When the Group Matters

A Game-Theoretic Analysis of Empathy, Team Reasoning, and Social Ties

E. Lorini[1] F. Moisan[1]

[1]Institut de Recherche en Informatique de Toulouse (IRIT)

IRIT, February 2013

Problems of Co-operation: Prisoner's Dilemma

- 2 suspects (Alice and Bob) arrested for a crime
- Prisoners isolated from each other
- Individual choice to betray the other (D) or stay silent (C)

		Bob: C	Bob: D
Alice	C	$(-2,-2)$	$(-10,0)$
	D	$(0,-10)$	$(-5,-5)$

- What should each prisoner do?
  - $\Rightarrow$ they should betray each other (choose D)!
  - $\Rightarrow$ In real-life: 30-40% of people choose C!

Problems of Co-operation: Prisoner's Dilemma

- 2 suspects (Alice and Bob) arrested for a crime
- Prisoners isolated from each other
- Individual choice to betray the other (D) or stay silent (C)

		Bob C	Bob D
Alice	C	$(-2,-2)$	$(-10,0)$
Alice	D	$(0,-10)$	$(-5,-5)$

- What should each prisoner do?
  - $\Rightarrow$ they should betray each other (choose D)!
  - $\Rightarrow$ In real-life: 30-40% of people choose C!

Problems of Co-operation: Prisoner's Dilemma

- 2 suspects (Alice and Bob) arrested for a crime
- Prisoners isolated from each other
- Individual choice to betray the other (D) or stay silent (C)

		Bob C	Bob D
Alice	C	$(-2, -2)$	$(-10, 0)$
Alice	D	$(0, -10)$	$(-5, -5)$

- What should each prisoner do?
  - $\Rightarrow$ they should betray each other (choose D)!
  - $\Rightarrow$ In real-life: 30-40% of people choose C!

Problems of Co-operation: Hi-Lo matching game

- Alice and Bob have to paint a room together
- Each individual is responsible to buy a tin of paint
- Individual choice to buy blue color paint (B) or green color paint (G)

		Bob B	Bob G
Alice	B	$(10, 10)$	$(0, 0)$
Alice	G	$(0, 0)$	$(5, 5)$

- What should each individual do?
  - $\Rightarrow$ No unique rational solution!
  - $\Rightarrow$ Counter-intuitive!

Problems of Co-operation: Hi-Lo matching game

- Alice and Bob have to paint a room together
- Each individual is responsible to buy a tin of paint
- Individual choice to buy blue color paint (B) or green color paint (G)

		Bob B	Bob G
Alice	B	$(10, 10)$	$(0, 0)$
Alice	G	$(0, 0)$	$(5, 5)$

- What should each individual do?
  - $\Rightarrow$ No unique rational solution!
  - $\Rightarrow$ Counter-intuitive!

Formalizing social interactions

Definition (Standard Strategic Game)

$G = \langle Agt, \{S_i | i \in Agt\}, \{U_i | i \in Agt\}\rangle$ where:

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

Theories of social (other-regarding) preferences

e.g., A model of fairness [Charness & Rabin,2002]:

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

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

- $\Rightarrow$ Can predict cooperation in Prisoner's Dilemma game!
- $\Rightarrow$ Remains indecisive in Hi-Lo game!

$\Rightarrow$ No existing valid theory based on social preferences!

Theories of social (other-regarding) preferences

e.g., A model of fairness [Charness & Rabin,2002]:

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

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

- $\Rightarrow$ Can predict cooperation in Prisoner's Dilemma game!
- $\Rightarrow$ Remains indecisive in Hi-Lo game!

$\Rightarrow$ No existing valid theory based on social preferences!

Outline

1. Empathetic Preferences
2. Team Reasoning
3. Social Ties
4. Conclusion

Defining empathy

- Many definitions!
  - $\Rightarrow$ Here: economic concept [Binmore,1994,2005]

- Empathy $\Leftrightarrow$ combining my own preferences with my preferences when imagining myself to be in another agent's position
  - $\Rightarrow$ I must consider the other's preferences!
  - $\neq$ Golden Rule: *"One should treat others as one would like others to treat oneself"*
  - $\Rightarrow$ I must separate my preferences from the other's ($\neq$ sympathy)

- How to define empathetic preferences?
  - $\Rightarrow$ behind the *veil of ignorance* [Rawls,1971]
  - $\Rightarrow$ while ignoring one's personal identity

Defining empathy

- Many definitions!
  - $\Rightarrow$ Here: economic concept [Binmore,1994,2005]

- Empathy $\Leftrightarrow$ combining my own preferences with my preferences when imagining myself to be in another agent's position
  - $\Rightarrow$ I must consider the other's preferences!
  - $\neq$ Golden Rule: *"One should treat others as one would like others to treat oneself"*
  - $\Rightarrow$ I must separate my preferences from the other's ($\neq$ sympathy)

- How to define empathetic preferences?
  - $\Rightarrow$ behind the *veil of ignorance* [Rawls,1971]
  - $\Rightarrow$ while ignoring one's personal identity

Defining empathy

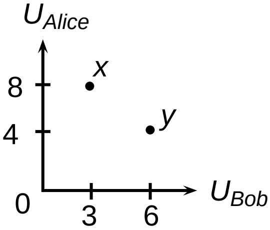
- Many definitions!
  - $\Rightarrow$ Here: economic concept [Binmore,1994,2005]

- Empathy $\Leftrightarrow$ combining my own preferences with my preferences when imagining myself to be in another agent's position
  - $\Rightarrow$ I must consider the other's preferences!
  - $\neq$ Golden Rule: *"One should treat others as one would like others to treat oneself"*
  - $\Rightarrow$ I must separate my preferences from the other's ($\neq$ sympathy)

- How to define empathetic preferences?
  - $\Rightarrow$ behind the *veil of ignorance* [Rawls,1971]
  - $\Rightarrow$ while ignoring one's personal identity

Interpersonal Comparison of Utility

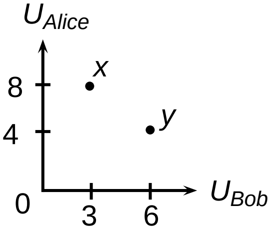
Let x = “playing squash” and y = “playing tennis”:



- Does one prefer (1) playing squash while being Alice or (2) playing tennis while being Bob?
  - (1)>(2) if one is indifferent for being either Alice or Bob
  - (1)<(2) if one always prefers to be Bob

Interpersonal Comparison of Utility

Let x = "playing squash" and y = "playing tennis":



- Does one prefer (1) playing squash while being Alice or (2) playing tennis while being Bob?
  - (1)$>$(2) if one is indifferent for being either Alice or Bob
  - (1)$<$(2) if one always prefers to be Bob

Modeling empathetic preferences

Definition (Game with Empathetic Preferences)

$EM = \langle Agt, \{S_i | i \in Agt\}, \{U_i | i \in Agt\}, \{U^E_{i,j} | i, j \in Agt\}\rangle$ where:

- $\langle Agt, \{S_i | i \in Agt\}, \{U_i | i \in Agt\}\rangle$ is a standard strategic game;
- $U^E_{i,j} : S \to \mathbb{R}$ is a total function defining agent i's empathetic utility for being agent j such that:
  - **C1** there exists $\alpha \in \mathbb{R}_+ \backslash \{0\}$ and $\beta \in \mathbb{R}$ such that, for every $s \in S$, $U^E_{i,j}(s) = \alpha \times U_j(s) + \beta$.

Modeling empathetic preferences

How to combine *i*'s empathetic preferences for every solution *s* behind the *veil of ignorance*?

- According to Harsanyi [Harsanyi,1986], *i* assigns (equal) probabilities to each event:

$$U^{H}_{i,J}(s) = \frac{1}{|J|} \cdot \sum_{j \in Agt} U^{E}_{i,j}(s)$$

- According to Rawls [Rawls,1971], *i* is not able to assign probabilities:

$$U^{R}_{i,J}(s) = \min_{j \in J} U^{E}_{i,j}(s)$$

- Who is right?
  - $\Rightarrow$ Depends on context!
  - $\Rightarrow$ Theory of external enforcement [Binmore,2005]

Modeling empathetic preferences

How to combine *i*'s empathetic preferences for every solution *s* behind the *veil of ignorance*?

- According to Harsanyi [Harsanyi,1986], *i* assigns (equal) probabilities to each event:

$$U^H_{i,J}(s) = \frac{1}{|J|} \cdot \sum_{j \in Agt} U^E_{i,j}(s)$$

- According to Rawls [Rawls,1971], *i* is not able to assign probabilities:

$$U^R_{i,J}(s) = \min_{j \in J} U^E_{i,j}(s)$$

- Who is right?
  - $\Rightarrow$ Depends on context!
  - $\Rightarrow$ Theory of external enforcement [Binmore,2005]

Modeling empathetic preferences

How to combine *i*'s empathetic preferences for every solution *s* behind the *veil of ignorance*?

- According to Harsanyi [Harsanyi,1986], *i* assigns (equal) probabilities to each event:

$$U^{H}_{i,J}(s) = \frac{1}{|J|} \cdot \sum_{j \in Agt} U^{E}_{i,j}(s)$$

- According to Rawls [Rawls,1971], *i* is not able to assign probabilities:

$$U^{R}_{i,J}(s) = \min_{j \in J} U^{E}_{i,j}(s)$$

- Who is right?
  - $\Rightarrow$ Depends on context!
  - $\Rightarrow$ Theory of external enforcement [Binmore,2005]

Empathy Equilibrium

- Behind the veil of ignorance:
  - $\Rightarrow$ Assumption that everybody shares the same empathetic preferences
  - $\Rightarrow$ No need for strategic thinking!

Definition (Empathy Equilibrium)

$s \in S$ is an empathy equilibrium in EM iff:

$$s \in argmax_{s' \in S} U_i^X(s') \text{ for every } i \in Agt$$

where $U_i^X = U_{i,Agt}^H$ or $U_i^X = U_{i,Agt}^R$

Empathy Equilibrium

- Behind the veil of ignorance:
  - $\Rightarrow$ Assumption that everybody shares the same empathetic preferences
  - $\Rightarrow$ No need for strategic thinking!

Definition (Empathy Equilibrium)

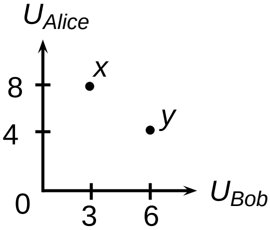
$s \in S$ is an empathy equilibrium in EM iff:

$$s \in argmax_{s' \in S} U_i^X(s') \text{ for every } i \in Agt$$

where $U_i^X = U_{i,Agt}^H$ or $U_i^X = U_{i,Agt}^R$

Empathy Equilibrium

Choose between x (“playing squash”) and y (“playing tennis”):



- Assuming indifference between being either Alice or Bob:
  - x = empathy equilibrium à la Harsanyi
  - y = empathy equilibrium à la Rawls

Remarks & Limitations

Binmore's model of empathetic preferences:

- allows to explain cooperative behavior!
- does not incorporate strategic reasoning!
  - $\Rightarrow$ it cannot explain mutual defection in the PD game!
  - $\Rightarrow$ it cannot model competing coalitions in larger games!
- does not allow for a clear quantification of empathy!
  - $\Rightarrow$ How thick is the *veil of ignorance*?
  - $\Rightarrow$ Nash equilibrium Vs. empathy equilibrium

Formalizing games with group utility

Definition (Game with Group Utility)

$G' = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}\rangle$ where:

- $Agt = \{1, \ldots, n\}$ 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.

Examples of group utility functions U_J:

- pure utilitarianism: $U_J(s) = \sum_{i \in J} U_i(s)$
- the *maximin* principle: $U_J(s) = \min_{i \in J} U_i(s)$
- induced by aligned empathetic preferences:
  - i.e., $U^E_{i,k} = U^E_{j,k}$ for every $i, j, k \in Agt$
  - $\Rightarrow U_J(s) = U^X_{i,J}(s)$ for some $i \in J$

Formalizing games with group utility

Definition (Game with Group Utility)

$G' = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}\rangle$ where:

- $Agt = \{1, \ldots, n\}$ 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.

Examples of group utility functions U_J:

- pure utilitarianism: $U_J(s) = \sum_{i \in J} U_i(s)$
- the *maximin* principle: $U_J(s) = \min_{i \in J} U_i(s)$
- induced by aligned empathetic preferences:
  - i.e., $U^E_{i,k} = U^E_{j,k}$ for every $i, j, k \in Agt$
  - $\Rightarrow U_J(s) = U^X_{i,J}(s)$ for some $i \in J$

Outline

Tuomela's theory of team reasoning

Tuomela's *I-mode* / *we-mode* distinction [Tuomela,2010]:

- *I-mode* = conceive the situation as a decision making problem for individual agents (reasoning as a private person)
  - *plain I-mode* = make a decision with the **individual intention** to maximize self-interest (cf., classical economic theory)
  - *pro-group I-mode* = make a decision with the **individual intention** to maximize the group utility (cf., theories of social preferences)
- *we-mode* = conceive the situation as a decision making problem for the group conceived as an agent
  - ⇒ **collective intention**!

Tuomela's theory of team reasoning

Tuomela's *I-mode* / *we-mode* distinction [Tuomela,2010]:

- *I-mode* = conceive the situation as a decision making problem for individual agents (reasoning as a private person)
    - *plain I-mode* = make a decision with the **individual intention** to maximize self-interest (cf., classical economic theory)
    - *pro-group I-mode* = make a decision with the **individual intention** to maximize the group utility (cf., theories of social preferences)
- *we-mode* = conceive the situation as a decision making problem for the group conceived as an agent
    - $\Rightarrow$ **collective intention**!

Tuomela's theory of team reasoning

Tuomela's *I-mode* / *we-mode* distinction [Tuomela,2010]:

- *I-mode* $=$ conceive the situation as a decision making problem for individual agents (reasoning as a private person)
  - *plain I-mode* $=$ make a decision with the **individual intention** to maximize self-interest (cf., classical economic theory)
  - *pro-group I-mode* $=$ make a decision with the **individual intention** to maximize the group utility (cf., theories of social preferences)
- *we-mode* $=$ conceive the situation as a decision making problem for the group conceived as an agent
  - $\Rightarrow$ **collective intention**!

Sugden's theory of team reasoning

According to [Sugden,2003,2007]:

Statement (Simple team reasoning)

*If **I** believe that:*

- ***I** am a member of group J.*
- *It is common knowledge among all members of J that **we** all identify with J.*
- *It is common knowledge among all members of J that **we** all want U_J to be maximized.*
- *It is common knowledge among all members of J that solution s uniquely maximizes U_J.*

*Then **I** should choose **my** strategy in s.*

Bacharach's theory of team reasoning

Concept of *unreliable team interaction* [Bacharach,1999]:

- $\Rightarrow$ A game theoretic model of team reasoning!
- Agents "know" their own type of reasoning (e.g., *I-mode*/*we-mode*)
  - $\Rightarrow$ a psychological factor, prior to any rational choice
- Agents can be uncertain about others' types of reasoning!
- $\Rightarrow$ Agents maximize their expected utility depending on their type

Bacharach's theory of team reasoning

Definition (Unreliable Team Interaction)

$UTI = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}, \{\Omega_i | i \in Agt\}\rangle$ where:

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

- + Notion of a *team protocol* $\alpha \in \Delta$
  - $\Rightarrow \alpha$ specifies a strategy for each team $J \in 2^{Agt*}$
  - $\neq$ strategy profile s in classical game theory

Bacharach's theory of team reasoning

Definition (Unreliable Team Interaction)

$UTI = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}, \{\Omega_i | i \in Agt\}\rangle$ where:

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

- + Notion of a *team protocol* $\alpha \in \Delta$
  - $\Rightarrow \alpha$ specifies a strategy for each team $J \in 2^{Agt*}$
  - $\neq$ strategy profile s in classical game theory

Bacharach's theory of team reasoning

Expected value of protocol $\alpha \in \Delta$ for some team $J \in 2^{Agt*}$:

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

Definition (Uti Equilibrium)

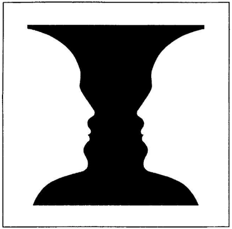
A protocol α is an *UTI* equilibrium if and only if:

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

- Equilibrium solution $\Leftrightarrow$ no individual **AND** no team can increase expected value by unilaterally deviating
- $\Rightarrow$ Equivalent to finding a Nash equilibrium in a transformed n-player game with $n = |2^{Agt*}|$

Limitations to Bacharach's model

- Complexity of computing an equilibrium solution
- Interpretation of exogenous probability distribution Ω_i
  - e.g., intrinsic to game structure? the players? social ties?
- Interpretation of the group utility function
  - e.g., pure utilitarianism? *maximin* principle?
- Only binary types of reasoning (*I-mode*/*we-mode*)
  - $\Rightarrow$ No gradual group identification (at best vacillations between modes)!



Limitation to all theories of team reasoning

A counter-example to all theories of team reasoning:

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

- In *I-mode* $\Rightarrow$ Bob will play A
- In *we-mode* $\Rightarrow$ Bob will play B
  - Following either utilitarianism or the *maximin* principle
- $\Rightarrow$ Bob will *never* play C (counter-intuitive!)

Limitation to all theories of team reasoning

A counter-example to all theories of team reasoning:

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

- In *I-mode* $\Rightarrow$ Bob will play A
- In *we-mode* $\Rightarrow$ Bob will play B
  - Following either utilitarianism or the *maximin* principle
- $\Rightarrow$ Bob will *never* play C (counter-intuitive!)

Outline

Definition of social ties

- No formal definition so far!
- $\Rightarrow$ friends, married couples, family relatives, colleagues, classmates, etc. . .
- Some psychological foundations:
  - Social features that define one's social identity
  - e.g., to identify as a student of Toulouse university, a supporter of Barcelona's soccer team, a Democrat, . . .
- Some epistemic foundations
  - *Minimal criterion*: a social tie between i and $j \Leftrightarrow i$ and j commonly believe that they share the same social features defining their social identities.

Definition of social ties

- No formal definition so far!
- $\Rightarrow$ friends, married couples, family relatives, colleagues, classmates, etc. . .
- Some psychological foundations:
  - Social features that define one's social identity
  - e.g., to identify as a student of Toulouse university, a supporter of Barcelona's soccer team, a Democrat, . . .
- Some epistemic foundations
  - *Minimal criterion*: a social tie between i and j $\Leftrightarrow$ i and j commonly believe that they share the same social features defining their social identities.

Definition of social ties

- No formal definition so far!
- $\Rightarrow$ friends, married couples, family relatives, colleagues, classmates, etc. . .
- Some psychological foundations:
  - Social features that define one's social identity
  - e.g., to identify as a student of Toulouse university, a supporter of Barcelona's soccer team, a Democrat, . . .
- Some epistemic foundations
  - *Minimal criterion*: a social tie between i and $j \Leftrightarrow i$ and j commonly believe that they share the same social features defining their social identities.

Definition of social ties

How to quantify the social tie between *i* and *j*?

- Quantity and importance of shared social features that define both *i* and *j*'s social identities
  - How many social features *i* and *j* share?
  - Are shared social features as important for *i* as for *j*?
- Quantity and quality of past interactions between *i* and *j*
  - How often *i* and *j* had meaningful interactions with each other?
  - e.g., exchanging ideas, opinions, sharing positive emotions, ...

Definition of social ties

How to quantify the social tie between i and j?

- Quantity and importance of shared social features that define both i and j's social identities
  - How many social features i and j share?
  - Are shared social features as important for i as for j?

- Quantity and quality of past interactions between i and j
  - How often i and j had meaningful interactions with each other?
  - e.g., exchanging ideas, opinions, sharing positive emotions, …

Modeling social ties

Definition (Social Ties Game)

$ST = \langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\}, \{k_i | i \in Agt\} \rangle$ where:

- $\langle Agt, \{S_i | i \in Agt\}, \{U_J | J \in 2^{Agt*}\} \rangle$ is a strategic game with group utility;
- Every k_i is a total function $k_i : 2^{Agt \setminus \{i\}} \to [0, 1]$, such that:
  C3 $\sum_{J \in 2^{Agt \setminus \{i\}}} k_i(J) = 1$
  C4 if $i, j \in Agt$, $i \neq j$, and $J \subseteq Agt \setminus \{i, j\}$, then $k_i(J \cup \{j\}) = k_j(J \cup \{i\})$

- Intuitions:
  - $k_i(J)$ = agent i's social tie with group J
  - **C3** $\Rightarrow$ a distribution of social ties!
  - **C4** $\Rightarrow$ social ties are bilateral!

Modeling social ties

Assumptions:

- Social ties affect social preferences!
  - $\Rightarrow$ individual intentionality ($\neq$ team reasoning)
- Inspired by Binmore's theory of empathetic preferences!
  - social ties $=$ measure of thickness of the *veil of ignorance*

Definition (Social Ties Utility)

Given a social ties game ST, for every strategy profile $s \in S$, the social ties utility function of player i is given by:

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

Intuition:

- $\Rightarrow$ the more i is tied with J, the more J's welfare matters in i's preferences

Modeling social ties

Assumptions:

- Social ties affect social preferences!
  - $\Rightarrow$ individual intentionality ($\neq$ team reasoning)
- Inspired by Binmore's theory of empathetic preferences!
  - social ties $=$ measure of thickness of the *veil of ignorance*

Definition (Social Ties Utility)

Given a social ties game ST, for every strategy profile $s \in S$, the social ties utility function of player i is given by:

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

Intuition:

- $\Rightarrow$ the more i is tied with J, the more J's welfare matters in i's preferences

Illustration: Social Ties in 2-player Games

$\Rightarrow$ A simplified utility function:

Given a social ties game ST with $Agt = \{i, j\}$, for every $s \in S$:

$$U_i^{ST}(s) = (1 - k_{ij}) \cdot U_i(s) + k_{ij} \cdot \max_{s'_j \in S_j} U_{\{i,j\}}(s_i, s'_j)$$

(where $k_{ij} = k_i(\{j\})$)
If $k_{ij} = 1$:

- $\Rightarrow$ assumes partner would "do the right thing for the group"
- $=$ individual decision problem!

Illustration: Social Ties in 2-player Games

$\Rightarrow$ Transformation of utilities (assuming $st_{ij} = 1$):

	Bob: C	Bob: D
Alice: C	$(-2, -2)$	$(-10, 0)$
Alice: D	$(0, -10)$	$(-5, -5)$

$\Rightarrow$

	Bob: C	Bob: D
Alice: C	$(-2, -2)$	$(-2, -5)$
Alice: D	$(-5, -2)$	$(-5, -5)$

	Bob: B	Bob: G
Alice: B	$(10, 10)$	$(0, 0)$
Alice: G	$(0, 0)$	$(5, 5)$

$\Rightarrow$

	Bob: B	Bob: G
Alice: B	$(10, 10)$	$(10, 5)$
Alice: G	$(5, 10)$	$(5, 5)$

Relationship with empathetic preferences

Theorem

Given:

- *a game EM with aligned empathetic preferences (i.e., $U^E_{i,k} = U^E_{j,k}$ for every $i, j, k \in Agt$)*
- *the strategic game G^{em} induced by EM (i.e., $U_J(s) = U^X_{i,J}(s)$ for some $i \in Agt$)*
- *the social ties game $ST = \langle G^{em}, \{k_i | i \in Agt\}\rangle$ s.t. $k_i(Agt \setminus \{i\}) = 1$ for every $i \in Agt$*

Finding an empathy equilibrium in EM

$\Leftrightarrow$

Finding a Nash equilibrium in the game induced by ST.

Relationship with Bacharach's team reasoning

Definition (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 $i \in Agt$, $\Omega_i(t_i) = 1$;

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

- for every $i \in Agt$ and every $J \subseteq Agt$, $k_i(J) \in \{0, 1\}$.

Relationship with Bacharach's team reasoning

Theorem

Given:

- *a strategic game with group utility G' with $|Agt| = 2$,*
- *a binary social ties game $BST = \langle G', \{k_i | i \in Agt\}\rangle$,*
- *a binary UTI structure $BUTI = \langle G', \{\Omega_i | i \in Agt\}\rangle$,*

if $k_i(Agt \backslash \{i\}) = \Omega_i(Agt)$ for every $i \in J$, then:

Finding a unique Nash equilibrium in game induced by BST
$\Leftrightarrow$
Finding a unique UTI equilibrium in BUTI

Relationship with Bacharach's team reasoning

Theorem

Given:

- *a strategic game with group utility G' with $|Agt| > 2$,*
- *a binary social ties game $BST = \langle G', \{k_i | i \in Agt\}\rangle$,*
- *a binary UTI structure $BUTI = \langle G', \{\Omega_i | i \in Agt\}\rangle$,*

if $k_i(J \backslash \{i\}) = \Omega_i(J)$ for every $J \in 2^{Agt*}$ and $i \in J$, then:

Finding a unique Nash equilibrium in game induced by BST
$\Rightarrow$
Finding a unique UTI equilibrium in BUTI

Relationship with Bacharach's team reasoning

Theorem

Given:

- *a strategic game with group utility G' with $|Agt| > 2$,*
- *a binary social ties game $BST = \langle G', \{k_i | i \in Agt\} \rangle$,*
- *a binary UTI structure $BUTI = \langle G', \{\Omega_i | i \in Agt\} \rangle$,*

if $k_i(J \setminus \{i\}) = \Omega_i(J)$ for every $J \in 2^{Agt*}$ and $i \in J$, then:

Finding a unique UTI equilibrium in BUTI
$\not\Rightarrow$
Finding a unique Nash equilibrium in game induced by BST

$\Rightarrow$ cf., games with ambiguous group intentions

Relationship with Bacharach's team reasoning

Different predictions in a 2-player game G':

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

- Assumption: group utility function = pure utilitarianism or *maximin* principle
- Can Bob play C?
  - According to any structures $UTI = \langle G', \{\Omega_i | i \in Agt\}\rangle$: No!
  - According to some game $ST = \langle G', \{k_i | i \in Agt\}\rangle$: Yes!

Relationship with Bacharach's team reasoning

Different predictions in a 2-player game G':

$$\begin{array}{cc|c|c|} & & \multicolumn{2}{c}{\text{Alice}} \\ & & A & B \\ \hline \text{Bob} & A & (2,-1) & (0,1) \\ \hline & B & (1,1) & (1,0) \\ \hline \end{array}$$

- Assumption: group utility function = pure utilitarianism
- Can Bob play A?
  - According to any game $ST = \langle G', \{k_i | i \in Agt\}\rangle$: No!
  - According to some structure $UTI = \langle G', \{\Omega_i | i \in Agt\}\rangle$: Yes!

Outline

Conclusion

- Theories of empathetic preferences and team reasoning:
  - $\Rightarrow$ can explain co-operation
  - $\Rightarrow$ limited for modeling complex collective behavior!
- Our theory of social ties:
  - $\Rightarrow$ simple and intuitive
  - $\Rightarrow$ an alternative theory of social preferences
  - $\Rightarrow$ collective reasoning based on individual intentionality
  - $\Rightarrow$ fills the gap between defection and cooperation
  - $\Rightarrow$ can model competing coalitions (whose intersection may be non-empty)

Conclusion

- Future work:
  - interpreting social ties towards groups in terms of social ties between individuals
  - Experimental study to distinguish models of social ties Vs. team reasoning
  - Collective reasoning in sequential games
  - Epistemic analysis of social ties
  - Dynamics of social ties & group formation
  - Consider other solution concepts in social ties game
  - ...