Fixing group STIT

LIRa Seminar

07.09.2023

Outline

- Intro to group STIT
- Axiomatizability and decidability issues
- Philosophical issues: shared agency, group formation
- Technical solutions: decidable fragments, enriched language
- Plan-restrcited group STIT logic (this name is awful, I know)

Intro to group STIT

Language

For a finite set of agents $Ags = \{1, \ldots, n\}$ and a countable set of propositions $Var = \{p_1, p_2, \ldots\}$ we define a language of group STIT $\mathcal{STIT}_G$:

$$\phi := p \mid \neg\phi \mid \phi \vee \phi \mid \Box\phi \mid [i]\phi \mid [G]\phi$$

where $p \in Var, i \in Ags, G \subseteq Ags$

Models

A class of group STIT frames $\mathbf{STIT}_G$ consists of structures of the form

$$F = (W, \{R_i\}_{i \in Ags}, R_{\Box})$$

- $W \neq \emptyset$ is a set of possible worlds
- $R_{\Box}$ is an universal relation
- R_i is an equivalence relation on W
  - Independence of agents: $\bigcap\limits_{i \in Ags}^{w \in W} R_i(w) \neq \emptyset$
- $R_G = \bigcap\limits_{i \in G} R_i$ for all $G \subseteq Ags$. We will presuppose that $R_{Ags} = Id_W$

rock, paper	paper, paper	scissors, paper
rock, rock	paper, rock	scissors, rock
rock, scissors	paper, scissors	scissors, scissors

What makes it cool: expressivity

- In group STIT we can express that a group has an/no abilitiy to ensure ϕ, just like in Coalition logic: $\Diamond[\{blue\}]rock \land \neg\Diamond[\{blue\}]win$
- But not only we can say about group's abilities, but we can say what group has exactly done: $[Ags]win \land \Diamond[Ags]\neg win$
- Given temporal extensions, we can embed Coalition logic andd Alternating-time temporal logic in temporal group STIT

What makes it not that cool: meta-properties 😢

- The logic cannot be finitely axiomatized if $Card(Ags) \geq 3$
- SAT problem is undecidable in the same cases
- **But WHY?**

Axiomatizibility and decidability issues

Smells like product

Let's think about a set of every agent's available actions - $\{X \subseteq W \mid \exists w \in W\big(X = R_i(w)\big)\}$ - as of a dimension of some multidimensional space. Then, a set of points of that space would be $\{X \mid \exists w \in W(R_{Ags}(w) = X)\}$. We can construct a product Kripke frame out of it.

Recap: product Kripke frames

Given two Kripke frames $F_1 = (W_1, R_1)$ and $F_2 = (W_2, R_2)$, the product Kripke frame is $F_1 \times F_2 = (W_1 \times W_2, \vec{R}_1, \vec{R}_2)$, where

$$\vec{R}_1 = \{((x, y), (z, y)) \mid x R_1 z \land y = y\}$$

$$\vec{R}_2 = \{((x, s), (x, d)) \mid x = x \land s R_2 d\}$$

Adding dimensions

Assume we have n-many Kripke frames $F_1, \ldots, F_n$. Their product is defined as $\prod_{1 \leq i \leq n} F_n = (W_1 \times \ldots \times W_n, \vec{R}_1, \ldots, \vec{R}_n)$, where

$$\vec{w}_1 \vec{R}_i \vec{w}_2 \text{ iff } \forall j \neq i : \vec{w}_1(j) = \vec{w}_2(j) \ \wedge \ \vec{w}_1(i) R_i \vec{w}_2(i)$$

Putting product S5 in STIT

So assume we have some group STIT frame $F = (W, \{R_i\}_{i \in Ags}, R_\Box)$. Construct $Card(Ags)$-many frames from it as

$$F_i = (W_i, R_i^\forall)$$

- $W_i = \{R_i(w) \mid w \in W\}$
- $R_i^\forall = W_i \times W_i$

And now take a product of this frames!

The tricky part of $\prod_{i \in Ags} F_i$ is what $\vec{R}_i$ are. Informally, if $\vec{w}_1 \vec{R}_i \vec{w}_2$, then in both worlds all agents, except i, are doing the same thing. Basically, it represents a set of possible outcomes of some action by $Ags \setminus \{i\}$. So $[i]$ modality in product S5 corresponds to $[Ags \setminus \{i\}]$ modality. Luckily, we have this nice validity in group STIT:

$$\mathbf{STIT}_G \models [G_1 \cap G_2]\phi \leftrightarrow [G_1][G_2]\phi$$

So by nesting $[i]$ modalities from product S5 language, we can always get any $[G]$ modality from group STIT, including singletons and $[Ags]$ (think of $\Box$ as of $[\emptyset]$). After that it is a matter of routine to show that validity in product S5 and validity in group STIT are actually equal

$S5^3$ 😱

- No finite axiom system
- You can encode tiling problem in it, so it is undecidable

Philosophical issues: shared agency, group formation

To recall: in STIT a group is any subset of agents, and the repertoire of group actions is all possible combinations of actions available for group's members. Formally, it is defined as $R_G = \bigcap_{i \in G} R_i$. It leads for the next validity:

$$\mathbf{STIT}_G \models [G]\phi \rightarrow [G']\phi \text{ for all } G, G' \text{ such that } G \subseteq G'$$

Example

- 🌎 $\models [\{Russel, Whitehead\}]\text{Principia}$
- 🌎 $\models [\{\text{Mouvement de Libé ration du Congo}\}]\text{XXI century genocide}$

So that,

- 🌎 $\models [\{\text{Mouvement de Libé ration du Congo}\} \cup \{\text{me}\}]\text{XXI century genocide}$
- 🌎 $\models [\{Russel, Whitehead\} \cup \{\text{me}\}]\ \text{Principia}$
- 🧑‍⚖️ at Den Haag: Daniil has participated in some war crimes, but at least he has made an outstanding contribution to math, philosophy and logic

The rest of my slides is an attempt not to go to The United Nations Detention Unit near Den Haag

Technical solutions

Extending the language with difference operator

All the features of group STIT frames correspond to some single Sahlqvist group STIT formula, except the only one:

Frame feature	Formula
$R_i, R_G, R_\Box$ are equivalence relations	S5 for $[i], [G], \Box$
$R_i \subseteq R_\Box$	$\Box\phi \to [i]\phi$
Independence of agents	$\Diamond[i_1]\phi_1 \wedge \ldots \Diamond \wedge [i_n]\phi_n \to \Diamond([i_1]\phi_1 \wedge \ldots \wedge [i_n]\phi_n)$
$R_G \subseteq \bigcap_{i \in G} R_i$	$[G]\phi \to [G']\phi \text{ for any } G \subseteq G'$
$\bigcap_{i \in G} R_i \subseteq R_G$	😔

What makes $R_G = \bigcap_{i \in G} R_i$ troublesome?

It corresponds to the next FO-formula

$$\forall x, w_1, \dots, w_n (x R_G w_1 \leftrightarrow (\bigwedge_{i \in G}^{1 \leq j \leq n} x R_i w_j \wedge w_1 = \dots = w_n))$$

and this formula would sit in the guarded fragment if only we could use $=$ as a guard

Group STIT with elsewhere modality

- $\phi := p \mid \neg\phi \mid \phi \vee \phi \mid \Box\phi \mid [i]\phi \mid [G]\phi \mid D\phi$
- $M, w \models D\phi \text{ iff } \forall w'(w \neq w' \rightarrow M, w' \models \phi)$

Finite axiom system with unorthodox rules

- S5 for $[i]$, $[G]$, $\Box$
- $\Box\phi \to X\phi$, where $X \in \{[i], [G], D\}$
- Ind: $(\Diamond[i_1]\phi_1 \wedge \ldots \wedge \Diamond[i_n]\phi_n) \to \Diamond([i_1]\phi_1 \wedge \ldots \wedge [i_n]\phi_n)$
- $[G]\phi \to [G']\phi$ for all $G \subseteq G$'
- Axioms for D:
  - (D1) $\phi \to D\hat{D}\phi$
  - (D2) $\hat{D}\hat{D}\phi \to (\phi \vee \hat{D}\phi)$
  - (D3) $\Diamond\phi \to (\phi \vee \hat{D}\phi)$
  - (IRD) $p \wedge D\neg p \to \phi \vdash \phi$ if $p \notin \phi$
- $(\Diamond(\phi \wedge D\neg\phi) \wedge \bigwedge_{i \in G} \langle i \rangle \phi) \to \langle G \rangle \phi$

Rigid solution: restrict the language

Let $Ags = \{1. \dots, n\}$ be a set of agents and $\Gamma \subseteq 2^{Ags}$ be a collection of Ags subsets that constitute groups. Then we can define a language $\mathcal{STIT}_G^-$:

$$\phi := p \mid \neg\phi \mid \phi \vee \phi \mid \Box\phi \mid [i]\phi \mid [G]\phi$$

where $i \in Ags, G \in \Gamma$. Given some properties of Γ -- vaguely stated, if Γ is small enough -- we get finitely axiomatizable and decidable fragments of group STIT. For futher details: *Schwarzentruber, F. "Complexity results of STIT fragments"*.

Philosophical problem: what is a social group?

One of the obvious objections we can make is that not every set of agents should be regarded as a collective, just like not every combination of individual actions is a collective act. But what makes agency collective in nature? Let's meditate at it a bit. For example, we say that two people are going to the museum together. What do we mean?

- The first person is going and the second is going and that's all? $\textbf{Not yet}$, dozens of people are going to the museum as well, but they do not go there $\textit{together}$
- They both should know/believe that they are going? $\textbf{Not yet}$, I may know that some person is going to the same museum as me, and they might know it as well (it may even be a common knowledge), but we still may not be going $\textit{together}$

Philosophical problem: what is a social group?

- They both intend to go? They both intend that the other is going? They both intend that the other intends that the other intends that...?
- They both $jointly\ intend$ to go to the museum? But what does it mean to jointly intend?

Collective intentionality

There is a whole philosophical movement inside social ontology - collective intentionality theorists. These guys argue that collective intentions are special attitudes which could not be reduced to group members' individual intentions/beliefs/desires or any aggregation of them. From this point of view, no social group can be regarded as such if their members have no joint/collective intention whatsoever. But what are those joint intentions and how groups adopt them?

Witty quote

"The formation of a joint intention (or plan) is based on their various personal and, especially joint desires and mutual and other beliefs. In this sense a joint intention can be said to “summarize” or reflect the motivation underlying joint action. Of course, this final motivation underlying the joint intention need not be anything like an aggregation of private motivations but may instead be a compromise based on discussion, negotiation, or bargaining"

R. Tuomela, *We-Intentions Revisited*

Collective intentionality, simplified

In order for a set of agents to be an actual active group, they should

- Have at least some common ground, be it some common beliefs or desires
- Adopt some joint intention based on that common ground
- Commit to it and coordinate their action in accordance to it

Collective intentionality, simplified

Doing without explicit intentions

- The most straightforward way to formalize such a view in STIT framework would be adding some modality for joint intentions. We would like to do without it and stay causal only
- If a group has collectively adopted some goal, it will establish some explicit or implicit code of conduct: group members at least should not act against the joint intention. Otherwise the very nature of their collective act will be broken

Implicit joint intentions in STIT: example

Imagine that two agents - blue and red - were captured by some cruel creature. The creature tells them: "Every morning you are to play rock-paper-scissors. If one of you loses, he/she will be killed. If you both shoot the same figure, we will wait till the next day and play again after the sunrise".

If blue and red have a possibility to communicate and a common ground - mutual belief that both of them should survive - they will adopt a joint intention to shoot the same figure (let's say, paper) every round to win as much time as possible to escape.

Syntax

- $M, w_2 \models \langle Ags \rangle \top \wedge [Ags] p \wedge p$
- $M \models \neg(p \wedge p) \rightarrow [Ags] \bot$

rock, paper	paper, paper	scissors, paper
rock, rock	paper, rock	scissors, rock
rock, scissors	paper, scissors	scissors, scissors

Restricted group STIT frames

We define a class of frames $\mathbf{STIT}^p_G$ that consists of structures of the next form:

$$F = (W, R_{\Box}, \{R_i\}_{i \in Ags}, \{R_G\}_{G \subseteq Ags})$$

- $R_G \subseteq \bigcap_{i \in G} R_i$ such that:
  - $\forall w \in W(R_G(w) = \bigcap_{i \in Ags} R_i(w) \vee R_G(w) = \emptyset)$
  - $R_{\{i\}} = R_i$
  - $R_G(w) \neq \emptyset \rightarrow R_{G'}(w) \neq \emptyset$ for all $G' \subseteq G$
  - *Independence of groups is not forced by* R_G *relations!*

Restricted group STIT frames

Another way to present restricted group STIT frames:

$$F = (W, R_\Box, \{R_i\}_{i \in Ags}, \{U_G\}_{G \subseteq Ags})$$

- $U_G = \{w \in W \mid R_G(w) \neq \emptyset\}$
- $U_{\{i\}} = W$
- $U_G \subseteq U_{G'}$ for all $G' \subseteq G$
- In other words, every set of agents has a associated subframe, which is nothing but possible worlds where these agents act as a group

Expressivity

We are now able to express new things without extending the group STIT language

- $[G]\bot$ - agents in G do not act as a group
- $\Box[G]\bot$ - agents in G are (historically) necessary not a group
- $\langle G\rangle\top$ - agents in G are acting together as a group
- $\langle G\rangle\top \bigwedge_{i\in G} \Diamond[i][G]\bot$ - agents in G act as a group and this is voluntary: every member of G could have prevented his/her membership in G
- $\langle G \cup \{i\}\rangle\top \wedge \neg\Diamond[i][G \cup \{i\}]\bot$ - i is forced to be a member of G
- $\neg\phi \rightarrow \langle G\rangle\top$ - ϕ is a precondition for G to act as a group

Why this expressivity is cool: responsibility example

- 🧑‍⚖️: Herr Gentzen, you were a part of Third Reich crimes against humanity $[Gentzen, Nazi\ Germany]crimes$
- Gentzen: But I was born as a German citizen, I had no choice! $\neg\Diamond[Gentzen][Nazi\ Germany, Gentzen]\bot$
- 🧑‍⚖️: Oh, really? But you've joined NSDAP party, that was your deliberate choice! Why have you done that? $\Diamond[Gentzen][Gentzen, Nazis]\bot$
- Gentzen: %Dies of starvation in prison%

Axioms for restrcited group STIT

- All classical propositional tautologies
- S5 for $[i]$ modalities
- Box: $\Box\phi \rightarrow [i]\phi$, $\Box\phi \rightarrow [G]\phi$
- Ind: $\Diamond[i_1]\phi_1 \wedge \ldots \wedge \Diamond[i_n]\phi_n \rightarrow \Diamond([i_1]\phi_1 \wedge \ldots \wedge [i_n]\phi_n)$
- KB4 for $[G]$ modalities
- $\neg([G]\bot \vee [G']\bot) \leftrightarrow ([G][G']\phi \rightarrow [G \cap G']\phi)$
- $[\{i\}]\phi \leftrightarrow [i]\phi$
- $\neg[G']\bot \rightarrow \neg[G]\bot$ for all $G \subseteq G'$

Completeness proof-sketch

- Fact 1. The logic of arbitrary subframes of product S5 frames is just a fusion S5
  Kurucz, Ágnes, and Michael Zakharyaschev. "A Note on Relativised Products of Modal Logics."
- Fact 2. The logic of disjoint union of S5 frames and dead-point frames is KB4
  Pietruszczak, Andrzej, Mateusz Klonowski, and Yaroslav Petrukhin."Simplified Kripke-style semantics for some normal modal logics."

Constructing $\mathbf{STIT}^p_G$ frame

- Take full group STIT frame F_{STIT}
- Make a product S5 frame from it $\vec{F}$
- Take relativizations, i.e. subframes of $\vec{F}$, corresponding to every U_G. Lets name them $\vec{U_G}$
- For every U_G take a subframe of dead-points as $W \setminus U_G$. Let's name them $Dummy_G$
- Take a frame of individual relations without any groups whatsoever, i.e. $F_{Ind} = (W, \{R_i\}_{i \in Ags})$

Constructing $\mathbf{STIT}^p_G$ frame

- We already know how to axiomatize F_{Ind}: S5 for $[i]$ + Box + Ind axiom
- We know axioms for every $\vec{U_G}$: it is a fusion of S5
- Take disjoint unions of $\vec{U_G}$ and $Dummy_G$ - we know the logic for such frames is fusion of KB4
- Use fibring technique to "glue" F_{Ind} with every $\vec{U_G} \bigoplus Dummy_G$. No bridge axioms arise yet

Bridging axioms

- We need to show that active groups are closed under subsets: add $\neg[G']\bot \to \neg[G]\bot$ for all $G \subseteq G'$
- We need to show that $R_i = R_{\{i\}}$: add $[\{i\}]\phi \leftrightarrow [i]\phi$
- We forgot that $\vec{U}_G$ is a relativization of a product frame, where $R_{Ags\setminus\{i\}}$ are base relations instead of R_i. We need another bridge axiom to adequately translate them: $\neg([G]\bot \vee [G']\bot) \to ([G][G']\phi \leftrightarrow [G \cap G']\phi)$

Approximating decidability

Logic of Functional Dependence (LFD)

Fix a finite set of variables V and a relational signature (i.e., a set of relation symbols with associated arity) Σ. The formulas of LFD without dependence atoms are recursively generated by the following grammar:

$$\phi := P(v_1, \ldots, v_n) \mid \neg\phi \mid \phi \vee \phi \mid E_G\phi$$

where with $v_1, \ldots, v_n \in V$, $G \subseteq V$, and $P \in \Sigma$ an n-ary relation symbol.

Semantics for LFD

The semantics of LFD is based on dependence models, or ‘generalized assignment models’. These are pairs $\mathbf{M} = (M, A)$ where M is a standard FO-model over Σ, and $A \subseteq Dom(M)^V$ is a set of admissible variable assignments. I.e. dependence models is nothing else but FO-models, where not all assignmets are admissible. They are called dependence models since they allow to model dependence between variables.

LFD and FO-translation

$$tr(P(v_1, \ldots, v_n)) = P(v_1, \ldots, v_n)$$

$$tr(\neg\phi) = \neg tr(\phi)$$

$$tr(\phi \vee \phi) = tr(\phi) \vee tr(\phi)$$

$$tr(E_G\phi) = \exists v_1, \ldots . v_n(A(v_1, \ldots v_n) \wedge tr(\phi))$$

where $A(v_1, \ldots, v_n)$ is a special predicate saying that $v_1, \ldots, v_n$ are interpreted in an admissible assignment

FO-translation of LFD and restricted group STIT

- We may associate variables with agents, admissible assignments with U_G (i.e. every subset of variables will have its one collection of admissible assignments) and E_G with $\langle G \rangle$. Given FO-translation, $\neg A(v_1, \ldots, v_n)$ will be asssociated with $[G]\bot$.
- We must notice that since we have saved independence of individual agents, E_x will be nothing else but standard FO quantifier. Nevertheless, it will not be a problem, since we know that individual STIT is reducable to 2-variable fragment of FO, which is decidable.
- For the rest, we will have a fusion of LFD models (a model for every G), and LFD was proven to be decidable as well. The only thing we must check is that our two bridge principles, i.e. $[\{i\}]\phi \leftrightarrow [i]\phi$, $\neg[G]'\bot \rightarrow \neg[G]\bot$, will not create any additional problems.

A number of concerns and further directions

- Restrcited group STIT logic cannot distinguish between two situations: 1) when the group had a joint intention, but one or more members betrayed it and acted against it and 2) when the set of agents had no joint intention whatsoever. We need to explicitly introduce joint intentions in our logic
- Since $[G]$ is normal modality, whenever $[G]\bot$ is true, $[G]\psi$ is vacuously true as well. The proper way to encode that an active group sees it it that ϕ is $\neg[G]\bot \wedge [G]\phi$. Why won't we try to treat it as a single modality?
- It would be interesting to look on epistemic and temporal extensions: how groups form and dissolve in time, how should we model agent's knowledge of her status as a member of a certain groups?
- Explore connections between Schawrzentruber's fragments and restricted group STIT

Thank you for attention!

Some literature and stuff

- Herzig, Andreas, and François Schwarzentruber. "Properties of logics of individual and group agency." Advances in modal logic 7 (2008): 133-149.
- Schwarzentruber, François. "Complexity results of STIT fragments." Studia logica 100 (2012): 1001-1045.
- van Benthem, Johan, Balder ten Cate, and Raoul Koudijs. "Local Dependence and Guarding." arXiv preprint arXiv:2206.06046 (2022).
- Kurucz, Ágnes, and Michael Zakharyaschev. "A Note on Relativised Products of Modal Logics." Advances in modal logic 4 (2002): 221-242.