Chapter 9
Preplay Communication in Multi-Player Sequential Games: An Overview of Recent Results

Andrea Celli

9.1 Introduction

The computational study of game-theoretic solution concepts is fundamental to describe the optimal behavior of rational agents interacting in a strategic setting, and to predict the most likely outcome of a game. Equilibrium computation techniques have been applied to numerous real-world problems. Among other applications, they are the key building block of the best poker-playing AI agents [5, 6, 27], and have been applied to physical and cybersecurity problems (see, e.g., [18, 20, 21, 30–32]).

In this section, we start by presenting a simple example and reviewing the fundamental model of sequential games with imperfect information. Then, we give a brief overview of problems involving preplay communication, which will be the focus of this summary.

9.1.1 Motivating Example

The following simple example shows that, even in toy problems, preplay communication can give a significant advantage to players that decide to implement it.

Example 1 (A simple coordination game) There are three players and each player has two available actions (which are denoted by LEFT and RIGHT). Moreover, player 1 and player 2 have identical goals, while player 3 has the opposite objectives. Therefore, player 1 and 2 will try to maximize their expected utilities and player 3 will act as an adversary, whose goal is minimizing player 1 and 2's expected utility.

A. Celli (✉)

Politecnico di Milano, Milano, Italy

e-mail: andrea.celli@polimi.it

   The strategic interaction goes as follows: as the game starts, each player has to select one action, without observing the choice of the others. Player 1 and 2 receive a payoff of $K > 0$ if they correctly guess the action taken by player 3 (e.g., when player 3 plays LEFT, player 1 and 2 are rewarded K if they both play LEFT), and they receive payoff 0 otherwise. Player 3 receives a payoff of $-K$ whenever player 1 and 2 receive payoff K, and 0 otherwise.

   The best that player 3 can do in order to fool the other players is selecting an action between LEFT/RIGHT with equal probability. Intuitively, this is the strategy of player 3 which makes it more difficult for the others to guess the action he/she chose. On the other hand, since player 1 and 2 have equal objectives, it may be profitable for them to jointly plan their strategies before the beginning of the game. Specifically, they could decide to play actions (LEFT, LEFT) with probability 0.5, and actions (RIGHT, RIGHT) with probability 0.5. In this way, they would be able to avoid outcomes with rewards always equal to 0 in which they play different actions from each other. Before the beginning of the game, player 1 and 2 could toss a coin and select whether to play (LEFT, LEFT) or (RIGHT, RIGHT) depending on the outcome of the coin toss. In this way, the expected utility of player 1 and 2 would be $K/2$ as they would reach each outcome with payoff K with probability 0.25. If player 1 and 2 decided not to communicate before the beginning of the game the best that they could do is employing player 3's strategy (i.e., selecting one of their actions randomly). In this way, player 1 and player 2 would get a payoff of $K/4$, since each outcome with payoff K is reached with probability 0.125.

   Even in such a simple game the possibility of exploiting preplay communication yield an increase of 50% in the expected utility of players who adopted it. This increase can be arbitrarily large in game instances which are slightly more complex [13]. In the remainder of this summary, we will discuss how to model such strategic scenarios and review some recent results related to these problems.

9.1.2   Sequential Games with Imperfect Information

We focus on a powerful model of strategic interaction which can model sequential moves, imperfect information, and outcome uncertainty. An *extensive-form game* (EFG) (*a.k.a.* a *sequential game*) models a sequential interaction among players. An EFG is represented as a game tree, where each node is identified by the ordered sequence of actions leading to it from the root node. Each node represents a decision point of the game and is associated to a single player, who has a set of available actions at that node represented by its branches. A payoff for each player is associated to each leaf node (*terminal node*) of the game tree. Finally, exogenous stochasticity is modeled via a virtual player (*a.k.a. nature* or *chance*) that plays non-strategically (i.e., it plays according to a fixed strategy). Chance is used to describe, e.g., the probability of receiving a certain hand in a card playing game. We say that a two-player game is *zero-sum* if, for each terminal node, the sum of the utilities of the
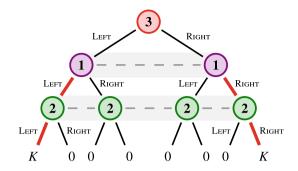
Fig. 9.1 The game described in Example 1. Decision nodes of each player are denoted by the number of the player taking action at that decision point. Each player has a single information set. Nodes belonging to the same information set are within the same grey area. Terminal nodes show the payoffs of player 1 and 2. Branches of the tree in red highlight the pairs of actions that player 1 and 2 should play when using preplay communication

two players equals 0. A game is *constant-sum* if, for each terminal node, the sum of players' utilities is equal to a certain constant. If this does not hold, we say that the game is *general-sum*.

In general, a player may not be able to observe all the other players' actions, and players may have information on the state of the game which is not shared (i.e., in poker each player does not know other players' hands). Imperfect information is represented via *information sets* (or *infosets*), which group together decision nodes of a certain player that are indistinguishable to her. We assume players have *perfect recall*, that is they have perfect memory of their past actions and observations. Figure 9.1 shows the EFG representing the game described in Example 1.

The most widely adopted notion of equilibrium is that of *Nash equilibrium* [29]: each player should not have incentives in deviating from his/her strategy, assuming the other players do not deviate either. In Example 1, players reach a Nash equilibrium when each of them selects an action according to a uniform probability distribution over the available actions.

9.1.3   Preplay Communication

In the context of imperfect-information games, a vast body of literature focuses on the computation of Nash equilibria in two-player, zero-sum games (see, e.g., [19, 33, 34]), where recent results demonstrated that it is possible to compute strong solutions in theory and practice.

While relevant, two-player, zero-sum games are rather restrictive, as many practical scenarios are not zero-sum and involve more than two players. Moreover, especially in general-sum games, the adoption of a Nash equilibrium may present some difficulties when used as a prescriptive tool. Indeed, when multiple Nash equilibria

coexist, the model prevents players from synchronizing their strategies, since communication between players is prohibited. In real-world scenarios, where some form of communication among players is usually possible, different solution concepts are required as communication allows players for coordinated behaviors.

We focus on scenarios where players can exploit *preplay communication* [24, 25], i.e., players have an opportunity to discuss and agree on tactics before the game starts, but will be unable to communicate during the game. Consider, as an illustration, the case of a poker game where multiple players are colluding against an identified target player. Colluders can agree on shared tactics before the beginning of the game, but are not allowed any explicit communication while playing. In other settings, players might be forced to cooperate by the nature of the interaction itself. This is the case, for instance, in Bridge. Preplay players' coordination introduces new challenges with respect to the case in which agents take decisions individually, as understanding how to coordinate before the beginning of the game requires reasoning over the entire game tree. It is easy to see that this causes an exponential blowup in the agents' action space and, therefore, even relatively small game instances are usually deemed intractable in this setting.

When modeling preplay communication, it is instructive to introduce an additional agent, called the *mediator*, that does not take part in the game, but may send signals (usually actions' recommendations) to other players just before the beginning of the game. In the following sections, we explore different forms of preplay coordination in sequential games. The scenarios we consider can be classified through the following questions: (i) *who is receiving the mediator's recommendations?* (ii) *do players have similar goals?* (iii) *is the mediator self-interested? and does the mediator have more information on the state of the game than other players?* When the mediator is sending signals only to members of the same team we talk about *adversarial team games* (Sect. 9.2). In more general settings, players may not have identical objectives. In these scenarios, the mediator has to accurately plan incentives for each individual player and the problem becomes the computation of a *correlated equilibrium* (Sect. 9.3). Finally, the mediator may hold more information than the players of the game, and he/she may be willing to use this information asymmetry to achieve his/her own goals. This setting is modeled via the Bayesian persuasion framework (Sect. 9.4).

9.2 Adversarial Team Games

A recent line of research focuses on preplay communication in *team games*. A *team* of agents is defined as a set of players sharing the same objectives. Following this simple definition, player 1 and 2 of Example 1 form a team. An interesting problem is understanding how team members can coordinate their actions when facing an opponent (e.g., player 3 in the example). We call these games *adversarial team games*. Even without communication during the game, the planning phase gives the team members an advantage: for instance, the team members could skew their strategies to

use certain actions to signal about their state (for example, in card-playing games, the current hands they're holding). In other words, by having agreed on each member's planned reaction under any possible circumstance of the game, information can be silently propagated in the clear, by simply observing public information.

Initially, adversarial team games where studied in games with simultaneous moves [2, 4]. Celli and Gatti [13] first studied the setting in which a team of agents faces an adversary in a sequential interaction. This work formally defines the game model and shows that different forms of intra-team communication result in different models of coordination: (i) a mediator that can send and receive *intraplay* signals (i.e., messages are exchanged during the execution of the game); (ii) a mediator that only exploits preplay communication, sending recommendations just before the beginning of the game; (iii) team members jointly plan their strategies, but have no access to a mediator to synchronize action execution. The main focus has been on the second scenario, where only preplay communication is possible. Scenarios (i) and (iii) are instructive to understand the advantages of different forms of intra-team communication. These different coordination capabilities are compared via the analysis of inefficiency indexes measuring the relative losses in the team's expected utility. Interestingly, their experimental evaluation shows that, in practice, preplay communication is often enough to reach near-optimal performances. An application of these techniques is the work of Basilico et al. [3], where team games are used to coordinate patrollers in environments at risk.

Motivated by the complexity of the problem of coordinating team members with preplay communication, Farina et al. [23] present a scalable learning algorithm to compute an approximate solution to this problem. In doing so, the authors highlight a strong analogy with imperfect-recall games, and propose a new game representation, called *realization form*, which can also be applied to this setting. Then, they exploit the new representation to derive an auxiliary construction that allows one to map the problem of finding an optimal coordinated strategy for the team to the well-understood Nash equilibrium-finding problem in a (larger) two-player zero-sum perfect-recall extensive-form game. By reasoning over the auxiliary game, they devise an anytime algorithm, *fictitious team-play*, that is guaranteed to converge to an optimal coordinated strategy for the team against an optimal opponent. Then, they demonstrate the scalability of the learning algorithm on standard imperfect-information test instances (such as, Leduc hold'em poker and Goofspiel).

9.3 Correlated Equilibria in Sequential Games

The members of a team share the same objectives. Therefore, the mediator does not have to enforce any incentive-compatibility constraint. However, it may happen that agents receiving mediator's recommendations do not share the same objectives. In this case, the mediator has to incentivize each player to follow moves' recommendations. Here, the mediator is assumed to be *benevolent*, i.e., she aims at maximizing the expected social welfare of the game.

We briefly discuss some works investigating whether correlation can be reached efficiently even in settings where players have limited communication capabilities (i.e., they can only observe signals before the beginning of the game). Therefore, we focus on sequential games in which only preplay communication is admitted, and study correlated equilibria that allow the mediator to recommend actions just *before* the playing phase of the game (namely, the *correlated equilibrium* (CE) [1] and the *coarse correlated equilibrium* (CCE) [28]).

These problems have been proved to be computationally hard in most settings. Celli et al. [11] provide several results characterizing the complexity of computing optimal (i.e., social welfare maximizing) CEs and CCEs, and their approximation complexity. First, in an extended version of the paper, they prove that approximating an optimal (i.e., social welfare maximizing) CE is not in Poly-APX even in two player games without chance moves, unless $P = NP$. Next, they identify the conditions for which finding an optimal CCE is NP-hard. However, they show that an optimal CCE can be found in polynomial-time in two-player extensive-form games without chance moves. Finally, Celli et al. [14] complete the picture on the computational complexity of finding social-welfare-maximizing CCEs by showing that the problem is not in Poly-APX, unless $P = NP$, in games with three or more players (chance included).

There are various algorithms for computing CCEs in general-sum, multi-player, sequential games. Celli et al. [11] provide a column generation framework to compute optimal CCEs in practice, and show how to generalize it to the *hard cases* of the problem. Celli et al. [14] focus on the problem of computing *an ε-CCE* (i.e., an approximate CCE). The authors present an enhanced version of CFR [34] which computes an average correlated strategy which is guaranteed to converge to an approximate CCE with a bound on the regret which is sub-linear in the size of the game tree.

9.4 Bayesian Persuasion with Sequential Games

Finally, it may happen that the mediator is self-interested, and may exploit asymmetries in the availability of information to design a signaling scheme, in order to persuade players to select favorable actions. In this setting, the mediator is looking for a way to coordinate the individual behavior of each player in order to reach a preferred outcome of the game.

Celli et al. [12] examine information-structure design problems as a means of forcing coordination towards a certain objective. More precisely, they start from the usual scenario where a mediator can communicate action recommendations to players before the beginning of a sequential game. Suppose that parties (i.e., the mediator and the players) are asymmetrically informed about the current state of the game. Specifically, the mediator is able to observe more information than the other players. Celli et al. [12] pose the following question: *can the mediator exploit the information asymmetry to coordinate players' behavior toward a favorable outcome?*

This problem can be accurately modeled via the *Bayesian persuasion* framework [26]. Celli et al. [12] investigate private persuasion problems with multiple

receivers interacting in a sequential game, and study the continuous optimization problem of computing a private signaling scheme which maximizes the sender's expected utility. The authors show how to address sequential, multi-receiver settings algorithmically via the notion of *ex ante* persuasive signaling scheme, where the receivers commit to following the sender's recommendations having observed only the signaling scheme. They show that an optimal *ex ante* signaling scheme may be computed in polynomial time in settings with two receivers and independent action types, which makes *ex ante* persuasive signaling schemes a persuasion tool which is applicable in practice. Moreover, they show that this result cannot be extended to settings with more than two receivers, as the problem of computing an optimal *ex ante* signaling scheme becomes NP-hard.

9.5 Discussion and Future Research

The research on equilibrium computation in general-sum, multi-player, sequential games has not yet reached the level of maturity reached in the two-player, zero-sum setting, where it is possible to compute strong solutions in theory and practice. In these settings, equilibrium selection problems may render the choice of the appropriate solution concept not obvious, since the Nash equilibrium may not be the appropriate one. Many practical scenarios allow for some form of communication, mitigating the equilibrium selection issue. In this summary, we presented some multi-player problems where players can reach some form of coordination via preplay communication.

There are many interesting questions that need to be addressed in the future. We outline some of them for each of the settings we described. First, it would be interesting to develop a scalable end-to-end approach to learning an optimal team coordinated strategy without prior domain knowledge. Some works going in this direction are Chen et al. [17], Celli et al. [10]. Moreover, as pointed out by Celli et al. [16], the study of algorithms for team games could shed further light on how to deal with imperfect-recall games, which are receiving increasing attention in the community due to the application of imperfect-recall abstractions to the computation of strategies for large sequential games. As for the computation of correlated equilibria in sequential games, it would be interesting to further investigate whether it is possible to define regret-minimizing procedures for general EFGs leading to refinements of the CCEs, such as EFCCEs [22]. A recent work studying a related problem is Celli et al. [15]. Finally, it would be interesting to complement recent works on Bayesian persuasion problems such as Castiglioni et al. [7–9] with scalable algorithms that can be effectively applied to real-world problems.

References

1. Aumann R (1974) Subjectivity and correlation in randomized strategies. J Math Econ 1(1):67–96
2. Basilico N, Celli A, De Nittis G, Gatti N (2017) Team-maxmin equilibrium: efficiency bounds and algorithms. In: AAAI conference on artificial intelligence (AAAI)
3. Basilico N, Celli A, De Nittis G, Gatti N (2017) Coordinating multiple defensive resources in patrolling games with alarm systems. In: Proceedings of the 16th conference on autonomous agents and multiagent systems (AAMAS), pp 678–686
4. Basilico N, Celli A, De Nittis G, Gatti N (2017) Computing the team–maxmin equilibrium in single–team single–adversary team games. Intell Artif
5. Brown N, Sandholm T (2017) Superhuman AI for heads-up no-limit poker: libratus beats top professionals. Science eaao1733:2017
6. Brown N, Sandholm T (2019) Superhuman ai for multiplayer poker. Science 365(6456):885–890
7. Castiglioni M, Celli A, Gatti N (2019) Persuading voters: it's easy to whisper, it's hard to speak loud. arXiv:1908.10620
8. Castiglioni M, Celli A, Gatti N (2020) Public bayesian persuasion: being almost optimal and almost persuasive. arXiv:2002.05156
9. Castiglioni M, Celli A, Marchesi A, Gatti N (2020) Signaling in bayesian network congestion games: the subtle power of symmetry. arXiv:2002.05190
10. Celli A, Ciccone M, Bongo R, Gatti N (2019) Coordination in adversarial sequential team games via multi-agent deep reinforcement learning. arXiv:1912.07712
11. Celli A, Coniglio S, Gatti N (2019) Computing optimal ex ante correlated equilibria in two-player sequential games. In: AAMAS
12. Celli A, Coniglio S, Gatti N (2020) Private bayesian persuasion with sequential games. In: AAAI
13. Celli A, Gatti N (2018) Computational results for extensive-form adversarial team games. In: AAAI
14. Celli A, Marchesi A, Bianchi T, Gatti N (2019) Learning to correlate in multi-player general-sum sequential games. In: NeurIPS
15. Celli A, Marchesi A, Farina G, Gatti N (2020) No-regret learning dynamics for extensive-form correlated and coarse correlated equilibria. arXiv:2004.00603
16. Celli A, Romano G, Gatti N (2019) Personality-based representations of imperfect-recall games. In: Proceedings of the 18th international conference on autonomous agents and multi-agent systems, AAMAS, pp 1868–1870
17. Chen L, Guo H, Du Y, Fang F, Zhang H, Zhu Y, Zhou M, Zhang W, Wang Q, Yu Y (2019) Signal instructed coordination in cooperative multi-agent reinforcement learning
18. Chen X, Han Z, Zhang H, Xue G, Xiao Y, Bennis M (2017) Wireless resource scheduling in virtualized radio access networks using stochastic learning. IEEE Trans Mobile Comput 17(4):961–974
19. Daskalakis C, Deckelbaum A, Kim A (2015) Near-optimal no-regret algorithms for zero-sum games. Games Econ Behav 92:327–348
20. Fang F, Nguyen TH, Pickles R, Lam WY, Clements GR, An B, Singh A, Tambe M, Lemieux A (2016) Deploying paws: Field optimization of the protection assistant for wildlife security. In: Twenty-eighth IAAI conference
21. Fang F, Stone P, Tambe M (2015) When security games go green: designing defender strategies to prevent poaching and illegal fishing. In: Twenty-fourth international joint conference on artificial intelligence
22. Farina G, Bianchi T, Sandholm T (2019) Coarse correlation in extensive-form games. arXiv:1908.09893
23. Farina G, Celli A, Gatti N, Sandholm T (2018) Ex ante coordination and collusion in zero-sum multi-player extensive-form games. In: NeurIPS

24. Forges F (1986) An approach to communication equilibria. Econometrica 54:1375–1385
25. Forges F (2006) Correlated equilibrium in games with incomplete information revisited. Theory Decision 61(4):329–344
26. Kamenica E, Gentzkow M (2011) Bayesian persuasion. Am Econ Rev 101(6):2590–2615
27. Matej M, Martin S, Neil B, Viliam L, Dustin M, Nolan B, Trevor D, Kevin W, Michael J, Michael B (2017) Deepstack: expert-level artificial intelligence in heads-up no-limit poker. Science 356(6337):aam6960
28. Moulin H, Vial J-P (1978) Strategically zero-sum games: the class of games whose completely mixed equilibria cannot be improved upon. Int J Game Theory 7(3):201–221
29. Nash J (1951) Non-cooperative games. Ann Math 54:286–295
30. Píbil R, Lisỳ V, Kiekintveld C, Bošanskỳ B, Pěchouček M (2012) Game theoretic model of strategic honeypot selection in computer networks. In: International conference on decision and game theory for security, pp 201–220. Springer, Berlin
31. Schlenker A, Xu H, Guirguis M, Kiekintveld C, Sinha A, Tambe M, Sonya SY, Balderas D, Dunstatter N (2017) Don't bury your head in warnings: a game-theoretic approach for intelligent allocation of cyber-security alerts. In: IJCAI, pp 381–387
32. Milind T (2011) Security and game theory: algorithms, deployed systems, lessons learned. Cambridge University Press, Cambridge
33. von Stengel B (1996) Efficient computation of behavior strategies. Games Econ Behav 14(2):220–246
34. Zinkevich M, Bowling M, Johanson M, Piccione C (2007) Regret minimization in games with incomplete information. In: Proceedings of the annual conference on neural information processing systems (NIPS)