2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

Learning of equilibria and misperceptions in hypergames with perfect observations

Bahman Gharesifard Jorge Cortés

Abstract— **This paper studies the learning of equilibria in adversarial situations when players may have misperceptions about the game they are involved in with their opponents. We use the concept of high-level hypergames to model these scenarios. By drawing connections with the theory of ordinal potential games, we establish that players in a hypergame can individually learn their perceived equilibria using any improving adjustment scheme. We investigate how players can incorporate the information gained from observing the opponents' actions by updating different levels of her perception. We introduce high-level perception updating algorithms for resolving inconsistencies in perception using self-blaming or opponent-blaming strategies. Finally, we establish that when all players are rational and have perfect observation about past outcomes, repeated play converges to an equilibrium.**

I. Introduction

In adversarial situations, imperfect or incomplete information may lead to misperception about the opponent's true intentions. Imperfect information refers to the fact that players may only partially observe the actions taken by other players. Incomplete information refers to the fact that the true payoffs of the opponents may be only partially known to the players. In this paper, we deal with a special class of games of incomplete information called hypergames. Our goal is to study learning of equilibria and analyze the dynamics of the repeated play of hypergames, when players are rational, perfectly observe the actions taken by other players, and use their perceptions to select their actions.

Literature review: The notion of hypergame goes back to [1] and is mostly used in the context of conflict analysis [2], [3]. The benefit of using hypergames lies in the capability to model the perceptions of individual players. Hypergames are particularly useful in scenarios where players are absolutely certain about their opponents' perceptions, while these certainties may be mutually inconsistent. In a hypergame, players can have different levels of perception about their opponents' game, in the sense that they might have perceptions about the opponents' preferences or about what the opponents think about their preferences and so on.

In game theory, learning typically refers to the synthesis and analysis of equilibria, used introspectively by a player presuming that the opponent is using a certain strategy. Players are, instead, playing a game when they individually make decisions about their next action based on their perception of the game state and the strategy they have chosen to follow. The literature on learning is vast (we refer the interested readers to [4], [5] and references therein). In most of the existing learning methods, it is assumed that complete information is available when players learn the equilibria. These equilibria can be different from the outcome of actually playing the game if players have misperceptions about the payoffs of the opponents or when the opponents use a strategy different from the one used to learn the equilibria. Learning has also been studied in the framework of Bayesian games, see [5], [6], where games of incomplete information are studied as games of imperfect information. To our knowledge, with a few exceptions [7], [8], learning strategies and their convergence have not been formally studied in the framework of hypergames. In particular, we are interested in characterizing the properties that make a strategy successfully converge to equilibria, and the features enjoyed by these equilibria.

Statement of contributions: Our contributions pertain to the learning of equilibria when players reason introspectively about the hypergame and to the convergence to equilibria when players play repeatedly, observe the opponents' actions, and update their perceptions. Regarding learning, we show that players can learn the equilibria using any improving adjustment scheme, i.e., any strategy in which players take a feasible action if it improves their payoff. Our technical approach is based on studying the graph-theoretic properties of H-digraphs, a concept that captures the stability properties of hypergames. Specifically, we show that the H-digraph associated to the perceived game of each player contains no weak improvement cycle. This observation draws an interesting analogy with ordinal potential games, and plays an instrumental role in deriving the contributions on convergence to equilibria. Regarding the repeated play of hypergames, we introduce the `high-order perception update algorithm`, which prescribes how players employ the information obtained by observing the opponents' actions to update their perceptions. We show that players may run into inconsistencies in their perceptions and, based on their understanding of the opponents, can use self-blaming or opponent-blaming strategies to make them consistent. We demonstrate that the repeated play of the hypergame defines a dynamical system in the space of outcomes which converges to an equilibrium if players are rational, are able to perfectly observe past outcomes, and use the `high-order perception update algorithm`.

II. Preliminaries

We denote the set of real numbers by $\mathbb{R}$. We denote by $\mathbb{R}_{\geq k}$ and $\mathbb{Z}_{\geq k}$ the set of real numbers and positive integers greater than or equal to $k \in \mathbb{R}$, respectively. We denote by $\mathbf{I}_{n\times n}$ the identity matrix in $\mathbb{R}^{n\times n}$, $n \in \mathbb{Z}_{\geq 1}$. A nonempty

Bahman Gharesifard and Jorge Cortés are with the Department of Mechanical and Aerospace Engineering, University of California San Diego, {bgharesifard,cortes}@ucsd.edu

978-1-4577-0079-8/11/$26.00 ©2011 AACC

set X along with a preorder $\succeq$, i.e., a reflexive and transitive binary relation, is called a *directed set* if for every pair of elements in X there exists an upper bound with respect to the preorder. A *string* σ on X is a finite sequence of elements in X. The *length* of σ is the number of elements in σ.

A. Basic graph notions

A directed graph, or simply digraph, G is a pair (V, E), where V is a finite set, called the vertex set, and $E \subseteq V \times V$, called the edge set. Given $(u, v) \in E$, u is an *in-neighbor* of v and v is an *out-neighbor* of u. A *directed path* in a digraph, or in short path, is an ordered sequence of vertices so that any two consecutive vertices are an edge of the digraph. A vertex u is reachable from v if there exists a path starting at v and ending at u. A *cycle* in a digraph is a directed path that starts and ends at the same vertex and has no other repeated vertex. A digraph without any cycle is *acyclic*.

B. Ordinal potential games and strategic paths

A *(finite) game* [9], [4] is a triplet $\mathbf{G} = (V, \mathbf{S}_{\text{outcome}}, \mathbf{P})$ with the following elements: V is a set of $n \in \mathbb{Z}_{\geq 1}$ players, $\mathbf{S}_{\text{outcome}} = S_1 \times \ldots \times S_n$ is the outcome set with cardinality $N = |\mathbf{S}_{\text{outcome}}| \in \mathbb{Z}_{\geq 1}$, where S_i is a finite set of actions available to player $v_i \in V$, and $\mathbf{P} = (P_1, \ldots, P_n)$, with $P_i = (x_1, \ldots, x_N)^T \in \mathbf{S}_p$, the *preference vector* of player v_i, $i \in \{1, \ldots, n\}$. Here, $\mathbf{S}_p \subset \mathbf{S}_{\text{outcome}}^N$ denotes the set of all elements of $\mathbf{S}_{\text{outcome}}^N$ with pairwise different entries. We denote by π_i the natural projection of $\mathbf{S}_{\text{outcome}}$ onto the strategy set S_i of the ith player. We also use π_{-i} to denote the natural projection of $\mathbf{S}_{\text{outcome}}$ onto $S_{-i} = S_1 \times S_2 \times \cdots \times \hat{S}_i \times \cdots \times S_n$, where the hat notation denotes that S_i is excluded from the product. Note that, for each $i \in \{1, \ldots, n\}$, the outcome set $\mathbf{S}_{\text{outcome}}$ is a directed set under the preorder $\succeq_{P_i}$ induced by the preference vector P_i of player v_i as follows: $x \succeq_{P_i} y$ iff x has a lower entry index that y in P_i.

A *strategic path* in $\mathbf{S}_{\text{outcome}}$ is a sequence of outcomes $\mathfrak{S} = (x_1, x_2, \ldots)$, with $x_j \in \mathbf{S}_{\text{outcome}}$, such that, for each $j \in \mathbb{Z}_{\geq 1}$, there exists $i(j) \in \{1, \ldots, n\}$ with $v_{i(j)} \in V$, $\pi_{i(j)}(x_j) \neq \pi_{i(j)}(x_{j+1})$, and $\pi_{-i(j)}(x_j) = \pi_{-i(j)}(x_{j+1})$. A strategic path $\mathfrak{S} = (x_1, x_2, \ldots)$ is *nondeteriorating* if $x_{j+1} \succeq_{P_{i(j)}} x_j$, for all $j \in \mathbb{Z}_{\geq 1}$ and is a *better reply path* if $x_{j+1} \succ_{P_{i(j)}} x_j$, for all $j \in \mathbb{Z}_{\geq 1}$. A finite strategic path $\mathfrak{S} = (x_1, x_2, \ldots, x_m, x_1)$, $m \in \mathbb{Z}_{\geq 1}$, is called a *weak improvement cycle* if it is nondeteriorating and $x_{j+1} \succ_{P_{i(j)}} x_j$ for some $j \in \{1, \ldots, m-1\}$.

For later use, we also recall the notion of ordinal potential games [10]. A game $\mathbf{G} = (V, \mathbf{S}_{\text{outcome}}, \mathbf{P})$ is called an *ordinal potential game* if there exists a real-valued function $\mathcal{P} : \mathbf{S}_{\text{outcome}} \to \mathbb{R}$ such that for all $v_i \in V$ and $a_i, b_i \in S_i$, we have $(a_i, a_{-i}) \succ_{P_i} (b_i, a_{-i})$ if and only if $\mathcal{P}(a_i, a_{-i}) > \mathcal{P}(b_i, a_{-i})$. The function $\mathcal{P}$ is called the *ordinal potential function* for $\mathbf{G}$. One can establish [10] that $\mathbf{G}$ is ordinal potential iff it does not have any weak improvement cycle.

III. Hypergame theory

In this section, we review the basic notions of hypergame theory [3], [11], [1]. Our exposition follows [12]. Some well-known examples of the application of hypergame analysis include the Normandy invasion and the Cuban missile crisis, see [3]. A 0-level hypergame is simply a finite game. A 1-*level hypergame* with n players is a set $H^1 = \{\mathbf{G}_1, \ldots, \mathbf{G}_n\}$, where $\mathbf{G}_i = (V, (\mathbf{S}_{\text{outcome}})_i, \mathbf{P}_i)$, for $i \in \{1, \ldots, n\}$, is the subjective finite game of player $v_i \in V$, and V is a set of n players; $(\mathbf{S}_{\text{outcome}})_i = S_{1i} \times \ldots \times S_{ni}$, with S_{ji} the finite set of strategies available to v_j, as perceived by v_i; $\mathbf{P}_i = (P_{1i}, \ldots, P_{ni})$, with P_{ji} the preference vector of v_j, as perceived by v_i.

In a 1-level hypergame, each player $v_i \in V$ plays the game $\mathbf{G}_i$ with the perception that she is playing a game with complete information, which is not necessarily true. This is in sharp contrast with Bayesian games [5], [6], where a 'nature' player determines, according to some probability distribution a priori known to all players, the preferences of each one. The definition of a 1-level hypergame can be extended to higher-level hypergames, where some of the players have access to additional information that allow them to form perceptions about other players' perceptions, other players' perceptions about them, and so on. One can, inductively, extend the definition of 1-level hypergame as follows: a *k-level hypergame* with n players, $k \geq 1$, is a set $H^k = \{H_1^{k_1}, \ldots, H_n^{k_n}\}$, where $k_i \leq k-1$ and at least one k_i is equal to $k-1$. The hypergame H^k is called *homogeneous* if $k_i = k-1$ for all $i \in \{1, \ldots, n\}$.

A. Equilibria and stability

Here, we discuss the notions of equilibria and stability. Consider a k-level hypergame H^k between players $\{A_1, \ldots, A_n\}$ with outcome set $\mathbf{S}_{\text{outcome}}$. Without loss of generality and for simplicity, we assume that H^k is homogeneous. For $x \in \mathbf{S}_{\text{outcome}}$, we denote by $\mathbf{S}_{\text{outcome}}|_{\pi_{A_i}(x)}$ the set of outcomes $y \in \mathbf{S}_{\text{outcome}}$ such that $\pi_{A_i}(y) = \pi_{A_i}(x)$. For a string σ of length k on the set $\{A_1, \ldots, A_n\}$, let $\mathrm{P}_{A_i\sigma}$, $i \in \{1, \ldots, n\}$, denote the preferences of A_i as perceived by σ in H^k. For instance, $\mathrm{P}_{A_1A_2A_1}$ corresponds to what player A_1 perceives that player A_2 thinks about player A_1's preferences in a 2-level hypergame H^2. We use the preference vectors $(\mathrm{P}_{A_1\sigma}, \ldots, \mathrm{P}_{A_n\sigma})$ to denote the 0-level hypergame H_σ^0, often referred to as the *subjective* hypergame perceived by σ. With a slight abuse of notation, $\succeq_{A_i\sigma}$ denotes the binary relation $\succeq_{\mathrm{P}_{A_i\sigma}}$ on $\mathbf{S}_{\text{outcome}}$ induced by $\mathrm{P}_{A_i\sigma}$.

Given two distinct outcomes $x, y \in \mathbf{S}_{\text{outcome}}$, y is an *improvement* from x for player A_i perceived by σ in H_σ^0 if and only if $\pi_{A_{-i}}(y) = \pi_{A_{-i}}(x)$ and $y \succ_{A_i\sigma} x$. An outcome $x \in \mathbf{S}_{\text{outcome}}$ is *rational* for player A_i in H_σ^0 if there exists no improvement from x for this player. Finally, $x \in \mathbf{S}_{\text{outcome}}$ is *sequentially rational* for A_i in H_σ^0 if and only if for each improvement y from x for A_i in H_σ^0 there exists $z \in \mathbf{S}_{\text{outcome}}$ which *sanctions* y, i.e., $\pi_i(z) = \pi_i(y)$ and $x \succ_{A_i\sigma} z$ such that for all $j \in \{1, \ldots, n\}$ with $j \neq i$, either $\pi_j(z) = \pi_j(y)$, or the outcome $z_{A_j} \in \mathbf{S}_{\text{outcome}}|_{\pi_i(y)}$, where $\pi_j(z_{A_j}) = \pi_j(z)$ and $\pi_l(z_{A_j}) = \pi_l(y)$, for all $l \in \{1, \ldots, n\}$, $l \neq j$, is an improvement from y in $\mathbf{S}_{\text{outcome}}|_{\pi_i(y)}$ for A_j. By this definition, any sanction against the improvement y from x is due to actions taken by some players $\{A_{j_1} \ldots A_{j_l}\}$, where $\{j_1, \ldots, j_l\} \subset \{1, \ldots, n\}$, $j_p \neq i$, for all $p \in \{1, \ldots, l\}$. A rational outcome is also sequentially rational.

A player is rational if she only takes actions associated to sanction-free improvements. It turns out that all 0-level hypergames have at least one sequentially rational outcome [3], [11]. An outcome $x \in \mathbf{S}_{\text{outcome}}$ is *unstable* for A_i, perceived by σ, in H^0_σ, $i \in \{1, \ldots, n\}$ if it is not sequentially rational and is an *equilibrium of* H^0_σ if it is sequentially rational for all players A_i, $i \in \{1, \ldots, n\}$, with respect to H^0_σ. Note that more than one equilibrium might exist. An outcome is called an equilibrium of H^k if it is sequentially rational in all H^0_σ, where $\sigma = A_i A_i \ldots A_i$, $i \in \{1, \ldots, n\}$ is a string of length k on $\{A_1, \ldots, A_n\}$. One can similarly define the notion of equilibrium for any intermediate level $H^{k_1}_\eta$, where $k_1 < k$ and η is sequence of length at most $k-1$ on $\{A_1, \ldots, A_n\}$. For brevity, we sometimes omit the wording 'with respect to H^0_σ' and 'perceived by σ' when it is clear from the context.

B. H-digraphs

The notion of H-digraph [12], generalized here to n players, contains the information about the possible improvements from an outcome to another outcome, the equilibria, and the sanctions. Consider a homogeneous k-level hypergame H^k between players $\{A_1, \ldots, A_n\}$. Given σ and $i \in \{1, \ldots, n\}$, we assign to each $x \in \mathbf{S}_{\text{outcome}}$ a positive number $\text{rank}(x, \mathrm{P}_{A_i\sigma}) \in \mathbb{R}_{>0}$, called *rank*, such that, for each $\mathbf{S}_{\text{outcome}} \ni y \neq x$, we have $\text{rank}(y, \mathrm{P}_{A_i\sigma}) > \text{rank}(x, \mathrm{P}_{A_i\sigma})$ if and only if $x \succ_{A_i\sigma} y$. The n-dimensional digraph $\mathcal{G}_{H^0_\sigma} = (\mathbf{S}_{\text{outcome}}, \mathcal{E}_{H^0_\sigma})$ is the *H-digraph* associated to the 0-level hypergame H^0_σ, where each vertex $x \in \mathbf{S}_{\text{outcome}}$ is labeled with $(\text{rank}(x, \mathrm{P}_{A_1\sigma}), \ldots, \text{rank}(x, \mathrm{P}_{A_n\sigma}))$, and (x, y) belongs to $\mathcal{E}_{H^0_\sigma}$ iff $\pi_i(x) \neq \pi_i(y)$, $i \in \{1, \ldots, n\}$, $\pi_{-i}(x) = \pi_{-i}(y)$, and there exists a perceived improvement y from x for player A_i in H^0_σ for which there exists no sanction of players A_{-i}, perceived by σ.

IV. Acyclic structure of H-digraphs

We study the structure of H-digraphs and examine the implications on the equilibria of hypergames. This allows us to draw an interesting analogy with ordinal potential games. The following definitions adapt the notions of nondeteriorating paths and weak improvement cycles for hypergames.

Definition 4.1: (Nondeteriorating paths and weak improvement cycles in subjective hypergames): A strategic path $\mathfrak{S} = (x_1, x_2, \ldots)$ in $\mathbf{S}_{\text{outcome}}$ is *nondeteriorating* for H^0_σ if $(x_j, x_{j+1}) \in \mathcal{E}_{H^0_\sigma}$, for all $j \in \mathbb{Z}_{\geq 1}$. A finite strategic path $\mathfrak{S} = (x_1, x_2, \ldots, x_m, x_1)$, $m \in \mathbb{Z}_{\geq 1}$, is a *weak improvement cycle* for H^0_σ if it is nondeteriorating and $x_{j+1} \succ_{A_i\sigma} x_j$ for some $j \in \{1, \ldots, m-1\}$ and $i \in \{1, \ldots, n\}$.

An *improving adjustment scheme* in H^0_σ is any method that, given an initial outcome $x_1 \in \mathbf{S}_{\text{outcome}}$, generates a nondeteriorating strategic path $\mathfrak{S} = (x_1, x_2, \ldots)$. A best-response scheme is a special case of this notion, see [13] for more details. Next, we present our first result.

Theorem 4.2: (Subjective hypergames with two players contain no weak improvement cycle): Consider a k-level hypergame H^k between players A_1 and A_2. Let H^0_σ be a 0-level subjective hypergame perceived by σ, a string of length at most k on $\{A_1, A_2\}$. Then H^0_σ contains no weak improvement cycle.

Proof: We reason by contradiction. Suppose $\mathfrak{S} = (x_1, x_2, x_3, \ldots, x_p, x_1)$ is a weak improvement cycle for H^0_σ. Without loss of generality, we assume that players take alternate turns to take actions along the path. In other words, for $1 \leq j \leq p-2$, if $\pi_{A_1}(x_j) \neq \pi_{A_1}(x_{j+1})$ (resp. $\pi_{A_2}(x_j) \neq \pi_{A_2}(x_{j+1})$), then $\pi_{A_2}(x_{j+1}) \neq \pi_{A_2}(x_{j+2})$ (resp. $\pi_{A_1}(x_{j+1}) \neq \pi_{A_1}(x_{j+2})$). Our assumption is justified by the fact that, if $\pi_{A_1}(x_j) \neq \pi_{A_1}(x_{j+1}) \neq \pi_{A_1}(x_{j+2})$, then x_{j+2} is a perceived improvement from x_j for player A_1 and thus the outcome x_{j+1} can be removed from the path $\mathfrak{S}$, which still would correspond to a weak improvement cycle for H^0_σ. Note that, in particular, our assumption implies $p \in 2\mathbb{Z}_{\geq 2}$.

Suppose A_2 is the first player to take an action, i.e., $\pi_{A_2}(x_1) \neq \pi_{A_2}(x_2)$ (the reasoning for the case when the first player is A_1 is analogous). Since $\mathfrak{S}$ is a weak improvement cycle, $x_2 \succeq_{A_2\sigma} x_1$. Moreover, since $\pi_{A_1}(x_2) \neq \pi_{A_1}(x_3)$, we have that $x_3 \succeq_{A_1\sigma} x_2$. As a result, we deduce that $x_3 \succeq_{A_2\sigma} x_1$; otherwise, A_2's perceived improvement x_2 from x_1 is not sanction-free. With a similar argument, one can deduce that, for $j \in \{1, \ldots, \frac{p-2}{2}\}$,

(i) $x_{2j}, x_{2j+1} \succeq_{A_2\sigma} x_{2j-1}$;
(ii) $x_{2j+1}, x_{2j+2} \succeq_{A_1\sigma} x_{2j}$;
(iii) $x_p, x_1 \succeq_{A_2\sigma} x_{p-1}$ and $x_1, x_2 \succeq_{A_1\sigma} x_p$.

Since $\mathfrak{S}$ is an improvement cycle, there must exist at least one $l \in \{1, \ldots, p-1\}$ such that either $x_{l+1} \succ_{A_1\sigma} x_l$ with $\pi_{A_1}(x_l) \neq \pi_{A_1}(x_{l+1})$ or $x_{l+1} \succ_{A_2\sigma} x_l$ with $\pi_{A_2}(x_l) \neq \pi_{A_2}(x_{l+1})$. Assume we are in the first case, i.e., l is odd, (the argument for the second case, i.e., l is even, is the same). Then, using (ii), one concludes that $x_p \succ_{A_1\sigma} x_2$, which contradicts (iii). ■

We generalize the result above to the case of an arbitrary number of players using an inductive procedure.

Theorem 4.3: (Subjective hypergames contain no weak improvement cycle): Consider a k-level hypergame H^k with n players $\{A_1, \ldots, A_n\}$. Then none of the subjective 0-level hypergame H^0_σ, where σ is a string of length at most k on $\{A_1, \ldots, A_n\}$, contains a weak improvement cycle.

Proof: Let $\{A_1, A_2, \ldots, A_n\}$ be a set of $n \in \mathbb{Z}_{\geq 3}$ players and $H^0_\sigma = (\mathrm{P}_{A_1\sigma}, \ldots, \mathrm{P}_{A_n\sigma})$. We denote by $\mathbf{S}^{\text{reachable}}_{\text{outcome}}|_{\pi_{A_i}(x)} \subseteq \mathbf{S}_{\text{outcome}}|_{\pi_{A_i}(x)}$ the set of all outcomes in $\mathbf{S}_{\text{outcome}}|_{\pi_{A_i}(x)}$ which can be reached from $x \in \mathbf{S}_{\text{outcome}}$ in the digraph $\mathcal{G}_{H^0_\sigma}$ by a directed path in $\mathbf{S}_{\text{outcome}}|_{\pi_{A_i}(x)}$.

Consider a strategic path $\mathfrak{S} = (x_1, x_2, \ldots, x_m)$ for H^0_σ, with $m \in \mathbb{Z}_{\geq 1}$. Similar to the two players' case, without loss of generality, we assume that if player A_i takes an action that changes the outcome from x_j to x_{j+1}, then player A_l takes an action next, where $i, l \in \{1, \ldots, n\}$ and $i \neq l$.

Without loss of generality, we assume that player A_2 is the first player that takes an action that changes the outcome from x_1 to x_2. Note that $\mathfrak{S}$ induces a sequence of outcomes, denoted by $\mathfrak{S} \supseteq \mathfrak{S}|_{\pi_{A_2}} = \{x_1, x_2, \ldots, x_{m'}\}$, $m' \in \mathbb{Z}_{\geq 1}$, with $\pi_{A_2}(x_{j'}) \neq \pi_{A_2}(x_{j'+1})$, for all $j' \in \{1, \ldots, m'-1\}$.

We proceed with the proof by induction on n. By Theorem 4.2, the claim holds for $n = 2$. Suppose that the claim holds for any subjective 0-level hypergame with $n = N-1$ players, and let us show that it also holds when $n = N$. If we fix the action of one player, say A_i, then players A_{-i}

are playing a 0-level hypergame with $N-1$ players, which contains no weak improvement cycle by the assumption of induction. Thus it is enough to show that $\mathfrak{S}|_{\pi_{A_2}}$ cannot be a weak improvement cycle.

We claim that $x_{j'+1} \succeq_{A_2\sigma} x_{j'}$, for all $x_{j'}, x_{j'+1} \in \mathfrak{S}|_{\pi_{A_2}}$. For any outcome $x_l \in \mathfrak{S} \cap \mathbf{S}_{\text{outcome}}^{\text{reachable}}|_{\pi_{A_2}(x_{j'})}$, we have that $x_l \succeq_{A_2\sigma} x_j$. In particular, there exists an outcome $x_l^* \in \mathfrak{S} \cap \mathbf{S}_{\text{outcome}}^{\text{reachable}}|_{\pi_{A_2}(x_{j'})}$ such that $\pi_{A_{-2}}(x_l^*) = \pi_{A_{-2}}(x_{j'+1})$, $x_{j'+1} \succeq_{A_2\sigma} x_l^*$, and $x_l^* \succeq_{A_2\sigma} x_j$. Thus we conclude that $x_{j'+1} \succeq_{A_2\sigma} x_{j'}$, as claimed. By a similar argument, one can conclude that $x_1 \succeq_{A_2\sigma} x_{m'}$. But since $\mathfrak{S}|_{\pi_{A_2}}$ is a weakly improvement cycle, there exists at least two consecutive outcomes $x, y \in \mathfrak{S}|_{\pi_{A_2}}$, such that player A_2 is perceived to strictly prefer y to x. But this gives a contradiction, with an argument similar to the one in Theorem 4.2. ■

Remark 4.4: (Connection to ordinal potential games): Suppose $\mathcal{G}_{H_\sigma^0}$ is the H-digraph associated to a subjective hypergame H_σ^0 with n players $V = \{A_1, \ldots, A_n\}$ and let $\mathbf{G} = (V, \mathbf{S}_{\text{outcome}}, \mathbf{P})$ be the game defined by $x_2 \succeq_{P_i} x_1$ with $\pi_{-i}(x_1) = \pi_{-i}(x_2)$, $\pi_i(x_1) \neq \pi_i(x_2)$ for $v_i \in V$ if and only if $(x_1, x_2) \in \mathcal{G}_{H_\sigma^0}$. Then, $\mathbf{G}$ is an ordinal potential game since, by Theorem 4.3, the digraph $\mathcal{G}_{H_\sigma^0}$ is acyclic. •

We state an immediate consequence of Theorem 4.3 which captures how each individual player learns the equilibrium of her subjective hypergame.

Corollary 4.5: (Learning in subjective hypergames): Any improving adjustment scheme used for learning the hypergame H_σ^0 will converge to an equilibrium.

We finish this section by revealing some interesting structural properties of the H-digraphs as a corollary of Theorem 4.3. In particular, we present a necessary condition for a digraph to be associated to a subjective hypergame.

Corollary 4.6: (Necessary conditions for a digraph to be an H-digraph): Suppose G is a H-digraph associated to a subjective hypergame with player set V and outcome set $\mathbf{S}_{\text{outcome}}$. Then all the eigenvalues of $\text{Adj}(G) + \mathbf{I}_{n\times n}$, where $\text{Adj}(G)$ is the adjacency matrix associated to G, are equal to 1. If $\mathcal{S}_{\text{sub-hyp}}(\mathbf{S}_{\text{outcome}})$ is the space of all subjective hypergames of a player in V, with outcome set $\mathbf{S}_{\text{outcome}}$, then $|\mathcal{S}_{\text{sub-hyp}}(\mathbf{S}_{\text{outcome}})| \leq \mathbf{N}_{\text{acyclic}}(|\mathbf{S}_{\text{outcome}}|)$, where

$$\mathbf{N}_{\text{acyclic}}(n) = \sum_{i=1}^{n} (-1)^{i-1} \binom{n}{i} 2^{i(n-i)} \mathbf{N}_{\text{acyclic}}(n-i).$$

Proof: By Theorem 4.3, G is acyclic. Thus all eigenvalues of $\text{Adj}(G)+\mathbf{I}_{n\times n}$ are equal to 1, see [14]. The second part follows from a combinatorial result on the number of acyclic digraphs with labeled vertices [15, Corollary 2]. ■

Note that there are, however, acyclic digraphs which cannot be associated to a hypergame. In fact, when each player action set has at least cardinality 2, the inequality in Corollary 4.6 is strict. We demonstrate this by an example.

Example 4.7: (An acyclic digraph which is not an H-digraph): Consider the following digraph,

$$\begin{array}{cc} x_1 \longrightarrow x_2 \\ \\ x_3 \qquad x_4 \end{array}$$

Suppose this digraph can be associated to H_σ^0, where σ is a string on $\{A_1, A_2\}$, with $\mathbf{S}_{\text{outcome}} = \{x_1, x_2, x_3, x_4\}$ and two players, where A_1 plays rows and A_2 plays columns. First, note that $x_4 \succ_{A_1\sigma} x_2$, since otherwise, there must exist an edge (x_4, x_2), because the improvement x_2 from x_4 is sanction free for A_1. Since there exists a perceived sanction by A_2 against the improvement x_4 from x_2 for A_1, one also concludes that $x_2 \succ_{A_1\sigma} x_3$. Next, notice that $x_3 \succ_{A_1\sigma} x_1$, since otherwise, there exists a sanction-free improvement x_1 from x_3 for A_1. If $x_4 \succ_{A_2\sigma} x_3$, then the improvement x_4 from x_3 is perceived as sanction-free for A_2, since the outcome x_4 is perceived as rational for A_1. This is a contradiction with the nonexistence of the edge (x_3, x_4). Conversely, suppose $x_3 \succ_{A_2\sigma} x_4$. Then the perceived improvement x_3 from x_1 for A_1 is sanction free, since x_3 is rational for A_2, but this is also a contradiction to the nonexistence of the edge (x_1, x_3). Thus this digraph cannot be associated to a subjective hypergame H_σ^0 with two players A_1 and A_2. •

V. Convergence of the Repeated Play of Hypergames under Perfect Observation

In this section, we study the equilibria of k-level, n-player homogeneous hypergames when players are playing repeatedly and perfectly observe the actions taken by the opponents. We assume that players take their actions sequentially, one after each other. This assumption is consistent with the notions of sequential rationality and improving adjustment schemes. In our scenario, each player updates her perception about the opponents' preferences using the information contained in the actions taken by them.

A. `High-order perception update algorithm`

Here, we introduce the `high-order perception update algorithm`. Before getting into the specific details, we make a few observations to motivate the discussion. Consider a 3-level hypergame between A_1 and A_2. Notice that $\mathrm{P}_{A_1A_2A_2A_1}$ in $H_{A_2A_2A_1}^0$ and $\mathrm{P}_{A_1A_1A_2A_1}$ in $H_{A_1A_2A_1}^0$ are the same vector. However, the stability properties of an outcome for A_1 can actually be different in $H_{A_2A_2A_1}^0$ and $H_{A_1A_2A_1}^0$. Any change in one preference vector also affects the other one. This observation motivates the following definitions. For a string σ of length k on $\{A_1, \ldots, A_n\}$, let $\widetilde{\sigma}$ the string that results from replacing any repeated sequence of adjacent characters in σ by only one such character. We refer to $\widetilde{\sigma}$ as the associated *reduced string*. For instance, the associated reduced string of $\sigma = A_1A_2A_2A_1$ is $\widetilde{\sigma} = A_1A_2A_1$. Two strings σ and η are *equivalent*, denoted $\sigma \cong \eta$, if and only if $\widetilde{\sigma} = \widetilde{\eta}$. Note that the preference vectors associated to two equivalent strings are equal.

Let us now discuss the elements that come into play when updating players' perceptions. Suppose A_{i^*}, $i^* \in \{1, \ldots, n\}$, takes an action that changes the outcome from x_1 to x_2. If A_i, $i \neq i^*$, believes that A_{i^*} is rational, then she will update her perception to reflect the fact that A_{i^*} prefers x_2 to x_1. In a k-level hypergame, A_i updates her perception $\mathrm{P}_{A_{i^*}^k A_i}$, with $A_{i^*}^k = A_{i^*} \dots A_{i^*}$ (k copies), to reflect

$$x_2 \succ_{A_{i^*}^k A_i} x_1. \tag{1}$$

The choice of $A_{i^*}^k A_i$ is determined by the fact that $A_{i^*}^k A_i$ is the unique string in $\{\sigma A_{i^*} A_i \mid \sigma \text{ is a string of length } k-1 \text{ on } \{A_1, \dots, A_n\}\}$ which is equivalent to $A_{i^*} A_i$.

The update (1) does not guarantee that x_1 will be perceived as unstable for A_{i^*} by A_i, which is the additional piece of information contained in the action taken by A_{i^*}. Since players are rational, A_i needs to adjust her perception to make it compatible with this observation. In this case, it would be enough for A_i to update her perception such that the improvement x_2 from x_1 is perceived as sanction free. In general, we do not infer any information on the stability of x_2, since, according to the H-digraph, A_{i^*} can take an action that changes the outcome to any sanction-free improvement, even though this outcome might be unstable for her.

The above discussion motivates the design of the `high-order perception update algorithm` in Table I. This strategy allows each player to incorporate the observations about the opponent's actions into her preferences and remove the inconsistencies.

Name:	`high-order perception update algorithm`
Goal:	Incorporate observations into perceptions of A_i in a consistent manner
Input:	Action of A_{i^*}, observed by A_i, changing the outcome from $x_1 \in \mathbf{S}_{\text{outcome}}$ to $x_2 \in \mathbf{S}_{\text{outcome}}$, with $i \in \{1, \dots, n\} \setminus \{i^*\}$
Output:	Updated $H^0_{\sigma A_i}$, σ sequence of length $k-1$

update $\mathrm{P}_{A_{i^*}^k A_i}$ with $x_2 \succ_{A_{i^*}^k A_i} x_1$
if improvement x_2 form x_1 is sanction-free for A_{i^*} in $H^0_{A_{i^*}^{k-1} A_i}$ **then**
  no update is required
else
  for each sanction $z \in \mathbf{S}_{\text{outcome}}$ against x_2 **do**
    if self-blaming **then**
      update perception with $z \succ_{A_{i^*}^k A_i} x_1$
    else if opponent-blaming **then**
      set $\{A_{j_1} \dots A_{j_\ell}\} \subset \{A_1, \dots, A_n\}$ with $\pi_{j_p}(z) \neq \pi_{j_p}(x_2)$ and $j_p \neq i^*$, for $p \in \{1, \dots, \ell\}$
      select non-empty $\mathcal{M} \subset \{1, \dots, \ell\}$
      for each $p \in \mathcal{M}$ **do**
        let z_{j_p} be the unique outcome given by $\pi_{j_p}(z_{j_p}) = \pi_{j_p}(z)$ and $\pi_l(z_{j_p}) = \pi_l(x_2)$, for all $l \in \{1, \dots, n\} \setminus \{j_p\}$
        update $\mathrm{P}_{A_{j_p} A_{i^*}^{k-1} A_i}$ with $x_2 \succ_{A_{j_p} A_{i^*}^{k-1} A_i} z_{j_p}$
      end for
    end if
  end for
end if

TABLE I

THE HIGH-ORDER PERCEPTION UPDATE ALGORITHM

Remark 5.1 (Preference update mechanism): There are many methods to change a player's preference to make (1) hold. Let us describe the formal requirements that such methods should satisfy. Let σ be a string on $\{A_1, \dots, A_n\}$ and let $\mathrm{P}_{\sigma A_{i^*} A_i}$ be a preference vector of player A_{i^*}, perceived by player A_i. We define the *observation set* $\mathcal{O}_{\sigma A_{i^*} A_i}$ as the set of all binary relations observed by player A_i about player A_{i^*}. We say that the preference vector $\mathrm{P}_{\sigma A_{i^*} A_i}$ is compatible with an observation set $\mathcal{O}_{\sigma A_{i^*} A_i}$ if all the binary relations in $\mathcal{O}_{\sigma A_{i^*} A_i}$ hold with the order $\succeq_{\sigma A_{i^*} A_i}$. A *preference update mechanism* compatible with an observation set $\mathcal{O}_{\sigma A_{i^*} A_i}$ is a map $\Psi_{\mathcal{O}_{\sigma A_{i^*} A_i}} : \mathbf{S}_p \to \mathbf{S}_p$ such that $\Psi_{\mathcal{O}_{\sigma A_{i^*} A_i}}(\mathrm{P})$ is compatible with $\mathcal{O}_{\sigma A_{i^*} A_i}$ for $\mathrm{P} \in \mathbf{S}_{\text{outcome}}$. Throughout this section, when we say a player updates her preferences with some binary relation, we mean that this player adds this binary relation to her associated observation set and uses a preference update mechanism to generate a preference vector compatible with the observation set. Particular instances of such updating mechanisms are given for instance in [12]. •

The next example shows how players can use the `high-order perception update algorithm`.

Example 5.2: (Algorithm execution): Let A_1 and A_2 play a 2-level hypergame with $\mathbf{S}_{\text{outcome}} = \{x_1, x_2, x_3, x_4\}$ and

$$\begin{aligned}
\mathrm{P}_{A_1 A_1 A_1} &= (x_3, x_2, x_4, x_1)^T, \\
\mathrm{P}_{A_2 A_1 A_1} &= (x_1, x_4, x_3, x_2)^T = \mathrm{P}_{A_2 A_2 A_1}, \\
\mathrm{P}_{A_1 A_2 A_1} &= (x_2, x_4, x_1, x_3)^T, \\
\mathrm{P}_{A_2 A_2 A_2} &= (x_4, x_2, x_3, x_1)^T, \\
\mathrm{P}_{A_1 A_2 A_2} &= (x_4, x_3, x_2, x_1)^T = \mathrm{P}_{A_1 A_1 A_2}, \\
\mathrm{P}_{A_2 A_1 A_2} &= (x_1, x_4, x_2, x_3)^T.
\end{aligned}$$

Let x_3 be the initial outcome. Observe that x_3 is perceived as unstable for A_2 in $H^0_{A_2 A_2}$. Thus A_2 takes an action which changes the outcome to x_4 and this is observed by A_1. This action is aligned with A_1's perception about the ranking of x_3 and x_4 since $x_4 \succ_{A_2 A_1 A_1} x_3$, but not with the fact that A_1 perceives x_3 as sequentially rational for A_2 in $H^0_{A_2 A_1}$. According to the `high-order perception update algorithm`, A_1 can either conclude that the inconsistency is due to her misperception about A_2's true game; or conclude that A_2 has an incorrect perception about A_1's true game. In the former case, A_1 updates her perception using the partial order $x_2 \succ_{A_2 A_2 A_1} x_3$, for example, by swapping the order of these outcomes in $\mathrm{P}_{A_2 A_2 A_1}$; in the latter case, A_1 updates her perception with the partial ordering $x_4 \succ_{A_1 A_2 A_1} x_2$. •

B. Analysis of the `high-order perception update algorithm`

Here we analyze the properties of the `high-order perception update algorithm` and its impact on the stability of the outcomes. The following result is an immediate consequence of the definition of the algorithm.

Lemma 5.3: (`high-order perception update algorithm` removes the inconsistencies): If player A_{i^*}, $i^* \in \{1, \dots, n\}$, takes an action that changes the outcome of a k-level, n-player hypergame from x_1 to x_2 and A_i, $i \in \{1, \dots, n\} \setminus \{i^*\}$, updates her perception about A_i^* using the `high-order perception update algorithm`, then x_1 is perceived as unstable for A_{i^*} in $H^0_{A_{i^*}^{k-1} A_i}$.

The next results describe in detail those parts of the perception of each player that are affected by the updates of the algorithm. We begin with the case where the inconsistencies are perceived to arise because of a player's own misperception, and consequently, she only uses `6:` in Table I.

Proposition 5.4: (Self-blaming the inconsistencies): Suppose A_{i^*}, $i^* \in \{1,\dots,n\}$, takes an action that changes the outcome of a k-level, n-player hypergame from x_1 to x_2 and A_i, $i \in \{1,\dots,n\} \setminus \{i^*\}$, updates her perception about A_i^* using the `high-order perception update algorithm` resolving the inconsistencies via `6:` Table I. Then, the only subjective hypergames whose stability can change are H^0_σ with $\sigma \cong A_{i^*}A_i$ or $\sigma = A_i^k$.

Proof: By assumptions and by the `high-order perception update algorithm`, A_i only updates the ranks of the outcomes in $\mathrm{P}_{A^k_{i^*}A_i}$. Such updates affect the stability of the outcomes in H^0_σ, σ a string of length k, where $\sigma \cong A_{i^*}A_i$ or $\sigma = A_i^k$. This is because for any such σ, the ordering in the preference vector $\mathrm{P}_{A_{i^*}\sigma}$ changes as $\mathrm{P}_{A^k_{i^*}A_i}$ changes. Moreover, the changes in the rankings in $\mathrm{P}_{A^k_{i^*}A_i}$ do not have any impact on the stability of outcomes in H^0_η, $\eta \neq \sigma$ a string on $\{A_1,\dots,A_n\}$ of length k, since $A_l\eta$ is not equivalent to $A^k_{i^*}A_i$ for any $l \in \{1,\dots,n\}$. ■

The next result characterizes the case in which a player believes that inconsistencies arise because of the misperception of other players about her game, and consequently, only uses `8:` Table I. The proof is similar to the one for the previous result and is omitted here.

Proposition 5.5: (Opponent-blaming for the inconsistencies): Suppose A_{i^*}, $i^* \in \{1,\dots,n\}$, takes an action that changes the outcome of a k-level, n-player hypergame from x_1 to x_2 and A_i, $i \in \{1,\dots,n\}\setminus\{i^*\}$ updates her perception about A_i^* using the `high-order perception update algorithm` resolving the inconsistencies via only `8:` Table I, by assuming that A_{i^*}'s perception about players $\{A_{j_1},\dots,A_{j_\ell}\} \subset \{A_1,\dots,A_n\}$, where $\pi_{j_p}(z) \neq \pi_{j_p}(x_2)$ and $j_p \neq i^*$, for all $p \in \{1,\dots,\ell\}$, is incorrect. Then, the only subjective hypergames whose stability can change are H^0_σ with $\sigma \cong A_{i^*}A_i$, $\sigma \cong A_{j_p}A_{i^*}A_i$, or $\sigma = A_i^k$.

Our final result states that the repeated play of a hypergame in which all players use the `high-order perception update algorithm` to update their perceptions is guaranteed to converge to an equilibrium.

Theorem 5.6: (Convergence to an equilibrium under the `high-order perception update algorithm`): Consider a k-level, n-player hypergame. Suppose all players are rational, can fully observe the actions of their opponents, play sequentially, and update their perceptions according to the `high-order perception update algorithm`. Then the repeated play of this hypergame converges to an equilibrium.

Proof: Since we assumed that the outcome set is a finite set and all players are playing a rational game, each player A_i, $i \in \{1,\dots,n\}$, will eventually fully learn the preferences of her opponents' in $\mathrm{P}_{A^k_{i^*}A_i}$, for all $i^* \in \{1,\dots,n\} \setminus \{i\}$, unless the evolution of the hypergame finishes in an equilibrium. If the evolution of $\mathrm{P}_{A^k_{i^*}A_i}$ converges to $\mathrm{P}_{A^{k+1}_{i^*}}$, for all $i \in \{1,\dots,n\}$, then any equilibrium of the 0-level subjective hypergame $H^0_{A_i^k}$ of A_i is also, by definition, an equilibrium of H^k. Since players are rational and play sequentially, the repeated play of the hypergame is an improving adjustment scheme for the subjective hypergame of A_i, which by Corollary 4.5 converges to an equilibrium. ■

VI. Conclusions

We have studied the learning of equilibria in hypergames. By drawing a connection with ordinal potential games, we have shown that the H-digraph associated to a finite subjective hypergame contains no weak improvement cycle. This property has allowed us to show that players can use any improving adjustment scheme to learn the equilibria of their subjective hypergames. We have designed the `high-order perception update algorithm` that allows players to consistently update their perceptions with the information contained in their observations and using either self-blaming or opponent-blaming strategies. We have characterized the properties of the `high-order perception update algorithm` and, more importantly, we have proved that if players are rational, have perfect observation about the past outcomes of the game, and use this information to update their perceptions about the opponents' preferences, any repeated play of the hypergame will converge to an equilibrium. Future work will explore the extension of the results of the paper to situations with imperfect observations, and their application to deception.

Acknowledgments

This work was supported by Award FA9550-10-1-0499.

References

[1] P. G. Bennett, "Towards a theory of hypergames," *OMEGA*, vol. 5, no. 6, pp. 749–751, 1977.

[2] P. G. Bennett, "Hypergames, the development of an approach to modeling conflicts," *Futures*, vol. 12, pp. 489–507, 1980.

[3] M. Wang, K. W. Hipel, and N. M. Fraser, "Solution concepts in hypergames," *Applied Mathematics and Computation*, vol. 34, pp. 147–171, 1989.

[4] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA: MIT Press, 1991.

[5] H. P. Young, *Strategic Learning and its Limits*. Oxford, UK: Oxford University Press, 2004.

[6] L. Smith and P. Sorensen, "Pathological outcomes of observational learning," *Econometrica*, vol. 68, no. 2, pp. 371–398, 2000.

[7] B. Chaib-dara, "Hypergame analysis in multiagent environments," in *AAAI Spring Symposia*, (Stanford, CA), pp. 147–149, Mar. 2001.

[8] R. Vane and P. Lehner, "Using hypergames to increase planned payoff and reduce risk," *Autonomous Agents and Multi-Agent Systems*, vol. 5, pp. 365–380, 2002.

[9] T. Basar and G. Oldser, *Dynamic Noncooperative Game Theory*. Academic Press, 1982.

[10] M. Voorneveld and H. Norde, "A characterization of ordinal potential games," *Games and Economic Behavior*, vol. 19, pp. 235–242, 1997.

[11] N. M. Fraser and K. W. Hipel, *Conflict Analysis: Models and Resolutions*. Systems Science and Engineering, Oxford, UK: Elsevier, 1984.

[12] B. Gharesifard and J. Cortés, "Evolution of the perception about the opponent in hypergames," in *IEEE Conf. on Decision and Control*, (Atlanta, Georgia), pp. 1076–1081, Dec. 2010.

[13] Y. M. Ermoliev and S. D. Flaam, "Repeated play of potential games," *Cybernetics and Systems Analysis*, vol. 38, pp. 355–367, 2002.

[14] B. D. McKay, F. E. Oggier, G. F. Royle, N. J. A. Sloane, I. M. Wanless, and H. S. Wilf, "Acyclic digraphs and eigenvalues of $(0, 1)$-matrices," *Journal of Integer Sequences*, vol. 7, 2004. Article 04.3.3.

[15] R. W. Robinson, "Counting unlabeled acyclic digraphs," in *New Directions in Graph Theory*, (Academic press, NY), pp. 239–273, 1973.