INFORMATION AGGREGATION IN A BEAUTY CONTEST GAME

Ceyhun Eksin, Pooya Molavi, Alejandro Ribeiro, and Ali Jadbabaie

Department of Electrical and Systems Engineering, University of Pennsylvania

ABSTRACT

We consider a repeated game in which a team of agents share a common, but only partially known, task. The team also has the goal to coordinate while completing the task. This creates a trade-off between estimating the task and coordinating with others reminiscent of the kind of trade-off exemplified by the Keynesian beauty contest game. The agents thus can benefit from learning from others. This paper provides a survey of results from [1–4]. We first present a recent result that states repeated play of the game by myopic but Bayesian agents, who observe the actions of their neighbors over a connected network, eventually yield coordination on a single action. Furthermore, the coordinated action is equal to the mean estimate of the common task given individual's information. This indicates that agents in the network have the same mean estimate in the limit despite the differences in the quality of local information. Finally, we state that if the space of signals is a finite set, the coordinated action is equal to the estimate of the common task given full information, that is, agents eventually aggregate the information available throughout the network on the common task optimally.

Index Terms— Repeated Bayesian games, coordination games, learning, social networks

1. INTRODUCTION

We consider asymptotic behavior in a repeated coordination game with incomplete information played among a population of agents in a network. For a motivating example consider a team of agents within an organization that are assigned to the same project. Agents receive noisy information about the given task and are in departments with only few agents in each department communicating with other departments of the organization. While it is easier to attain agreement on the task within the department, the agreement is harder to achieve interdepartmentally. The question we would like to answer is whether all agents agree on the task assigned and aggregate information optimally when all agents myopically do the best they can.

We model the exemplified organization problem with individual payoffs that include estimation and coordination terms in line with the *beauty contest game* proposed by [5, 6]. The task is defined as the *state of the world* and agents repeatedly receive (possibly uninformative) signals about the task and take actions to maximize individual payoffs. Furthermore, they observe the actions of their peers. We use a network to model the peer interactions (Section 2.1). We characterize individual behavior using the weak perfect Bayesian Nash equilibrium as the solution concept in which no single agent has a profitable deviation from its strategy given its belief (Section 2.2) – see [7, Chs. 4,5] for a discussion of different equilibrium concepts. In Section 3, we survey some of the results presented in [1–3] that show consensus in actions and ex-ante payoffs. We remark that these results extend beyond the beauty contest game payoffs to quadratic symmetric payoffs with strategic complementarities [1]. Next, we show that consensus in actions implies that agents agree on their individual estimates on the state of the world [2, 3]. However, this does not necessarily mean that the agreed estimate is optimal. Finally, we survey a recent result from [4] that shows the coordinated action is generically optimal, that is, it is the action that agents would play given all the signals are commonly known (Section 4). This means that agents eventually learn the best estimate of the task while playing the myopically optimal action. We conclude by a discussion on extensions of some of the results (Section 5). Specifically, we remark that it is possible to obtain the consensus results with endogenous signals – signals that depend on previous actions – and independently generated time varying networks with some minor connectivity assumptions.

The beauty contest game is relevant to research on social value of information [6, 8]. Unlike the coordination concern over a time horizon, these papers focus on the effect of information on the welfare for a single stage beauty contest game. The results in this paper are relevant to literature on Bayesian learning in networks [9–11]. While the papers in this literature are concerned with asymptotic behavior and optimal information aggregation, payoffs of agents in these studies only depend on the state of the world. Hence, there is only information externality among agents, that is, actions of others do not affect individual payoffs directly. In our setup, there is payoff externality caused by actions affecting each others' payoffs. Another line of related research is learning in networks with non-Bayesian agents [12, 13]. In these studies, agents act according to formulations that are ad hoc and are not driven by payoffs. Here, our interest is on learning while agents play according to equilibrium strategy, i.e., agents regard their incentives. Finally,

Work in this paper is supported by ARO P-57920-NS, NSF CAREER CCF-0952867, NSF CCF-1017454, ONR MURI "Next Generation Network Science", AFOSR MURI FA9550-10-1-0567 and ARO MURI W911NF-12-1-0509.

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

the work in this paper is related to learning in games [14–16]. In this literature the concern is whether agents learn to play a Nash equilibrium. The results in this paper contribute to this literature by considering interactions over a network.

2. BEAUTY CONTEST GAME

There are N agents repeatedly playing a game with payoffs depending on the unknown state of the world $\theta \in \Theta$. Agents have common prior denoted by μ on the state of the world. We make the following technical assumption on the set Θ.

Assumption 1 *The set Θ is bounded and measurable subset of $\mathbb{R}$.*

The game is played over discrete time stages $t = 1, 2, \ldots$ and at each stage t agent i takes action $a_{it} \in A_i = \mathbb{R}$. We use $\mathbf{a}_t = \{a_{1t}, \ldots, a_{Nt}\} \in A = \mathbb{R}^N$ to denote the action profile at time t. Agent i receives payoff depending on the state of the world θ and the action profile $\mathbf{a}$,

$$u_i(\mathbf{a}, \theta) = -(1-\lambda)(a_i - \theta)^2 - \frac{\lambda}{N-1} \sum_{j \neq i} (a_i - a_j)^2 \quad (1)$$

where $\lambda \in [0, 1)$ is a constant. The first term is the quadratic loss in estimation of the true state of the world. The second term is the coordination term representing the loss due to miscoordination with the rest of the population. This term is also referred to as the 'beauty contest term' in [6] representing the agents' goal to second guess actions of each other. The constant λ gauges the importance between estimation and coordination terms. When $\lambda = 0$, agents' actions do not affect payoff of each other and we are left with a payoff ubiquitous in distributed estimation literature. In this case, agents only have information externality whereas when $\lambda > 0$, there is both payoff and information externality.

2.1. Network, Signals and Histories

Agents interact according to a network G with agents in the nodes and edge set E. The neighbor set of agent i including himself is denoted by $\mathcal{N}_i := \{i\} \bigcup \{j : (j, i) \in E\}$. We make the following connectivity assumption on G.

Assumption 2 *The network G is strongly connected.*

The strong connectivity assumption is required to make sure that the information received by an agent is able to eventually flow to all the other agents in the network.

The information that agent i receives at time t consists of signal $s_{it} \in S_i$ and actions of its neighbors $a_{\mathcal{N}_i t-1}$ from the previous stage. We denote the signal profile at time t as $\mathbf{s}_t = \{s_{1t}, \ldots, s_{Nt}\} \in S = \times_{j=1}^{N} S_j$. Signal profile at time t is generated with respect to a probability distribution on S that depends on the state of the world and previous signals, $\pi_t(\theta, \mathbf{s}_{1:t-1})[\cdot] : \Theta \times S^{t-1} \rightarrow [0, 1]$. We define the measurable space consisting of state of the world and signals as $\Xi := \Theta \times S^{\mathbb{N}}$ and let $\mathcal{Z}$ be the corresponding σ-algebra.

A play consists of the true state of the world, signals and actions over the time horizon, hence belongs to the space $\Omega = \Theta \times (S \times A)^{\mathbb{N}}$. Let $(\Omega, \mathcal{B})$ be the measurable space of play with the Borel σ-algebra $\mathcal{B}$. The state of the world, signals and actions until time t make up the history of the play up to time t and can be defined by the following recursion,

$$h_t = (h_{t-1}, \mathbf{a}_{t-1}, \mathbf{s}_{t-1}) \quad (2)$$

where $h_1 = \theta$.

Agent i observes a subset of the history determined by private signals and neighboring actions,

$$h_{it} = \{h_{i,t-1}, \mathbf{a}_{\mathcal{N}_i t-1}, s_{it-1}\} \quad (3)$$

with $h_{i1} = \emptyset$. The private history of agent i at time t belongs to the set H_{it}. The history of agent i over the whole horizon belongs to the set $H_i := \bigcup_{t=1}^{\infty} H_{it}$. We denote σ-algebras defined by the Borel sets of H_{it} and H_i with $\mathcal{H}_{it}$ and $\mathcal{H}_i$, respectively. The union over all $\mathcal{H}_i$ is denoted by $\mathcal{H}$.

2.2. Equilibrium

The equilibrium play of agent i at time t is determined by his strategy and belief.

A strategy is an action plan for any possible observation of the play by i, that is, $\sigma_{it} : H_{it} \rightarrow A_i$. The strategy profile of agent i is denoted by σ_i. The strategy profile of the network $\sigma := \{\sigma_1, \ldots, \sigma_N\}$ determines the path of play together with realization of the state of the world and signals. Any prior μ and the $\{\pi_t\}_{t \in \mathbb{N}}$ induces a probability distribution P over the measurable space of state and signals $(\Xi, \mathcal{Z})$. The strategy profile σ determines the path of play given P, that is, strategy profile σ induces a probability distribution P^σ on the space of plays $(\Omega, \mathcal{B})$ together with the prior μ and signal distributions $\{\pi_t\}_{t \in \mathbb{N}}$. We let E^σ denote the expectation with respect to P^σ.

Agent i's belief at time t is a probability distribution on the space of play Ω conditional to his observed private history h_{it} denoted by $q_{it}(\cdot) = P^\sigma(\cdot | \mathcal{H}_{it})$. In words, the belief of agent i measures the likelihood of a path of play given the information set of i. The conditional probability and the recursive nature of the private histories in (3) indicate that agents follow Bayes' rule to form their beliefs.

Next, we define our equilibrium notion which requires expected payoff maximization and a consistency between the strategies and beliefs. Note that the payoff of agent i depends on actions of other agents in (1), hence he needs to reason about strategies of other agents in order to compute his expected payoff. Given the strategy profile σ, reasoning about others' strategies means reasoning about their histories. Given the strategy profile σ and a realization of the state and signals $\xi \in \Xi$, the realization of history of agent i is a function of ξ, that is, there exists a function $\tilde{h}_{it}(\cdot) : \Xi \rightarrow H_{it}$. Hence, the strategy of agent j can be expressed a function that maps ξ to his action space, $\sigma_{jt}(\tilde{h}_{jt}(\cdot)) : \Xi \to A_i$. In consequence, agent i's belief on the space of state and signals suffices to form beliefs on strategy of j given strategy profile σ and $\tilde{h}_{jt}(\cdot)$. In light of this discussion, we define the weak perfect Bayesian Nash equilibrium strategy profile as follows – see [15] for a similar definition.

Definition 1 *The strategy profile σ^* is an equilibrium strategy profile such that for all t and i,*

$$E^{\sigma^*}[u_i(\sigma^*_{it}(h_{it}),\sigma^*_{-it},\theta) \,\big|\, \mathcal{H}_{it}] \geq E^{\sigma^*}[u_i(\sigma_{it}(h_{it}),\sigma^*_{-it},\theta) \,\big|\, \mathcal{H}_{it}] \quad (4)$$

where $\sigma^*_{-it} := \{\sigma^*_{jt}\}_{j \neq i}$.

According to the equilibrium definition, there is no unilateral deviation for agent i that provides a higher expected payoff given his belief. Note that the agents consider present payoffs only, hence they are myopic. The definition requires a consistency between the way agents form their beliefs and the way strategies are selected. The expectation in condition (4) is with respect to the belief of agent i, $q_{it} = P^{\sigma^*}(\cdot|\mathcal{H}_{it})$. We emphasize that q_{it} is the conditional probability distribution induced by σ^*. Hence, the belief is based on the equilibrium strategy which is determined by the payoff and the belief itself in (4).

In this paper, we rule out equilibria where each agent's expected payoff is equal to minus infinity, regardless of his own strategy by restricting our attention to square integrable strategies. Given that the strategies are square integrable, the equilibrium condition in (4) at time t can be equivalently written as follows by the first order condition of optimality,

$$\sigma^*_{it} = (1-\lambda)E^{\sigma^*}[\theta|\mathcal{H}_{it}] + \frac{\lambda}{N-1}\sum_{j\neq i} E^{\sigma^*}[\sigma^*_{jt}|\mathcal{H}_{it}]. \quad (5)$$

Note that first order condition is necessary and sufficient because the expected payoff of an individual is concave with respect to his own action. We observe that the equilibrium actions are a linear combination of the expectation of the state and actions of others. Using this form of the equilibrium strategy at time t together with the fact that $\lambda < 1$, it can be shown that the equilibrium strategy is almost surely unique [4].

Proposition 1 *If Assumption 1 hold then a square integrable equilibrium strategy profile σ^* exists and is P^{σ^*}-almost surely unique for all time t.*

The result above follows by first observing that the equivalent representation of the equilibrium in (5) is a fixed point equation at all time t and then showing that the right hand side of the equation is a contraction mapping for any distribution. Finally, we resort to a fixed point theorem to show uniqueness. The result here can be generalized to other symmetric quadratic games with strategic complementarities, that is, quadratic games in which an increase in someone's action reinforces an increase in others' action and the reinforcement level is symmetric [1]. In the following sections, we use the properties of the game and its equilibrium strategies to characterize individuals' asymptotic behavior.

3. CONSENSUS

In this section, we show asymptotic consensus in ex-ante payoffs, actions and conditional expectation on the true state of the world. The first results shows that agents' limit strategy actually converges.

Proposition 2 *If Assumption 1 holds and the equilibrium strategy σ^* is square integrable then for all $i = 1, \dots, N$, as $t \to \infty$*

$$E^{\sigma^*}[(\sigma^*_{it} - \varsigma_i)^2] \longrightarrow 0$$

where $\varsigma_i : \Xi \to A_i$.

This result shows that agent i's strategy converges to some limit strategy ς_i in L^2 sense. This plays a key role in neighbors identifying limit strategies of their neighbors. In the next theorem we state the main consensus results of this section and discuss implications of Proposition 2 on these results.

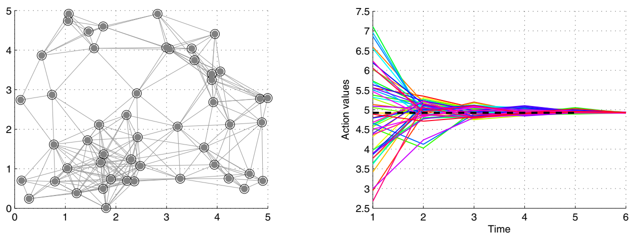
Theorem 1 *If Assumptions 1 and 2 hold and σ^* is a square integrable equilibrium strategy profile then as $t \to \infty$, for all $i = 1, \dots, N$ and $j = 1, \dots, N$,*

$$E^{\sigma^*}[(\sigma^*_{it} - \sigma^*_{jt})^2] \to 0 \quad (6)$$

$$E^{\sigma^*}[|u_i(\sigma^*_t, \theta) - u_j(\sigma^*_t, \theta)|] \to 0. \quad (7)$$

The first statement in (6) indicates that agents' limit actions are the same eventually (see [4] for a proof). The result relies on the imitation principle which is first introduced by [9]. Agent i observes limit strategy of $j \in \mathcal{N}_i$ infinitely often hence he can identify j's limit strategy, i.e., ς_j in Proposition 2 becomes measurable with respect to information of i. As a result, agent i can imitate the limit action of j. However, agent i would not unilaterally deviate to j's limit action as it is not profitable to deviate by the definition of equilibrium profile (4). The same argument applies to j and his neighbors and his neighbors' neighbors eventually ending up at an agent that observes i's limit action by the strong connectivity of the network. Consequently, we obtain a chain of agents starting and ending at i where each agent's ex-ante expected utility is greater than or equal to his neighbor's expected utility. Using the equilibrium strategy profile for the beauty contest game in (5), it can be shown that this can only happen if all limit actions are equal. The second statement of Theorem 1 in (7) shows that prior to the start of the game the agents are expected to receive the same utility asymptotically. This result follows from the consensus of the limit actions in (6). Intuitively, the coordination term in (1) disappears when all agents take the same action and the estimation term is identical for all again by using (6).

Even though the ex-ante expected utilities are equal, the conditional expected utilities are not necessarily equal. That is, an agent might think that he is receiving a higher payoff



(a) (b)

Fig. 1. (a) Geometric network with $N = 50$ agents on a 5×5 area and a connection threshold of 1.5 units. (b) Individual action values over time. Dashed line indicates the optimal action given all the information.

even if the utility function is symmetric and they all play the same action. The examples where agents are expecting different payoffs can occur when their beliefs on the state of the world disagree as we show in the following example.

Example 1 *There are $N = 2$ agents with uniform priors on the state of the world $\theta \in \{-2, -1, 1, 2\}$. Agent 1 at time $t = 2$ observes a signal that reveals him the absolute value of θ, that is, $s_{12} = |\theta|$. Agent 2 does not observe any informative signal at any point in time. Given the priors and signal structure, agent i's equilibrium strategy is to play $\sigma^*_{it} = 0$ for $i = 1, 2$ and all time t. The conditional expected utility of agent 1 is equal to $E^{\sigma^*}[u_1(\sigma^*_t, \theta) \,|\, s_{12})] = -|\theta|^2(1 - \lambda)$, whereas agent 2's is equal to $E^{\sigma^*}[u_2(\sigma^*_t, \theta) \,|\, s_{22})] = -5/2(1 - \lambda)$ for $t \geq 1$.*

4. OPTIMAL INFORMATION AGGREGATION

The previous section showed that agents eventually agree on their limit actions but so far we have not specified what the agreed limit action is and whether it is the optimal action or not. When the limit action in (6) is substituted into the equilibrium strategy profile in (5), we obtain that the limit action is equal to

$$\sigma^*_{it} \rightarrow E^{\sigma^*}[\theta | \mathcal{H}_i] \quad \text{as} \quad t \rightarrow \infty \tag{8}$$

for $i = 1, \ldots, N$. Moreover, by the above statement and the consensus in actions, it must be that

$$E^{\sigma^*}[\theta | \mathcal{H}_i] = E^{\sigma^*}[\theta | \mathcal{H}_j] \tag{9}$$

for all i and j. By (9), agents' mean estimate on the state of the world is asymptotically identical even if they have asymmetric information. Note that the optimal action given common knowledge of all the signals is to play $E^{\sigma^*}[\theta | \mathcal{H}]$. This result in (9) does not imply that the agents necessarily converge to the optimal action. As it turns out it need not be that the limit action is the optimal action. In [4], we provide an example where the limit action is different than the optimal action. However, the following result shows that agents aggregate information and agree on the optimal action generically.

Before we state the result, we define what is meant by generically. We consider the set of all probability distributions $\mathbf{P}$ over $(\Xi, \mathcal{Z})$. As mentioned in Section 2.2, any prior μ and signal function $\{\pi_t\}_{t \in \mathbb{N}}$ induces a probability distribution P over $(\Xi, \mathcal{Z})$. Any P is uniquely induced by some prior and signal function up to sets of measure zero. A generic set in the space $\mathbf{P}$ is the complement of a union of countably many nowhere dense subsets of $\mathbf{P}$. Hence, a generic point P belongs to a generic set in the space $\mathbf{P}$. Next, we state the main result of this section – see Theorem 2 in [4] for a proof and more detailed discussion.

Theorem 2 *If Assumptions 1 and 2 hold and S is a finite set then for a generic point P in $\mathbf{P}$ and all $i = 1, \ldots, N$*

$$E[|\sigma^*_{it} - E[\theta | \mathcal{H}]|] \rightarrow 0 \quad \textit{as} \quad t \rightarrow \infty \tag{10}$$

Theorem 2 states that agents asymptotically aggregate information optimally and learn the best estimate of θ given all the signals even when they are myopic. Note that this does not necessarily imply that agents learn all the private signals. We exemplify this result in a numerical experiment.

Example 2 *Let $\lambda = 2/3$ and $\theta = 5$. $N = 50$ agents on a geometric network receive an initial signal Gaussian with mean θ and variance 1. Fig. 1(a) depicts the geometric network. Fig. 1(b) illustrates the individual action values over time – see [17] for a tractable algorithm to recursively compute equilibrium strategies and beliefs. We observe that agents actions converge to the mean of the private signals that is marked by black dashed lines in Fig. 1(b). Moreover, convergence is fast in the order of the diameter of the network.*

5. CONCLUSION

We considered a repeated beauty contest game played among a networked population. The agents acted according to their equilibrium strategies and the interest was whether they aggregate information optimally on equilibrium play. We surveyed a set of results that showed equilibrium actions eventually converge to a consensus strategy which is their estimate of the true state of the world. This meant that agents eventually agree on their estimate. Finally, we showed that in general, this agreement implies that agents aggregated information optimally. Throughout the paper we made the assumption that the network is strongly connected. It is possible to obtain the consensus results in Section 3 when this assumption is relaxed to random time varying networks where there exists an underlying strongly connected network [4]. Secondly, the consensus results hold when the signals are endogenously generated, that is, the signal function depends on previous actions [4]. On the other hand, exogenous signals are required for information aggregation. Finally, we considered the beauty contest game to show our results which is a surrogate for games with symmetric quadratic payoffs with strategic complementarities.

6. REFERENCES

[1] C. Eksin, P. Molavi, A. Ribeiro, and A. Jadbabaie, "Learning in linear games over networks," in *Proceedings of the 50th Annual Allerton Conference on Communications, Control, and Computing (to appear)*, Allerton, Illinois, USA., 2012.

[2] C. Eksin, P. Molavi, A. Ribeiro, and A. Jadbabaie, "Games with side information in economic networks," in *Proceedings of the 46th Asilomar Conference on Signals, Systems and Computers (invited)*, Pacific Grove, California, USA., 2012.

[3] P. Molavi, C. Eksin, A. Ribeiro, and A. Jadbabaie, "Learning to coordinate in a beauty contest game," in *Proc. Control and Decision Conference*, Florence, Italy, 2013.

[4] P. Molavi, C. Eksin, A. Ribeiro, and A. Jadbabaie, "Learning to coordinate in social networks," *Available at SSRN: http://ssrn.com/abstract=2292124*, July 2013.

[5] J.M. Keynes, *The General Theory of Interest, Employment and Money*, Macmillan, London, UK, 1936.

[6] S. Morris and H.S. Shin, "The social value of public information," *American Economic Review*, vol. 92, pp. 1521–1534, 2002.

[7] R.B. Myerson, *Game Theory: Analysis of Conflict*, Harvard Univ. Press, Cambridge, MA, 1991.

[8] G.M. Angeletos and A. Pavan, "Efficient use of information and social value of information," *Econometrica*, vol. 75, no. 4, pp. 1103–1142, 2007.

[9] D. Gale and S. Kariv, "Bayesian learning in social networks," *Games Econ. Behav.*, vol. 45, pp. 329–346, 2003.

[10] D. Rosenberg, E. Solan, and N. Vieille, "Informational externalities and emergence of consensus," *Games Econ. Behav.*, vol. 66, pp. 979–994, 2009.

[11] P.M. Djuric and Y. Wang, "Distributed bayesian learning in multiagent systems," *IEEE Signal Process. Mag.*, vol. 29, pp. 65–76, March, 2012.

[12] A. Jadbabaie, P. Molavi, A. Sandroni, and A. Tahbaz-Salehi, "Non-Bayesian social learning," *Games and Economic Behavior*, vol. 76, no. 4, pp. 210–225, 2012.

[13] P. M. DeMarzo, D. Vayanos, and J. Zwiebel, "Persuasion bias, social influence, and unidimensional opinions," *The Quarterly Journal of Economics*, vol. 118, pp. 909–968, 2003.

[14] J.S. Jordan, "Bayesian learning in normal form games," *Games and Economic Behavior*, vol. 3, pp. 60–81, 1991.

[15] M.O. Jackson and E. Kalai, "Social learning in recurring games," *Games and Economic Behavior*, vol. 21, pp. 102–134, 1997.

[16] D. P. Foster and H. P. Young, "Learning, hypothesis testing, and nash equilibrium," *Games and Economic Behavior*, vol. 45, pp. 73–96, 2003.

[17] C. Eksin, P. Molavi, A. Ribeiro, and A. Jadbabaie, "Bayesian quadratic network game filters," *ArXiv e-prints - arXiv:1302.0249v1*, February 2013.