A Bayesian framework for adaptive learning in educational games

April 1, 2015

1 Introduction

As technology advances, educational games and adaptive learning methods are increasingly implemented to enhance education. Adaptive learning promises to help educators support the wide variety of learning needs and goals in our current educational system, and it has been shown that personalization improves learning (e.g., Connor, Morrison, Fishman, Schatschneider, & Underwood, 2007). In addition, there is ample evidence that game-based learning improves engagement (Steinkuehler, Squire, & Barab, 2012).

An educational game can be adaptive in many ways. Learners may be assessed initially and then assigned to a fixed learning path, or their path may change after each “level” of the game. At the extreme is real time adaptivity, where the game adapts every time a learner completes a challenge. This high level of adaptivity is intended to keep the game at a “Goldilocks” level of difficulty for the learner (neither too easy nor too hard) to optimize engagement and learning. Real time adaptivity of educational games, when combined with a psychometric framework, can help educational games meet two goals: (a) adapt to the level of the learner to keep the game motivating and challenging (b) involve parents and teachers in children’s digital learning by giving them feedback on how a learner is doing.

One way to achieve the continuous assessment required for real time adaptivity is to use item response theory (IRT) models as developed for computerized adaptive testing (CAT, e.g., van der Linden & Glas, 2010; Wainer et al., 2000). IRT provides the tools to estimate item (i.e., game challenge) difficulty and learner ability on a single scale. Hence, the difficulty of the next game challenge can easily be matched to the estimated ability level of the learner. In addition, the estimated ability level can be used to report on the learner’s progress to parents or teachers.

Games differ in important ways from tests, however. Tests consist of a considerable number of test items, administered at one specific point in time. They generate a lot of information about the ability of a learner at that moment, resulting in a highly accurate measurement. Educational games, however, typically consist
of multiple short interactions with a few responses each. In addition, games encourage learning over time, with the result that ability is assumed to change frequently. These features cause trouble for traditional computerized adaptive testing models when they are applied to educational games: There is not enough information to yield an accurate assessment of a learner’s ability within one gameplay session, but taking all measurements together to estimate ability leaves no room to monitor growth.

Recently, several approaches have been developed to address these concerns in computerized adaptive learning (CAL; e.g., Brinkhuis, 2014; Brinkhuis & Maris, 2009; Eggen, 2012; Klinkenberg, Straatemeier & van der Maas, 2011; Wauters, Desmet & Van Den Noortgate, 2010; 2011). One approach is to use the Elo rating system (Glickman, 1995), originally designed to rate chess players. In this system items and learners are paired with each other, and the change in both learner ability score and item parameter is based on the difference between the expected and predicted outcome, sometimes weighted by the uncertainty about the previous outcome (e.g. Klinkenberg, Straatemeier & van der Maas, 2011).

What is missing from these approaches, however, is a flexible way to combine previous information about a learner from other sources into the estimation of ability, enabling the game to start at an engaging level right away. In addition, uncertainty about the estimated scores is not always well defined, which is desirable in reports to teachers and parents.

Our approach is to use Bayesian IRT models (for more on Bayesian IRT models, see Fox, 2010 and van der Linden & Glas, 2010) for computer adaptive learning. In Bayesian IRT models, prior knowledge about a person’s ability can be considered in the estimation process, so previous gameplay results, or other information, can be taken into account when learners begin a subsequent play session. To monitor change in performance over time, this prior information should be weighted less heavily when some time has passed between play sessions than if the sessions had been played in a single assessment situation. In this way, the effect of short “test” length can be countered by incorporating additional information, while still allowing measures of ability to change over time.

We will present two different models taking into account time dependency through prior distributions. The first model considers each play session as a separate measurement moment. The second model, drawing inspiration from various fields of science, will weigh the influence of previous results as a (decay) function of the time passed.

Two applications of the presented models in games for preschoolers will be presented. The first application is the implementation of Bayesian IRT models to drive an adaptive game on patterns and shapes for preschoolers for the iPad, in which pretest scores were used to determine the starting level, and in which
learners received two or three items per play session (the maximum number for preschoolers) with days or even weeks in between play sessions. We will present some results on the effectiveness of the application, the effects on engagement and learning, and use the data to illustrate the different time decay methods. The second application is the implementation of an ability estimation procedure in a game for preschoolers covering many different early learning dimensions. The ability estimates are used to give feedback to parents in an accompanying app. The parents receive a notification when their child reaches a “milestone” on the ability scale with enough certainty.

Many possible avenues for extension will be outlined in the discussion.

2 Bayesian IRT models for computer adaptive learning

The main advantage of using Bayesian models for computer adaptive learning is the possibility to use prior information for the estimation of parameters. This makes it possible, for example, to use initial guesses about item parameters to speed up their calibration, and to update the item parameter estimates after each new learner comes in. The focus in this paper, however, will be on the advantages for ability estimation when ability is changing and the information about ability obtained at each individual measurement occasion is scarce. We will therefore, for the time being, consider the item parameters of the items to be known from previous calibration and stable over time (even though these assumptions are violated in many practical applications, which is content for its own presentation).

Games will be considered in which the learner encounters challenges (items) to which he can respond correctly or incorrectly. The responses in this game can be modeled with a simple 1 parameter logistic IRT model, where the probability of a learner answering an item correctly at moment \( t \) can be modeled as a function of the learner ability at that moment \( \theta_t \) and the item difficulty of the specific item \( b \):

\[
P(Y_t = 1 \mid \theta_t, b) = \frac{e^{\theta_t - b}}{1 + e^{\theta_t - b}}.
\]

Assume an answer \( Y_t \) is observed and we want to make an inference about the learner ability \( \theta_t \), given that the value of \( b \) is known. To be able to do Bayesian inference, we first have to choose a prior distribution for the ability \( \theta_t \). If nothing is known about the learner, it makes sense to use the assumed ability distribution in the population, often a standard normal distribution \( N(0, 1) \).

The posterior distribution of the ability parameter \( \theta_t \) is proportional to the likelihood of observing \( Y_t \)
given \( \theta \) multiplied by the prior density of \( \theta \):

\[
p(\theta_t|Y_t) \propto P(Y_t \mid \theta_t, b)p(\theta_t)
\]

Because a normal prior distribution is not conjugate to the logistic likelihood function of the IRT model, the resulting posterior distribution is not normal. We choose to approximate this posterior distribution as a normal distribution with mean and standard deviation found via integration by quadrature (Bock & Mislevy, 1982; Bock & Aitkin, 1981). For most parameter combinations, this results in an adequately low approximation error. We are working on methods to make item-wise updating of the ability distribution more accurate and more efficient (which is outside the scope of this presentation).

Figure 1 shows how the posterior distribution for \( \theta_t \) looks if a standard normal prior distribution is used with one observation \( Y_t \), for a correct and incorrect answer to an item with item difficulty \( b = .5 \). As can be inferred from the relative width of the posterior distributions, one observation doesn’t increase the certainty about the ability much.

When there is information about the ability of a learner beforehand, the uninformative prior distribution \( N(0, 1) \) can be replaced by a more informative prior distribution. For example, when it is known that different age groups have different ability distributions, this can be represented in the prior distributions for the learners. When a pretest has been taken by the learners, the prior mean can be based on the score on the pretest. This situation is described in paragraph 4.

Figure 1: Posterior distributions to for an item with difficulty \( b = .5 \).
More often, however, there will be previous play sessions of the same game, or previous items in the same play session, from which information about ability can be inferred. In a computer adaptive test, the posterior after the previous item can be used as the prior for the next item. In this way, the information acquired from sequential test items accumulates naturally to make the posterior distribution of ability at that time point more and more precise (Figure 2).

In game situations, however, there are often not enough items to make such a precise estimate for individual play sessions. But while it is desirable to take information from previous play moments into account, room is needed to detect learner ability growth, or decline, in between play sessions. Accumulating information of sequential items over play sessions as shown above will soon converge to an estimate with quite high precision, after which the fluctuations in estimates possible are very restricted. The next paragraph will outline two strategies to handle this problem.

3 Time dependency in computer adaptive learning models

When assessing the predictiveness of the ability score \( t - 1 \) for the ability at time \( t \), two factors are important: How certain were we about the previous ability, and how much time has passed since the previous assessment? This corresponds to the uncertainty factors identified by Glickman (1995) in the Elo rating system for weighting previous information: frequency and recency. Klinkenberg, Straatemeier and van der Maas (2011)
transformed these ideas into an uncertainty measure, which linearly decreases with the number of items played up to a maximum of 40 and increases with the number of days which have passed since the last play session up to a maximum of 30.

Using Bayesian inference, the uncertainty about the previous ability is accumulated in the width of the posterior distribution of this ability, as represented by the posterior variance. When many items have been played, the posterior distribution is small and the weight it will receive relative to the new information that comes in will be high. We will use this principle to handle the uncertainty about the previous estimate. To accommodate for recency, we will propose two strategies: 1) to reset the prior variance after each sequential play session 2) to combine the posterior after the previously played item with the original uninformative prior, with weights derived from a time decay function.

3.1 Resetting the prior variance after each session

This model assumes that each time a learner starts a new game, the best guess for the learner’s current ability will be the estimated ability from the previous game, the posterior mean \( \mu_{\theta_{t-1}} \). Within the same session, the posterior variance after the previous item \( \sigma_{\theta_{t-1}}^2 \) is taken as the prior variance for the next item. To allow for learning, or forgetting, in between sessions, the prior variance is reset to a typical posterior variance after one item administration \( \sigma_1^2 \) (e.g. .65), making the prior more informative than the original uninformative prior, but leaving ample room for change between play sessions. Basically, each play session is taken as different measurement occasion.

If \( t \) and \( t-1 \) are in the same play session:

\[
\theta_t \sim N(\mu_{\theta_{t-1}}, \sigma_{\theta_{t-1}}^2)
\]

If \( t \) and \( t-1 \) are not the same play session

\[
\theta_t \sim N(\mu_{\theta_{t-1}}, \sigma_1^2)
\]

3.2 Weighted prior with time decay function

The aim is to combine the previous score with the new information in such a way that the previous score will have more weight if we are more certain about this score, and in which the weight of the previous score decreases when more time has passed. We drew inspiration from models used in learning systems,
signal processing, econometric forecasting and memory and decision making to come up with a model which combines these properties in a natural way. This model will assume that all relevant information about previous gameplay is contained in the posterior distribution of $\theta$ at time $t - 1$.

First, we write our IRT model as a combination between the model of the outcome $Y_t$ given the ability $\theta_t$ and the model of the of the latent ability, in a similar way to what is common in state-space models (e.g. Kalman filters). The dependent variable $Y_t$ depends on the ”state” $\theta_t$ and the item difficulty $b$ of the item administered at time $t$ through the item response function in equation 1. Then $\theta_t$ can be modeled by a ”transition equation” (which is similar to an autoregressive function), where $\theta_t$ is modeled as a linear function of the predicted score at time $t - 1$, in our case the posterior mean $\mu_{\theta_{t-1}}$, and a ”noise” component representing external factors which could influence $\theta_t$. This transition equation can be implemented by adjusting the normal prior for $\theta_t$ as follows:

$$\theta_t \sim N(g_t \mu_{\theta_{t-1}}, \sigma_t^2),$$

where

$$\sigma_t^2 \equiv g_t^2 \sigma_{\theta_{t-1}}^2 + \sigma_{\eta_t}^2.$$  

A time dependent transition weight factor $g_t$ and a variance factor $\sigma_{\eta_t}^2$ are added to the normal prior. The variance factor $\sigma_{\eta_t}^2$ is often called the noise factor in signal processing, introducing external noise to the process.

Now what do we expect the transition process to look like when we consider ability in educational games? We expect that in between play sessions, learners can either improve or decline in their ability. Hence, over time, we are less certain of the predictive value of the performance of a learner at the previous play session, and we want to add more noise to the process to enable forgetting or learning in between play sessions.

We therefore propose to model the prior for $\theta_t$ as a linear combination of the posterior distribution at $t - 1$ with an uninformative normal distribution which holds no information about the data (noise) and is independent of the posterior. Because this is a linear combination of two (approximately) normal distributions, this will result in a normal distribution. It makes sense to set the noise distribution equal to the uninformative prior we assume when we have no information about the data (e.g. $N(0,1)$). This will result in a ”back to the prior” forgetting strategy, as described by Van Vaerenberg, Lazaro-Gredilla and Santamaria (2012), where over time the prior will go back to the original prior, which indicates the state
when we don’t know anything about the learner. The resulting prior would be rewritten as a combination between the posterior at time \( t - 1 \) and a normal \( N(\mu_0, \sigma_0^2) \) distribution, with weight \( g_t \) given to the posterior at time \( t - 1 \) and weight \( 1 - g_t \) given to the uninformative prior distribution:

\[
\theta_t \sim N(g_t \mu_{θ_{t-1}} + (1 - g_t)\mu_0, \sigma_t^2),
\]

where

\[
\sigma_t^2 \equiv g_t^2 \sigma_{θ_{t-1}}^2 + (1 - g_t^2)\sigma_0^2.
\]

The result is that, over time, the prior variance will go to the variance of the uninformative prior, but also that the prior mean will go towards the uninformative mean. This can be seen as a form of regression to the mean: over time, as we become less certain of a learner’s score, a safe bet is that it will have moved closer to the overall mean, either by forgetting or by learning in between play sessions.

How fast the prior will go back to the uninformative prior depends on the weight \( g_t \). To model this weight, we will use a time decay function, also used in (machine) learning systems, decision making processes and economic forecasting. The main principle is that older examples are less important for predicting new information than newer examples, so we want to weigh information according to the ”age” of that information (Gama et al., 2014; Klinkenberg, 2004). This is also known as ”temporal discounting” in the decision making literature. Pelanek (2014) applied several decay functions to ability scoring, although within a different modeling framework. He found that the exponential decay function and hyperbolic decay function produced similar results. For now, we will stick with the exponential decay function on \( g_t \) (we are still researching the best form of the decay function for our applications):

\[
g_t = e^{-\lambda t},
\]

where \( t_\Delta \) is the time passed since the previous play item was played.

The parameter \( \lambda \) can be anywhere between 0 and 1. Figure 3 shows the weight \( g_t \) as a function of time for different values of \( \lambda \). Pelanek (2014) found \( \lambda = .1 \) to be an optimal value for \( \lambda \) in his real datasets. The parameter \( \lambda \) can be assumed to be game dependent, however, and we will investigate in the next section whether it can be estimated.

Once the decay parameter is known, the prior mean and variance for \( \theta_t \) at time \( t \) can be computed as a
function of the posterior mean and variance and the time passed after the previous play session. Updating from prior to posterior at time $t$ can then proceed as described in paragraph 2.

### 3.3 Simulation results

We executed a (currently still very small) simulation study to see whether the parameter $\lambda$ can be estimated from the data. 50 datasets with varying values of $d$, drawn from a uniform distribution between 0 and .2 were simulated. For now, $N$ was kept at 3000 observations. All independent single item responses were simulated with a different time passed since previous play session, different prior means, but a prior variance of .1. We used the program JAGS (Plummer, 2004) to estimate the decay parameter $\lambda$. Figure 4 shows the correlation between the simulated and estimated $\lambda$’s. It appears that the estimation of $\lambda$ works quite well for this combination of conditions, as there is a correlation of .93 between the simulated and estimated $\lambda$’s.

*Figure 3: Decay function for different values of $\lambda$*

*I am planning to include another plot showing how well the time decay and reset prior recover simulated $\theta$ values.*
4 Shapes and patterns: a real time adaptive game to increase learning and engagement

4.1 Study description

An iPad app was created for preschool-aged children, to advance shape understanding and to teach pattern recognition and extension to preschoolers using two shape games and two pattern games. In all four games, participants were given corrective feedback and hints after incorrect responses, with multiple chances to provide the right answer (although only the first response counted towards ability estimation). In addition, participants were shown short video clips between the games to reinforce the concepts they had just been working on. We conducted a field experiment to measure engagement and learning in adaptive and non-adaptive versions of this game. All participants began by completing a pretest with nine questions about shapes and nine about patterns. Participants were then randomly assigned to the adaptive (n = 44), non-adaptive (n = 47) or control (n = 48) condition, with condition assignment stratified by pretest score. For both the adaptive and non-adaptive conditions, each game was designed to continuously measure participant ability. For only the adaptive condition, participants were then presented challenges (i.e., items) with an expected 70% probability of correct response, where the first item was presented based on the pretest score.

Figure 4: Correlation between simulated and estimated λ’s
For participants in the non-adaptive condition, challenge difficulty was increased at the beginning of the third and then every other lesson (i.e., every four to five days), regardless of the participant’s ability or whether the participant had played once, multiple times, or not at all during that lesson. After six weeks, all participants were asked to complete an 18-question posttest (a parallel form of the pretest), which resulted in 36 (adaptive), 39 (non-adaptive), and 40 (control) completed posttests. Participants in the control group were given access to the game after completing the posttest.

4.2 Evaluating the adaptive mechanism

In the adaptive condition, we reset the prior variance after each play session, as described in 3.1. To evaluate whether the adaptive version of the game was able to adapt to the level of the participant adequately, we looked at the in-game ability estimates and the difficulties of the items offered to the participants in the adaptive condition. Because participants were given items that they were expected to answer correctly with 70% probability, the threshold they needed to reach for moving up in item difficulty was equal to the item’s difficulty + 0.5. Figure 7 shows that item difficulty thresholds (bold lines) and in-game participant-ability estimates (dotted lines) matched quite well for the shapes games. For the patterns games, however, the adaptivity of the games was not optimal: The range of item difficulties for this domain was too narrow, overlapping only with a small percentage of the actual participant abilities. Therefore, participants with very low or high ability got stuck on one level of the game.

One important lesson from this experiment for designers of adaptive learning games is to include challenges (i.e., items) spanning a wide enough range of difficulties to match the full range of learners’ abilities. This requires some form of item calibration (testing challenges with diverse learners to assess their difficulties) during the design of the game so that gaps in challenge difficulty can be filled with appropriately difficult new challenges.

4.3 Adaptive versus non-adaptive

As the initial goal of the experiment was to assess whether adaptivity increases learning and engagement, we will present some of our results regarding these hypotheses.

4.3.1 Change from pretest to posttest

Unfortunately, no differences in the change from pre to posttest were found among the adaptive, non-adaptive and control conditions, quite possibly due to the low reliability of the instruments with our participants. It
is also possible that the instruments were not sensitive enough to overcome the effects of parental support in the tests: In a post-study survey, 75% of parents reported having assisted at least a little with either the pre- or posttest, even though they were explicitly instructed not to. Another possible explanation could be that staying focused and completing an 18-item test was simply too ambitious for our preschoolers (e.g., Jones, Rothbart & Posner, 2003).

4.3.2 Learning

Next we looked at in-game performance measures, using ability estimates calculated throughout the study to see how ability estimates at the end of each play session changed over time. Time was defined by lesson number, which represents the subsequent 2-3 day periods in which the participants were supposed to play all four games at least once. For both domains (shapes and patterns), we ran a linear mixed model on the estimated ability scores with random effects for the participants, and fixed effects for lesson number, condition, and the lesson number-by-condition interaction. The results confirm what we have seen above for the working of the adaptive mechanism.

For the shapes game, there was an interaction between condition and lesson number. The Bayes factor in favor of the interaction indicated that the data were $7 \times 10^5$ times more likely to have occurred under the alternative hypothesis that there is an interaction than under the null hypothesis that there is no interaction.
Figure 6 shows that the learners in the adaptive condition increased more in their estimated ability than the learners in the non-adaptive condition.

Figure 6: Change in learner ability estimates per condition

For the Patterns game, there was also an interaction between condition and lesson number, but in the opposite direction. The Bayes factor in favor of the interaction indicated that the data were $16 \times 10^3$ times more likely to have occurred under the alternative hypothesis that there is an interaction than under the null hypothesis that there is no interaction. Figure 6 show that the learners in the adaptive condition increased less in their estimated ability than the learners in the non-adaptive condition. As the range of difficulty levels for the games’ challenges in this game did not match the actual abilities of the participants playing the games, the majority of adaptive-condition participants got stuck in either the easiest or hardest game challenges. In the non-adaptive condition, participants received challenges of increasing difficulty regardless of their performance, which could explain why these learners showed increased learning over time.

4.3.3 Engagement

The first measure of engagement we looked at was the duration of play sessions, defined as the time between opening and closing the app that contained the games. Because the distributions are skewed (as is often the
case with measures of time or duration), we took the log of the durations and assessed whether there was a difference in log(duration) between conditions. The average duration of play sessions higher in the adaptive condition (9.9 minutes) than in the non-adaptive condition (8.7 minutes) (Bayes factor results to follow).

As a way to look at retention, we considered the number of times participants played within each lesson period. There was a decrease in the number of playthroughs per lesson for the adaptive condition, but this decrease was steeper for the non-adaptive condition. Figure 7 illustrates these results. The dotted lines represent the observed average number of playthroughs in each lesson, and the solid lines represent the number of playthroughs predicted by a negative binomial regression model. Retention is significantly lower in the non-adaptive than in the adaptive condition (Bayes factor results to follow).

![Figure 7: Retention of learners over time: Number of playthroughs in each of the lesson periods.](image)

4.4 Time decay prior versus the reset prior

Unfortunately, the experiment did not consist of enough participants to be able to estimate the parameter $\lambda$ reliably. To illustrate the difference between the two priors, a $\lambda$ of .05 was used here. The ability estimates produced by incorporating the two different time-dependent priors are shown in Figure 8 and Figure 9. What stands out in Figure 8 is that the estimates from the reset prior fluctuate a lot, while the time decay prior produces more stable results. In addition, the time decay prior produces less extreme estimates, which stay closer to zero. Looking at Figure 9 shows that the reason for the less extreme estimates is that the time decay prior pulls the extreme estimates towards zero as more time passes. Except from the high ability
estimates, however, the ability scores seem to converge after a few items in each play session, however. This is also reflected in the high correlation between the scores, which was .94.

Figure 8: Ability estimates resulting from reset and time decay priors against the number of items played

Figure 9: Ability estimates resulting from reset and time decay priors against the time passed
5 Learner mosaic: Reporting scores back to parents

One of our goals with adaptive educational games was to give feedback to parents and teachers. We implemented this in our game Leo’s pad, a game for preschoolers covering multiple learning dimensions.

5.1 Parent and teacher feedback

Play data and continuously updated learner ability estimates not only help drive adaptivity in educational games. They can also help drive more personalized feedback for a learner’s parents and teachers. We combine these efforts through our Leo’s Pad Enrichment Program (LPEP) and Learner Mosaic, a mobile application designed to provide parents with meaningful and actionable insights about their learner’s efforts and progress. For example, say a learner played games in LPEP that feature items on Ordinality. After a period of struggling to succeed, the learner masters the challenges and consistently succeeds. This change would be reflected in a positive change in her Ordinality ability estimate as well as in higher confidence levels in that estimate. These changes can in turn be used to send the learner’s mother an alert indicating that the learner has achieved a new milestone in Ordinality along with suggestions for activities she could try that would build on this success. Figure 10 shows how the game looks, and Figure 11 shows how feedback to parents was provided and how the suggestions for activities looked.

Figure 10
5.2 Collecting additional information

In addition to providing this kind of interpreted guidance about changes in learner ability estimates, Learner Mosaic also provides an alternative source of measurement data. Parents can offer their own observations of learner performance guided through question items provided in the application. These items can then also contribute to the emerging learner ability model as an additional source of evidence. Figure 12 shows an example of an additional question asked in Learner Mosaic.
6 Discussion

A Bayesian adaptive framework for educational games was presented, with applications to adaptive learning and parent feedback. There are many more things to research within the Bayesian IRT framework for adaptive games. More complex IRT models like categorical models, multidimensional models and models including response time can be adjusted to adaptive learning environments. Different decay functions can be investigated, and ways of calibrating item parameters that might change over time. Other ways of predicting the prior distribution can be investigated, possibly by methods derived from machine learning and data science fields, using different types of information about environmental factors, gameplay in other area’s, or even modeling of individual growth paths. In addition, we need to work on optimizing the estimation procedures to run these models on mobile devices and to make them scalable for a large number of users.

References


Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. (2014). A survey on concept drift adaptation. ACM Computing Surveys (CSUR), 46, 44.


