Some Observations on Automated Strategy Game Design

Cameron Browne (RIKEN Institute, Tokyo)

The program LUDI, developed for my PhD studies [1], evolved combinatorial board games as symbolic expressions representing structured rule trees. The generation of games proved easy -- the system could evolve thousands of candidates per second using standard genetic programming operators -- but their evaluation proved much more difficult. Evaluation required approximately 30 minutes of self-play analysis per game, and the derivation of appropriate aesthetic indicators made up the bulk of the research effort.

The quality of a game, i.e. its potential to interest human players, cannot be evaluated by its rules alone. Much of a game's appeal lies in the unexpected emergent behaviour that can occur as it is played, which may not be deducible from the rules. For example, the most popular game evolved by LUDI, which it named Yavalath [2], contains two apparently contradictory rules: players win by making 4-in-a-row of their colour but lose by making 3-in-a-row of their colour beforehand. This combination produces an interesting forced move mechanism when played, adding tension and drama to the game and making it interesting for players.

Such emergent strategies can be not just entertaining for players, but crucial for their comprehension of the game. For example, in the recently invented game Omega [3], players' scores are based on the product of group sizes, requiring a degree of calculation that made the game mentally exhausting and difficult to plan ahead. Players found the game opaque and typically made uniformed moves until forced to count the score at the end, removing any tension and making it initially unpopular. It was not until Omega was implemented in the Axiom game system [4] that an emergent strategy was observed. The system's *Monte Carlo tree search* (MCTS) [5] move planner made moves consistent with a strategy of forming its own pieces into groups as close to size 3 as possible, while forming enemy pieces into groups as far from size 3 as possible. This was later proven to be an optimal strategy for the game [3].

This optimal strategy of "form groups of size 3" provided a convenient meme for players that made the game more comprehensible and outlined a concrete plan of action. It also revealed the game to fundamentally be both a connection game and an anti-connection game [6], as players sought to connect enemy groups into larger configurations while blocking their own groups from being so extended, which imported a whole slew of implicit sub-strategies. Players immediately found Omega more accessible and enjoyable through a simple change of perception, and the discovery of this simple strategy saved the game.

Lantz *et al.* introduce the notion of the *strategy ladder* [7], in which players learn increasingly complex strategies relevant to a game as they play it, that build upon each other. They posit that the most interesting games are those with a constant and linearly increasing strategy ladder. This makes good sense as such games would give players both something to play towards (the strategies they know) and something to learn (the strategies they don't know). Games in which winning strategies are trivially learnt and applied would be too simple to be of interest to players, while those in which even the simplest strategies are excessively difficult to learn would be too intractable to be of interest to players. This resonates with the observation by Allis *et al.* that games which survive do so because they provide an intellectual challenge at a level which is neither too simple to be solvable, nor too complex to be incomprehensible [8].

The question then arises: when automatically evaluating games for their potential to interest human players through AI self-play, what level of playing strength will best capture an authentic experience of the game as played between human players? Random play will obviously not simulate the experience of the game as played between intelligent players. Conversely, superhuman AI play could go too far to the other extreme, and give an equally unrepresentative experience of the game as played by human players. For example, Draughts is drawish when played at even the human champion level -- international tournaments have ended with a whimper when finalists drew all 20 games in the final [9] -- but remains an engaging and hugely popular game for the average player worldwide.

There is constant and understandable pressure in the game AI research community to strive for superhuman results in all cases. However, I argue that capping the playing strength of AI agents at a lower (strong human) level for the purposes of game evaluation, is more likely to capture a realistic ``human'' experience of the game. But even estimating what constitutes ``strong human level'' for a given game remains an open question.

References

1. C. Browne, ``Automatic Generation and Evaluation of Recombination Games'', PhD thesis, QUT, Brisbane, 2008.

2. C. Browne, Evolutionary Game Design, Springer, 2011.

3. N. Romeral Andres, ``From Mathematical Insight to Strategy'', Game & Puzzle Design, 1:2, pp. 53--54, 2015.

4. G. Schmidt, ``The Axiom General Purpose Game Playing System", IEEE Transactions on Computational Intelligence and AI in Games (T-CIAIG), 6:4, pp. 332--342, 2014.

5. C. Browne, *et al.*, ``A Survey of Monte Carlo Tree Search Methods'', IEEE Transactions on Computational Intelligence and AI in Games (T-CIAIG), 4:1, pp. 1--43, 2012.

6. C. Browne, Connection Games: Variations on a Theme, AK Peters, 2005.

7. F. Lantz, A. Isaksen, A. Jaffe, A. Nealen, and J. Togelius, ``Depth in Strategic Games'', in The AAAI-17 Workshop on What's Next for AI in Games? (WS-17-15), pp. 967--974, 2017.

8. L. V. Allis, H. J. van den Herik, and I. S. Herschberg, ``Which Games Will Survive?'', in Heuristic Programming in Artificial Intelligence 2: The Second Computer Olympiad, Ellis Horwood, pp. 232--243, 1991.

9. C. Freeling, ``Dameo: A New Step in the Evolution of Draughts?'', Game & Puzzle Design, 1:2, pp. 28--37, 2015.