Boolean Games: Inferring Agents' Goals Using Taxation Queries

Extended Abstract

Abhijin Adiga
Biocomplexity Institute and Initiative
University of Virginia
Charlottesville, VA, USA
abhijin@virginia.edu

Sarit Kraus
Department of Computer Science
Bar-Ilan University
Ramat Gan, Israel
sarit@cs.biu.ac.il

S. S. Ravi
Biocomplexity Institute and Initiative
University of Virginia and
University at Albany – SUNY
Charlottesville, VA, USA
ssravi0@gmail.com

ABSTRACT

In Boolean games, each agent controls a set of Boolean variables and has a goal represented by a propositional formula. We initiate a study of inference in Boolean games assuming the presence of a Principal who has the ability to control the agents and impose taxation schemes. Previous work used taxation schemes to guide a game towards certain equilibria. We show how taxation schemes can also be used to infer agents' goals. In our formulation, agents' goals are assumed to be unknown and the objective of the Principal is to infer the goals of all the agents using appropriate taxation queries. Using an undirected graph (called the goal overlap graph) associated with a Boolean game, we establish necessary and sufficient conditions for the existence of a Nash equilibrium for any taxation query. Using these conditions, we develop an algorithm that uses taxation queries to learn agents' goals. Using a valid node coloring of the goal overlap graph, we show that goals of many agents can be inferred simultaneously. We also present more efficient (in terms of number of queries) goal inference algorithms for two special classes of Boolean functions, namely threshold and symmetric functions.

KEYWORDS

Boolean games; Goal inference; Taxation scheme; Node coloring

ACM Reference Format:
Abhijin Adiga, Sarit Kraus, and S. S. Ravi. 2020. Boolean Games: Inferring Agents' Goals Using Taxation Queries. 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

Boolean games [11] are a class of strategic games where each agent's goal is represented by a propositional logic formula. Each agent i controls a distinct set of Boolean variables Φ_i, and there is a cost associated with each assignment of values to variables. Agent i's formula or goal γ_i is composed of a set Γ_i of variables that are not necessarily in its control. Each agent's first priority is to achieve its goal and its second priority is to minimize its total cost. The general appeal of such games is that they are well-structured and expressive, and the well-developed theory of propositional logic makes them amenable to computational and structural analysis. Much of the work on Boolean games is theoretical in nature [4, 6, 10, 11, 17, 22, 23]. However, Boolean games have recently been used to model problems such as charging strategies for electric vehicles and traffic signaling [17].

In many applications of multi-agent systems, agent functions or goals are unknown and need to be inferred. Recently, a number of works have emerged where a user either actively queries the system or relies on passive observations to infer the functions [2, 3, 12, 14, 20]. Wooldridge et al. [23] introduced the notion of a Principal – an external agent– who can influence the agents' decisions through *taxation schemes* (additional costs for assigning values to variables). Their objective was to achieve a desirable equilibrium by choosing an appropriate taxation scheme. Under this framework, we study the following inference problem on Boolean games. The Principal has knowledge regarding every variable, the agent that controls it and the goals in which it appears but does not know the goal of any agent. The objective of the Principal is to infer these goals by repeatedly "querying" the system (through taxation schemes) and observing the outcomes.

Here, we consider inference of goal functions in *cost-free* Boolean games; thus, the only cost incurred by an agent for assigning values to its variables is due to the taxation scheme. We also restrict our attention to pure Nash equilibria. In such a scenario, it is known that some taxation schemes have no Nash equilibria while others have one or more equilibria [17]. Such scenarios might make it impossible to infer some or all goals. To overcome this problem, we provide additional control to the Principal. Each agent i is associated with an inhibitor variable ψ_i, such that its overall goal is $\gamma_i' = \gamma_i \wedge \psi_i$, where γ_i is the actual goal of the agent. If ψ_i set to zero, the agent can never achieve its goal. This notion of the Principal inhibiting certain agents has been used in applications of Boolean games. For example, Levit et al. [16] discuss a Boolean game model for charging of electric vehicles where some vehicles are not allowed to charge at certain time intervals to avoid overloading. This is similar to our notion of the Principal inhibiting certain agents. A *taxation query* or simply a query consists of (i) assignments of values to ψ_i and (ii) a taxation scheme. While this framework seems to provide an unrealistic amount of control to the Principal, we show scenarios where without such control inference might not be possible. Our approach is to strategically inhibit some agents and infer the goals of remaining agents by simultaneously querying them through taxation schemes. To this end, we construct an undirected graph (called *goal overlap graph*) representing dependencies between agents' goals and apply vertex coloring on the graph. Then, for each query, the Principal observes a Nash Equilibrium (NE) to carry out goal inference. In this context, the questions of

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.

primary interest are: (i) does there exist a set of queries so that the Principal can infer all goals? (ii) if yes, can the size of such a query set be estimated?

2 SUMMARY OF RESULTS

1. Necessary and sufficient conditions for the existence of an NE for a {0,1}-taxation query. Using an undirected graph (called the goal overlap graph) that captures the overlaps between the sets of variables used in agents' goals, we establish necessary and sufficient conditions for the existence of an NE for a Boolean game and any {0,1}-taxation query (i.e., taxation queries with only 0 and 1 costs).
2. Ability to evaluate goals of selected agents at an NE. When there is an NE for a {0,1}-taxation query, we show that the value of the goals of selected agents for the zero cost assignment (i.e., the assignment whose cost is zero for a selected agent) can be determined regardless of which NE is reached by the agents. This result in conjunction with the sufficient condition mentioned in Item 1 forms the basis for our goal inference algorithms.
3. Goal inference algorithms. Using the results in Items 1 and 2 above, we show that any algorithm that learns a Boolean function f using **membership queries** (i.e., queries which specify an input α to the function f and the required response is the value of the function $f(\alpha)$) can be used to develop an algorithm that uses taxation queries to learn f when f is the goal of an agent in a Boolean game. Further, using a valid node coloring of the goal overlap graph, we show that the goals of many agents can be inferred simultaneously. We observe that this scheme can be significantly more efficient than inferring the goals one at a time. We also obtain more efficient (in terms of the number of taxation queries) goal inference algorithms for two special classes of goal functions, namely threshold functions and symmetric functions.
4. Hardness of counting the number of equilibria. When the condition mentioned in Item 1 holds, there is at least one NE for any {0,1}-taxation query. However, we show that the problem of counting the number of equilibria is **#P**-hard, even if there is only one agent in the game. This result is established using a reduction from the problem of counting the number of minimum vertex covers of an undirected graph which is known to be **#P**-hard [21]. This result, which shows that the number of equilibria can be large in general, brings up an important challenge faced by goal inference algorithms. Such algorithms must be able to correctly infer all the goals regardless of which of the many equilibria is reached.

Related work. Boolean games were introduced by Harrenstein et al. [11] as a class of two-player games and were later generalized to n players [5]. The structural and computational properties of Nash equilibria in Boolean games are well-studied. Bonzon et al. [6] define a "dependency graph" (similar to our goal overlap graph) which takes advantage of the structure of agents' goals. In our work, since the Principal has no such knowledge, the goal overlap graph is constructed purely based on the variables that appear in agents' goals. A substantial body of work has been devoted to the problem of manipulating a Boolean game to achieve a desired outcome, which is typically a Nash equilibrium (NE) or a variant there of. Wooldridge et al. [23] use taxation schemes to achieve a desirable equilibrium. Levit et al. [17, 18] proposed methods for finding taxation schemes that incentivize the agents to reach a stable state. They provide necessary and sufficient conditions for the existence of an NE. However, unlike our work, their aim is to find an NE when *all the agents are uninhibited.* Grant et al. [10] propose a framework where along with the Boolean variables under an agent's control, there are environmental variables whose values are selectively announced by the Principal to influence agents' beliefs. Boolean games where players have incomplete information about each other's goals have been considered. De Clercq et al. [7] studied games where possibilistic propositional logic was used to capture this uncertainty. Ågotnes et al. [4] propose a different formulation where each agent can observe a subset of the variables (called the visibility set for the agent), making some equilibria verifiable and others not verifiable. To the best of our knowledge, ours is the first work on inference problems in the domain of Boolean games.

Learning in game theory has been a popular topic albeit from a different perspective. How agents learn strategies by repeatedly playing a game, and what type of equilibria result from such strategies is well-studied [9]. Fearnley et al. [8] consider the problem of finding Nash equilibria by submitting strategy profiles as queries and observing the payoffs for the agents.

There is an emerging body of similar work on inferring agent functions in other multi-agent systems contexts. Adiga et al. [3] consider learning thresholds of nodes in a network propagation model where a user (like the Principal) actively queries the system. Kleinberg et al. [14] consider the problem of inferring comparison-based choice functions that capture a broad range of online behaviors. Inferring agent functions from passive observations has been considered under the probably approximately correct (PAC) framework [2, 12, 20]. The problem of learning Boolean functions using membership queries has received much attention in the learning theory literature. For example, Abasi et al. [1] establish upper and lower bounds on the number of membership queries for learning monotone disjunctive normal form (DNF) expressions. Under the PAC model, the problems of learning some conjunctive normal form (CNF) and DNF functions have also been studied [13]. Some early work on learning considered inferring finite automata from passive observations [19] and learning monomial functions (that model biological phenomena) using experimental data [15].

3 FUTURE WORK

We presented upper bounds on the number of taxation queries used for inference. It is of interest to develop appropriate lower bounds. Our work uses representation-dependent learning; in that setting, if a goal function is from a class $\mathbb{C}$, then the inference algorithm must produce the correct function from that class. It is of interest to extend the results to the representation-independent setting where the learner may produce an equivalent function from another class. Finally, one can also study inference problems for other components of the Boolean game (e.g., control variables).

Acknowledgments. We thank the AAMAS 2020 reviewers for their comments. This work was partially supported by NSF Grants ACI-1443054 (DIBBS), IIS-1633028 (BIG DATA), CMMI-1745207 (EAGER), IIS-1908530 and by the Ministry of Science & Technology, Israel and the Ministry of Education, Science, Research and Sport of the Slovak Republic.

REFERENCES

[1] Hasan Abasi, Nader H. Bshouty, and Hanna Mazzawi. 2014. On Exact Learning Monotone DNF from Membership Queries. *CoRR* abs/1405.0792 (2014), 1–16.

[2] Abhijin Adiga, Chris J Kuhlman, Madhav Marathe, S Ravi, and Anil Vullikanti. 2019. PAC Learnability of Node Functions in Networked Dynamical Systems. In *International Conference on Machine Learning*. PMLR, Nagoya, Japan, 82–91.

[3] Abhijin Adiga, Chris J. Kuhlman, Madhav V. Marathe, S. S. Ravi, Daniel J. Rosenkrantz, and Richard E. Stearns. 2018. Learning the Behavior of a Dynamical System Via a "20 Questions" approach. In *Thirty second AAAI Conference on Artificial Intelligence*. AAAI Press, Palo Alto, CA, 4630–4637.

[4] Thomas Ågotnes, Paul Harrenstein, Wiebe Van Der Hoek, and Michael Wooldridge. 2013. Verifiable equilibria in boolean games. In *Twenty-Third International Joint Conference on Artificial Intelligence*. AAAI Press, Palo Alto, CA, 689–695.

[5] Elise Bonzon, Marie-Christine Lagasquie-Schiex, and Jérôme Lang. 2006. Boolean games revisited. In *Proc. ECAI*. IOS Press, Amsterdam, The Netherlands, 265–269.

[6] Elise Bonzon, Marie-Christine Lagasquie-Schiex, and Jérôme Lang. 2007. Dependencies between players in Boolean games. In *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*. Springer, Hidelberg, Germany, 743–754.

[7] Sofie De Clercq, Steven Schockaert, Ann Nowé, and Martine De Cock. 2015. Multilateral negotiation in Boolean games with incomplete information using generalized possibilistic logic. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*. AAAI Press, Palo Alto, CA, 2890–2896.

[8] John Fearnley, Martin Gairing, Paul W Goldberg, and Rahul Savani. 2015. Learning equilibria of games via payoff queries. *The Journal of Machine Learning Research* 16, 1 (2015), 1305–1344.

[9] Drew Fudenberg and David K Levine. 1998. *The theory of learning in games*. Vol. 2. MIT press, Cambridge, MA.

[10] John Grant, Sarit Kraus, Michael Wooldridge, and Inon Zuckerman. 2014. Manipulating games by sharing information. *Studia Logica* 102, 2 (2014), 267–295.

[11] Paul Harrenstein, Wiebe van der Hoek, John-Jules Meyer, and Cees Witteveen. 2001. Boolean games. In *Proceedings of the 8th conference on Theoretical aspects of rationality and knowledge*. Morgan Kaufmann Publishers Inc., Burlington, MA, 287–298.

[12] Xinran He, Ke Xu, David Kempe, and Yan Liu. 2016. Learning influence functions from incomplete observations. In *Advances in Neural Information Processing Systems*. Neural Information Systems Processing Foundation, San Diego, CA, 2073–2081.

[13] Michael J Kearns and Umesh Virkumar Vazirani. 1994. *An Introduction to Computational Learning Theory*. MIT Press, Cambridge, MA.

[14] Jon Kleinberg, Sendhil Mullainathan, and Johan Ugander. 2017. Comparison-based choices. In *Proceedings of the 2017 ACM Conference on Economics and Computation*. ACM, New York, NY, 127–144.

[15] R. Laubenbacher and B. Stigler. 2004. A computational algebra approach to the reverse engineering of gene regulatory networks. *J. Theoretical Biology* 229 (2004), 523–537.

[16] Vadim Levit, Tal Grinshpoun, and Amnon Meisels. 2013. Boolean games for charging electric vehicles. In *Proceedings of the 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)-Volume 02*. IEEE Computer Society, Washington, DC, 86–93.

[17] Vadim Levit, Tal Grinshpoun, Amnon Meisels, and Ana LC Bazzan. 2013. Taxation search in boolean games. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, USA, 183–190.

[18] V. Levitt, Z. Komarovsky, T. Grinshpoun, and A. L. C. Bazzan. 2019. Incentive-based search for equilibria in Boolean games. *Constraints* 24 (2019), 288–319.

[19] Kevin P Murphy. 1996. *Passively learning finite automata*. Technical Report 96-04-017. Santa Fe Institute, Santa Fe, NM.

[20] Harikrishna Narasimhan, David C Parkes, and Yaron Singer. 2015. Learnability of influence in networks. In *Advances in Neural Information Processing Systems*. Neural Information Systems Processing Foundation, San Diego, CA, 3186–3194.

[21] J. S. Provan and M. O. Ball. 1983. The Complexity of Counting Cuts and of Computing the Probability that a Graph is Connected. *SIAM J. Computing* 12, 4 (1983), 777–788.

[22] Luigi Sauro and Serena Villata. 2013. Dependency in cooperative boolean games. *Journal of Logic and Computation* 23, 2 (2013), 425–444.

[23] Michael Wooldridge, Ulle Endriss, Sarit Kraus, and Jérôme Lang. 2013. Incentive engineering for Boolean games. *Artificial Intelligence* 195 (2013), 418–439.