Safe Pareto improvements for delegated game playing

Anonymous Author(s)
Submission Id: 558

ABSTRACT

A set of players delegate playing a game to a set of representatives, one for each player. We imagine that each player trusts their respective representative's strategic abilities. Thus, we might imagine that per default, the original players would simply instruct the representatives to play the original game as best as they can. In this paper, we ask: are there safe Pareto improvements on this default way of giving instructions? That is, we imagine that the original players can coordinate to tell their representatives to only consider some subset of the available strategies and to assign utilities to outcomes differently than the original players. Then can the original players do this in such a way that the payoff is guaranteed to be weakly higher than under the default instructions for all the original players? In particular, can they Pareto-improve without probabilistic assumptions about how the representatives play games? In this paper, we give some examples of safe Pareto improvements. We further prove that the notion of safe Pareto improvements is closely related to a notion of outcome correspondence between games. We also show that under some specific assumptions about how the representatives play games, finding safe Pareto improvements is NP-complete.

KEYWORDS

program equilibrium, delegation, bargaining, Pareto efficiency, smart contracts

ACM Reference Format:
Anonymous Author(s). 2021. Safe Pareto improvements for delegated game playing. In *Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), London, UK, May 3–7, 2021*, IFAAMAS, 17 pages.

1 INTRODUCTION

Between Aliceland and Bobbesia lies a sparsely populated desert. Until recently, neither of the two countries had any interest in the desert. However, geologists have recently discovered that it contains large oil reserves. Now, both Aliceland and Bobbesia would like to annex the desert. They worry, however, about a military conflict that would ensue if both countries insist on annexing.

Table 1 models this strategic situation as a normal-form game. The strategy DM (short for "Demand with Military") denotes a military invasion of the desert, demanding annexation. If both countries send their military with such an aggressive mission, the countries fight a devastating war. The strategy RM (for "Refrain with Mility") denotes yielding the territory to the other country, but building defenses to prevent an invasion of one's original territories. Alternatively, the countries can choose to not raise a military force at all, while potentially still demanding control of the desert by sending only its leader (DL, short for "Demand with Leader"). In this case, if both countries demand the desert, war does not ensue. Finally, they could neither demand nor build up a mility (RL). If one of the two countries has their military ready and the other does not, the militarized country will know and will be able to invade the other country. In game-theoretical terms, militarizing therefore strictly dominates not militarizing.

Instead of making the decision directly, the parliaments of Aliceland and Bobbesia appoint special commissions for making this strategic decision, led by Alice and Bob, respectively. The parliaments can instruct these *representatives* in various ways. They can explicitly tell them what to do – for example, Aliceland could directly tell Alice to play DM. However, we imagine that the parliaments trust the commissions' judgments more than they trust their own and hence they might prefer to give an instruction of the type, "make whatever demands you think are best for our country". They might not know what that will entail, i.e., how the commissions decide what demands to make given that instruction. However, – based on their trust in their representatives – they might still believe that this leads to better outcomes than giving an explicit instruction.

We will also imagine these instructions are (or at least can be) given publicly and that the commissions are bound (as if by a contract) to follow these instructions. In particular, we imagine that the two commissions can see each other's instructions. Thus, in instructing their commissions, the countries play a game with bilateral precommitment. When instructed to play a game as best as they can, we imagine that the commissions play that game in the usual way, i.e., without further abilities to credibly commit or to instruct subcommittees and so forth.

It may seem that without having their parliaments ponder equilibrium selection, Aliceland and Bobbesia cannot do better than leave the game to their representatives. Unfortunately, in this default equilibrium, war is still a possibility. Even the brilliant strategists Alice and Bob may not always be able to resolve the difficult equilibrium selection problem to the same pure Nash equilibrium.

In the literature on commitment devices and in particular the literature on program equilibrium, important ideas have been proposed for avoiding such bad outcomes. Imagine for a moment that Alice and Bob will play a Prisoner's Dilemma (rather than the Demand Game of Table 1). Then the default of (Defect, Defect) can be Pareto-improved upon as follows. Both original players (Aliceland and Bobbesia) use the following instruction for their representatives: "If the opponent's instruction is equal to this instruction, Cooperate; otherwise Defect." [11, 16, 31] Then it is a Nash equilibrium for both players to use this instruction. In this equilibrium,

Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), U. Endriss, A. Nowé, F. Dignum, A. Lomuscio (eds.), May 3–7, 2021, London, UK.

		Player 2			
		DM	RM	DL	RL
Player 1	DM	$-5, -5$	2, 0	$5, -5$	$5, -5$
	RM	0, 2	1, 1	$5, -5$	$5, -5$
	DL	$-5, 5$	$-5, 5$	1, 1	2, 0
	RL	$-5, 5$	$-5, 5$	0, 2	1, 1

Table 1: The Demand Game

		Player 2's rep.	
		DL	RL
Player 1's rep.	DL	$-5, -5$	2, 0
	RL	0, 2	1, 1

Table 2: A safe Pareto improvement for the Demand Game

(Cooperate, Cooperate) is played and it is thus Pareto-optimal and Pareto-better than the default.[1]

In cases like the Demand Game, it is more difficult to apply this approach to improve upon the default of simply delegating the choice. Of course, if one could calculate the expected utility of submitting the default instructions, then one could similarly commit the representatives to follow some (joint) mix over the Pareto-optimal outcomes ((RM,DM), (DM, RM), (RM, RM), (DL, DL), etc.) that Pareto improve on the default expected utilities. However, we will assume that the original players are unable or unwilling to form probabilistic expectations about how the representatives play the Demand Game, i.e., about what would happen with the default instructions. If this is the case, then this type of Pareto-improvement on the default is unappealing.

The goal of this paper is to show and analyze how even without forming probabilistic beliefs about the representatives, the original players can Pareto-improve on the default equilibrium. We will call such improvements *safe Pareto improvements* (SPIs). We here briefly give an example in the Demand Game.

The key idea is for the original players to instruct the representatives to select only from $\{\text{DL}, \text{RL}\}$, i.e., to not raise a military. Further, they tell them to disvalue the conflict outcome (DL, DL) as they would disvalue the original conflict outcome of war in the default equilibrium. Overall, this means telling them to play the game of Table 2. Importantly, Aliceland's instruction to play that game must be conditional on Bobbesia also instructing their commission to play that game, and vice versa. Otherwise, one of the countries could profit from deviating by instructing their representative to always play DM or RM (or to play by the original utility function).

The game of Table 2 is isomorphic to the DM-RM part of the original Demand Game of Table 1. Of course, the original players know neither how the original Demand Game nor the game of Table 2 will be played by the representatives. However, since these games are isomorphic, one should arguably expect them to be played in isomorphic ways. For example, one should expect that (RM, DM) would be played in the original game if and only if (RL, DL) would be played in the modified game. However, the conflict outcome (DM, DM) is replaced in the new game with the outcome (DL, DL). This outcome is harmless (Pareto-optimal) for the original players.

[1]The case where commitment is unilateral – i.e., one player commits and another responds to that commitment – corresponds to a *Stackelberg* model [34]. Algorithms for computing solutions to such games [5, 14, 21, 35] have been of great interest to the multiagent systems community, particularly due their use in *security games* [18, 23, 24, 29, 30]. However, in this paper, the focus is on games in which *all* players can simultaneously commit themselves, to having a representative play on their behalf.

Our paper generalizes this idea to arbitrary normal-form games and is organized as follows. In Section 2, we introduce some notation for games and multivalued functions that we will use throughout this paper. In Section 3 we introduce the setting of delegated game playing for this paper and motivate the search for safe Pareto improvements in more detail. Section 4 provides further justification and relates our setting to the literature on program equilibrium. In Section 5, we introduce a notion of outcome correspondence between games which expresses the original players' beliefs about similarities between how the representatives might play different games. For example, in our example, the Demand Game of Table 1 is (arguably) equivalent to the game of Table 2 in that the representatives (arguably) would play (DM, DM) in the original game if and only if they play (DL, DL) in the new game, and so forth. We also show some basic results (reflexivity, transitivity, etc.) about the outcome correspondence relation on games. In Section 6 we show that the notion of outcome correspondence is central to deriving safe Pareto improvements. In particular, we show that some game Γ^s is a safe Pareto improvement on some other game Γ if and only if there is a Pareto-improving outcome correspondence relation between Γ^s and Γ.

To derive safe Pareto improvements, we need to make some assumptions about outcome correspondence, i.e., about which games are played in similar ways by representatives. We give two very weak assumptions of this type in Section 7. The first is that the representatives play isomorphic games isomorphically. The second is that the representatives' play is invariant under the removal of strictly dominated strategies. For example, we assume that in the Demand Game the representatives only play DM and RM. Moreover we assume that we could remove DL and RL from the game and the representatives would still play the same strategies as in the original Demand Game with certainty. Our safe Pareto improvement for the Demand Game can be proven using these assumptions. Section 8 shows that determining whether there exists a safe Pareto improvement based on these assumptions is NP-complete. Section 9 considers a different setting in which we allow the original players to let the representatives choose from newly constructed strategies whose corresponding outcomes map arbitrarily onto feasible payoff vectors from the original game. In this new setting, finding safe Pareto-improvements can be done in polynomial time. We conclude by discussing the problem of selecting between different safe Pareto improvements on a given game (Section 10) and giving some ideas for directions for future work (Section 11).

2 PRELIMINARIES

2.1 Games

We here recall some basic definitions from game theory. An *n-player game* is a tuple $(A, \mathbf{u})$ of a set $A = A_1 \times ... \times A_n$ of *(pure) strategy profiles* (or outcomes) and a function $\mathbf{u}\colon A \rightarrow \mathbb{R}^n$ that assigns to each outcome a utility for each player. We will also write games as $(A_1, ..., A_n, u_1, ..., u_n)$. We say that $a_i \in A_i$ *strictly dominates* $a_i' \in A_i$ if for all $a_{-i} \in A_{-i}$, $u_i(a_i, a_{-i}) > u_i(a_i', a_{-i})$. For any given game $\Gamma = (A, \mathbf{u})$, we will call any game $\Gamma' = (A', \mathbf{u}')$ a *subset game* of Γ if $A_i' \subseteq A_i$ for $i = 1, ..., n$. Note that a subset game may assign different utilities to outcomes than the original game. For any set of strategies S, we denote by $\Gamma - S \coloneqq ((A_1 - S) \times ...., (A_n - S), \mathbf{u})$ the game that arises from Γ by removing the strategies S for all players.

We say that some utility vector $\mathbf{y} \in \mathbb{R}^n$ is a Pareto-improvement on (or is Pareto-better than) $\mathbf{y}' \in \mathbb{R}^n$ if $y_i \geq y_i'$ for $i = 1, ..., n$. We will also denote this by $\mathbf{y} \geq \mathbf{y}'$. Note that, contrary to convention, we allow $\mathbf{y} = \mathbf{y}'$. Whenever we require one of the inequalities to be strict, we will say that $\mathbf{y}$ is a strict Pareto improvement on $\mathbf{y}'$. In a given game, we will also say that an outcome $\mathbf{a}$ is a Pareto-improvement on another outcome $\mathbf{a}'$ if $\mathbf{u}(\mathbf{a}) \geq \mathbf{u}(\mathbf{a}')$. We say that $\mathbf{y}$ is Pareto-optimal or Pareto-efficient relative to some $S \subset \mathbb{R}^n$ if there is no element of S that strictly Pareto-dominates $\mathbf{y}$.

The Demand Game of Table 1 is an example of a game that we will use throughout this paper. As noted earlier, DM and RM strictly dominate DL and RL. The game of Table 2 is a subset game of the Demand Game.

2.2 Multivalued functions

or sets M and N, a *multi-valued function* $\Phi\colon M \multimap N$ is a function which maps each element $m \in M$ to a set $\Phi(m) \subseteq N$. For a subset $Q \subseteq M$, we define $\Phi(Q) \coloneqq \bigcup_{m \in Q} \Phi(m)$. Note that $\Phi(Q) \subseteq N$ and that $\Phi(\emptyset) = \emptyset$. For any set M, we define the identity function $\mathrm{id}_M\colon M \multimap M\colon m \rightarrow \{m\}$. Also, for two sets M, N, we define $\mathrm{all}_{M,N}\colon M \multimap N\colon m \mapsto N$. We define the inverse $\Phi^{-1}\colon N \multimap M\colon n \mapsto \{m \in M \mid n \in \Phi(m)\}$. Note that $\Phi^{-1}(\emptyset) = \emptyset$ for any multi-valued function Φ. For sets M, N, Q and functions $\Phi\colon M \multimap N$, $\Psi\colon N \multimap Q$, we define the composite $\Psi \circ \Phi\colon M \multimap Q\colon m \mapsto \Psi(\Phi(m))$. As with regular functions, composition of multi-valued functions is associative.

We say that $\Phi\colon M \multimap N$ is *single-valued* if $|\Phi(m)| = 1$ for all $m \in M$. Whenever a multi-valued function is single-valued, we can apply many of the terms for regular functions. For example, we will take injectivity, surjectivity, and bijectivity for single-valued functions to have the usual meaning. We will never apply these notions to non-single-valued functions.

3 DELEGATION AND SAFE PARETO IMPROVEMENTS

We consider a setting in which a given game Γ is played through what we will call *representatives*. For example, the representatives could be humans whose behavior is determined or incentivized by some contract à la the principal–agent literature [15].

We imagine that one way in which the representatives can be instructed is to in turn play a subset game $\Gamma^s = (A_1^s \subseteq A_1, ..., A_n^s \subseteq A_n, \mathbf{u}^s)$ of the original game, *without necessarily specifying a strategy or algorithm for solving such a game*. We emphasize, again, that $\mathbf{u}^s$ is allowed to be a vector of entirely different utility functions. For any subset game Γ^s, we denote by $\Pi(\Gamma^s)$ the outcome that arises if the representatives play the subset game Γ^s of Γ. Because in many games, it is not clear what the right choice is, the original players might might be uncertain about $\Pi(\Gamma^s)$ for many games Γ^s. We will therefore model each $\Pi(\Gamma^s)$ as a random variable.

The original players trust their representatives to the extent that we take $\Pi(\Gamma)$ to be a default way for the game to played for any Γ. For example, in the Game of Chicken, it is not clear what the right action is. Thus, if one can simply delegate the decision to someone with more relevant expertise, that is the first option one would consider.

We are interested in whether and how the original players can jointly Pareto-improve on the default. Of course, one option is to compute the expected utilities in the default ($\mathbb{E}\left[\mathbf{u}(\Pi(\Gamma))\right]$) and then let the representatives play a distribution over outcomes whose expected utility exceeds that default expected utility. However, this is unrealistic if Γ is a complex game with multiple Nash equilibria. For one, the precise point of delegation is that the original players are unable or unwilling to properly evaluate Γ. Second, there is no widely agreed upon, universal procedure for selecting an action in the face of equilibrium selection problems.

We address this problem in a typical way. Essentially, we require of any attempted improvement that it incurs no regret in the worst-case. That is, we are interested in subset games Γ^s that are Pareto improvements *with certainty* under weak and purely qualitative assumptions about Π.

Definition 3.1. Let Γ^s be a subset game of Γ. We say Γ^s is a *safe Pareto improvement (SPI)* on Γ if $\mathbf{u}(\Pi(\Gamma^s)) \geq \mathbf{u}(\Pi(\Gamma))$ with certainty. We say that Γ^s is a *strict* SPI if furthermore, there is a Player i s.t. $u_i(\Pi(\Gamma^s)) > u_i(\Pi(\Gamma^s))$ with positive probability.

4 PROGRAM EQUILIBRIUM

So far, we have been vague about the details of the strategic situation that the original players face in instructing their representatives. From what set of actions can they choose? How can they jointly let the representatives play some new subset game Γ^s? Are SPIs Nash equilibria of the meta game played by the representatives? In this section, we will describe one way in which one can fill this gap. We will thereby also discuss the concept of program equilibrium [3, 7, 8, 19, 31], which is the strand of work most closely related to the present paper. This section is essential to understanding why SPIs are relevant. However, the remaining technical content of this paper does not rely on this section and the main ideas presented here are straightforward from previous work. We therefore leave formal details to Appendix A.

For any game Γ, the program equilibrium literature considers the meta game in which each player chooses from a set of computer programs. Each program for Player i takes as input any vector containing everyone else's chosen program and returns an action from A_i. As a real-life example, we could imagine that each player i selects a contract that can refer to everyone's else contract. The meta game can be analyzed like any other game. Its Nash equilibria are called *program equilibria*. Since the chosen programs can refer to one another, the program equilibria can implement payoffs not implemented by any Nash equilibria of Γ itself. For example, in the Prisoner's Dilemma, both players can submit a program that says, "If the opponent's chosen computer program is equal to this computer program, Cooperate; otherwise Defect." [11, 16, 31] This is a Nash equilibrium which implements mutual cooperation.

In the setting for our paper, we similarly imagine that players can choose contracts that bind the representatives. However, our representatives are themselves able to competently choose an action for player i in any given game Γ^s, without the original player having to describe how this choice is to be made. For example, a human representative is capable of maximizing a utility function u^s if a payment proportional to the value of u^s is contractually guaranteed. For such representatives, we can similarly imagine that the original players submit computer programs. However, in addition to the elementary instructions used in a typical computer program, we now additionally allow Player i to use an instruction "Play $\Pi_i(\Gamma^s)$". To jointly let the representatives play, e.g., the SPI Γ^s of Table 2 on the Demand Game of Table 1, the representatives can both use an instruction that says, "If the opponent's chosen contract is analogous to this one, play $\Pi_i(\Gamma^s)$; otherwise play DM". More generally, we can prove the following result; see Appendix A for details.

Theorem 4.1. *Let Γ be a game and Γ^s be an SPI of Γ. Now consider a program game on Γ, where each player i can choose from a set of computer programs. In addition to the normal kind of instructions, we allow the use of the command "play $\Pi_i(\Gamma')$" for any subset game Γ' of Γ. Finally, assume that $\Pi(\Gamma)$ guarantees each player i at least that player's minimax utility (a.k.a. threat point) in Γ. Then $\Pi(\Gamma^s)$ is played in a program equilibrium, i.e., in a Nash equilibrium of the program game.*

As an alternative to having the original players choose contracts separately, we could imagine the use of jointly signed contracts which only come into effect once signed by all players [cf. 12, 17]. Also compare a paper by Raub [25], which we discuss in Appendix B.

5 OUTCOME CORRESPONDENCE BETWEEN GAMES

In this section, we introduce a very general notion of outcome correspondence, which we will use throughout this paper to construct safe Pareto improvements.

Definition 5.1. Consider two games $\Gamma = (A_1, ..., A_n, \mathbf{u})$ and $\Gamma' = (A'_1, ..., A'_n, \mathbf{u}')$. We write $\Gamma \sim_\Phi \Gamma'$ for $\Phi\colon A \multimap A'$ if $\Pi(\Gamma') \in \Phi(\Pi(\Gamma))$ with certainty.

Note that aside from trivial relations like $\Gamma \sim_{\text{id}} \Gamma$, a relationship $\Gamma \sim_\Phi \Gamma'$ expresses a qualitative fact (or a piece of the players' knowledge) about Π, i.e., about how the representatives choose. They are therefore dependent on the specific representatives being used. In Section 7, we describe two general circumstances under which it seems plausible that $\Gamma \sim_\Phi \Gamma'$. For example, if two games Γ and Γ' are isomorphic, then one might expect $\Gamma \sim_\Phi \Gamma'$, where Φ is constructed from the n isomorphisms of the particular action spaces.

We now state some basic facts about the relation $\sim$, many of which we will use throughout this paper.

Lemma 5.2. *Let $\Gamma = (A, \mathbf{u}), \Gamma' = (A', \mathbf{u}'), \hat{\Gamma} = (\hat{A}, \hat{\mathbf{u}})$ and $\Phi, \Xi\colon A \multimap A'$, $\Psi\colon A' \multimap \hat{A}$.*

1. *Reflexivity: $\Gamma \sim_{\text{id}_A} \Gamma$, where $\text{id}_A\colon A \multimap A\colon \mathbf{a} \mapsto \{\mathbf{a}\}$.*
2. *Symmetry: If $\Gamma \sim_\Phi \Gamma'$, then $\Gamma' \sim_{\Phi^{-1}} \Gamma$.*
3. *Transitivity: If $\Gamma \sim_\Phi \Gamma'$ and $\Gamma' \sim_\Psi \hat{\Gamma}$, then $\Gamma \sim_{\Psi \circ \Phi} \hat{\Gamma}$.*
4. *If $\Gamma \sim_\Phi \Gamma'$ and $\Phi(\mathbf{a}) \subseteq \Xi(\mathbf{a})$ for all $\mathbf{a} \in A$, then $\Gamma \sim_\Xi \Gamma'$.*
5. *$\Gamma \sim_{\text{all}_{A,A'}} \Gamma'$, where $\text{all}_{A,A'}\colon A \multimap A'\colon \mathbf{a} \mapsto A'$.*
6. *If $\Gamma \sim_\Phi \Gamma'$ and $\Phi(\mathbf{a}) = \emptyset$, then $\Pi(\Gamma) \neq \mathbf{a}$ with certainty.*

We prove these in Appendix C. Items 1–3 show that $\sim$ has properties resembling those of an equivalence relation. Note, however, that since $\sim$ is not a binary relationship, $\sim$ itself cannot be an equivalence relation in the usual sense. Item 4 states that we can make an outcome correspondence claim less precise and it will still hold true. Item 5 states that in the extreme, it is always $\Gamma \sim_{\text{all}_{A,A'}} \Gamma'$, where $\text{all}_{A,A'}$ is the trivial, maximally imprecise outcome correspondence function that confers no information. Item 6 shows that $\sim$ can be used to express the elimination of outcomes, i.e., the belief that a particular outcome (or strategy) will never occur.

6 SAFE PARETO IMPROVEMENTS THROUGH OUTCOME CORRESPONDENCE

We now show that as advertised, outcome correspondence is closely tied to safe Pareto improvements. The following theorem shows not only how outcome correspondences can be used to find (and prove) safe Pareto improvements. It also shows that any safe Pareto improvement requires an outcome correspondence relation with what we will call a *Pareto-improving* correspondence function.

Theorem 6.1. *Let $\Gamma = (A, \mathbf{u})$ be a game and $\Gamma^s = (A^s, \mathbf{u}^s)$ be a subset game of Γ. Then Γ^s is an SPI on Γ if and only if there is Φ such that $\Gamma \sim_\Phi \Gamma^s$ and for all $\mathbf{a} \in A$ it is for all $\mathbf{a}^s \in \Phi(\mathbf{a})$ the case that $\mathbf{u}(\mathbf{a}^s) \geq \mathbf{u}(\mathbf{a})$.*

Proof. $\Leftarrow$: By definition, it holds with certainty that $\Pi(\Gamma^s) \in \Phi(\Pi(\Gamma))$. Hence, for $i = 1, 2$, it holds with certainty that $u_i(\Pi(\Gamma^s)) \in u_i(\Phi(\Pi(\Gamma)))$. (Note that by Lemma 5.2.6, $\Phi(\Pi(\Gamma))$ and therefore $u_i(\Phi(\Pi(\Gamma)))$ is non-empty with certainty.) Hence, by assumption about Φ, with certainty, $u_i(\Pi(\Gamma^s)) \geq u_i(\Pi(\Gamma))$.

$\Rightarrow$: Assume that $u_i(\Pi(\Gamma)) \geq u_i(\Pi(\Gamma^s))$ with certainty for $i = 1, 2$. We define

$$\Phi\colon A \to A^s\colon \mathbf{a} \mapsto \left\{\mathbf{a}^s \in A^s \mid \mathbf{u}(\mathbf{a}^s) \geq \mathbf{u}(\mathbf{a})\right\}. \tag{1}$$

It is immediately obvious that Φ is Pareto-improving as required.

Then whenever $\Pi(\Gamma) = \mathbf{a}$ and $\Pi(\Gamma^s) = \mathbf{a}^s$ for any $\mathbf{a} \in A$ and $\mathbf{a}^s \in A^s$, it is (by assumption) with certainty $\mathbf{u}(\mathbf{a}^s) \geq \mathbf{u}(\mathbf{a})$. Thus, by definition of Φ, it holds that $\mathbf{a}^s \in \Phi(\mathbf{a})$. We can conclude that $\Gamma \sim_\Phi \Gamma^s$ as claimed. □

Note that the theorem concerns weak SPIs and therefore allows the case where with certainty $\mathbf{u}(\Pi(\Gamma)) = \mathbf{u}(\Pi(\Gamma^s))$. To show that some Γ^s is a *strict* SPI, we need additional information about which outcomes occur with positive probability.

We now illustrate how outcome correspondences can be used to derive the SPI for the Demand Game from the introduction as per Theorem 6.1. Of course, at this point we do not have any assumptions about when games are equivalent. We will introduce some in the following section. Nevertheless, we can already sketch the argument. Let Γ be the Demand Game of Table 1. First, it seems plausible that Γ is in some sense equivalent to Γ', where $\Gamma' = \Gamma - \{\text{DL}, \text{RL}\}$ is the game that results from removing DL and RL for both players from Γ. Again, strict dominance could be given as an argument. We can formalize this as $\Gamma \sim_\Phi \Gamma'$, where $\Phi(a_1, a_2) = \{(a_1, a_2)\}$ if $a_1, a_2 \in \{\text{DM}, \text{RM}\}$ and $\Phi(a_1, a_2) = \emptyset$ otherwise. In a second step, it seems plausible that $\Gamma' \sim_\Psi \Gamma^s$, where Γ^s is the game of Table 2 and Ψ is the isomorphism between Γ' and Γ^s. Finally, we can use transitivity to obtain $\Gamma \sim_{\Psi\circ\Phi} \Gamma^s$. To see that $\Psi \circ \Phi$ is Pareto-improving for the original utility functions of Γ, notice that Φ does not change utilities at all. Ψ maps the conflict outcome (DM, DM) onto the outcome (DL, DL), which is better for both original players. Other than that, Ψ, too, does not change the utilities. Hence, $\Psi \circ \Phi$ is Pareto-improving. By Theorem 6.1, Γ^s is therefore an SPI on Γ.

In principle, Theorem 6.1 does not hinge on $\Pi(\Gamma)$ and $\Pi(\Gamma^s)$ resulting from playing games. An analogous result holds for any random variables over A and A^s. In particular, this means that Theorem 6.1 applies also if the representatives receive other kinds of instructions (cf. Section 4). However, it seems hard to establish non-trivial outcome correspondences between $\Pi(\Gamma)$ and other types of instructions. Still, the use of more complicated instructions can be used to derive different kinds of SPIs. For example, if there are different game SPIs, then the original players could tell their representatives randomize between them in a coordinated way.

7 ASSUMPTIONS ABOUT OUTCOME CORRESPONDENCE

To make any claims about how the original players should play the meta-game, i.e., about what instructions they should submit, we have to make assumptions about how the representatives choose and (by Theorem 6.1) about outcome correspondence in particular. We here make two fairly weak assumptions.

The first is that the representatives play two isomorphic games in isomorphically.

Assumption 1. *Let $\Gamma = (A, \mathbf{u})$ and $\Gamma' = (A', \mathbf{u}')$ be two games such that there are single-valued bijections $\Phi_i : A_i \rightarrow A'_i$ for $i = 1, ..., n$ such that $\mathbf{u}(a_1, ..., a_n) = \mathbf{u}'(\Phi_1(a_1), ..., \Phi_n(a_n))$ for all $\mathbf{a} \in A$. Then $\Gamma \sim_\Phi \Gamma'$.*

Similar desiderata have been discussed in the context of equilibrium selection, e.g., by Harsanyi and Selten [10, Chapter 3.4]. In fact, they consider a generalization in which the utilities are allowed to be linear transformations of each other. Although this generalization is extremely plausible, we omit it here for simplicity.

One could criticize Assumption 1 by referring to focal points (introduced by Schelling [26, pp. 54–58]) as an example where context and labels of strategies matter. A possible response might be that in games where context plays a role, that context should be included as additional information and not be considered part of $(A, \mathbf{u})$. Assumption 1 would then either not apply to such games with (relevant) context or would require one to, in some way, translate the context along with the strategies. However, in this paper we will not formalize context, and assume that there is no decision-relevant context.

Assumption 2. *Let $\Gamma = (A, \mathbf{u})$ be an arbitrary n-player game where $A_1, ..., A_n$ are pairwise disjoint, and $\tilde{a}_i \in A_i$ be strictly dominated by some other strategy $\hat{a}_i \in A_i$. Then $\Gamma \sim_\Phi \Gamma - \{\tilde{a}_i\}$, where for all $a_{-i} \in A_{-i}$, $\Phi(\tilde{a}_i, a_{-i}) = \emptyset$ and $\Phi(a_i, a_{-i}) = \{(a_i, a_{-i})\}$ whenever $a_i \neq \tilde{a}_i$.*

Assumption 2 expresses that representatives should never play strictly dominated strategies. Moreover, it states that we can remove strictly dominated strategies from a game and the resulting game will be played in the same way by the representatives. For example, this implies that when evaluating a strategy a_i, the representatives do not take into account how many other strategies a_i strictly dominates. Assumption 2 also allows (via Transitivity of $\sim$ as per Lemma 5.2.3) the iterated removal of strictly dominated strategies. The notion that we can (iteratively) remove strictly dominated strategies is common in game theory [13, 20, 22, Section 2.9, Chapter 12] and has rarely been questioned. It is also implicit in the solution concept of Nash equilibrium – by iterated removal of strictly dominated strategies on some game Γ we never remove a strategy played in a Nash equilibrium of Γ. However, like the concept of Nash equilibrium, the elimination of strictly dominated strategies becomes implausible if the game is not played in the usual way. In particular, for Assumption 2 to hold, we will in most games Γ have to assume that the representatives cannot in turn make credible precommitments (or delegate to further subrepresentatives) or play the game iteratively [2].

We can now use Assumptions 1 and 2 to formally prove two SPIs: the trivial case of mutual cooperation in the Prisoner's Dilemma and the SPI of Table 2 for the Demand Game. Our proofs are in Appendix D.

Proposition (Example) 1. *Let Γ be the Prisoner's Dilemma and $\Gamma^s = (A_1^s, A_2^s, u_1^s, u_2^s)$ be any subset game of Γ with $A_1^s = A_2^s = \{\text{Cooperate}\}$. Then under Assumption 2, Γ^s is a strict SPI on Γ.*

Proposition (Example) 2. *Let Γ be the Demand Game of Table 1 and Γ^s be the subset game described in Table 2. Under Assumptions 1 and 2, Γ^s is an SPI on Γ. Further, if $P(\Pi(\Gamma)=(DM, DM)) > 0$, Γ^s is a strict safe Pareto improvement.*

8 COMPUTING SAFE PARETO IMPROVEMENTS

In this section, we ask how computationally costly it is for the original players to identify for a given game Γ a non-trivial SPI Γ^s. In particular, we ask whether a given game Γ has a non-trivial SPI that can be proved using only Assumptions 1 and 2, Transitivity (Lemma 5.2.3) and Theorem 6.1. Formally:

Definition 8.1. The *SPI decision problem* consists in deciding for any given Γ, whether there is a sequence of outcome correspondences $\Phi^1, ..., \Phi^k$ and a sequence of subset games $\Gamma^0 = \Gamma, \Gamma^1, ..., \Gamma^k$ of Γ s.t.:

(1) (Non-triviality:) If we fully reduce Γ^k and Γ using iterated strict dominance (Assumption 2), the two resulting games are not equal. (Of course, they are allowed to be isomorphic.)
(2) For $i = 1, ..., k$, $\Gamma^{i-1} \sim_{\Phi^i} \Gamma^i$ is valid by a single application of either Assumption 1 or Assumption 2.
(3) For all $\mathbf{a} \in A$, and whenever $\mathbf{a}^s \in (\Phi^k \circ \Phi^{k-1} \circ ... \circ \Phi^1)(\mathbf{a})$, it is the case that $u(\mathbf{a}^s) \geq \mathbf{u}(\mathbf{a})$.

For the *strict* SPI decision problem, we further require:

(4) There is a player i and an outcome $\mathbf{a}$ that survives iterated elimination of strictly dominated strategies from Γ s.t. $u_i((\Phi^k \circ \Phi^{k-1} \circ ... \circ \Phi^1)(\mathbf{a})) > u_i(\mathbf{a})$.

Theorem 8.2. *The (strict) SPI decision problem is NP-complete, even for 2-player games.*

Proposition 8.3. *For games Γ with $|A_1| + ... + |A_n| = m$ that can be reduced (via iterative application of Assumption 2) to a game Γ' with $|A'_1| + ... + |A'_n| = l$, the (strict) SPI decision problem can be solved in $O(m^l)$.*

We prove this in Appendix E. Besides being about a specific set of assumptions about $\sim$, note that this result also assumes that the game is represented explicitly in normal form as a payoff matrix. If we changed the game representation, this could affect the complexity of the SPI problem. Compare Gabarró et al. [9], who show how the related problem of deciding whether two games are isomorphic depends on the game representation. In fact, even reducing a game using strict dominance by pure strategies – which contributes only insignificantly to the complexity of the SPI problem for normal-form games – is difficult in some game representations [4, Section 6].

9 SAFE PARETO IMPROVEMENTS UNDER IMPROVED COORDINATION

In this section, we imagine that the players are able to simply invent new token strategies with new payoffs that arise from mixing existing feasible payoffs. To define this formally, we first define for any game $\Gamma = (A, \mathbf{u})$,

$$C(\Gamma) := \mathbf{u}(\Delta(A)) = \left\{ \sum_{\mathbf{a} \in A} p_{\mathbf{a}} \mathbf{u}(\mathbf{a}) \;\middle|\; \forall \mathbf{a} \in A : p_{\mathbf{a}} \in [0,1], \sum_{\mathbf{a} \in A} p_{\mathbf{a}} = 1 \right\}$$

to be the set of feasible coordinated payoff vectors of Γ, which is exactly the convex closure of $\mathbf{u}(A)$, i.e., of the deterministically achievable utilities of the original game.

For any game Γ, we then imagine that in addition to subset games, the players can let the representatives play a *perfect-coordination token game* $(A^s, \mathbf{u}^s, \mathbf{u}^e)$, where for all i, $A_i^s \cap A_i = \emptyset$ and $u_i^s : A^s \to \mathbb{R}$ are arbitrary utility functions to be used by the representatives and $\mathbf{u}^e : A^s \to C(\Gamma)$ are the utilities that the original players assign to the token strategies.

The instruction $(A^s, \mathbf{u}^s, \mathbf{u}^e)$ lets the representatives play the game $(A^s, \mathbf{u}^s)$ as usual. However, the strategies A^s are imagined to be meaningless token strategies which do not resolve the given game Γ. Once some token strategies $\mathbf{a}^s$ are selected, these are translated into some probability distribution over A, i.e., over outcomes of the original game, thus giving rise to (expected) utilities $\mathbf{u}^e(\mathbf{a}^s) \in C(\Gamma)$. These distributions and thus utilities are specified by the original players. We here imagine in our definition of $C(\Gamma)$ that these distributions over A could require the representatives to correlate their choices for the original game for any given $\mathbf{a}^s$.

Definition 9.1. Let Γ be a game. A *perfect-coordination SPI* for Γ is a perfect-coordination token game $(A^s, \mathbf{u}^s, \mathbf{u}^e)$ for Γ s.t. $\mathbf{u}^e(\Pi(A^s, u^s)) \geq \mathbf{u}(\Pi(\Gamma))$ with certainty. We call $(A^s, \mathbf{u}^s, \mathbf{u}^e)$ a *strict* perfect-coordination SPI if there furthermore is a player i for whom $u_i^e(\Pi(A^s, u^s)) > u_i(\Pi(\Gamma))$ with positive probability.

As an example, imagine that Γ is just the DM-RM subset game of the Demand Game of Table 1. Then, intuitively, an SPI under improved coordination could consist of the original players telling the representatives, "Play as if you were playing the DM-RM subset game of the Demand Game, but whenever you find yourself playing (DM, DM), randomize [according to some given distribution] between the other (Pareto-optimal) outcomes instead". Formally, $A_1^s = \{\hat{D}, \hat{R}\}$, $A_2^s = \{\hat{D}, \hat{R}\}$ would then consist of tokenized versions of the original strategies. The utility functions u_1^s, u_2^s are then simply the same as in the original Demand Game except that they are applied to the token strategies. E.g., $\mathbf{u}^s(\hat{D}, \hat{R}) = (2, 0)$. The utilities for the original players remove the conflict outcome. For example, the original players might specify $\mathbf{u}^e(\hat{D}, \hat{D}) = (1, 1)$, representing that the representatives are supposed to play (RM,RM) in the $(\hat{D}, \hat{D})$ case. For all other outcomes $(\hat{a}_1, \hat{a}_2)$, it must be the case that $\mathbf{u}^e(\hat{a}_1, \hat{a}_2) = \mathbf{u}^s(\hat{a}_1, \hat{a}_2)$ because the other outcomes cannot be Pareto-improved upon. As with our earlier SPIs for the Demand Game, Assumption 1 implies that $\Gamma \sim_\Phi \Gamma^s$, where Φ maps the original conflict outcome (DM,DM) onto the Pareto-optimal $(\hat{D}, \hat{D})$.

Relative to the SPIs considered up until now, these new types of instructions put significant additional requirements on how the representatives interact. They now have to engage in a two-round process of first choosing and observing one another's token strategies and then playing the corresponding distribution over outcomes from the original game. Further, it must be the case that this additional coordination does not affect the payoffs of the original outcomes. For example, we may imagine this is not the case in the Game of Chicken. That is, we could imagine a Game of Chicken in which coordination is possible but that the rewards of the game change if the players (or representatives) coordinate. After all, the underlying story in the Game of Chicken is that the positive reward (admiration from peers) is attained precisely for accepting a grave risk.

With these more powerful ways to instruct representatives, we can now replace individual outcomes of the default game *ad libitum*. For example, in the reduced Demand Game, we singled out the outcome (DM,DM) as Pareto-suboptimal and replaced it by a Pareto-optimal outcome, while keeping all other outcomes the same. This allows us to construct SPIs in many more games then before.

Definition 9.2. The *strict full-coordination SPI decision problem* consists in deciding for any given Γ whether under Assumption 1 there is a perfect-coordination SPI Γ^s for Γ.

Lemma 9.3. *For a given n-player game Γ and payoff vector $\mathbf{y} \in \mathbb{R}^n$, it can be decided by linear programming and thus in polynomial time whether $\mathbf{y}$ is Pareto-optimal in $C(\Gamma)$.*

This is easy, but for completeness the linear program is given in Appendix F. Based on Lemma 9.3, Algorithm 1 is our algorithm for deciding whether there is a strict perfect-coordination SPI for a given game Γ.

It is easy to see that this algorithm runs in polynomial time (in the size of, e.g., the normal form representation of the game). It is also correct: if it returns True, simply replace the Pareto-suboptimal outcome while keeping all other outcomes the same; if it returns False, then all outcomes are Pareto-optimal within $C(\Gamma)$ and so

Algorithm 1: An algorithm for deciding the strict perfect-coordination SPI problem.

Data: Game Γ, set $\text{supp}(\Pi(\Gamma)$
for $\mathbf{a} \in \text{supp}(\Pi(\Gamma))$ **do**
  if $\mathbf{u}(\mathbf{a})$ *is Pareto-suboptimal within* $C(\Gamma)$ **then**
    Return True;
Return False;

there can be no strict SPI. We summarize this result in the following proposition.

Proposition 9.4. *Assuming* $\text{supp}(\Pi(\Gamma))$ *is known and that Assumption 1 holds, it can be decided in polynomial time whether there is a strict perfect-coordination SPI.*

From the problem of deciding whether there are strict SPIs under improved coordination at all, we move on to the question of what different perfect-coordination SPIs there are. In particular, one might ask what the cost is of only considering *safe* Pareto improvements relative to acting on a probability distribution over $\Pi(\Gamma)$ and the resulting expected utilities $\mathbb{E}\left[\mathbf{u}(\Pi(\Gamma))\right]$. We start with a lemma that directly provides a characterization. So far, all the perfect-coordination SPIs $(A^s, \mathbf{u}^s, \mathbf{u}^e)$ for a game $(A, \mathbf{u})$ we have discussed have consisted in letting the representatives play a game $(A^s, \mathbf{u}^s)$ that is isomorphic to the original game, but Pareto-improves (from the original players' perspectives, i.e., $\mathbf{u}^e$) at least one of the outcomes. It turns out that we can restrict attention to this very simple type of SPI under improved coordination.

Lemma 9.5. *Let* $\Gamma = (\{a_1^1, ..., a_1^{l_1}\}, ..., \{a_n^1, ..., a_n^{l_n}\}, \mathbf{u})$ *be any game. Let* Γ' *be a perfect-coordination SPI on* Γ. *Then we can define* $\mathbf{u}^e$ *with values in* $C(\Gamma)$ *such that under Assumption 1 the game*

$$\Gamma^s = \Big(\hat{A}_1 \coloneqq \{\hat{a}_1^1, ..., \hat{a}_1^{l_1}\}, ..., \hat{A}_n \coloneqq \{\hat{a}_n^1, ..., \hat{a}_n^{l_n}\}, \quad \hat{\mathbf{u}} \colon (\hat{a}_1^{i_1}, ..., \hat{a}_n^{i_n}) \mapsto \mathbf{u}(a_1^{i_1}, ..., a_n^{i_n}), \mathbf{u}^e\Big) \tag{2}$$

is also an SPI on Γ, *with*

$$\mathbb{E}\left[\mathbf{u}(\Pi(\Gamma^s)) \mid \Pi(\Gamma){=}\mathbf{a}\right] = \mathbb{E}\left[\mathbf{u}(\Pi(\Gamma')) \mid \Pi(\Gamma){=}\mathbf{a}\right] \tag{3}$$

for all $\mathbf{a} \in A$ *and consequently* $\mathbb{E}\left[\mathbf{u}(\Pi(\Gamma^s))\right] = \mathbb{E}\left[\mathbf{u}(\Pi(\Gamma'))\right]$.

We prove this in Appendix G. Because of this result, we will focus one these particular types of safe Pareto improvements, which simply create an isomorphic game with different (Pareto-better) utilities. Note, however, that without assigning exact probabilities to the distributions of $\Pi(\Gamma), \Pi(\Gamma')$, the original players will in general not be able to *construct* a Γ^s that satisfies the expected payoff equalities. For this reason, one could still conceive of situations in which a different type of SPI would be chosen by the original players and the original players are unable to instead choose an SPI of the type described in Lemma 9.5.

Lemma 9.5 directly implies a characterization of the expected utilities than can be achieved with perfect-coordination SPIs. Of course, this characterization depends on the exact distribution of $\Pi(\Gamma)$. We omit the statement of this result. However, we state the following implication.

Corollary 9.6. *Under Assumption 1, the set of Pareto improvements that are safely achievable with perfect coordination* $\{\mathbb{E}[\mathbf{u}(\Gamma')] \mid \Gamma'$ *is perfect-coordination SPI on* $\Gamma\}$ *is a convex polygon.*

Because of this result, one can also efficiently optimize convex functions over the set of perfect-coordination SPIs. Even without referring to the distribution $\Pi(\Gamma)$, many interesting questions can be answered efficiently. For example, we can efficiently identify the perfect-coordination SPI that maximizes the minimum improvements across players and outcomes $\mathbf{a} \in A$.

In the following, we aim to give maximally strong positive results about what Pareto improvements can be safely achieved, without referring to exact probabilities over $\Pi(\Gamma)$. To keep things simple, we will do this only for the case of two players. To state our results, we first need some notation: We use

$$\text{PF}(C) \coloneqq \left\{\mathbf{y} \in C \mid \nexists \mathbf{y}'{\in}C, i \in \{1, ..., n\} : \mathbf{y}' \geq \mathbf{y}, y_i' > y\right\}$$

to denote the Pareto frontier of a convex polygon C (or more generally convex, closed set). For any real number $x \in \mathbb{R}$, we use $\pi_i(x, C(\Gamma))$ to denote the $\mathbf{y}' \in C(\Gamma)$ which maximizes y_{-i}' under the constraint $y_i' = x$. (Recall that we consider 2-player games, so y_{-i}' is a single real number.) Note that such a $\mathbf{y}'$ exists if and only if x is i's utility in some feasible payoff vector. We first state our result formally. Afterwards, we will give a graphical explanation of the result, which we believe is easier to understand.

Theorem 9.7. *Make Assumption 1. Let* Γ *be a two-player game. Let* $\mathbf{y} \in \mathbb{R}^2$ *be some potentially unsafe Pareto improvement on* $\mathbb{E}\left[\mathbf{u}(\Pi(\Gamma))\right]$. *For* $i = 1, 2$, *let* $x_i^{\min/\max} = \min/\max u_i\left(\text{supp}(\Pi(\Gamma))\right)$. *Then:*
A) If there is some element in $C(\Gamma)$ *which Pareto-dominates all of* $\text{supp}(\Pi(\Gamma))$ *and if* $\mathbf{y}$ *is Pareto-dominated by an element of at least one of the following three sets:*

- L_1 $\coloneqq$ *the line segment between* $\pi_1(x_1^{\min}, \text{PF}(C(\Gamma))$ *and* $\pi_1(x_1^{\max}, \text{PF}(C(\Gamma))$;
- L_2 $\coloneqq$ *the segment of the curve* $\text{PF}(C(\Gamma))$ *between* $\pi_1(x_{\max}^1, \text{PF}(C(\Gamma))))$ *and* $\pi_2(x_{\max}^2 \text{PF}(C(\Gamma))))$;
- L_3 $\coloneqq$ *the line segment between* $\pi_2(x_2^{\max}, \text{PF}(C(\Gamma))$ *and* $\pi_2(x_2^{\min}, \text{PF}(C(\Gamma))$.

Then there is an SPI under improved coordination Γ^s *such that* $\mathbb{E}\left[\mathbf{u}(\Pi(\Gamma^s))\right] = \mathbf{y}$.
B) If there is no element in $C(\Gamma)$ *which Pareto-dominates all of* $\text{supp}(\Pi(\Gamma))$ *and if* $\mathbf{y}$ *is Pareto-dominated by an element each of* L_1 *and* L_3 *as defined above, then there is a perfect-coordination SPI* Γ^s *such that* $\mathbb{E}\left[\mathbf{u}(\Pi(\Gamma^s))\right] = \mathbf{y}$.

We prove this result in Appendix H. We now illustrate the result graphically. We start with Case A, which is illustrated in Figure 1. The Pareto-frontier is the solid line in the north and east. The points marked x indicate outcomes in $\text{supp}(\Pi(\Gamma))$. The point marked by a filled circle indicates the expected value of the default equilibrium $\mathbb{E}\left[\mathbf{u}(\Pi(\Gamma))\right]$. For some $\mathbf{y} \in \mathbb{R}^2$ to be a Pareto-improvement, it must be to the north-east of the filled circle. The vertical dashed lines starting at the two extreme x marks illustrate the application of π_1 to project $x_1^{\min/\max}$ onto the Pareto frontier. The dotted line between these two points is L_1. Similarly, the horizontal dashed lines starting at x marks illustrate the application of π_2 to project $x_2^{\min/\max}$ onto the Pareto frontier. The line segment between these

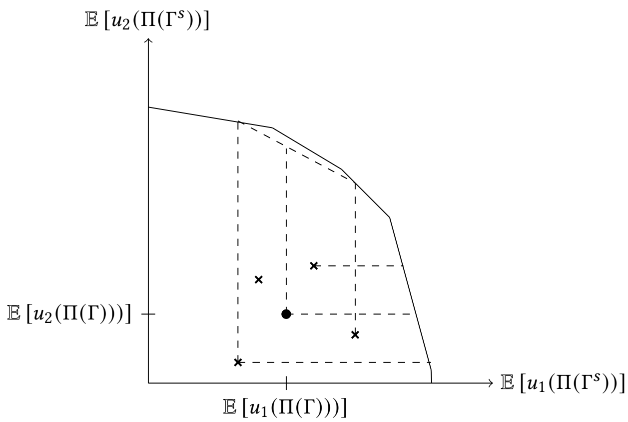

Figure 1: This figure illustrates Theorem 9.7, Case A.

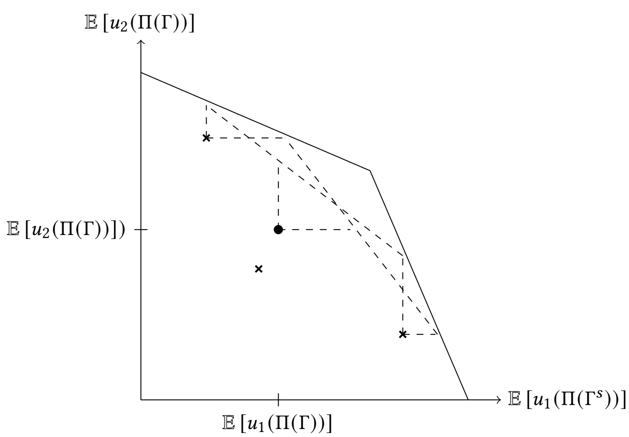

Figure 2: This figure illustrates Theorem 9.7, Case B.

two points is L_3. In this case, this line segments lies on the Pareto frontier. The set L_2 is simply that part of the Pareto frontier, which Pareto-dominates all elements of $\text{supp}(\Pi(\Gamma))$, i.e., the part of the Pareto frontier to the north-east between the two intersections with the northern horizontal dashed line and eastern vertical dashed line.

Case B of Theorem 9.7 is depicted in Figure 2. Note that here the two line segments L_1 and L_3 intersect. To ensure that a Pareto improvement is safely achievable, the theorem requires that it is below both of these lines.

As a corollary of Theorem 9.7, we can see that all (potentially unsafe) Pareto improvements in the DM-RM subset game of the Demand Game of Table 1 are equivalent to some perfect-coordination SPI.

However, there are also games in which some unsafe Pareto improvements (i.e., Pareto-improvements that depend on the distribution of Π) cannot be made safe, even under perfect coordination, which we prove in Appendix I.

Proposition 9.8. *There is a game* $\Gamma = (A, \mathbf{u})$, *representatives* Π *that satisfy Assumptions 1 and 2, and an outcome* $\mathbf{a} \in A$ *s.t.* $u_i(\mathbf{a}) > u_i(\Pi(\Gamma))$ *for all players i, but there is no perfect-coordination SPI* $(A^s, \mathbf{u}^s, \mathbf{u}^e)$ *s.t. for all players i,* $\mathbb{E}\left[u_i^e(\Pi(A^s, \mathbf{u}^s)\right] = u_i(\mathbf{a})$.

10 THE SPI SELECTION PROBLEM

In the Demand Game, there happens to be a single non-trivial SPI. However, in general there may be multiple incomparable safe Pareto improvements that result in different payoffs for the players. This is clear in the setting studied in Section 9. However, it's easy to come up with a variant of the Demand Game (Table 1) which involves such ambiguity, too. If multiple safe Pareto improvements are available, the original players would be left with the difficult decision of which safe Pareto improvement to demand in their instruction.

The difficulty of choosing what SPI to demand cannot be denied. However, we would here like to emphasize that players can profit from the use of SPIs even without addressing this SPI selection problem. To do so, a player picks an instruction that is very compliant ("dove-ish") w.r.t. what SPI is chosen, e.g., one that simply goes with whatever SPI the other players demand as long as that SPI cannot further be safely Pareto-improved upon. Note that in many cases, all such SPIs benefit both players. For example, SPIs in bargaining scenarios like the Demand Game remove the conflict outcome, which benefits all parties. Thus, a player can expect a safe improvement even under such maximally compliant demands on the selected SPI.

11 CONCLUSION AND FUTURE DIRECTIONS

Safe Pareto improvements are a promising new idea for delegating strategic decision making. To conclude this paper, we discuss some ideas for further research on SPIs.

Straightforward technical questions arise in the context of the complexity results of Section 8. First, what impact on the complexity does varying the assumptions have? Our NP-completeness proof is easy to generalize at least to some other types of assumptions. It would be interesting to give a generic version of the result. We also wonder whether there are plausible assumptions under which the complexity changes in interesting ways. Second, one could ask how the complexity changes if we use more sophisticated game representations (see the remarks at the end of that section). Third, one could impose additional restrictions on the sought SPI. For example, some of the players may be unable to have their representative maximize arbitrary utility functions. We could then ask whether there is an SPI in which only a given subset of the players adopt different utility functions and restrictions on the set of available strategies. Fourth, we could restrict the games under consideration. Are there games in which it becomes easy to decide whether there is an SPI?

It would also be interesting to see what real-world situations can already be interpreted as utilizing SPIs, or could be Pareto-improved upon using SPIs.

REFERENCES

[1] Krzysztof R. Apt. 2004. Uniform Proofs of Order Independence for Various Strategy Elimination Procedures. *The B.E. Journal of Theoretical Economics* 4, 1 (2004). https://doi.org/10.2202/1534-5971.1141

[2] Robert Axelrod. 1984. *The Evolution of Cooperation.* Basic Books.

[3] Mihaly Barasz, Paul Christiano, Benja Fallenstein, Marcello Herreshoff, Patrick LaVictoire, and Eliezer Yudkowsky. 2014. Robust Cooperation in the Prisoner's Dilemma: Program Equilibrium via Provability Logic. https://arxiv.org/abs/1401.5577

[4] Vincent Conitzer and Tuomas Sandholm. 2005. Complexity of (Iterated) Dominance. In *Proceedings of the 6th ACM conference on Electronic commerce.* Vancouver, Canada, 88–97. https://doi.org/10.1145/1064009.1064019

[5] Vincent Conitzer and Tuomas Sandholm. 2006. Computing the Optimal Strategy to Commit to. In *Proceedings of the ACM Conference on Electronic Commerce (EC).* Ann Arbor, MI, USA, 82–90.

[6] Stephen A. Cook. 1971. The complexity of theorem-proving procedures. In *STOC '71: Proceedings of the third annual ACM symposium on Theory of computing.* 151–158. https://doi.org/10.1145/800157.805047

[7] Andrew Critch. 2019. A Parametric, Resource-Bounded Generalization of Löb's Theorem, and a Robust Cooperation Criterion for Open-Source Game Theory. *Journal of Symbolic Logic* 84, 4 (12 2019), 1368–1381. https://doi.org/10.1017/jsl.2017.42

[8] Lance Fortnow. 2008. *Program equilibria and discounted ComputationTime.* Technical Report 1473. Kellogg School of Management – Center for Mathematical Studies in Economics and Management Science, Northwestern University. https://www.econstor.eu/bitstream/10419/31197/1/587667478.PDF

[9] Joaquim Gabarró, Alina García, and Maria Serna. 2011. The complexity of game isomorphism. *Theoretical Computer Science* 412, 48 (11 2011), 6675–6695. https://doi.org/10.1016/j.tcs.2011.07.022

[10] John C. Harsanyi and Reinhard Selten. 1988. *A General Theory of Equilibrium Selection in Games.* The MIT Press.

[11] J. V. Howard. 1988. Cooperation in the Prisoner's Dilemma. *Theory and Decision* 24 (5 1988), 203–213. https://doi.org/10.1007/BF00148954

[12] Adam Tauman Kalai, Ehud Kalai, Ehud Lehrer, and Dov Samet. 2010. A commitment folk theorem. *Games and Economic Behavior* 69 (2010), 127–137. https://doi.org/10.1016/j.geb.2009.09.008

[13] Elon Kohlberg and Jean-Francois Mertens. 1986. On the Strategic Stability of Equilibria. *Econometrica* 54, 5 (9 1986), 1003–1037. https://doi.org/10.2307/1912320

[14] Dmytro Korzhyk, Vincent Conitzer, and Ronald Parr. 2010. Complexity of Computing Optimal Stackelberg Strategies in Security Resource Allocation Games. In *Proceedings of the National Conference on Artificial Intelligence (AAAI).* Atlanta, GA, USA, 805–810.

[15] Jean-Jacques Laffont and David Martimort. 2002. *The Theory of Incentives – The Principal-Agent Model.* Princeton University Press.

[16] R. Preston McAfee. 1984. Effective Computability in Economic Decisions. (5 1984). https://www.mcafee.cc/Papers/PDF/EffectiveComputability.pdf

[17] Dov Monderer and Moshe Tennenholtz. 2009. Strong mediated equilibrium. *Artificial Intelligence* 173, 1 (1 2009), 180–195. https://doi.org/10.1016/j.artint.2008.10.005

[18] Thanh H. Nguyen, Arunesh Sinha, Shahrzad Gholami, Andrew J. Plumptre, Lucas Joppa, Milind Tambe, Margaret Driciru, Fred Wanyama, Aggrey Rwetsiba, Rob Critchlow, and Colin Beale. 2016. CAPTURE: A New Predictive Anti-Poaching Tool for Wildlife Protection. In *Proceedings of the Fifteenth International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS).* Singapore, 767–775.

[19] Caspar Oesterheld. 2019. Robust Program Equilibrium. *Theory and Decision* 86, 1 (2 2019), 143–159.

[20] Martin J. Osborne. 2004. *An Introduction to Game Theory.* Oxford University Press.

[21] Praveen Paruchuri, Jonathan P. Pearce, Janusz Marecki, Milind Tambe, Fernando Ordóñez, and Sarit Kraus. 2008. Playing games for security: An efficient exact algorithm for solving Bayesian Stackelberg games. In *Proceedings of the Seventh International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS).* Estoril, Portugal, 895–902.

[22] David G. Pearce. 1984. Rationalizable Strategic Behavior and the Problem of Perfection. *Econometrica* 54, 4 (7 1984), 1029–1050.

[23] James Pita, Manish Jain, Craig Western, Christopher Portway, Milind Tambe, Fernando Ordóñez, Sarit Kraus, and Praveen Parachuri. 2008. Deployed ARMOR protection: The application of a game-theoretic model for security at the Los Angeles International Airport. In *Proceedings of the Seventh International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS).* Estoril, Portugal, 125–132.

[24] James Pita, Milind Tambe, Chris Kiekintveld, Shane Cullen, and Erin Steigerwald. 2011. GUARDS - Game Theoretic Security Allocation on a National Scale. In *Proceedings of the Tenth International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS).* Taipei, Taiwan, 37–44.

[25] Werner Raub. 1990. A General Game-Theoretic Model of Preference Adaptions in Problematic Social Situations. *Rationality and Society* 2, 1 (1 1990), 67–93.

[26] Thomas C. Schelling. 1960. *The Strategy of Conflict.* Harvard University Press.

[27] Alexander Schrijver. 1998. *Theory of Linear and Integer Programming.* John Wiley & Sons.

[28] Amartya Sen. 1974. Choice, orderings and morality. In *Practical Reason*, Stephan Körner (Ed.). Basil Blackwell, Chapter II, 54–67.

[29] Eric Shieh, Bo An, Rong Yang, Milind Tambe, Craig Baldwin, Joseph DiRenzo, Ben Maule, and Garrett Meyer. 2012. PROTECT: A Deployed Game Theoretic System to Protect the Ports of the United States. In *Proceedings of the Eleventh International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS).* Valencia, Spain, 13–20.

[30] Milind Tambe. 2011. *Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned.* Cambridge University Press.

[31] Moshe Tennenholtz. 2004. Program equilibrium. *Games and Economic Behavior* 49, 2 (11 2004), 363–373.

[32] Wiebe van der Hoek, Cees Witteveen, and Micheal Wooldridge. 2013. Program equilibrium – a program reasoning approach. *International Journal of Game Theory* 42 (8 2013), 639–671. Issue 3.

[33] John von Neumann. 1928. Zur Theorie der Gesellschaftsspiele. *Math. Ann.* 100 (1928), 295–320. https://doi.org/10.1007/BF01448847

[34] Heinrich von Stackelberg. 1934. *Marktform und Gleichgewicht.* Springer, Vienna, 58–70.

[35] Bernhard von Stengel and Shmuel Zamir. 2010. Leadership Games with Convex Strategy Sets. *Games and Economic Behavior* 69 (2010), 446–457.