Chapter 9 - GVGAI: What's Next?

Diego Perez-Liebana

In the previous chapters of this book we have described the GVGAI framework and competitions, and also some of the most interesting lines of research that have used this benchmark to tackle GVGAI problems. This is, however, a non exclusive list, as a recent survey shows [6]. We have also seen how this framework has been employed in education, with higher education institutions using GVGAI for undergraduate and postgraduate assignments, final year, master and PhD projects. The more technical chapters also provide a list of exercises that can serve as project ideas to take this further.

Research-wise, most of the attention of the GVGAI framework has been devoted to the single and two-player version of the framework. Chapter 4 in this book outlines the main outstanding problems of this challenge: No studies have achieved better than approximately 50% victory rate with any technique when applying it to sets of 10 or more games. Improvements on methods proposed typically boost performance in a subset of games, but not in all. Therefore, one of the most challenging directions at the moment is to increase play performance in a broad variety of games.

Given that, in most cases adding features that are GVGAI focused leads to uneven improvements, it is sensible to think that a potential way forward is to work on feature extraction that is independent from the framework. An example of this is the work described in Chapter 4, where features are based on the agent's experience instead of the game. Recent advances in deep convolutional neural networks could also be a fruitful line of future work. We have described other possible improvements for single-player planning also in Chapter 4.

In addition, the two-player planning problem adds an extra complexity on the GVGAI challenge. Not only the agents need to adapt to any game is given, but also they need to compete or collaborate with another agent they know nothing about. The current literature shows no study that tries to identify the behaviour of the opponent, with regards to the other player being cooperative or competitive. This type of analysis taps on research on game theory and opponent modelling, being a field of study on its own right. So far, the most involved studies on GVGAI have tried to model the action distribution observed by the other player, using this model for the opponent actions in the forward model [3]. Investigation on more complex opponent models is another line of future work that could yield better results in this problem.

Using GVGAI for AI-assisted game design, as seen in Chapter 7, is another area of fruitful future research. Using agents to automatically test and debug games have attracted the interest in researchers in the last years, for instance using Relative Algorithm Performance Profiles (RAPP), where how good a game is gets evaluated by a measure of the agent performance in them. In a general setting, this can be enhanced using factors like the amount of level being explored by agent, the decisiveness or convergence rate of the agent [8] or the entropy of the moves made.

The GVGAI framework can also be subject to further enhancements. For instance, one of the potential areas of future work is to expand the range of games that VGDL can build. At present, there is no support for games with large state spaces, games that are not avatar-centric (i.e. like Tetris or Candy Crush), board and card games. Although the concepts described in Chapter 8 can alleviate this, there is a scientific interest in making VGDL more complete. Adding more games can also be achieved by making use of the rule generation tracks (see Chapter 6) and the game parameterization tool (Chapter 7).

This can be complemented with adding an integration with other systems. Different general frameworks like like OpenAI Gym [2], ALE [1] or Microsoft Malmö [4] already count on a great number of games (single, multi-player, model free and model based). Interfacing with these systems would increment the number of available games which all GVGAI agents could play via a common API.

Multi-agent games have also drawn people's attention, for instance, real-time strategy games (e.g. StarCraft) and board games (e.g. Mahjong). The study of multi-agent GVGAI is a fruitful research topic. Atari games can also be extended to multi-agent games. In particular, the Pac-Man can be seen as a multi-agent game and related competitions have been held since 2011 [7]. The most recent Ms. Pac-Man vs Ghost Team Competition [9], which included partial observability, was held at CIG in 2016. Nevertheless, a more general multi-agent track is favorable. Another interesting possibility is to enhance these games by providing a larger range of available actions, either simultaneously (as in Atari games or others were game-pad controllers are used) or in a continuous action space (like joystick or steering wheel controllers in racing games).

The upcoming years could also see the development of the GVGAI competition, potentially adding a few more tracks set up as challenges. The following is a non-exclusive list of possibilities for future tracks.

- Game generation track: the goal of this track would be to provide systems that automatically generate completely new games. Concrete installments of this track could provide specific themes, or a sub-space of rules to narrow down the space of possibilities. Ideally, generators would create games from scratch, also this

challenge is open to modification on existing games as a stepping stone towards full-game generation.
- Automatic game design: there is scope for running a competition track on game tuning and play-testing (player experience, game feeling, fun, etc.). This can be in combination to other generation tracks, such as the level, rule or the game generation challenge mentioned in the previous point.
- Multi-Agent GVGAI: In all VGDL games used in the GVGAI competition, the player controls only one agent. An interesting extension of this challenge would be a setting in which the controller decides the moves of a set of agents, for instance in games with more strategic depth like Hero Academy or Starcraft.
- Multi-player GVGAI: Scaling up the two-player GVGAI track to accommodate multiple players is another potential future track for the competition. This would allow for an extended research framework where agents must account for the actions of more than one external agent. Games like Bomberman (where four players play with an avatar each) or Civilization (N players controlling M units each, for a multi-agent multi-player setting) would take part in this track.
- Turing Test GVGAI: The idea of applying a turing test, which in this case determines if a game is being played by a human or by a player, is not new [5]. Adding such a track to GVGAI would allow the community to investigate the general features that make an agent play like a human in any game.

Apart from the framework and the competition, a point of future work is the website that holds the competition[1]. The feedback provided to competition participants can be enhanced with more data analysis on results, visualizations, action entropy, exploration and other features that can be extracted from gameplay, providing a more flexible support for gameplay metric logging. Including the possibility of playing (and replaying) games on the web browser and being able to analyze these features in real-time can help analyze the performance of the agents to build AI improvements.

As a final word, and reflecting on the multiple lines of future work available, we can safely say that GVGAI is very much alive. There is plenty of potential for further development and ideas to take this research further. The framework, being open source, is available for anyone to provide improvements and propose new challenges. We have discussed a few here, and provided lists of extensions and exercises on each chapter, but there is no limit in terms of what other things can be done.

[1] `www.gvgai.net`

References

1. M. G. Bellemare, Y. Naddaf, J. Veness, and M. Bowling, "The arcade learning environment: an evaluation platform for general agents," *Journal of Artificial Intelligence Research*, vol. 47, no. 1, pp. 253–279, 2013.
2. G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba, "Openai Gym," *arXiv preprint arXiv:1606.01540*, 2016.
3. J. M. Gonzalez-Castro and D. Perez-Liebana, "Opponent Models Comparison for 2 Players in GVGAI Competitions," in *Computer Science and Electronic Engineering Conference, 2017 9th*. IEEE, 2017.
4. M. Johnson, K. Hofmann, T. Hutton, and D. Bignell, "The Malmo Platform for Artificial Intelligence Experimentation," in *IJCAI*, 2016, pp. 4246–4247.
5. J. Lehman and R. Miikkulainen, "General Video Game Playing as a Benchmark for Human-Competitive AI," in *AAAI-15 Workshop on Beyond the Turing Test*, 2015.
6. D. Perez-Liebana, J. Liu, A. Khalifa, R. D. Gaina, J. Togelius, and S. M. Lucas, "General video game ai: a multi-track framework for evaluating agents, games and content generation algorithms," *arXiv preprint arXiv:1802.10363*, 2018.
7. D. P.-L. Philipp Rohlfshagen, Jialin Liu and S. M. Lucas, "Pac-Man Conquers Academia: Two Decades of Research Using a Classic Arcade Game," *IEEE Transactions on Computational Intelligence and AI in Games*, 2017.
8. V. Volz, D. Ashlock, S. Colton, S. Dahlskog, J. Liu, S. Lucas, D. Perez-Liebana, and T. Thompson, "Gameplay Evaluation Measures," *Dagstuhl Follow-Ups*, 2017.
9. P. R. Williams, D. Perez-Liebana, and S. M. Lucas, "Ms. Pac-Man Versus Ghost Team CIG 2016 Competition," in *Computational Intelligence and Games (CIG), 2016 IEEE Conference on*, 2016.