Real-time Adaptation of Augmented-Reality Games for Optimizing Player Satisfaction

Georgios N. Yannakakis, *Member, IEEE*, and John Hallam

***Abstract*— A first endeavor for optimizing player satisfaction in augmented-reality games through the 'Playware' physical interactive platform is presented in this paper. Constructed user models, reported in the literature, map individual playing characteristics to reported entertainment preferences of augmented-reality game players. An adaptive mechanism then adjusts controllable game parameters in real-time in order to improve the entertainment value of the game for the player. The basic approach presented here applies gradient ascent to such a model to reveal the direction toward games of higher entertainment value while a rule-based system exploits the derivative information to adjust specific game parameters to augment the entertainment value. Those adjustments take place frequently during the game in small time intervals that maintain the constructed model's accuracy. Performance of the adaptation mechanism is evaluated using a game survey experiment. Results reveal that children show a notable preference for the adaptive versus the static Bug-Smasher ('Playware' test-bed) game variant even when simple adaptive approaches like the one proposed are used. The limitations and the use of the methodology as a baseline effective adaptive mechanism to entertainment augmentation are discussed.**

I. Introduction

Cognitive user models of playing experience promise significant potential for the design of digital interactive entertainment systems such as screen-based computer or augmented-reality games. Quantitatively modeling entertainment or satisfaction as a class of user experiences may reveal game features or user features of play that relate to the level of satisfaction perceived by the user (player). That relationship can then be used to adjust digital entertainment systems according to individual user preferences to optimize player satisfaction in real-time. This paper introduces an adaptive mechanism for achieving this in augmented-reality games. A game designed on the 'Playware' [1] intelligent interactive physical playground is used as the test-bed for the experiments presented here.

Entertainment models constructed in our previous studies [2], [3] are based on quantitative measures of Malone's intrinsic qualitative factors for engaging gameplay [4], namely *challenge* and *curiosity*. In that prior work, a mapping between the aforementioned factors (game features), the player's individual characteristics (player features), such as response time and foot pressure, and the children's notion of 'fun' or entertainment (the two terms are used interchangeably herein) is derived for a game (Bug-Smasher) developed on the Playware playground as a test-bed. Preference learning (neuro-evolution) models are trained on these gameplay experimental data to construct a function mapping the examined game and player features to the reported player satisfaction preferences. Feature selection methodology applied in conjunction with preference learning yielded an artificial neural network (ANN) model that predicts the entertainment preferences of children with an accuracy of 77.77% [3], built on four features: the player's average response time with the playground, the variance of the pressure force instances on the playground, the number of interactions with the playground and the level of curiosity generated by the game opponents.

Following from the reported success with entertainment modeling in physical interactive games [3], an attempt to optimize the entertainment value of those games in real-time is presented in this paper. A real-time adaptation mechanism using gradient ascent on the constructed model is tested through a survey experiment of children playing variants of the Bug-Smasher game. Results demonstrate the efficacy of the adaptive mechanism since the children who participated in the experiment expressed a considerable preference for the adaptive version of the game under investigation. These positive indications for simple hill-climbing mechanisms applied to augmenting player satisfaction suggest the promise of future implementations of adaptive learning in real-time for the same problem.

The work reported here is novel in that it demonstrates a way of constructing a subjective model (a predictor of user preferences) of reported entertainment grounded in statistical features obtained from child-game interaction and introduces an approach for exploiting the knowledge represented by that model towards enhancing player satisfaction in real-time. The limitations of the proposed methodology and its generic use as an efficient baseline for optimizing entertainment in physical interactive games in real-time are discussed.

II. Entertainment Modeling and Optimization

We classify approaches for capturing the level of player satisfaction into *qualitative* and *quantitative*. The first includes qualitative features/criteria that collectively contribute to engaging experiences in entertainment systems, derived from experimental psychology studies, whereas the latter

Georgios N. Yannakakis is with Center for Computer Games Research, IT-University of Copenhagen, Rued Langgaards Vej 7, DK-2300 Copenhagen (phone: +45-7218-5078; fax: +45-7218-5001; email: yannakakis@itu.dk). John Hallam is with the Mærsk Mc-Kinney Møller Institute, University of Southern Denmark, Campusvej 55, DK-5230, Odense (phone: +45-6550-3546; fax: +45-6550-7697; email: john@mmmi.sdu.dk)

978-1-4244-2974-5/08/$25.00 ©2008 IEEE

includes studies for quantifying the reported qualitative criteria of entertainment and constructing models that relate to the complicated mental state of satisfaction perceived while interacting with digital interactive systems. Related work on methodologies for improving player satisfaction in real-time is presented at the end of this section.

A. Qualitative Approaches

There have been several psychological studies to identify what is 'fun' in a game and what engages people playing computer games. Theoretical approaches include Malone's principles of intrinsic qualitative factors for engaging game play [4], namely *challenge*, *curiosity* and *fantasy* as well as the well-known concepts of the theory of flow [5] incorporated in computer games as a model for evaluating player enjoyment, namely *GameFlow* [6]. A comprehensive review of the literature on qualitative approaches for modeling player enjoyment demonstrates a tendency of overlapping with Malone's and Csikszentmihalyi's foundational concepts. Many of these approaches are based on Lazzaro's 'fun' clustering which uses four entertainment factors based on facial expressions and data obtained from game surveys on players [7]: hard fun, easy fun, altered states and socialization. Koster's [8] theory of fun, which is primarily inspired by Lazzaro's four factors, defines 'fun' as the act of mastering the game mentally. An alternative approach to fun capture is presented in [9] where fun is composed of three dimensions: endurability, engagement and expectations.

B. Quantitative approaches

Iida's work on metrics of entertainment in board games was the first attempt in the area of quantitative 'fun' modeling. He introduced a general metric of entertainment for variants of chess games depending on average game length and possible moves [10]. Other work in the field of quantitative entertainment capture is based on the hypothesis that the player-opponent interaction — rather than the audiovisual features, the context or the genre of the game — is the property that contributes the majority of the quality features of entertainment in a computer game [11]. Based on this fundamental assumption, a metric for measuring the real time entertainment value of predator/prey games was designed, and established as efficient and reliable by validation against human judgement [12], [13]. Further studies by Yannakakis and Hallam [14] have shown that ANNs and fuzzy neural networks can extract a better estimator of player satisfaction than a human-designed one, given appropriate estimators of the challenge and curiosity of the game and data on human players' preferences.

A step further to entertainment capture is towards games of richer human-computer interaction and affect recognizers which are able to identify correlations between physiological signals and the human notion of entertainment. Experiments by Yannakakis et al. [15], [16] have isolated features of heart rate (HR) signal recordings of children attributed to entertainment from those features that correspond to pure physical activity in action games played in interactive physical playgrounds. In that study, a highly accurate subjective model (a predictor of entertainment preferences) of reported "fun" grounded in statistical features of HR signal dynamics was constructed. Moreover, Rani et al. [17] propose a methodology for detecting anxiety level of the player and appropriately adjusting the level of challenge (e.g. speed) in the game of 'Pong'. Physiological state (hear-rate, galvanic skin response) prediction models have also been proposed for potential entertainment augmentation in computer games [18].

C. Optimizing Player Satisfaction

Approaches towards optimizing player satisfaction can be classified to *implicit* and *explicit*. Within the first class of approaches we meet use of machine learning techniques for adjusting a game's difficulty — based on the assumption that challenge is the only factor that contributes to enjoyable gaming experiences — which implies entertainment augmentation. Such approaches include applications of reinforcement learning [19], genetic algorithms [20], probabilistic models [21] and dynamic scripting [22], [23]. Moreover, user models have been constructed for the generation of adaptive interactive narrative systems that potentially optimize the experience of the user [24], [25], [26]. User preference modeling towards content (race track) creation in racing games has also shown a potential for enhancing the quality of playing experience in those games [27], [28]. However, human survey experiments that cross-verify the assumptions of player satisfaction enhancement have not been reported in all aforementioned approaches.

Within the *explicit* methods for optimizing player satisfaction, robust adaptive learning mechanisms have been built to optimize the human-verified ad-hoc 'interest' (entertainment) metric for prey/predator games introduced in [11], [12]. Experiments showed that an on-line neuro-evolution mechanism [13], [29], [30], [31] and a player modeling technique through bayesian learning [32] were each capable of maintaining or increasing the game's entertainment value while the game was being played. Effectiveness and robustness of the adaptive (neuro-evolution) learning mechanism in real-time has been evaluated via human survey experiments [12]. Furthermore, studies with the "Playware" playground have shown that real-time adaptive ad-hoc rule-based mechanisms may improve children's gameplay experience in physical interactive playgrounds [33].

Following the theoretical principles reported from Malone [4], Koster [8] and Yannakakis [34], this paper is primarily focused on the contributions of game opponents' behavior to the real-time entertainment value of the game. We argue that among the three dimensions of 'fun' (endurability, engagement, expectations) defined in [9] it is only *engagement* that is affected by the opponent since both *endurability* and *expectations* are based primarily on the game design *per se*. Given a successful interactive game design that yields high expectations and endurability, we focus only on the level of engagement that generates 'fun' (entertainment). However,

instead of being based on empirical observations of children's entertainment, the work presented here uses quantitative entertainment models already constructed using experimental data obtained from a survey experiment with children playing with Playware playground [3]. An adaptive mechanism is proposed for augmenting children's satisfaction in real-time and an additional survey experiment validates its efficacy in the Bug-Smasher test-bed game.

III. Playware Playground

The Playware [1] prototype playground consists of several building blocks (i.e. tangible tiles) that allow for the game designer (e.g. the child) to develop a significant number of different games within the same platform. The overall technological concept of Playware is based on embodied AI [35] where intelligent physical identities (tiles) incorporate processing power, communication, input and output, focusing on the role of the morphology-intelligence interplay in developing game platforms. See [2], [3], [1] for further details on Playware playground.

A. Bug-Smasher Game

The test-bed game used for the experiments presented here is called 'Bug-Smasher'. The game is developed on a 6 x 6 square tile topology. During the game, different 'bugs' (colored lights) appear on the game surface and disappear sequentially after a short period of time by turning a tile's light on and off respectively. A bug's position is picked randomly according to the predefined level of the bugs' spatial diversity. Spatial diversity is measured by the entropy (H) of the bug-visited tiles.

The child's goal is to smash as many bugs as possible by stepping on the lighted tiles. Bug-smasher has been used as a test-bed in previous work; further details can be found in [2], [15], [36], [3] and [16] .

IV. Experimental Data for Entertainment Modeling

In [3] the Bug-Smasher game was used to acquire data of platform-child interaction and children's judgement on entertainment. Three states ('Low', 'Average', and 'High') were used for each of the two entertainment factors of challenge and curiosity summing up to 9 different game variants. In that study the speed (S — in sec^{-1}) that the bugs appear and disappear from the game and their spatial diversity (H) on the game's plane were considered as appropriate measures to represent the level of challenge and the level of curiosity (unpredictability) respectively [4] during gameplay. The former provides a notion of a goal whose attainment is uncertain and the latter effectively portrays a notion of unpredictability in the subsequent events of the game — the higher the H value the higher the bug appearance unpredictability and therefore the higher the curiosity.

Seventy-two normal-weighted (based on their body mass index) children whose age covered a range between 8 and 10 years participated in the experiment reported [3]. In that experiment, each subject played two game variants (A and B) for 90 seconds each; the two games differed in the levels of one or both entertainment factors of challenge and curiosity. Each time a pair of games was finished, the child was asked whether the first game was more 'fun' than the second game i.e. whether A or B generated a more entertaining game (pairwise preference — 2-alternative forced choice). Further details regarding the protocol used in that experimental survey and the statistical analysis on the effect of order of play can be found in [3].

Pressed tile events during the experiment presented in [3] were recorded in real-time and a selection of nine personalized (individual) player features were extracted for each child. These include the number of smashed bugs over the total number of bugs presented, P (i.e. child's score); the number of interactions with the game environment N_I; the average and the variance of the response times $(E\{r_t\}, \sigma^2\{r_t\})$; the average and the variance of the distance between the pressed tile and the bugs appearing on the game $(E\{D_b\}, \sigma^2\{D_b\})$; the average and the variance of the pressure recorded from the FSR sensor $(E\{p\}, \sigma^2\{p\})$; and the entropy of the tiles that the child visited H_C. The total number of game pairs played was 144; however, data from 137 game pairs were used due to hardware (communication ports) failure during seven games. The child's answers and recorded data were used to guide the training of an ANN model of reported entertainment (see Section V).

A. Statistical Analysis

A statistical analysis introduced in [3] was applied to identify statistically significant correlations between children's notion of entertainment and any of the aforementioned individual player features and/or the quantitative entertainment factors (game features): challenge and curiosity. Among the obtained significant effects presented in that study we report the effect of N_I (p-value $= 0.0298$) and $E\{r_t\}$ (p-value $= 0.0050$) on reported entertainment. These linear correlations are used as background knowledge for the construction of the rules of the adaptation mechanism (see Section IX). The reader may refer to [3] for further details on the statistical analysis briefly presented here.

V. Preference Learning

The proposed approach to entertainment modeling is based on selecting a (constrained) minimal subset (see Section VI) of individual features and constructing a quantitative user model that predicts the subject's reported entertainment preferences. The assumption is that the entertainment value y of a given game, which models the subject's internal response to playing the game, that is, how much "fun" it is, is an unknown function of individual features which a machine learning mechanism can learn. The subject's expressed preferences constrain but do not specify the values of y for individual games but we assume that the subject's expressed preferences are consistent.

Constraint satisfaction algorithms cannot solve the problem since the variable y under the constraint $y_A > y_B$ for any two given games A and B has no specific domain

values. Likewise, any machine learning which is based on learning a target output is inapplicable since target outputs are unknown. By the use of a ranking approach numerical values for the y variable could be made available; however, ranking is an undesired method for the self-report design of comparative "fun" analysis for the disadvantages mentioned earlier. Preference learning [37] is the only applicable type of machine learning for this constrained classification problem. There are several techniques that learn from a set of pairwise preferences such as algorithms based on support vector machines [38], gaussian processes [39] and evolving ANNs [40].

A. Neuro-evolution

Among the aforementioned preference learning approaches a fully-connected feedforward ANN has successfully been used for learning the relation between the selected game and player features (ANN inputs) and the "entertainment value" (ANN output) of a game [3]. Since there are no prescribed target outputs for the learning problem (i.e. no differentiable output error function), ANN training algorithms such as back-propagation are inapplicable. Learning is achieved through artificial evolution [40].

A generational genetic algorithm (GA) [41] is implemented, which uses a fitness function that measures the difference between the children's reported preferences of entertainment and the model output value y. The ANN is itself evolved. In the algorithm presented in [3], the ANN topology is fixed and the GA chromosome is a vector of ANN connection weights. For further information on the neuro-evolution algorithm the reader may refer to [2] and [3].

VI. Feature Selection

There were two different feature selection schemes applied and compared for the construction of entertainment models in [3]: the *n Best Features Selection* (nBest) and the *Sequential Forward Selection* (SFS). Features selected by each method constitute the input vector of the evolving ANN. The feature selection procedure followed in that study evaluates the usability of each one of the features available and obtains the minimal feature subset that performs best in the classification between games reported as entertaining and games reported as non-entertaining (see Section V). To evaluate the performance of each feature subset the available data is randomly divided into three subsets which are combined to give three training and, independent, validation data sets each consisting of 2/3 and 1/3 of the data respectively. The performance of an ANN model is measured through the average classification accuracy of the ANN in three independent runs using 3-fold cross-validation technique on the training and validation data sets. The reader may refer to [3] for further information on the motivation behind the use of feature selection and the details of each method used.

VII. Entertainment Model Constructed

This section presents briefly the model generated and introduced in [3]. Experiments for finding the minimal feature subset that yields the highest classification performance in matching the ANNs output with children's reported answers on entertainment in unknown data (validation data set) reveal a 3-fold cross-validation performance of 77.77% (average of 70%, 73.33% and 90%) when the ANN input (selected features) contains $E\{r_t\}$, $\sigma^2\{p\}$, H and N_I. The binomial-distributed probability of this performance to occur at random is 0.0019 demonstrating statistical significance and providing evidence for this solution's robustness.

Difficulties in obtaining higher classification accuracy were found in experimental noise in both the recorded features and the children's answers on self reports. Even though comparative fun analysis is a reliable and established method for capturing reported entertainment in computer [34] and augmented-reality [2], [3] games, it generates a significant amount of uncertainty in subjects' reported answers. Uncertainty appears when the two games played are not significantly different with regards to the entertainment value they generate for the player and therefore cannot be distinguished.

A further analysis of the feature subset $\{E\{r_t\}, \sigma^2\{p\}, N_I, H\}$ with the highest validation performance (90.00%, in one of the three learning attempts) reveals that fast responding children ($E\{r_t\} \approx 0$) tend to enjoy average and high curiosity values whereas slow children ($E\{r_t\} \approx 1$) appear to prefer games that generate low curiosity levels [3].

VIII. "Fun" during the game

The main assumption guiding this research is that entertainment is constant during the selected gameplay window (i.e. 90 seconds). In other words, the expressed entertainment preference for a game is valid for the game as a whole. Obtaining entertainment preferences during play is a protocol an experiment designer may follow; however, it was not preferred since it could have a significant impact on the validity of the obtained data. According to the ideal conditions for effective data collection, subjects (children) should not be questioned during the task under examination (game) [42].

To test the extent to which the entertainment value of a game is constant over the entire playing time window we proceed as follows. ANNs models of entertainment built through 3-fold cross-validation on data derived from the whole game play (e.g. average response time over 90 seconds) are available from [3] (see Section VII). Those three ANN models are tested on the entire data set of 137 game pairs (see Section IV) and the average classification performance is calculated. Then the game is divided in two equal parts and all features are recalculated for those two 45-second time windows. The models are tested on the entire new data set (2 times 137 data) assuming that the expressed entertainment preferences are valid for both 45-second segments. The performance is calculated and the gameplay data are further divided into smaller time periods.

This procedure continues up to the point where a small time window is reached (e.g. 9 seconds). This test indicates the minimum length time window for which the model can still predict reported entertainment with high accuracy. That time window can then be used to set the frequency of any real-time adaptation mechanism applied.

Fig. 1 illustrates the classification performance of the ANN models with respect to the time window selected. The percentage of data lost are also presented. A data point is lost when a child does not interact with the platform within a specific time window, so no calculation of interaction data is possible since there are no recorded instances. Results show that there appear to be gameplay intervals which yield better classification performance than others and that player satisfaction may vary in between those intervals. This is demonstrated by the immediate decrease of the model's performance (62.23%) as soon as the game's time window is split in half (i.e. 45 seconds). The performance stabilizes around 58% when the time interval is further reduced. Even though the difference between the ANN model's evaluation performance on the 90-second play window and the respective performance on the 9-second time window is approximately only 16%, it demonstrates that the hypothesis of constant entertainment is ruled out for small time intervals.

Fig. 1 shows that data are lost when gameplay time windows become smaller than 22.5 seconds. This suggests that the minimum acceptable time window lies between 30 and 45 seconds since the performance of the models, even though reduced, is still above 60.0%. To further investigate the 30–45 second interval we introduce time-shifted offsets (TSOs) for those two time windows. TSO is a time interval that is added to the selected time window to yield a new time interval. For instance, a 45-second window with a 15-second TSO generates the following four time intervals: 0 to 45, 15 to 60, 30 to 75 and 45 to 90 seconds. Given that definition, if the TSO equals the time window length there is no overlap between the generated time intervals (no offset effect). As seen in Table I, TSOs of 45, 22.5 and 15 seconds are used for the 45-second time window whereas TSOs of 30 and 15 seconds are used for the 30-seconds time window. Table I presents the classification performances for all time window-TSO combinations mentioned above. The amount of data that ANN models are evaluated on and data lost due to the lack of interaction instances with the game platform are also presented.

TABLE I
TIME-SHIFTED OFFSET (TSO) IMPACT ON 45 AND 30-SECOND WINDOWS

Time Window	TSO	Data	Data Lost	Performance (%)
45	45	274	0	62.23
45	22.5	411	0	63.17
45	15	548	0	64.53
30	30	407	4	59.34
30	15	680	5	60.88

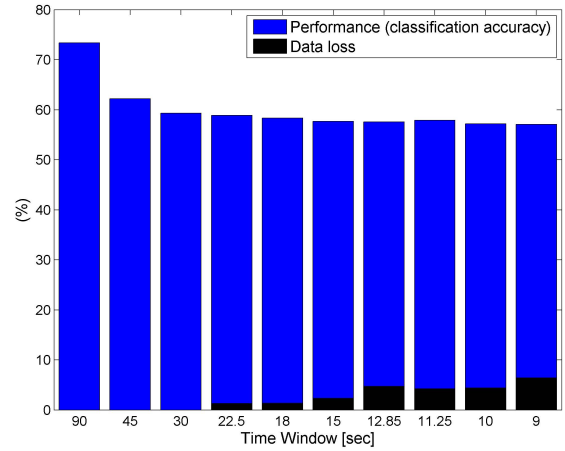

Fig. 1. ANN model performance and respective data loss percentage over different gameplay time windows. The two percentage values (columns) are independent in this illustration.

ANN models evaluated on the 45:15 (45-second time window combined with a 15-second TSO) intervals yield a performance of 64.53% while interaction is maintained within that time period (data are not lost). This suggests that the 45:15 time scheme is the most appropriate — among the ones tested — to be used for real-time adaptation of the Bug-Smasher game. This, furthermore, implies that action of any adaptation mechanism designed occurs at 45 seconds and each time a 15 second interval elapses thereafter — i.e. after 45, 60, and 75 seconds of play.

IX. REAL-TIME ADAPTATION MECHANISM

Given the ANN model of entertainment preferences built on individual gameplay features (see Section VII), the question that arises is how this model can be used for improving children's gameplay experience in real-time. There are several ways of exploiting the model's built-in knowledge (discussed below in section XI of this paper) and adapting Playware games for enhancing the level of entertainment. In this initial study we start with a simple hill-climbing mechanism presented in this section.

The idea behind real-time adaptation is to use the metric evaluation function constructed using machine learning directly to enhance the entertainment provided by the game. The key to this is the observation that the model (ANN) relates game features to an entertainment value y. It is therefore possible in principle to infer what changes to game features will cause an increase in the entertainment value of the game, and to adjust game parameters to make those changes. Thus, adaptation of the game opponents (e.g. bugs) may occur according to the player's individual playing style, based on reaction time, recorded pressure on tiles and amount of interactions, and as far as the curiosity factor (H) of entertainment is concerned. Given the real-time average response time ($E\{r_t\}$) of a child, the variance of his/her pressure forces on the tiles ($\sigma^2\{p\}$) and the number

of times he/she interacts with the environment (N_I), the partial derivative of the model output $\vartheta y/\vartheta H$ can be used to appropriately adjust the level of entropy (curiosity) of the opponent for the entertainment value y to be augmented (see equation (1)).

$$\vartheta y/\vartheta H = y(1-y)\sum_{i=1}^{M}\{o_i(1-o_i)w^1_{H,i}w^2_{i,1}\} \qquad (1)$$

In equation (1), o_i is the output of the i-th hidden neuron of the ANN model; $w^1_{H,i}$ is the connection weight between the H input and the i-th hidden neuron and $w^2_{i,1}$ is the connection weight between the i-th hidden neuron and the output neuron of the ANN model.

Given that the $\vartheta y/\vartheta H$ value indicates the change in entertainment for a small change in the curiosity level, one could use gradient ascent to attempt to improve entertainment with such a model.

Previous studies [3] have shown that the number of interactions and the average response time are features correlated linearly with entertainment preferences (see Section IV-A). Our assumption is that these two features are strongly related to the speed (challenge) of the game under investigation. Therefore, in addition to curiosity level adjustment, the game speed is adapted in real-time. Analogously to (1), the partial derivatives $\vartheta y/\vartheta N_I$ and $\vartheta y/\vartheta E\{r_t\}$ are calculated and the game's speed is altered if those values have different signs: higher speed for positive $\vartheta y/\vartheta N_I$ and negative $\vartheta y/\vartheta E\{r_t\}$; lower speed for negative $\vartheta y/\vartheta N_I$ and positive $\vartheta y/\vartheta E\{r_t\}$. Table II presents the rules used for adjusting the curiosity and challenge levels of the game using a 45:15 second time window scheme in a 90 second game — that is on the 45th, the 60th and the 75th second (see Section VIII). Adjustments are implemented by altering the state ('Low', 'Average', 'High') of the internal controls (challenge, curiosity) by one level up (+) or down (−). Note that when $|\vartheta y/\vartheta H| < \epsilon$ (third row of Table II), curiosity is either increased or decreased (50% probability for each action to occur).

TABLE II
ADAPTATION MECHANISM RULES. ϵ EQUALS 0.5 IN THIS PAPER.

Condition	Action
$\vartheta y/\vartheta H > \epsilon$	H +
$\vartheta y/\vartheta H < -\epsilon$	H −
$\lvert\vartheta y/\vartheta H\rvert < \epsilon$	H + or H −
$(\vartheta y/\vartheta E\{r_t\} < -\epsilon)$ AND $(\vartheta y/\vartheta N_I > \epsilon)$	S +
$(\vartheta y/\vartheta E\{r_t\} > \epsilon)$ AND $(\vartheta y/\vartheta N_I < -\epsilon)$	S −

X. ADAPTATION EXPERIMENT

The Bug-Smasher game has been used to test the efficiency of the adaptation mechanism proposed. Two variants of this game have been constructed: the *static* and the *adaptive*. The bugs' speed (S) and spatial diversity (H) in the game plane for the *static* game are adjusted to the average values of the three different levels ('Low', 'Average', and 'High') of challenge and curiosity respectively used in the Bug-Smasher experiments. The *adaptive* game is initialized with the same values for speed and spatial diversity but the challenge and curiosity levels (states) are adjusted on the 45th, 60th and 75th second of the game according to the adaptation rules presented in Table II.

For the adaptation experiment, we asked 24 naïve normal-weighted children (13 boys and 11 girls) aged 8 to 10 years to play 4 games each on the Playware platform. The set of 4 games played comprised 2 games of static and 2 games of adaptive Bug-Smasher in all combinations. Thus, the number of children participated in the experiment is derived from $4 \cdot C^4_2 = 24$ being four times the required number of all combinations of 2 out of 4 games. Subjects play games in pairs (game A and game B) and each time a pair of games is finished, the child is asked to choose among the following alternatives (note that, children are not interviewed but are asked to fill in a questionnaire, minimizing the interviewing effects reported in [43])

- the first [second] game was more 'fun' (see [9] for terminology used in experiments with children) than the second [first] game (*cf.* 2-alternative forced choice)
- both games were equally 'fun' or
- neither of the two games was 'fun.'

Children complete the comparison questionnaire above after games 2, 3, and 4, resulting in three fun comparisons (expressed preferences) between games 1–2, 2–3, and 3–4, for each child to report. That provides a total of 72 (24 children times 3 comparisons) "fun" comparisons. The 4-alternative forced choice (4-AFC) protocol above is used since it offers several advantages for subjective entertainment capture: it minimizes the assumptions made about children's notions of "fun" and allows a fair comparison between the answers of different children. The 4-AFC protocol provides the same preference information as the 2-AFC used in previous experiments [3], [16] for any machine learning process applied to construct entertainment models while also making explicit the "no preference" cases concealed by 2-AFC.

A. Adaptive vs. Static Bug-Smasher

Given the experimental protocol there are 50 out of 72 'fun' comparisons between the static and the adaptive Bug-Smasher. Results demonstrate that within those 50 comparisons there are 20 choices for the 'both games were equally fun' alternative and not a single choice for the 'neither was fun' alternative. The first indicates the difficulty for the respective children of distinguishing between the two games whereas the latter shows that both games offered an enjoyable experience to all children participated to the experiment. Moreover, in 18 and 12 out of the 30 comparisons remaining, children expressed a preference for the adaptive and the static game respectively. This results in a percentage of 60% and 76% for respectively higher, and higher or equal, preference for the adaptive game versus the static game. Even though not statistically significant, the percentages of

children's preference over the adaptive game provide promise for the effectiveness of the adaptation mechanism proposed here.

XI. Conclusions & Discussion

This paper introduced an adaptive approach for augmenting the player's satisfaction in Playware games in real-time. The first step in applying adaptive learning in those games is to make use of the player satisfaction formulae derived from entertainment modeling processes to adjust Bug-Smasher game parameters on-line. Previous studies [3] on preference learning through the combination of neuro-evolution and feature selection generated ANN models which derive from a feature subset of four features: $\{E\{r_t\},\ \sigma^2\{p\},\ N_I,\ H\}$. This model managed to map between children's average response time, the variance of their force pressure on the tiles, the number of interactions with the playground, the game feature of curiosity and the children's notion of gameplay entertainment with a 3-fold cross-validation accuracy of 77.77% (binomial-distributed p-value = 0.0019).

The mechanism presented here uses gradient ascent on the constructed entertainment model with regard to the number of interactions, average response time and curiosity metric inputs of that model. Thus, for instance, if the partial derivative of the entertainment value (model output) with respect to curiosity (model input) is positive, one could increase the curiosity level control a little and expect a positive increment in player enjoyment. Continual adjustments of both the speed (challenge) and spatial diversity (curiosity) of the game opponents (i.e. bugs) are controlled by simple rules and they take place in a time window of 45 seconds combined with a shifted time offset of 15 seconds. That time interval is extracted from experiments that evaluate the accuracy of the entertainment model over different time intervals testing the assumption of a constant entertainment value during the game.

A survey experiment for evaluating the performance of the adaptation mechanism was designed in which 24 children were asked to compare the standard (static) versus the adaptive variant of the Bug-Smasher game. Results reveal a positive preference of children for the adaptive Bug-Smasher game in 60% of those 'fun' comparisons. These results provide some first indications that augmenting player satisfaction in Playware games in real-time is possible even with simple hill-climbing adaptive mechanisms. Further analysis of the entertainment value generated might lead to safer conclusions regarding the effectiveness of the adaptation mechanism.

One of the limitations of the proposed entertainment modeling approach lies in the complexity of entertainment as a mental state. The generated y value cannot be regarded as a mental affective state approximator but as a correlate of expressed children's preferences of entertainment. However, this correlate serves the purpose of this work well as far as entertainment modeling and optimization is concerned. Using the proposed entertainment augmentation scheme, knowledge of the direction (from the partial derivative) that specific controllable features should be adjusted in is available through the model; however, the magnitude of such an adjustment is not known a priori. Thus, applying gradient search with a fixed step (difference of two states) in the search map between playing features and entertainment value may unexpectedly lead to lower values of entertainment. This constitutes a major limitation of the adaptive mechanism which might be resolved by injecting more controllable feature (challenge, curiosity) states in the search space or by introducing machine learning in real-time as proposed below.

The proposed adaptation mechanism can be used as a baseline approach for future implementations of adaptive learning on Playware games. The next obvious step is to apply reinforcement learning, most likely via classifier systems. In such an approach, the game internal controls (challenge, curiosity) could be adjusted (within limits) and the effect on player satisfaction monitored using the entertainment model. The observed satisfaction can act as reinforcement for the actor process adjusting the controls. To speed up learning, the reinforcement learning system could be seeded with the gradient information from the satisfaction measure. The results of learning could be saved so that experience with a given player accumulates: the game then adapts to the particular player over time.

Acknowledgments

The authors would like to thank Henrik Jørgensen and all children of Henriette Hørlücks School, Odense, Denmark that participated in the experiments.

The tiles were designed by C. Isaksen from Isaksen Design and parts of their hardware and software implementation were collectively done by A. Derakhshan, F. Hammer, T. Klitbo and J. Nielsen. KOMPAN, Mads Clausen Institute, and Danfoss Universe also participated in the development of the tiles.

This work was in part supported by the Danish National Research Council (project no: 274-05-0511).

References

[1] H. H. Lund, T. Klitbo, and C. Jessen, "Playware technology for physically activating play," *Artifical Life and Robotics Journal*, vol. 9, no. 4, pp. 165–174, 2005.

[2] G. N. Yannakakis, H. H. Lund, and J. Hallam, "Modeling Children's Entertainment in the Playware Playground," in *Proceedings of the IEEE Symposium on Computational Intelligence and Games*. Reno, USA: IEEE, May 2006, pp. 134–141.

[3] G. N. Yannakakis and J. Hallam, "Game and Player Feature Selection for Entertainment Capture," in *Proceedings of the IEEE Symposium on Computational Intelligence and Games*. Hawaii, USA: IEEE, April 2007, pp. 244–251.

[4] T. W. Malone, "What makes computer games fun?" *Byte*, vol. 6, pp. 258–277, 1981.

[5] M. Csikszentmihalyi, *Flow: The Psychology of Optimal Experience*. New York: Harper & Row, 1990.

[6] P. Sweetser and P. Wyeth, "GameFlow: A Model for Evaluating Player Enjoyment in Games," *ACM Computers in Entertainment*, vol. 3, no. 3, July 2005.

[7] N. Lazzaro, "Why we play games: Four keys to more emotion without story," XEO Design Inc.," Technical Report, 2004.

[8] R. Koster, *A Theory of Fun for Game Design*. Paraglyph Press, 2005.

[9] J. Read, S. MacFarlane, and C. Cassey, "Endurability, engagement and expectations," in *Proceedings of International Conference for Interaction Design and Children*, 2002.

[10] H. Iida, N. Takeshita, and J. Yoshimura, "A metric for entertainment of boardgames: its implication for evolution of chess variants," in *IWEC2002 Proceedings*, R. Nakatsu and J. Hoshino, Eds. Kluwer, 2003, pp. 65–72.

[11] G. N. Yannakakis and J. Hallam, "Evolving Opponents for Interesting Interactive Computer Games," in *From Animals to Animats 8: Proceedings of the 8th International Conference on Simulation of Adaptive Behavior (SAB-04)*, S. Schaal, A. Ijspeert, A. Billard, S. Vijayakumar, J. Hallam, and J.-A. Meyer, Eds. Santa Monica, LA, CA: The MIT Press, July 2004, pp. 499–508.

[12] ——, "Towards Optimizing Entertainment in Computer Games," *Applied Artificial Intelligence*, vol. 21, pp. 933–971, 2007.

[13] ——, "A Generic Approach for Obtaining Higher Entertainment in Predator/Prey Computer Games," *Journal of Game Development*, vol. 1, no. 3, pp. 23–50, December 2005.

[14] ——, "Towards Capturing and Enhancing Entertainment in Computer Games," in *Proceedings of the 4th Hellenic Conference on Artificial Intelligence, Lecture Notes in Artificial Intelligence*, vol. 3955. Heraklion, Greece: Springer-Verlag, May 2006, pp. 432–442.

[15] G. N. Yannakakis, J. Hallam, and H. H. Lund, "Capturing Entertainment through Heart-rate Dynamics in the Playware Playground," in *Proceedings of the 5th International Conference on Entertainment Computing, Lecture Notes in Computer Science*, vol. 4161. Cambridge, UK: Springer-Verlag, 2006, pp. 314–317.

[16] ——, "Entertainment Capture through Heart Rate Activity in Physical Interactive Playgrounds," *User Modeling and User-Adapted Interaction, Special Issue: Affective Modeling and Adaptation*, vol. 18, no. 1-2, pp. 207–243, February 2008.

[17] P. Rani, N. Sarkar, and C. Liu, "Maintaining optimal challenge in computer games through real-time physiological feedback," in *Proceedings of the 11th International Conference on Human Computer Interaction*, 2005.

[18] S. McQuiggan, S. Lee, and J. Lester, "Predicting User Physiological Response for Interactive Environments: An Inductive Approach," in *Proceedings of the 2nd Artificial Intelligence for Interactive Digital Entertainment Conference*, 2006, pp. 60–65.

[19] G. Andrade, G. Ramalho, H. Santana, and V. Corruble, "Extending reinforcement learning to provide dynamic game balancing," in *Proceedings of the Workshop on Reasoning, Representation, and Learning in Computer Games, 19th International Joint Conference on Artificial Intelligence (IJCAI)*, August 2005, pp. 7–12.

[20] M. A. Verma and P. W. McOwan, "An adaptive methodology for synthesising Mobile Phone Games using Genetic Algorithms," in *Congress on Evolutionary Computation (CEC-05)*, Edinburgh, UK, September 2005, pp. 528–535.

[21] R. Hunicke and V. Chapman, "AI for Dynamic Difficulty Adjustment in Games," in *Proceedings of the Challenges in Game AI Workshop, 19th Nineteenth National Conference on Artificial Intelligence (AAAI'04)*, 2004.

[22] P. Spronck, I. Sprinkhuizen-Kuyper, and E. Postma, "Difficulty Scaling of Game AI," in *Proceedings of the 5th International Conference on Intelligent Games and Simulation (GAME-ON 2004)*, 2004, pp. 33–37.

[23] J. Ludwig and A. Farley, "A learning infrastructure for improving agent performance and game balance," in *Proceedings of the AIIDE'07 Workshop on Optimizing Player Satisfaction, Technical Report WS-07-01*, G. N. Yannakakis and J. Hallam, Eds. AAAI Press, 2007, pp. 7–12.

[24] H. Barber and D. Kudenko, "A user model for the generation of dilemma-based interactive narratives," in *Proceedings of the AIIDE'07 Workshop on Optimizing Player Satisfaction, Technical Report WS-07-01*, G. N. Yannakakis and J. Hallam, Eds. AAAI Press, 2007, pp. 13–18.

[25] D. L. Roberts, C. R. Strong, and C. L. Isbell, "Estimating player satisfaction through the author's eyes," in *Proceedings of the AIIDE'07 Workshop on Optimizing Player Satisfaction, Technical Report WS-07-01*, G. N. Yannakakis and J. Hallam, Eds. AAAI Press, 2007, pp. 31–36.

[26] D. Thue, V. Bulitko, M. Spetch, and E. Wasylishen, "Learning player preferences to inform delayed authoring," in *Papers from the AAAI'07 Fall Symposium on Intelligent Narrative Technologies*. AAAI Press, 2007.

[27] J. Togelius, R. D. Nardi, and S. M. Lucas, "Making racing fun through player modeling and track evolution," in *Proceedings of the SAB Workshop on Adaptive Approaches to Optimizing Player Satisfaction*, G. N. Yannakakis and J. Hallam, Eds., Rome, 2006, pp. 61–70.

[28] ——, "Towards automatic personalised content creation for racing games," in *Proceedings of the IEEE Symposium on Computational Intelligence and Games*. Hawaii, USA: IEEE, April 2007, pp. 252–259.

[29] G. N. Yannakakis and J. Hallam, "A generic approach for generating interesting interactive pac-man opponents," in *Proceedings of the IEEE Symposium on Computational Intelligence and Games*, G. Kendall and S. M. Lucas, Eds., Essex University, Colchester, UK, 4–6 April 2005, pp. 94–101.

[30] ——, "A scheme for creating digital entertainment with substance," in *Proceedings of the Workshop on Reasoning, Representation, and Learning in Computer Games, 19th International Joint Conference on Artificial Intelligence (IJCAI)*, August 2005, pp. 119–124.

[31] G. N. Yannakakis, J. Levine, and J. Hallam, "An Evolutionary Approach for Interactive Computer Games," in *Proceedings of the Congress on Evolutionary Computation (CEC-04)*, June 2004, pp. 986–993.

[32] G. N. Yannakakis and M. Maragoudakis, "Player modeling impact on player's entertainment in computer games," in *Proceedings of the 10th International Conference on User Modeling; Lecture Notes in Computer Science*, vol. 3538. Edinburgh: Springer-Verlag, 24–30 July 2005, pp. 74–78.

[33] F. Hammer, A. Derakhshan, and H. H. Lund, "Adapting Playgrounds for Children Play using Ambient Playware," in *Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS06)*, Beijing, China, October 9–15 2006, pp. 5625–5630.

[34] G. N. Yannakakis, "AI in Computer Games: Generating Interesting Interactive Opponents by the use of Evolutionary Computation," Ph.D. thesis, University of Edinburgh, November 2005.

[35] R. Pfeifer and C. Scheier, *Understanding Intelligence*. Cambridge, MIT Press, 1999.

[36] G. N. Yannakakis, J. Hallam, and H. H. Lund, "Comparative Fun Analysis in the Innovative Playware Game Platform," in *Proceedings of the 1st World Conference for Fun 'n Games*, 2006, pp. 64–70.

[37] J. Doyle, "Prospects for preferences," *Computational Intelligence*, vol. 20, no. 2, pp. 111–136, May 2004.

[38] C.-N. Fiechter and S. Rogers, "Learning subjective functions with large margins," in *ICML '00: Proceedings of the Seventeenth International Conference on Machine Learning*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2000, pp. 287–294.

[39] W. Chu and Z. Ghahramani, "Preference learning with gaussian processes," in *Proceedings of the 22nd International Conference on Machine Learning (ICML-2005)*, 2005.

[40] X. Yao, "Evolving artificial neural networks," in *Proceedings of the IEEE*, vol. 87, no. 9, 1999, pp. 1423–1447.

[41] J. H. Holland, *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI: University of Michigan Press, 1975.

[42] R. W. Picard, E. Vyzas, and J. Healey, "Toward Machine Emotional Intelligence: Analysis of Affective Physiological State," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 10, pp. 1175–1191, 2001.

[43] R. L. Mandryk, K. M. Inkpen, and T. W. Calvert, "Using Psychophysiological Techniques to Measure User Experience with Entertainment Technologies," *Behaviour and Information Technology (Special Issue on User Experience)*, vol. 25, no. 2, pp. 141–158, 2006.