

# Safe Pareto Improvements for Delegated Game Playing

Caspar Oesterheld

Department of Computer Science, Duke University  
caspar.oesterheld@duke.edu

Vincent Conitzer

Department of Computer Science, Duke University  
conitzer@cs.duke.edu

## 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 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:

Caspar Oesterheld and Vincent Conitzer. 2021. Safe Pareto Improvements for Delegated Game Playing. In *Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021)*, Online, May 3–7, 2021, IFAAMAS, 9 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, but they worry 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 Military”) 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 military (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” (perhaps contractually guaranteeing a reward in proportion to the utility of the final outcome). 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. Both original players (Aliceland and Bobbesia) can use the following instruction for their representatives: “If the opponent’s instruction is equal to this instruction, Cooperate; otherwise Defect.” [9, 13, 20] Then it is a Nash equilibrium for both players to use this instruction. In this equilibrium, (Cooperate, Cooperate) is played and it is thus Pareto-optimal and Pareto-better than the default.

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, Online. © 2021 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

		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**

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-improves 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 {DL, 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. (Again, we could imagine that the instructions specify Table 2 to be how Aliceland and Bobbesia financially reward Alice and Bob.) 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 isomorphically. 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.

*Contributions.* 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 define and motivate the concept of safe Pareto improvements in more detail. In Section 4,

		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**

we briefly review the the concepts of program games and program equilibrium and show that SPIs can be implemented as program equilibria. In Section 5, we introduce a notion of outcome correspondence between games. This relation expresses the original players’ beliefs about similarities between how the representatives play different games. For example, in our example, the Demand Game of Table 1 (arguably) corresponds 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 SPIs. In particular, we show that a game  $\Gamma^s$  is an SPI on another game  $\Gamma$  if and only if there is a Pareto-improving outcome correspondence relation between  $\Gamma^s$  and  $\Gamma$ .

To derive SPIs, 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 SPI for the Demand Game can be proven using these assumptions. Section 8 shows that determining whether there exists an SPI 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 SPIs can be done in polynomial time. We conclude by discussing the problem of selecting between different SPIs on a given game (Section 10) and giving some ideas for directions for future work (Section 11). For lack of space, we omit some technical detail. See [users.cs.duke.edu/~ocaspar/SPIAAMAS.pdf](https://users.cs.duke.edu/~ocaspar/SPIAAMAS.pdf) for a copy of this paper with an appendix containing detailed proofs (including easy ones), etc.

## 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  $A = A_1 \times \dots \times A_n$  of (pure) strategy profiles (or outcomes) and a function  $\mathbf{u}: A \rightarrow \mathbb{R}^n$  that assigns to each outcome a utility for each player. Instead of,  $(A, \mathbf{u})$  will also write  $(A_1, \dots, A_n, u_1, \dots, u_n)$ . 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  $\Gamma$  if  $A'_i \subseteq A_i$  for  $i = 1, \dots, n$ . Note that a subset game may assign different utilities to outcomes than the original game. For any set of strategies  $S$ , we denote by  $\Gamma - S := ((A_1 - S) \times \dots, (A_n - S), \mathbf{u})$  the game that arises from  $\Gamma$  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  $i = 1, \dots, n$ . 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

For sets  $M$  and  $N$ , a *multi-valued function*  $\Phi: M \multimap N$  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  $\Phi(Q) := \bigcup_{m \in Q} \Phi(m)$ . Note that  $\Phi(Q) \subseteq N$  and that  $\Phi(\emptyset) = \emptyset$ . For any set  $M$ , we define the identity function  $\text{id}_M: M \multimap M: m \mapsto \{m\}$ . Also, for two sets  $M, N$ , we define  $\text{all}_{M,N}: M \multimap N: m \mapsto N$ . We define the inverse  $\Phi^{-1}: N \multimap M: n \mapsto \{m \in M \mid n \in \Phi(m)\}$ . Note that  $\Phi^{-1}(\emptyset) = \emptyset$  for any multi-valued function  $\Phi$ . For sets  $M, N, Q$  and functions  $\Phi: M \multimap N$ ,  $\Psi: N \multimap Q$ , we define the composite  $\Psi \circ \Phi: M \multimap Q: m \mapsto \Psi(\Phi(m))$ . As with regular functions, composition of multi-valued functions is associative. We say that  $\Phi: M \multimap N$  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  $\Gamma$  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 [12].

We imagine that one way in which the representatives can be instructed is to in turn play a subset game  $\Gamma^s = (A_1^s \subseteq A_1, \dots, A_n^s \subseteq A_n, \mathbf{u}^s)$  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  $\Gamma^s$ , we denote by  $\Pi(\Gamma^s)$  the outcome that arises if the representatives play the subset game  $\Gamma^s$  of  $\Gamma$ . Because in many games, it is not clear what the right choice is, the original players might be uncertain about  $\Pi(\Gamma^s)$  for many games  $\Gamma^s$ . 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 be played for any  $\Gamma$ .

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 ( $\mathbb{E}[\mathbf{u}(\Pi(\Gamma))]$ ) and then let the representatives play a distribution over outcomes whose expected utility exceeds that default expected utility. However, this is unrealistic if  $\Gamma$  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  $\Gamma$ . 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  $\Gamma^s$  that are Pareto improvements *with certainty* under weak and purely qualitative assumptions about  $\Pi$ .

*Definition 3.1.* Let  $\Gamma^s$  be a subset game of  $\Gamma$ . We say  $\Gamma^s$  is a *safe Pareto improvement (SPI)* on  $\Gamma$  if  $\mathbf{u}(\Pi(\Gamma^s)) \geq \mathbf{u}(\Pi(\Gamma))$  with certainty. We say that  $\Gamma^s$  is a *strict SPI* if furthermore, there is a player  $i$  s.t.  $u_i(\Pi(\Gamma^s)) > u_i(\Pi(\Gamma))$  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  $\Gamma^s$ ? Are SPIs Nash equilibria of the meta game played by the representatives? In this section, we briefly describe one way to fill this gap by discussing the concept of program games and program equilibrium [2, 4, 6, 15, 20]. 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 avoid introducing notation.

For any game  $\Gamma = (A, \mathbf{u})$ , the program equilibrium literature considers the following meta game. First, each player  $i$  chooses from a set of computer programs. Each program then receives as input a vector containing everyone else's chosen program. Each player  $i$ 's program then returns an action from  $A_i$ , player  $i$ 's set of actions in  $\Gamma$ . Together these actions then form an outcome  $\mathbf{a} \in A$  of the original game. Finally, the utilities  $\mathbf{u}(\mathbf{a})$  are realized according to the utility function of  $\Gamma$ . The meta game can be analyzed like any other game. Its Nash equilibria are called *program equilibria*. Importantly, the program equilibria can implement payoffs not implemented by any Nash equilibria of  $\Gamma$  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." [9, 13, 20] This is a program equilibrium which implements mutual cooperation.

In the setting for our paper, we similarly imagine that each player  $i$  can choose from a set of programs that in turn choose from  $A_i$ . However, the types of program that we have in mind here are more

sophisticated than those typically considered in the program equilibrium literature. Specifically we imagine that the programs are executed by intelligent *representatives* who are themselves able to competently choose an action for player  $i$  in any given game  $\Gamma^s$ , without the original player having to describe how this choice is to be made. The original player may not even understand much about this program other than that it generally plays well. Thus, in addition to the elementary instructions used in a typical computer program (branches, comparisons, arithmetic operations, etc.), we allow player  $i$  to use an instruction “Play  $\Pi_i(\Gamma^s)$ ” in the program she submits. To jointly let the representatives play, e.g., the SPI  $\Gamma^s$  of Table 2 on the Demand Game of Table 1, the representatives can both use an instruction that says, “If the opponent’s chosen program is analogous to this one, play  $\Pi_i(\Gamma^s)$ ; otherwise play DM”. Assuming some minimal rationality requirements on the representatives (i.e., on how “play  $\Pi_i(\Gamma^s)$ ” is implemented), this is a Nash equilibrium. More generally, we can prove the following result (see the full version for a proof).

**THEOREM 4.1.** *Let  $\Gamma$  be a game and  $\Gamma^s$  be an SPI of  $\Gamma$ . Now consider a program game on  $\Gamma$ , where each player  $i$  can choose from a set of computer programs that output actions for  $\Gamma$ . In addition to the normal kind of instructions, we allow the use of the command “play  $\Pi_i(\Gamma^s)$ ” for any subset game  $\Gamma'$  of  $\Gamma$ . Finally, assume that  $\Pi(\Gamma)$  guarantees each player  $i$  at least that player’s minimax utility (a.k.a. threat point) in the base game  $\Gamma$ . 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. 10, 14].

## 5 OUTCOME CORRESPONDENCE BETWEEN GAMES

In this section, we introduce a notion of outcome correspondence, which we will see is essential to constructing SPIs.

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

Note that  $\Gamma \sim_{\Phi} \Gamma'$  is a statement about  $\Pi$ , i.e., about how the representatives choose. Whether such a statement holds generally depends 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  $\Gamma$  and  $\Gamma'$  are isomorphic, then one might expect  $\Gamma \sim_{\Phi} \Gamma'$ , where  $\Phi$  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  $\Phi, \Xi: A \rightarrow A', \Psi: A' \rightarrow \hat{A}$ .*

- (1) *Reflexivity:*  $\Gamma \sim_{\text{id}_A} \Gamma$ , where  $\text{id}_A: A \rightarrow A: \mathbf{a} \mapsto \{\mathbf{a}\}$ .
- (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  $\text{all}_{A,A'}: A \rightarrow A': \mathbf{a} \mapsto A'$ .
- (6) *If  $\Gamma \sim_{\Phi} \Gamma'$  and  $\Phi(\mathbf{a}) = \emptyset$ , then  $\Pi(\Gamma) \neq \mathbf{a}$  with certainty.*

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 SPIs. The following theorem shows not only how outcome correspondences can be used to find (and prove) SPIs. It also shows that any SPI 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  $\Gamma$ . Then  $\Gamma^s$  is an SPI on  $\Gamma$  if and only if there is  $\Phi$  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,  $\Pi(\Gamma^s) \in \Phi(\Pi(\Gamma))$  with certainty. Hence, for  $i = 1, 2$ ,  $u_i(\Pi(\Gamma^s)) \in u_i(\Phi(\Pi(\Gamma)))$  with certainty. Hence, by assumption about  $\Phi$ , 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  $\Phi: A \rightarrow A^s: \mathbf{a} \mapsto \{\mathbf{a}^s \in A^s \mid \mathbf{u}(\mathbf{a}^s) \geq \mathbf{u}(\mathbf{a})\}$ . It is immediately obvious that  $\Phi$  is Pareto-improving as required. Also, 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  $\Phi$ , it holds that  $\mathbf{a}^s \in \Phi(\mathbf{a})$ . We conclude that  $\Gamma \sim_{\Phi} \Gamma^s$  as claimed.  $\square$

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  $\Gamma^s$  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  $\Gamma$  be the Demand Game of Table 1. First, it seems plausible that  $\Gamma$  is in some sense equivalent to  $\Gamma'$ , where  $\Gamma' = \Gamma - \{\text{DL}, \text{RL}\}$  is the game that results from removing DL and RL for both players from  $\Gamma$ . 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  $\Gamma^s$  is the game of Table 2 and  $\Psi$  is the isomorphism between  $\Gamma'$  and  $\Gamma^s$ . 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  $\Gamma$ , notice that  $\Phi$  does not change utilities at all.  $\Psi$  maps the conflict outcome (DM, DM) onto the outcome (DL, DL), which is better for both original players. Other than that,  $\Psi$ , too, does not change the utilities. Hence,  $\Psi \circ \Phi$  is Pareto-improving. By Theorem 6.1,  $\Gamma^s$  is therefore an SPI on  $\Gamma$ .

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 to 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 isomorphically.

**ASSUMPTION 7.1.** *Let  $\Gamma = (A, \mathbf{u})$  and  $\Gamma' = (A', \mathbf{u}')$  be two games for which there are single-valued bijections  $\Phi_i: A_i \rightarrow A'_i$  for  $i = 1, \dots, n$  such that  $\mathbf{u}(a_1, \dots, a_n) = \mathbf{u}'(\Phi_1(a_1), \dots, \Phi_n(a_n))$  for all  $\mathbf{a} \in A$ . Then for one tuple  $\Phi$  of such bijections,  $\Gamma \sim_{\Phi} \Gamma'$ .*

Similar desiderata have been discussed in the context of equilibrium selection, e.g., by Harsanyi and Selten [8, 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 7.1 by referring to focal points (introduced by Schelling [19, 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 7.1 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 7.2.** *Let  $\Gamma = (A, \mathbf{u})$  be an arbitrary  $n$ -player game where  $A_1, \dots, A_n$  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 7.2 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  $a_j$  strictly dominates. Assumption 7.2 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 [11, 16, 17, Section 2.9, Chapter 12] and has rarely been questioned. It is also implicit in the solution concept of Nash equilibrium – if a strategy is removed

by iterated strict dominance, that strategy is played in no Nash equilibrium. 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 7.2 to hold, we will in most games  $\Gamma$  have to assume that the representatives cannot in turn make credible precommitments (or delegate to further subrepresentatives) or play the game iteratively [1].

With Assumptions 7.1 and 7.2 we can finally state our example SPIs formally:

**PROPOSITION (EXAMPLE) 7.1.** *Let  $\Gamma$  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  $\Gamma$  with  $A_1^s = A_2^s = \{\text{Cooperate}\}$ . Then under Assumption 7.2,  $\Gamma^s$  is a strict SPI on  $\Gamma$ .*

**PROPOSITION (EXAMPLE) 7.2.** *Let  $\Gamma$  be the Demand Game of Table 1 and  $\Gamma^s$  be the subset game described in Table 2. Under Assumptions 7.1 and 7.2,  $\Gamma^s$  is an SPI on  $\Gamma$ . Further, if  $P(\Pi(\Gamma) = (\text{DM}, \text{DM})) > 0$ ,  $\Gamma^s$  is a strict SPI.*

## 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  $\Gamma$  a non-trivial SPI  $\Gamma^s$ . In particular, we ask whether a given game  $\Gamma$  has a non-trivial SPI that can be proved using only Assumptions 7.1 and 7.2, Transitivity (Lemma 5.2.3) and Theorem 6.1. Formally:

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

- (1) (Non-triviality:) If we fully reduce  $\Gamma^k$  and  $\Gamma$  using iterated strict dominance (Assumption 7.2), the two resulting games are not equal. (Of course, they are allowed to be isomorphic.)
- (2) For  $i = 1, \dots, k$ ,  $\Gamma^{i-1} \sim_{\Phi^i} \Gamma^i$  is valid by a single application of either Assumption 7.1 or Assumption 7.2.
- (3) For all  $\mathbf{a} \in A$ , and whenever  $\mathbf{a}^s \in (\Phi^k \circ \Phi^{k-1} \circ \dots \circ \Phi^1)(\mathbf{a})$ , 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  $\Gamma$  s.t.  $u_i((\Phi^k \circ \Phi^{k-1} \circ \dots \circ \Phi^1)(\mathbf{a})) > u_i(\mathbf{a})$ .

Many variants of this problem may be considered. For example, we might generalize it to allow imposing additional properties on the SPI. This will generally not change the computational complexity of the problem. One may also wish to compute all SPIs, or – in line with multi-criteria optimization [5, 21] – of all SPIs that cannot in turn be safely improved upon. However, in general there may exist exponentially many such SPIs. To retain any hope of developing an efficient algorithm, one would therefore have to first develop a more efficient representation scheme [cf. 18, Sect. 16.4].

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

**PROPOSITION 8.3.** *For games  $\Gamma$  with  $|A_1| + \dots + |A_n| = m$  that can be reduced (via iterative application of Assumption 7.2) to a game*

$\Gamma'$  with  $|A'_1| + \dots + |A'_n| = l$ , the (strict) SPI decision problem can be solved in  $O(m^l)$ .

The full proof is tedious (see the full version of this paper), but the main idea is simple. To find an SPI on  $\Gamma$  based on Assumptions 7.1 and 7.2, one has to first iteratively remove all strictly dominated actions to obtain a reduced game  $\Gamma'$ , which the representatives would play the same as the original game. This can be done in polynomial time. One then has to map the actions  $\Gamma'$  onto the original  $\Gamma$  in such a way that each outcome in  $\Gamma'$  is mapped onto a weakly Pareto-better outcome in  $\Gamma$ . Our proof of NP-hardness works by reducing from the subgraph isomorphism problem, where the payoff matrices of  $\Gamma'$ ,  $\Gamma$  represent the adjacency matrices of the graphs.

Besides being about a specific set of assumptions about  $\sim$ , note that Theorem 8.2 and Proposition 8.3 also assume that the utility function of the game is represented explicitly in normal form as a payoff matrix. If we changed the game representation (e.g., to boolean circuits, extensive form game trees, quantified boolean formulas, or even Turing machines), this can affect the complexity of the SPI problem [cf. 7]. 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 [3, 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  $\Gamma = (A, \mathbf{u})$ ,

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

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

For any game  $\Gamma$ , 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  $u_i^s: A^s \rightarrow \mathbb{R}$  are arbitrary utility functions to be used by the representatives and  $\mathbf{u}^e: A^s \rightarrow C(\Gamma)$  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  $\Gamma$ . 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  $\Gamma$  be a game. A *perfect-coordination SPI* for  $\Gamma$  is a perfect-coordination token game  $(A^s, \mathbf{u}^s, \mathbf{u}^e)$  for  $\Gamma$  s.t.  $\mathbf{u}^e(\Pi(A^s, \mathbf{u}^s)) \geq \mathbf{u}(\Pi(\Gamma))$  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  $u_i^e(\Pi(A^s, \mathbf{u}^s)) > u_i(\Pi(\Gamma))$  with positive probability.

As an example, imagine that  $\Gamma$  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 7.1 implies that  $\Gamma \sim_{\Phi} \Gamma^s$ , where  $\Phi$  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. The latter may not be the case in, e.g., 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 do 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 than before.

*Definition 9.2.* The *strict full-coordination SPI decision problem* consists in deciding for any given  $\Gamma$  whether under Assumption 7.1 there is a perfect-coordination SPI  $\Gamma^s$  for  $\Gamma$ .

**LEMMA 9.3.** For a given  $n$ -player game  $\Gamma$  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  $C(\Gamma)$ .

Based on Lemma 9.3, Algorithm 1 decides whether there is a strict perfect-coordination SPI for a given game  $\Gamma$ .

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  $\Gamma$ , set  $\text{supp}(\Pi(\Gamma))$

```

1 for  $\mathbf{a} \in \text{supp}(\Pi(\Gamma))$  do
2   if  $\mathbf{u}(\mathbf{a})$  is Pareto-suboptimal within  $C(\Gamma)$  then
3     Return True;
4 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 7.1 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  $\mathbb{E}[\mathbf{u}(\Pi(\Gamma))]$ . We start with a lemma that directly provides a characterization. So far, all the considered perfect-coordination SPIs ( $A^s, \mathbf{u}^s, \mathbf{u}^e$ ) for a game  $(A, \mathbf{u})$  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  $\Gamma = (\{a_1^1, \dots, a_1^l\}, \dots, \{a_n^1, \dots, a_n^l\}, \mathbf{u})$  be any game. Let  $\Gamma^s$  be a perfect-coordination SPI on  $\Gamma$ . Then we can define  $\mathbf{u}^e$  with values in  $C(\Gamma)$  such that under Assumption 7.1 the game*

$$\Gamma^s = \left( \hat{A}_1 := \{\hat{a}_1^1, \dots, \hat{a}_1^l\}, \dots, \hat{A}_n := \{\hat{a}_n^1, \dots, \hat{a}_n^l\}, \right. \\ \left. \hat{\mathbf{u}}: (\hat{a}_1^1, \dots, \hat{a}_n^l) \mapsto \mathbf{u}(a_1^1, \dots, a_n^l), \mathbf{u}^e \right)$$

is also an SPI on  $\Gamma$ , with

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

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

Because of this result, we will focus on these particular types of SPIs, 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  $\Gamma^s$  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 7.1, the set of Pareto improvements that are safely achievable with perfect coordination  $\{\mathbb{E}[\mathbf{u}(\Gamma') \mid \Gamma' \text{ is perfect-coordination SPI on } \Gamma]\}$  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 use Lemma 9.5 and Corollary 9.6 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  $\text{PF}(C) := \{\mathbf{y} \in C \mid \nexists \mathbf{y}' \in C, i \in \{1, \dots, n\} : \mathbf{y}' \geq \mathbf{y}, y'_i > y_i\}$  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 7.1. Let  $\Gamma$  be a two-player game. Let  $\mathbf{y} \in \mathbb{R}^2$  be some potentially unsafe Pareto improvement on  $\mathbb{E}[\mathbf{u}(\Pi(\Gamma))]$ . For  $i = 1, 2$ , let  $x_i^{\min/\max} = \min/\max u_i(\text{supp}(\Pi(\Gamma)))$ . 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:*

- $L_1 :=$  the line segment between  $\pi_1(x_1^{\min}, \text{PF}(C(\Gamma)))$  and  $\pi_1(x_1^{\max}, \text{PF}(C(\Gamma)))$ ;
- $L_2 :=$  the segment of the curve  $\text{PF}(C(\Gamma))$  between  $\pi_1(x_1^{\max}, \text{PF}(C(\Gamma)))$  and  $\pi_2(x_2^{\max}, \text{PF}(C(\Gamma)))$ ;
- $L_3 :=$  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  $\Gamma^s$  such that  $\mathbb{E}[\mathbf{u}(\Pi(\Gamma^s))] = \mathbf{y}$ .*

*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  $\Gamma^s$  such that  $\mathbb{E}[\mathbf{u}(\Pi(\Gamma^s))] = \mathbf{y}$ .*

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  $\mathbb{E}[\mathbf{u}(\Pi(\Gamma))]$ . 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  $\pi_1$  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  $\pi_2$  to project  $x_2^{\min/\max}$  onto the Pareto frontier. The line segment between these two points is  $L_3$ . In this case, this

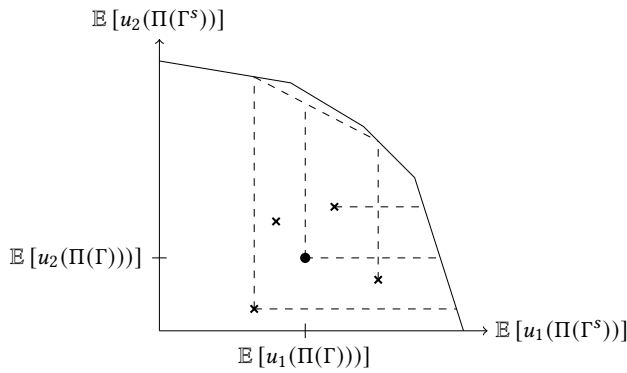


Figure 1: This figure illustrates Theorem 9.7, Case A.

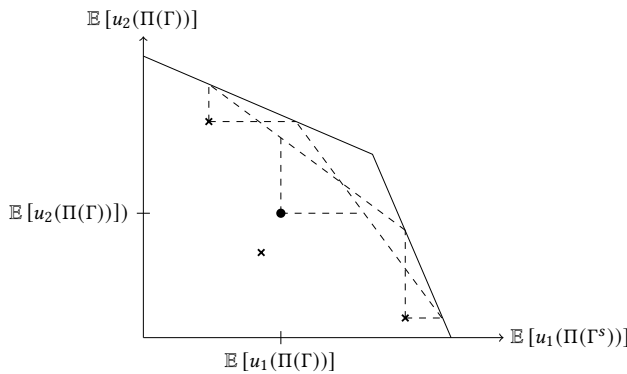


Figure 2: This figure illustrates Theorem 9.7, Case B.

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.

Theorem 9.7 is proven by re-mapping each of the outcomes of the original game as per Lemma 9.5. For example, the projection of the default equilibrium  $\mathbb{E}[\mathbf{u}(\Pi(\Gamma))]$  (i.e., the filled circle) onto  $L_1$  is obtained as an SPI by projecting all the outcomes (i.e., all the  $x$  marks) onto  $L_1$ . In Case A, any utility vector  $\mathbf{y} \in L_2$  that Pareto-improves on all outcomes of the original game can be obtained by re-mapping all outcomes onto  $\mathbf{y}$ . Other kinds of  $\mathbf{y}$  are handled similarly. For brevity, we only give a detailed proof in the full version.

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, this is not always the case:

**PROPOSITION 9.8.** *There is a game  $\Gamma = (A, \mathbf{u})$ , representatives  $\Pi$  that satisfy Assumptions 7.1 and 7.2, 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  $i$ ,  $\mathbb{E}[u_i^e(\Pi(A^s, \mathbf{u}^s))] = u_i(\mathbf{a})$ .*

## 10 THE SPI SELECTION PROBLEM

In the Demand Game, there happens to be a single non-trivial SPI. However, in general (even without the type of coordination assumed in Section 9) there may be multiple incomparable SPIs that result in different payoffs for the players. If multiple SPIs are available, the original players would be left with the difficult decision of which SPI to demand in their instruction.

This 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. 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.

## ACKNOWLEDGMENTS

This work was supported by NSF under Award IIS-1814056 and the Foundational Research Institute (now the Center on Long-Term Risk). For valuable comments and discussions, we thank our anonymous referees, Keerti Anand, Tobias Baumann, Jesse Clifton, Max Daniel, Lukas Gloor, Adrian Hutter, Anni Leskelä, Brian Tomasik and Johannes Treutlein.



## REFERENCES

- [1] Robert Axelrod. 1984. *The Evolution of Cooperation*. Basic Books, New York.
- [2] 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>
- [3] Vincent Conitzer and Tuomas Sandholm. 2005. Complexity of (Iterated) Dominance. In *Proceedings of the 6th ACM conference on Electronic commerce*. Association for Computing Machinery, Vancouver, Canada, 88–97. <https://doi.org/10.1145/1064009.1064019>
- [4] 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>
- [5] Matthias Ehrgott. 2005. *Multicriteria Optimization* (2 ed.). Springer, Berlin.
- [6] 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>
- [7] 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>
- [8] John C. Harsanyi and Reinhard Selten. 1988. *A General Theory of Equilibrium Selection in Games*. The MIT Press, Cambridge, MA.
- [9] J. V. Howard. 1988. Cooperation in the Prisoner's Dilemma. *Theory and Decision* 24 (5 1988), 203–213. <https://doi.org/10.1007/BF00148954>
- [10] 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>
- [11] 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>
- [12] Jean-Jacques Laffont and David Martimort. 2002. *The Theory of Incentives – The Principal-Agent Model*. Princeton University Press, Princeton, NJ.
- [13] R. Preston McAfee. 1984. Effective Computability in Economic Decisions. (5 1984). <https://www.mcafee.cc/Papers/PDF/EffectiveComputability.pdf>
- [14] 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>
- [15] Caspar Oesterheld. 2019. Robust Program Equilibrium. *Theory and Decision* 86, 1 (2 2019), 143–159.
- [16] Martin J. Osborne. 2004. *An Introduction to Game Theory*. Oxford University Press, New York.
- [17] David G. Pearce. 1984. Rationalizable Strategic Behavior and the Problem of Perfection. *Econometrica* 54, 4 (7 1984), 1029–1050.
- [18] Guillaume Perez. 2017. *Decision diagrams: constraints and algorithms*. Ph.D. Dissertation. Université Côte d'Azur. <https://tel.archives-ouvertes.fr/tel-01677857/document>
- [19] Thomas C. Schelling. 1960. *The Strategy of Conflict*. Harvard University Press, Cambridge, MA.
- [20] Moshe Tennenholtz. 2004. Program equilibrium. *Games and Economic Behavior* 49, 2 (11 2004), 363–373.
- [21] Luc N. van Wassenhove and Ludo F. Gelders. 1980. Solving a bicriterion scheduling problem. *European Journal of Operational Research* 4 (1980), 42–48.