Proceedings, The Thirteenth AAAI Conference on
Artificial Intelligence and Interactive Digital Entertainment (AIIDE-17)

A Skill-Based Framework for the Generation and Presentation of Educational Videogame Content

Britton Horn
Northeastern University
College of Computer and Information Science
bhorn@ccs.neu.edu

Abstract

We regularly encounter complex activities consisting of basic skills—both conscious and subconscious. Adequately performing these complex activities involves mastering the individual basic skills and having the ability to seamlessly integrate them together. Games are one such example of a complex activity that is difficult to break down into the basic skills required, but engagement in games relies on designers introducing challenges proportionate to a player's skill. Procedurally generated levels cause additional problems since it is hard to estimate level difficulty for a particular player. This proposal suggests a framework for determining the skills necessary to successfully complete a game, creating AI-based bots with those skills to reflect players with the same skills, and identifying and generating optimal orderings of levels to promote learning each skill of a game. The proposed framework will be implemented in three citizen science games—*Paradox*, *Foldit*, and *Nanocrafter*—and one computer science educational game called *GrACE*.

Introduction

Each day we perform complex tasks requiring both mastery of complex and individual underlying skills. Combining and mastering skills is integral to learning and performing in all aspects of life including driving, reading, or playing a game. Driving requires independently mastering braking, steering, and switching gears, but also integrating and fluently switching between them all. Complex skills have an ideal hierarchy in which to learn the constituent basic skills: we have to learn counting before we can learn addition and subtraction, and it is easier and more useful to learn addition before multiplication (Baralt, Gilabert, and Robinson 2014; Duschl, Maeng, and Sezen 2011; Gagne 1968; Méheut and Psillos 2004; White and Gagné 1974). Often it is difficult to identify the individual component skills required to complete a complex cognitive task (e.g. playing a videogame), how these component skills integrate, and the optimal order to present challenges in a learning environment to promote mastery of each skill.

A challenge for both game designers and artificial intelligence (AI) researchers is accurately understanding the skills a player has mastered in a particular game. Designers care about a player's specific set of skills because they want to create levels to match that player's skills and create level progressions that promote learning key features and mechanics of the game. If levels are too difficult or too easy, the player can become frustrated, bored and disinterested. AI researchers model player skill to predict player performance in new environments and challenges. Accurate player modeling helps create more engaging games by creating content tailored for a specific player at a particular point in time (Bakkes et al. 2014).

Many educational games exist and evaluating their educational success is a challenging endeavor (Harteveld et al. 2014). Games come in many genres with different target populations and various research goals such as promoting learning, engagement, enjoyment, or diversity. Evaluation and assessment techniques and materials that work for one game may not be appropriate for others. There are implications to the order of skill introduction and assumption of pre-existing skills on player behavior and performance in educational games. This proposal aims to create a framework that helps designers and researchers outline expected incoming knowledge and the desired learning outcomes of a game as well as methods to evaluate if the process is successful.

The proposed work primarily contributes to the fields of procedural content generation and game design by developing a framework implemented in multiple games for extracting the individual skills required for a game and their relationships, designing AI's with those skills (Holmgaard et al. 2014), identifying and generating optimal orderings of challenges that promote mastering a particular set of skills, and evaluating player performance and learning outcomes. Additionally, this research can identify skills present in multiple games that transfer between environments (e.g. using a mouse and keyboard) and information about why certain players struggle at key points in games where others may not. Contained in this research is a survey of existing Computer Science educational games including the CS material taught, game genre, and extent the player is required to grasp each CS concept in the game.

This proposal aims to answer three main research questions:

- How can the skills necessary to complete a videogame be extracted and organized by dependencies?
- How can video game levels be generated to accommodate a particular set of skills and promote learning?
- How can player skills in video games be abstracted such that they can be compared across games or are there a certain set of identical skills present across multiple games?

The proposed work of this thesis is to create a game agnostic method for authoring, evaluating and generating levels with AI-based bots called StrataBots (hierarchical bots with skills ranging from optimal to sub-optimal) that is tested and evaluated across multiple games. Generating levels for specific strategies is a relatively new field so the first step of this proposal will be creating a level generation system in an educational game called *GrACE* (Horn, Clark, et al. 2016; Horn, Hoover, et al. 2016; Folajimi et al. 2016) that produces increasingly difficult levels based on StrataBot assessment and player performance. Preliminary work analyzing player and StrataBot performance suggests that levels requiring advanced strategies to complete correlate with lower player success rates. Future work will devise methods to create ideal progressions for each player by first generating levels that players can complete with simple or novice strategies. Then based on player performance, the level generation system will advance players along a difficulty progression to allow better players to be served more difficult levels at an earlier stage than novice players.

Level difficulty will then be assessed in *Paradox*—a citizen science game—with StrataBots based on skill atoms, rather than hand authored strategies. To create *Paradox* StrataBots, we will first develop a search-based StrataBot framework that finds strategies through novelty search and designer-specified game mechanics. Novelty search differs from traditional search optimization techniques by searching the space based on a distance metric and archive of novel solutions designed to avoid converging to a particular local-optima. Additionally, having the designer specify mechanics rather than a complete list of potential strategies will help alleviate demand on the designer and allow a novelty search to find potentially more possible strategies than the designer can imagine alone. We will evolve neural network controllers that input game state and output a desired move to find novel controllers. We will run the controllers on a large selection of game puzzles and each time a controller is created that produces sufficiently different behavior, we will add it to a pool of controllers and restart the process until no additional controllers can be produced. Variation between controllers will be measured with edit distance similar to the player strategy analysis conducted on *GrACE* player data.

Finally, we will create a level generator for *Paradox* with StrataBots to demonstrate the application of a framework that formalizes the authoring and evaluation of StrataBots outside of *GrACE*. To further refine this model and prove it is viable for other game types, we will repeat this process for two additional citizen science games—*Nanocrafter* and *Foldit*. We will produce skill chains (Church 2006; Cook 2007; Deterding 2015; Koster 2013) for *Nanocrafter* and *Foldit* during this process which we will analyze to find similarities in skill atoms and determine the skills that transfer between games.

The StrataBot framework will be quantitatively evaluated on its ability to accurately reflect expected strategies as well as produce novel strategies. StrataBots produced with the framework outline in this proposal will be compared to player data collected from players playing online versions of *GrACE*, *Paradox*, *Nanocrafter*, and *Foldit* and compared using a modified edit distance. When comparing the edit distances between players and StrataBots, we expect to see that at least one player's strategy is reflected by a StrataBot through a low edit distance. To determine the extent to which novel StrataBots can be produced, we also expect to find some StrataBots that do not closely match any player. Playtraces from these StrataBots will be analyzed qualitatively in more detail to understand if the produced strategy can be described and replicated manually.

Levels generated with the help of the StrataBot framework for *GrACE* will be evaluated for accurate difficulty progression by quantitatively comparing player success rates with StrataBot predicted success rates. Levels that allow more StrataBots to complete them are assumed to be easier and increase expected player success. Additionally, a player's mastery of previous levels requiring a highly complex strategy increases expected success.

Finally, we will qualitatively analyze skill chains produced for *Paradox*, *Nanocrafter*, and *Foldit* to determine the skills that appear in each game's skill chain. We expect that the repeated skills will be interaction based skills such as using a mouse, clicking buttons, and understanding high scores. The skill atoms we expect to not see repeat are atoms that describe combinations of game-specific mechanics such as particular strategies for completing *Foldit* levels.

References

Andersen, Erik, Sumit Gulwani, and Zoran Popovic. 2013. A Trace-Based Framework for Analyzing and Synthesizing Educational Progressions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 773–782. ACM.

Bakkes, Sander, Shimon Whiteson, Guangliang Li, George Viorel Vişniuc, Efstathios Charitos, Norbert Heijne, and Arjen Swellengrebel. 2014. Challenge Balancing for Personalised Game Spaces. In *Games Media Entertainment (GEM), 2014 IEEE*, 1–8. IEEE.

Baralt, Melissa, Roger Gilabert, and Peter Robinson. 2014. An Introduction to Theory and Research in Task Sequencing and Instructed Second Language Learning. *Task Sequencing and Instructed Second Language Learning*, 1–50.

Church, Doug. 2006. Formal Abstract. The Game Design Reader: A Rules of Play Anthology. MIT Press, Cambridge

Clark, Richard E., D. Feldon, Jeroen JG van Merriënboer, Kenneth Yates, and Sean Early. 2008. Cognitive Task Analysis. *Handbook of Research on Educational Communications and Technology* 3: 577–593.

Cook, Daniel. 2007. *The Chemistry of Game Design*. http://www.gamasutra.com/view/feature/1524/the_chemistry_of_game_design.php.

Deterding, Sebastian. 2015. The Lens of Intrinsic Skill Atoms: A Method for Gameful Design. *Human-Computer Interaction* 30 (3–4): 294–335.

Duschl, Richard, Seungho Maeng, and Asli Sezen. 2011. Learning Progressions and Teaching Sequences: A Review and Analysis. *Studies in Science Education* 47 (2): 123–182.

Folajimi, Yetunde, Britton Horn, Jacqueline Barnes, Amy Hoover, Gillian Smith, and Casper Harteveld. 2016. A Cross-Cultural Evaluation of a Computer Science Teaching Game. In *Proceedings of Games+Learning+Society*. Pittsburgh, PA: ETC Press.

Gagne, Robert M. 1968. Presidential Address of Division 15 Learning Hierarchies. *Educational Psychologist* 6 (1): 1–9.

Harpstead, Erik, and Vincent Aleven. 2015. Using Empirical Learning Curve Analysis to Inform Design in an Educational Game. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*, 197–207. ACM.

Harteveld, Casper, Gillian Smith, Gail Carmichael, Elisabeth Gee, and Carolee Stewart-Gardiner. 2014. A Design-Focused Analysis of Games Teaching Computer Science. In *Proceedings of Games+Learning+Society 10*. Madison, WI.

Holmgaard, Christoffer, Antonios Liapis, Julian Togelius, and Georgios N Yannakakis. 2014. Evolving Personas for Player Decision Modeling. In *Computational Intelligence and Games (CIG), 2014 IEEE Conference on*, 1–8. IEEE.

Horn, Britton, Christopher Clark, Oskar Strom, Hilery Chao, Amy J Stahl, Casper Harteveld, and Gillian Smith. 2016. Design Insights into the Creation and Evaluation of a Computer Science Educational Game. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, 576–581. ACM.

Horn, Britton, Amy K Hoover, Jackie Barnes, Yetunde Folajimi, Gillian Smith, and Casper Harteveld. 2016. Opening the Black Box of Play: Strategy Analysis of an Educational Game. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*, 142–153. ACM.

Koster, Raph. 2013. *Theory of Fun for Game Design*. O'Reilly Media, Inc.

Méheut, Martine, and Dimitris Psillos. 2004. Teaching–learning Sequences: Aims and Tools for Science Education Research. *International Journal of Science Education* 26 (5): 515–535.

Pitkanen, Jori. 2015. Studying Thoughts: Stimulated Recall as a Game Research Method. In *Game Research Methods*, edited by Petri Lankoski and Staffan L Björk, 117–132. ETC Press.

Smith, Adam M, Erik Andersen, Michael Mateas, and Zoran Popović. 2012. A Case Study of Expressively Constrainable Level Design Automation Tools for a Puzzle Game. In *Proceedings of the International Conference on the Foundations of Digital Games*, 156–163. ACM.

White, Richard T, and Robert M Gagné. 1974. Past and Future Research on Learning Hierarchies 1. *Educational Psychologist* 11 (1): 19–28.

Yates, Kenneth A., and David F. Feldon. 2011. Advancing the Practice of Cognitive Task Analysis: A Call for Taxonomic Research. *Theoretical Issues in Ergonomics Science* 12 (6): 472–495.