Modeling the Acquisition of Fluent Skill in Educational Action Games

Ryan S.J.d. Baker, M.P. Jacob Habgood, Shaaron E. Ainsworth

Learning Sciences Research Institute
University of Nottingham
Nottingham, UK
ryan@educationaldatamining.org, jake@gamelearning.net, Shaaron.Ainsworth@nottingham.ac.uk

Albert T. Corbett

Human-Computer Interaction Institute, Carnegie Mellon University,
Pittsburgh, PA, USA
corbett@cmu.edu

Abstract. There has been increasing interest in using games for education, but little investigation of how to model student learning within games [cf. 6]. We investigate how existing techniques for modeling the acquisition of fluent skill can be adapted to the context of an educational action game, Zombie Division. We discuss why this adaptation is necessarily different for educational action games than for other types of games, such as turn-based games. We demonstrate that gain in accuracy over time is straightforward to model using exponential learning curves, but that models of gain in speed over time must also take gameplay learning into account.

1 Introduction

Over the last decades, a number of very effective techniques have been developed for modeling student learning within interactive learning environments. Bayesian Knowledge-Tracing [3] and Bayes Nets [6] have proven to be very effective at modeling student knowledge at a specific point in time. Another technique, empirical learning curves [cf. 1,5] have proven successful for assessing students' gains in both accuracy and speed over time, as they use a learning environment.

These techniques have been generally very successful at modeling knowledge and learning within the environments where they have been used, and have contributed to making these environments more educationally effective. However, there are many types of environments where these techniques are underused – in particular educational games [6]. Almost since the advent of the personal computer, educational games have been an important part of many students' educational experiences. It has been hypothesized by many researchers that games have the potential to make education more fun, and to improve student learning by improving student engagement [cf. 7]. Yet the development of educational games has generally not benefited from the analytical tools that have been used to study and improve the educational effectiveness of other types of learning environments, such as intelligent tutoring systems [cf. 1,2,3,5].

One important and popular type of educational game is the educational action game. Educational action games incorporate educational material into fast-paced game environments where the student must respond quickly to continual challenges. Unlike many other forms of interactive learning environments, and turn-based educational games (studied in [6]), educational action games offer little time for reflection, at least during main gameplay. Because educational action games do not offer time for reflection, it has been suggested that they are more appropriate for building skill fluency (i.e. speed and accuracy at exercising a skill) than for the acquisition of new and complex concepts [cf. 9].

In intelligent tutoring systems, empirical learning curves have been found to be an appropriate method for assessing gain in speed and accuracy [cf. 1,2,5]. Two challenges will need to be surmounted, however, in order to use exponential learning curves for fluency assessment in educational action games.

The first challenge is that the relationship between performance and knowledge is more complex in games than tutors. Unlike tutors, educational games generally do not attempt to explicitly make student thinking visible and communicate domain goal structure [cf. 1], design goals which result in environments where it is comparatively easy to assess student knowledge. However, Manske and Conati [6] have successfully developed Bayes Nets which can make appropriate assessments of knowledge in turn-based educational games. In this paper, we will discuss what additional challenges to assessing knowledge are present within educational action games.

A second challenge, particularly important within educational action games, is that some portion of students' gain in speed is likely due to learning how to play the game, rather than domain learning. In this paper, we will investigate how gameplay learning affects our ability to assess the development of fluent skill.

Within this paper, we will investigate how these challenges can be addressed, so that student fluency gain can be accurately modeled within an educational action game, Zombie Division.

1.1 Zombie Division

Zombie Division, shown in Figure 1, is an educational game designed to help elementary school students learn about division [cf. 4]. Zombie Division is at its core a third-person action game, though it also has adventure-game elements.

Within Zombie Division, the player is a hero from Ancient Greece, who must defeat skeletal enemies in hand-to-hand combat in order to progress. Each skeleton has a number on its chest. The player has a set of weapons, each of which corresponds to a divisor number. Each weapon is linked to a key on the keyboard – the 2 weapon is used by pressing the "F2" key, the 3 weapon is used by pressing the "F3" key, and so on. If the player attacks (attempts to divide) a skeleton by a number which divides that skeleton's number (i.e. skeleton modulus weapon = 0), the skeleton dies. If the player attacks (attempts to divide) a skeleton using a number which is not a divisor of the skeleton's number, the skeleton counter-attacks, causing the player to lose health.

As the player proceeds from level to level of the game, his or her weapons (set of potential divisors) change, requiring the player to use different divisors to divide the same skeleton at different times (for example, needing to use 2 or 4 on different levels to divide a 32 skeleton). Some skeletons are not divisible by any of the student's weapons and must be avoided. The mathematical skills involved in Zombie Division (and which a student will hopefully know more about after playing Zombie Division) are, therefore, being able to determine whether a number is divisible by 2 (e.g. even), 3, or 5, and being able to determine whether a number is divisible by 4, 6, 8, or 10 when small divisors are not available.

Beyond the mathematical features of the gameplay, there are also aspects to the game which are included purely to support enjoyable and challenging gameplay: some skeletons move from place to place, other skeletons hold special keys that enable the student to move on to new game regions, some skeletons pursue the student, and some skeletons (increasingly on higher levels) attack spontaneously if the player delays. Hence, Zombie Division is designed for the joint purposes of teaching mathematics and providing the student with a fun experience.

Later in the paper, we will also discuss data from an alternate ("extrinsic") version of Zombie Division. In the "intrinsic" version discussed above, the mathematical content of Zombie Division is integrated into the gameplay. The "extrinsic" version, has the same mathematical content and the same gameplay, but these two components of the student's experience are separated. Mathematical problems are given at the completion of each game level, and the student plays a game which is identical to the intrinsic version of the game described above, but where the mathematical content has been removed. The same keys on the keyboard are used to kill skeletons, but no divisors are associated with those keys. Instead of having numbers on their chests, the skeletons have pictures of the weapons that can kill them. The student encounters exactly the same skeletons at the same times and locations in each version of Zombie Division; the only difference is the conceptual meaning of the key the student must press to kill each skeleton.

In this paper, we will focus predominantly on studying learning and gameplay in the intrinsic condition of Zombie Division, where these two components are mixed (as in most educational action games) – however, we will in some cases consider evidence from the extrinsic condition in order to better understand the pattern of student performance in the intrinsic condition.

The data we will discuss is drawn from four classes in a large primary school in a low-income area on the outskirts of a medium-size city in northern England. The school has an average number of students with special educational needs, but with significantly below-average scores on national assessments. 17 students used the intrinsic condition of Zombie Division; 18 students used the extrinsic condition of Zombie Division – two additional students were removed from each condition for missing the post-test. Each student used Zombie Division for 135 total minutes of class-time, across 6 class days spread across 4 weeks. Log files were used to distill measures of students' learning and performance as they used Zombie Division.

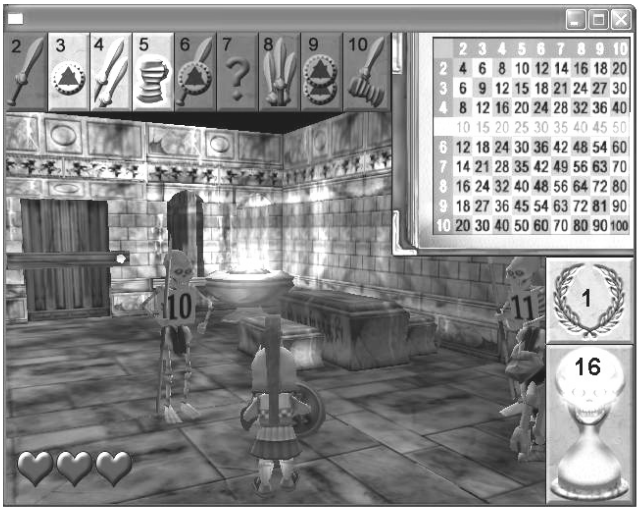

Fig. 1. Zombie Division (intrinsic version)

2 Accuracy-Based Models of Student Learning Within Zombie Division

In this section, we will study how to adapt empirical learning curves [cf. 1,2,5] to the context of Zombie Division in order to study students' gain in accuracy over time.

2.1 Mapping Game Actions to Evidence on Learning

In order to plot learning curves, we need to map the student actions and their consequences within Zombie Division to a conceptual framework which allows us to define opportunities to practice a mathematical skill and whether a student has correctly demonstrated the skill or not. Such a conceptual framework has been created for intelligent tutoring systems [1] and, more recently, for turn-based educational games [6].

One challenge that does not occur in intelligent tutoring systems and is substantially less common in turn-based educational games is that not all "errors" from the perspective of the game give evidence about the student's mathematical skill. For example, if a student walks into a skeleton and does not attack, the skeleton attacks the student and the student loses health; though this is an error in gameplay (and results in negative consequences within the game), it gives little evidence on the student's mathematical knowledge.

However, many events within the game do give information on the student's mathematical knowledge. Attacking a skeleton and killing it (for example, using the "2" weapon to divide a "26" skeleton), is evidence that the student knows how to determine whether a number is divisible by 2. Correspondingly, unsuccessfully attacking a specific skeleton (for example, trying to use the "2" weapon to divide a

“15” skeleton), is evidence that the student does not know how to determine whether a number is divisible by 2. In addition, avoiding certain actions may also give information about the student’s knowledge. If a student flees from a skeleton (defined as leaving the room the skeleton is in) which he/she could not have killed (the student has “2”, “3”, and “5” weapons but the skeleton is “49”), there is evidence that the student knows how to determine if a number is divisible by 2,3, and 5; if a student flees from a skeleton which he/she could have killed, on the other hand (the student has a “2” weapon and flees from a “16” skeleton), there is evidence that the student does not know how to determine if a number is divisible by 2.

In the analyses that follow, we consider any given skeleton as a single opportunity for a student to demonstrate a skill [cf. 3]: multiple attempts to kill a skeleton with different weapons may indicate process of elimination rather than mathematical knowledge.

2.2 Details of Analysis

In the analyses that will follow, we focus on four mathematical skills: the student’s ability to determine if a number is divisible by 2, 3, 4, and 5. 2, 3, and 5 are all prime numbers, introduced as weapons/divisors early in the game, and are thus reasonably straightforward to analyze. 10 is also introduced as a weapon/divisor early in the game, but 10 is co-present with 2 or 5 in all early opportunities to use 10. This creates considerable risk of bias: with both 2 and 10 available, a student having difficulty deciding if an even number was divisible by 10 can automatically revert to using 2 – hence, a number of situations where the student did not know how to divide by 10 could be missed during analysis. 4, on the other hand, is introduced as a weapon/divisor later in the game, when 2 has been removed as a weapon/divisor. Hence, students’ ability to determining if a number is divisible by 4 is not occluded by the presence of 2. 6, 8, and 9 also occur in the game as divisors, but only on later levels and thus with insufficient frequency to analyze.

Since time was controlled in this study, some students are able to complete more of Zombie Division than others. Hence, some students will encounter more skeletons than others. Since students who get further in Zombie Division and encounter more skeletons are likely to be better at mathematics, it would introduce bias to use all data from all student actions in our analyses. Hence, we set a cut-off, and do not analyze opportunities to practice a mathematical skill which were reached by less than half of the students (in practice, this gives us from 50 actions per student for the divisible-by-2 skill, to 19 actions per student for the divisible-by-4 skill).

2.3 Results

Graphs showing students’ accuracy over time at using some of the mathematical skills needed to play Zombie Division are shown in Figure 2. Each of these graphs shows the average performance at each opportunity to practice each skill, with the best-fitting exponential curve overlaid on each graph. An exponential learning curve will fit the data if students have a fairly high error rate at the beginning, improve fairly rapidly, and show slowing improvement over time.

The skill of determining whether a number is divisible by 2, shown in the top-left graph of Figure 2, appears to fit this pattern very well. The best-fitting exponential function to this data achieves a very healthy r^2 of 0.52. The skill of determining whether a number is divisible by 5, shown in the top-right graph of Figure 2, also appears to fit this pattern very well. The best-fitting exponential function to this data achieves a respectable r^2 of 0.32. The skill of determining whether a number is divisible by 4, shown in the bottom-right graph of Figure 2, also appears to fit this pattern, though the best-fitting exponential function to this data achieves a relatively low r^2 of 0.09. At first glance, it appears that the leap in difficulty at the seventh opportunity to divide by 4 may indicate that two different skills are being combined together. However, many skeletons can be killed by multiple weapons and students often encounter skeletons in different orders, so it is not immediately possible to interpret a spike in difficulty as a second skill being encountered, unlike in learning curve analyses of intelligent tutors [cf. 1]. In fact, that point represents 12 different skeletons encountered by 19 students, with no two errors made on the same skeleton.

The skill of determining whether a number is divisible by 3, shown in the bottom-left graph of Figure 2, does not appear to fit an exponential curve. The best-fitting exponential function to this data achieves an r^2 under 0.01, and even points in the wrong direction, going very slightly up over time. This suggests that students are having more difficulty determining if a number is divisible by 3 than if a number is divisible by 2,4, and 5 – interestingly, division by 3 had not yet been discussed in class before the students used Zombie Division, whereas the other divisors had been. This serves as a valuable reminder that fluency-building learning environments will probably be most effective if used after appropriate conceptual instruction.

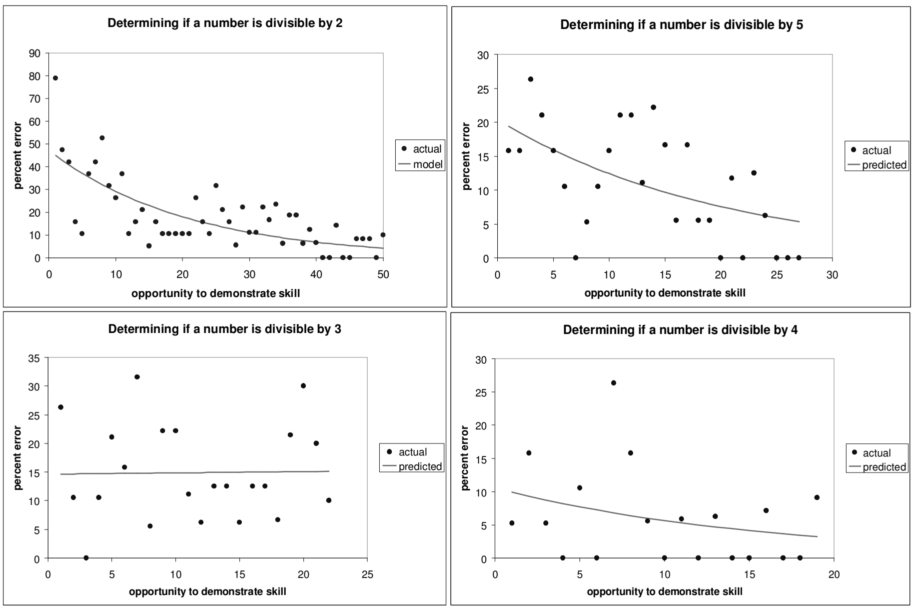

Fig. 2. Students' change in percent correct over time, for specific mathematical skills

3 Time-Based Learning Models of Student Learning Within Zombie Division

In this section, we investigate whether students learn to use mathematical skills with greater speed during the time they use Zombie Division, again using empirical learning curves. Empirical learning curves have been used successfully to model gain in speed in intelligent tutoring systems [cf. 1,5]. However, the relationship between speed and learning is different in educational action games than in intelligent tutors. Intelligent tutors generally involve interface actions which are common to most computer applications (such as clicking on a blank and typing a number, or pointing and clicking). It is reasonable to assume that most students will have experience applying these basic interface skills; and therefore most of the speed gains a student has while using an intelligent tutor should involve the relevant domain skills, not gains in speed in interacting with the user interface.

By contrast, an educational action game like Zombie Division involves several novel interface skills which must be learned, such as how to move the correct distance away from a skeleton, and how to use each divisor. While many students will have had considerable prior gaming experience, Zombie Division's gameplay will be subtly different from games students have played in the past (for example: how much movement is obtained by pressing a movement key, and which keys correspond to different divisors). Therefore, some of each student's speed gains while using Zombie Division may be attributable to learning gameplay skills instead of domain skills.

In this section, we will first present an analysis which ignores interface and gameplay learning. We will then explicitly account for interface and gameplay learning, and show how accounting for gameplay learning affects the results.

3.1 Details of Analysis

In the analyses which follow, we will consider only a subset of actions within Zombie Division. Specifically, we will analyze the time a student takes to attack a skeleton with an appropriate weapon, on their first attempt to respond to that skeleton. We eliminate errors (attacking with the wrong number) from consideration, since these actions will not be representative of the student's gain in efficiency at using correct knowledge over time. We also eliminate second and subsequent attempts to respond to a skeleton, since they are likely to involve error-correction instead of simply exercising a known skill. Finally, we eliminate fleeing actions from consideration, since the amount of time required to flee may be governed primarily by the size of the room and presence of other skeletons in the room .

In addition, as in the previous section, we do not analyze opportunities to practice a mathematical skill which were reached by less than half of the students. For brevity, we will focus on the skill of determining whether a number is divisible by 2; however, the pattern we will show in this section is the same pattern as is found when the skills of dividing by 4 and 5 are analyzed.

3.2 Results

A graph showing students' accuracy over time at determining whether a number is divisible by 2 is shown on the left of Figure 3. This graph shows the average time taken at each opportunity to practice each skill, with the best-fitting exponential curve overlaid. An exponential learning curve will fit the data if students work fairly slowly at the beginning, improve fairly rapidly, and show slowing improvement over time.

The skill of determining whether a number is divisible by 2, shown in the left graph of Figure 3, appears to fit this pattern very well. The best-fitting exponential function to this data achieves a reasonably high r^2 of 0.23. Hence, using this approach suggests that students are getting faster at dividing by 2 over time, and therefore that they are gaining fluency in this skill.

However, it is not clear from this approach whether the students' gain in fluency is a gain in fluency with mathematics or a gain in fluency at playing Zombie Division. In many cases, this distinction would be difficult to tease apart. However, in this case, data from the extrinsic condition can be used. As discussed earlier, the two conditions have identical gameplay but in the extrinsic condition the mathematics is given separately. The extrinsic condition data can therefore be used to determine how much of the speed-up seen in the intrinsic condition is explained by gameplay learning, and therefore how much domain learning occurred. This will in turn give evidence on the appropriateness of computing time learning curves which do not account for gameplay.

A graph showing students' accuracy over time at killing the skeletons using F2 in the extrinsic condition (equivalent to dividing by 2 in the intrinsic condition) is shown on the right of Figure 3. This graph shows the average time at each opportunity to practice each skill, with the best-fitting exponential curve overlaid. The best-fitting exponential function to this data achieves only a modest r^2 of 0.05, but interestingly, the best-fitting functions have a fairly similar appearance between conditions.

We can now use this gameplay-only curve to calculate whether there is mathematics learning occurring in the intrinsic condition. If there is both gameplay and mathematics learning in the intrinsic condition, the learning curve in the intrinsic condition should actually be a composite of two curves: a gameplay learning curve, and a mathematics learning curve. The gameplay learning curve derived in the

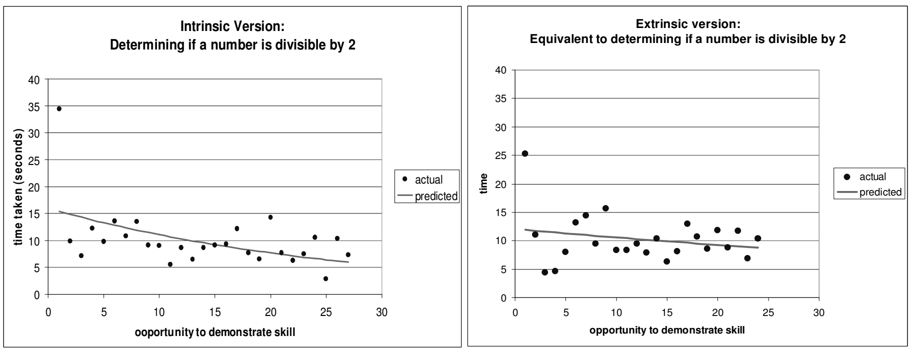

Fig. 3. Students' change in speed at exercising a skill over time.

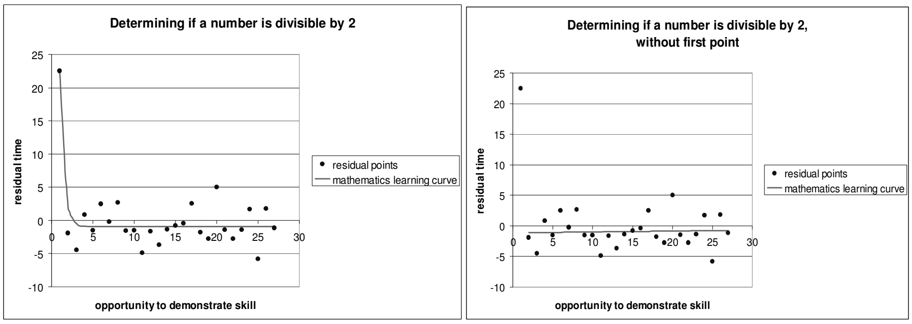

Fig. 4. Students' mathematics learning in the intrinsic condition, with gameplay factored out. The graph on the left includes the first point in curve calculation; the graph on the right omits that point.

extrinsic condition should be a reasonably accurate estimate of the gameplay learning curve in the intrinsic condition, since gameplay is identical across conditions. Hence, we subtract the extrinsic gameplay learning curve from the data in the intrinsic condition. If the resultant points still fit an exponential learning curve, we can be certain that mathematics learning actually was occurring in the intrinsic condition.

The resultant residual data points, and learning curve, are shown on the left side of Figure 4. This curve achieves a spectacular r^2 of 0.77 – at first glance suggesting that considerable domain learning is indeed occurring. However, note that the curve is completely flat after the earliest opportunities to practice the skill, and that the first point is a major outlier. If we eliminate the first point from consideration when fitting the curve, the slope flips around (see the right side of Figure 4), going upwards (though with r^2 under 0.01). Hence, it appears that non-gameplay learning is occurring in the intrinsic condition, but only between the first and second opportunities to practice the skill. The additional learning does not appear to be a gain in mathematics fluency over time. It may instead be the student learning to apply his or her existing mathematical knowledge within Zombie Division.

Hence, when we look at the overall picture, it does not appear that students are gaining fluency, at least in terms of speed, at deciding if a number is divisible by 2, while playing Zombie Division. At minimum, if the students are gaining fluency, the effect is much smaller and more variable than the effects of gameplay learning. This suggests that plotting learning curves of student speed, without taking gameplay into account, is not an appropriate way to model fluency gain in educational action games.

4 Conclusions

In this paper, we have analyzed two ways of studying student gains in fluency over time, within an educational action game: studying gains in accuracy over time, and studying gains in speed over time. We have found that studying gains in speed over time is not straightforward to do correctly; in particular, different results are obtained, depending on whether or not gameplay is explicitly accounted for. Because of this, simply computing gain in speed over time, without accounting for gameplay, does not appear to be an appropriate way to model learning in educational action games. In the absence of a measure of gameplay, an alternate and probably more reliable way to assess whether students gain speed at applying a skill is to time students' responses on the pre-test and post-test.

It does appear, though, that existing methods for modeling gain in accuracy over time are appropriate for use in educational action games, with only minor modifications. This will make it possible to quickly and effectively determine which skills students gain and fail to gain while using an educational action game, solely from their behavior within the system. In the case of Zombie Division, we see that students successfully gained fluency in determining if a number is divisible by 2,4, or 5, but that students did not gain fluency in determining if a number is divisible by 3; in developing future versions of Zombie Division, we now know that this skill will require extra support and scaffolding. Generally, analysis of accuracy learning curves has been found to be a useful technique for making formative assessments which can be used to drive rapid re-design and improvement of intelligent tutoring systems [cf. 2]. The results presented here suggest that, properly used, this technique will be useful for formative assessment in educational action games as well.

Acknowledgements. We would like to thank Jenny Habgood, Lucy Button and Lizzie Evans for their assistance in organizing and conducting the studies reported here. The work presented in this paper was funded by research fellowships from the Learning Sciences Research Institute, at the University of Nottingham.

References

1. Anderson, J.R., Conrad, F.G., Corbett, A.T. Skill Acquisition and the LISP Tutor. *Cognitive Science,* 13 (1989), 467-505.
2. Beck, J.E. Using learning decomposition to analyze student fluency development. *Proceedings of the workshop on Educational Data Mining at the 8th International Conference on Intelligent Tutoring Systems* (2006)*,* 21-28.
3. Corbett, A.T., Anderson, J.R. Knowledge Tracing: Modeling the Acquisition of Procedural Knowledge. *User Modeling and User-Adapted Interaction,* 4 (1995), 253-278.
4. Habgood, M.P.J. Zombie Division: Intrinsic Integration in Digital Learning Games. *Proceedings of the 2005 Human Centered Technology Workshop.*
5. Martin, B., Koedinger, K., Mitrovic, A., Mathan, S. On Using Learning Curves to Evaluate ITS. *Proceedings of the 12th International Conference on Artificial Intelligence in Education (AIED-2005),* 419-426.
6. Manske, M., Conati, C. Modeling Learning in Educational Games. *Proceedings of the 12th International Conference on Artificial Intelligence in Education (AIED-2005),* 411-418.
7. Prensky, M. Digital game-based learning. *Computers in Entertainment,* 1 (1) (2003), 1-4.
8. Repenning, A., Clayton, L. Playing a game: the ecology of designing, building, and testing games as educational activities. *Proceedings of ED-MEDIA: World Conference on Educational Multimedia, Hypermedia, and Telecommunications (2005).*