Strategic Negotiations for Extensive-Form Games*

JAAMAS Track

Dave de Jonge
IIIA-CSIC
Barcelona, Catalonia, Spain
davedejonge@iiia.csic.es

Dongmo Zhang
Western Sydney University
Penrith, NSW, Australia
d.zhang@westernsydney.edu.au

ABSTRACT

When studying extensive-form games it is typically assumed that players make their decisions individually and that it is not possible for them to make formally binding agreements about future moves. As a consequence, many non-zero-sum games have been shown to have equilibria that are suboptimal and arguably counter-intuitive. For this reason we explore a new line of research in which game-playing agents are allowed to negotiate binding agreements. We analyze what happens under such assumptions and define a new equilibrium solution concept to capture this (the *Negotiation Value*). We show that the outcomes predicted by this new solution concept are more efficient than the Subgame Perfect Equilibrium and, therefore, arguably more realistic. Furthermore, we demonstrate experimentally that a bounded rational agent is able to approximate our solution concept in several games and that it strongly outperforms non-negotiating rational players.

KEYWORDS

Automated Negotiations, Non-Zero-Sum Games, Extensive-form Games, General Game Playing, Monte Carlo Tree Search

ACM Reference Format:

Dave de Jonge and Dongmo Zhang. 2020. Strategic Negotiations for Extensive-Form Games. In *Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand, May 9–13, 2020,* IFAAMAS, 3 pages.

1 INTRODUCTION

Research on game-playing algorithms has mostly focused on zero-sum-games, like Chess, Poker, and Go, even though many real-life problems are better modeled as non-zero-sum games. For example, a market economy is essentially a non-zero-sum game. Traditional game-playing algorithms such as Minimax and Monte Carlo Tree Search (MCTS), therefore, do not take into account the ability of an agent to negotiate its actions with its opponent(s), which may result in inefficient outcomes when applied to non-zero-sum games.

Non-zero-sum games have been studied extensively from a formal point of view, but commonly used solution concepts such as the Nash Equilibrium and the Subgame Perfect Equilibrium (SPE) are based on the assumption that players choose their actions individually. Again, one does not take into account that players may be able to negotiate binding agreements about their actions. We argue that for this reason such equilibrium concepts often predict outcomes that are suboptimal compared to the outcomes one could expect to achieve in reality. For example, when two people find themselves in a real-life situation comparable to the Prisoner's Dilemma, the natural thing to do would be to discuss the situation and agree to cooperate. Even if they do not trust each other they could still cooperate, by signing a legally binding contract.

The field of Automated Negotiations, on the other hand, has given little attention to scenarios with the complexity of extensive-form games. Most work in this field assumes the utility of any deal can always be determined exactly and without much computational effort, while for games like Chess and Go it is generally unfeasible to determine the exact value of any action.

Therefore, we aim to bridge the gap between Automated Negotiations and extensive-form games. We aim to answer the following question: *"Given some extensive-form game, we know that without negotiations two perfectly rational players would play the SPE, but what would happen if we did allow those players to negotiate binding agreements about their actions?"* To answer this question we define a new solution concept for extensive-form games (*the Negotiation Value*), which takes into account the possibility that players make binding agreements. We have calculated the Negotiation Values of three traditional games, namely the Iterated Prisoner's Dilemma [1], the Centipede Game [21], and the Dollar Auction [24], and show that they dominate the outcomes prescribed by the traditional SPE.

Furthermore, we have defined another new solution concept to predict the outcome of a bilateral negotiation, which is an alternative to the Nash Bargaining Solution [20]. The difference is that our solution works for discrete agreement spaces, whereas Nash's solution assumes convex agreement spaces.

2 RELATED WORK

The earliest work on Automated Negotiations mainly focused on proving formal properties of idealized scenarios [20]. Later, focus shifted more towards heuristic approaches [4, 5]. However, such heuristic approaches still often assumed there is only a small set of possible agreements, that the utility functions are linear and that they can be calculated without much computational cost [2].

Recently, more attention has been given to large domains with non-linear utility functions [9, 18, 19], but these works still assumed the value of any given contract can be calculated easily. Therefore, [11] studied domains with utility functions that are computationally hard to calculate. An even more complex negotiation scenario is the game of Diplomacy [3, 6, 10, 12, 22]. This is an extensive-form game that involves negotiations before each round. These negotiations are complex, because the players' utility functions are not directly defined in terms of the agreements they make, but more indirectly

*This is an extended abstract of our full paper [15] to which we refer for more details.

Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), B. An, N. Yorke-Smith, A. El Fallah Seghrouchni, G. Sukthankar (eds.), May 9–13, 2020, Auckland, New Zealand.

through the moves they make in the game, which in turn are subject to the negotiated agreements.

The field of General Game Playing (GGP) [8] studies algorithms for game-playing agents that only know the rules of the game at run-time. Therefore, when developing a GGP agent, one cannot use any game-specific heuristics. Common techniques applied in GGP are minimax [25], alpha-beta pruning [16] and MCTS [7, 17, 23]. The idea of combining GGP with Automated Negotiations was brought forward by us in [13] and [14].

3 ASSUMPTIONS

We consider a setting in which two agents play a non-zero-sum extensive-form game that takes place over multiple rounds. Before each round the players have the opportunity to negotiate with each other and make binding agreements about which strategies they will follow throughout the rest of the game. Our goal is to implement an algorithm for one of these two agents that aims to maximize that agent's utility by choosing the right moves, and by negotiating the right deals with its opponent. Furthermore, we make the following assumptions: 1) Each agent is purely **self-interested**, so they are not interested in any form of 'social welfare'. 2) **The opponent is unknown**, so our agent cannot make any assumptions about the opponent, except that it is (bounded) rational and aims to maximize its own utility. 3) **Agreements are always obeyed**, because there is some mechanism in place that can enforce them, even if it would be rational for the agents not to obey them. This makes the analysis of our games-with-negotiations different from the classical analysis of extensive-form games. 4) A **bounded rational** agent should be able to approximate the theoretically optimal outcomes. 5) The agents are **general game players**, so the rules of the games they are playing are only known to them at run-time. 6) We only consider games of **full information** and with no random events. This means that agents have a complete description of each others' utility functions. These descriptions, however, may be very complex, so in general it may be too hard to calculate utilities exactly.

4 GAMES WITH NEGOTIATIONS

Let G be an extensive-form game for two players denoted α_0 and α_1 respectively. For any state w of G and any joint strategy $\vec{\sigma}$ for G we define a Negotiation Domain $N_{w,\vec{\sigma}}$ for which the agreement space is the set of all joint strategies of G. This negotiation domain represents the situation that the game has reached state w, while the players have previously agreed to obey the joint strategy $\vec{\sigma}$, and they are currently negotiating whether they can agree on some better joint strategy $\vec{\sigma}'$. If negotiations in $N_{w,\vec{\sigma}}$ fail, the players remain committed to $\vec{\sigma}$. We use $\vec{\tau}$ to denote the 'empty commitment', so $N_{w,\vec{\tau}}$ represents the situation where the agents have not yet made any agreements at all.

For any given extensive-form game G we define the corresponding **Extensive-Form Game with Negotiations** NG to consist of G together with a negotiation domain $N_{w,\vec{\sigma}}$ for every state w of G and every joint strategy $\vec{\sigma}$ of G. We use $rv_{w,\vec{\sigma},i}$ to denote the reservation value of agent α_i in negotiation domain $N_{w,\vec{\sigma}}$, and $spe_{w,\vec{\sigma},i}$ to denote the utility that α_i obtains under the SPE of the subgame starting at state w, while the players have committed themselves to joint strategy $\vec{\sigma}$.

Theorem 4.1. *Let G be an extensive-form game. Then for any state w, any joint strategy $\vec{\sigma}$, and any player α_i we have: $rv_{w,\vec{\sigma},i} \geq spe_{w,\vec{\sigma},i}$.*

Intuitively, this result states that if negotiations fail the agents should each still be able to obtain at least the utility associated with the SPE, which should be obvious. However, what is more interesting, is that in some cases this inequality is strict. This is because even if negotiations fail in the current round, the agents may still come to an agreement in the following rounds, so they can still expect to achieve utility values higher than the SPE.

For any negotiation domain N we define $nv_i(N)$ to be the expected utility that player α_i can expect to obtain from the negotiation. This value is determined by modeling the negotiations as a normal-form game and then calculating its Nash-equilibrium.

Given an extensive-form game G and a player α_i we define its **Negotiation Value** as $NV_i(G) := nv_i(N_{w_0,\vec{\tau}})$, where w_0 is the initial state of G. Similarly, we define $SPE_i(G) := spe_{w_0,\vec{\tau},i}$.

Theorem 4.2. *For any extensive-form game G we have $NV_i(G) \geq SPE_i(G)$, and for some games this inequality is strict.*

In our full paper [15] we show that for the Centipede Game and the Iterated Prisoner's Dilemma this inequality is strict.

5 THE CENTIPEDE GAME

We have analyzed the Centipede Game (CG) under the assumption that players are able to negotiate binding agreements. We formally show that, unlike in the traditional CG, the players no longer have the incentive to break the game off at an early stage. Furthermore, we make two interesting observations.

(1) *Rational players may continue playing **without making any agreements** until they are close to the end of the game.*
(2) *Although the game has two terminal states that are Pareto-optimal, and which both dominate the SPE, only one of them is a rational outcome for negotiating players.*

The first fact is surprising because we know that without the possibility to make binding agreements, rational players would always break the game off as quickly as possible. However, it turns out that when they do have the possibility to negotiate, it becomes perfectly rational to continue playing for several rounds *without actually making such agreements.* It is the mere *possibility* to make such agreements *later on* that makes this rational.

The second fact is surprising, because in classical automated negotiations any outcome that is pareto-optimal and dominates the conflict outcome is considered feasible. Here, however, we see that these conditions are not sufficient, because the structure of the game imposes extra constraints on the feasibility.

6 EXPERIMENTS

We have implemented an algorithm that we call Monte Carlo Negotiation Search (MCNS), which builds a negotiation algorithm on top of an MCTS-based GGP player. We have tested it on the three aforementioned traditional games. In the case of the Centipede Game and the Dollar Auction it always agrees with its opponent to play the theoretically optimal joint strategy (the one that yields the Negotiation Values), and in the case of the Iterated Prisoner's Dilemma it comes to an agreement that closely approximates it.

REFERENCES

[1] R Axelrod and WD Hamilton. 1981. The evolution of cooperation. *Science* 211, 4489 (1981), 1390–1396. https://doi.org/10.1126/science.7466396 arXiv:http://science.sciencemag.org/content/211/4489/1390.full.pdf

[2] Tim Baarslag, Koen Hindriks, Catholijn M. Jonker, Sarit Kraus, and Raz Lin. 2010. The First Automated Negotiating Agents Competition (ANAC 2010). In *New Trends in Agent-based Complex Automated Negotiations, Series of Studies in Computational Intelligence*, Takayuki Ito, Minjie Zhang, Valentin Robu, Shaheen Fatima, and Tokuro Matsuo (Eds.). Springer-Verlag.

[3] Angela Fabregues and Carles Sierra. 2011. DipGame: a challenging negotiation testbed. *Engineering Applications of Artificial Intelligence* (2011).

[4] Peyman Faratin, Carles Sierra, and Nicholas R. Jennings. 1998. Negotiation decision functions for autonomous agents. *Robotics and Autonomous Systems* 24, 3-4 (1998), 159 – 182. https://doi.org/10.1016/S0921-8890(98)00029-3 Multi-Agent Rationality.

[5] Peyman Faratin, Carles Sierra, and Nicholas R. Jennings. 2000. Using Similarity criteria to make negotiation trade-offs. In *International Conference on Multi-Agent Systems, ICMAS'00*. 119–126.

[6] André Ferreira, Henrique Lopes Cardoso, and Luís Paulo Reis. 2015. DipBlue: A Diplomacy Agent with Strategic and Trust Reasoning. In *7th International Conference on Agents and Artificial Intelligence (ICAART 2015)*. 398–405.

[7] Hilmar Finnsson. 2012. *Simulation-Based General Game Playing*. Ph.D. Dissertation. School of Computer Science, Reykjavik University.

[8] M. Genesereth, N. Love, and B. Pell. 2005. General game playing: Overview of the AAAI competition. *AI Magazine* 26(2) (2005), 62–72.

[9] Takayuki Ito, Mark Klein, and Hiromitsu Hattori. 2008. A multi-issue negotiation protocol among agents with nonlinear utility functions. *Multiagent Grid Syst.* 4 (January 2008), 67–83. Issue 1. http://dl.acm.org/citation.cfm?id=1378675.1378678

[10] Dave de Jonge, Tim Baarslag, Reyhan Aydoğan, Catholijn Jonker, Katsuhide Fujita, and Takayuki Ito. 2019. The Challenge of Negotiation in the Game of Diplomacy. In *Agreement Technologies 2018, Revised Selected Papers*, Marin Lujak (Ed.). Springer International Publishing, Cham, 100–114.

[11] Dave de Jonge and Carles Sierra. 2015. NB3: a multilateral negotiation algorithm for large, non-linear agreement spaces with limited time. *Autonomous Agents and Multi-Agent Systems* 29, 5 (2015), 896–942. https://doi.org/10.1007/s10458-014-9271-3

[12] Dave de Jonge and Carles Sierra. 2017. D-Brane: a diplomacy playing agent for automated negotiations research. *Applied Intelligence* (2017), 1–20. https://doi.org/10.1007/s10489-017-0919-y

[13] Dave de Jonge and Dongmo Zhang. 2016. Using GDL to Represent Domain Knowledge for Automated Negotiations. In *Autonomous Agents and Multiagent Systems: AAMAS 2016 Workshops, Visionary Papers, Singapore, Singapore, May 9-10, 2016, Revised Selected Papers*, Nardine Osman and Carles Sierra (Eds.). Springer International Publishing, Cham, 134–153. https://doi.org/10.1007/978-3-319-46840-2_9

[14] Dave de Jonge and Dongmo Zhang. 2017. Automated Negotiations for General Game Playing. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems, AAMAS 2017, São Paulo, Brazil, May 8-12, 2017*, Kate Larson, Michael Winikoff, Sanmay Das, and Edmund Durfee (Eds.). ACM, 371–379.

[15] Dave de Jonge and Dongmo Zhang. 2020. Strategic negotiations for extensive-form games. *Autonomous Agents and Multi-Agent Systems* 34, 1 (Apr 2020). https://doi.org/10.1007/s10458-019-09424-y

[16] Donald E. Knuth and Ronald W. Moore. 1975. An analysis of alpha-beta pruning. *Artificial Intelligence* 6, 4 (1975), 293 – 326. https://doi.org/10.1016/0004-3702(75)90019-3

[17] Levente Kocsis and Csaba Szepesvári. 2006. Bandit Based Monte-carlo Planning. In *Proceedings of the 17th European Conference on Machine Learning* (Berlin, Germany) *(ECML'06)*. Springer-Verlag, Berlin, Heidelberg, 282–293. https://doi.org/10.1007/11871842_29

[18] Ivan Marsa-Maestre, Miguel A. Lopez-Carmona, Juan R. Velasco, and Enrique de la Hoz. 2009. Effective bidding and deal identification for negotiations in highly nonlinear scenarios. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems - Volume 2* (Budapest, Hungary) *(AAMAS '09)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 1057–1064. http://dl.acm.org/citation.cfm?id=1558109.1558160

[19] Ivan Marsa-Maestre, Miguel A. Lopez-Carmona, Juan R. Velasco, Takayuki Ito, Mark Klein, and Katsuhide Fujita. 2009. Balancing Utility and Deal Probability for Auction-based Negotiations in Highly Nonlinear Utility Spaces. In *Proceedings of the 21st International Jont Conference on Artifical Intelligence* (Pasadena, California, USA) *(IJCAI'09)*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 214–219. http://dl.acm.org/citation.cfm?id=1661445.1661480

[20] J.F. Nash. 1950. The Bargaining Problem. *"Econometrica"* "18" (1950), 155–162.

[21] Robert W Rosenthal. 1981. Games of perfect information, predatory pricing and the chain-store paradox. *Journal of Economic Theory* 25, 1 (1981), 92 – 100. https://doi.org/10.1016/0022-0531(81)90018-1

[22] E. Ephrati S. Kraus, D. Lehman. 1989. An automated diplomacy player. In *Heuristic Programming in Artificial Intelligence: The 1st Computer Olympia*, D. Levy and D. Beal (Eds.). Ellis Horwood Limited, 134–153.

[23] Stephan Schiffel and Michael Thielscher. 2007. M.: Fluxplayer: A successful general game player. In *In: Proceedings of the AAAI National Conference on Artificial Intelligence*. AAAI Press, 1191–1196.

[24] Martin Shubik. 1971. The Dollar Auction Game: A Paradox in Noncooperative Behavior and Escalation. *The Journal of Conflict Resolution* 15, 1 (1971), 109–111.

[25] John von Neumann. 1959. On the Theory of Games of Strategy. In *Contributions to the Theory of Games*, A.W. Tucker and R.D. Luce (Eds.). Princeton University Press, 13–42.