Plan-restricted group STIT logic

Daniil Khaitovich
ILLC, University of Amsterdam

Introduction

STIT (acronym for *sees to it that*) is a modal logic widely studied by philosophers and multi-agent systems researchers. STIT formal language describes agency via sentences of the form "agent i sees to it that ϕ holds", where "agent i sees to it that" is treated as a modality $[i]$. The main advantage of the logic is its expressive power: unlike other modal logics of agency, STIT can deal with both actions and abilities. This is useful in case with counterfactual statements of the form "agent i could have done ϕ, but did ψ". The crucial feature of the way multi-agent interactions are presented in STIT is independence of agents: everyone is free to do anything they can no matter what other agents choose to do.

A richer version of STIT logic, which contains not only modalities for individuals, but for groups as well, was introduced by John Horty [5]. Unfortunately, group STIT logic suffers from both technical and philosophical disadvantages. From the mathematical side, STIT with groups is neither decidable nor finitely axiomatizable [8]. From the philosophical side, STIT treats groups as simple as all sets of agents. This view on group agency is criticized, since it is common to argue that agents form a group only if they share some mental attitudes such as beliefs, desires or intentions, and that is why not any set of agents is a group [3,9].

Strikingly, mathematical and conceptual problems are linked: given n-many agents, there are 2^n-many group modalities, all of which are tightly connected with each other. It is precisely the reason why group STIT is undecidable and finitely unaxiomatizable in case with at least 3 agents. As it was shown in [1,8], by restricting which sets of agents form groups we may get finitely axiomatizable and decidable fragments. Nevertheless, these fragments lack any explanation why we treat some sets as groups and why other sets are not granted this status. Moreover, the fragments state what groups are in the rigid manner: the group status of the set is safe on the whole class of frames, while in fact it may be contingent even in the model. The latter case correspond to the natural language expressions as "agents a, b, c can act as a group, but they do not".

This paper presents a decidable and finitely axiomatizable version of group STIT logic that escapes these drawbacks. The main idea is to treat a set of agents as an active group iff 1) there is some *joint intention* among agents and 2) actions of group members do not contradict that joint intention. By joint intention we mean members' commitment to a certain course of their actions, which they have collectively adopted via agreement, bargaining, public declaration or any other way.

In order to express this view, we will relax independence of agents constraint on the group level: a certain degree of dependence between group members will allow us to implicitly illustrate their joint intentions. If some set of agents has no joint intentions, we will not consider them as a group.

Group STIT and its problems

For a given countable set of propositional variables $\Phi = \{p_1, p_2, \ldots\}$ and a finite set of agents $Ags = \{1, \ldots, n\}$, we recursively define the language of group STIT logic $STIT_G$:

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

where $p \in \Phi, G \subseteq Ags$.

Definition 0.1 $\mathbf{STIT}_G$ Kripke frames

A class of group STIT frames $\mathbf{STIT}_G$ consists of structures of the form: $\mathscr{F} = (W, R_\Box, \{R_i\}_{i \in Ags})$, where:

(i) $W \neq \emptyset$

(ii) $R_\Box \subseteq W^2$ is an equivalence relation

(iii) For every $i \in Ags$, $R_i \subseteq W^2$ is an equivalence relation, such that:
  (a) $R_i \subseteq R_\Box$
  (b) *Independence of agents:* if $Card(Ags) = n$, then for any sequence of $R_\Box$-connected worlds $(w_1, \ldots, w_n)$, $\bigcap\limits_{i \in Ags} R_i(w_i) \neq \emptyset$

(iv) *Group actions:* for any $G \subseteq Ags$, for arbitrary $w \in W$, $R_G(w) = \bigcap\limits_{i \in G} R_i(w)$

Valuation function and satisfiability definitions for the formulas are standard for normal modal logic. It was demonstrated [1] that from the fact that every R_i is an equivalence relation and iii it follows that for every $1, 2 \in Ags$:

$$\models [1][2]\phi \leftrightarrow [2][1]\phi$$

$$\models \langle 1 \rangle [2]\phi \rightarrow \langle 2 \rangle [1]\phi$$

So as soon as we have a fragment with individual agents and no groups, the logic is nothing else than a fusion of $\mathbf{S5} \times \mathbf{S5}$. The logic $\mathscr{L}(\mathbf{STIT}_G)$ itself is neither decidable nor finitely axiomatizable if $Card(Ags) \geq 3$, since $\mathscr{L}(\mathbf{STIT}_G)$ is reducible to $\mathbf{S5^n}$, where $n = Card(Ags)$ [1].

1 STIT with plan-restricted groups

The rest of the paper presents plan-restrcited group STIT logic. It is a version of group STIT, where groups are active in worlds compatible with the group's joint intention. Consequently, if some set of agents have no joint intention, they are not active as a group.

It results in the relaxation of the independence of agents principle on the group level: not every combination of group members' actions will constitute a group action. Group members *while acting as members* are not free to do anything they can. If some of the group members are acting against their joint intention, it breaks the group action as such: there will be no group G anymore, only members of G acting by their own. In group STIT formal language, it will be expressed as follows: if the set G is not active as a group at w, then $w \models [G]\bot$. Consequently, if some set of agents does not constitute a group at all, such "group" will not be active in the whole model: $\mathscr{M} \models [G]\bot$.

We end up with the next definition of the class of plan-restricted group STIT frames $\mathbf{STIT}^p_G$:

Definition 1.1 A Kripke frame $\mathscr{F} = (W, R_\Box, \{R_i\}_{i\in Ags}, \{R_G\}_{G\subseteq Ags})$ is a plan-restricted group STIT frame iff (i for any $i \in Ags$, G for any $G \subseteq Ags$):

(i) $W \neq \emptyset$

(ii) Every R_i and $R_\Box$ are equivalence relations on W

(iii) $R_i \subseteq R_\Box$

(iv) *Idividual independence of agents:* if $Card(Ags) = n$, then for any sequence of $R_\Box$-connected worlds $(w_1, \ldots, w_n)$, $\bigcap\limits_{i\in Ags} R_i(w_i) \neq \emptyset$

(v) *Group actions:* $R_G \subseteq \bigcap\limits_{i\in G} R_i$, such that R_G is a transitive and symmetric relation, which satisfies $\forall w \in W(R_G(w) = \emptyset \vee R_G(w) = \bigcap\limits_{i\in G} R_i)$ and for which $R_{\{i\}} = R_i$ holds. Informally, if $R_G(w) = \emptyset$, then G is not active as a group. We pressupose that groups are closed under subsets: if $R_G(w) \neq \emptyset$, then $R_{G'}(w) \neq \emptyset$ for every $G' \subseteq G$. Independence of agents fails for groups.

(PL)	All tautologies of classical propositional logic
(S5)	**S5** for every $[i]$ and $\Box$
(Nec)	$\Box\phi \to [i]\phi$
(Ind)	$\Diamond[1]\phi_1 \wedge \ldots \wedge \Diamond[n]\phi_n \to \Diamond([1]\phi_1 \wedge \ldots \wedge [n]\phi_n)$
(KB4)	**KB4** for every $[G]$
(i-G)	$[i]\phi \leftrightarrow [\{i\}]\phi$
(Inter)	$[G_1 \cap G_2]\phi \leftrightarrow [G_1][G_2]\phi$
(SubG)	$\Diamond\langle G\rangle\top \to \Diamond\langle G'\rangle\top$ for every $G' \subseteq G$

Table 1
Plan-restricted group STIT logic

The completeness and decidability proof-sketch. We decompose every $\mathbf{STIT}^p_G$ frame in two parts: 1) individual frame $(W, \{R_i\}_{i\in Ags})$, where we have no group relations; 2) and for every $G \subseteq Ags$ there is a pre-frame $(W, \{R_J\}_{J\subseteq G})$, where we have nothing but a group and its subgroups relations.

For the individual part axiom system is already known [4]. For the pre-frames, we need to divide each of them once again. Fix a group G. Take two subframes: $(U_G, \{R_J\}_{J\subseteq G})$, where $U_G = \{w \in W \mid R_G(w) \neq \emptyset\}$; and $Dummy_G = (W \setminus U_G, R_G)$. The first part is a subframe where relation R_G is non-empty, while the second is exactly a subframe where it is empty.

Group STIT frames are transformable to $\mathbf{S5}^n$ frames. Moreover, we know that the logic of (not necessarily generated) subframes of $\mathbf{S5}^n$ is finitely axioamatizable and nothing else but a fusion of **S5** [6]. U_G is exactly a subframe of group STIT frame, where $Ags = G$. It is straightforward how to transform every U_G into a subframe of some $\mathbf{S5}^{Card(G)}$ frame, so that every U_G is axiomatized by the fusion of **S5** logics (for every $[J]$ modality). Nevertheless, we need to "glue" them with their "dummy" frames, where every R_J relation is empty. Luckily, we know from [7] that it will result in **KB4** for every $[J]$. The unification of all the dismembered frame's parts may be done via fibring procedure, and we know it will not result in additional validities [2].

On top of axioms for individual frames and fusion of **KB4** we need to add three additional axioms (see Table 1): (i-G), reflecting on the fact that $R_{\{i\}} = R_i$; (SubG), showing that groups are closed under subsets; and (Inter), reflecting upon the translation from $\mathbf{S5}^n$ to $\mathbf{STIT}^p_G$ (see [1] for details).

It is not hard to see that the logic coincides with a fusion of individual STIT and **KB4** logic with intersection modality (from (Inter), S5 for every $[i]$ and (i-G) we get $[G]\phi \to [G']\phi$ for every $G \subseteq G'$). For the individual STIT, decidability is proved by standard filtration argument, while the proof for **KB4** with intersection modality may be obtained via eliminating Hintikka sets by analogy with **S5** with intersection modality [12,11]. Another way to show that satisfiability problem is decidable is by reducing it to SAT problem of LFD logic without dependency atoms, where admissible assignments are treated as worlds where the group is active [10].

References

[1] Balbiani, P., A. Herzig and N. Troquard, *Alternative axiomatics and complexity of deliberative stit theories*, Journal of Philosophical Logic **37** (2008), pp. 387–406.

[2] Gabbay, D., "Fibring logics," Clarendon Press, 1998.

[3] Gilbert, M., "On social facts," Princeton University Press, 1992.

[4] Herzig, A. and F. Schwarzentruber, *Properties of logics of individual and group agency.*, Advances in modal logic **7** (2008), pp. 133–149.

[5] Horty, J. F., "Agency and deontic logic," Oxford University Press, 2001.

[6] Kurucz, A. and M. Zakharyaschev, *A note on relativised products of modal logics* (2002).

[7] Pietruszczak, A., M. Klonowski and Y. Petrukhin, *Simplified kripke-style semantics for some normal modal logics*, Studia Logica **108** (2020), pp. 451–476.

[8] Schwarzentruber, F., *Complexity results of stit fragments*, Studia logica **100** (2012), pp. 1001–1045.

[9] Tuomela, R., "Social ontology: Collective intentionality and group agents," Oxford University Press, 2013.

[10] van Benthem, J., B. t. Cate and R. Koudijs, *Local dependence and guarding*, arXiv preprint arXiv:2206.06046 (2022).

[11] Wáng, Y. N. and T. Ågotnes, *Simpler completeness proofs for modal logics with intersection*, in: *International Workshop on Dynamic Logic*, Springer, 2020, pp. 259–276.

[12] Wáng, Y. N. and T. Ågotnes, *Public announcement logic with distributed knowledge: expressivity, completeness and complexity*, Synthese **190** (2013), pp. 135–162.