Visualizations for the Assessment of Learning in Computer Games

Lori L. Scarlatos Department of Technology & Society Stony Brook University Stony Brook, NY, USA Lori.Scarlatos@stonybrook.edu

Abstract — Mounting evidence suggests that current trends in global energy usage are leading to global warming, which will likely change our climate irreparably. Yet, as noted in IPCC reports, most people do not take this danger seriously enough to change their behaviors. Computer games, which are increasingly being used for educational purposes, have the potential to change people's understanding and attitudes toward critical issues such as energy use and global climate change. Yet it remains unclear how well serious games achieve these ends, and what, exactly, it is that makes them effective. We propose that by looking at data collected by these games, and correlating it with instruments that measure changes in attitudes, we can determine what game scenarios and activities are actually changing people's minds. This will help us to design more effective games for educating the public in a way that yields tangible results. In this paper we describe a novel strategy for classifying and visualizing the dynamic, multivariate data generated by serious games. Our contribution is a framework for categorizing these data, corresponding to layered visualizations that help to reveal the patterns in what players are doing over time. Specifically, this paper introduces the concept of Action Shapes, which are glyphs that are automatically generated using a variation on parallel coordinates. As elements in the layered visualization, Action Shapes represent the "benificence" of students' choices seen in the contexts of student progress and the overall game state. As proof of concept, we are applying this visualization to Energy Choices, a multiplayer game that teaches people about the interrelated issues of global warming and energy use. Although the examples provided in this paper are specific to this particular game, this strategy may be readily applied to a wide variety of other educational games designed to help people to be smarter about energy use and the planet.

Information visualization; learning assessment; serious games; glyph-based techniques; parallel coordinates

I. INTRODUCTION

It is widely accepted that educational video and computer games have the potential to greatly enhance learning [8, 23]. As a result, numerous organizations have invested heavily in the creation of games for learning [4, 5]. Yet data supporting the efficacy of these investments is currently insufficient. A significant problem is that most studies of learning with these games base their conclusions either on subjective data obtained from surveys, or traditional paper tests [17]. Although tests can

This work has been supported, in part, by an AERTC seed grant.

Tony Scarlatos Department of Computer Science Stony Brook University Stony Brook, NY, USA tony@cs.sunysb.edu

measure what students know, they can't measure how students learned it or to what degree they can apply that knowledge in different scenarios. Meanwhile, the rich set of dynamic data generated by the game is generally ignored. In the few cases where game data is being considered for assessment purposes, the approaches taken - e.g. model tracing, Bayesian, or artificial neural networks - provide only limited insights into how the learning is taking place. The problems with extracting useful assessment data using current toolsets are compounded when considering games that have multiple players, and that are played multiple times.

We are addressing this deficiency by developing visualizations that enable instructors and instructional designers to see patterns in how students acquire and apply knowledge from games. A novel aspect of our approach is the way that we categorize game data and represent those categories in visual layers, enabling the user of the toolset to see interrelationships among the variables. Although we are developing this toolset to initially work with our Energy Choices game, this approach may be applied to a broad range of learning games, and even other applications where the combined choices of multiple players impact both personal and global properties. Extended use of this toolset will contribute to an emerging understanding of how students learn from games, and how we can use that knowledge to create more effective learning tools. Ultimately, the insights gained by using this toolset will lead to the development of algorithms that can automatically detect patterns that represent degrees of learning.

The contributions of this paper are twofold. First, we present a framework for categorizing educational game data, so that it may be viewed simultaneously in layers in a visualization. Although we are currently working with one particular game, this framework is general enough to be applicable to all learning games. Second, we describe an algorithm for automatically generating what we call Action Shapes. These glyphs use a variation on parallel coordinates to form a readily recognizable, and intuitively understandable, representation of players' actions in the game. Unlike star glyphs, our action shapes are distinguishable by how well rounded they are, and are therefore not orientation dependent. Altogether this work has the potential to transform the way that instructors and instructional designers assess learning in computer games, by giving them a new way to "see" how player choices, game outcomes, and learning outcomes all fit together in the game environment.

II. BACKGROUND

Traditionally, the educational impact of learning games has been assessed primarily using either standardized tests or selfreporting mechanisms such as surveys. For decades, Item Response Theory has provided the most reliable means of measuring what students know [10]. It is therefore easy to see why this continues to be the most common way of determining what students know after playing a game. Yet most of these testing methods fail to conclusively show what about the gaming experience leads to learning [9]. The problem is that these assessments can only gauge levels of knowledge and comprehension; other essential learning objectives defined in Bloom's Taxonomy (application, analysis, synthesis and evaluation) are ignored [1].

This deficiency has led to efforts to learn more from the data generated by the game. In Intelligent Tutoring Systems, model tracing is typically used to determine how well a student is doing, based on choices made [29]. This has been extended to the realm of games, where student performance is reflected in "information trails" based on the sequence of choices that students make [14]. Bayesian networks have also been used to characterize what is being learned [15]. Yet building a Bayesian network that considers all variables is NP-complete [16], and so the assessment of game data must start with clear assumptions to guide development. The IMMEX project at UCLA gets around this by using a combination of item response theory, artificial neural networks, and hidden Markov modeling to estimate students' problem-solving abilities based on what information resources they use, and how successful they are at solving specific problems with a simulator [24]. Yet although each of these methods can indicate how well students are doing on a given task, they do not reflect what actions make them most successful, and which aspects of the game are most effective for enhancing their learning.

We propose that visualizations of game data can lead to greater understanding of the complex information space than automated approaches alone. This is because visualizations can reveal patterns among the gameplay actions and outcomes to those who are most familiar with the learning objectives: instructors and instructional designers. Seeing these patterns can then lead to insights that might be used to improve the algorithms for automatically analyzing the data. Although the range of available visualization techniques is vast, we focus on a few that are most promising for our particular problem.

Glyph-based techniques are useful because they can encapsulate multivariate data in a simple symbol or shape. They can also support quantitative analysis by taking advantage of hybrid visualizations to provide context [19]. In addition to varying color or size, glyph shapes can represent values such as student activity in an e-learning system [6] or point data on a map [12]. Star glyphs incorporate even more dimensions by using star points to represent differing data values [28].

Parallel coordinates [11] provide an elegant way of representing multidimensional data in a single view. This technique has been adapted to a variety of visualizations, from visualizing trends in computer science retention [30] to the visualization of multivariate particle acceleration data [21].

Layering and separation allow viewers to see very complex data in a way that fosters focusing on one particular aspect while keeping the context in sight [27]. This technique most commonly appears in maps, where topographic, hydrographic, transportation line and political boundary data can be seen together using different color schemes to represent the different layers.

Changes over time can be visualized using either timebased animations or small multiples [27]. Animation is an effective way of providing a viewer with an overview of the temporal coherence of data values in relation to one another. This has been used to great effect in a variety of projects, from Rosling's Gapminder [20] to Langelier et. al.'s visualizations of software quality [13]. However, because they change so rapidly, animations are not necessarily well suited for detailed analysis. Robertson et. al. have, in fact, found that static views with small multiples are better for this [18]. It is therefore useful to make both approaches available.

III. FRAMEWORK FOR VISUALIZING GAME DATA

Every computer game keeps track of multivariate data, in a variety of categories, while the game is being played. We contend that all of these data might be relevant to discovering what is being learned, and how, as the game is played. We have therefore identified five broad categories of educational game data which can be used to partition the complete set of data. They may also be represented in differing layers of a visualization. These categories are:

- **Player choices** What the player does is evidence of the player's approach to solving problems posed by the game. Player choices may include purchases made, an approach to solving a problem, or even choosing to tackle a problem at all.
- **Player state** Properties associated with the player represent the current state of the player in the game. They may include wealth, health, power, and resources held.
- Player performance How well a player is doing is often reflected in a score which may or may not be multidimensional. This can include the time it takes to solve given tasks, as well as factors based on player properties.
- Game state Every game is played within some virtual environment with its own set of properties. These properties may include availability of resources, as well as factors such as danger or temperature. Over time, actions of the players as well as random events generated by the game can collectively change those properties.
- External events Gameplay is also impacted by various events that can cause players to take different strategies or view the game in a different way. These events may be player-initiated, such as conversations between players and the viewing of supplemental information. They may also include random events introduced by the game system, such as natural disasters.

IV. VISUALIZATION DESIGN

We are applying our framework to a multiplayer online game that we developed called Energy Choices [22]. In this game, players take on the roles of enlightened despots ruling the 25 most populous countries in the world. Given the ability to decide how their gross domestic product (GDP) is spent, and what types of energy are purchased, each player's goal is to make their country more prosperous (i.e. increasing GDP per



capita) without destroying the planet (i.e. raising the average global temperature through carbon emissions). Figure 1 shows a snapshot of the main gameplay screen.

This game is designed to be played as part of a high school environmental course or an undergraduate course in general science, where students are learning about the science behind global climate change. In assessing both the students' learning and the effectiveness of the game as an educational tool, we would like to be able to "see" how 1) choices made by multiple players impact the world (game environment) and each other (individual countries); 2) player choices, player performance, and game state change over the course of time as the game is being played; and 3) student choices and resulting outcomes change over time, as the game is played multiple times.

The Energy Choices game is an extension of an agent-based simulation developed with Repast [3]. The advantage of basing the game on such a simulation is that all 25 countries always participate in determining the outcome of the game Energy Choices, no matter how many players are participating. In addition, Energy Choices can be run as a simulation by a single operator. The back-end of the game is written in Java and runs on a server, where state is maintained in a MySQL database. This enables us to save a complete history of every game that is played. The front-end was developed using Adobe Flash, and runs locally in a web browser, receiving XML data from the server. To facilitate rapid prototyping, we have also generated the visualizations for this paper with Flash.

Our approach is to create dynamic visualizations composed of multiple layers of information representing the different types of game data identified above. Visual distinctions between the layers allow analysts to focus on a particular type of data without losing its context within the other layers. Animation allow for the detection of trends as glyphs move around the data space. Following are descriptions of the various layers being visualized.

A. Player Choices

We represent player choices with a variation on parallel coordinates (Inselberg) that we call Action Shapes. We use shapes instead of lines because they are easier to recognize. For a set $C = \{c_1, c_2, ..., c_N\}$ of all choices available to the player, we represent the choices made by player i with an Ndimensional point $P_i = \{p_{il}, p_{i2}, ..., p_{iN}\}$ where each coordinate p_{ii} is normalized to the range $[0.0 \rightarrow 0.1]$. These coordinates are mapped onto N axes $X_1, X_2, ..., X_N$, each perpendicular to the Xaxis on the *xy*-plane with an equal distance of ∂_h between each axis, forming the top of a 2-D polygon. The bottom of the polygon is formed by the line $y=y_{min}$, where $y_{min}=-\partial_{y}$, which ensures that the resulting shape will be perceived as a single polygon even if $p_{ii}=0$. The sides of the polygon are formed by the line segments $\{P_0P_1\}$ and $\{P_NP_{N+1}\}$, where $P_0=(x_1,-\partial_v)$ and $P_{N+1}=(x_{N},-\partial_{\nu})$. The Y-axis is scaled such that the distance from $0 \rightarrow 1$ is

$$\frac{(N-1)\delta}{y_{max} - y_{min}}$$
(1)

so that a polygon with $p_{ij} = y_{max} = l$ can be bounded by a square. This is illustrated in Figure 2.

We found that the ordering of axes X_1 , X_2 , ..., X_N , is important for helping analysts to quickly distinguish "beneficent shapes" (i.e. shapes representing actions leading to positive outcomes) from "maleficent shapes" (i.e. shapes representing actions leading to negative outcomes). Convex or



rounded shapes appear balanced and harmonious, while undulating concave shapes appear agitated and discordant. We therefore place axes representing positive choices near the center of the shape, with less desirable choices at the extreme ends. For example, in Energy Choices, players choose how to spend their GDP (on consumption, savings, or energy) and what energy sources to use (renewable versus fossil fuels). Spending more on energy increases overall GDP; investing in infrastructure (savings) improves the factor determining how efficiently energy expenditures are used to achieve this; and using more renewable energy sources reduces carbon footprint. Therefore, spending more on savings, energy, and renewable energy yields a better score for the player. Figure 3 shows two examples of possible action shapes, representing beneficent choices versus maleficent choices.

Action shapes may also be used to represent the aggregate choices of the players over the course of the game. Figure 4 shows an example.

B. Player Properties and Performance

Using the action shapes as glyphs representing individual players, the current state of each player's properties and performance can be shown by mapping the action shapes onto a 2D grid. After experimenting with different representations of the passage of time in a static image, we found that this is most readily understood with small multiples shown moving from left to right. Therefore, we use the horizontal axis to represent



time. We use the vertical axis to represent the player's score (a summary of the player's overall performance at that instance) because that is the most important factor for an instructor to consider in relation to the player's choices. In most cases, the score at time t can be represented as a sum of M weighted factors added to a base score g(t) as follows:

$$s(t) = g(t) + \sum_{i=1}^{M} w_i f_i(t)$$
 (2)

where $f_i(t)$ is the factor at time t and w_i is the weight. The base score depends on the application: for example, it might be that g(t) = s(t-1) or g(t) = 0.

Individual factors, or other player properties, can be shown in context by using them to determine the size and color of the glyphs. Color of the background can also provide context. To create a continuous scale for the background color, we generate HSB values with hues ranging from H_{min} to H_{max} . To keep these colors in the background, we maintain brilliance at 100% and saturation at 25%. Thus, for parameter $e \in [e_{min} \dots e_{max}]$, the background hue is calculated as

$$H(e) = H_{max} \left(1 - \frac{e}{e_{max}} \right)$$
(3)

Figure 5 shows one example of how we applied this to the Energy Choices visualization, producing a static view of the progress of a single player. Here, time corresponds to "years" in the game, with the game starting in the recent past and players making choices after every 5 years or iterations of the simulation. In calculating the score, g(t) = s(t-1), f_1 represents the change in GDP per capita $(f_1(t) = (GPC(t) - GPC(t-1)) / GPC(t-1))$ and f_2 represents a reduction in carbon emissions $(f_2(t) = (CE(t-1) - CE(t)) / CE(t-1))$. In both cases, $w_i = 0.5$. Because $f_1(t)$ and $f_2(t)$ can be negative, it is possible then for the score to drop over time.

Factors contributing to this overall score are also reflected separately in the visualization. Here, the background color represents carbon emissions added to the atmosphere by that player, with H_{min} = green for e = 0 pounds of CO₂ added, and H_{max} = red for $e = e_{max}$ pounds of CO₂ added. Size of the action shapes represents the "size" of the player's GDP per capita, with the initial size representing the initial value, and subsequent changes in size reflecting changes over time.



C. State of the Game Environment

Putting this all together, we can view the actions of multiple players (and their outcomes) several different ways. One way is to vertically stack the visualizations of individual players as small multiples. A second option is to animate the action shapes as they change over time, changing their positions and background colors to represent performance parameters. Actions of all players may be shown simultaneously in an animation that shows what is happening in the game over time. Each player is represented by an action shape representing the current set of choices or actions taken by that player, shown spatially in the context of his/her current performance in the game. Autonomous agents that are not being controlled by a player can be aggregated because they all make the same choices.

In any case, data representing the overall state of the game environment appears quietly in the background. Here, we use the background color to represent an increase in global temperature *t*, measured in Centigrade. As before, we generate HSB values with hues ranging from green for 0°C to red for t_{max} °C. This is calculated using (3), substituting *t* for *e* in the equation.

We can also draw a grey line with the slope representing the change in the price of the fossil fuels. At each iteration of the game, the current price P_i is re-calculated after all countries have made their fuel purchases, using the formula $P_i = 2.9$ $\exp(E_i/E_{rem})$ where E_i is the initial fossil fuel reserve and E_{rem} is the remaining reserve. In the visualization, the slope *m* of the line representing the change is calculated based on current price (P_i) and the previous year's price (P_{i-1}) as follows:

$$m = \left(\frac{P_i - P_{i-1}}{P_{i-1}}\right) \tag{4}$$

D. Aggregate Data

Analyzing game performance data often requires looking at many sessions involving many students over time. This makes it necessary to aggregate the data. Aggregate action shapes can be created by combining the action shapes of all players, as shown in figure 4. An aggregate action shape may also be generated by getting the sum of all player choices, weighted by their influence on the system. In other words, if a player's country generates 15% of the global GDP, then the weight for that player's choices is 0.15. So, the cumulative actions of all players can be represented by an N-dimensional point WP ={ $wp_1, wp_2, ..., wp_N$ } where value of each world point coordinate wp_i is calculated as

$$wp_{j} = \sum_{i=1}^{M} w_{i}p_{ij}$$
(5)

We can then visualize the entire game using the same approach we use for individual players. Once again, several games may be viewed simultaneously by stacking them vertically. Figure 6, for example, compares data from two runs of the Energy Choices simulation, where the choices only change when necessary. The top shows the results of always choosing the cheapest fuel; the bottom shows the results of always spending half of the energy budget on fossil fuels, and half on renewable fuels. In both cases, the goal is to maintain the original GDP growth rate. In these visualizations, the background color represents an increase in average global temperature, the vertical axis represents global GDP per capita, and the horizontal axis represents time.

V. CONCLUSIONS

In this paper, we described 1) our novel framework for considering game data, 2) our algorithm for generating action shapes automatically, and 3) our approach to using this in visualizations where data categories are visually separated by layering.

We are currently in the process of building an interface for instructors that will enable them to view and analyze the data generated by one or more game sessions. With this interface, instructors will be able to focus in on the actions of a particular student, seeing all the details of his or her actions and their consequences. External events (such as class discussions and extra help obtained) will appear as annotations in the visualizations. Tools for controlling the animations will also be included. This interface will be used by instructors using Energy Choices in their undergraduate general science classes at two different universities in the fall semester. Feedback from the instructors will help us to fine-tune the interface as well as the visualizations. We believe that they will also help us to make the game a more effective learning tool.

We would also like to see this approach used in the assessment of other educational games and simulations. As we extend our Energy Choices game framework to other learning tasks, the visualization tools will be an integral part of it.

ACKNOWLEDGMENT

We wish to thank our collaborators on the Energy Choices game: Micha Tomkiewicz, Aneesha Bulchandani, Karthik Srinaivasan, and Parag Naik. We would also like to thank Dave Ferguson for his continuing support. This project was funded in part by a seed grant from the Advanced Energy Research and Technology Center (AERTC).



References

- [1] Bloom, B. S. (1956). Taxonomy of educational objectives. Handbook I: Cognitive domain. New York: David McKay.
- [2] Cheng, I., Rodriguez, S. and Basu, A. (2009) Multimedia and games incorporating student modeling for education, *International Workshop* on *Technology for Education*, IEEE, 91-94.
- [3] Collier, N. (2001) Repast: an estensible framework for agent simulation, *Natural Resources and Environmental Issues*, vol. 8, article 4.
- [4] Federation of American Scientists, Summit on Educational Games: Harnessing the Power of Video Games for Learning, 2006, downloaded 8/30/07 from http://www.fas.org/gamesummit/Resources/Summit%20on%20Educatio nal%20Games.pdf.
- [5] Fletcher, J.D. (2009) Education and Training Technology in the Military, *Science 2*, vol. 323 no. 5910, 72-75.
- [6] France, L., Heraud, J., Marty, J., Carron, T. and Heili, J. (2006) Monitoring virtual classroom: Visualization techniques to observe student activities in an e-learning system, *Proceedings of the Sixth IEEE International Conference on Advanced Learning Technologies*, 716-720.
- [7] Freeman, J.B., Rule, N.O. and Ambady, N. (2009) The cultural neuroscience of person perception, in Chiao, J.Y. (ed.), *Progress in Brain Research*, vol 178, Elsevier B.V., 191-201.
- [8] Gee, J.P. (2003). What Video Games Have to Teach Us About Learning and Literacy, Palgrave Macmillan, New York.
- [9] Gosen, J. and Washbush, J. (2004) A review of scholarship on assessing experiential learning effectiveness, *Simulation & Gaming*, 35(2), 270-293.
- [10] Hambleton, R.K. (1991) Fundamentals of Item Response Theory, Sage Publications.
- [11] Inselberg, A. and Dimsdale, B. (1990) Parallel coordinates: a tool for visualizing multi-dimensional geometry, *IEEE Visualization* '90, 361-378.
- [12] Karnick, P., Jeschke, S., Cline, D., Razdan, A., Wentz, E. and Wonka, P. (2009). A Shape Grammar for Developing Glyph-based Visualizations, *Computer Graphics Forum*, 28(8), 2176-2188.
- [13] Langelier, G., Sahraoui, H. and Poulin, P. (2008) Exploring the evolution of software quality with animated visualization, *IEEE Symposium on Visual Languages and Human-Centric Computing 2008*, 13-20.
- [14] Loh, C.S. (2007) Designing Online Games Assessment as "Information Trails", in *Games and Simulation in Online Learning*, Gibson, Aldrich and Prensky editors, Idea Group, Inc., 323-348.
- [15] Mislevy, R.J., Almond, R.G., Yan, D. and Steinberg, L.S. (2000) Bayes Nets in Educational Assessment: Where Do the Numbers Come From? UCLA Center for the Study of Evaluation, technical report 518.
- [16] Niedermayer, D. (1998) An Introduction to Bayesian Networks and their Contemporary Applications.

- [17] Quellmalz, E.S., Timms, M.J. and Schneider, S.A. (2009) Assessment of Student Learning in Science Simulations and Games, WestEd report to the National Academies, downloaded 12/3/09 from http://www7.nationalacademies.org/bose/Schneider_Gaming_Commissi onedPaper.pdf.
- [18] Robertson, G., Fernandez, R., Fisher, D., Lee, B. and Stasko, J. (2008) Effectiveness of Animation in Trend Visualization, *IEEE Transactions* on Visualization and Computer Graphics 14, 6 (2008), 1325–1332.
- [19] Ropinski, T. and Preim, B. (2008) Taxonomy and Usage Guidelines for Glyph-based Medical Visualization, *Proceedings of the 19th Conference* on Simulation and Visualization (SimVis08), 121-138.
- [20] Rosling, H., Rosling, A., Rosling, O. (2006) New software brings statistics beyond the eye. In *Statistics, Knowledge and Policy: Key Indicators to Inform Decision Making*, OECD Publishing, Paris, France, 522–530.
- [21] Rubel, O., Prabhat, Wu, K. Childs, H., Meredith, J., Geddes, C.G.R., Cormier-Michel, E., Ahern, S., Weber, G.H., Messmer, P., Haga, H., Hamann, B. and Bethel, E.W. (2008) High performance multivariate visual data exploration for extremely large data, *Proceedings of the 2008* ACM/IEEE conference on Supercomputing.
- [22] Scarlatos, L.L., Tomkiewicz, M., Bulchandani, A., Srinaivasan, K.A. and Naik, P. (2009). Intelligent Energy Choices, *Proceedings of the 20th IASTED International Conference on Modelling and Simulation* (MS 2009).
- [23] D.W. Shaffer (2006). *How Computer Games Help Children Learn*, Palgrave Macmillan, NY.
- [24] Stevens, R., Beal, M. and Sprang, M. (2009). Tracking the Development of Problem Solving Skills with Learning Trajectories, *Proceedings of the* 17th International Conference on Computers in Education [CDROM]. Hong Kong: Asia-Pacific Society for Computers in Education.
- [25] Stevens, R. (2007) Machine Learning Assessment Systems for Modeling Patterns of Student Learning, in *Games and Simulation in Online Learning*, Gibson, Aldrich and Prensky editors, Idea Group, Inc., 349-365.
- [26] Thomas, J.J. and Cook, K.A. editors (2005) Illuminating the Path: Research and Development Agenda for Visual Analytics, IEEE, downloaded 1/11/10 from http://nvac.pnl.gov/docs/RD Agenda VisualAnalytics.pdf.
- [27] Tufte, E.R. (1990) Envisioning Information, Graphics Press.
- [28] Ware, C. (2000) Information Visualization: Perception for Design, Academic Press, Morgan-Kaufman Publishers.
- [29] Woolf, B.P., Beck, J., Eliot, C. and Stern, M. (2001) Growth and Maturity of Intelligent Tutoring Systems: a Status Report, in *Smart Machines in Education*, Forbus, K.D. and Feltovich, P.J. editors, MIT Press.
- [30] Wortman, P. and Rheingans, P. (2007) Visualizing trends in student performance across computer science courses, *Proceedings of the 38th* SIGCSE technical symposium, 430-434.