International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET

A Review on Hanabi Game for Multiagent Learning using Artificial Intelligence

Shweta Pramodrao Sontakke[1], Dr. A. N. Thakare[2]
[1]PG Scholar, Computer Science and Engineering, Department of Computer Engineering, Bapurao Deshmukh College of Engineering, Sevagram, Wardha, Maharashtra, India
[2]Assistant Professor, Department of Computer Engineering, Bapurao Deshmukh College of Engineering, Sevagram, Wardha, Maharashtra, India

Article Info
Volume 8 Issue 2
Page Number: 127-133

Publication Issue :
March-April-2021

Article History
Accepted : 10 March 2021
Published : 15 March 2021

ABSTRACT
A popular board game Hanabi is a combination of cooperative gameplay with imperfect information. Partial observability makes the game, a challenging domain for AI research. Especially, when AI should cooperate with a human player. Imperfect information game is nontrivial due to complicated interplay of policies. The combination of cooperation, imperfect information, and limited communication make Hanabi an ideal challenge in both self-play and ad-hoc team settings. Ad-hoc team settings, where partners and strategies are not known in advance. In this paper, we are trying to review all such type of games, which is evaluated with the help of Artificial Intelligence and machine technique. We expect this article will help unify and motivate future research to take advantage of the abundant literature that exists to promote fruitful research in the multiagent community.

Keywords : Ad-Hoc Team, Communication, Cooperative, Imperfect Information

I. INTRODUCTION

Artificial Intelligence is the science of getting machines to thing and make decision like human beings do. Since the development of complex Artificial Intelligence algorithms, it has been able to accomplish this by creating machines and robots that are applied in a wide range of fields including agriculture, healthcare, robotics, marketing, business analytics, games and many more.

Artificial Intelligence has mastered some of the world's most complex games. Hanabi is different from the adversarial two-player zero-sum game where computers have reached superhuman skills, beating top player at chess[29,32], checkers[31], go[28], backgammon[27], two-player poker[25,26] and even starcraft-2[30], a real time strategy computer game. Artificial Intelligence stumbles in cracking some of the seemingly simple game once that required an ability to communicate and collaborate. Hanabi as a new challenge domain with novel problems that arise from its combination of purely cooperative gameplay.

A. Hanabi: the game

Hanabi is a cooperative card game created by French game designer Antoine Bauza (2010). In which players are aware of other's cards but not their own. To succeed, players must coordinate to efficiently reveal information to their teammates, however players can only communicate through grounded hint actions that point out all of a player's cards of a chosen rank or colour. So, everyone needs the advice of his fellow players and attempt to play a series of cards in a specific order to set off a simulated firework. Hanabi is a challenging benchmark problem for Artificial Intelligence. Hanabi an interesting multi-agent learning challenges [1] for both learning a good self-play policy and adapting to an ad-hoc team of players.

Source: James Goodman, 2019 IEEE

Figure 1 : A game of Hanabi in progress. The player at the camera's perspective can see the other player's cards, but not their own. The current score in the game is 12, from the sum of the top cards in each suit in the tableau.

B. Hanabi: the challenge

a. Challenge one: self-play learning

Self-play challenge is focused on finding a joint policy that achieves a high expected score entirely through self-play learning. Hanabi is difficult for current learning algorithm even when large amount of data and computation time, learning agents have trouble approaching the performance of hand-crafted rules in four player games, and fall far short of such rules for three and five players.

b. Challenge two: ad-hoc team play

In ad-hoc team play, the ultimate goal is agents that are capable of playing with other agents or human players. For this, a policy which achieves a high score in self-play is of little use if it must be followed exactly by teammates. good strategies are not unique, and a robust player must learn to recognise intent in other agent's actions and adapt to a wide range of possible strategies. We propose to evaluate ad-hoc team performance by measuring an agent's ability to play with a wide range of teammates it has never encountered before.

In the future we expect to see canonical agents pulls of pre-trend or hard-coded self-play agents be made available for training and hold-out sets to allow for consistent comparisons. To facilitate future research, the open-source Hanabi Learning Environment, propose an experimental framework for the research community to evaluate algorithmic advances, and assess the performance of current state-of-the-art techniques.

II. BACKGROUND AND RELATED WORK

Hanabi is a cooperative card game created by Antoine Bauza, a French game designer in 2010. Later in the same year Peter Stone et al. [22], work on ad-hoc human teams and ad-hoc autonomous agent teams, also give an example of human soccer and robot soccer with the ad-hoc team player. Game theory Leyton-Brown and Shoham provides a useful theoretical foundation for multiagent interaction. A good ad-hoc team player may need to make an explicit assumption that its teammates are observing and reacting to its actions.

Hanabi won the prestigious Spiel des Jarhes, Game of the year award in 2013. In 2015 Osawa [18], describes experiment with two players only. It showed that simulated strategies (ideal, random, internal state, outer-state, self-recognition) that try to recognize the intention of other players performed better than a fixed set of static strategies. Later Cox et al. [20], developed the hat strategy. The result of simulating the recommendation and information strategies and also simulate a cheating strategy for the comparison purpose.

Bruno Bouzy [8], developed a set of players name Hannibal, in which each player being either a knowledge-based simulator or a tree search player using a simulator. Using Hat principle, they reported achieving 24.92 points and a perfect score 92%, in that an information move informs all the players at once, not only the targeted player. joseph et al. [5], implement a number of rule-based agents. In addition to an Information Set Monte Carlo Tree Search (IS-MCTS) agent. M. Eger et al. [6], presented AI agent for the two-player version of the cooperative card game Hanabi that is based on intentionality and communication theory. C. Martens and M. Eger [12], presented the cooperative card game, which allows users to play the game with a variety of Ais in a web browser. Additionally, GUI has the capability to watch replays of previously completed games, and to take over control of these replays at any point, these are some techniques to achieve strong Hanabi play.

Rodrigo Canaan et al. [4], describes a two-track competition of agents for the game will take place in the 2018 CIG conference. In this author develop a genetic algorithm that builds rule-based agents by determining the best sequence of rules from a fixed rule set to use as strategy, it uses evolution in three steps to get better playing agents than the human-created baselines. Eva T. Gottwald et al. [11], implemented the two-player version of Hanabi in Unity with support for a Tobii eye tracker. There were two tracks, called “Mirror” and “Mixed”. Using eye tracker, they are able to determine where a player’s gaze lingers with reasonable precision to determine which card they are focusing on.

Another work on ad-hoc teamplay using Hanabi is by N. Bard et al. [1], who independently trained reinforcement learning agents that scored 20 to 22 points in self-play, but only 0 to 5 when paired with one another. They also proposed an ad-hoc setting where self-play playtraces of the partner agent are provided prior to gameplay for learning, but no agent currently takes advantage of this. Pablo S. Chacon et al. [7], describes proposed approach for AI agents that can play the game(pandemic) with human players without requiring explicit communication. Andy Nealen et al. [3], showed that, using MAP-Elites, it is possible to generate a pool of Hanabi-playing agents that differs in two important behavioural dimensions: risk aversion and communicativeness. Jungkyu Park [15], results with default max relay buffer, in which they stop the training of all models at iteration 10000 and compare the best evaluation performance. Results with small max replay buffer, in which the best model not only outperforms model with bigger replay buffer but also deep mind reported performance of Rainbow agent on two-player Hanabi. James Goodman [16], used ISMCTS as a base algorithm coupled with a re-deterministation technique that re-samples consistent world states for everyone but the player acting at a node. This entry also used a neural network to represent a policy and value function trained via self-play. Jacob N. Foerster et al. [17], presented a BAD, a novel algorithm for multi-agent reinforcement learning in cooperative partially observable settings. It uses a factorized, approximate belief state that allows agents to efficiently learn informative actions, leading to the discovery of conventions. For more on multi-agent deep reinforcement learning, P. Hernandez-Leal et al. [19], provides a recent survey.

Recently, M. Eger et al. [2], performed two online experiments, one in which Eager demonstrate that the intentional behaviour leads to an increase in score over a baseline agent, and another in which we demonstrate how taking timing into account can improve the agent's performance in games with unfavourable starting condition.

Now a day's deep reinforcement learning for multiagent system and its methods used by Thanh T. Nguyen et al. [9], Yuandong T. et al. [10], proposed a JPS, a general optimization technique to jointly optimize policy for multiagent collaborative agents in imperfect information efficiently. A survey of multi-agent strategy based on reinforcement learning by Liang Chen et al. [13], and introduces the basic methods of RL, the main method of single agent RL and multiagent RL. Rodrigo Canaan et al. [14] and Xianbo Gao et al. [21], evaluating RL agents in Hanabi with unseen partners and trained agents using the popular Rainbow DQN architecture in Hanabi using self-play, a single rule-based partner, and a mix of rule-based partners.

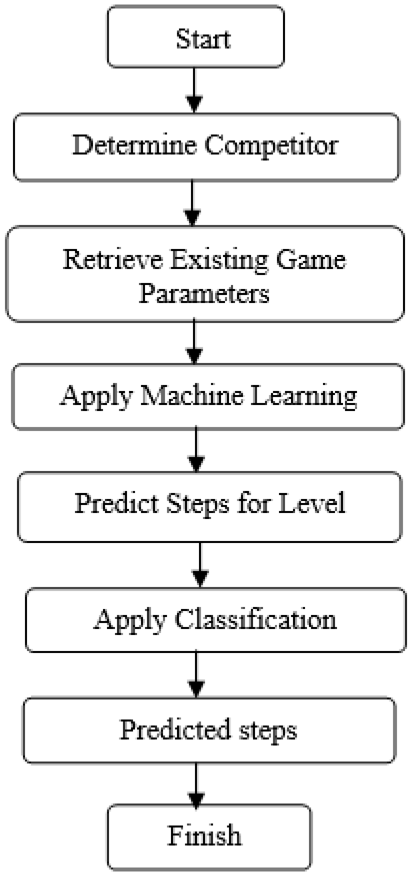

Fig. 2. Flowchart: steps involved in multiagent learning Hanabi game

III. PROPOSED METHODOLOGY

Ad-hoc team play is learning to play with a set of unknown partners, with only a few games of interaction. Ad-hoc team play's ultimate goal that is capable of playing with other agents or even human players. For this we propose an algorithm generation with the help of self-play is little bit in use. A robust player must learn to recognize intent in other agent's behaviour and adapt a wide range of possible strategies being played. Good strategies are developed by repeating the players by playing 1000 different random sets. The important aspect of ad-hoc team is to be recognizing or modelling the capabilities of one's teammates.

We propose to evaluate ad-hoc team performance by measuring an agent's ability to play with a wide range of teammates it has never done before. This performance is measured via score achieved by the agent when it paired with autonomous agents and then players exhibit a diverse strategy's which can be hard-coded or learned by self-playing.

It is possible to create the handcrafted program that plays this game well, as we humans already know good strategies, however this project is about getting several instances of an AI to learn new ways to communicate with each other effectively. Again, the goal is not to get a computer program that plays Hanabi well, the goal is to get an AI to learn to communicate effectively and work together towards a common goal.

IV. CONCLUSION

The combination of cooperative gameplay and imperfect information make Hanabi a compelling research challenge for Artificial Intelligence and machine learning techniques in multi-agent settings. This article briefly introduces about the multi-agent learning in different learning techniques, techniques to achieve a strong Hanabi play, best policies, and strategies. In ad-hoc team settings, where the agents must play with unknown teammates will help us to understand better the role theory of mind reasoning might play for AI systems that learn to collaborate with other agents and humans. Machine learning techniques used to simplify, which is based on AI technique, so is conclude that we try to find out feasible solution based on recent study. And the game becomes more playable as compare to existing. In future work, we are trying to design an algorithm for predicting multiagent learning method and check the performance analysis of design algorithm.

V. ACKNOWLEDGMENT

We would like to thank the authors, many people in board games designing and Antoine Bauza, who designed Hanabi, a cooperative card game. A special thank for those who writing clear, readable code for the Hanabi research environment used in our experiments. Dr. A. N. Thakare for help with coordinating across different time zones, and discussion with cooperative games.

VI. REFERENCES

[1]. N. Bard, J. N. Foerster, S. Chandar, N. Burch, M. Lanctot, H. F. Song, E. Parisotto, V. Dumoulin, S. Moitra, E. Hughes, I. Dunning, S. Mourad, H. Larochelle, M. G. Bellemare, M. Bowling, “The Hanabi challenge: A new frontier for AI research,” CoRR, vol. abs/1902.00506, 2019. [Online]. Available: http://arxiv.org/abs/1902.00506

[2]. M. Eager, C. Martens, P. Sauma Chacon, M. Alfaro Cordoba, J. Hidalgo-Cespedes, “Operationalizing Intentionality to Play Hanabi with Human Players,” DOI 10.1109/TG.2020.3009359, IEEE

[3]. R. Canaan, J. Togelius, A. Nealen, S. Menzel, “Diverse Agents for Ad-Hoc Cooperation in Hanabi,” IEEE 2019.

[4]. R. Canaan, J. Togelius, H. Shen, A. Nealen, R. Torrado, S. Menzel, “Evolving Agents for the Hanabi 2018 CIG Competition,” ArXiv e-prints, Sep. 2018.

[5]. J. Walton-Rivers, P. R. Williams, R. Bartle, D. Perez-Liebana, S. M. Lucas, “Evaluating and modelling hanabi-playing agents”, in Evolutionary Computation (CEC), 2017 IEEE Congress on. IEEE,2017, pp. 1382-1389.

[6]. R. Patil Rashmi, Y. Gandhi, V. Sarmalkar, P. Pund and V. Khetani, "RDPC: Secure Cloud Storage with Deduplication Technique," 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 2020, pp. 1280-1283, doi: 10.1109/I-SMAC49090.2020.9243442.

[7]. M. Eger, C. Martens, and M. Alfaro Cordoba, “An intentional AI for Hanabi,” in 2017 IEEE Conference on Computational Intelligence and Games (CIG), 2017, pp. 68-75.

[8]. P. Sauma Chacon and M. Eger, “Pandemic as a challenge for human-AI cooperation,” in Proceedings of the AIIDE workshop on Experimental AI in Games, 2019.

[9]. B. Bouzy “Playing Hanabi near-optimally,” in 15th International Conference on Advances in Computer Game (ACG15), ICGA. Cham: Springer International Publishing, 2017, pp. 51-62.

[10]. T. Thi Nguyen, N. Duy Nguyen, and Saeid Nahavandi, Senior Member, IEEE, “Deep Reinforcement Learning for Multiagent

Systems: A Review of Challenges, Solutions, and Applications," IEEE, April 28, 2020.

[11]. Y. Tian, Q. Gong, and T. Jiang, "Joint Policy Search for Multi-agent Collaboration with Imperfect Information," 34th Conference on Neural Information Processing Systems (NeuroIPS), 2020.

[12]. E. T. Gottwald, M. Eger and C. Martens, "I see what you see: Integrating eye tracking into Hanabi playing agents," CEUR-WS.org/Vol-2282/EXAG_112

[13]. M. Eger and C. Martens, "A Browser-based interface for the exploration and evaluation of Hanabi AIs," https://github.com/yawgmoth/pyhanabi, 2017.

[14]. L. Cheng, T. Guo, Yun-ting Liu and Jia-Ming Yang, "Survey of Multi-Agent Strategy Based on Reinforcement Learning," IEEE, 2020.

[15]. R. Canaan, X. Gao, Y. Chung, J. Togelius, A. Nealen, and S. Menzel, "Evaluating RL Agents in Hanabi with Unseen Partners," AAAI Workshop on Reinforcement Learning in Games ,2020.

[16]. J. (JP) Park "Advancing AI: Hanabi Challenge," Inference and Representation, 2019

[17]. J. Goodman, "Re-determinizing MCTS in Hanabi," arXiv preprint arXiv: 1902.06075, 2019.

[18]. J. N. Foerster, H. F. Song, E. Hughes, N. Burch, I. Dunning, S. Whiteson, M. Botvinick, and M. Bowling, "Bayesian action decoder for deep multi-agent reinforcement learning," arXiv preprint arXiv:1811.01458, 2018.

[19]. H. Osawa, "Solving Hanabi: Estimating hands by opponent's actions in cooperative Game with incomplete information," in Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence, 2015. [Online]. Available: https://aaai.org/ocs/index.php/WS/AAAIW15/paper/view/10167

[20]. P. Hernandez-Leal, B. Kartal, and M. E. Taylor, "A survey and critique of multiagent deep reinforcement learning," Springer,16 October 2019.

[21]. C. Cox, J. De Silva, P. Deorsey, F. H. Kenter, T. Retter, and J. Tobin, "How to make the perfect firework display: Two strategies for Hanabi," Mathematics Magazine, vol. 88, no. 5, pp. 323-336, 2015.

[22]. R. Canaan, X. Gao, Y. Chung, J. Togelius, A. Nealen, and S. Menzel, "Behavioural Evaluation of Hanabi Rainbow DQN Agents and Rule-Based Agents", 16th AAAI Conference on AIIDE, 2020

[23]. P. Stone, G. A. Kaminka, S. Kraus, and J. S. Rosenschein, "Ad-Hoc Autonomous Agent Teams: Collaboration without Pre-Coordination", American Association for Artificial Intelligence, 2010.

[24]. A. Bauza, "Hanabi" https://boardgamegeek.com/boardgame/98778/hanabi,2010.

[25]. G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, W. Zaremba, "Open AI Gym", arXiv:1606.01540v1 [cs. LG] 5 Jun 2016.

[26]. N. Brown and T. Sandholm, Science, "Superhuman AI for heads-up no-limit poker: Libratus beats top professionals", 10.1126/science. aao1733 (2017).

[27]. M. Moravcik, M. Schmid, N. Burch, V. Lisy, D. Morrill, N. Bard, T. Davis, K. Waugh, M. Johanson, M. Bowling, "DeepStack: Expert-Level Artificial Intelligence in heads-up no-limit Poker", arXiv: 1701.01724v3 [cs. AI] 3Mar 2017.

[28]. G. Tesauro "Temporal difference learning and TD-Gammon", ACM, March 1995/ Vol. 38, No.3.

[29]. D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglu, V. Panneershelvam, M. Lanctot, S.

Dielema, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel and D. Hassabis, "Mastering the game of GO with deep neural networks and tree search", doi:10.1038/nature16961 vol 529 Jan 2016.

[30]. N. Ensmenger "Is chess the drosophila of artificial intelligence A social history of an algorithm", DOI:10.1177/0306312711424596 sss.sagepub.com 2011.

[31]. A. A. Sanchez-Ruiz, M. Miranda, "A machine learning approach to predict the winner in StarCraft based on influence maps", http://dx.doi.org/10.1016/j.entcom.2016.11.005

[32]. J. Schaeffer, R. Lake, P. Lu, and M. Bryant, "Chinook: The world man-machine checkers champion", AI magazine volume 17 Nov 1 (1996) AAAI.

[33]. M. Campbell, A. J. Hoane Jr., Feng-hsiung Hsu, "Deep Blue", PII: S0004-3702(01)00129-1 2001 by Elsevier.

[34]. A. Iraci, "Convensions for Hanabi", http://hanabi.pythonanywhere.com/static/Hanabi.pdf, 2018.

Cite this article as :

Shweta Pramodrao Sontakke, Dr. A. N. Thakare, " A Review on Hanabi Game for Multiagent Learning using Artificial Intelligence, International Journal of Scientific Research in Science, Engineering and Technology(IJSRSET), Print ISSN : 2395-1990, Online ISSN : 2394-4099, Volume 8, Issue 2, pp.127-133, March-April-2021.

Journal URL : https://ijsrset.com/IJSRSET218245