When the Group Matters: A Game-Theoretic Analysis of Team Reasoning and Social Ties[*] (Preliminary and incomplete)

Emiliano Lorini[†] Frédéric Moisan [†]

1 Introduction

Many everyday tasks require individuals to act collectively and to coordinate for the pursuit of a common goal. Examples include musicians from an orchestra who need to act together in a specific way in order to play some intended symphony, or players from a soccer team who have to coordinate with each other so that they can eventually score a goal. Even among other tasks that could be done by a single individual, many would be achieved more effectively through teamwork, e.g., painting a house, carrying an heavy object. A common property of all these situations is that each individual in the team acts as a team member and intends to do his part in the joint action of the team. Collective intentionality has been, through the last decades, a central topic in social philosophy (see e.g., [20, 16, 6, 7]) as well as in economics (see e.g., [1, 2, 18, 19]).

The general aim of this work is to use game theory to bridge the gap between individually egoistic behavior and social cooperation in the context of strategic interactions. For this purpose, we provide an analysis of a central theory from the economics literature, which explains how agents, either human or artificial, can manage to solve coordination problems in the context of a joint activity: Bacharach's theory of team reasoning. After discussing the limitations of this theory in modeling some intuitive social behavior, we present our own original model of social ties and show the various advantages it offers compared to Bacharach's theory, especially in the context of social interactions in which different competing groups may coexist. The study we present in this article focuses on the central concept of group identification, which often allows to uncover the 'focal point' required to solve a given coordination problem [15]. For example, consider two agents, say Alice and her partner Bob who have to paint their room together. Let us assume that they can paint it either in blue or in green and that they both prefer to paint it in green. Moreover, each of them is responsible for buying a tin of paint. Which color of paint will Alice and Bob buy? The solution in which each of them buys a tin of green paint becomes the 'focal point' of their coordination problem, because it satisfies their common goal, i.e., to paint the room in green. There-

[*]Funding through the ANR- 2010 JCJC 1803 01 is gratefully acknowledged.
[†]Université de Toulouse, CNRS, Institut de Recherche en Informatique de Toulouse (IRIT)

fore, if Alice and Bob reason as team members, they will likely coordinate on the joint action of buying two tins of green paint, which is the most useful for the whole team.

2 Problems of co-operation

In this section, we consider the most well-known type of two player games that involves co-operation, and where classical economic theories fail to predict the behavior actually followed by human beings. A general form of this sort of games is depicted in Figure 1, in which case each player (Alice and Bob as respectively the row player and the column player) can either cooperate (i.e., play C) or defect (i.e., play D). The main obvious characteristics of this game is that both players can then obtain the same payoff if and only if they manage to coordinate with each other (i.e., they each get x or y depending on whether they both play C or D).

		Bob	
		C	D
Alice	C	(x,x)	$(0,z)$
	D	$(z,0)$	(y,y)

Figure 1: Game involving co-operation (with $0 < y < x$ and $0 \leq z$)

However, another important property of the game in Figure 1 is that the value of the payoff z can provide some incentives for each individual to defect. In fact, Table 1 shows that the game in Figure 1 can refer to three different well known games from the literature, depending on the payoff z.

Constraints on z	Game classifications
$z = 0$	Hi-Lo matching game
$y \leq z < x$	Stag-Hunt game
$x \leq z$	Prisoner's Dilemma

Table 1: Various classifications of games

In the particular case of the Hi-Lo matching game (i.e., when $z = 0$), one can detect no incentive for each individual to defect. However, the interesting observation is that traditional game theory interprets it as a social dilemma where no prediction can be made (the game indeed yields two distinct Nash equilibria, i.e., (C, C) and (D, D)), although it is shown in [2, 9] that people largely coordinate on the most rewarding outcome for both players, i.e., (C, C).

Concerning the Stag-Hunt game (i.e., when $y \leq z < x$), also known as the "assurance game" or the "trust dilemma", although it appears to have the same theoretical properties as the Hi-Lo game (the game then yields the same two Nash equilibria), it introduces some incentives for defecting: in fact, playing C involves a risk of losing if the other player defects, whereas playing D ensures a payoff of at least $y > 0$. However, as for the Hi-Lo game, experimental evidence has shown that people largely cooperate by playing C even if it is risky to do so [21].

Finally, the situation characterized by the Prisoner's Dilemma (i.e., when $x \leq z$) increases further the above incentives so that cooperating now becomes strictly dominated by defecting for both players (i.e., (D, D) then becomes the unique Nash equilibrium). Yet, various experimental studies have also shown a non-negligible cooperation rate (of reaching the (C, C) outcome) in such scenarios, which varies between 30-40% (see e.g., [17]).

In order to formally analyze such social dilemmas in more details, let us first define a classical strategic game structure G.

Definition 2.1 (Standard Strategic Game) *A strategic game is a tuple $G = \langle Agt, \{S_i | i \in Agt\}, \{U_i | i \in Agt\}\rangle$ where:*

- $Agt = \{1, \ldots, n\}$ *is the set of agents;*
- S_i *defines the set of strategies for agent i;*
- $U_i : \prod_{i \in Agt} S_i \to \mathbb{R}$ *is a total payoff function mapping every strategy profile to some real number for some agent i.*

Let us denote $S_J = \prod_{i \in J} S_i$ the set of joint strategies performed by every group of agents $J \in 2^{Agt*}$ where $2^{Agt*} = 2^{Agt} \setminus \{\emptyset\}$. For notational convenience, throughout the paper we write S instead of S_{Agt}.

One might then wonder whether current theories of social preferences proposed in the field of behavioral economics (see [10] for an overview of these theories) are able to explain empirical evidences of mutual cooperation in the previous types of games (i.e., coordinating on (C, C)). The main idea of these theories consist in 'transforming' the agents' utility in the original game on the basis of some social feature such as altruism, inequity aversion or fairness in order to obtain a new game in which equilibria can be computed using classical solution concepts (e.g., Nash equilibrium).

The theory of social preferences that seems to be most relevant to the above games relies on the notion of fairness. In [8], Charness & Rabin indeed propose a specific form of social preferences they call *quasi-maximin* preferences. In their model, group payoff is computed by means of a social welfare function which corresponds to a *weighted* combination of Rawls' *maximin* and of the utilitarian welfare function (i.e., summation of individual payoffs) (see [8, p. 851]).

Formally, for every solution $s \in S$, according to Charness & Rabin's *fairness* model, the utility function of player $i \in \{Alice, Bob\}$ is given by:

$$U_i^F(s) = (1 - \lambda) \cdot U_i(s) + \lambda \cdot SW_i(s)$$

where $\lambda \in [0, 1]$ and $SW_i(s)$ defines the social welfare function as follows:

$$SW_i(s) = \delta \cdot \min_{j \in Agt} U_j(s) + (1 - \delta) \cdot \sum_{j \in Agt} U_j(s)$$

where $\delta \in [0, 1]$.

The two parameters λ and δ can be interpreted as follows: λ measures how much player i cares about pursuing the social welfare versus his own self-interest. In the social welfare function, δ measures the degree of concern for helping the worst-off person versus maximizing the total social surplus. Setting $\delta = 1$ corresponds to the pure "*maximin*" principle (or "Rawlsian" criterion), while setting $\delta = 0$ corresponds to pure utilitarianism.

However, although such a model allows to explain why people choose to cooperate in the context of the Prisoner's Dilemma, it remains indecisive in the case of the Hi-Lo game or the Stag-Hunt game. In fact, in the latter games, such a model appears to be of no help because, for any possible value of δ and λ, the two solutions (C, C) and (D, D) remain Nash equilibria in the resulting transformed game. As a consequence, similarly to the original game considering pure self-interested agents, no prediction can be made.

Since current theories of social preferences are not sufficient to explain the behavior observed in situations like the Hi-Lo game and the Stag-Hunt game, we provide through the next sections some alternative theories that allow to correctly model and explain such observations[1]. As those theories refer to the concept of collective utility, let us first extend our previous formalization of a strategic game from Definition 2.1 so that it incorporates the group utility function.

Definition 2.2 (Game with Group Utility) *A strategic game with group utility is a tuple $G' = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}\rangle$ where:*

- $Agt = \{1, \dots, n\}$ *is the set of agents;*
- S_i *defines the set of strategies for agent i;*
- $U_J : \prod_{i \in Agt} S_i \to \mathbb{R}$ *is a total payoff function mapping every strategy profile to some real number for some team J.*

For notational convenience, throughout the paper, we write U_i instead of $U_{\{i\}}$.

While nothing specifies how to compute each group utility in Definition 2.2, one can define those in terms of individual payoffs. As some examples, the following principles can be considered, which are well known to be the most realistic.

For every $J \in 2^{Agt*}$ and every $s \in S$, let us define:

- pure utilitarianism: $U_J(s) = \sum_{i \in J} U_i(s)$
- the *maximin* principle: $U_J(s) = \min_{i \in J} U_i(s)$

Note, however, that such a collective utility function may however be defined differently.

[1]Other theories of social preferences such as altruism, reciprocity [13], and inequity aversion [10] are not discussed here because they simply reduce the game from Figure 1 to another type of Hi-Lo matching game.

3 Team reasoning

In this section, we consider the well-known concept of team reasoning, which was first introduced by Bacharach in [1] in order to explain observed behavior in the game from Section 2. In order to interpret the intuitive behavior in such games, some theorists have indeed proposed to incorporate new modes of reasoning into game theory. For instance, starting from the work of Gilbert [11] and Reagan [14], some economists and logicians [12] have studied team reasoning (also called *we-mode* reasoning) as an alternative to the best-response reasoning assumed in traditional game theory [19, 1, 9] (also called *I-mode* reasoning). Team-directed reasoning is the kind of reasoning that people use when they perceive themselves as acting as members of a group or team [19]. That is, when an agent i engages in team reasoning, he identifies himself as a member of a group of agents J and conceives J as a unit of agency acting as a single entity in pursuit of some collective objective. A team reasoning player acts for the interest of his group by identifying a strategy profile that maximizes the collective payoff of the group, and then, if the maximizing strategy profile is unique, by choosing the action that forms a component of this strategy profile.

3.1 Definitions

In order to formally illustrate Bacharach's original theory introduced in [1], let us extend the strategic game defined by Definition 2.2 in Section 2, and consider what Bacharach calls an *unreliable team interaction* (UTI) structure, that is, a game structure in which there is a probability that a given player identifies with a team and chooses the action which maximizes the team benefit (i.e., the player plays in the *we-mode*), and another probability that the player is a self-interested agent who tries to maximize his own benefit (i.e., the player plays in the *I-mode*)[2]. In this sense, the interaction is "unreliable" because there is no certainty that a player will reason and act as a team member.

Definition 3.1 (Unreliable Team Interaction) *An unreliable team interaction structure is a tuple* $UTI = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}, \{\Omega_i | i \in Agt\}\rangle$ *where:*

- $\langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}\rangle$ *is a strategic game with group utility according to Definition 2.2;*
- Ω_i *is a probability distribution over the set* $T_i = \{J \in 2^{Agt*} | i \in J\}$.

Intuitively, for any agent $i \in Agt$, T_i is the set of groups agent i may identify with. For every $J \in T_i$, $\Omega_i(J)$ is the probability that agent i identifies with team J (i.e., the probability that agent i reasons and acts as a member of team J). In the definition of an UTI structure it is implicitly assumed that an agent i identifies with a unique team at a given moment which can be either the singleton $\{i\}$, which corresponds to agent i playing in the *I-mode*, or some set of agents $J \in T_i$ such that $|J| > 1$, which corresponds to agent i playing in the *we-mode*.

[2] As a matter of simplicity in our model, we assume the absence of any outside signal observable by the players, since it does not appear to be relevant to our study (such signals are considered in Bacharach's original UTI structure from [1]).

The set of group identification states T is defined as $T = \prod_{i \in Agt} T_i$. Elements of T are denoted by $t, t', \ldots$. Given some group identification state $t \in T$, we write t_i to denote the element of t corresponding to agent i (i.e., the group agent i identifies with according to t). The probability distribution Ω over the set T is defined as $\Omega = \prod_{i \in Agt} \Omega_i$. Given some $\langle J_1, \ldots, J_n \rangle \in T$, $\Omega(\langle J_1, \ldots, J_n \rangle)$ is the probability that "agent 1 identifies with team J_1 *and* agent 2 identifies with team J_2 *and... and* agent n identifies with team J_n".

In [1], Bacharach further introduces the notion of a *protocol*, which consists in specifying a strategy for every group $J \in 2^{Agt*}$. Formally, a protocol α is a function mapping every team $J \in 2^{Agt*}$ to a strategy profile $s_J \in S_J$. The set of all protocols is denoted by Δ. Intuitively, given a protocol α and a team J, $\alpha(J)$ specifies the strategy that the agents in J *should* play according to the protocol α, when identifying themselves as members of team J. Note that if J is a singleton $\{i\}$, $\alpha(\{i\})$ is nothing but the action that the agent i *should* play according to the protocol α, when reasoning in the *I-mode*.

Given the probability distribution Ω on the set of group identification states T, one can then express the expected value of a given team protocol $\alpha \in \Delta$ for some team $J \in 2^{Agt*}$:

$$EV_J(\alpha) = \sum_{t \in T} \Omega(t) \cdot U_J(s_1^{\alpha,t}, \ldots, s_n^{\alpha,t}) \tag{1}$$

where, for every $i \in Agt$, $s_i^{\alpha,t}$ is the action of agent i in the strategy profile $\alpha(t_i)$. Intuitively, $EV_J(\alpha)$ measures how much utility team J can expect from the protocol α.

Furthermore, given two protocols α and β, and a group $J \in 2^{Agt*}$, we write $\alpha_J \cdot \beta_{-J}$ to denote the protocol such that $\alpha_J \cdot \beta_{-J}(J) = \alpha(J)$ and for all $H \in 2^{Agt*}$ such that $H \neq J$, $\alpha_J \cdot \beta_{-J}(H) = \beta(H)$.

This allows us to express the concept of an *UTI* equilibrium as follows:

Definition 3.2 (Uti Equilibrium) *A protocol α is an UTI equilibrium if and only if:*

$$\forall J \in 2^{Agt*}, \forall \beta \in \Delta, EV_J(\beta_J \cdot \alpha_{-J}) \leq EV_J(\alpha_J \cdot \alpha_{-J})$$

In order to illustrate the above UTI structure, let us consider the game presented in Section 2. The corresponding structure $UTI = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}, \{\Omega_i | i \in Agt\} \rangle$ can therefore be defined as follows:

- $Agt = \{a, b\}$ where a and b respectively stand for Alice and Bob;
- $S_a = S_b = \{C, D\}$;
  $S_{\{a,b\}} = \{(D,D), (D,C), (C,D), (C,C)\}$;
- The individual payoff functions are defined according to Figure 1 (where $0 < y < x$):
  $U_a(C,C) = U_b(C,C) = x$;
  $U_a(D,D) = U_b(D,D) = y$;
  $U_a(D,C) = U_b(C,D) = z$;

$U_a(C, D) = U_b(D, C) = 0$;
The group payoff function may then be freely defined according to the *maximin* principle [3], e.g.,
$U_{\{a,b\}}(C, C) = x$;
$U_{\{a,b\}}(D, D) = y$;
$U_{\{a,b\}}(C, D) = U_{\{a,b\}}(D, C) = 0$;

- for every $i \in \{a, b\}$, Ω_i is defined as follows:
  $\Omega_i(\{i\}) = \omega_i$
  $\Omega_i(\{a, b\}) = 1 - \omega_i$

  where ω_i and $1-\omega_i$ respectively characterize the probability that player i reasons in the *I-mode* / *we-mode*.

In the case of the preceding game, let us consider the protocol α such that $\alpha(\{a\}) = \alpha(\{b\}) = C$ and $\alpha(\{a, b\}) = (C, C)$. Protocol α simply specifies that both players choose strategy C independently of reasoning in the *I-mode* or in the *we-mode*. Similarly, let us consider the alternative protocol β such that $\beta(\{a\}) = \beta(\{b\}) = D$ and $\beta(\{a, b\}) = (C, C)$. In this case, protocol β specifies that both players choose strategy D when reasoning in the *I-mode* and choose strategy C when reasoning in the *we-mode*.

Table 2 then illustrates the conditions under which either of the above protocols (α or β) is the unique UTI equilibrium in the above UTI structure, depending on the value of payoff z and the group identification parameters ω_a and ω_b.

Group payoff functions	**Conditions on z**	**Conditions on ω_a and ω_b**	**Unique UTI equilibria**
maximin	$0 \leq z$	$\frac{\omega_a \cdot \omega_b}{\omega_a + \omega_b} > \frac{y}{x+y}$	α
pure utilitarianism	$0 \leq z < y$	$\frac{\omega_a \cdot \omega_b}{\omega_a + \omega_b} > \frac{2y-z}{2x+2y-2z}$	α
	$y \leq z < x$	$\max(\omega_a, \omega_b) > \frac{y}{x+y-z}$	α
	$x < z \leq 2x$	$0 \leq \omega_a, \omega_b \leq 1$	β
	$2x < z$	$\max(\omega_a, \omega_b) < \frac{z/2-y}{z-x-y}$	β

Table 2: Predictions in the UTI structure

The main observation that can be made from Table 2 is that, whenever $0 \leq z < x$, for some sufficiently high probability ω_a or ω_b that either agent a or b identifies as a member of team $\{a, b\}$, playing strategy C becomes the only optimal choice no matter whether a and b *actually* reason in the *I-mode* or in the *we-mode* (cf. protocol α).

However, there exist some notable differences between using the maximin principle of pure utilitarianism as the group payoff function. In fact, one should note that the determination of the unique equilibrium α is independent of the value of z when using the maximin principle. As an example, if one assume that $x = 10$ and $y = 5$ in the

[3] One could similarly use the pure utilitarian criterion as the group payoff function.

above UTI structure, and that both players similarly identify as a member of group $\{a, b\}$, i.e., $\omega_a = \omega_b$, then α is the unique UTI equilibrium whenever $\omega_a, \omega_b > 2/3$, independently of whether the UTI structure represents an Hi-Lo game, a Stag-Hunt game, or a prisoner's dilemma (i.e., no matter the value of z).

On the other hand, when using the pure utilitarianism criterion, the determination of a UTI equilibrium is more complex as it relies on the value of z. In the context of the Hi-Lo game or the Stag-Hunt game (i.e., when $0 \leq z < x$), protocol α can again be the unique equilibrium for some sufficiently high probability of identifying with the group. However, in the case of the prisoner's dilemma where mutual cooperation is the best outcome for the group (i.e., when $x < z < 2x$), protocol β is always the unique optimal solution, no matter the players' level of group identification[4]: if one identifies with the group, then one will play C, else one will play D. One should also note that, in the particular case where there is no unique best outcome for the group (i.e., when $z > 2x$), if both players sufficiently identify with the group and $\omega_a = \omega_b$, then there exists no unique UTI equilibrium (solutions (C, D) and (D, C) become equally good for both players).

Let us now give a more precise interpretation of a team reasoning structure in terms of a classical strategic form game.

Definition 3.3 (Induced Strategic Game) *Given an unreliable team interaction structure* $UTI = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}, \{\Omega_i | i \in Agt\}\rangle$*, the corresponding induced strategic game is a tuple* $G^{uti} = \langle Agt', \{S'_J | J \in Agt'\}, \{U'_J | J \in Agt'\}\rangle$ *where:*

- $Agt' = 2^{Agt*}$;
- *for each* $J \in Agt'$, $S'_J = S_J$;
- *for each* $J \in Agt'$, U'_J *is the utility function on* S' *such that*

$$U'_J(s) = EV_J(\alpha)$$

where $\alpha \in \Delta$ *is the protocol such that, for every* $H \in Agt'$, $\alpha(H) = s_H$.

According to Definition 3.3, every UTI structure induce a strategic form game with a player for every coalition in the original game. For example, in the above UTI structure for the game from Section 2, the induced strategic game would include three players, i.e., Alice, Bob, and the team whose members are Alice and Bob.

As already shown in [1], it is straightforward to demonstrate the following proposition.

Proposition 3.1 *Given an unreliable team interaction structure* $UTI = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}, \{\Omega_i | i \in Agt\}\rangle$*, the protocol* $\alpha \in \Delta$ *is an* UTI *equilibrium if and only if the strategy profile* s *such that* $s_J = \alpha(J)$ *for all* $J \in 2^{Agt*}$ *is a Nash equilibrium in the strategic game* G^{uti} *induced by* UTI.

[4]Note that, whenever $z = x$, then there exist no unique UTI equilibrium: both protocols α and β are UTI equilibria.

3.2 Limitations

Although Bacharach's theory of team reasoning improves over Binmore's theory of empathetic preferences, as it allows to model situations in which different competing coalitions coexist, it has some limitations that we want to discuss here.

As pointed out in Section 3, Bacharach's theory relies on the assumption that every agent identifies with a unique team at a given time. This is a strong assumption, as it prevents from modeling situations in which an agent plays as a member of more than one group. To illustrate this, consider a scenario where one faces the dilemma between cooperating with a close friend, and cooperating with a family member (assuming the friend and the family member cannot cooperate with each other). In this case, Bacharach's theory predicts that the individual will choose over these two options, even though a more egalitarian solution might exist that would satisfy equally all players. This limitation is therefore particularly relevant when modeling more flexible and heterogenous multiagent systems in which different coalitions might be formed whose intersections are non-empty.

Another problem of the theory of team reasoning concerns the exogenous probabilistic distributions Ω_i in the definition of an UTI structure. In fact, it is not completely clear how they should be interpreted. While such probabilities may depend on some intrinsic features of the game such as the payoff structure, they may also be determined by some pre-existent social relationships between the players. Bacharach's theory however remains vague regarding this issue.

Apart from the previous conceptual restrictions and ambiguities, the theory of team reasoning has a technical limitation which lies in the complexity of the problem of computing equilibria in the context of teamwork by using Nash equilibrium (see Proposition 3.1). In fact, as Definition 3.3 indicates, given a structure $UTI = \langle Agt, \{S_i | i \in Agt\}, \{U'_J | J \in 2^{Agt*}\}, \{\Omega_i | i \in Agt\}\rangle$, the size of the set of agent Agt' in the game G^{uti} induced by UTI is exponential in Agt, in particular we have $|Agt'| = 2^{|Agt|} - 1$. In other words, if one wants to use Nash equilibrium in order to compute UTI equilibria, he has to increase exponentially the size of the structure of interaction.

In the next section we present a model of social ties which provides a simpler approach to group reasoning and collective decision making. The interesting aspect of this model is that it allows to compute equilibria in the context of collaborative activity by using the concept of Nash equilibrium and without increasing the size of the structure of interaction.

4 A theory of social ties

In this section, we introduce a new model that characterizes well the agents' behavior in the presence of social ties. The aim of such a model is to explain the complex social dilemmas such as the game from Section 2 by simply following the main idea on which theories of social preferences are based (see e.g., [10, 8]), that is: (1) the idea of performing some utility transformation in a given game, and then (2) of applying classical solution concepts from game theory (e.g., Nash equilibrium) in the transformed game in order to find equilibria. Indeed, similarly to theories of social preferences, our starting assumption is that the existence of a social tie between two individuals may have an impact on their utilities, in the sense that the utility that an agent attaches to a given option may be affected by his social ties with other agents. In the next section, we explain in detail how social ties may affect an agent's utility function.

4.1 Definitions

Let us now introduce a social ties game, which extends the strategic game structure presented in Definition 2.2 (see Section 2).

Definition 4.1 (Social Ties Game) *A social ties game is a tuple* $ST = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}, k\rangle$ *where:*

- $\langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}\rangle$ *is a strategic game with group utility according to Definition 2.2;*
- k *is a total function* $k : 2^{Agt*} \to [0, 1]$*, such that:*

  C1 $\forall_{i \in Agt}, \sum_{J \subseteq Agt \setminus \{i\}} k(J \cup \{i\}) = 1$

Every parameter $k(J)$ in Definition 4.1 should be seen as a measure of the social closeness within the group J in the current game context. In particular, $k(J)$ measures how much every member of J is socially tied with the group J. Setting $k(J)$ to 0 corresponds to a non-existing tie with group J for all members of J, whereas setting $k(J)$ to 1 means that every member from J is strongly tied with group J. Moreover, $k(\{i\})$ stands for agent i's measure of selfishness.

According to Constraint **C1**, every agent's identifications with every (sub)group are exclusive, i.e., an agent can neither fully identify with different groups, nor identify with no group at all (in the most extreme case, the agent i is maximally selfish in the sense that $k(\{i\}) = 1$).

The following definition introduces the notion of social ties utility function, i.e., how an agent's utility is affected by his social ties.

Definition 4.2 (Social Ties Utility) *For every strategy profile* $s \in S$*, the social ties utility function of player* i *is given by:*

$$U_i^{ST}(s) = \sum_{J \subseteq Agt \setminus \{i\}} k(J \cup \{i\}) \cdot \max_{s'_J \in S_J} U_{J \cup \{i\}}(s_{-J}, s'_J)$$

In Definition 4.2, S_{-J} denotes the set of joint strategies for the coalition $Agt \setminus J$ (i.e., $S_{-J} = S_{Agt \setminus J}$), and $U_J(s)$ stands for group J's utility function, as it is described in Definition 2.2.

The general idea of our model, which is formally expressed by the preceding social ties utility function $U_i^{ST}(s)$, is that, in the presence of a strong tie with a group J, agent i does not face a full strategic problem anymore. Indeed, the utility of the strategy profile s for agent i becomes independent of what other agents from $J \setminus \{i\}$ do in this strategy profile (i.e., $s_{J \setminus \{i\}}$). Therefore, agent i only needs to reason strategically regarding the choices of every player outside of J, and choose the action from the strategy profile which maximizes group J's utility. As a result, i's way of reasoning can be interpreted as "do the right thing for the group J, assuming that all other players in J also do the right thing for the group J".

As in the previous section, let us now interpret the above social ties game in terms of a classical strategic form game.

Definition 4.3 (Induced Strategic Game) *Given a social ties game $ST = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}, k \rangle$, the corresponding induced strategic game is a tuple $G^{st} = \langle Agt, \{S_i | i \in Agt\}, \{U'_i | i \in Agt\} \rangle$ where:*

- $U'_i(s) = U_i^{ST}(s)$ *for every strategy profile* $s \in S$.

Moreover, as a concrete example of such a social ties game, let us again consider the type of interactions illustrated in Figure 1 from Section 2. Let us first restate the social ties utility function from Definition 4.2, which can be simplified as follows when applied to any two player game (i.e., $|Agt| = 2$). For every $i \in Agt$ and every $s \in S$:

$$U_i^{ST}(s) = (1 - k(Agt)) \cdot U_i(s) + k(Agt) \cdot \max_{s'_{-i} \in S_{-i}} U_{Agt}(s_i, s'_{-i})$$

Starting from the matrix payoff from Figure 1, Figure 2 represents the corresponding transformed utilities for each player based on the *maximin* principle as the group utility function, and where an extremely strong social tie exists between Alice and Bob (i.e., $k(Agt) = 1$).

		Bob	
		C	D
Alice	C	(x,x)	(x,y)
	D	(y,x)	(y,y)

Figure 2: Transformed utilities (with $k(Agt) = 1$)

In this case, it is easy to show, through iterated elimination of strictly dominated strategies, that the only Nash equilibrium resulting from this transformation is (C, C). One should note that each player's choice becomes independent of the opponent's choice. In other words, the initially strategic problem becomes a classical problem of individual decision making that each player has to solve. More generally, Table 3 depicts the predictions that can be made in this game depending on the value of z and social tie parameter $k(Agt)$.

The main observation one can make from Table 3 is that strategy C strictly dominates D in the context of some sufficiently strong social tie between both players in such a social ties game. However, as for team reasoning (see Table 2 in Section 3), some difference appear between the two types of group payoff functions. Under the assumption of the maximin principle, the agents' behavior is independent of whether the game characterizes an Hi-Lo game, a Stag-Hunt game, or a prisoner's dilemma

Group payoff functions	**Conditions on z**	**Conditions on $k(Agt)$**	**Unique Nash equilibria**
maximin	$0 \leq z$	$k(Agt) > y/x$	(C, C)
pure utilitarianism	$0 \leq z < 2x$	$k(Agt) > (z - x)/x$ and $k(Agt) > y/(2x - y)$	

Table 3: Predictions in the strategic game induced by ST

(i.e., it does not rely on the value of z). On the other hand, when using the pure utilitarianism criterion, the same prediction can be made only if the group has a unique favourite outcome (i.e., when $z < 2x$). In the case where this condition is not satisfied (i.e., when $z \geq 2x$), social ties cannot help the players to cooperate with each other: they are then unable to coordinate as both (C, D) and (D, C) become equally optimal solutions[5].

4.2 Analysis

The following theorem demonstrates the ability of social ties to converge towards playing a Nash equilibrium solution.

Theorem 4.1 *For any strategic game with group utility $G' = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}\rangle$, there exists a social ties game $ST = \langle G', k\rangle$ whose induced strategic game has a Nash equilibrium in pure strategies.*

One can show, through the proof of Theorem 4.1, that there exist games with group utility G' that do not yield a unique Nash equilibrium in the strategic game induced by any social ties game $ST = \langle G', k\rangle$. In this case, as several distinct equilibria are unable to make any clear prediction, one should note that exploiting the group utility as done by such a ST game therefore becomes irrelevant[6]. However, notice that, for any game with group utility $G = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}\rangle$ such that $argmax_{s' \in S} U_{Agt}(s')$ is a singleton (i.e., the group Agt has a unique goal), there exists a social ties game $ST = \langle G', k\rangle$ such that $k(Agt) = 1$ and the game induced by ST has a unique pure strategy Nash equilibrium.

4.3 Relationship with team reasoning

Let us now provide a comparative analysis of the social ties game with the previous theory of team reasoning (see Section 3). For this purpose, we first focus on the particular case of two-player games before extending the analysis to any n-player game such that $n \geq 2$.

[5]The resulting transformed game then corresponds to a version of the well known Battle of the Sexes game (see, e.g., Figure 3).

[6]This remark also applies to the theory of team reasoning from the previous section.

4.3.1 Comparison in any two-player game

In the context of two-player games, let us first specify the various similarities that can emerge when considering subclasses of the above models of team reasoning and social ties. In order to do so, we then consider some binary interpretation of those game structures, which can be defined as in Definition 4.4.

Definition 4.4 (Binary Games) *A binary unreliable team interaction $BUTI$ is a structure $UTI = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}, \{\Omega_i | i \in Agt\}\rangle$ where there exists $t \in T$ such that:*

- *for every $i \in Agt$, $\Omega_i(t_i) = 1$;*
- *for every $i, j \in Agt$ such that $i \neq j$, $j \in t_i$ if and only if $i \in t_j$.*

A binary social ties game BST is a game $ST = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}, k\rangle$ where:

- *for every $J \subseteq Agt$, $k(J) \in \{0, 1\}$.*

According to the $BUTI$ structure, there is no uncertainty about which group each agent identifies with. In such a structure, the concept of a protocol can be reduced to a simple strategy profile, which therefore allows the comparison with some ST game. Moreover, note that in this case, an agent identifies with a group if and only if other members of that group also identify with it. Similarly, the BST game represents an extreme interpretation of the ST game where every agent can only identify with a unique group.

Thus performing a detailed analysis of such binary two-player games can reveal their similarities under the previous constraints from Definition 4.4, as shown through Theorem 4.2.

Theorem 4.2 *Given a strategic game with group utility $G' = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}\rangle$ with $|Agt| = 2$, a binary social ties game $BST = \langle G', k\rangle$ and a binary unreliable team interaction structure $BUTI = \langle G', \{\Omega_i | i \in Agt\}\rangle$, such that $k(Agt) = \Omega_i(Agt) = 1$ for every $i \in Agt$:*

(1) if the game induced by BST has a unique Nash equilibrium $s \in S$, then s is the unique solution such that every protocol $\alpha \in \Delta$ defined by $\alpha(Agt) = s$ is an UTI equilibrium in $BUTI$.

(2) if there exists a unique solution $s \in S$ such that every protocol $\alpha \in \Delta$ defined by $\alpha(Agt) = s$ is an UTI equilibrium in $BUTI$, then s is the unique Nash equilibrium in the game induced by BST.

Theorem 4.2 therefore indicates that binary versions of both game structures always make the same predictions regarding the agents' behavior in any two-player game.

Moreover, another common property that both models share is their reliance on the concept of group intentions. In fact, in an UTI structure as well as in a ST game, the formation of a group implies that all members of this group seek to satisfy the group's objective, that is, they aim at reaching the highest possible payoff for the group. However, it appears that, depending on the type of interactive situation being considered, such a collective goal may not be clearly determined. As an example, one may consider the two-player game in Figure 3, which corresponds to the well-known Battle of the Sexes game.

	D	R
U	$(10,5)$	$(0,0)$
D	$(0,0)$	$(5,10)$

Figure 3: Battle of the Sexes

In the context of this game, it is straightforward to show that any UTI structure and any ST game will always be indecisive, no matter the type of group identification involved. Such an observation is obviously justified by the fact that the group made of the two players does not have a unique goal (i.e., both solutions (U,D) and (D,R) are equally good for the group according to either utilitarianism or the maximin principle). As a consequence, ST games and UTI structures simply become irrelevant in such a game.

Nevertheless, beside the previously highlighted similarity, one can observe some significant mathematical differences between both theories. In fact, the major difference existing between both game structures is related to the type of reasoning followed by the agents: unlike the UTI structure, the ST game considers agents that do not exclusively reason in any binary mode (e.g., either *I-mode* or *we-mode*), but instead assumes that they behave according to some subjective combination of selfishness and blind trust and cooperation. As a result of this important difference, both game structures can disagree about the predicted outcome. More formally, this distinction between both game structures leads to the following theorem:

Theorem 4.3 *There exists a strategic game with group utility* $G' = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}\rangle$ *with* $|Agt| = 2$*, and a social ties game* $ST = \langle G', k\rangle$ *whose induced game has a unique Nash equilibrium* $s \in S$*, such that, for any unreliable team interaction* $UTI = \langle G', \Omega\rangle$*, if a protocol* $\alpha \in \Delta$ *is an* UTI *equilibrium, then* $(s_1^{\alpha,t}, \ldots, s_{|Agt|}^{\alpha,t}) \neq s$ *for any group identification state* $t \in T$*.*

Following Theorem 4.3, and as already mentioned in Section 3.2, such a difference between both game structures can be emphasized by considering more complex multiagent systems that involve the formation of sub-coalitions whose intersections are non-empty.

Furthermore, it is worth noting that Bacharach's unreliable team interaction also differs from our social ties game as it considers opportunistic behavior. In fact, Bacharach's concept of a team protocol (see Section 3.1) specifies what each player should do in both modes of reasoning (i.e., in *I-mode* and in *we-mode*). As a consequence, assuming a situation where the probability ω of reasoning in *we-mode* for each player is sufficiently high (e.g., $\omega = 0.9$) implies that a player who reasons in *I-mode* will seek to maximize his own payoff, given what the other player does while reasoning in *we-mode*. In other words, such an *I-mode* reasoner, surrounded by *we-mode* reasoners, will act as an opportunist. On the other hand, our model does not characterize such a type of behavior as it uniquely specifies what the players should do based on the level with which they similarly identify with the group. As a result, this difference leads to the following theorem:

Theorem 4.4 *There exists a strategic game with group utility $G' = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}\rangle$ with $|Agt| = 2$, and an unreliable team interaction $UTI = \langle G', \Omega\rangle$ whose unique UTI equilibria $\alpha \in \Delta$ specifies a solution $s = (s_1^{\alpha,t}, \dots, s_{|Agt|}^{\alpha,t})$ for some group identification state $t \in T$, such that s cannot be a Nash equilibrium in the game induced by any social ties game $ST = \langle G', k\rangle$.*

4.3.2 Comparison in any n-player game

Let us now extend the previous analysis with considering more complex interactive situations that involve more than two players.

We have demonstrated in Section 4.3.1 that both models UTI and ST face the same limitation whenever the collective goal or intention (i.e., the most preferred outcome) is ambiguous in two-player games. We now intend to show that, although this remark remains true, diverging group intentions are interpreted differently by the two theories.

In fact, as a means to emphasize this difference, we define a restricted class of strategic games with unambiguous group intentions according to Definition 4.5.

Definition 4.5 (Game with Unambiguous Group Intentions) *A game with unambiguous group intentions is a game with group utility $G' = \langle Agt, \{S_i | i \in Agt\}, \{U_i | i \in Agt\}\rangle$ as defined in Definition 2.2, such that, for every $J \in 2^{Agt*}$:*

C2 *for every $J \in 2^{Agt*}$ s.t. $|J| > 1$ and every $s_{-J} \in S_{-J}$, if there exist $s'_J, s''_J \in S_J$ s.t. $U_J(s'_J, s_{-J}) = U_J(s''_J, s_{-J}) = \max_{s'''_J \in S_J} U_J(s'''_J, s_{-J})$, then $s'_J = s''_J$.*

In Definition 4.5, Constraint **C2** simply ensures that no group can have multiple goals at once, that is, every group has a unique preferred outcome. Note that this constraint indeed rules out interactive situations such as the game in Figure 3.

Thus performing a detailed analysis of any binary n-player game can reveal their equivalence under the previous constraint **C2**, as shown through Theorem 4.5.

Theorem 4.5 *Given a strategic game with group utility $G' = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}\rangle$ with $|Agt| > 2$, a binary social ties game $BST = \langle G', k\rangle$, and binary unreliable team interaction structure $BUTI = \langle G', \{\Omega_i | i \in Agt\}\rangle$, such that $k(J) = \Omega_i(J)$ for every $J \in 2^{Agt*}$ and every $i \in J$:*

(1) if the game induced by BST has a unique Nash equilibrium $s \in S$, then s is the unique solution such that the protocol $\alpha \in \Delta$ defined by $\alpha(J) = s_J$ for every $J \in 2^{Agt*}$ is an UTI equilibrium.

*(2) if G' satisfies Constraint **C2** from Definition 4.5, and there exists a unique solution $s \in S$ such that the protocol $\alpha \in \Delta$ defined by $\alpha(J) = s_J$ for every*

$J \in 2^{Agt*}$ is an UTI equilibrium in $BUTI$, then s is the unique Nash equilibrium in the game induced by BST.

Theorem 4.5 therefore generalizes Theorem 4.2 from Section 4.3.1 and shows that binary versions of both game structures can make the same predictions regarding the agents' behavior in any n-player game that satisfies Constraint **C2**.

However, the need for an additional constraint on the original game structure G' in Theorem 4.5 (i.e., Constraint **C2**) points out to another important conceptual difference concerning the type of transformation each model is based upon. Indeed, Bacharach's concept of team reasoning relies on what he calls agency transformation, which consists in conceiving the situation not as a decision making problem for individual agents, but as a decision making problem for the group as an agent. Alternatively, our concept of group identification relies on some payoff transformation, that is, each agent internalizes the intentions of the group before considering the situation as a classical decision making problem. As a consequence of such a conceptual difference, it can lead the above binary models to make different predictions whenever Constraint **C2** is not satisfied, as shown through Theorem 4.6.

Theorem 4.6 *There exists a strategic game with group utility $G' = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}\rangle$, a binary unreliable team interaction structure $BUTI = \langle G', \{\Omega_i | i \in Agt\}\rangle$, and a unique solution $s \in S$, such that the protocol $\alpha \in \Delta$ defined by $\alpha(J) = s_J$ for every $J \in 2^{Agt*}$ is an UTI equilibrium in $BUTI$, and s cannot be a Nash equilibrium in the game induced by any binary social ties game $BST = \langle G', k\rangle$.*

Moreover, another consequence of the different types of transformation each model relies on concerns the complexity for computing an equilibrium solution. In fact, unlike for Bacharach's UTI structures, ST games consider the standard concept of a strategy profile, as in traditional game theory (UTI structures instead consider protocols, as depicted in Section 3.1). As shown through Definition 4.3, a consequence is that the strategic game G^{st} induced by any social ties game ST does not increase the complexity of computing the equilibrium solution, which is indeed an important difference with Bacharach's UTI structure: given a game with group utility G', the game G^{st} induced by any $ST = \langle G', k\rangle$ only transforms the utility function of the original game G' (the number of agents remains unchanged here). Finding a Nash equilibrium in the game G^{st} therefore appears to be mathematically simpler than finding such a solution in the game G^{uti} induced by any $UTI = \langle G', \{\Omega_i | i \in Agt\}\rangle$.

5 Summary and future work

In this work, we provide a comparative analysis of two economic theories that allow to explain collective behavior in situations that involve co-operation. After discussing the limitations of Bacharach's theory of team reasoning, we introduce a new model of social ties, which appears to be well suited to model collective behavior in the context of complex strategic interactions. The advantages of our model compared to Bacharach's theory have been highlighted.

Future work include refining the measure of social ties in the above model. In fact, one may wonder how the social closeness of a group is determined (i.e., function k in Definition 4.1). It is reasonable to assume that a group is made of agents who share some individual ties between each of its members. Expressing the above social ties function k in terms of individual ties however appears not to be an easy task: a player may indeed be very close individually to every member of a group without being socially close to the group itself. We indeed believe that an agent i's identification with some group J does not only rely on i's individual ties with each member of J, but also on each other member's ties with each other within J. As an attempt to fill this gap, we will show that the above model of social ties can be interpreted in terms of an alternative game that incorporate such individual ties.

Furthermore, while we argue, through our model of social ties, that cooperation can be explained through the modification of the players' individual preferences, we will also discuss a different approach introduced by Binmore [3, 4, 5], which relies on the concept of empathetic preferences. Although this theory offers another relevant explanation of cooperative behavior, we will show its limitation to model complex social interactions as it does not incorporate strategic reasoning.

References

[1] M. Bacharach. Interactive team reasoning: a contribution to the theory of cooperation. *Research in economics*, 23:117–147, 1999.

[2] M. Bacharach. *Beyond individual choice: teams and frames in game theory*. Princeton University Press, Oxford, 2006.

[3] K. Binmore. *Game Theory and the Social Contract-Vol. 1: Playing Fair*, volume 1. MIT Press, 1994.

[4] K. Binmore. *Game Theory and the Social Contract-Vol. 2: Just Playing*, volume 2. MIT press, 1998.

[5] K. Binmore. *Natural justice*. Oxford University Press, USA, 2005.

[6] M. Bratman. Shared cooperative activity. *The philosophical review*, 101(2):327–341, 1992.

[7] M. Bratman. Shared intention. *Ethics*, 104(1):97–113, 1993.

[8] G. B. Charness and M. Rabin. Understanding social preferences with simple tests. *Quarterly Journal of Economics*, 117:817–869, 2002.

[9] A. Colman, B. Pulford, and J. Rose. Collective rationality in interactive decisions: Evidence for team reasoning. *Acta psychologica*, 128(2):387–397, 2008.

[10] E. Fehr and K. Schmidt. The economics of fairness, reciprocity and altruism–experimental evidence and new theories. *Handbook on the economics of giving, reciprocity and altruism*, 1:615–691, 2006.

[11] M. Gilbert. *On social facts*. Routledge, London, 1989.

[12] E. Lorini. From self-regarding to other-regarding agents in strategic games: a logical analysis. *J. of Applied Non-Classical Logics*, 21(3-4):443–475, 2011.

[13] M. Rabin. Incorporating fairness into game theory and economics. *The American Economic Review*, pages 1281–1302, 1993.

[14] D. Regan. *Utilitarianism and cooperation*. Clarendon Press, Oxford, 1980.

[15] T. Schelling. *The Strategy of Conflict*. Harvard University Press, 1960.

[16] J. Searle. Collective intentions and actions john r. searle. *Intentions in communication*, page 401, 1990.

[17] E. Shafir and A. Tversky. Thinking through uncertainty: Nonconsequential reasoning and choice. *Cognitive psychology*, 24(4):449–474, 1992.

[18] R. Sugden. Thinking as a team: towards an explanation of non-selfish behavior. *Social Philosophy and Policy*, 10(1):69–89, 1993.

[19] R. Sugden. Team preferences. *Economics and Philosophy*, 16:175–204, 2000.

[20] R. Tuomela and K. Miller. We-intentions. *Philosophical Studies*, 53(3):367–389, 1988.

[21] J. Van Huyck, R. Battalio, and R. Beil. Tacit coordination games, strategic uncertainty, and coordination failure. *The American Economic Review*, 80(1):234–248, 1990.

A Proofs of Theorems

A.1 Proof of Theorem 4.2

We here demonstrate that, given a strategic game with group utility $G' = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}\rangle$ with $|Agt| = 2$, the predictions made by both game structures $BST = \langle G', k\rangle$ and $BUTI = \langle G', \{\Omega_i | i \in Agt\}\rangle$ are equivalent whenever $k(Agt) = \Omega_i(Agt)$ for every $i \in Agt$.
Let us first consider the BST game.

From Definitions 4.4 and 4.2, it follows that, in BST, for every $i \in Agt, s \in S$, and for $J = Agt \backslash \{i\}$, we have:

$$U_i^{ST}(s) = \max_{s'_J \in S_J} U_{Agt}(s_i, s'_J) \tag{1}$$

Then (1) allows to state that, given $s \in S$ is a Nash equilibrium solution in the strategic game induced by BST, then for every agent $i \in J$, we have:

$$U_i^{ST}(s) = U_{Agt}(s) \tag{2}$$

where:

$$U_{Agt}(s) = \max_{s' \in S} U_{Agt}(s') \tag{3}$$

Moreover, the fact that s is a unique Nash solution in the game induced by BST implies that the group Agt has a unique goal such that, for every joint strategy $s' \in S$:

$$\text{if } U_{Agt}(s) = U_{Agt}(s'), \text{ then } s = s' \tag{4}$$

In other words, all agents in Agt will maximize the group utility given, and they will do so by performing the unique joint strategy $s \in S$ that allows it.

As a result, one can state that a solution $s \in S$ is a unique Nash equilibrium induced by BST if and only if the conditions (3) and (4) hold.

Let us now similarly consider the $BUTI$ structure.

Definition 4.4 means that every agent can only identify with a unique group. It follows from this definition that there exists a unique group identification state $t \in T$ such that:

$$t_i = Agt \text{ for every } i \in Agt \tag{5}$$

It follows from Definition 4.4 that for every team $J \in 2^{Agt*}$ and every protocol $\alpha \in \Delta$, the expected value can be simplified as follows:

$$EV_J(\alpha) = U_J(s_1^{\alpha,t}, \ldots, s_n^{\alpha,t}) \tag{6}$$

One can note from (6) that, if protocol α is an UTI equilibrium, then, for every $J \in A$, we have:

$$EV_J(\alpha) = \max_{s'_J \in S_J} U_J(s'_J, s_{-J}^{\alpha,t}) \tag{7}$$

where $s_{-J}^{\alpha,t}$ is the combination of actions of all agents $i \in Agt \backslash J$ specified by the strategy profile $\alpha(t_i)$. However, it follows from (5) and (7) that, for every coalition $J \in 2^{Agt*}$ such that $J \notin Agt$ (i.e., for every team nobody identifies with), and for every $\alpha \in \Delta$, $EV_J(\alpha)$ remains constant. This implies that every such group J is simply irrelevant to the computation of the UTI equilibrium. As a result, it is immediate to show that there exists a unique solution $s \in S$ such that every protocol $\alpha \in \Delta$ defined by $\alpha(Agt) = s$ is an UTI equilibrium if and only if the conditions (3) and (4) hold.

A.2 Proof of Theorem 4.3

We consider a two-player strategic game with group utility $G' = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}\rangle$ whose individual payoff matrix is represented in Figure 4(a) for both players in $Agt = \{a, b\}$. Moreover, we take as the group payoff function the classical utilitarian principle (i.e., for any $s \in S$, $U_{\{a,b\}}(s) = U_a(s) + U_b(s)$).

The obvious main characteristics of the game depicted in Figure 4(a) is that both players' outcome is uniquely determined by a single player (i.e., Player a).

We then define a corresponding social ties game $ST = \langle G', k\rangle$ such that $k(\{a, b\}) = 0.5$. One should then note that, in order to satisfy Definition 4.1, we must have that

	D
A	$(8,0)$
B	$(5,7)$
C	$(7,4)$

(a) Game **A**

	D
A	$(8,6)$
B	$(8.5,9.5)$
C	$(9,8)$

(b) Game **B**

Figure 4: Simple dictator games

$k(\{a\}) = k(\{b\}) = 0.5$. The resulting strategic game induced by ST is depicted in Figure 4(b). In this case, it is straightforward to show that the unique Nash equilibrium in this game is (C, D).

On the other hand, let us now consider a corresponding structure $UTI = \langle G', \{\Omega_i | i \in Agt\}\rangle$ such that ω defines the probability that Player a reasons in *we-mode*, that is $\Omega_a(\{a,b\}) = \omega$ and $\Omega_a(\{a\}) = 1 - \omega$ [7].

Table 4 illustrates the expected values for every possible protocol in the game from Figure 4 (note that all protocol specify the sam unique action for b, that is, $\alpha_i(\{b\}) = D$).

Protocols			**Expected values**	
α_i	$\alpha_i(\{a\})$	$\alpha_i(\{a,b\})$	$EV_a(\alpha_i)$	$EV_{\{a,b\}}(\alpha_i)$
α_1	A	(A,D)	8	8
α_2	B	(B,D)	5	12
α_3	C	(C,D)	7	11
α_4	A	(B,D)	$8-3\omega$	$8+4\omega$
α_5	A	(C,D)	$8-\omega$	$8+3\omega$
α_6	B	(A,D)	$5+3\omega$	$12-4\omega$
α_7	B	(C,D)	$5+2\omega$	$12-\omega$
α_8	C	(A,D)	$7+\omega$	$11-3\omega$
α_9	C	(B,D)	$7-2\omega$	$11+\omega$

Table 4: All protocols in the dictator game A (Figure 4(a))

It is then straightforward to show from Table 4 and Definition 3.2 that, for any value of ω such that $0 < \omega < 1$, the structure UTI yields the unique UTI equilibrium α_4. Indeed, for every $0 < \omega < 1$, we have:

$$\begin{array}{lcl|lcl} EV_{\{1,2\}}(\alpha_4) & > & EV_{\{1,2\}}(\alpha_1) & EV_{\{1\}}(\alpha_1) & > & EV_{\{1\}}(\alpha_6) \\ EV_{\{1\}}(\alpha_4) & > & EV_{\{1\}}(\alpha_2) & EV_{\{1\}}(\alpha_5) & > & EV_{\{1\}}(\alpha_7) \\ EV_{\{1\}}(\alpha_5) & > & EV_{\{1\}}(\alpha_3) & EV_{\{1\}}(\alpha_1) & > & EV_{\{1\}}(\alpha_8) \\ EV_{\{1,2\}}(\alpha_4) & > & EV_{\{1,2\}}(\alpha_5) & EV_{\{1\}}(\alpha_4) & > & EV_{\{1\}}(\alpha_9) \end{array}$$

[7] Note that the probability function Ω_b is irrelevant here as Player b cannot influence the outcome of the game.

Moreover, Player a will play A whenever $\omega = 0$ whereas Player a will play B whenever $\omega = 1$. As a result, there exist no UTI structure build on the game G' that has an UTI equilibrium specifying Player a to play C, as predicted by the above social ties game ST.

A.3 Proof of Theorem 4.4

We here consider a two-player strategic game with group utility $G' = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}\rangle$ whose individual payoff matrix is represented in Figure 5 for both players in $Agt = \{a, b\}$.

	L	R
U	$(2, -1)$	$(0, 1)$
D	$(1, 1)$	$(1, 0)$

Figure 5: Game with no pure strategy Nash equilibrium

Note that such a game does not yield any pure strategy Nash equilibrium in its original version. As before, we take as the group payoff function the same classical utilitarian principle for any $s \in S$.

We then define a corresponding structure $UTI = \langle G', \{\Omega_i | i \in Agt\}\rangle$ such that the probability ω that both players reason in *we-mode* is sufficiently high, e.g., $\Omega_a(\{a, b\}) = \Omega_b(\{a, b\}) = \omega = 0.9$. In this particular case, this leads to a unique UTI equilibrium protocol $\alpha \in \Delta$ defined by:

$$\begin{array}{lcl} \alpha(\{a\}) & = & U \\ \alpha(\{b\}) & = & L \\ \alpha(\{a, b\}) & = & (D, L) \end{array}$$

However, while considering any possible corresponding social ties game $ST = \langle G', k\rangle$, one can similarly remark that the unique Nash equilibrium solution (DL) may exist in the strategic game induced by ST whenever $k(\{a, b\}) \geq 1/2$ (otherwise, the game has no equilibrium solution). One can therefore observe that the solution (U, L) specified by the protocol α (when both a and b are in *I-mode*) cannot be a Nash equilibrium in the game induced by any social ties game ST.

However, it is worth pointing out that the strategy profile (U, L) can only be predicted by some UTI structure as a suboptimal solution, as it is subject to some opportunistic behavior of a few *I-mode* reasoners that are surrounded by *we-mode* players.

A.4 Proof of Theorem 4.5

We here demonstrate that, given a strategic game with group utility $G' = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}\rangle$ with $|Agt| > 2$, both game structures $BST = \langle G', k\rangle$ and $BUTI = \langle G', \{\Omega_i | i \in Agt\}\rangle$ can be transformed into simpler strategic games G^{bgi}

and G^{buti}.
Let us first transform the BST game.

From Definitions 4.4 and 4.2, it follows that, for every $i \in Agt, s \in S$, and for some $J \subseteq Agt \backslash \{i\}$:

$$U_i^{ST}(s) = \max_{s'_J \in S_J} U_{J \cup \{i\}}(s_{-J}, s'_J) \quad (1)$$

From (1), Definition 4.1 and Constraint **C1** allow to state that, for every agent $j \in J$ such that $j \neq i$:

$$U_j^{ST}(s) = \max_{s'_K \in S_K} U_{K \cup \{j\}}(s_{-K}, s'_K) \quad (2)$$
where $K = J \cup \{i\} \backslash \{j\}$

Let C denote the set of actual coalitions in the BST game:

$$C = \{J \in 2^{Agt*} | k(J) = 1\} \quad (3)$$

Note that $\bigcup_{J \in C} J = Agt$ and $\bigcap_{J \in C} = \emptyset$.

Then (1) and (2) allow to state that, given $s \in S$ is a Nash equilibrium solution in the strategic game induced by BST, then for every coalition $J \in C$ and every agent $i \in J$, we have:

$$U_i^{ST}(s) = U_J(s) = \max_{s'_J \in S_J} U_J(s'_J, s_{-J}) \quad (4)$$

Moreover, the fact that s is a unique Nash solution in the game induced by BST implies that every group $J \in C$ has a unique goal such that, for every joint strategy $s'_J \in S_J$:

$$\text{if } U_J(s_J, s_{-J}) = U_J(s'_J, s_{-J}), \text{ then } s_J = s'_J \quad (5)$$

In other words, agents in J will maximize the group utility given the strategies of other agents outside J, and they will do so by performing the unique joint strategy $s_J \in S_J$ that allows it.

Furthermore, one should note that, for every solution $s \in S$:

$$U_J(s) \leq \max_{s'_J \in S_J} U_J(s'_J, s_{-J}) \quad (6)$$

As a result from (4), (5) and (6), the current game $BST = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}, k\rangle$ can be transformed into a simplified strategic game $G^{bgi} = \langle Agt', \{S'_J | J \in Agt'\}, \{U'_J | J \in Agt'\}\rangle$ where each group $J \in C$ acts as a single agent, that is:

$$\begin{array}{lcl} Agt' & = & C \\ S'_J & = & S_J \\ U'_J(s') & = & U_J(s) \text{ where } s' \in S' \text{ and } s \in S \text{ are} \\ & & \text{such that } s'_K = s_K \text{ for every } K \in C. \end{array}$$

Note that through this game transformation, $1 \leq |Agt'| \leq |Agt|$. In the particular case where every agent in Agt is selfish (i.e., $k_i(\emptyset) = 1$ for every $i \in Agt$), then $|Agt'| = |Agt|$.

It is then straightforward to show, through (4), (5) and (6) and the definition of the Nash equilibrium, that, if a solution $s \in S$ is a unique Nash equilibrium in the game induced by BST, then a solution $s' \in S'$ is a unique Nash equilibrium in the transformed strategic game G^{bgi}, such that $s'_J = s_J$ for every $J \in C$. Conversely, as Constraint **C2** ensures that every group $J \in C$ has a unique goal in any game BST and G^{bst}, it implies that if a solution $s \in S$ is a unique Nash equilibrium in G^{bst}, then a solution $s' \in S'$ is a unique Nash equilibrium in the game induced by BST, such that $s'_J = s_J$ for every $J \in C$.

Let us now similarly transform the $BUTI$ structure.

Definition 4.4 means that every agent can only identify with a unique group. It follows from this definition that there exists a unique group identification state $t \in T$ such that:

$$\bigcup_{i \in Agt} t_i = Agt \text{ and } \bigcap_{i \in Agt} t_i = \emptyset \tag{7}$$

Given this unique group identification state t that satisfies (7), one can therefore define the set of active teams A as follows:

$$A = \{J \in 2^{Agt*} | \forall i \in J, t_i = J\} \tag{8}$$

It follows from (8) and Definition 4.4 that for every team $J \in 2^{Agt*}$ and every protocol $\alpha \in \Delta$, the expected value can be simplified as follows:

$$EV_J(\alpha) = U_J(s_1^{\alpha,t}, \ldots, s_n^{\alpha,t}) \tag{9}$$

One can note from (9) that, if protocol α is an UTI equilibrium, then, for every $J \in A$, we have:

$$EV_J(\alpha) = \max_{s'_J \in S_J} U_J(s'_J, s_{-J}^{\alpha,t}) \tag{10}$$

where $s_{-J}^{\alpha,t}$ is the combination of actions of all agents $i \in Agt \backslash J$ specified by the strategy profile $\alpha(t_i)$. Note that, for every coalition $K \in 2^{Agt*}$ such that $K \notin A$ (i.e., for every team nobody identifies with), and for every $\alpha \in \Delta$, $EV_K(\alpha)$ remains constant. This implies that every such group K is simply irrelevant to the computation of the UTI equilibrium.

As a result from (7), (8), (9) and (10), a structure $BUTI = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}, \{\Omega_i | i \in Agt\} \rangle$ can be transformed into a simplified strategic game $G^{buti} = \langle Agt'', \{S''_J | J \in Agt''\}, \{U''_J | J \in Agt''\} \rangle$ where every active team $J \in A$ acts as a single agent, that is:

$$\begin{array}{lcl} Agt'' & = & A \\ S''_J & = & S_J \\ U''_J(s') & = & U_J(s) \text{ where } s' \in S' \text{ and } s \in S \\ & & \text{such that } s'_K = s_K \text{ for every } K \in A. \end{array}$$

Note that through this game transformation, $1 \leq |Agt''| \leq |Agt|$. In the particular case where every agent is selfish (i.e., $t_i = \{i\}$ for every $i \in Agt$), then $|Agt''| = |Agt|$.

It is then straightforward to demonstrate from (10) and Definition 3.2 that if $s \in S''$ is the unique solution such that the protocol $\alpha \in \Delta$ defined by $\alpha(J) = s_J$ for every $J \in 2^{Agt*}$ is an UTI equilibrium in the $BUTI$ structure, then s is the unique Nash equilibrium in the transformed strategic game G^{buti}. Conversely, if $s \in S''$ is the unique Nash equilibrium in the transformed strategic game G^{buti}, then s is the unique solution such that the protocol $\alpha \in \Delta$ defined by $\alpha(J) = s_J$ for every $J \in 2^{Agt*}$ is an UTI equilibrium in the original $BUTI$ structure. Note that in the last case, a protocol α may not be the only UTI equilibrium in $BUTI$ as the strategy profile specified by $\alpha(K)$ is irrelevant and can therefore take any value from S_K for every $K \notin A$.

Finally, given a strategic game with group utility $G' = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}\rangle$ that satisfies Constraint **C2**, a binary unreliable team interaction structure $BUTI = \langle G', \{\Omega_i | i \in Agt\}\rangle$, and a social ties game $BST = \langle G', k\rangle$, it is immediate to show that the corresponding transformed games G^{buti} and G^{bst} are equivalent whenever $C = A$.

A.5 Proof of Theorem 4.6

We consider a three-player strategic game with group utility $G' = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}\rangle$ whose individual payoff matrix is represented in Table 5 for all players in $Agt = \{1, 2, 3\}$. Moreover, we take as the group payoff function the maximin principle (i.e., for every $J \in 2^{Agt*}$ and every $s \in S$, $U_J(s) = \min_{i \in J} U_i(s)$).

Actions			**Utilities**		
Player 1	Player 2	Player 3	Player 1	Player 2	Player 3
A	A	A	5	5	0
A	A	B	5	5	0
A	B	A	0	0	0
A	B	B	0	0	0
B	A	A	4	4	4
B	A	B	0	0	5
B	B	A	5	5	0
B	B	B	3	3	3

Table 5: A three-player coordination game

The main characteristics of this game is that G' does not satisfy Constraint **C2** from Definition 4.5. In fact, if Player 3 selects A, then the team $\{1, 2\}$ made of the two other players appears to have diverging goals, that is, solutions (A, A, A) and (B, B, A) are equally the best option for the group $\{1, 2\}$.

Let us now define a corresponding structure $BUTI = \langle G', \{\Omega_i | i \in Agt\}\rangle$ such that $\Omega_1(\{1, 2\}) = \Omega_2(\{1, 2\}) = 1$, and $\Omega_1(\{1, 3\}) = \Omega_3(\{1, 3\}) = \Omega_2(\{2, 3\}) = \Omega_3(\{2, 3\}) = \Omega_1(\{1, 2, 3\}) = \Omega_2(\{1, 2, 3\}) = \Omega_3(\{1, 2, 3\}) = 0$.

In this case, it is straightforward to show that every protocol $\alpha \in \Delta$ that is an UTI equilibrium is defined by $\alpha(\{1, 2\}) = (A, A)$ and $\alpha(\{3\}) = (B)$ [8].

[8] Note that every group $J \in 2^{Agt*}$ such that $\alpha(J) \neq \{1, 2\}$ and $\alpha(J) \neq \{3\}$ is irrelevant as J is an

Similarly, we define a binary social ties game $BST = \langle G', k \rangle$ and determine the various prediction such a game can make depending on the valuation of the function k. The corresponding sets of Nash equilibria in the game induced by BST can be found in Table 6.

k_{12}	k_{13}	k_{23}	k_{123}	Predicted outcomes
0	0	0	0	(A, A, A) (A, A, B) (B, B, B)
1	0	0	0	(A, A, A) (A, A, B) (A, B, A)
0	1	0	0	(B, B, B)
0	0	1	0	(A, A, A) (A, A, B) (B, B, B)
0	0	0	1	(B, A, A)

Table 6: Predictions (Nash equilibria in the induced game) based on social ties

One can therefore observe from Table 6 that there exits no BST structure such that the solution (A, A, B) is the unique Nash equilibrium in the game induced by BST.

inactive team.