Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)

Bayesian Markov Games with Explicit Finite-Level Types

Muthukumaran Chandrasekaran[1] and **Yingke Chen**[2] and **Prashant Doshi**[3]
539G Boyd GSRC, THINC Lab, University of Georgia
Athens, Georgia 30602
{mkran[1],pdoshi[3]}@uga.edu, yke.chen@gmail.com[2]

Abstract

In impromptu or ad hoc settings, participating players are precluded from precoordination. Subsequently, each player's own model is private and includes some uncertainty about the others' types or behaviors. Harsanyi's formulation of a Bayesian game lays emphasis on this uncertainty while the players each play exactly one turn. We propose a new game-theoretic framework where Bayesian players engage in a Markov game and each has private but imperfect information regarding other players' types. Consequently, we construct player types whose structure is explicit and includes a finite level belief hierarchy instead of utilizing Harsanyi's abstract types and a common prior distribution. We formalize this new framework and demonstrate its effectiveness on two standard ad hoc teamwork domains involving two or more ad hoc players.

Empirical findings in strategic games (Goodie, Doshi, and Young 2012) strongly suggest that humans reason about others' beliefs to finite and often low depths. In part, this explains why a significant proportion of human participants find it difficult to play Nash equilibrium profiles of games because it involves them reasoning about others' who in turn reason about the others and so on *ad infinitum* and such reasoning is typically beyond their conscious cognitive capacity. Consequently, Kets (2014) characterized a *finite-level* equilibrium between players having finite-level beliefs in a framework generalized from the standard Harsanyi's Bayesian games.

We further generalize Kets' single-stage framework to allow Bayesian players to play a *Markov game* with incomplete information where players with explicit types are situated in a dynamic environment whose states transitions stochastically to another state sequentially as all participants simultaneously perform their actions at each stage of game. Contextual to these types that induce finite belief hierarchies, we characterize and define the equilibria within this new framework for Bayesian Markov games (BMGs) with *explicit types* as *Markov-perfect finite-level equilibria* and formalize the computation of finding such equilibria as a constraint satisfaction problem solved using an adaptation of Soni et al. (2007)'s constraint satisfaction algorithm. Additionally, we exploit equivalences between these dynamic types, as motivated by behavioral equivalence (Zeng and Doshi 2012), in order to speed up computation of the equilibrium. Consequently, we model the multi-access broadcast channel and foraging problems – well-known evaluation domains for ad hoc teamwork – as a BMG and solve it.

BMG with Finite-Level Types

We *explicitly* define a type space for player i as, $\Theta_i^{\mathcal{K}} = \langle \Theta_i, \mathcal{S}_i, \Sigma_i, \beta_i \rangle$, where Θ_i is the non-empty set of types of player i; $\mathcal{S}_i$ is the collection of all sigma algebras on Θ_i; $\Sigma_i : \Theta_i \rightarrow \mathcal{S}_j$ maps each type in Θ_i to a sigma algebra in $\mathcal{S}_j$; and β_i gives the belief associated with each type of i, $\beta_i(\theta_i) \in \triangle(X \times \Theta_j, \mathcal{F}_X \times \Sigma_i(\theta_i))$, X is the states of nature (set of payoff functions), $\mathcal{F}_X$ is a sigma algebra on X.

We define a Bayesian Markov game (BMG) below which uses Kets' (2014) formalization of finite-level beliefs:

Definition 1 (BMG)**.** *A Bayesian Markov game with finite-level type spaces (BMG) is a collection:*

$$\mathcal{G}^* = \langle S, X, (A_i, R_i, \Theta_i^k)_{i \in N}, T, OC \rangle$$

- *S is the set of physical states of the game;*
- *A_i is the set of player i's actions;*
- *$R_i : S \times X \times \prod_{i \in N} A_i \rightarrow \mathbb{R}$ is i's reward function;*
- *Θ_i^k is the finite-level Kets type space of depth k;*
- *$T : S \times \prod_{i \in N} A \rightarrow \Delta(S)$ is a stochastic physical state transition function of the Markov game; and*
- *OC is the optimality criterion in order to optimize over finite or infinite steps with discount factor, $\gamma \in (0, 1)$.*

A BMG between two agents, i and j of some type θ_i and θ_j respectively, proceeds as follows: both agents initially start at s^t and perform actions a_i^t and a_j^t according to their markov strategies, respectively. The state then transitions to s^{t+1} according to the stochastic transition function of the game, T. Both agents now receive observations, $o_i^{t+1} = \langle s^{t+1}, a_j^t \rangle$ and $o_j^{t+1} = \langle s^{t+1}, a_i^t \rangle$, respectively, that perfectly inform them about current state and other's previous action. Based on these observations, the agents update their respective types and their next actions, a_i^{t+1} and a_j^{t+1}, are selected based on their strategies. Since the mixed-strategy space is continuous, we discretize it using a τ-grid as defined in Soni et al. (2007). We define the equilibrium next.

Definition 2 (Markov-perfect finite level equilibrium)**.** *A profile of strategies, $\boldsymbol{\pi}_k^h = \langle \boldsymbol{\pi}_{i,k}^h \rangle_{i \in N}$ is in Markov-perfect finite-level equilibrium (MPFLE) of level k if the following holds:*

1. *Each player has a Kets type space of level k;*
2. *Strategy $\boldsymbol{\pi}_{j,k}^h$, $j \in N$, $j \neq i$ and at every horizon is comprehensible for every type of player i.*
3. *Each player's strategy for every type is a best response to all other players' strategies in the profile and the equilibrium is subgame perfect.*

Notice that if, instead of condition 1 above, players possess the standard Harsanyi type space, then Def. 2 gives the Markov-perfect Bayes-Nash equilibrium. However, a profile of strategies in equilibrium may not lie on the τ-grid. Therefore, the discretization may introduce error and motivates relaxing the exact equilibrium to ϵ-MPFLE which relaxes the strict requirement of the exact equilibrium allowing a player in approximate equilibrium to deviate if her loss due to deviating to some other strategy is not more than ϵ. Interestingly, an upper bound on ϵ may be obtained for a given value of τ. Finally, a MPFLE may not always exist for a BMG of level k because the given Kets type space of a player may not admit any comprehensible and best response strategy as required by conditions 2 and 3 of Def. 2.

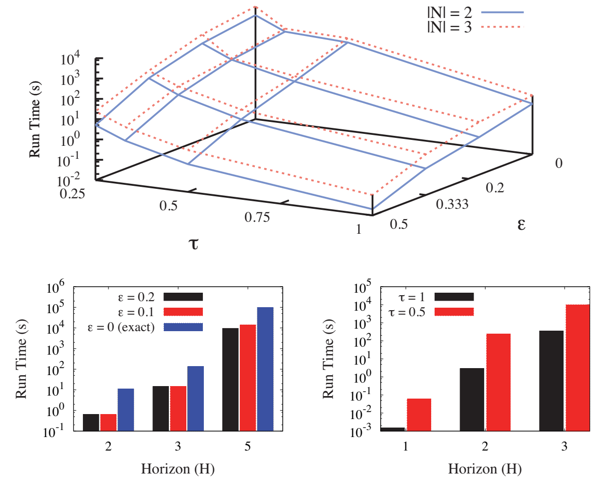
Experimental Results

We implemented the MAC3 constraint satisfaction algorithm for obtaining MPFLE and show the applicability of BMGs toward two benchmark domains used in the ad hoc team work literature: *n-agent multiple access broadcast channel* (nMABC) and *level-based foraging* ($m \times m$ Foraging) .

We begin by noting that the pareto-optimal MPFLE generated by our CSP coincides with the optimal joint policy obtained from a decentralized POMDP using DP-JESP (Nair et al. 2003) for the 2MABC problem. This empirically verifies the correctness of our approach. Multiple equilibria with pure and mixed comprehensible strategies were found for level 1 Kets type spaces. Next, in Fig. **??** we explore the runtime performance of BMG and investigate how varying the different parameters impacts the performance and scalability in the two domains. Next, we systematically and exactly compress large type spaces using exact behavioral equivalence. This *type equivalence (TE)* preserves the quality of the solutions obtained, which we verified empirically as well. Table 1 illustrates the reduction in type space due to TE in 2MABC for $H = 5$ and the computational savings in generating one pure-strategy profile in equilibrium.

Conclusion

To the best of our knowledge, BMG is the first formalization of incomplete-information Markov games played by Bayesian players, which integrates types that induce finite-level beliefs into an operational framework. In particular, BMGs are better suited for modeling ad hoc coordination in comparison to previous game-theoretic frameworks. The framework also offers a promising departure point for modeling empirical data on strategic interactions betwen humans. This further motivates its study and forms an important avenue of future work.



H	Without TE		With TE	
	$\lvert\Theta^{k=1}\rvert$	**Time (s)**	$\lvert\Theta^{k=1}\rvert$	**Time (s)**
5	16	1.56	4	1.26
	36	31.24	16	26.7
	64	>1 day	25	3161.7

Table 1: Computational savings due to TE while computing a pure-strategy MPFLE in 2MABC for Kets level 1 type spaces.

References

Goodie, A. S.; Doshi, P.; and Young, D. L. 2012. Levels of theory-of-mind reasoning in competitive games. *Journal of Behavioral Decision Making* 24:95–108.

Kets, W. 2014. Finite depth of reasoning and equilibrium play in games with incomplete information. Technical Report 1569, Northwestern University, Center for Mathematical Studies in Economics and Management Science.

Singh, S.; Soni, V.; and Wellman, M. 2004. Computing approximate bayes-nash equilibria in tree-games of incomplete information. In *5th ACM Conference on Electronic Commerce*, 81–90.

Soni, V.; Singh, S.; and Wellman, M. P. 2007. Constraint satisfaction algorithms for graphical games. In *Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 423–430.

Zeng, Y., and Doshi, P. 2012. Exploiting model equivalences for solving interactive dynamic influence diagrams. *Journal of Artificial Intelligence Research* 43:211–255.