Game Complexity vs Strategic Depth

Matthew Stephenson, Diego Perez-Liebana, Mark Nelson,
Ahmed Khalifa, and Alexander Zook

The notion of complexity and strategic depth within games has been a long-debated topic with many unanswered questions. How exactly do you measure the complexity of a game? How do you quantify its strategic depth objectively? This seminar answered neither of these questions but instead presents the opinion that these properties are, for the most part, subjective to the human or agent that is playing them. What is complex or deep for one player may be simple or shallow for another. Despite this, determining generally applicable measures for estimating the complexity and depth of a given game (either independently or comparatively), relative to the abilities of a given player or player type, can provide several benefits for game designers and researchers.

There are multiple possible ways of measuring the complexity or depth of a game, each of which is likely to give a different outcome. Lantz et al. propose that strategic depth is an objective, measurable property of a game, and that games with a large amount of strategic depth continually produce challenging problems even after many hours of play [1]. Snakes and ladders can be described as having no strategic depth, due to the fact that each player's choices (or lack thereof) have no impact on the game's outcome. Other similar (albeit subjective) evaluations are also possible for some games when comparing relative depth, such as comparing Tic-Tac-Toe against StarCraft. However, these comparative decisions are not always obvious and are often biased by personal preference. As such, we cannot always say for certain which games are more complex or deep than others. As an example, consider the board games Chess and Go. Chess has more piece types, each with differing movement rules and properties, whereas Go typically has a much larger board, providing a sizeable state and action space. It is unclear how much each of these factors impacts the complexity or depth of each game. Would playing Chess on a larger board make it more strategic to play? Would adding extra rules to Go increase the game's depth or be seen as ruining a beautiful and elegant game? While increasing the complexity of a game can also increase its depth, adjusting certain gameplay factors might have more of an effect than others. Browne suggests that strategic depth should be considered relative to a games complexity [2], and that games which are more complex than others should also possess additional strategic depth.

The number of factors that could potentially influence the complexity or depth of a game is likely to be vast. Common properties might be aspects such as the size of the state space, the branching factor (i.e. action space), the number of rules, deterministic or stochastic, discrete or continuous, the number of players, and so on. Even this small collection of properties poses some problems regarding how they are measured. When determining the number of rules for a game, what description language should be used? How do you compare single-player and two-player games? Should the response time of a human compared to that of an agent be taken into account? We do not have any answers to these questions and any individual opinions are likely to be highly subjective. This also holds for comparing the relative impact of each of these properties. One player might do very well at fully deterministic games that require long term planning, while a second can better deal with probability calculations, and a third is able to keep a straight face in bluffing games. The perceived complexity and depth of any given game is likely to vary between these players. This also applies to artificial agents depending on the AI techniques and approaches being employed. This makes it impossible to say that one game is more complex or deep than another, without taking into account the human or agent that is playing it.

While it is not yet clear how to accurately estimate the complexity or depth of games, doing so could have several benefits for game analysis and development. One application could be for identifying flaws or limitations in games. The original rules for several traditional board games, such as the ancient Viking game of Hnefatafl or the Maori game of Mu Torere, were incorrectly recorded, leading to unfairly balanced games [2]. Methods for analysing the depth of these games would allow such weaknesses to be detected and corrected. Such a case was demonstrated for the 1982 video game Q*bert, where a previously unknown glitch was discovered by a reinforcement learning agent [3]. Agents can also identify additional strategies or levels of depth not previously considered by humans, such as with DeepBlue and AlphaGo[4].

One idea for future work could be to select a suitable set of benchmark games and test how complex or deep each game is for a collection of agents and a variety of possible measures. Identifying any similarities between resource and performance curves across different game features would allow us to be more confident of which features most impact the complexity or depth of a game, particularly if several different empirical measures broadly align. It might also be worthwhile investigating or developing games that humans find easy to play but agents currently perform poorly on, as these likely represent limitations with current AI techniques.

References

1. F. Lantz, A. Isaksen, A. Jaffe, A. Nealen, J. Togelius, "Depth in Strategic Games", AAAI Workshops, North America, 2017, pp. 1-8.
2. C. Browne, "Modern Techniques for Ancient Games", IEEE Conference on Computational Intelligence and Games, Maastricht, 2018, pp. 1-8.
3. P. Chrabaszcz, I. Loshchilov, F. Hutter, "Back to Basics: Benchmarking Canonical Evolution Strategies for Playing Atari", IJCAI, Stockholm, 2018, pp. 1419-1426.
4. Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M. and Dieleman, S. (2016), "Mastering the game of Go with deep neural networks and tree search." Nature, 529(7587), 484.