is the prerequisite that the activities and the performance in the virtual worlds can really contribute to a formative assessment and thus a tailored support of students.

3 Learning Analytics with ProNIFA

ProNIFA is a tool to support teachers in the assessment process that has been developed in the context of the European Next-Tell project (www.next-tell.eu). The name stands for *probabilistic non-invasive formative assessment* and, in essence, establishes a handy user interface for related data aggregation and analysis services and functions. Conceptually, the functions are based on *Competence-based Knowledge Space Theory* (CbKST), originally established by Jean-Paul Doignon and Jean-Claude Falmagne [6, 7], which is a well elaborated set-theoretic framework for addressing the relations among problems (e.g., test items). It provides a basis for structuring a domain of knowledge and for representing the knowledge based on *prerequisite relations*. While the original idea considered performance only (the behavior; for example, solving a test item), extensions of the approach introduced a separation of observable performance and latent, unobservable competencies, which determine the performance [8, 9].

CbKST assumes a finite set of more or less atomic competencies (in the sense of some well-defined, small scale descriptions of some sort of aptitude, ability, knowledge, or skill) and a prerequisite relation between those competencies. A prerequisite relation states that competency *a* (e.g., to multiply two positive integers) is a prerequisite to acquire another competency *b* (e.g., to divide two positive integers). If a person has competency *b*, we can assume that the person also has competency *a*. To account for the fact that more than one set of competences can be a prerequisite for another competency (e.g., competency *a* or *b* are a prerequisite for acquiring competency *c*), *prerequisite functions* have been introduced, relying on and/or-type relations. A person's *competence state* is described by a subset of competencies. Due to the prerequisite relations between the competencies, not all subsets are admissible competence states.

By utilizing *interpretation and representation functions* the latent competencies are mapped to a set of tasks (or test items) covering a given domain. By this means, mastering a task correctly is linked to a set of necessary competencies and, in addition, not mastering a task is linked to a set of lacking competencies. This assignment induces a *performance structure*, which is the collection of all possible *performance states*. Recent versions of the conceptual framework are based on a *probabilistic mapping* of competencies and performance indicators, accounting for making lucky guesses or careless errors. This means, mastering a task correctly provides the evidence for certain competencies and competence states with a certain probability.

ProNIFA retrieves performance data (e.g., the results of a test or the activities in a virtual environment) and updates the probabilities of the competencies and competence states in a domain. When a task is mastered, all associated competencies are increased in their probability, vice versa, failing in a task decreases the probabilities of the associated competencies. A distinct feature in the context of formative assessment is the multi-source approach. ProNIFA allows for connecting the analysis features to a broad range of sources of evidence. This refers to direct interfaces (for example to *Google Docs*) and it refers to connecting, automatically or manually, to certain log files. Using this level of connectivity, multiple sources can be merged and can contribute to a holistic analysis of learners' achievements and activity levels. The interpretation of the sources of evidence occurs depending on a-priori specified and defined conditions, heuristics, and rules, which associate sets of available and lacking competencies to achievements exhibited in the sources of evidence. The idea is to define certain conditions or states in a given environment (regardless if it is a Moodle test or a status of a problem solving process in a learning game). Examples for such conditions may be the direction, pace, and altitude a learner is flying with a space ship in an adventure game or a combination of correctly and incorrectly ticked multiple choice tasks in a regular online school test. The specification of such state can occur in multiple forms, ranging from simply listing test items and the correctness of the items to complex heuristics such as the degree to which an activity reduced the 'distance' to the solution in a problem solving process. The next step is to assign a set of competencies that can be assumed to be available and also lacking when a certain state occurs. This assumption can be weighted with the strength of the probability updates. In essence, this approach equals the conceptual framework of *micro adaptivity* as, for example, described by [10]. Figure 2 is a screenshot of ProNIFA analyzed data from a *Second Life* activity.

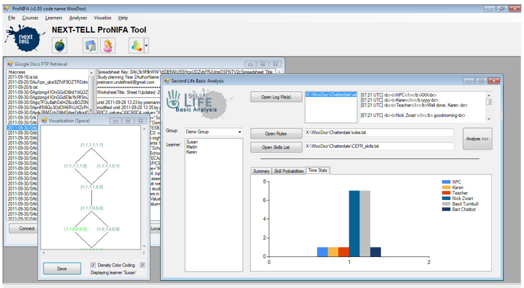

Fig. 2. Screenshot of ProNIFA

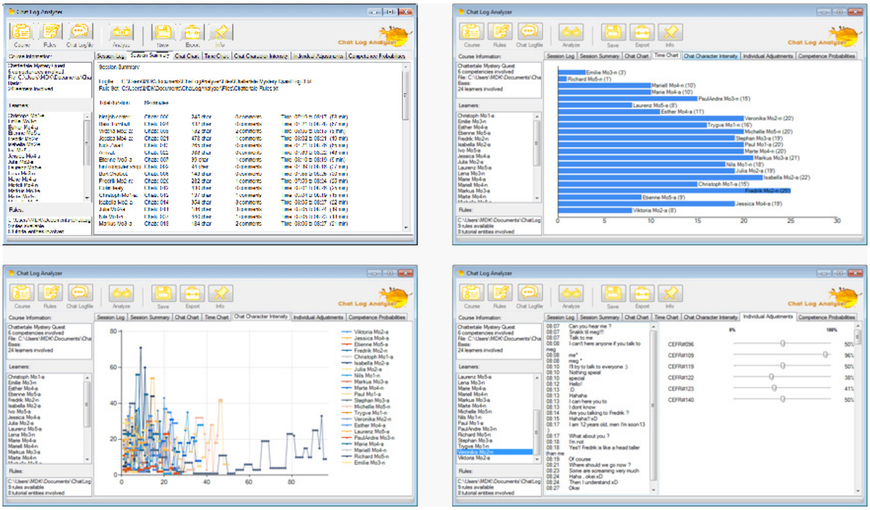

Fig. 3. Screenshots of the log file analysis module

3.1 Working with Log Files

Related to aforementioned English scenario for *OpenSimulator*, we developed a chat log analysis module (cf. Figure 3). The module allows using a course with a set of assigned students and a set of involved competencies (in this particular case 6 of aforementioned CEFR skills); in a second step the teacher may apply a set of rules to interpret the log files. The possibilities range from simple counting of certain events up to using scripting code to identify competencies. In the following box an example is given; this rule defines a time-based quest. The students have to listen to what the drunkard says; his talk finishes with "…, now go there!". The actual target is a box with a hidden letter in a hotel. If a student, indicated by "<NAME>", has understood the speech and manages to get there within 8 minutes, the system takes this as an indicator that the student has competence number 2 and consequently increases the probability in the competence model for this student by the value 0.2. In the same manner competencies could be decreased in the probability as well.

```
[Time1]
Type=1
WhoS=Drunkard
WhoE=Hotel Hint
WhatS=<NAME>, now go there!
WhatE=<NAME> arrived
Whom=Mario Wolf; Maria Wolf
Up=002
UpVal=0,2
Down=
DownVal=
```

On such basis, the entire log file with the data about all students is analyzed and the competence models for all students are updated incrementally. As a result, a teacher can access a broad range of information of the *OpenSimulator* session, the activities and learning performance. In Figure 3, some examples are shown. The upper left panel shows how a session summary is presented; this includes among other the log in and log out times, amount of chat contacts, activities, etc. for each student. The upper right images shows a bar chart visualization of the chat intensity (i.e., the number of characters typed by each student in the text chat, related to that, the lower left image shows the intensity of chat activities for all students over time. The lower right image illustrates another important feature. A teacher can access the chat text of each student extracted from the entire log file; alongside the competencies assigned to this course are listed with slider controls. These sliders indicate the competency level (in a percentage of the likelihood that this competency is available), in a way they mirror the system's adjustments of a student's competency model during the session. A teacher can now, in view of the real chat text, intervene and adjust the competency levels manually.

As mentioned above, ProNIFA operates on the basis of probability distributions over competence structures. In simple words this means that there is an order set of meaningful states a student can be in, ranging from having none of the competency to having all of them. However, due to the underlying prerequisite relations not all states are possible. For example, it is highly unlikely that a student cannot understand a written text in a foreign language and, at the same time, has the ability to understand the same text spoken by a native speaker. Since we are applying a probability distribution, the probabilities to be in one of the states sum up to 1. A teacher can access these information in form of so-called *Hasse diagrams* (an example is shown in the lower left part of Figure 2). The competence states are arrange according their structure and the assigned probabilities are displayed in form of color coding – the more salient a state appears the higher is its probability (ProNIFA allows custom color themes).

4 Conclusion

There is no doubt that learning scenarios for and in virtual worlds will be a part of classroom education. The advantages are convincing: immersion and fun, collaboration and interaction, exploration and active competence construction. The downside is that a reasonable and effective implementation of activities in the virtual world requires smart software solutions and sound approach to learning analytics to record, aggregate, analyze, visualize and store the data. ProNIFA is a tool that supports exactly those needs. Moreover, ProNIFA has a scientific, psycho-pedagogical framework in the background enabling a competence and learning performance oriented analysis of data.

Presently we are conducting classroom, studies to elucidate the usefulness of the existing features and to collect real-world demands for such learning analytics system – also beyond the focus on log files (as described here).

Acknowledgements. This project is supported by the European Community (EC) under the Information Society Technology priority of the 7th Framework Programme for R&D under contract no 258114 NEXT-TELL. This document does not represent the opinion of the EC and the EC is not responsible for any use that might be made of its content.

References

1. Dionisio, J., Burns, W., Gilbert, R.: 3D virtual worlds and the Metaverse: Current status and futre possibilities (2011) (unpublished manuscript)
2. Gibbs, C.B., Gibson, J.L.: Unpacking the concept of virtuality: The effects of geographic dispersion, electronic dependence, dynamic structure, and national diversity on team innovation. Administrative. Science Quarterly 51, 451–495 (2006)
3. Kratzer, J., Leenders, R.T., Engelen, J.M.: Managing creative team performance in virtual environments: an empirical study in 44 R&D teams. Technovation 26, 42–49 (2006)
4. Larach, D., Cabra, J.: Creative problem solving in Second Life: An action research study. Creativity and Innovation Management 19, 167–179 (2010)
5. Ward, T.B., Sonneborn, M.S.: Creative expression in virtual worlds: Imitation, imagination, and individualized collaboration. Psychology of Popular Media Culture 1, 32–47 (2011)
6. Falmagne, J., Cosyn, E., Doignon, J., Thiéry, N.: The Assessment of Knowledge. Theory and in Practice (2003), `http://www.scribd.com/doc/3155044/Science-Behind-ALEKS`
7. Doignon, J., Falmagne, J.: Knowledge Spaces. Springer, Berlin (1999)
8. Albert, D., Lukas, J. (eds.): Knowledge spaces: Theories, empirical research, and applications. Lawrence Erlbaum Associates, Mahwah (1999)
9. Korossy, K.: Modelling knowledge as competence and performance. In: Albert, D., Lukas, J. (eds.) Knowledge Spaces: Theories, Empirical Research, and Applications, pp. 103–132. Lawrence Erlbaum Associates, Mahwah (1999)
10. Kickmeier-Rust, M.D., Albert, D.: Micro adaptivity: Protecting immersion in didactically adaptive digital educational games. Journal of Computer Assisted Learning 26, 95–105 (2011)