Rational coordination with no communication or conventions

Valentin Goranko[*] Stockholm University Sweden

Antti Kuusisto[†] University of Bremen Germany

Raine Rönnholm[‡] University of Tampere Finland

Abstract

We study pure coordination games where in every outcome, all players have identical payoffs, 'win' or 'lose'. We identify and discuss a range of 'purely rational principles' guiding the reasoning of rational players in such games and analyse which classes of coordination games can be solved by such players with no preplay communication or conventions. We observe that it is highly nontrivial to delineate a boundary between purely rational principles and other decision methods, such as conventions, for solving such coordination games.

1 Introduction

Coordination games ([11]) are strategic form games which have several pure strategy Nash equilibria with the same or comparable payoffs for every player, and where all players have the mutual interest to select one of these equilibria. In *pure coordination games* ([11]), aka *games of common payoffs* ([12]), all the players in the game receive the same payoffs and thus all payers have fully aligned preferences to coordinate in order to reach the best possible outcome for everyone. Here we study one-step *pure win-lose coordination games* (WLC games) in which all payoffs are either 1 (i.e., *win*) or 0 (i.e., *lose*).

Clearly, if players can communicate when playing a pure coordination game with at least one winning outcome, then they can simply agree on a winning strategy profile, so the game is trivialised. What makes such games non-trivial is the limited, or no possibility of communication before the game is presented to the players. In this paper we assume *no preplay communication*[1] *at all*, meaning that the players must make their choices by reasoning individually, without any contact with the other players before (or during) playing the game.

There are many natural real-life situations where such coordination scenarios occur. For example, (A) two cars driving towards each other on a narrow street such that they can avoid a collision by swerving either to the right or to the left. Or, (B) a group of n people who get separated in a city and they must each decide on a place where to get together ('regroup'), supposing they do not have any way of contacting each other.

[*] valentin.goranko@philosophy.su.se

[†] kuusisto@uni-bremen.de

[‡] raine.ronnholm@uta.fi

[1] Note that, unlike the common use of 'preplay communication' in game theory to mean communication before the given game is played, here we mean communication before the players are even presented with the game.

Even if no preplay communication is possible, players may still share some *conventions* ([8], [11], [18]) which they believe everyone to follow. In (A), a collision could be avoided by using the convention (or rule) that cars should always swerve to the right (or, to the left). In (B), everyone could go to a famous meeting spot in the city, e.g., the main railway station. Conventions need not be explicit agreements, but they can also naturally emerge as so-called *focal points*, for example. The theory of focal points, originating from Schelling [17], has been further developed in the context of coordination games, e.g. in [13], [20].

In this paper we assume that the players share no conventions, either. Thus, in our setting players play independently of each other. They can be assumed to come from completely different cultures, or even from different galaxies, for that matter. However, we assume that *it is common belief* among the players that:

(1) *every player knows the structure of the game*;

(2) *all players have the same goal*, viz. selecting together a winning profile;

Initially in this paper we will only assume *individual rationality*, i.e. that every player acts with the aim to win the game. Later we will assume in addition *common belief in rationality*, i.e. that every player is individually rational and that it is commonly believed amongst all players that every player is rational.

Our main objective is to analyse what kinds of reasoning can be accepted as 'purely rational' and what kinds of WLC games can be solved by such reasoning. Thus, we try to identify '*purely rational principles*' that *every* rational player ought to follow in *every* WLC game. We also study the hierarchy of such principles based on classes of WLC games that can be won by following different reasoning principles. It is easy to see that coordination by pure rationality is not possible in the example situations (A) and (B) above. However, we will see that there are many natural pure coordination scenarios in which it seems clear that rational players can coordinate successfully.

One of the principal findings of our study is that it is highly nontrivial to demarcate the "purely rational" principles from the rest[2]. Indeed, this seems to be an open-ended question and its answer depends on different background assumptions. Still, we identify a hierarchy of principles that can be regarded as rational and we also provide justifications for them. However, these justifications have varying levels of common acceptability and a more in-depth discussion would be needed to settle some of the issues arising there. Due to space constraints, a more detailed discussion on these issues is deferred to a follow-up work.

Coordination and rationality are natural and interesting topics that have been studied in various contexts in, e.g., [3], [6], [7], [8], [19]. We note the close conceptual relationship of the present study with the notion of *rationalisability* of strategies [2], [5], [15], which is particularly important in epistemic game theory. We also mention two recent relevant works related to logic to which the observations and results in the present paper could be directly applied: in [10], two-player coordination games were related to a variant of *Coalition Logic*[3], and in [1], coordination was analysed with respect to the game-theoretic semantics of *Independence Friendly Logic*.

In addition to the theoretical work presented here, we also run some empirical experiments on people's behaviour in certain (2-player) WLC games. One of our tests can be accessed from the link given in [9]. We suggest to the reader to do that test before reading further this paper and then to do it again after reading the paper.

[2]Schelling shares this view on pure coordination games (see [17], pg. 283, ftn. 16).

[3]In fact, the initial motivation for the present work came from concerns with the semantics of Alternating time temporal logic ATL, extending Coalition Logic.

2 Pure win-lose coordination games

2.1 The setting

A *pure win-lose coordination game* G is a strategic form game with n players $(1, \ldots, n)$ whose available *choices* (*moves*, *actions*) are given by sets $\{C_i\}_{i \leq n}$. The set of winning *choice profiles* is presented by an n-ary *winning relation* W_G. For technical convenience and simplification of some definitions, we present these games as *relational structures* (see, e.g., [4]). A formal definition follows.

Definition 2.1. An n-player **win-lose coordination game** (WLC **game**) is a relational structure $G = (A, C_1, \ldots, C_n, W_G)$ where A is a finite domain of **choices**, each $C_i \neq \emptyset$ is a unary predicate, representing the choices of player i, s.t. $C_1 \cup \cdots \cup C_n = A$, and W_G is an n-ary relation s.t. $W_G \subseteq C_1 \times \cdots \times C_n$. Here we also assume that the players have pairwise disjoint choice sets, i.e., $C_i \cap C_j = \emptyset$ for every $i, j \leq n$ s.t. $i \neq j$. A tuple $\sigma \in C_1 \times \cdots \times C_n$ is called a **choice profile** for G and the choice profiles in W_G are called **winning**.

We use the following terminology for any WLC game $G = (A, C_1, \ldots, C_n, W_G)$.

- The **losing relation of** G is the relation $L_G := C_1 \times \cdots \times C_n \setminus W_G$. A choice profile $\sigma \in L_G$ is called **losing**.

- The **complementary game of** G is the game $\overline{G} := (A, C_1, \ldots, C_n, L_G)$.

- Let $A_i \subseteq C_i$ for every $i \leq n$. The **restriction** of G to $(A_1, \ldots, A_n)$ is the game $G \upharpoonright (A_1, \ldots, A_n) := (A_1 \cup \cdots \cup A_n,\ A_1, \ldots, A_n,\ W_G \upharpoonright A_1 \times \cdots \times A_n)$.

- For every choice $c \in C_i$ of a player i, the **winning extension of** c **in** G is the set $W^i_G(c)$ of all tuples $\tau \in C_1 \times \cdots \times C_{i-1} \times C_{i+1} \times \cdots \times C_n$ such that the choice profile obtained from τ by adding c to the i-th position is winning. We define the **losing extension of** c **in** G analogously.

- A choice $c \in C_i$ of a player i is **(surely) winning**, respectively **(surely) losing**, if it is guaranteed to produce a winning (respectively losing) choice profile regardless of what choices the other player(s) make. Note that c is a winning choice iff $W^i_G(c) = C_1 \times \cdots \times C_{i-1} \times C_{i+1} \times \cdots \times C_n$. Similarly, c is a losing choice iff $W^i_G(c) = \emptyset$.

- A choice $c \in C_i$ is **at least as good as** (respectively, **better than**) a choice $c' \in C_i$ if $W^i_G(c') \subseteq W^i_G(c)$ (respectively, $W^i_G(c') \subsetneq W^i_G(c)$). A choice $c \in C_i$ is **optimal** for a player i if it is at least as good as any other choice of i.

Note that a choice $c \in C_i$ is better than a choice $c' \in C_i$ precisely when c *weakly dominates* c' in the usual game-theoretic sense (see e.g. [12], [16]), and a choice $c \in C_i$ is an optimal choice of player i when it is a weakly dominant choice. Note also that c *strictly dominates* c' (*ibid.*) if and only if c is surely winning and c' is surely losing. Thus, strict domination is a too strong concept in WLC games. Also the concept of *Nash equilibrium* is not very useful here.

Example 2.2. We present here a 3-player coordination story which will be used as a running example hereafter. The three robbers Casper, Jesper and Jonathan[4] are

[4]This example is based on the children's book *When the Robbers Came to Cardamom Town* by Thorbjørn Egner, featuring the characters Casper, Jesper and Jonathan.

planning to quickly steal a cake from the bakery of Cardamom Town while the baker is out. They have two possible plans to enter the bakery: either (a) to break in through the front door or (b) to sneak in through a dark open basement. For (a) they need a *crowbar* and for (b) a *lantern*. The baker keeps the cake on top of a high cupboard, and the robbers can only reach it by using a *ladder*.

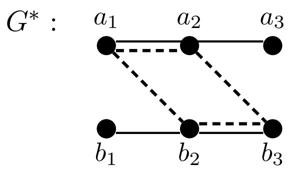
When approaching the bakery, Casper is carrying a crowbar, Jesper is carrying a ladder and Jonathan is carrying a lantern. However, the robbers cannot agree whether they should follow plan (a) or plan (b). While the robbers are quarreling, suddenly Constable Bastian appears and the robbers all flee to different directions. After this the robbers have to individually decide whether to go to the front door (by plan (a)) or to the basement entrance (by plan (b)). They must do the right decision fast before the baker returns.

The scenario we described here can naturally be modeled as a WLC game. We relate Casper, Jesper and Jonathan with players 1, 2 and 3, respectively. Each player i has two choices a_i and b_i which correspond to either going to the front door or to the basement entrance, respectively. The robbers succeed in obtaining the cake if both Casper and Jesper go to the front door (whence it does not matter what Jonathan does). Or, alternatively, they succeed if both Jonathan and Jesper go to the basement (whence the choice of Casper is irrelevant). Hence this coordination scenario corresponds to the following WLC game $G^* = (\{a_1, b_1, a_2, b_2, a_3, b_3\}, C_1, C_2, C_3, W_{G^*})$, where for each player i, $C_i = \{a_i, b_i\}$ and $W_{G^*} = \{(a_1, a_2, a_3), (a_1, a_2, b_3), (a_1, b_2, b_3), (b_1, b_2, b_3)\}$. (For a graphical presentation of this game, see Example 2.3 below.)

2.2 Presenting WLC games as hypergraphs

The n-ary winning relation W_G of an n-player WLC game G defines a *hypergraph* on the set of all choices. We give visual presentations of hypergraphs corresponding to WLC games as follows: The choices of each player are displayed as columns of nodes starting from the choices of player 1 on the left and ending with the column with choices of player n. The winning relation consists of lines that go through some choice of each player[5]. This kind of graphical presentation of a WLC game G will be called a *game graph (drawing) of* G. (Note that game graphs of 2-player WLC games are simply bipartite graphs.)

Example 2.3. The WLC game G^* in Example 2.2 has the following game graph:



We now define several simple types of WLC games and introduce a uniform notation for them. Since the names of choices are not given for these games, each game given here actually corresponds to a class of games with the same structure. However, in this paper we usually consider all these games to be equivalent[6]. Let $k_1, \ldots, k_n \in \mathbb{N}$.

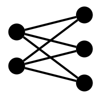
- $G(k_1 \times \cdots \times k_n)$ is the n-player WLC game where the player i has k_i choices and the winning relation is the *universal relation* $C_1 \times \cdots \times C_n$.

[5]In pictures these lines can be drawn in different styles or colours, to tell them apart.

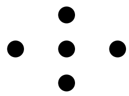
[6]If a player reasons by pure rationality, the names of the choices should not have an effect on that player's reasoning. We will discuss further this issue later on.

- $G(\overline{k_1 \times \cdots \times k_n})$ is the n-player WLC game where the player i has k_i choices and the winning relation is the *empty relation*. Note that with this notation we have $G(\overline{k_1 \times \cdots \times k_n}) = \overline{G(k_1 \times \cdots \times k_n)}$. Some examples:

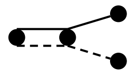
$G(2 \times 3)$



$G(\overline{1 \times 3 \times 1})$

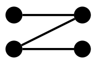


$G(1 \times 1 \times 2)$

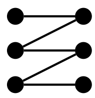


- Let $k \in \mathbb{N}$. We write $G(Z_k)$ for the *2-player* WLC game in which both players have k choices and the winning relation forms a single path that goes through all the choices (see below for an example). Similarly, $G(O_k)$, where $k \geq 2$, denotes the 2-player WLC game where the winning relation forms a $2k$-cycle that goes through all the choices. These are exemplified by the following:

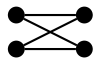
$G(Z_2)$



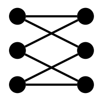
$G(Z_3)$



$G(O_2) = G(2 \times 2)$

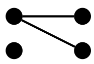


$G(O_3)$

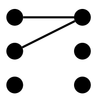


- Suppose that $G(A)$ and $G(B)$ have been defined, both having the same number of players. Then $G(A + B)$ is the *disjoint union* of $G(A)$ and $G(B)$, i.e., the game obtained by assigning to each player a disjoint union of her choices in $G(A)$ and $G(B)$, and where the winning relation for $G(A + B)$ is the union of the winning relations in $G(A)$ and $G(B)$. Some examples:

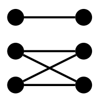
$G(1 \times 2 + 1 \times 0)$



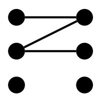
$G(2 \times 1 + \overline{1 \times 2})$



$G(1 \times 1 + 2 \times 2)$

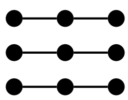


$G(Z_2 + \overline{1 \times 1})$

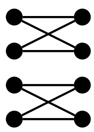


- Let $m \in \mathbb{N}$. Then $G(mA) := G(A + \cdots + A)$ (m times). Examples:

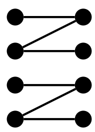
$G(3(1 \times 1 \times 1))$



$G(2(2 \times 2))$



$G(2Z_2)$



- Recall our "regrouping scenario" (B) from the introduction. If there are n people in the group and there are m possible meeting spots in the city, then the game is of the form $G(m(1^n))$, where $1^n := 1 \times \cdots \times 1$ (n times).

2.3 Symmetries of WLC games and structural protocols

A **protocol** is a mapping Σ that assigns to every pair (G, i), where G is a WLC game and i a player in G, a nonempty set $\Sigma(G, i) \subseteq C_i$ of choices. Thus a protocol gives global nondeterministic strategy for playing any WLC game in the role of any player. Intuitively, a protocol represents a global mode of acting in any situation that involves playing WLC games. Hence, protocols can be informally regarded as global "reasoning styles" or "behaviour modes". Thus, a protocol can also be identified with an agent who acts according to that protocol in all situations that involve playing different WLC games in different player roles.

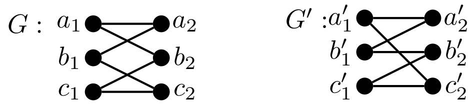
Assuming a setting based on pure rationality with no special conventions or preplay communication, a protocol will only take into account the *structural properties* of the game and its winning relation. Thus the names of the choices and the names (or ordering) of the players should be of no relevance. In this section we make this issue precise.

Definition 2.4. An isomorphism[7] between games G and G' is called a **choice-renaming**. An automorphism of G is called a **choice-renaming of** G.

Let $G = (A, C_1, \ldots, C_n, W_G)$ be a WLC game. For a player i, we say that the choices $c, c' \in C_i$ are **i-equivalent**, denoted by $c \simeq_i c'$, if there is a choice-renaming of G that maps c to c'. For each $i \leq n$, the relation $\simeq_i$ is an equivalence relation on the set C_i. We denote the equivalence class of $c \in C_i$ by $[\![c]\!]_i$.

Supposing that a player i does not use names or labels of her choices (or she has no preferences over them), then she should be *indifferent* about the choices that are in the same equivalence class.

Example 2.5. Let $A = \{a_1, b_1, c_1, a_2, b_2, c_2\}$ and $A = \{a'_1, b'_1, c'_1, a'_2, b'_2, c'_2\}$. Consider the WLC games G and G' whose game graphs are given below.



A function $\pi : A \to A'$, which maps b_2 to c'_2, c_2 to b'_2 and c to c' for all the other choices $c \in A$, is choice-renaming from G to G'. Note that actually both G and G' are of the form $G(O_3)$. A function that maps a_i to b_i, b_i to c_i and c_i to a_i (for $i \in \{1, 2\}$) is a choice-renaming of G. Therefore $a_1 \simeq_1 b_1 \simeq_1 c_1$ and $a_2 \simeq_2 b_2 \simeq_2 c_2$.

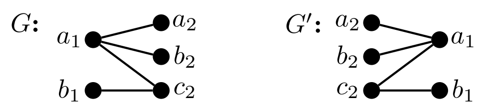
Definition 2.6. Consider n-player WLC games $G = (A, C_1, \ldots, C_n, W_G)$ and $G' = (A, C'_1, \ldots, C'_n, W'_G)$. A permutation $\beta : \{1, ..., n\} \to \{1, ..., n\}$ is called a **player-renaming** between G and G' if the following conditions hold:

(1) $C_{\beta(i)} = C'_i$ for each $i \leq n$.

(2) $W'_G = \{\, (c_{\beta(1)}, \ldots, c_{\beta(n)}) \mid (c_1, \ldots, c_n) \in W_G \,\}$.

If there is a player-renaming between two WLC games, the games are essentially the same, the only difference being the ordering of the players. Furthermore, the game graph of G' is simply obtained by permuting the columns of the game graph of G.

Example 2.7. Consider the following WLC games:



A permutation β, which swaps 1 and 2, is a player-renaming between G and G'.

[7] Isomorphism is defined as usual for relational structures (see, e.g., [4]).

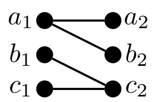
Definition 2.8. Consider WLC games G and G'. A pair (β, π) is a **full renaming** between G and G' if there is a WLC game G'' such that β is a player-renaming between G and G'' and π is a choice-renaming between G'' and G'. If $G = G'$, we say that (β, π) is a **full renaming of** G. We say that choices $c \in C_i$ and $c' \in C_j$ in the same game are **structurally equivalent**, denoted by $c \sim c'$, if there is a full renaming (β, π) of G such that $\beta(i) = j$ and $\pi(c) = c'$. It is quite easy to see that $\sim$ is an equivalence relation on the set A of all choices. We denote the equivalence class of a choice c by $[c]$.

We also make the following observations:

- If $c \simeq_i c'$ for some i, then also $c \sim c'$.
- Suppose that there is a tuple $G_1, \ldots, G_n$ of WLC games such that for any i there is either a choice-renaming or a player-renaming between G_i and G_{i+1}. Then it is easy to show that there is a full renaming from G_1 to G_n.

Example 2.9. Consider a WLC game of the form $G(1 \times 2 + 2 \times 1)$.

$G(1 \times 2 + 2 \times 1)$:



Let β be the permutation which swaps (players) 1 and 2, and let π be the bijection $\{(a_1, c_2), (b_1, b_2), (c_1, a_2), (a_2, c_1), (b_2, b_1), (c_2, a_1)\}$. Now the pair (β, π) is a full renaming of $G(1 \times 2 + 2 \times 1)$. It is easy to see that $\simeq_1$ has the equivalence classes $\{a_1\}$ and $\{b_1, c_1\}$, and similarly, $\simeq_2$ has the equivalence classes $\{c_2\}$ and $\{a_2, b_2\}$. Furthermore, $\sim$ has the equivalence classes $\{a_1, c_2\}$ and $\{b_1, c_1, a_2, b_2\}$. Likewise, in the game G^* from Example 2.2 the relation $\sim$ has the equivalence classes $\{a_1, b_3\}$, $\{b_1, a_3\}$, $\{a_2, b_2\}$.

We say that a protocol Σ is **structural** if it is "indifferent" with respect to full renamings, which means that, given any WLC games G, G' for which there exists a full renaming (β, π) between G and G', for any i and any choice $c \in C_i$, it must hold that $c \in \Sigma(G, i)$ iff $\pi(c) \in \Sigma(G', \beta(i))$. Intuitively, this reflects the idea that when following a structural protocol, one acts independently of the names of choices and names (or ordering) of player roles. Thus, following a structural protocol, one cannot tell the difference between choices that are structurally equivalent. Hereafter, unless otherwise specified, we only consider structural protocols.

It is worth noting that if we considered a framework where WLC games were presented so that the names of the choices and players could be used to define an ordering (of the players and their choices), things would trivialize because it would be easy to win all games by the prenegotiated agreement to always choose the lexicographically least tuple from the winning relation.

3 Purely rational principles in WLC games

By a **principle** we mean any nonempty class of protocols. Intuitively, these are the protocols "complying" with that principle. If protocols are regarded as "reasoning styles" (or "behaviour modes"), then principles are *properties* of such reasoning styles (or behaviour modes). Principles that contain only structural protocols are called **structural principles.**

A player i **follows a principle P** in a WLC game G if she plays according to some protocol in P. We are mainly interested in structural principles which describe "purely rational" reasoning that involves neither preplay communication nor conventions and which are rational to follow in *every* WLC game. Such principles will be called **purely rational principles**. Intuitively, purely rational principles should always be followed by all rational players. Consider:

$$\begin{aligned} \mathrm{P}_1 &:= \{\Sigma \mid \Sigma(G,i) \text{ does not contain any surely losing choices when } W_G \neq \emptyset\}, \\ \mathrm{P}_2 &:= \{\Sigma \mid \Sigma(G,i) \text{ contains all choices } c \in C_i \text{ such that } |W_G^i(c)| \\ &\qquad \text{is a prime number. If there are no such choices, } \Sigma(G,i) = C_i.\}. \end{aligned}$$

If player i follows P_1, then she always uses some protocol which does not select surely losing choices, if possible. This seems a principle that any rational agent would follow. If player i follows P_2, then she always plays choices whose degree (in the game graph) is a prime number, if possible. Note that both principles are structural, but P_1 can be seen as a purely rational principle, while P_2 seems arbitrary; it could possibly be some seemingly odd convention, for example.

We say that a **principle P solves** a WLC game G (or G **is P-solvable**), if G is won whenever every player follows some protocol that belongs to P. Formally, this means that $\Sigma_1(G,1) \times \cdots \times \Sigma_n(G,n) \subseteq W_G$ for all protocols $\Sigma_1, \ldots, \Sigma_n \in \mathrm{P}$. The class of all P-solvable games is denoted by $s(\mathrm{P})$.

In this paper we try to identify (a hierarchy of) principles that can be considered to be purely rational and analyse the classes of games that they solve.

3.1 Basic individual rationality

Hereafter we describe principles by the properties of protocols that they determine. We begin by considering the case where players are individually rational, but there is no common knowledge about this being the case. It is safe to assume that any individually rational player would follow at least the following principle.

Fundamental individual rationality (FIR):
Never play a strictly dominated choice.[8]

As noted before, strict domination is a very weak concept with WLC games. Following FIR simply means that a player should never prefer a losing choice to a winning one. Therefore FIR is a very weak principle that can solve only some quite trivial types of games such as $G(1 \times 2 + 1 \times 0)$ (See figure 1). In general, FIR-solvable games have a simple description: at least one of the players has (at least one) winning choice, and all non-winning choices of that player are losing. Thus, for example all games of the form $G(k \times l + m \times 0)$ are FIR-solvable. FIR has two natural strengthenings which can be considered purely rational:

1. **Non-losing principle (NL):** Never play a losing choice, if possible.

2. **Sure winning principle (SW):** Always play a winning choice, if possible.

Since losing choices cannot be winning choices, these principles can naturally be put together (by taking the intersection of these principles):

Basic individual rationality (BIR): $\mathrm{NL} \cap \mathrm{SW}$.

When following BIR, a player plays a winning choice if she has one, and else she plays a non-losing choice. We make the following observations. See the corresponding pictures in Figure 1.

[8]Recall, that a choice a is strictly dominated by a choice b if the choice b guarantees a strictly higher payoff than the choice a in every play of the game (see e.g. [12], [16]).

1. NL and SW do not imply each other and neither of them follows from FIR. This can be seen by the following examples.
   - The game $G(1 \times 1 + \overline{1 \times 1})$ is NL-solvable but not SW-solvable. This is because neither of the players has a winning choice, but the non-losing moves form a cartesian product.
   - The game $G(Z_2)$ is SW-solvable but not NL-solvable. This is because both players have a winning choice, but there are no losing choices. Note that in this game both players can force winning and thus they both would be sure of winning even without knowing that the other player follows SW.
2. FIR-solvable games are solvable by *both* SW and NL. This is because in FIR-solvable games, at least one player i has a winning choice and all the other choices of that player are losing. Hence by following either SW or NL, the player i will select a winning choice.
3. Every BIR-solvable game is *either* NL or SW-solvable. This is because a BIR-solvable game G is won when every player selects a winning choice, if they have one, or else if they each play a non-losing choice. If at least one player has a winning choice in G, then it is SW-solvable, else it is NL-solvable.

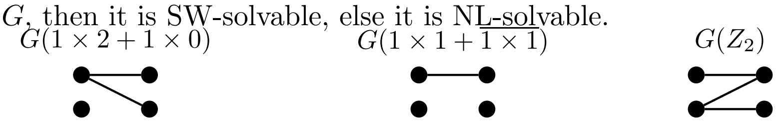
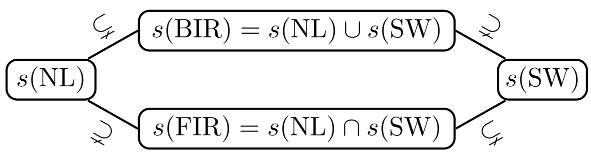


Figure 1:

Therefore we can see that the sets of games solvable by FIR, NL, SW, BIR form the following lattice:



SW-solvable and NL-solvable games have simple descriptions: In SW-solvable games, at least one player has a surely winning choice. In NL-solvable games, the winning relation forms a nonempty *Cartesian product* between all non-losing choices. BIR-solvable games have (at least) one of these two properties.

Note that in order to follow BIR, the players do not make any assumptions on the behavior or rationality of each other. In fact, the players do not even need to know that everyone has a mutual goal in the game; that is, following BIR would be equally rational even in the games that are *not cooperative*.

3.2 Common beliefs in rationality and iterated reasoning

In contrast to individual rationality, collective rationality allows players to make assumptions on each other's rationality. Let P be a (purely rational) principle. When *all players believe that everyone follows* P, they can reason as follows:

($\star$) Suppose that by following P each player i must play a choice from $A_i \subseteq C_i$ (that is, A_i is the smallest set such that $\Sigma(G, i) \subseteq A_i$ for every $\Sigma \in$ P). By this assumption, the players may collectively assume that the game that is played is actually $G' := G \upharpoonright (A_1, \ldots, A_n)$, and therefore all P-compliant protocols should only prescribe choices in G'.

If players have *common belief* in P being followed, then the reasoning ($\star$) above can be repeated for the game G' and this iteration can be continued until a fixed point is reached. By $\mathsf{cir}(\mathrm{P})$ we denote the principle of **collective iterated reasoning of P** which prescribes that P is followed in the reduced game obtained by the iterated reasoning of ($\star$). Note that after every iteration of ($\star$), the sets of choices for each player become smaller (or stay the same). And since each protocol in any principle P must give nonempty set of choices for any WLC game, $\mathsf{cir}(\mathrm{P})$ cannot make the set of choices empty for any player. By these observations it is easy to see that $s(\mathrm{P}) \subseteq s(\mathsf{cir}(\mathrm{P}))$ for any principle P.

When considering principles of *collective* rationality, we will apply collective iterated reasoning. It may be debated whether such reasoning counts as purely rational, so a question arises: if P is a purely rational principle, is $\mathsf{cir}(\mathrm{P})$ always purely rational as well? For the lack of space we will not discuss this issue here. We note, however, the extensive literature relating common beliefs and knowledge with individual and collective rationality, see e.g. [5], [11], [14], [21].

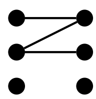
3.3 Basic collective rationality

Here we extend individually rational principles of Section 3.1 by adding common belief in the principles (as described in Section 3.2) to the picture. We first analyse what happens with principles NL and SW. It is easy to see that the collective iterated reasoning of NL reaches a fixed point in a single step by simply removing the losing choices of every player. Hence $s(\mathrm{NL}) = s(\mathsf{cir}(\mathrm{NL}))$. Collective iterated reasoning of SW also reaches a fixed point in a single step by eliminating all non-winning choices of every player who has a winning choice. But if even one player has a winning choice, then the game is already SW-solvable. Therefore $s(\mathrm{SW}) = s(\mathsf{cir}(\mathrm{SW}))$.

However, assuming common belief in BIR, some games which are not BIR-solvable may become solvable. See the following example.

Example 3.1. The game $G(Z_2 + \overline{1 \times 1})$ cannot be solved with NL or SW. However, if the players can assume that neither of them selects a losing choice (by NL) and eliminate those choices from the game, then they (both) have a winning choice in the reduced game and can win in it by SW.

$G(Z_2 + \overline{1 \times 1})$:



Thus, we define the following principle:

Basic collective rationality (BCR): $\mathsf{cir}(\mathrm{BIR})$.

The above example shows that $s(\mathrm{BIR}) \subsetneq s(\mathrm{BCR})$, i.e. BCR is *stronger* than BIR. The games solvable by BCR have the following characterisation: *after removing all surely losing choices of every player, at least one of the players has a surely winning choice.* It is worth noting that common belief in SW is not needed for solving games with BCR because a *single* iteration of $\mathsf{cir}(\mathrm{NL})$ suffices. Thus, players could solve BCR-solvable games simply by following BIR and believing that everyone follows NL. We also point out that the principle BCR is equivalent to the principle applied in [10] for Strategic Coordination Logic.

3.4 Principles using optimal choices

If a rational player has optimal choices (that are at least as good as all other choices), it is natural to assume that she selects such a choice. Note that players may have several optimal choices or none at all. For example, in game $G(2 \times 2)$ both players have two optimal choices while in $G(Z_3)$ neither of the players has optimal choices. We now introduce the following principle:

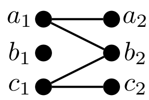
Individual optimal choices (IOC): Play an optimal choice, if possible.

Example 3.2. Recall the WLC game G^* from Example 2.2. For Casper (who is carrying the crowbar) it is a better choice to go to the front door than to the basement. Likewise, for Jonathan (who is carrying the lantern) it is a better choice to go to the basement than to the front door. Therefore the choice a_1 is (the only) optimal choice for player 1 and b_3 is (the only) optimal choice for the player 3. The player 2 (Jesper) does not have any optimal choices, but if both 1 and 3 play their optimal choices, then the game is won regardless of the choice of 2. Therefore, the game G^* is solvable with IOC. But since no player has winning or losing choices in this game, it is easy to see that it is not BCR-solvable.

Note that if a player has winning choices, then the set of optimal choices is the set of winning choices, and therefore IOC $\subseteq$ SW. From the description of BIR-solvable games, we see that they are also IOC-solvable. The next example shows that IOC is *incomparable* with BCR (with respect to solvable games).

Example 3.3. Consider a WLC game G^* from Example 2.2. In this game, none of the players has winning or losing choices, and therefore it is not solvable with BCR. However, both player 1 and 3 have an optimal choice a_1 and b_3, respectively. If 1 and 3 play a optimal choice, then the game is won regardless of the choice of player 2. Hence $G^* \in s(\text{IOC})$.

Consider the following WLC game G_Σ.

G_Σ:


By following BCR, player 1 chooses either a_1 or c_1 and player 2 chooses b_2, whence the game is won. However, G_Σ is not solvable with IOC since player 1 does not have any optimal choices (and may thus end up choosing the losing choice b_1).

As we saw above, if a player does not have optimal choices, following only IOC might lead to playing a losing choice. In order to avoid pathological cases like this, we should at least add NL to IOC.

Improved basic individual rationality (BIR$^+$): IOC $\cap$ NL

Since IOC $\subseteq$ SW, we have BIR$^+ \subseteq$ BIR. Note that, unlike BCR, the principle BIR$^+$ is only based on *individual* reasoning. However, BIR$^+$ is nevertheless stronger than BCR as shown by the following proposition.

Proposition 3.4. $s(\text{BCR}) \subsetneq s(\text{BIR}^+)$.

Proof. Suppose first that $G \in s(\text{BCR})$. Then player i has a nonempty set A_i of winning choices in the reduced game after removing all losing choices of all the other players. But now every choice in A_i must be an optimal choice of i in the original game G. Hence, by following BIR$^+$, the player i will play a choice from A_i (by IOC) while all the other players play a non-losing choice (by NL), whence the game is won. Therefore $G \in s(\text{BIR}^+)$ and thus $s(\text{BCR}) \subseteq s(\text{BIR}^+)$. In Example 3.3 we saw that G^* is solvable with IOC but not with BCR. Therefore $s(\text{BCR}) \subsetneq s(\text{BIR}^+)$. □

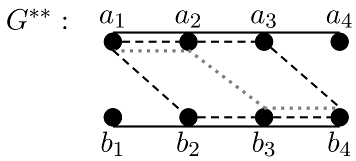
We now consider the collective version of IOC:

Collective optimal choices (COC): cir(IOC)

Proposition 3.5. $s(\text{BIR}^+) \subsetneq s(\text{COC})$.

Proof. We first show that $s(\text{BIR}^+) \subseteq s(\text{COC})$. Suppose that WLC game G is BIR$^+$-solvable, i.e. the game is won when every player plays an optimal choice, if they have any, else they play a non-losing choice. Let G' be the game that is obtained after the first collective iteration of IOC. Now all the remaining non-losing choices of every player in G' must be winning choices. Since winning choices are also optimal choices, all losing choices are eliminated in the second iteration of cir(IOC). After this, all combinations of the remaining choices are winning. Thus the game is won by following COC.

Consider the following WLC game G^{**}.



Here only players 1 and 4 have optimal choices a_1 and b_4, respectively, and no player has losing choices. Hence we see that G cannot be solved with BIR$^+$. (By following BIR$^+$, players may end up selecting the choice profile (a_1, b_2, a_3, b_4) which is not winning.) However, after the first iteration of cir(IOC), the players 2 and 3 have optimal choices a_2 and b_3, respectively. Hence, by following COC, the players end up choosing a winning choice profile (a_1, a_2, b_3, b_4).

Note that we can construct a similar game for $2n$ players, where it takes n iterations of cir(IOC) for solving it. □

Finally, let us consider what happens in the special case of 2-player WLC games. We first observe that *the only optimal choices in 2-player* WLC *game G (where $W_G \neq \emptyset$) are those that are winning against all non-surely losing choices of the other player.* Consequently, when considering 2-player WLC games, we have $s(\text{IOC}), s(\text{COCs}) \subseteq s(\text{BCR})$. By combining this with the results of Propositions 3.4 and 3.5, we obtain the following result.

Proposition 3.6. *For 2-player* WLC *games:* $s(\text{BCR}) = s(\text{BIR}^+) = s(\text{COC})$.

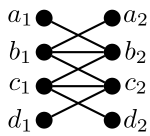
3.5 Elimination of weakly dominated choices

Usually in game-theory, rationality is associated with the elimination of strictly or weakly dominated strategies. As noted in Section 3.1, *strict* domination is a too strong concept for WLC games. Weak domination, on the other hand, gives the following principle when applied individually.

Individually rational choices (IRC): Do not play a choice a when there is a better choice b available, i.e., if $W^i_G(a) \subsetneq W^i_G(b)$, then i does not play a.

Note that by the definition, IRC $\subseteq$ NL $\cap$ IOC and therefore $s(\text{BIR}^+) \subseteq s(\text{IRC})$. The inclusion here is proper since there are WLC games that are solvable with IRC but not with BIR$^+$. See the following example.

Example 3.7. Consider the following WLC game $G_{\#}$.

$G_{\#}$: 

In $G_{\#}$ none of the players has losing choices nor optimal choices and therefore it cannot be solved with BIR$^+$. But the choice b_1 is better than a_1 and likewise c_1 is better than d_1. (Note that b_1 and c_1 are not comparable with each other.) Therefore, by following IRC, the player on will play either b_1 or c_1. With a symmetric reasoning the player 2 will play either b_2 or c_2, which leads to win. Therefore $G^{\#} \in s(\text{IRC})$.

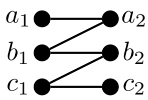
The COC-solvable game G^{**} (in the proof of Proposition 3.5) is unsolvable with IRC. (This is because neither of the moves a_2 and b_2 (respectively a_3 and b_3) is better than the other.) On the other hand, the game $G_{\#}$ in Example 3.7 is unsolvable with COC, and therefore IRC is incomparable with COC in the general case. However, in the 2-player case $s(\text{COC}) \subsetneq s(\text{IRC})$, since then $s(\text{COC}) = s(\text{BIR}^+)$ by Proposition 3.6.

We next assume common belief in IRC. As commonly known (see e.g. [14]), *iterated elimination of weakly dominated strategies* eventually stabilises in some reduced game but different elimination orders may produce different results. However, when applying cir(IRC), the process will stabilise to a unique reduced game since all weakly dominated choices are always removed *simultaneously.* By following the next principle, players will play a choice within this reduced game.

Collective rational choices (CRC): cir(IRC)

The following example shows that $s(\text{IRC}) \subsetneq s(\text{CRC})$.

Example 3.8. Consider the following WLC game.

$G(Z_3)$: 

We first note that b_1 is a better choice than a_1 and likewise b_2 is a better choice than c_2. Therefore, by following IRC, player 1 will play a choice from $\{b_1, c_1\}$ and player 2 will play from $\{a_2, b_2\}$, which does not guarantee winning. However, after eliminating a_1 and c_2, then b_1 is better than c_1 and b_2 is better than a_2. Thus, by following CRC and doing one more iteration of cir(IRC), player 1 and 2 have only the choices b_1 and b_2 which are winning.

In $G(Z_3)$, we needed two iterations of cir(IRC). It is easy to see that in the game $G(Z_4)$ the iterations are done analogously and the fixed point is reached in 3 iterations. Furthermore, we can see that $n-1$ iterations of cir(IRC) are needed for solving $G(Z_n)$. Therefore different numbers of iterations of cir(IRC) form a proper hierarchy of CRC-solvable 2-player WLC games.

3.6 Symmetry-based principles

By only following the concept of rationality from game-theory, one could argue that CRC reaches the border of purely rational principles. However, we now define more principles which are incomparable with CRC but can still be regarded as purely rational. These principles are based on *symmetries* in WLC games and the assumption that players follow only structural protocols is central here.

We begin with auxiliary definitions. We say that a choice profile $(c_1, \dots, c_n)$ **exhibits a bad choice symmetry** if $[\![c_1]\!]_1 \times \cdots \times [\![c_n]\!]_n \not\subseteq W_G$ (recall Definition 2.4), and that a choice c **generates a bad choice symmetry** if σ_c exhibits bad choice symmetry for *every* choice profile σ_c that contains c.

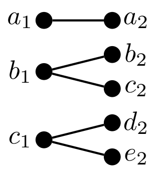
Elimination of bad choice symmetries (ECS):

Never play choices that generate a bad choice symmetry, if possible.

Why should this principle be considered rational? Suppose that a player i plays a choice c_i which generates a bad choice symmetry. It is now possible to win only if some tuple $(c_1, \dots, c_{i-1}, c_i, c_{i+1}, \dots, c_n) \in W_G$ is eventually chosen. However, the players have *exactly the same reason* (based on structural principles) to play so that any other tuple in $[\![c_1]\!]_1 \times \cdots \times [\![c_n]\!]_n$ is selected, and such other tuple may possibly be a losing one since $[\![c_1]\!]_1 \times \cdots \times [\![c_n]\!]_n \not\subseteq W_G$.

Here is a typical example of using ECS. Suppose that the game graph of G has two (or more) connected components that are isomorphic to each other. Since no player can see a difference between those components, all players should avoid playing choices from them. See the following example.

Example 3.9. Consider the WLC game $G(1 \times 1 + 2(1 \times 2))$:



In this game $b_1 \simeq_1 c_1$ and $b_2 \simeq_2 c_2 \simeq_2 d_2 \simeq_2 e_2$. Since all the choice profiles in $\{b_1, c_1\} \times \{b_2, c_2, d_2, e_2\}$ are not winning, we see that both b_1 and c_1 generate a bad choice symmetry. Likewise, b_2, c_2, d_2 and e_2 generate a bad choice symmetry. Therefore, by following ECS, the players will choose a_1 and a_2.

While ECS only considers symmetries between similar choices, the next principle takes symmetries *between players* into account. Consider a choice profile $\vec{c} = (c_1, \dots, c_n)$ and let $S_i^p(\vec{c}) := \{c_i\} \cup (C_i \cap \bigcup_{j \neq i} [c_j])$ for each i (recall Definition 2.8). We say that $(c_1, ..., c_n)$ **exhibits a bad player symmetry** if $S_1^p(\vec{c}) \times \cdots \times S_n^p(\vec{c}) \not\subseteq W_G$ and a choice c **generates a bad player symmetry** if σ_c exhibits a bad player symmetry for every choice profile σ_c that contains c.

Elimination of bad player symmetries (EPS):

Never play choices that generate bad player symmetries, if possible.

Here the players assume that all players reason similarly, or alternatively, each player wants to play so that she would at least coordinate with herself in the case she was to use her protocol to make a choice in each player role of a WLC game. Suppose that the players have some reasons to select a choice profile $(c_1, \dots, c_n)$. Now, if there are players $i \neq j$ and a choice $c'_j \in C_j$ such that $c'_j \sim c_i$, then the player j should have the same reason to play c'_j as i has for playing c_i. Hence, if the players have their reasons to play $(c_1, \dots, c_n)$, they should have the same reasons to play any choice profile in $S_1^p(\vec{c}) \times \cdots \times S_n^p(\vec{c})$. Winning is not guaranteed if $S_1^p(\vec{c}) \times \cdots \times S_n^p(\vec{c}) \not\subseteq W_G$.

Example 3.10. Consider EPS in the case of a two-player game WLC game G. If for a given choice $c \in C_1$, there is a structurally equivalent choice $c' \in C_2$ such that $(c, c') \notin W_G$, then by following EPS, player 1 does not play the choice c (and likewise player 2 does not play the choice c'). With this kind of reasoning, some CRC-unsolvable games like $G(1 \times 1 + 1 \times 2 + 2 \times 1)$ become solvable.

Note also that the game G^* (recall Example 2.2) is EPS-solvable since both choices b_1 and a_3 generate a bad player symmetry.

Example 3.11. In Example 3.8 we showed that in order to solve $G(Z_n)$ by CRC it takes $n-1$ collective iterations and after that the "middle choices" are selected by both of the players. The game $G(Z_n)$ can also be solved by ELR and the players will end up choosing the same choices as with CRC. This is because every other choice—except the middle choice—generates a bad player symmetry.

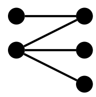
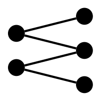
Finally, we introduce a principle that takes both types of symmetries into account. For a choice profile $\vec{c} = (c_1, ..., c_n)$ let $S_i(\vec{c}) := C_i \cap \bigcup_j [c_j]$ for each i. We say that $(c_1, ..., c_n)$ **exhibits a bad symmetry** if $S_1(\vec{c}) \times \cdots \times S_n(\vec{c}) \nsubseteq W_G$, and a choice c **generates a bad symmetry** if σ_c exhibits a bad symmetry for every choice profile σ_c that contains c.

Elimination of bad symmetries (ES):

Never play choices that generate bad symmetries, if possible.

By the definition of bad choice symmetry it is easy to see that if a choice c generates either a bad choice symmetry or a bad player symmetry, then c also generates a bad symmetry. Therefore by using Claim I—which is presented in Section 3.9—it is easy to show that $s(\text{ECS}), s(\text{EPS}) \subseteq s(\text{ES})$.

By the definitions of ECS and EPS, it is clear that they can solve all NL solvable games and therefore also ES can solve all NL solvable games. Furthermore, we can show that the games solvable by ECS, EPS and CRC are *completely independent* of each other. See the following table:

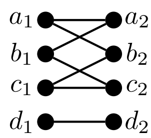
A class $\mathcal{G}$ of games	Example of a game in the class $\mathcal{G}$
$s(\text{ECS}) \setminus (s(\text{EPS} \cup s(\text{CRC}))$	$G(1 \times 1 + 2(1 \times 2))$
$s(\text{EPS}) \setminus (s(\text{ECS} \cup s(\text{CRC}))$	$G(1 \times 1 + 1 \times 2 + 2 \times 1)$
$s(\text{CRC}) \setminus (s(\text{ECS} \cup s(\text{EPS}))$	
$(s(\text{ECS}) \cup s(\text{EPS})) \setminus s(\text{CRC})$	$G(1 \times 1 + 2(2 \times 2))$
$(s(\text{ECS}) \cup s(\text{CRC})) \setminus s(\text{EPS})$	
$(s(\text{EPS}) \cup s(\text{CRC})) \setminus s(\text{ECS})$	$G(Z_3)$

The WLC game in class $s(\text{CRC}) \setminus (s(\text{ECS} \cup s(\text{EPS}))$ above is also unsolvable with ES and therefore ES and CRC are incomparable with each other. This particular game is also SW-solvable, and thus it follows that all symmetry based principles are incomparable with SW. Since ECS and EPS are incomparable and $s(\text{ECS}), s(\text{EPS}) \subseteq s(\text{ES})$, it also follows that ES is stronger than both ECS and EPS.

So far we have only presented examples of such ECS-solvable games with contain isomorphic connected components. In the following example we see how ECS can be used for eliminating moves from a single component. This particular example can also be solved with EPS but not with CRC.

Example 3.12. In the WLC game $G(O_3 + 1 \times 1)$, there are no weakly dominated choices and thus it is not CRC-solvable. However, by applying ECS or EPS, players will play choises d_1 and d_2 which are winning.

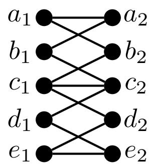
$G(O_3 + 1 \times 1)$:



So far we have only seen ECS and EPS solving games whose game graphs consist of several connected components. It is easy to see that none of these kinds of games can be solved with CRC. In the next Example we present an ECS and EPS-solvable game whose game graph is connected, but which is not CRC-solvable.

Example 3.13. In the WLC game G below, there are no weakly dominated choices. However, by applying ECS or EPS players will pick choices c_1 and c_2 which are winning. (Note here that G is almost of the type s $G(O_5)$, the only difference being a single extra edge that "forms a diagonal of the 10-cycle".)

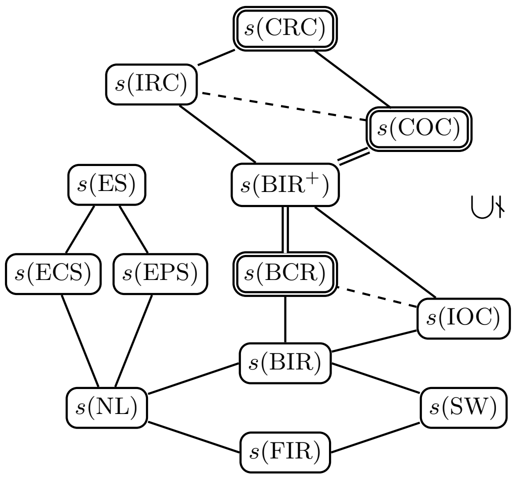
G :



In a follow-up work we will address questions about compatibility of the symmetry principles ECS and EPS with each other and with the other principles considered so far, in particular with CRC which is the strongest of them.

3.7 Hierarchy of the principles presented so far

The partially ordered diagram below presents the hierarchy of solvable games with the principles we have presented in this paper. The principles that only use individual reasoning have normal frames and the ones that use collective reasoning have double frames.



- Normal lines represent proper inclusions in both the general *and* 2-player case.
- *Double* lines represent proper inclusions in the general case. In the 2-player case there is an identity.
- *Dashed* lines represent proper inclusions in the 2-player case. In the general case the two sets are not comparable.

Note that the diagram is complete in the sense that no new lines can be added to it (in the general nor the 2-player case).

3.8 Beyond the limits of pure rationality

How far can we go up the hierarchy of rational principles? This seems a genuinely difficult question to answer. We now mention—without providing precise formal definitions—two structural principles for which it would seem somewhat controversial to claim them rational in our sense, but they are definitely meaningful and natural nevertheless.

The first one is the **principle of probabilistically optimal reasoning (PR)**. Informally put, this principle prescribes to always play a choice that have as large winning extension as possible. These choices have the highest probability of winning, supposing that all the other players play randomly (but *not* if the others follow PR, too: consider e.g. $G(1 \times 2 + 2 \times 1)$).

With PR one can solve games like $G(1 \times 1 + 2 \times 2)$ that are unsolvable with all other principles presented here. However, in $G(1 \times 1 + 2 \times 2)$ one could also reason (perhaps less convincingly) that both players should pick their choices from the subgame $G(1 \times 1)$ since that is the '*simplest*' (and, also the only '*unique*') winning choice profile. We call this kind of reasoning the **Occam razor principle (OR)**. In fact, it generalises the idea of *focal point* [13], [17], [20].

Note that $G(1 \times 1 + 2 \times 2)$ can be won if both players follow PR or if both follow OR, but not if one follows PR while the other follows OR. Moreover, in this game it is impossible for a player to follow *both* PR and OR. Hence, *at least one* of these principles is not purely rational. Actually, it can be argued that *none of them* is purely rational. It is also interesting to note that following PR can violate the symmetry principles, as demonstrated by the game $G(2(2 \times 2) + 1 \times 1)$.

3.9 Characterising structurally unsolvable games

So far we have characterised several principles with different levels of justification for being purely rational. It seems difficult to pinpoint a single strongest principle of pure rationality, but even if such a principle existed, certain games would nevertheless be unsolvable (assuming that purely rational principles must be *structural*). The simplest nontrivial example of such a game is $G(2(1 \times 1))$.

We now characterise the class of WLC games that are **structurally unsolvable**, i.e., *unsolvable by any structural principle*. We say that G is **structurally indeterminate** if all choice profiles in W_G exhibit a bad symmetry (recall the definition of the principle ES). For an example the game $G(1 \times 2 + 2 \times 1)$ is structurally indeterminate, whereas the game $G(1 \times 1 + 2 \times 2)$ is not.

Claim I. *No structural principle can solve a structurally indeterminate game.*

Proof. For the sake of contradiction, suppose that there is a structural principle P and a structurally indeterminate WLC game G such that $G \in s(\mathrm{P})$. Let Σ be any protocol in P. Since P is a structural principle, Σ must be a structural protocol. Since $\mathrm{P}' \subseteq \mathrm{P}$ implies $s(\mathrm{P}) \subseteq s(\mathrm{P}')$, the also the singleton principle $\{\Sigma\}$ solves G.

Let $(u_1, \dots, u_n) \in \Sigma(G, 1) \times \cdots \times \Sigma(G, n)$. Since G is structurally indeterminate, $(u_1, \dots, u_n)$ must exhibit a global losing symmetry. Therefore there is a choice profile $(u'_1, \dots, u'_n) \in U_1 \times \cdots \times U_n$ such that $(u'_1, \dots, u'_n) \notin W_G$. Since Σ is a structural protocol, we must have $(u'_1, \dots, u'_n) \in \Sigma(G, 1) \times \cdots \times \Sigma(G, n)$. Since $(u'_1, \dots, u'_n) \notin W_G$, we have $\Sigma(G, 1) \times \cdots \times \Sigma(G, n) \nsubseteq W_G$. Therefore $\{\Sigma\}$ does not solve G, which is a contradiction. □

This characterisation is optimal in the sense that all games that are not structurally indeterminate, can be solved by some structural principle. This follows from the following even stronger claim.

Claim II. *There exists a protocol* Σ *such that the principle* $\{\Sigma\}$ *can solve all* WLC *games that are not structurally indeterminate.*

Proof. The idea is simply to define a protocol that chooses, on an arbitrary input (G, i) where G is not structurally indeterminate, a node from a tuple of G that does not exhibit global losing symmetry. The only difficulty is that there may be several such tuples in G, and these tuples do not necessarily form a Cartesian product. We now briefly describe how to circumvent this problem.

Firstly, we use some standard encoding of relational structures by binary strings. Furthermore, we require that this encoding is based on a linear order in the standard way, so that every encoding defines a lexicographic order of the tuples of the structure encoded. We note that a single structure can have several encodings, as each linear ordering of the domain can potentially define a different encoding.

Now, when presented with an input (G, i), we do the following. We first define a finite set $\mathcal{G}$ that contains one represenation of each player renaming of G. Then we define the finite set C of strings that encode, for all possible linear orderings, all the structures in $\mathcal{G}$. Then, we choose the string $s \in C$ with the smallest binary number. Using this encoding, we choose the lexicographically smallest tuple that does not exhibit global losing symmetry. Thus we obtain a renaming G' of G together with a tuple $\overline{w} \in W_{G'}$ that does not exhibit global losing symmetry.

Let w_j denote some coordinate of $\overline{w}$ such that there exists a full renaming (β, π) from G' to G that sends the player role number j to the player role number i. There may be several such coordinates j and several renamings (β, π) for j. Let S be the subset of G that contains exactly all points $\pi(w_j)$ for all such renamings (β, π) and coordinates j. The desired protocol outputs S on the input (G, i). □

There are many games that are not structurally unsolvable, but in order to solve them, the players need to follow structural principles that seem arbitrary and certainly cannot be considered purely rational. We call such principles *structural conventions.* However, it is difficult to separate some rational principles from structural conventions. This and other related conceptual issues will be discussed in an extended version of this paper.

4 Concluding remarks

We have proposed and studied a hierarchy of principles of rational players' reasoning about how to act in WLC games. We have compared their strength in terms of the classes of WLC games solvable by them. One major conclusion we draw is that the boundary of (pure) rationality is rather uncertain.

In this paper we have focused on scenarios where players look for choices that guarantee winning if a suitable rational principle is followed. But it is very natural to ask how players should act in a game which seems not solvable by any purely rational principle. If players cannot guarantee a win, it is natural to assume that they should at least try to maximize somehow their collective chances of winning, say, by considering protocols involving some probability distribution between their choices. Another natural extension of our framework is to consider non-structural principles based on limited preplay communication and use of various types of conventions. Also, studying pure dis-coordination games and combinations of coordination/dis-coordination are major lines for further work. We also plan to continue empirical testing in parallel to this work in order to better understand people's reasoning and coordination abilities.

Acknowledgements

The work of Valentin Goranko was partly supported by a research grant 2015-04388 of the Swedish Research Council. The work of Antti Kuusisto was supported by the ERC grant 647289 "CODA."

References

[1] Fausto Barbero. Cooperation in games and epistemic readings of independence-friendly sentences. *to appear in the Journal of Logic, Language and Information*, 2017.

[2] B. Douglas Bernheim. Rationalizable strategic behavior. *Econometrica*, pages 1007–1028, 1984.

[3] Cristina Bicchieri. *Rationality and Coordination.* Cambridge University Press, 1994.

[4] Heinz-Dieter Ebbinghaus, Jörg Flum, and Wolfgang Thomas. *Mathematical logic (2. ed.).* Undergraduate texts in mathematics. Springer, 1994.

[5] Drew Fudenberg and Jean Tirole. *Game theory.* MIT Press, 1991.

[6] David Gauthier. Coordination. *Dialogue*, 14(02):195–221, 1975.

[7] Michael R. Genesereth, Matthew L. Ginsberg, and Jeffrey S. Rosenschein. Cooperation without communication. In *Proc. of AAAI'86, vol. 1*, pages 51–57, 1986.

[8] Margaret Gilbert. Rationality, coordination, and convention. *Synthese*, 84(1):1–21, 1990.

[9] Valentin Goranko, Antti Kuusisto, and Raine Rönnholm. Online tests on pure coordination games, 2017. Available from `https://form.jotformeu.com/70735718307358`.

[10] Peter Hawke. The logic of joint ability in two-player tacit games. *The Review of Symbolic Logic*, pages 1–28, 2017.

[11] D. Lewis. *Convention, A Philosophical Study.* Harvard University Press, 1969.

[12] Kevin Leyton-Brown and Yoav Shoham. *Essentials of Game Theory: A Concise Multidisciplinary Introduction.* Morgan & Claypool Publishers, 2008.

[13] Judith Mehta, Chris Starmer, and Robert Sugden. Focal points in pure coordination games: An experimental investigation. *Theory and Decision*, 36(2):163–185, 1994.

[14] Martin Osborne and Ariel Rubinstein. *A Course in Game Theory.* MIT Press, 1994.

[15] David G. Pearce. Rationalizable strategic behavior and the problem of perfection. *Econometrica*, pages 1029–1050, 1984.

[16] Hans Peters. *Game Theory: A Multi-Leveled Approach.* Springer, 2008.

[17] Thomas C. Schelling. *The strategy of conflict*. Harvard University Press, 1960.

[18] Robert Sugden. Spontaneous order. *J. of Economic Perspectives*, 3(4):85–97, 1989.

[19] Robert Sugden. Rational Choice: A Survey of Contributions from Economics and Philosophy. *Economic Journal*, 101(407):751–785, July 1991.

[20] Robert Sugden. A theory of focal points. *The Economic Journal*, pages 533–550, 1995.

[21] Paul Syverson. *Logic, Convention, and Common Knowledge. A Conventionalist Account of Logic*. CSLI Lecture Notes 142. CSLI Publications, 2002.