Data Games : Towards generating content using the real world

Gabriella A. B. Barros

IT University of Copenhagen, Denmark,
gbar@itu.dk

Abstract. New data is added to the Internet by the minute [10, 33]. This project aims at expanding the concept of data games, which explore the use of open data in procedural content generation. To do so, it tackles challenges in data acquisition, transformation and balancing. Data games have the potential of serving as data visualization tools for open data, while also providing new sources of inspirations for content generation.

Keywords: data games, procedural content generation, data analysis and visualization

1 Introduction

The use the real-world as inspiration for content generation is not novelty. Designers of strategy and action games often look at our world as a muse when constructing environments. Simulation games, like *"Flight Simulator"* (Microsoft, 1982), also mimic real life (e.g. their vehicles and the way players interact with them). However, all of these required designers to manually translate real ideas into the game world, usually reshaping data in order to make it more coherent with a game [12].

Although using our physical world as inspiration is not a new concept, the idea of games that allow players to freely interact with, explore and learn from real data was only introduced in 2013 [11]. Data games use real-world information – in special, contextual information and open data [13] – for creating their game's content and/or mechanics. They allow for players to view, learn and interact with the original data in a playful manner [12, 6]. On the other hand, the amount of information available on the Internet make of it a nearly inexhaustible source of inspiration for Procedural Content Generation.

The significance of this goes beyond making games more fun. If we base our games on real data, we might also learn about the world as we play. A game could showcase particular parts or facets of the world's geography, simply tweaking data selection. Interacting with the game could mean interacting with a faithful representation of aspects of the real world. In this way, games can act as visualization tools, providing more playful means to learn and interact with information that would otherwise be tedious or difficult to decipher.

1.1 Problem statement

This project aims at exploring the use of open data to automatic generate content for games, and how players can view and learn about said data via the games. A series of challenges emerge from this assumption, of which we can highlight: (1) data acquiring (i.e. where and how to obtain data), (2) data transformation (i.e. how to best transform data in game content, which can be very dependent on the data type, the content type and/or the game genre), (3) experience/learning balancing (i.e. how to provide an interesting game experience while ensuring that data is understandable, and minimize adding or modifying the data in a "imprecise" way).

1.2 Contributions

The possible contributions of this work are threefold. From the point of view of PCG research, the use of open data can provide nearly infinite source of inspiration for creating content, contributing to the PCG research in the Academia as well as practical methods for game development. On the other hand, learning and analyzing information is hardly a trivial matter, and games can help as more intuitive and easily understood tools, which is not only in the Academia best interest, but also useful for government and industrial purposes. Finally, this subject has hardly been explored before, with only a couple of years of research. Therefore, there are many opportunities within this area to be explored.

2 Related Work

This project is related to multiple areas, including procedural content generation and interactive narrative. This section presents a brief overview of the following fields: Procedural level generation, procedural complete game generation, interactive narrative and data games. It will not cover all areas of PCG, for it is a too broad subject for the purpose of this paper. For a comprehensive review on the field, please refer to (Shaker et al, 2015) [31].

2.1 Procedural Level Generation

The use of procedural content generation (PCG) in different game genres has been studied extensively [31]. Few of those genres do not require a level or map to be played upon – however abstract these may be. To use games as means of visualizing and playing with data, it will be necessary to create levels or maps that can present said information in a meaningful way.

Many different methods for map generation have been proposed before. One popular and simple way, yielding interesting results and fast runtime, is the use of fractals[23, 24], however this does not allow for much control. Dungeon layouts can also be produced by placing various sized rooms and hallways on a two-dimensional area, using as fitness function the length from start to finish [32]. Togelius et al [38] use search based PCG through multiobjective evolution to create maps for StarCraft. StarCraft map generation was also attempted by Uriarte and Ontan, using Voronoi diagrams to define the

initial terrain layout and populating it using metrics [41]. Search-based algorithms [19, 15], specially evolutionary approaches[5, 25, 4], are also popular for map generation, specially for RTS games.

2.2 Complete Game Generation

As mentioned previously, a large part of the work done in PCG involves creating levels and maps; the generation of complete games has revolved around more constrained game genres, such as board- or arcade-games [37]. However, to use data for content generation requires, at times, that the whole game be built automatically from the data itself.

When it comes to generating complete games, most efforts employ search based, solver-based or constructive methods [37]. Evolutionary computation, among others optimization algorithms, are used in search-based game generation, in order to find feasible games within a search space. Browne's Ludi generates board games in this manner, using a weighted sum of many different heuristics to evaluate generated games [3]. It has also been used to generate different arcade-like mechanics-heavy games such as Pac-man-like games [39], platformers [8], or simple arcade games using game definition languages as PuzzleScript [17].

2.3 Interactive Narrative

Stories can be important facilitators of entertainment and education not only in games, because complex information cannot explained without stories [9]. In most games, experience is built over conflict, where resolution is dictated by player's choices. The collection of conflicts, resolutions and choices can be seen as a representation of the game's story. Therefore, a part of the attention given while generating a complete game should also be given to how the story (if any) is conveyed.

Interactive Narrative involves the creation or adaptation of a interactive experience, such as a game or a interactive digital story, in response to the user. It is commonly associated with an omniscient and intelligent agent that can view the narrative's world and modify it to drive the story forward, usually by selecting plot points in a story graph taking into account the user's actions. This process is usually done either by using manually authored models and simulations [29].

Examples of systems focused on authored stories are Haunt 2 [18] and PaSSAGE [35]. Haunt 2 attempts to predict the player's actions to avoid those that lead to problems (e.g. killing a fundamental character). On the other hand, PaSSAGE uses player modelling to select pre-authored story pieces. While these approaches give the user a sense of agency, creating the illusion of control, they are still constrained to the set of designed story pieces. To decrease these constraints, the EM can also contain a complete story generator, which have been extensively researched in the last decades [14, 22, 40, 16, 42, 28, 34, 7].

Relationships between characters have also been explored, mostly in case-based approaches. Prom Week [21] creates the story by using rule based components. MÉXICA

used emotions and relationship models of virtual agents for plot generation [26]. NetworkING [27] develops a plot according to the expressed relationships between characters, while Façade uses a mix of simulation and authored approaches [20].

2.4 Data Games

Although much research on PCG is currently available [31], and its commercial use increases by the day, the use of real data to generate such content is still scarce. Most of the work was done by or involving the same group of individuals, of whom we can highlight Julian Togelius and Marie Gustafsson Friberger, the duo responsible for introducing the concept of Data Games[11, 12]. Examples of data games include Bar Chart Ball [36], a game about a ball atop a bar chart that players can move by choosing different demographic indicators, which alter the chart's appearance. Open Data Monopoly use real demographic and geographical information to create *Monopoly* (Parker Brothers 1935) boards[11]. Open Trumps evolve cards for *Top Trumps* (Dubreq 1977) based on open data about countries.

3 Methodology

To study how different types of data and content interact with each other, we chose to develop smaller projects. To start, we selected a somewhat simple transformation: from maps to map levels. The game genre selected was strategy games, and we used *FreeCiv*[1], a game based on the popular Civilization series. Data was obtained from OpenStreetMaps[2] (OSM) and online image about resource deposits (e.g. oil, gold, fished, etc). OSM renders aerial map images of the world, which are transformed into *FreeCiv* maps. These maps were evolved using a 12+18 evolution strategy to balance game resources positions. This work was published in (Barros and Togelius, 2015) [2].

We then explored a different genre of games: adventure games. The goal of this project is to build a complete adventure game using the relations between Wikipedia articles as the plot. A crawler was developed to discover relations between two given articles, and a parser transform this relations into game objects. We also use OSM and Wikimedia Commons to obtain the geographical and visual content for the game. A preliminary work on this project was published in [1]. Our next steps involve expanding this project, both to improve the game itself, but also to investigate different sub-genres of adventure games, such as dating simulators and mystery-solving adventure games, as well as testing new data types and game mechanics and genres.

Another side-project aimed at creating music and lyrics from academic papers [30], which can be used as soundtrack for a game yet to be implemented.

The experience gathered from the before mentioned projects lead to the proposal of a general framework for procedural content generation in data games, as shown in Figure 1. It consists of two main parts: a *Crawler* and a *Parser*. The Crawler module is responsible for data acquiring, and depends on the type of original information to be

[1] FreeCiv: https://play.freeciv.org/

[2] http://wiki.openstreetmap.org/wiki/JMapViewer

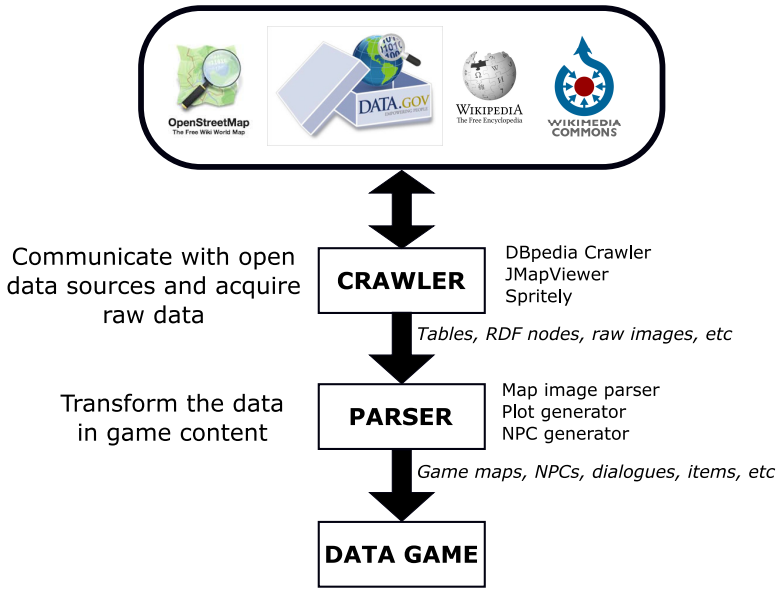

Fig. 1: General framework for content generation in a data game.

obtained. It acts as an interface between the parser and the source of information (e.g. Wikipedia). The Parser must be able to take the raw data obtained, and understand and transform it into whatever game object it shall represent (e.g. items, dialogue, game locations, etc). The final game created depends essentially on the parser, and different parsers can use a same crawler to generate different types of game content.

4 Future work

There is still much to be done in this field. The next steps in this project include further developing the current data adventures' project and studying how other types of content can be generated, in different game genres. Firstly, new mechanics for the Data Adventures must be developed, as its gameplay still has much space for improvement. Secondly, it is necessary to study more types of game genres. At the moment, we ventured more deeply in adventure games, but there are many other genres to be explored. There are also many different types of content not incorporated yet, such as Twitter feeds and governmental data. This will help the validation of our proposed framework – also a future step –, while providing new methods for transforming and acquiring data. This can also shed light in a different challenge: how can absurd non-trivial contents emerge from data? Some transformations are more straightforward than others, such as real world geographical data to game maps. But data types can be applied to a range of different content types. What happens if one uses twitter feed to algorithmically evolve game mechanics? Or economic data numbers to create quests?

This project can also possibly bear contributions to the field of Education. To verify this assumption, it is necessary to perform user studies to evaluate to which level the original data can be understood from the game, and how fun the experience is for the players.

References

1. Barros, G.A.B., Liapis, A., Togelius, J.: Data adventures. In: Proceedings of the FDG Workshop on Procedural Content Generation (2015)
2. Barros, G.A.B., Togelius, J.: Balanced civilization map generation based on open data. In: Proceedings of the 2015 IEEE Congress on Evolutionary Computation. IEEE (2015)
3. Browne, C.: Automatic generation and evaluation of recombination games. Ph.D. thesis, Queensland University of Technology (2008)
4. Cachia, W., Liapis, A., Yannakakis, G.N.: Multi-level evolution of shooter levels (2015)
5. Cardamone, L., Yannakakis, G.N., Togelius, J., Lanzi, P.L.: Evolving interesting maps for a first person shooter. In: Applications of Evolutionary Computation, pp. 63–72. Springer (2011)
6. Cardona, A.B., Hansen, A.W., Togelius, J., Gustafsson, M.: Open trumps, a data game. In: Proceedings of the 9th Conference on Foundations of Digital Games (2014)
7. Cavazza, M., Charles, F., Mead, S.J.: Character-based interactive storytelling. IEEE Intelligent systems (2002)
8. Cook, M., Colton, S.: Multi-faceted evolution of simple arcade games. In: CIG. pp. 289–296 (2011)
9. Crawford, C.: Chris Crawford on interactive storytelling. New Riders (2012)
10. Domo: Data never sleeps 3.0. https://www.domo.com/blog/2015/08/data-never-sleeps-3-0/, Last access: 23/09/2015 (2015)
11. Friberger, M.G., Togelius, J.: Generating interesting monopoly boards from open data. In: Proceedings of the IEEE Conference on Computational Intelligence and Games. pp. 288–295. IEEE (2012)
12. Friberger, M.G., Togelius, J., Cardona, A.B., Ermacora, M., Mousten, A., Møller Jensen, M., Tanase, V.A., Brøndsted, U.: Data games. In: Proceedings of the 8th Conference on Foundations of Digital Games (2013)
13. Kitchin, R.: The data revolution: Big data, open data, data infrastructures and their consequences. Sage (2014)
14. Klein, S., Aeschlimann, J., Balsiger, D., et al.: Automatic novel writing: A status report. Wisconsin University (1973)
15. Lara-Cabrera, R., Cotta, C., Fernandez-Leiva, A.J.: Procedural map generation for a rts game. In: 13th International GAME-ON Conference on Intelligent Games and Simulation. pp. 53–58 (2012)
16. Lebowitz, M.: Story-telling as planning and learning. Poetics 14(6), 483–502 (1985)
17. Lim, C.U., Harrell, D.F.: An approach to general videogame evaluation and automatic generation using a description language. In: Proceedings of the IEEE Conference on Computational Intelligence and Games (CIG). IEEE (2014)
18. Magerko, B.: Evaluating preemptive story direction in the interactive drama architecture. Journal of Game Development 2(3), 25–52 (2007)
19. Mahlmann, T., Togelius, J., Yannakakis, G.N.: Spicing up map generation. In: Applications of evolutionary computation, pp. 224–233. Springer (2012)
20. Mateas, M., Stern, A.: Façade: An experiment in building a fully-realized interactive drama. In: Game Developers Conference. vol. 2 (2003)

21. McCoy, J., Treanor, M., Samuel, B., Mateas, M., Wardrip-Fruin, N.: Prom week: social physics as gameplay. In: Proceedings of the 6th International Conference on Foundations of Digital Games. pp. 319–321. ACM (2011)
22. Meehan, J.R.: Tale-spin, an interactive program that writes stories. In: IJCAI. vol. 77, pp. 91–98 (1977)
23. Miller, G.S.: The definition and rendering of terrain maps. In: ACM SIGGRAPH Computer Graphics. vol. 20, pp. 39–48. ACM (1986)
24. Olsen, J.: Realtime procedural terrain generation (2004)
25. Olsted, P.T., Ma, B., Risi, S.: Interactive evolution of levels for a competitive multiplayer fps. In: Evolutionary Computation (CEC), 2015 IEEE Congress on. pp. 1527–1534 (May 2015)
26. Pérez, R.P.Ý., Sharples, M.: Mexica: A computer model of a cognitive account of creative writing. Journal of Experimental & Theoretical Artificial Intelligence 13(2), 119–139 (2001)
27. Porteous, J., Charles, F., Cavazza, M.: Networking: using character relationships for interactive narrative generation. In: Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems. pp. 595–602. International Foundation for Autonomous Agents and Multiagent Systems (2013)
28. Riedl, M.O., Young, R.M.: Narrative planning: Balancing plot and character. Journal of Artificial Intelligence Research 39(1), 217–268 (2010)
29. Riedl, M.O., Bulitko, V.: Interactive narrative: An intelligent systems approach. AI Magazine 34(1), 67 (2012)
30. Scirea, M., Barros, G.A., Shaker, N., Togelius, J.: Smug: Scientific music generator. In: Proceedings of the Sixth International Conference on Computational Creativity June. p. 204 (2015)
31. Shaker, N., Togelius, J., Nelson, M.J.: Procedural Content Generation in Games: A Textbook and an Overview of Current Research. Springer (2015)
32. Sorenson, N., Pasquier, P.: Towards a generic framework for automated video game level creation. In: Applications of Evolutionary Computation, pp. 131–140. Springer (2010)
33. Stats, I.L.: Total number of websites. http://www.internetlivestats.com,Last access: 23/09/2015 (2015)
34. Theune, M., Faas, S., Heylen, D., Nijholt, A.: The virtual storyteller: Story creation by intelligent agents (2003)
35. Thue, D., Bulitko, V., Spetch, M.: Passage: A demonstration of player modeling in interactive storytelling. In: AIIDE. pp. 227–228 (2008)
36. Togelius, J., Friberger, M.G.: Bar chart ball, a data game. In: Proceedings of the 8th Conference on Foundations of Digital Games (2013)
37. Togelius, J., Nelson, M.J., Liapis, A.: Characteristics of generatable games. In: Proceedings of the FDG Workshop on Procedural Content Generation (2014)
38. Togelius, J., Preuss, M., Beume, N., Wessing, S., Hagelback, J., Yannakakis, G.N.: Multiobjective exploration of the starcraft map space. In: Computational Intelligence and Games (CIG), 2010 IEEE Symposium on. pp. 265–272. IEEE (2010)
39. Togelius, J., Schmidhuber, J.: An experiment in automatic game design. In: Proceedings of the IEEE Symposium On Computational Intelligence and Games. pp. 111–118. IEEE (2008)
40. Turner, S.R.: Minstrel: a computer model of creativity and storytelling (1993)
41. Uriarte, A., Ontanón, S.: Psmage: Balanced map generation for starcraft. In: IEEE Conference on Computational Intelligence in Games (CIG). pp. 1–8. IEEE (2013)
42. Young, R.M., Riedl, M.O., Branly, M., Jhala, A., Martin, R., Saretto, C.: An architecture for integrating plan-based behavior generation with interactive game environments. Journal of Game Development 1(1), 51–70 (2004)