Safe Pareto Improvements for Expected Utility Maximizers in Program Games

Anthony DiGiovanni
Center on Long-Term Risk
London, United Kingdom
anthony.digiovanni@longtermrisk.org

Jesse Clifton
Center on Long-Term Risk
London, United Kingdom
jesse.clifton@longtermrisk.org

Nicolas Macé
Center on Long-Term Risk
London, United Kingdom
nicolas.mace@longtermrisk.org

ABSTRACT

Agents in mixed-motive coordination problems such as Chicken may fail to coordinate on a Pareto-efficient outcome. Safe Pareto improvements (SPIs) were originally proposed to mitigate miscoordination in cases where players lack probabilistic beliefs as to how their agents will play a game; agents are instructed to behave so as to guarantee a Pareto improvement on how they would play by default. More generally, SPIs may be defined as transformations of strategy profiles such that all players are necessarily better off under the transformed profile. In this work, we investigate the extent to which SPIs can reduce downsides of miscoordination between expected utility-maximizing agents. We consider games in which players submit computer programs that can condition their decisions on each other's code, and use this property to construct SPIs using programs capable of *renegotiation*. We first show that under mild conditions on players' beliefs, each player always prefers to use renegotiation. Next, we show that under similar assumptions, each player always prefers to be willing to renegotiate at least to the point at which they receive the lowest payoff they can attain in any efficient outcome. Thus subjectively optimal play guarantees players at least these payoffs, without the need for coordination on specific Pareto improvements.

KEYWORDS

program equilibrium; bargaining; Pareto efficiency; Cooperative AI

ACM Reference Format:
Anthony DiGiovanni, Jesse Clifton, and Nicolas Macé. 2025. Safe Pareto Improvements for Expected Utility Maximizers in Program Games. In *Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Detroit, Michigan, USA, May 19 – 23, 2025*, IFAAMAS, 9 pages.

1 INTRODUCTION

Artificially intelligent (AI) systems will increasingly advise or make decisions on behalf of humans, including in interactions with other agents. Thus there is a need for research on *cooperative AI* [2, 5]: How can we design AI systems that are capable of interacting with other players in ways that lead to high social welfare? One way that AI systems assisting humans could fail to cooperate is by failing to coordinate on one of several Pareto-efficient equilibria. This risk is especially large in *bargaining problems*, where players have different preferences over Pareto-efficient equilibria (think of the game of Chicken). These problems are particularly prone to miscoordination, where each player uses a strategy that is part of some Pareto-efficient equilibrium, but collectively the players' strategies are not an equilibrium. Bargaining problems are ubiquitous, including in high-stakes negotiations over climate change, nuclear proliferation, or military disputes, making them a crucial area of study for cooperative AI.

We will explore how the ability of AI systems to condition their decisions on each other's inner workings could reduce downsides of miscoordination in bargaining problems. The literature on *program equilibrium* has shown how games played by computer programs that can read each other's source code admit more cooperative equilibria in other challenges for cooperation such as the Prisoner's Dilemma [16, 18, 26]. *Safe Pareto improvements (SPIs)* [19] were proposed as a mitigation for inefficiencies in settings where players have delegates play a game on their behalf, and have Knightian uncertainty (i.e., lack probabilistic beliefs [15]) about how their delegates will play. Under an SPI, players change their default policies so as to guarantee Pareto improvement on the default outcome. For example, consider two parties A and B who would by default go to war over some territory. They might instruct their delegates to, instead, accept the outcome of a lottery that allocates the territory to A with the probability that A would have won the war.

We will consider the extent to which SPIs can mitigate inefficiencies from miscoordination when (i) players do have probabilistic beliefs and maximize subjective expected utility and (ii) games are played by computer programs that can condition on their counterparts' source code. Our goal is to establish guarantees against miscoordination in the well-studied program game setting. Relaxations of standard assumptions in this setting — e.g., players can precisely read each other's programs' source code, can syntactically verify if a program follows some template [26], and participate in the program game in the first place — are left to future work. While this is an idealized framework, insights from studying program games could be applied to more realistic interactions between actors with some degree of conditional commitment ability. For example, countries engaging in climate negotiations might write bills that specify when the country would be bound to some policies conditional on the terms of other countries' bills [11]. And, smart contracts implemented on a blockchain could execute commitments to transactions conditional on other actors' contracts [25, 27].

Our contributions are as follows:

(1) We construct SPIs in the program game setting using programs that *renegotiate*. Such programs have a "default" program; check if their default played against their counterparts' defaults results in an inefficient outcome; and, if so, call a

Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Y. Vorobeychik, S. Das, A. Nowé (eds.), May 19 – 23, 2025, Detroit, Michigan, USA. © 2025 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

renegotiation routine in an attempt to Pareto-improve on the default outcome. We examine when renegotiation would be used by players who optimize expected utility given their beliefs about what programs their counterparts will use (i.e., in *subjective equilibrium* [14]). Under mild assumptions on players' beliefs, we show that SPIs are always used in subjective equilibrium (Propositions 1 and 2).

(2) We show that due to the ability to renegotiate, under mild assumptions on players' beliefs, players always weakly prefer programs that guarantee them at least the lowest payoff they can obtain on the Pareto frontier (Theorem 3). Following Rabin [21], we call this payoff profile the *Pareto meet minimum (PMM)*. Thus we provide for this setting a (partial) solution to the "SPI selection problem" identified by Oesterheld and Conitzer [19] (hereafter, "OC"), i.e., the problem that players must coordinate among SPIs in order to Pareto-improve on default outcomes. The intuition for this is: The PMM is the most efficient point such that, no matter how aggressively the players bargain, no one expects to risk getting a worse deal by being willing to renegotiate to that point.

2 RELATED WORK

Program equilibrium and commitment games. We build on program games, where computer players condition their actions on each other's source code. Prior work has shown that the ability of computer-based agents to condition their decisions on their counterparts' programs can enable more efficient equilibria [4, 6, 12, 16–18, 22, 26]. For example, McAfee [17]'s program *"If other player's code == my code: Cooperate; Else: Defect"* is a Nash equilibrium of the program game version of the one-shot Prisoner's Dilemma in which both players cooperate. (See also the literature on commitment games, e.g., Forges [9], Kalai et al. [13].) However, this literature focuses on the Nash equilibria of program games, rather than studying failure to coordinate on a Nash equilibrium as we do.

Coordination problems and equilibrium selection. There are large theoretical and empirical literatures on how agents might coordinate in complete information bargaining problems (see Schuessler and Van der Rijt [24] and references therein). Most closely related to this paper is the literature on whether communication before playing a simultaneous-move game can improve coordination [3, 7, 8, 10, 21]. Rabin [21] considers solution concepts for games with pre-play communication called *negotiated equilibrium (NGE)* and *negotiated rationalizability (NGR)*, where NGE assumes that players know their counterpart's strategies exactly (up to randomization). Rabin shows that under NGE players are guaranteed at least their PMM payoff in bargaining problems, whereas under NGR they are not. NGR is closer to the notion of subjective equilibrium used in our paper, which allows players to have possibly-inaccurate beliefs about what programs their counterparts will use. Santos [23] shows results analogous to Rabin [21]'s under cheap talk with alternating (rather than simultaneous) announcements. Finally, OC proposed safe Pareto improvements for mitigating inefficiencies from coordination failures. We discuss OC and its connections to the present work at greater length in Section 3.2.

3 MISCOORDINATION AND SAFE PARETO IMPROVEMENTS IN PROGRAM GAMES

In this section, we introduce the program games framework and subjective equilibrium, the solution concept that is our focus in this paper. Then we review OC's safe Pareto improvements, and show how they can be constructed in our setting using renegotiation. Section 5 contains a table summarizing the notation used in this section and Section 4. Throughout the paper, our formalism will be for games with two players, for ease of exposition. See supplementary material[1] for full proofs of our results in the more general n-player formalism. The extension to n players doesn't introduce qualitatively new challenges. Intuitively, since players submit programs independently of each other, we can apply the same arguments to the profile of counterparts for a given player, as we did to the single counterpart in the two-player case.

3.1 Setup: Program Games and Subjective Equilibrium

Two players $i = 1, 2$ will play a "base game" of complete information $G = (\mathcal{A} = A_1 \times A_2, (u_1, u_2))$. Let A_i be the set of possible actions for player i, and let $u_i(\boldsymbol{a})$ be player i's payoff in G when the players follow an action profile $\boldsymbol{a} = (a_1, a_2)$. Write $\mathbf{u}(\boldsymbol{a}) = (u_1(\boldsymbol{a}), u_2(\boldsymbol{a}))$, and refer to the set of payoff profiles attainable by some $\boldsymbol{a}$ in $\mathcal{A}$ as the **feasible set**. Throughout, we use the index j for the player $j \neq i$. For payoff profiles $\mathbf{x}$ and $\mathbf{y}$, write $\mathbf{x} \geq \mathbf{y}$ if $x_i \geq y_i$ for all i, and $\mathbf{x} > \mathbf{y}$ if $x_i > y_i$ for all i.

A program game $G(\mathcal{P})$ is a game in which a strategy is a program that maps the profile of other players' programs to an action in G.[2] This way, each player's program implements a commitment to an action conditional on the others' programs. Assume the action sets of G are continuous; this is practically without loss of generality, because our program game setting can be extended to a setting where players can use correlated randomization (see, e.g., Kalai et al. [13]). Here, $\mathcal{P} = P_1 \times P_2$, where P_i is a set of computable functions from P_j to A_i. We assume that all programs in P_i halt against all programs in P_j, for each i, as is standard in program game literature (see, e.g., Oesterheld [18], Oesterheld and Conitzer [19], Tennenholtz [26]). (Each P_i can be viewed as player i's "default" program set, which we will extend in Section 3.2 with a set of programs that have a special structure.)

Player i's program is $p_i \in P_i$. For a program profile $\mathbf{p} = (p_1, p_2)$, abusing notation, let the action profile played in the base game by players with a given program profile be $\boldsymbol{a}(\mathbf{p}) = (p_1(p_2), p_2(p_1))$. After all programs are simultaneously submitted, the induced action profile $\boldsymbol{a}(\mathbf{p})$ is played in G. Thus the payoff for player i in $G(\mathcal{P})$ resulting from the program profile $\mathbf{p}$ is $U_i(\mathbf{p}) = u_i(\boldsymbol{a}(\mathbf{p}))$.

To capture the possibility of miscoordination, we do not assume a Nash equilibrium is played. Instead, each player i has beliefs as to what program p_j the other player will use, distributed according to a probability distribution β_i (whose support may be a superset of P_j).[3] Then, a subjective equilibrium [14] is a profile of programs

[1] Available online at https://arxiv.org/abs/2403.05103.

[2] We restrict to deterministic programs for ease of exposition; the extension to probabilistic programs, as in, e.g., Kalai et al. [13], is straightforward.

[3] Allowing for β_i to be supported on a *superset* of P_j will be important when we consider extensions of players' program sets with SPIs in Section 3.2.

Table 1: Payoff matrix for the Scheduling Game

	Slot 1	Slot 2	Slot 3
Slot 1	3, 1	0, 0	0, 0
Slot 2	0, 0	1, 3	0, 0
Slot 3	0, 0	0, 0	1, 1

and beliefs such that each player's program maximizes expected utility with respect to their beliefs:

Definition 1. Let $\mathbf{p}^* = (p_1^*, p_2^*)$ and $\boldsymbol{\beta} = (\beta_1, \beta_2)$ be profiles of programs and beliefs, respectively, in $G(\mathcal{P})$. We say $(\mathbf{p}^*, \boldsymbol{\beta})$ is a **subjective equilibrium** of $G(\mathcal{P})$ if, for each i,

$$p_i^* \in \arg\max_{p_i \in P_i} \mathbb{E}_{p_j \sim \beta_i} U_i(\mathbf{p}).$$

Subjective equilibrium is, of course, a weaker solution concept than Nash equilibrium (or even rationalizable strategies [1, 20]). The results in this paper that follow will be stronger than showing that a given strategy is used in *some* subjective equilibrium. Instead, we will construct strategies such that, for *any* beliefs players might have under some assumptions, and *any* program profile they consider using, our strategies are individually (weakly) preferred by players over that program profile — and are thus used in a subjective equilibrium associated with those beliefs. Therefore, considering subjective equilibrium will make our results stronger than if we had assumed players' beliefs satisfied a Nash equilibrium assumption.

The base games we are interested in are *bargaining problems*, where players can miscoordinate in subjective equilibrium if they are sufficiently confident their counterparts will play favorably to them. This is possible even when players are capable of conditional commitments as in program games:

Example 3.1. (Miscoordination in subjective equilibrium) Suppose two principals delegate to AI assistants to negotiate on their behalf over the time for a meeting. Call this the Scheduling Game (Table 1). The principals meet if and only if the AIs agree on one of three possible time slots. Each principal i most prefers slot i, but would rather meet at slot 3 than not at all. Suppose each player i thinks j is sufficiently likely to use[4] p_j^C = *"Slot i if other player's code == 'always Slot i'; Else: Slot j"*. Intuitively, this program "demands" the player's best outcome, except against p_i^D = *"always Slot i"*, which exploits this program. Each player might believe the other is likely to use p_j^C because it can both exploit programs that yield to its demand and avoid miscoordinating with p_i^D. Then it is subjectively optimal for each player to submit p_i^D. The pair of programs (p_1^D, p_2^D) played in a subjective equilibrium under these beliefs results in the maximally inefficient (Slot 1, Slot 2) outcome.

3.2 Constructing Safe Pareto Improvements via Renegotiation

Informally, safe Pareto improvements (SPIs) [19] are transformations $\mathbf{f}$ of strategy profiles — in our case, program profiles $\mathbf{p}$ — such that, for any $\mathbf{p}$, all players are at least as well off under $\mathbf{f}(\mathbf{p})$ as under $\mathbf{p}$. OC focus on transformations induced by *payoff* transformations, and they formally define SPIs accordingly. However, they note that probability-1 Pareto improvements on players' default strategies can be achieved with other kinds of instructions besides having delegates play a game with transformed payoffs (see OC, pg. 14). Thus in this paper we define SPIs to be general transformations of strategy profiles that guarantee Pareto improvement:

Definition 2. For a program game $G(\mathcal{P})$, let $\mathbf{f} : \mathcal{P} \to \mathcal{P}'$ be a function of program profiles, written $\mathbf{f}(\mathbf{p}) = (f_1(p_1), f_2(p_2))$, for some joint program space $\mathcal{P}' = P_1' \times P_2'$.[5] Then $\mathbf{f}$ is an **SPI for** $G(\mathcal{P})$ if, for all program profiles $\mathbf{p}$, we have $\mathbf{U}(\mathbf{f}(\mathbf{p})) \geq \mathbf{U}(\mathbf{p})$; and for some program profile $\mathbf{p}$, there is some i' such that $U_{i'}(\mathbf{f}(\mathbf{p})) > U_{i'}(\mathbf{p})$.

A natural approach to constructing an SPI is to construct programs that, when they are all used against each other, map the action profile returned by default programs to a Pareto improvement whenever the default programs would have otherwise miscoordinated (i.e., the action profile is inefficient). We call this construction "renegotiation," and call mappings of action profiles to Pareto improvements "renegotiation functions."[6]

Definition 3. Call $\mathbf{rn} : \mathcal{A} \to \mathcal{A}$ a **renegotiation function** if:

(1) For every $\boldsymbol{a}$, $\mathbf{u}(\mathbf{rn}(\boldsymbol{a})) \geq \mathbf{u}(\boldsymbol{a})$.
(2) For some $\boldsymbol{a}$ and some i', $u_{i'}(\mathbf{rn}(\boldsymbol{a})) > u_{i'}(\boldsymbol{a})$.

And let $\mathcal{R}$ be the set of all renegotiation functions for the given game G.

We jointly define the spaces of **renegotiation programs** $P_i^{\mathrm{rn}}(\mathbf{rn}^i)$ for $i = 1, 2$ as those programs with the structure of Algorithm 1, for some:

- renegotiation function $\mathbf{rn}^i$ and
- "default program" $p_i^{\mathrm{def}} \in P_i \setminus P_i^{\mathrm{rn}}(\mathbf{rn}^i)$.

(Note that the definition of Algorithm 1 for a given player i references the sets of programs given by Algorithm 1 for the *other* player j, so this definition is not circular.) For any program profile $\mathbf{p} \in P_1^{\mathrm{rn}}(\mathbf{rn}^1) \times P_2^{\mathrm{rn}}(\mathbf{rn}^2)$ and any renegotiation function $\mathbf{rn}$, we write $\mathbf{p}^{\mathrm{def}} = (p_1^{\mathrm{def}}, p_2^{\mathrm{def}})$ and $\mathbf{rn}(\boldsymbol{a}) = (\mathrm{rn}_1(\boldsymbol{a}), \mathrm{rn}_2(\boldsymbol{a}))$.

Renegotiation programs work as follows: Consider the "default outcome," the action profile given by all players' default programs if they all use renegotiation programs (line 2). Against any program p_j such that the players' renegotiation functions (if any) don't all return the same Pareto improvement on the default outcome, $p_i \in P_i^{\mathrm{rn}}(\mathbf{rn}^i)$ plays according to its default program p_i^{def} (lines 6 and 8 in Algorithm 1). Against a program profile that is willing to renegotiate to the same Pareto improvement, however, p_i plays its part of the Pareto-improved outcome (line 4).

It is easy to see that any possible Pareto improvement (i.e., any possible mapping provided by a renegotiation function) can be implemented as an SPI via renegotiation programs:

Proposition 1. Let $\mathbf{rn}$ be a renegotiation function. For $i = 1, 2$, define $f_i : P_i \to P_i^{\mathrm{rn}}(\mathbf{rn})$ such that, for each $p_i \in P_i$, $f_i(p_i)$ is of the form given in Algorithm 1 with $f_i(p_i)^{\mathrm{def}} = p_i$. Then, the function $\mathbf{f} : \mathbf{p} \mapsto (f_1(p_1), f_2(p_2))$ is an SPI.

[4] Abusing notation, we write "p_i = ⟨pseudocode for p_i⟩" to describe programs p_i.

[5] The assumption above that programs halt against each other extends to $\mathcal{P}'$.

[6] Compare to section "Safe Pareto improvements under improved coordination" in OC.

Algorithm 1 Renegotiation program $p_i \in P_i^{\text{rn}}(\mathbf{rn}^i)$, for some p_i^{def}

Require: Counterpart program p_j

if $p_j \in P_j^{\text{rn}}(\mathbf{rn}^j)$ for some $\mathbf{rn}^j \in \mathcal{R}$ **then** ▷ Check that p_j renegotiates
  $\widehat{\boldsymbol{a}} \leftarrow \boldsymbol{a}(\mathbf{p}^{\text{def}})$
  if $\mathbf{rn}^i(\widehat{\boldsymbol{a}}) = \mathbf{rn}^j(\widehat{\boldsymbol{a}})$ **then**
    return $\text{rn}_i^i(\widehat{\boldsymbol{a}})$ ▷ Play renegotiation action
  else
    return $\widehat{a}_i$ ▷ Play default against others' defaults
else
  return $p_i^{\text{def}}(p_j)$ ▷ Play default

Proof. This follows immediately from the definitions of renegotiation function, Algorithm 1, and SPI. □

Example 3.2. (SPI using renegotiation) In Example 3.1, the players miscoordinated in the Scheduling Game. However, each player i might reason that, if they were to renegotiate with j, a renegotiation function that is fair enough to both players that they would both be willing to use it is: Map each outcome where players choose different slots to the symmetric (Slot 3, Slot 3) outcome. So, they could be better off using transformed versions of their defaults that renegotiate in this way.

3.3 Incentives to Renegotiate

When is the use of renegotiation guaranteed in subjective equilibrium in program games? SPIs by definition make all players (weakly) better off *ex post*, but it remains to show that players prefer to renegotiate *ex ante*. Intuitively, one might worry that players will choose not to accept a Pareto improvement in order to avoid losing bargaining power.

It is plausible that, all else equal, players prefer strategies that admit more opportunities for coordination. So, suppose players always prefer a renegotiation program over a non-renegotiation program if their expected utility is unchanged. Then Proposition 2 shows that, under a mild assumption on players' beliefs, each player always prefers to transform their default program into *some* renegotiation program. That is, each player i always prefers to use a program in $P_i^{\text{rn}} = \bigcup_{\mathbf{rn}} P_i^{\text{rn}}(\mathbf{rn})$ (so their program profile is in $\mathcal{P}^{\text{rn}} = P_1^{\text{rn}} \times P_2^{\text{rn}}$).

For this result, we assume (Assumption 4) the following holds for any program profile $\mathbf{p}$ given by (a) a program used by some player i in subjective equilibrium and (b) a program in the support of player i's beliefs: If the programs in $\mathbf{p}$ don't renegotiate with each other, then, a program should respond equivalently to any renegotiation program as it would respond to that program's default. This is because it seems implausible that players would respond differently to renegotiation programs that do not respond differently to them (in particular, "punish" renegotiation), all else equal.

Assumption 4. We say that players with beliefs $\boldsymbol{\beta}$ **are certain that renegotiation won't be punished** if the following holds. Take any renegotiation function $\mathbf{rn} \in \mathcal{R}$; any renegotiation program $p_i \in P_i^{\text{rn}}(\mathbf{rn})$; and any p_j in the support of β_i such that the programs in $\mathbf{p}$ don't renegotiate with each other. (I.e., there is no $\mathbf{rn}^j$ such that $p_j \in P_j^{\text{rn}}(\mathbf{rn}^j)$ where $\mathbf{rn}^j(\boldsymbol{a}(\mathbf{p}^{\text{def}})) = \mathbf{rn}(\boldsymbol{a}(\mathbf{p}^{\text{def}}))$.) Then:

(1) $p_j(p_i) = p_j(p_i^{\text{def}})$.
(2) If p_i^{def} is used in subjective equilibrium with respect to β_i, and $p_j \in P_j^{\text{rn}}$, we have $p_i^{\text{def}}(p_j) = p_i^{\text{def}}(p_j^{\text{def}})$.

Proposition 2. Let $G(\mathcal{P})$ be any program game. Let $\boldsymbol{\beta}$ be any belief profile satisfying the assumption that players are certain that renegotiation won't be punished (Assumption 4). And, for some arbitrary renegotiation function $\mathbf{rn}$, for each i and $p_i \in P_i$, let $f_i^*(p_i)$ be the program of the form in Algorithm 1 with $f_i^*(p_i)^{\text{def}} = p_i$. Then, for every subjective equilibrium $(\mathbf{p}^*, \boldsymbol{\beta})$ of $G(\mathcal{P} \cup \mathcal{P}^{\text{rn}})$ where $p_{i'}^* \notin P_{i'}^{\text{rn}}$ for some i', there exists $\mathbf{p}' \in \mathcal{P}^{\text{rn}}$ such that:

(1) For all i,

$$p_i' = \begin{cases} p_i^*, & \text{if } p_i^* \in P_i^{\text{rn}}(\mathbf{rn}^i) \text{ for some } \mathbf{rn}^i; \\ f_i^*(p_i^*), & \text{else.} \end{cases}$$

(2) $(\mathbf{p}', \boldsymbol{\beta})$ is a subjective equilibrium of $G(\mathcal{P} \cup \mathcal{P}^{\text{rn}})$.

Proof Sketch. For any non-renegotiation program for player i, construct a renegotiation program by letting this program be the default of Algorithm 1. If the other players' programs don't renegotiate to the same outcome as i's program, then i uses their default, so by the no-punishment assumption they achieve the same payoff as in the original subjective equilibrium. Otherwise, renegotiation Pareto-improves on the default, so the player is better off using the renegotiation program. □

4 THE SPI SELECTION PROBLEM AND CONDITIONAL SET-VALUED RENEGOTIATION

To Pareto-improve on the default outcome, the renegotiation programs defined in Section 3.2 require players to coordinate on the renegotiation function. So does renegotiation just reproduce the same coordination problem it was intended to solve? This is a general problem for SPIs, referred to by OC as the "SPI selection problem."[7]

Here, we argue that, although in part the players' initial bargaining problem recurs in SPI selection, players will always renegotiate so that each attains at least the worst payoff they can get in any efficient outcome. Following Rabin [21] we call the profile of these payoffs the **Pareto meet minimum (PMM)**. Player i's **Pareto meet projection (PMP)** (Fig. 1) maps each outcome to the set of Pareto improvements such that, first, each player's payoff is at least the PMM, and second, the payoff of $j \neq i$ is not increased except up to the PMM. We will prove our bound by arguing that if players attempt to negotiate a Pareto improvement on an outcome, they always at least weakly prefer to be willing to negotiate to the PMP of that outcome.

Definition 5. Let E be the set of Pareto-efficient action profiles in G. Then the **Pareto meet minimum (PMM)** payoff profile is

[7] OC give a brief informal characterization of an idea similar to our proposed partial solution to SPI selection (p. 39): "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." However, our approach does not require complying with whatever SPI the other player demands.

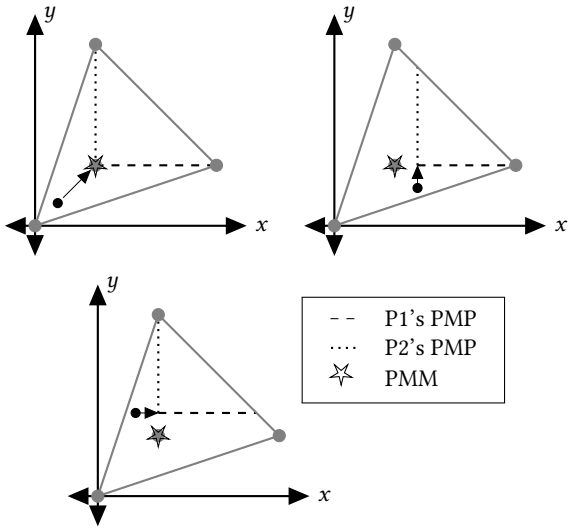

Figure 1: Illustration of the Pareto meet projection (PMP) of three different outcomes (black points) in the Scheduling Game, for each player. Gray points represent payoffs at each pure strategy profile. Each black point is mapped via a player's PMP (black arrows) to a set containing a) the "nearest" point in the Pareto meet and b) all points better for the given player and no better for the other player than (a).

$\mathbf{u}^{\text{PMM}} = (\min_{\boldsymbol{a}\in E} u_1(\boldsymbol{a}), \min_{\boldsymbol{a}\in E} u_2(\boldsymbol{a}))$. Player i's **Pareto meet projection (PMP)** of an action profile $\boldsymbol{a}$ is the set $\mathbf{PMP}_i(\boldsymbol{a})$ of action profiles $\tilde{\boldsymbol{a}}$ such that $u_i(\tilde{\boldsymbol{a}}) \geq \max\{u_i^{\text{PMM}}, u_i(\boldsymbol{a})\}$ and $u_j(\tilde{\boldsymbol{a}}) = \max\{u_j^{\text{PMM}}, u_j(\boldsymbol{a})\}$.

We'll start by giving an informal description of the algorithm we will use to prove the guarantee, called **conditional set-valued renegotiation (CSR)**. Next, we'll describe different components of the algorithm in more depth. Finally, we'll formally present the algorithm and the guarantee.

4.1 Overview of CSR

If players want to increase their chances of Pareto-improving via renegotiation, without necessarily accepting renegotiation outcomes that heavily favor their counterpart, they can report to each other *multiple* renegotiation outcomes they each would find acceptable and take Pareto improvements on which they agree. CSR implements such an approach. Like Algorithm 1, CSR involves default programs, and checks whether the default programs of a profile of CSR algorithms result in an efficient outcome. If not, CSR moves to a renegotiation procedure that works as follows:

(1) **Renegotiation using conditional sets.** At this stage, programs "announce" sets of points that Pareto-improve on the default and that they are willing to renegotiate to, conditional on the other player's program (see shaded regions in Fig. 2). If these sets overlap, the procedure continues to the second step; otherwise the players revert to their defaults.

The intuition for using *sets* at this stage is that (we will argue) this way a player can use a program that is willing to renegotiate to a payoff above their PMM payoff, without risking miscoordination if the other player does not also choose this new payoff precisely (see Fig. 2). Renegotiation sets that *condition* on the other player's renegotiation set function are crucial to the result that players are guaranteed their PMM payoff. This is because unconditionally adding an outcome to the renegotiation set might provide Pareto improvements against some possible counterpart program, but make the outcome worse against some *other* possible counterpart program (see Example 4.2).

(2) **Choosing a point in the agreement set.** Call the intersection of the sets players announced at the previous stage the "agreement set." At this stage, a "selection function" chooses an outcome from the Pareto frontier of the agreement set, which the players play instead of their miscoordinated default outcome. (Section 4.2 discusses how players coordinate on the selection function, without needing to solve a further bargaining problem.)

4.2 Components of CSR

4.2.1 Set-valued renegotiation. To avoid the need to coordinate on an exact renegotiation function, players can use functions that map miscoordinated outcomes to sets of Pareto improvements they each find acceptable. (In examples, we'll abuse terminology by referring to action profiles by their corresponding payoff profiles.) Then, we suppose the players follow some rule (a **selection function**) for choosing an efficient outcome from their agreement set.

Definition 6. Let $\mathbf{C}(\mathcal{A})$ be the set of closed subsets of $\mathcal{A}$.[8] Letting $\mathcal{R}^i$ be a set of functions from $\mathcal{R}^j \times \mathcal{A}$ to $\mathbf{C}(\mathcal{A})$, a function $\mathbf{RN}^i \in \mathcal{R}^i$ is a **set-valued renegotiation function** if, for all $\mathbf{RN}^j \in \mathcal{R}^j$:

(1) For all $\boldsymbol{a} \in \mathcal{A}$ and $\boldsymbol{a'} \in \mathbf{RN}^i(\mathbf{RN}^j, \boldsymbol{a})$, we have $\mathbf{u}(\boldsymbol{a'}) \geq \mathbf{u}(\boldsymbol{a})$.
(2) For some $\boldsymbol{a}$ and some $\boldsymbol{a'} \in \mathbf{RN}^i(\mathbf{RN}^j, \boldsymbol{a})$, we have $u_{i'}(\boldsymbol{a'}) > u_{i'}(\boldsymbol{a})$ for some i'.

A function $\mathbf{sel} : \mathbf{C}(\mathcal{A}) \to \mathcal{A}$ is a **selection function** if $\mathbf{sel}(S)$ is Pareto-efficient among points in S.[9] A selection function is **transitive** if, for all S, S' such that $\mathbf{u}(\mathbf{x}) \geq \mathbf{u}(\mathbf{sel}(S))$ for all $\mathbf{x} \in S'$, we have $\mathbf{u}(\mathbf{sel}(S \cup S')) \geq \mathbf{u}(\mathbf{sel}(S))$.

One might worry that by assuming a fixed selection function, we still haven't avoided the need for coordination. However, note that there is no *bargaining* problem involved in coordinating on a selection function. To see this, consider two players who intended to use renegotiation programs with different selection functions. Each player could switch to using a program that used the other player's selection function, and modify their set-valued renegotiation function so as to guarantee the same outcome as if the other player switched to *their* selection function. (See Appendix B in supplementary material[10] for a formal argument.) So the players are indifferent as to which selection function is used. (Coordinating on a selection function is a *pure* coordination problem, however; compare to the problem of coordinating on the programming language used in syntactic comparison-based program equilibrium [26].) In

[8] I.e., closed with respect to the topology on $\mathcal{A}$ induced by the Euclidean distance $d(\boldsymbol{a}, \boldsymbol{a'}) = ||\mathbf{u}(\boldsymbol{a}) - \mathbf{u}(\boldsymbol{a'})||$.
[9] Because each $S \in \mathbf{C}(\mathcal{A})$ is closed, some points in S are guaranteed to be Pareto-efficient among points in S.
[10] Available online at https://arxiv.org/abs/2403.05103.

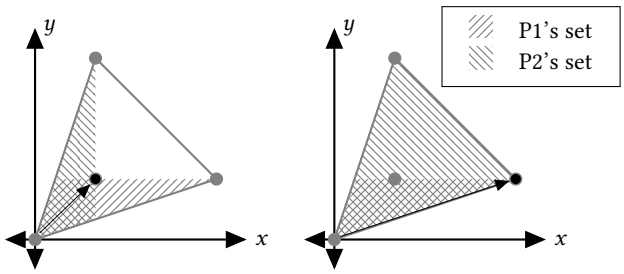

Figure 2: Set-valued renegotiation in the Scheduling Game, for two possible player 2 renegotiation sets. Black points represent renegotiation outcomes (mapped from the miscoordination outcome $(0, 0)$). If player 1 uses the renegotiation set shown here, they can achieve a Pareto improvement even if players don't reach the Pareto frontier (left), while still allowing for their best possible outcome (right).

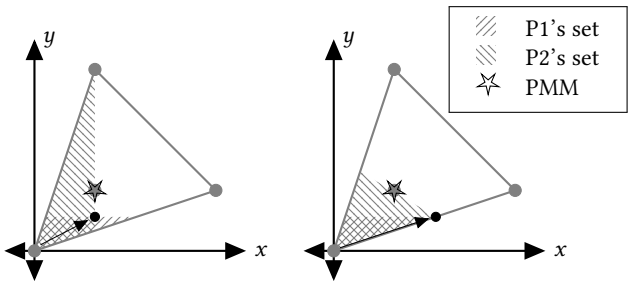

Figure 3: Two possible renegotiation procedures in the Scheduling Game, for different player 2 renegotiation sets. Player 1 might add the PMM (star) to their unconditional renegotiation set. In the case in the left plot, player 1 is no worse off by adding the PMM to their set. But in the case in the right plot, if player 1 adds the PMM, they might do worse if the selection function chooses the PMM instead of the black point that would have otherwise been achieved.

the results that follow, we will show that players can guarantee the PMM no matter which (transitive) selection function they use.

Example 4.1. (Set-valued renegotiation) Suppose players in the Scheduling Game (Table 1) miscoordinate at $\boldsymbol{a} = (0, 0)$. The two plots in Fig. 2 illustrate set-valued renegotiation for two possible player 2 renegotiation sets $\mathbf{RN}^2(\mathbf{RN}^1, \boldsymbol{a})$, and a fixed player 1 renegotiation set $\mathbf{RN}^1(\mathbf{RN}^2, \boldsymbol{a})$. Black points indicate the corresponding renegotiation outcomes. Player 1 thinks it's likely that the only efficient outcome player 2 is willing to renegotiate to is their own most preferred outcome (1, 3) (topmost gray point, left plot). But player 1 believes that with positive probability player 2's renegotiation set will also include player 1's most preferred outcome (3, 1) (black point, right plot). Player 1's best response given these beliefs may be to choose a set-valued renegotiation function $\mathbf{RN}^1$ that maps (0, 0) to a set including both (3, 1) and all outcomes Pareto-worse than (3, 1), i.e., the set depicted in Fig. 2. This way, they still achieve a Pareto improvement if player 2 has the smaller set (left plot), and get their best payoff if player 2 has the larger set (right plot).

4.2.2 Conditional renegotiation sets. We saw that renegotiation sets allow a player to achieve Pareto improvements against a wider variety of other players than is possible with renegotiation functions. However, suppose a player could not condition their renegotiation set on the other player's program. Then, by adding a point to their renegotiation set in attempt to Pareto-improve against some possible players, they might lock themselves out of a better outcome against other possible players.

Example 4.2. (Failure of unconditional renegotiation sets) Suppose that in the Scheduling Game, Player 1 uses an unconditional set-valued renegotiation function $\mathbf{RN}^1$. Fig. 3 shows their set $\mathbf{RN}^1(\mathbf{RN}^2, \boldsymbol{a})$ for the miscoordination outcome $\boldsymbol{a} = (0, 0)$. Suppose player 1 instead considers using $\mathbf{RN}^{1\prime}$ such that for all $\mathbf{RN}^2$, their renegotiation set is $\mathbf{RN}^{1\prime}(\mathbf{RN}^2, \boldsymbol{a}) = \mathbf{RN}^1(\mathbf{RN}^2, \boldsymbol{a}) \cup \{\mathbf{u}^{\text{PMM}}\}$. For both player 2 renegotiation sets shown in the figure, the players renegotiate to (i.e., the selection function chooses) the PMM (star). Then, player 1 is better off than under the default renegotiation outcome (black point) in the case in the left plot, but worse off in the case in the right plot. But if player 1 had access to conditional renegotiation sets, they could instead use an $\mathbf{RN}^{1\prime}$ that includes the PMM *only* against the player 2 set in the left plot.

4.2.3 Renegotiation sets that guarantee the PMM. How can a player guarantee a payoff better than some miscoordination outcome, without losing the opportunity to bargain for their most-preferred outcome? Suppose player i considers using a set-valued renegotiation function that doesn't guarantee the PMM. That is, against some counterpart program, the resulting outcome $\boldsymbol{a}$ is worse for at least one player than their least-preferred efficient outcome (i.e., their PMM payoff). Then:

(1) As we will argue in Theorem 3, under mild assumptions, player i is no worse off also including $\mathbf{PMP}_i(\boldsymbol{a})$ in their renegotiation set. So, if player j also follows the same incentive to include $\mathbf{PMP}_j(\boldsymbol{a})$ in their renegotiation set, these programs will guarantee at least the PMM. (Notice that players' ability to guarantee the PMM depends on conditional renegotiation sets, for the reasons discussed in Example 4.2.)
(2) On the other hand, if i thinks the selection function might not choose their optimal outcome in the agreement set, i will *not* prefer to include outcomes strictly better for player j than those in $\mathbf{PMP}_i(\boldsymbol{a})$. For example, in Fig. 2, $\mathbf{RN}^1(\mathbf{RN}^2, (0, 0))$ includes all outcomes Pareto-worse than $\mathbf{PMP}_1((0, 0))$. This set safely guarantees the PMM against a player who also uses a set of this form, and gives player 1 their best possible outcome against the $\mathbf{RN}^2$ in the right plot. But if $\mathbf{RN}^1(\mathbf{RN}^2, (0, 0))$ in the right plot included additional outcomes, which would be worse for player 1 than player 1's best possible outcome, the selection function might choose an outcome that is worse for player 1 than otherwise. (This is why, when we construct strategies for the proof of Theorem 3, we add the entire PMP even though it is sufficient to only add the point in the PMP that minimizes the player's payoff.)

Finally, here is the formal definition of CSR programs. For a set-valued renegotiation function $\mathbf{RN}^i$, we define the space of CSR

programs $P_i^{\text{RN}}(\mathbf{RN}^i)$ as the space of programs with the structure of Algorithm 2, for some default p_i^{def}. (Let $\mathcal{R}^{\text{RN}}$ be the set of all set-valued renegotiation functions, and for each i let the space of all CSR programs be $P_i^{\text{RN}} = \bigcup_{\mathbf{RN}^i \in \mathcal{R}^{\text{RN}}} P_i^{\text{RN}}(\mathbf{RN}^i)$. Let $\mathcal{P}^{\text{RN}} = P_1^{\text{RN}} \times P_2^{\text{RN}}$.) The selection function **sel** is given to the players (and we suppress dependence of P_i^{RN} on **sel** for simplicity).

Algorithm 2 Conditional set-valued renegotiation program $p_i \in P_i^{\text{RN}}(\mathbf{RN}^i)$, for some p_i^{def}

Require: Counterpart program p_j

if $p_j \in P_j^{\text{RN}}(\mathbf{RN}^j)$ for some $\mathbf{RN}^j \in \mathcal{R}^{\text{RN}}$ **then**
  $\widehat{\boldsymbol{a}} \leftarrow \boldsymbol{a}(\mathbf{p}^{\text{def}})$
  $I \leftarrow \mathbf{RN}^1(\mathbf{RN}^2, \widehat{\boldsymbol{a}}) \cap \mathbf{RN}^2(\mathbf{RN}^1, \widehat{\boldsymbol{a}})$ ▷ Agreement set
  if $I \neq \emptyset$ **then**
    $\widehat{\boldsymbol{a}} \leftarrow \mathbf{sel}(I)$ ▷ Renegotiation outcome
  return $\widehat{a}_i$ ▷ Play renegotiation outcome, or default
else
  return $p_i^{\text{def}}(p_j)$

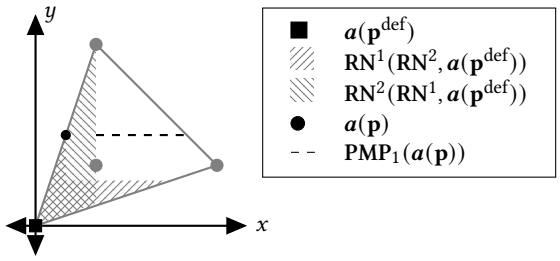

Figure 4: Illustration of the argument for Theorem 3. By default, the renegotiation outcome is the black circle, $\boldsymbol{a}(\mathbf{p})$. Player 1 considers whether to add to their renegotiation set $\mathbf{RN}^1(\mathbf{RN}^2, \boldsymbol{a}(\mathbf{p}^{\text{def}}))$ the black striped segment $\mathbf{PMP}_1(\boldsymbol{a}(\mathbf{p}))$. Player 1 is certain that player 2 would not change their set $\mathbf{RN}^2(\mathbf{RN}^1, \boldsymbol{a}(\mathbf{p}^{\text{def}}))$ in response to this addition in a way that would make player 1 worse off (Assumption 8). This is because the only change player 1 has made is to add outcomes that make both players weakly better off than $\boldsymbol{a}(\mathbf{p})$ and do not make player 2 strictly better off.

4.3 Guaranteeing PMM Payoffs Using CSR

Similar to the assumption in Section 3.3, suppose players always include more outcomes in their renegotiation sets if their expected utility is unchanged. So in particular, to show that in subjective equilibrium players use programs that guarantee at least the PMM, it will suffice to show that they weakly prefer these programs.

Then, we will show in Theorem 3 that under mild assumptions on players' beliefs, for any program that does not guarantee a player at least their PMM payoff, there is a corresponding CSR program the player prefers that *does* guarantee their PMM payoff. We prove this result by constructing programs identical to the programs players would otherwise use, except that these new programs' renegotiation sets for each outcome include their PMP of the outcome they would have otherwise achieved. For a program p_i, we call this modified program the **PMP-extension** of p_i (Definition 7).

This result requires two assumptions on players' beliefs and the structure of programs used in subjective equilibrium (Assumptions 8i and 8ii), analogous to Assumption 4 of Proposition 2:

(1) Assumption 8i is equivalent to Assumption 4 applied to CSR programs rather than renegotiation programs: For any program used in subjective equilibrium or in the support of a player's beliefs, if that program never renegotiates, it responds identically to counterpart CSR programs as to their defaults.
(2) Informally, Assumption 8ii says that players believe that, with probability 1: If a CSR program is modified only by adding PMP points to its renegotiation set, the only changes the counterparts would prefer to make are those that also add PMP points. The intuition for this assumption is: For any possible default renegotiation outcome, the PMP-extension, by definition, doesn't add any points that make the counterpart strictly better off than that outcome while making the focal player worse off (see Fig. 4). So, similar to Assumption 8i, the counterpart doesn't have an incentive to make changes to their renegotiation set that would make the focal player worse off. (This argument wouldn't work if player i also added outcomes that are *better* for j than their PMP-extension. This is because, as noted in the previous section, j would then have an incentive to exclude i's most-preferred outcome from j's renegotiation set.)

For Theorem 3 we also assume the selection function is transitive. This is an intuitive property: If outcomes are added to the agreement set that make *all* players weakly better off than the default renegotiation outcome, the new renegotiation outcome should be weakly better for all players.

The remainder of this subsection provides the formal details for the statement of Theorem 3, and a sketch of the proof.

Definition 7. For any $p_i \in P_i^{\text{RN}}(\mathbf{RN}^i)$ for some $\mathbf{RN}^i$, the **PMP-extension** $\widetilde{p}_i \in P_i^{\text{RN}}(\widetilde{\mathbf{RN}}^i)$ is the program identical to p_i except: for all $p_j \in P_j^{\text{RN}}(\mathbf{RN}^j)$ for some $\mathbf{RN}^j$, writing $\widetilde{\mathbf{p}}^i = (\widetilde{p}_i, p_j)$, we have

$$\widetilde{\mathbf{RN}}^i(\mathbf{RN}^j, \boldsymbol{a}(\widetilde{\mathbf{p}}^{i^{\text{def}}})) = \mathbf{RN}^i(\mathbf{RN}^j, \boldsymbol{a}(\widetilde{\mathbf{p}}^{i^{\text{def}}})) \cup \mathbf{PMP}_i(\boldsymbol{a}(\mathbf{p})).$$

Assumption 8. We say that players with beliefs $\boldsymbol{\beta}$ **are (i) certain that CSR won't be punished and (ii) certain that PMP-extension won't be punished** if the following hold:

(i) Suppose either p_i is in a subjective equilibrium of $G(\mathcal{P} \cup \mathcal{P}^{\text{RN}})$, or p_i is in the support of β_j. Suppose $p_i \notin P_i^{\text{RN}}$. Then for any $p_j \in P_j^{\text{RN}}$, we have $p_i(p_j) = p_i(p_j^{\text{def}})$.
(ii) Let $p_j \in P_j^{\text{RN}}(\mathbf{RN}^j)$ be in the support of β_i, and take any $p_i \in P_i^{\text{RN}}(\mathbf{RN}^i)$ with PMP-extension $\widetilde{p}_i$. For all $\boldsymbol{a}$, we have that $\mathbf{RN}^j(\widetilde{\mathbf{RN}}^i, \boldsymbol{a}) = \mathbf{RN}^j(\mathbf{RN}^i, \boldsymbol{a}) \cup V$ for some $V \subseteq \mathbf{PMP}_i(\boldsymbol{a}(\mathbf{p}))$.

Theorem 3. Let $G(\mathcal{P})$ be a program game, and **sel** be any transitive selection function. Suppose the action sets of G are continuous, so that for any $\boldsymbol{a} \in \mathcal{A}$, player i's PMP of that action profile $\mathbf{PMP}_i(\boldsymbol{a})$ is nonempty. Let $\boldsymbol{\beta}$ be any belief profile satisfying the assumption that players are (i) certain that CSR won't be punished and (ii) certain that PMP-extension won't be punished (Assumption 8).

Then, for any subjective equilibrium $(\mathbf{p}, \boldsymbol{\beta})$ of $G(\mathcal{P} \cup \mathcal{P}^{\text{RN}})$ where $U_i(\mathbf{p}) < u_i^{\text{PMM}}$ for some i, there exists $\mathbf{p}'$ such that:

(1) For all i, p_i' is the PMP-extension of p_i.
(2) $\mathbf{U}(\mathbf{p}') \geq \mathbf{u}^{\text{PMM}}$.
(3) $(\mathbf{p}', \boldsymbol{\beta})$ is a subjective equilibrium of $G(\mathcal{P} \cup \mathcal{P}^{\text{RN}})$.

Proof Sketch. Assumption 8i implies that players always use CSR programs. Consider any renegotiation outcome $\boldsymbol{a}$ worse for some player than the PMM, which is achieved by player i's "old" program against some counterpart. By Assumption 8ii, player j doesn't punish i for adding their PMP of that outcome, $\mathbf{PMP}_i(\boldsymbol{a})$, to their renegotiation set (in their "new" program). So the renegotiation outcome of the new program against j is only different from that of the old program if j is also willing to renegotiate to some outcome in $\mathbf{PMP}_i(\boldsymbol{a})$. But in that case, because the selection function is transitive, the new renegotiation outcome is no worse for i than under the old program. Therefore, each player always prefers to replace a given program with its PMP-extension, and when all players use PMP-extended programs, the Pareto frontier of their agreement set only includes outcomes guaranteeing each player their PMM payoff. □

Remark: Notice that the argument above does not require that players refrain from using programs that implement other kinds of SPIs, besides PMP-extensions. First, the PMP-extension can be constructed from any default program, including, e.g., a CSR program whose renegotiation set is only extended to include the player's best outcome, not their PMP (call this a "self-favoring extension"). And if a player's final choice of program is their self-favoring extension, they are still incentivized to use the PMP-extension within their default program.

Second, while it is true that an analogous argument to the proof of Theorem 3 could show that a player is weakly better off *ex ante* using a self-favoring extension than not extending their renegotiation set at all, this does not undermine our argument. This is because, as we claimed at the start of this section, it is reasonable to assume that among programs with equal expected utility, each player prefers to also include their PMP. But wouldn't the player also prefer an even larger renegotiation set that includes outcomes that Pareto-dominate the PMM as well? No, because those outcomes will be worse for that player *and* better for their counterpart than the player's most-preferred outcome, such that the counterpart would have an incentive to make the player worse off (i.e., it's plausible that Assumption 8ii would be violated).

We can now formalize the claim that CSR is an SPI that partially solves SPI selection: The mapping from programs $\mathbf{p}$ to instances of Algorithm 2 with$\mathbf{p}$ as defaults, for *any profile* $(\mathbf{RN}^1, \mathbf{RN}^2)$ used in subjective equilibrium under the assumptions of Theorem 3, is an SPI that guarantees players their PMM payoffs.

Proposition 4. For $i = 1, 2$, for some selection function $\mathbf{sel}$, define $f_i^{\mathbf{RN}^i} : P_i \to P_i^{\text{RN}}(\mathbf{RN}^i)$ such that, for each $p_i \in P_i$, $f_i^{\mathbf{RN}^i}(p_i)$ is of the form given in Algorithm 2 with $f_i^{\mathbf{RN}^i}(p_i)^{\text{def}} = p_i$. Then, under the assumptions of Theorem 3, for any $(\mathbf{RN}^1, \mathbf{RN}^2)$, $\mathbf{p}^{\text{def}}$ such that for all $\mathbf{RN}^j$, $\mathbf{PMP}_i(\boldsymbol{a}(\mathbf{p})) \subseteq \mathbf{RN}^i(\mathbf{RN}^j, \boldsymbol{a}(\mathbf{p}))$:

(1) The function $\mathbf{f}^{\mathbf{RN}} : \mathbf{p} \mapsto (f_1^{\mathbf{RN}^1}(p_1), f_2^{\mathbf{RN}^2}(p_2))$ is an SPI.
(2) For all i, $U_i(\mathbf{f}^{\mathbf{RN}}(\mathbf{p}^{\text{def}})) \geq \max\{U_i(\mathbf{p}^{\text{def}}), u_i^{\text{PMM}}\}$.

Proof. This follows immediately from the argument used to prove Theorem 3. □

Table 2: Key notation

Symbol	*Description (page introduced)*
$\boldsymbol{a}(\mathbf{p})$	action profile in the base game played by players with the given program profile (2)
$\mathbf{rn}^i$	renegotiation function for player i (maps an action profile to a Pareto-improved action profile) (3)
$\mathbf{RN}^i$	set-valued renegotiation function for player i (maps j's set-valued renegotiation function and an action profile to a set of Pareto-improved action profiles) (5)
$\mathcal{R}, \mathcal{R}^{\text{RN}}$	sets of all renegotiation functions and set-valued renegotiation functions, respectively (3, 6)
$P_i^{\text{rn}}(\mathbf{rn}^i)$	set of renegotiation programs (Algorithm 1) that use the renegotiation function $\mathbf{rn}^i$ (3)
$P_i^{\text{RN}}(\mathbf{RN}^i)$	set of conditional set-valued renegotiation programs (Algorithm 2) that use the set-valued renegotiation function $\mathbf{RN}^i$ (6)
p_i^{def}	default program for a program p_i in P_i^{rn} or P_i^{RN} (3)
$\mathbf{sel}$	selection function (maps a set of action profiles to an action profile that is efficient within that set) (5)
$\mathbf{u}^{\text{PMM}}$	Pareto meet minimum (5)
$\mathbf{PMP}_i$	Pareto meet projection for player i (maps an action profile to a particular set of Pareto-improved action profiles) (5)

5 DISCUSSION

Using renegotiation to construct SPIs in program games is a rich and novel area, with many directions to explore. To name a few:

- Which plausible conditions would violate our assumptions about players' beliefs used for the PMM guarantee?
- What do *unilateral* SPIs [19] look like in this setting?
- When are SPIs used in sequential, rather than simultaneous-move, settings? In particular, in sequential settings, the first-moving player's decision whether to use a renegotiation program could signal private information to the second-moving player.
- We have assumed complete information about payoffs; using ideas from DiGiovanni and Clifton [6]'s framework for program games in the presence of private information, it should also be possible to construct SPIs in incomplete information settings.
- How can this theory inform real-world AI system design?

ACKNOWLEDGMENTS

For valuable comments and discussions, we thank our anonymous referees, Caspar Oesterheld, Lukas Finnveden, Vojtech Kovarik, and Alex Kastner.

REFERENCES

[1] B. Douglas Bernheim. 1984. Rationalizable Strategic Behavior. *Econometrica* 52, 4 (1984), 1007–28.

[2] Vincent Conitzer and Caspar Oesterheld. 2023. Foundations of Cooperative AI. *AAAI* 37, 13 (June 2023), 15359–15367. https://doi.org/10.1609/aaai.v37i13.26791

[3] Vincent P Crawford. 2017. Let's talk it over: Coordination via preplay communication with level-k thinking. *Research in Economics* 71, 1 (2017), 20–31.

[4] Andrew Critch. 2019. A parametric, resource-bounded generalization of Löb's theorem, and a robust cooperation criterion for open-source game theory. *The Journal of Symbolic Logic* 84, 4 (2019), 1368–1381.

[5] Allan Dafoe, Edward Hughes, Yoram Bachrach, Tantum Collins, Kevin R. McKee, Joel Z. Leibo, Kate Larson, and Thore Graepel. 2020. Open Problems in Cooperative AI. arXiv:2012.08630 [cs.AI]

[6] Anthony DiGiovanni and Jesse Clifton. 2023. Commitment games with conditional information disclosure. *Proceedings of the AAAI Conference on Artificial Intelligence* 37, 5 (2023), 5616–5623.

[7] Joseph Farrell. 1987. Cheap talk, coordination, and entry. *The RAND Journal of Economics* (1987), 34–39.

[8] Joseph Farrell and Matthew Rabin. 1996. Cheap talk. *Journal of Economic perspectives* 10, 3 (1996), 103–118.

[9] Françoise Forges. 2013. A folk theorem for Bayesian games with commitment. *Games and Economic Behavior* 78 (2013), 64–71. https://doi.org/10.1016/j.geb.2012.11.004

[10] Simin He, Theo Offerman, and Jeroen van de Ven. 2019. The power and limits of sequential communication in coordination games. *Journal of Economic Theory* 181 (2019), 238–273.

[11] Jobst Heitzig, Jörg Oechssler, Christoph Pröschel, Niranjana Ragavan, and Yat Long Lo. 2023. Improving International Climate Policy via Mutually Conditional Binding Commitments. arXiv:2307.14267 [cs.CY] https://arxiv.org/abs/2307.14267

[12] J. V. Howard. 1988. Cooperation in the Prisoner's Dilemma. *Theory And Decision* 24, 3 (May 1988), 203–213. https://doi.org/10.1007/BF00148954

[13] Adam Tauman Kalai, Ehud Kalai, Ehud Lehrer, and Dov Samet. 2010. A commitment folk theorem. *Games and Economic Behavior* 69, 1 (2010), 127–137.

[14] Ehud Kalai and Ehud Lehrer. 1995. Subjective games and equilibria. *Games Econom. Behav.* 8, 1 (Jan. 1995), 123–163. https://doi.org/10.1016/S0899-8256(05)80019-3

[15] Frank H. Knight. 1921. *Risk, Uncertainty, and Profit.* Houghton Mifflin Company, Boston, MA, USA.

[16] Patrick LaVictoire, Benja Fallenstein, Eliezer Yudkowsky, Mihaly Barasz, Paul Christiano, and Marcello Herreshoff. 2014. Program Equilibrium in the Prisoner's Dilemma via Löb's Theorem. In *Workshops at the Twenty-Eighth AAAI Conference on Artificial Intelligence*.

[17] R. Preston McAfee. 1984. Effective Computability in Economic Decisions. Mimeo, University of Western Ontario.

[18] Caspar Oesterheld. 2019. Robust program equilibrium. *Theory and Decision* 86, 1 (2019), 143–159.

[19] Caspar Oesterheld and Vincent Conitzer. 2022. Safe Pareto improvements for delegated game playing. In *Autonomous Agents and Multi-Agent Systems*, Vol. 36. Springer, 1–47.

[20] David G. Pearce. 1984. Rationalizable Strategic Behavior and the Problem of Perfection. *Econometrica* 52, 4 (1984), 1029–50.

[21] Matthew Rabin. 1991. A Model of Pre-game Communication. *Journal of Economic Theory* 63 (1991), 370–391.

[22] Ariel Rubinstein. 1998. *Modeling Bounded Rationality.* The MIT Press.

[23] Vasco Santos. 2000. Alternating-announcements cheap talk. *Journal of Economic Behavior & Organization* 42, 3 (July 2000), 405–416. https://doi.org/10.1016/S0167-2681(00)00094-9

[24] Rudolf Schuessler and Jan-Willem Van der Rijt. 2019. *Focal points in negotiation.* Springer.

[25] Xinyuan Sun, Davide Crapis, Matt Stephenson, Barnabé Monnot, Thomas Thiery, and Jonathan Passerat-Palmbach. 2023. Cooperative AI via Decentralized Commitment Devices. arXiv:2311.07815 [cs.AI] https://arxiv.org/abs/2311.07815

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

[27] Hal R Varian. 2010. Computer mediated transactions. *American Economic Review* 100, 2 (2010), 1–10.