Plan-restricted group STIT logic

Daniil Khaitovich[a].

[a] Institute for Logic, Language and Computation, University of Amsterdam, Science Park 107, Amsterdam, 1098 XG, The Netherlands

ARTICLE HISTORY
Compiled July 15, 2026

ABSTRACT
The paper presents a refinement of group STIT logic that captures the distinction between genuine group actions and mere aggregations of independent actions by individuals. In the presented framework, whether a set of agents constitutes a group is conditional on the existence of a mutually agreed-upon plan and compliance of the agents' actions with the plan. The framework is demonstrated to be beneficial for formal models of (collective) responsibility attribution. From the technical side, a sound and complete axiom system is presented for a family of fragments, where all the groups are mutually disjoint. Further, a class of structures is defined that allows reasoning under the assumption that a designated group always acts in accordance with its plan. A sound and complete axiom system for the logic of such a class of structures is presented as well.

KEYWORDS
STIT logic; multi-agent systems; responsibility; group agency; modal logic

Email: d.khaitovich@uva.nl. The authors report there are no competing interests to declare

1. Introduction

The concept of group agency has been studied in philosophy (Gilbert, 1992; Tollefsen, 2015; Tuomela, 2013), economics (Peleg & Sudhölter, 2007; Sugden, 1993), computer science and artificial intelligence (de Neufville & Baum, 2021). Numerous logical formalisms have been developed to model different aspects of group agency, varying in their conceptual grounds and metalogical properties, such as expressive and deductive power and computational complexity (Alur, Henzinger, & Kupferman, 2002; Danecki, 1984; Horty, 2001; Pauly & Parikh, 2003).

To the best of our knowledge, all of these logics suffer from one conceptual disadvantage: any set of agents is treated as a group, while any aggregation of actions by individuals is treated as a group action. This view is overly simplistic. According to a widely accepted philosophical argument, a group of agents exists only if its members share a plan and act in accordance with it (Gilbert, 1992; Searle, 1990; Tuomela, 2013). Consequently, *group actions are those performed by group members to fulfill their plan.*

In this paper, we propose a refinement of group STIT logic[1] that captures the distinction between genuine group actions and mere aggregations of independent individual acts. In our framework, a set of agents is treated as a group only when they share a plan, and their actions are regarded as a group action only if they conform to the plan.

Related work. Although we are not aware of any multi-agent modal logics that capture exactly what our framework does, the present proposal is inspired by many related prior works. The only logics of group agency we know that restrict the definition of groups are in the spirit of (Schwarzentruber, 2012), where the restriction is purely syntactic and introduced for technical reasons. Specifically, (Schwarzentruber, 2012) studies fragments of group STIT logic in which group modalities are available only for certain sets of agents. These fragments enjoy better metalogical properties, such as finite axiomatizability and decidability of the satisfiability problem. Namely, given that the set of agents is defined as $Ags = \{1, \ldots, k \times m + m - 1\}$ for some natural numbers $k \geq 0$ and $m \geq 2$, Schwarzentruber (2012) provides a sound and complete axiomatization for logics, where we have modalities only for groups that are included

[1] STIT (seeing to it that) logic is a multi-agent logic of agency that originates from the articles by Belnap, Perloff and Xu (Belnap, Perloff, & Xu, 2001). For a detailed overview of the STIT-type logics and their history, see (Segerberg, Meyer, & Kracht, 2009). A richer version of STIT logic with groups is introduced in (Horty, 2001).

in the following collection (Schwarzentruber, 2012, p.1015):

$$
\begin{aligned}
&Lat = \{\\
&\{1\}, \dots, \{m-1\}\\
&\{1, \dots, m-1\},\\
&\{1, \dots, m\},\\
&\{1, \dots, m\} \cup \{m+1\}, \dots, \{1 \dots, m\} \cup \{m+m+1\},\\
&\{1, \dots, m+m+1\},\\
&\dots\\
&\{1, \dots, k \times m\},\\
&\{1, \dots, k \times m\} \cup \{k \times m+1\}, \dots, \{1, \dots, k \times m\} \cup \{k \times m+m-1\},\\
&\{1, \dots, k \times m+m-1\}\\
&\}
\end{aligned}
$$

However, the restriction is not conceptually motivated, and whenever the language contains a modality for a given set of agents, any aggregation of those agents' actions is still treated as a group action. It is also limiting, since there are no mutually disjoint non-singleton groups in Lat. We need to have group modalities for disjoint sets if we want to reason about concurrent behaviour of competing groups, e.g., sports teams, political parties. For example, if we have five agents $Ags = \{1, \dots, 5\}$ (i.e., $k = m = 2$) that are partitioned into two opposing teams $\{1, 2, 3\}, \{4, 5\}$, then fragments presented in (Schwarzentruber, 2012) are not expressive enough to reason about individuals' and teams' actions: $\{\{i\} \mid i \in Ags\} \cup \{\{1, 2, 3\}, \{4, 5\}\} \not\subseteq Lat$.

Some papers make individual epistemic agency contingent in a similar way we make group agency contingent. Namely, the epistemic logic of *impure simplicial complexes* was proposed, where agents may not be sure whether other agents are "dead" or "alive", i.e., whether one should regard others as epistemic agents (Goubault, Kniazev, Ledent, & Rajsbaum, 2024; Van Ditmarsch & Kuznets, 2024). Nevertheless, the group epistemic notions, such as common and distributed knowledge, are treated there standardly, given that all the individuals of the group are agents. A more comprehensive study of the connection between impure simplicial complexes and our framework is to be explored in future work.

Some variations of coalition logics model certain aspects of socially oriented reasoning in multi-agent settings. In (Goranko & Enqvist, 2018; Goranko & Ju, 2022), the focus is on conditional and socially cooperative abilities of coalitions. These logics allow reasoning about coalitions' abilities to achieve their goals conditional on others acting towards their goals, as well as their abilities towards cooperative behaviour that enables others to achieve their goals. While coalitions are still modeled as arbitrary sets of agents, such frameworks can implicitly reveal which agents have compatible goals and can benefit from cooperation.

Social groups can also be studied via logics of social networks, as in (Christoff, 2016; Pedersen, Smets, & Ågotnes, 2020; Smets & Velázquez-Quesada, 2017). These formalisms can capture the emergence and dynamics of social groups—for example, which agents might establish connections with which others, based on their properties. However, unlike our approach, logics of social networks operate entirely on abstract properties of agents rather than on their plans and actions.

Our contributions. The framework presented in this paper makes the ascription of group agency contingent, while, unlike in (Goubault et al., 2024; Van Ditmarsch & Kuznets, 2024), all individuals are always treated as agents. Whether a set of agents constitutes a group is conditional on the existence of a mutually agreed-upon plan and compliance of their actions with the plan. We demonstrate the framework to be beneficial for formal models of (collective) responsibility attribution. From the technical side, we provide a sound and complete axiom system for the fragment of our formal language, where we have modalities for all individuals and a collection of mutually disjoint groups, which goes beyond the results in (Schwarzentruber, 2012). We also axiomatize a special class of structures that allow us to reason under the assumption that a designated group always acts in accordance to its plan.

Structure of the paper. The paper is organized as follows. In Section 2 we recall the standard group STIT logic and explicitly demonstrate its drawbacks in modeling group agency. In Section 3 we introduce our proposal, *plan-restricted group STIT logic*: we define its formal language $\mathcal{L}_{pSTIT}$, semantics, and demonstrate the advantages we gain via the use-case of collective responsibility attribution. Section 4 contains the investigation of the metalogical properties of plan-restricted group STIT logic. We show that the full set of validities over the class of models is not finitely axiomatizable. We finitely axiomatize the set of validities in the fragment of $\mathcal{L}_{pSTIT}$ with disjoint groups. We also axiomatize the logic of substructures that allow us to reason under the assumption that a designated group always follows its plan. We conclude and outline directions for future work in Section 5.

2. Group STIT logic and its problems

In this section, we introduce standard group STIT logic à la (Horty, 2001) and emphasize the problematic treatment of groups and group actions in it. We use a slightly different semantics than the original Horty's framework, namely, Kripke STIT models, as they are introduced in (Lorini, 2013). If the reader is familiar with these formalisms, the section could be safely skipped.

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

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \vee \varphi) \mid \Box\varphi \mid [stit]_G\varphi$$

where $p \in Var, \emptyset \neq G \subseteq Ags$. Negation and disjunction have their standard meanings, as well as other derivable connectives and constants $(\wedge, \rightarrow, \leftrightarrow, \top, \bot)$. $\Box\varphi$ stands for "*it is historically necessary that* φ"; $[stit]_G\varphi$ stands for "*group* G *sees to it that* φ *holds*". By seeing to it that φ we understand the behaviour which forces φ to be true no matter what all other agents do. Group modalities with singleton groups $[stit]_{\{i\}}$ are abbreviated as $[stit]_i$. Every modality has its dual, abbreviated as $\Diamond := \neg\Box\neg$, $\langle stit\rangle_G := \neg[stit]_G\neg$, where $\Diamond\varphi$ reads as "*it is historically possible that* φ" and $\langle stit\rangle_G\varphi$ – "*G's choice of action does not rule out* φ".

STIT formalisms model agents' choices of action in indeterministic environments, *branching time structures.*

Definition 2.1 (**BT** structures)**.** A branching time structure is a tuple of the form $(T, <)$, where $T \neq \emptyset$ is a non-empty set of moments and $< \subseteq T \times T$ is an irreflexive, transitive and backward-linear[2] relation on T, ordering moments by precedence/succession.

Histories are maximal $<$-ordered chains of moments: $h \subseteq T$ is a history iff for any $m \in h$, if $m' < m$ or $m < m'$, then $m' \in h$ and for any $m, m' \in h$, either $m < m'$ or $m' < m$ or $m = m'$.

The original semantics for STIT as in (Belnap et al., 2001) defines agents' choices by extending branching time structures with choice functions. The semantics we use is simpler. The difference lies in the way we represent branching time: namely, what is treated as a moment/history pair in (Belnap et al., 2001) becomes an atomic object, a *time point*. This alternative semantics for STIT, first introduced in (Lorini, 2013), is closer to the standard Kripke semantics of modal logic. We use it solely for the sake of simplicity. Given that $\mathcal{L}_{STIT}$ is *atemporal* (i.e., we do not have any temporal modalities such as $G\varphi$ (it will always be true that φ), $H\varphi$ (it was always true that φ), $\psi U \varphi$ (ψ is true until φ), etc.), this choice does not affect the logical properties of our formalism.[3] For comparison of different semantics for temporal STIT logics, see (Ciuni & Lorini, 2018).

Definition 2.2 (Kripke STIT frames)**.** A Kripke STIT frame is a tuple $\mathcal{F} = (W, R_<, R_\Box, \{R_i\}_{i \in Ags})$, where:

(1) $W \neq \emptyset$ is a set of time points. For brevity, "*(time) points*" and "*(possible) worlds*" are to be used interchangeably;

(2) $R_\Box, \{R_i\}_{i \in Ags}$ are equivalence relations[4] on W. $wR_\Box v$ represents historical equivalence, i.e., $wR_\Box v$ iff w and v represent different ways the same moment in time could be, depending on the agents' current and future choice of actions. $wR_i v$ represents equivalence w.r.t. the current choice of action by agent i. $R_\Box, \{R_i\}_{i \in Ags}$ have the next additional properties:
   - (a) $R_i \subseteq R_\Box$ for every $i \in Ags$: agent's choice of actions only distinguishes between historically equivalent alternatives;
   - (b) *Independence of agents:* if $Card(Ags) = n$, then for any sequence of $R_\Box$-connected time points $(w_1, \ldots, w_n)$, $\bigcap\limits_{i \in Ags} R_i(w_i) \neq \emptyset$. At the given moment, whatever action each agent chooses, these actions could be performed simultaneously, i.e., agents choose their actions independently from one another;

(3) $R_<$ is a serial transitive relation on W. $wR_< v$ iff time point v succeeds w. $R_<$ has the next properties:
   - (a) *No branching backwards*: for all $w, v, z \in W$, if $vR_< w$ and $zR_< w$, then either $vR_< z$ or $zR_< v$ or $z = v$;
   - (b) *No branching forwards*: for all $w, v, z \in W$, if $wR_< v$ and $wR_< z$, then either $vR_< z$ or $zR_< v$ or $z = v$;
   - (c) *No choice between undivided histories*: $R_< \circ R_\Box \subseteq \bigcap_{i \in Ags} R_i \circ R_<$;[5]
   - (d) For all $w, v \in W$: if $wR_\Box v$, then $(w, v) \notin R_<$ and $(v, w) \notin R_<$: if time

[2] By backward linearity we mean that for any $m_1, m_2, m_3 \in T$, if $m_3 < m_1$ and $m_2 < m_1$, then either $m_2 < m_3$ or $m_3 < m_2$ or $m_3 = m_2$.

[3] There exist richer STIT logics that contain temporal modalities (Ciuni & Lorini, 2018). In this paper, we stick to the atemporal fragment for the sake of parsimony: the temporal aspects of group agency are not in the focus of the paper, while the modal characterization of branching time structures require some complicated non-trivial techniques.

[4] A binary relation $R \subseteq W \times W$ is an equivalence relation iff it is reflexive ($\forall w \in W(wRw)$), symmetric ($\forall w, v \in W(wRv \Rightarrow vRw)$) and transitive ($\forall w, v, u \in W((wRv \wedge vRu) \Rightarrow wRu)$).

[5] Given any two relations R_1, R_2, $R_1 \circ R_2 = \{(w, v) \in W \times W \mid \exists t \in W : wR_1 t, tR_2 v\}$.

points w and v are historically equivalent, then neither of them succeeds the other.

(4) *Group actions*: for every $G \subseteq Ags$, $R_G = \bigcap_{i \in G} R_i$;

From 3d together with reflexivity of $R_\Box$ it follows that $R_<$ is irreflexive. For every $w \in W$, $T_w = \{w\} \cup R_<(w) \cup R_<^{-1}(w)$ is the *history* of w,[6] i.e. the set of time points that are temporally related to w, either as predecessors or as successors. From 3a and 3b together with irreflexivity and transitivity of $R_<$ it follows that for every history T_w, $(T_w, R_<)$ is a strict linear order. From this fact it follows that every time point w has a unique history passing through it. $wR_\Box v$ means that w and v are time points that represent the same moment in time in different histories. In other words, a tuple $(W, R_\Box, R_<)$ represents a branching time structure, where moments are represented as $R_\Box$-equivalence classes. Given two moments $R_\Box(w)$ and $R_\Box(v)$, $R_\Box(w)$ precedes $R_\Box(v)$ iff there exist time points $x \in R_\Box(w), y \in R_\Box(v)$, such that $xR_<y$. For a graphic representation of the relation between branching time structures and their representation in Kripke STIT frames, see Figure 1.

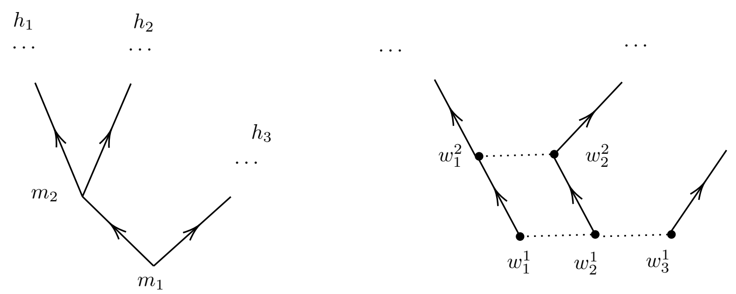

Figure 1. A branching time structure (on the left) and its representations in Kripke STIT frame (on the right). Moment/history pairs on the left correspond to time points on the right: e.g., m_1/h_1 corresponds to w_1^1, w_2^1 corresponds to m_1/h_2, etc. Arrowed lines represent $<$ and $R_<$ relations, respectively, and dotted lines represent $R_\Box$.

R_i relation models how agent i can choose among her possible actions. Her choice of action determines which *set* of histories remains possible. Condition 3c guarantees that agents cannot choose between histories that do not branch yet. By 2b, individuals' choice of actions is independent. By 4, a group action is nothing but the simultaneous actions of group members. See Figure 2 for an example of a Kripke STIT frame.

Given a frame $\mathcal{F} = (W, R_<, R_\Box, \{R_i\}_{i \in Ags})$, we denote the domain W of $\mathcal{F}$ as $Dom(\mathcal{F})$. Every Kripke STIT frame $\mathcal{F}$ could be extended to a model $\mathcal{M} = (\mathcal{F}, V)$, where $V : Var \to 2^W$ is a valuation function, which takes a propositional variable and returns a set of time points, where this proposition is interpreted as true in the given model.

Definition 2.3 (Semantic clauses for $\mathcal{L}_{STIT}$)**.** Given a Kripke STIT model $\mathcal{M} = (W, R_<, R_\Box, \{R_i\}_{i \in Ags}, V)$ and $w \in W$, the truth of a formula $\varphi \in \mathcal{L}_{STIT}$ is defined recursively as follows:

[6]For any binary relation $R \subseteq W \times W$, $R(w) = \{v \in W \mid (w, v) \in R\}$.

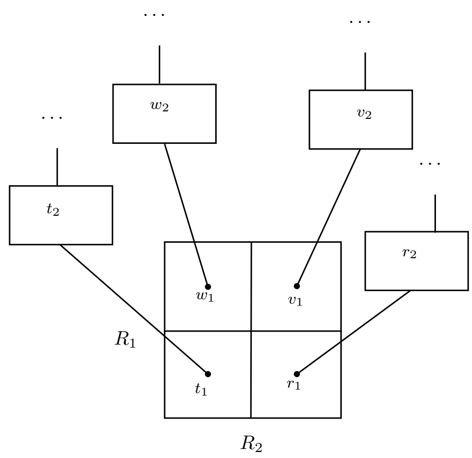

Figure 2. Kripke STIT frame $\mathcal{F} = (W, R_<, R_\Box, R_1, R_2)$. $W = \{w_n, v_n, t_n, r_n \mid n \geq 1\}$. $R_\Box(w_1) = \{w_1, v_1, t_1, r_1\}$. For every $x \in \{w_n, v_n, t_n, r_n\}$, s.t. $n > 1$: $R_\Box(x) = \{x\}$. $R_1 = \{(w_1, v_1), (v_1, w_1), (t_1, r_1), (r_1, t_1)\} \cup \{(x, x) \mid x \in W\}$, $R_2 = \{(w_1, t_1), (t_1, w_1), (v_1, r_1), (r_1, v_1)\} \cup \{(x, x) \mid x \in W\}$, $R_< = \{(x_n, x_m) \mid n < m, x \in \{w, v, r, t\}\}$. Informally, for two agents, $\{1, 2\}$, there are two available actions at the initial moment. Afterwards, for every possible outcome, each of the agents has a single action available in any future moment.

$$\mathcal{M}, w \models p \Leftrightarrow w \in V(p)$$
$$\mathcal{M}, w \models \neg\varphi \Leftrightarrow \mathcal{M}, w \not\models \varphi$$
$$\mathcal{M}, w \models \varphi \vee \psi \Leftrightarrow \mathcal{M}, w \models \varphi \text{ or } \mathcal{M}, w \models \psi$$
$$\mathcal{M}, w \models \Box\varphi \Leftrightarrow \forall v \in W(wR_\Box v \Rightarrow \mathcal{M}, v \models \varphi)$$
$$\mathcal{M}, w \models [stit]_G\varphi \Leftrightarrow \forall v \in W(wR_G v \Rightarrow \mathcal{M}, v \models \varphi)$$

By $\mathcal{M} \models \varphi$ we mean that $\mathcal{M}, w \models \varphi$ for any $w \in W$; $\mathcal{F}, w \models \varphi$ means $(\mathcal{F}, V), w \models \varphi$ for any valuation function V; $\mathcal{F} \models \varphi$ means $\mathcal{F}, w \models \varphi$ for any $w \in W$; $\models \varphi$ means $\mathcal{F} \models \varphi$ for any Kripke STIT frame. In the aforementioned cases we say that φ is *valid* in the model ($\mathcal{M} \models \varphi$)/ frame ($\mathcal{F} \models \varphi$)/class of frames ($\models \varphi$), respectively. Given a set of formulae $\Gamma \subseteq \mathcal{L}_{STIT}$, by $\models \Gamma$ we mean that $\models \varphi$ for any $\varphi \in \Gamma$. By $\Gamma \models \varphi$ we mean that for every model $\mathcal{M}$ and every world w, $\mathcal{M}, w \models \Gamma$ implies $\mathcal{M}, w \models \varphi$.

From Definition 2.3 it follows that any set of agents $G \subseteq Ags$ can be treated as a group. Moreover, any individual actions can be aggregated into the group action:

$$\models ([stit]_{i_1}\varphi_1 \wedge \ldots \wedge [stit]_{i_n}\varphi_n) \rightarrow [stit]_{\{i_1,\ldots,i_n\}}(\varphi_1 \wedge \ldots \wedge \varphi_n)$$

The logic $\{\varphi \in \mathcal{L}_{STIT} \mid \models \varphi\}$ has a finite sound and complete axiom system only for the case when $Card(Ags) \leq 2$: it is **L**, as it is defined in Table 1. Otherwise the logic of Kripke STIT frames is neither axiomatizable nor decidable (Herzig & Schwarzentruber, 2008).

(S5)	**S5** axioms and rules for $[stit]_a, [stit]_b, \Box, [stit]_{Ags}$
(Incl)	$\Box\varphi \to [stit]_G\varphi$, where $G \subseteq Ags$
(Ind)	$(\Diamond[stit]_a\varphi \wedge \Diamond[stit]_b\psi) \to \Diamond([stit]_a\varphi \wedge [stit]_b\psi)$
(i-[Agt])	$[stit]_i\varphi \to [stit]_{Ags}\varphi$

Table 1. Logic **L**

3. STIT logic with plan-restricted groups

This section presents our proposal: *plan-restricted group STIT logic*. It is a version of group STIT logic, where we treat agents' actions as a *group* action only if these agents share some mutual agreed-upon plan and act upon it. If some agents have no such plan, they do not form a group, hence, their actions will be treated as nothing but simultaneous acts by individuals.

3.1. Conceptual motivation

As we have just shown, the standard group STIT logic has a weak definition of groups and group actions: every set of agents $G \subseteq Ags$ is a group, while a group action is any combination of group members' actions. Moreover, the independence of agents constraint presupposes that agents are completely free to act independently from one another. We find this view counterintuitive. Agents constitute a group only if they restrict their choices of actions by the plan, hence are not free to act independently from one another *as members of the group*. So either all agents should be free to act independently and there should be no groups, or group membership should entail a possible restriction of the independence of agents.

We adopt the latter view. At the individual level, each agent is indeed independent in their choice of action. However, not any simultaneous choice of actions by any set of agents counts as a group action. As Raimo Tuomela notes:

> In the strongest sense of acting together we require acting on a joint, agreed-upon plan. [...] This kind of joint action is a collective social action in its most central sense, acting as a team. (Tuomela, 2005, p.83)

In other words, for a combination of individual actions to be counted as a group action, the individuals should have a "*joint, agreed-upon plan*", and their actions in question should adhere to that plan.

For example, members of a running club going for a run *as a group* differ from runners who simultaneously run at the same park by a mere coincidence, because the members of the running club have planned their run and acted accordingly. Importantly, the mere existence of a joint plan among some agents is not enough to count any combination of their individual actions as a group action: such actions should follow the plan. For example, think of two friends who decided to organize a running club and agreed to go on weekly runs every Sunday at 10:00 at a local park. Every time they show up at the given time and place and go for a run, we can count it as a group action of the running club. But if they do not show up for a run and stay home without notifying one another, it seems strange to regard it as a group action of their club: at this point, those are just two individuals who defy their mutual agreement.

From now on, we will use the following terminology. By a *shared action* of the set of agents $G \subseteq Ags$ we mean any simultaneous actions by members of G, independently of the existence of any agreed-upon plan among them. By a *group action* of G we mean their "collective social action in its most central sense" as put by Tuomela, i.e., shared action of G that follows their plan. All group actions are shared actions, but not all shared actions qualify as group actions. As we have shown in Section 2, standard group STIT logic does not distinguish between shared and group actions.

According to (Bratman, 1987), a plan's function is to limit the scope of future choice of action, so that the planner will not have to constantly deliberate what is the best thing to do at every given moment:

> [Plans] constrain solutions to these problems [of deliberation], providing a *filter of admissibility* for options. [...] They narrow the scope of the deliberation to a limited set of options. (Bratman, 1987, p.33)

Following Bratmanian ideas, a group plan should limit their choice of present and future actions. Given a set of agents $G \subseteq Ags$, we can abstractly represent G's plan as a subset of G's possible current and future shared actions that adhere to G's plan, i.e., that are *group* actions. In case this set is empty, G has no plan and does not constitute a group.[7]

To introduce these ideas in STIT framework, we will associate every subset of agents $G \subseteq Ags$ with some possibly empty set of time points $U_G \subseteq W$, where members of G act in accordance with their plan. If members of G are not a group, i.e. they do not coordinate their actions in any way, then the corresponding set of worlds will be empty. In this case, any behaviour of G's members is just simultaneous acting of individuals rather than acting of a group.

We will impose a number of constraints on these sets of points. First, U_G should be $\bigcap_{i \in G} R_i$-closed: if $w \in U_G$, wR_iv for all $i \in G$, then $v \in U_G$. Informally, the group plan is a subset of members' shared actions: whether a group's plan will be fulfilled is only up to the group members' choice of actions. Second, we will associate the set of all points with singleton groups to preserve the independence of individual agents: for every $i \in Ags : U_{\{i\}} = W$.

3.2. Formal language and semantics

Having specified our understanding of group actions, we proceed with definitions of the formal language and its semantics. Given a countable set of propositional variables Var and a finite set of agents Ags, we define an extension of $\mathcal{L}_{STIT}$, a formal language $\mathcal{L}_{pSTIT}$ of plan-restricted STIT logic, as follows:

$$\mathcal{L}_{pSTIT} \ni \varphi ::= p \mid \neg\varphi \mid (\varphi \vee \varphi) \mid \Box\varphi \mid [stit]_G\varphi \mid [G]\varphi$$

where $p \in Var$, $\emptyset \neq G \subseteq Ags$. Propositional variables, Boolean connectives, $\Box$ and $[stit]_G$-modalities have their standard meaning, as in $\mathcal{L}_{STIT}$. $[G]\varphi$ reads as "*if G performs a group action, then G sees to it that φ*". A dual $\langle G\rangle\varphi := \neg[G]\neg\varphi$ reads as "*G performs a group action that does not rule out φ*". Consequently, $\langle G\rangle\top$ is read as "*G performs a group action*", and $[G]\bot$ – as "*G does not perform a group action*", since (as we are to see after we introduce the semantics) it is tautological that no one sees to it that $\bot$/rules out $\top$ by their actions. We introduce the following abbreviations: for any agent $i \in Ags$, $[i]\varphi := [\{i\}]\varphi$; $[pstit]_G\varphi := \langle G\rangle\top \wedge [G]\varphi$. $[pstit]_G\varphi$ reads as "*G performs a group action and sees to it that φ*", or for brevity "*G jointly sees to it that φ*". By $Gr = \{G \subseteq Ags \mid Card(G) \geq 2\}$ we will denote the set of non-singleton groups.

[7] A similar way to formalize group plans in game-theoretic setting was used in (Duijf, Broersen, Kuncova, & Abarca, 2021).

It is important to note that we interpret $[G]$ and $[stit]_G$-modalities in a *causal* manner. For example, $[stit]_i\varphi$ simply means that i chooses an action that forces φ. It means neither that i does that knowingly nor that i does that intentionally. Similarly, $[pstit]_G\varphi$ means that members of G choose actions that adhere to their plan and that force φ, which means neither that it is their collective intention nor that it is commonly known among them.

We propose the next definition of the class of plan-restricted group STIT frames **pSTIT**, which models our intuitions about group behaviour explained above.

Definition 3.1 (**pSTIT**-frames)**.** A tuple $(W, R_<, R_\Box, \{R_i\}_{i\in Ags}, \{U_G\}_{G\subseteq Ags})$ is a **pSTIT**-frame iff:

(1) $(W, R_<, R_\Box, \{R_i\}_{i\in Ags})$ is a Kripke STIT frame in accordance with Definition 2.2. Standardly, $R_G = \bigcap_{i\in G} R_i$;
(2) For every $G \subseteq Ags$, $U_G \subseteq W$ is a possibly empty set of time points, where G acts in accordance with their plan. $\{U_G\}_{G\subseteq Ags}$ have the next properties:
  (a) Every U_G is a R_G-closed subset of worlds.[8]
  (b) For every agent $i \in Ags$: $U_{\{i\}} = W$.

Condition 2a expresses that G's plan is a subset of G's shared actions. Condition 2b expresses that individuals are not restricted in their choice of action: there is no difference between shared and group actions of a "group" $\{i\}$.

For technical reasons, it is useful to define derivable relations $\{\sim_G\}_{G\subseteq Ags}$: $w \sim_G v$ iff w and v are equivalent with respect to the choice of group actions by G.

Definition 3.2 ($\sim_G$-relations)**.** Given any **pSTIT**-frame $(W, R_<, R_\Box, \{R_i\}_{i\in Ags}, \{U_G\}_{G\subseteq Ags})$, we define a set of relations $\{\sim_G\}_{G\subseteq Ags}$ as follows: for any $G \subseteq Ags$, $\sim_G = R_G \cap (U_G \times U_G)$. Since every R_G is an equivalence relation, $\sim_G$ is a transitive and symmetric relation.[9] Since every U_G is R_G-closed, if $w \sim_G v$, then wR_Gv for any $w, v \in W$; if $\sim_G (w) = \emptyset$, then for every v, such that wR_Gv, $\sim_G (v) = \emptyset$. Since $U_i = W$ for any $i \in Ags$, $\sim_i = R_i$ for any $i \in Ags$.

$\{\sim_G\}_{G\subseteq Ags}$ and $\{U_G\}_{G\subseteq Ags}$ are interdefinable: $\sim_G$ is defined via U_G and R_G, while $U_G = \{w \in W \mid \sim_G (w) \neq \emptyset\}$.

Definition 3.3 (Semantic clauses for $\mathcal{L}_{pSTIT}$)**.** Given any formula $\varphi \in \mathcal{L}_{pSTIT}$, a model $M = (W, R_<, R_\Box, \{R_i\}_{i\in Ags}, \{U_G\}_{G\subseteq Ags}, V)$ and a world $w \in W$, we define $\models$ relation as follows:

[8] The subset of worlds $U_G \subseteq W$ is R_G-closed iff the following holds: $\forall w \in U_G \forall v \in W (wR_Gv \Rightarrow v \in U_G)$.

[9] As it was shown in, e.g., (Goubault et al., 2024, p.90), a restriction of an equivalence relation $R \subseteq W \times W$ to any subdomain $X \subseteq W$ yields a transitive symmetric relation, which is sometimes called "partial equivalence relation" or "local equivalence relation".

$$M, w \models p \Leftrightarrow w \in V(p)$$
$$M, w \models \neg\varphi \Leftrightarrow M, w \not\models \varphi$$
$$M, w \models \varphi \vee \psi \Leftrightarrow M, w \models \varphi \text{ or } M, w \models \psi$$
$$M, w \models \Box\varphi \Leftrightarrow \forall v \in W(wR_{\Box}v \Rightarrow M, v \models \varphi)$$
$$M, w \models [stit]_G\varphi \Leftrightarrow \forall v \in W(wR_G v \Rightarrow M, v \models \varphi)$$
$$M, w \models [G]\varphi \Leftrightarrow (w \in U_G \Rightarrow M, w \models [stit]_G\varphi)$$

Equivalently, the last clause can be defined via $\sim_G$ relation:

$$M, w \models [G]\varphi \Leftrightarrow \forall v \in W(w \sim_G v \Rightarrow M, v \models \varphi)$$

$M \models \varphi$, $F, w \models \varphi$, $F \models \varphi$ and **pSTIT** $\models \varphi$ are defined standardly, as in Definition 2.3. By $\Gamma \models_{\mathbf{pSTIT}} \varphi$ we mean that for every **pSTIT**-model M and every world w, $M, w \models \Gamma$ implies $M, w \models \varphi$.

3.3. Examples: collective responsibility, complicity

We can use **pSTIT**-models for a better analysis of causal responsibility[10] attribution. Unlike standard group STIT logic, our formalism can distinguish between proper group actions and mere shared actions. Therefore, we can meaningfully attribute *collective* or *corporate* responsibility:

> *[...] It is important to distinguish collective and corporate responsibility. Both concepts apply to responsibility as apportioned to groups. But collective responsibility is the idea that individual persons within a group are responsible for an outcome produced collectively. That is, responsibility is apportioned to individuals and to them alone. With corporate responsibility the group is treated as a being distinct from its members and responsibility for wrongdoing.* (Sverdlik, 1987, p.62)

Example 3.4 (Market collusion)**.** *GameStop*, a slowly dying gaming retail company, rose sharply in price in early 2021: because of the unprecedented demand, its stock which cost \$20 skyrocketed up to \$400 at peak. There was no financial stimulus for people to suddenly begin investing in *GameStop*: such behaviour was mostly driven by ironic discussions and jokes online. As a result, hedge funds have lost billions of dollars from following the dismissive market signals.

The question arose: was it a planned market manipulation of some organized group or just an instance of simultaneous economically irrational behaviour of individuals? If the former is the case, the criminal offense of illegal collusion took place, while no laws have been broken otherwise.

To simplify the example, assume there were just two GameStop buyers, a and b, and that the hedge funds have lost their money on the 100th day after GameStop began to grow. We assume that a and b were coordinating their actions and had a plan to keep on buying GameStop until hedge funds begin to lose money.

[10] An agent or a group of agents is held causally responsible for an outcome iff their actions have caused that outcome, regardless of intent or awareness.

Let $Ags = \{a, b\}$ be the set of GameStop buyers[11] and $Var = \{p_a, p_b, l, d_1, d_2 \dots ,\}$ be the set of propositional variables, where p_a reads as *a buys GameStop stocks*, p_b – *b buys GameStop stocks*, l – *hedge funds have lost money*, d_i – *today is the ith day of observing the market anomaly.*

Let $\Lambda = (\{1_a, 0_a\} \times \{1_b, 0_b\})^\omega$ be a set of infinite strings that represent different possible timelines of a and b's market behaviour: for any $i \in \{a, b\}$, 1_i means that i buys GameStop stocks, while 0_i means that i does not. For example, $(1_a, 1_b)^\omega$ is a timeline where both a and b are buying GameStop stocks every day. For any $\alpha \in \Lambda$, $\alpha[0, \dots, n]$ represents the corresponding initial segment of the first $n + 1$ days of the timeline. For example, if $\alpha = (1_a, 1_b)^\omega$ (i.e. both a and b buy stocks every day), $\alpha[0, \dots, 2] = \langle(1_a, 1_b), (1_a, 1_b), (1_a, 1_b)\rangle$. By $\alpha(n)$ we mean the $n + 1$th day of the timeline. For example, if $\alpha = \langle(1_a, 0_b), (0_a, 1_b)\rangle^\omega$, then $\alpha(0) = \alpha(2) = (1_a, 0_b)$, $\alpha(1) = \alpha(3) = (0_a, 1_b)$. With a slight abuse of notation, $\alpha_a(n) \in \{1, 0\}$ represents a's action at the $n+1$th day in the timeline α. For example, if $\alpha_a(3) = 1$, then a is buying on the fourth day in α, i.e. $\alpha(3) \in \{(1_a, 0_b), (1_a, 1_b)\}$. $\alpha_b(n)$ is defined accordingly.

Let $\Pi = \{\pi \in \Lambda \mid \pi[0, \dots, 99] = \{(1_a, 1_b)\}^{100}\}$ be a set of timelines, where both a and b buy GameStop for the first 100 days. We model Example 3.4 with the following **pSTIT**-model:

$$M = (W, R_<, R_\Box, R_a, R_b, U_{Ags}, V)$$

(1) $W = \{w^i_\alpha \mid i \in \mathbb{N}, \alpha \in \Lambda\}$, where w^i_α represents the $i+1$th day of the timeline $\alpha \in \Lambda$. For example, $\alpha := (1_a, 0_b)^\omega$ represents the timeline where a buys GameStop every day, while b never does that. w^3_α represents the fourth day of such timeline;

(2) $R_< = \{(w^i_\alpha, w^j_\alpha) \mid i < j\}$, only points of the same timeline constitute a history and hence are temporally related;

(3) $R_\Box = \{(w^i_\alpha, w^i_\beta) \mid \alpha[0, \dots, i-1] = \beta[0, \dots, i-1]\}$, i.e. points are historically equivalent iff they represent the same day of timelines with the same past;

(4) $R_a = \{(w^i_\alpha, w^i_\beta) \mid w^i_\alpha R_\Box w^i_\beta, \alpha_a(i) = \beta_a(i)\}$, i.e. a can choose to buy or not to buy GameStop stocks every single day. $R_b = \{(w^i_\alpha, w^i_\beta) \mid w^i_\alpha R_\Box w^i_\beta, \alpha_b(i) = \beta_b(i)\}$, i.e. analogously for b;

(5) $U_{Ags} = \{w^i_\alpha \mid i \leq 99, \forall \pi \in \Pi : w^i_\alpha R_a w^i_\pi \text{ and } w^i_\alpha R_b w^i_\pi\}$, i.e. according to $\{a, b\}$'s plan, both a and b should buy GameStop stocks until the hedge funds face losses, i.e. for 100 days straight. After that, they stop coordinating their actions and do not behave as a group.[12]

(6) $V(p_a) = \{w^i_\alpha \mid \alpha_a(i) = 1\}$, $V(p_b) = \{w^i_\alpha \mid \alpha_b(i) = 1\}$, $V(d_{i+1}) = \{w^i_\alpha \mid \alpha \in \Lambda\}$, $V(l) = \{w^{99}_\pi \mid \pi \in \Pi\}$, i.e. p_a, p_b are true only in time points where a and b buy GameStop stocks, respectively. d_i is true at any time point that represents the ith day and l is true on the 100th consecutive day of a and b buying GameStop stocks.

For a graphic illustration of the model, see Figure 3.

We can express that a and b's plan presupposes that both of them should buy for

[11] For the sake of simplicity, we do not model hedge funds as agents, assuming that their actions are functionally dependent on the market signals that are created by actions of a and b.

[12] Note that all timelines in Π divide only after 100th day. Therefore, by 3c of Definition 2.2, for any $\pi_1, \pi_2 \in \Pi$ and any $i \in \{0, \dots, 99\}$, $(w^i_{\pi_1}, w^i_{\pi_2}) \in R_a \cap R_b$. Hence, U_{Ags} is $R_a \cap R_b$-closed, as required by Definition 3.1.

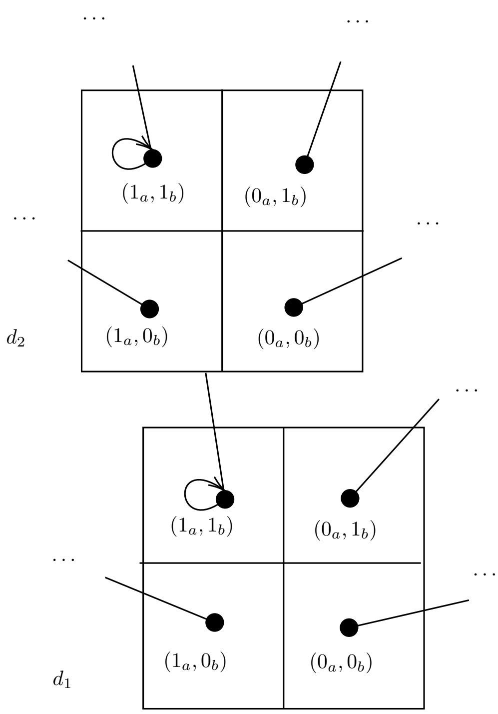

Figure 3. **pSTIT**-frame for Example 3.4. Points on the figure represent $R_a \cap R_b$-equivalence classes, i.e. worlds with the same past and the same current choice of both agents. The reflexive arrows decorate $R_a \cap R_b$-equivalence classes that are included in U_{Ags}, i.e. choices of a and b that go in accordance with the plan.

the first 100 days:

$$M \models \langle a, b\rangle \top \to (\bigvee_{1\leq i\leq 100} d_i \to [a]p_a \wedge [b]p_b)$$

We can also express that on the 100th day of following their plan, a and b jointly see to it that hedge funds lose their money: for any $\pi \in \Pi$, $M, w_\pi^{99} \models [pstit]_{\{a,b\}}l$.

Moving to collective responsibility attribution, we first recall how it is done in standard group STIT logic. In (Sergot, 2021), it is argued that the group G is responsible for the outcome φ iff 1) G sees to it that φ, 2) φ is not necessary true and 3) there is no proper subgroup $J \subset G$, such that J sees to it that φ. In other words, actions of all members of G are (1) sufficient and (2-3) necessary to force φ. Formally, these conditions are expressed in $\mathcal{L}_{STIT}$ as follows:

$$\Delta_G^{\min}\varphi := [stit]_G\varphi \wedge \neg\Box\varphi \wedge \bigwedge_{J\subset G} \neg[stit]_J\varphi$$

which, in the case of two agents, is equivalent to

$$\Delta_{\{a,b\}}^{\min}\varphi := [stit]_{\{a,b\}}\varphi \wedge \neg\Box\varphi \wedge \bigwedge_{i\in\{a,b\}} \neg[stit]_i\varphi$$

Since the conditions are formulated in terms of standard group STIT modalities, where any set of agents is a group, our example is challenging for this definition. We argue that the responsibility of G should be formalized as follows:

$$\Sigma_G^{\min}\varphi := \Delta_G^{\min}\varphi \wedge \bigvee_{G\subseteq J\subseteq Ags} [pstit]_J\varphi$$

i.e., actions of G are necessary and sufficient to guarantee φ, and their actions are coordinated by the plan of some group.

In our example, the group $\{a, b\}$ is causally responsible for φ iff

$$\Sigma_{\{a,b\}}^{\min}\varphi := [stit]_{\{a,b\}}\varphi \wedge \neg\Box\varphi \wedge \bigwedge_{i\in\{a,b\}} \neg[stit]_i\varphi \wedge [pstit]_{\{a,b\}}\varphi$$

To see why this formalization suits better, let M' be a **pSTIT**-model that differs with M only in the definition of U_{Ags}: namely, $U_{Ags} = \emptyset$ in M'. In other words, M' models the same situation from Example 3.4, where a and b do not act as a group, but may simultaneously show the same pattern of economically irrational behaviour. It is straightforward to see that the standard group STIT language $\mathcal{L}_{STIT}$ cannot distinguish between M and M': for any $\varphi \in \mathcal{L}_{STIT}$ and any $w \in W$, $M, w \models \varphi$ iff $M', w \models \varphi$. Moreover, if we reason from the standard group STIT point of view, we face a *responsibility void*.[13] Namely, a and b see to it that the hedge funds lose money, but neither a nor b is causally responsible for that outcome, since neither of them has caused that financial loss by themselves:

$$M', w_\pi^{99} \models \Delta_{\{a,b\}}^{\min} l \wedge \neg[stit]_a l \wedge \neg[stit]_b l$$

[13] A responsibility void is a gap in blameworthiness for an outcome of agents' shared action, when none of the agents can be individually blamed for that outcome (Duijf, 2022).

Informally, even though the shared action of a and b buying GameStop stocks caused the money loss on the 100th day, actions of neither a nor b caused it alone: if a would not buy GameStop stocks, given the very same actions of b, hedge funds would not lose their money, and analogously for b.

The responsibility void is genuinely problematic only if we deal with M', i.e. if a and b are just two independent agents who happen to simultaneously behave in a similar economically irrational way. In such a case, we cannot attribute collective responsibility to a and b, since there is no such collective as a and b! In the meantime, responsibility attribution appears to be unproblematic in the case of M. If both a and b agree upon a plan to manipulate the market and jointly see to it that hedge funds lose their money, it seems natural to ascribe collective responsibility to both a and b. Since standard group STIT logic cannot distinguish between M and M', this important difference cannot be accounted for.

Unlike $\mathcal{L}_{STIT}$, $\mathcal{L}_{pSTIT}$ can distinguish between M and M': for any $\pi \in \Pi$, $M, w_\pi^{99} \models \langle a, b \rangle \top$ and $M', w_\pi^{99} \not\models \langle a, b \rangle \top$. We can attribute collective responsibility to a and b for hedge funds' financial loss in case they acted as a group:

$$M, w_\pi^{99} \models \Sigma_{\{a,b\}}^{\min} l$$

At the same time, if a and b were not acting as a group, we can reason only about their shared, but not group actions, avoiding the ungrounded attribution of collective forms of responsibility. Formally:

$$M', w_\pi^{99} \models [stit]_{\{a,b\}} l \wedge \neg [pstit]_{\{a,b\}} l$$

So far, we have been reasoning about responsibility in a *causal* manner, i.e. a group G is held collectively responsible for the outcome φ iff G has *caused* φ. In $\mathcal{L}_{pSTIT}$, we can go further and express the weaker form of responsibility, *complicity.* As several philosophers argue, an agent is complicit in bringing about that φ iff they have merely participated in a group action that has resulted in φ, even if their actions were not necessary to achieve φ (Bazargan, 2013; Kutz, 2000). For instance, consider three agents: a killer, a lookout, and a passerby (we denote them as k, l, and p, respectively). The killer and the lookout form a group, while the passerby does not cooperate with any of them. According to the plan of the killer and the lookout, the former is to commit murder, and the latter is to guard the area. In case the lookout sees a passerby, he is to make sure that the passerby will not witness the murder. As an outcome, the killer commits murder, while the lookout guards the location, and the passerby takes a different route, by chance not walking past the location where the crime takes place.

Intuitively, the lookout is clearly complicit with the murder, while the passerby is not. If we think of the causal relations of actions of the lookout and the passerby to the murder in terms of $\mathcal{L}_{STIT}$, they are completely symmetrical. The killer has seen to it that someone is murdered. Neither the lookout nor the passerby has causally contributed to the murder at all:

$$\Delta_{\{k\}}^{\min} murder \wedge \neg(\Delta_{\{l\}}^{\min} murder \vee \Delta_{\{p\}}^{\min} murder)$$

Their actions are also symmetrical with respect to the cover-up of the murder. The killer and the lookout have seen to it that the murder was not witnessed, but the killer and the passerby have seen to it as well. If, ceteris paribus, the passerby had done otherwise and decided to pass by the crime scene, the murder would still have happened and remained unwitnessed, given that the murderer had killed and the lookout had guarded the area, hence would have prevented the passerby from witnessing the crime. Analogously, if the lookout had done otherwise and had not guarded the area, the murder would still have happened and remained unwitnessed, since the murderer had killed and the passerby had decided not to walk past the area. Moreover, neither the killer nor the lookout nor the passerby alone has seen to it that the covered murder has happened. Therefore, in terms of $\mathcal{L}_{STIT}$, both duos of the killer with the lookout and the killer with the passerby are responsible for the covered murder, which seems wrong:

$$\Delta^{\min}_{\{k,l\}}(murder \wedge covered) \wedge \Delta^{\min}_{\{k,p\}}(murder \wedge covered)$$

We cannot adequately formalize the notion of complicity in $\mathcal{L}_{STIT}$, since we cannot distinguish group actions. The following formula is valid for any groups $G \subseteq J \subseteq Ags$:

$$\models [stit]_G\varphi \rightarrow [stit]_J\varphi$$

i.e. if a group sees to it that φ, then any supergroup of it sees to it that φ. Therefore, if a single agent i has seen to it that φ, $[stit]_i\varphi$, then i and any other agent j has seen to it that φ, $[stit]_{\{i,j\}}\varphi$. Therefore, j has "participated" in a shared action that has resulted in φ. Was it a group action or not – $\mathcal{L}_{STIT}$ lacks expressive power to answer that question.

We can make the distinction between group actions and mere shared actions in $\mathcal{L}_{pSTIT}$. Reflecting the minimal account of (Kutz, 2000), where complicity is reduced to the participation in the group action that brings about that φ, an agent $i \in Ags$ is complicit in bringing about that φ iff φ is not necessary and there exists a group $G \subseteq Ags$, such that $i \in G$ and G jointly sees to it that φ. Formally:

$$Comp_i\varphi := \neg\Box\varphi \wedge \bigvee_{i \in G \subseteq Ags} [pstit]_G\varphi$$

Applying this definition to the example, the lookout, unlike the passerby, is complicit in the murder and its coverage:

$$\begin{aligned} &Comp_l murder \wedge Comp_l(murder \wedge covered) \wedge \\ &\qquad\qquad \neg(Comp_p murder \vee Comp_p(murder \wedge covered)) \end{aligned}$$

where *murder* and *covered* are propositions that read as *"the murder has happened"* and *"the murder was not witnessed"*, respectively.

The covered murder was not necessary (given that the killer could have refused to kill). The lookout is complicit in both the murder and its coverage, since he is a member of a group that has jointly seen to it that the covered murder happened. On the contrary, the passerby is not complicit, since he is not a member of any group that has jointly seen to it that the murder or its coverage has happened.

To summarize, we demonstrated how plan-restricted group STIT logic may refine the formal models of responsibility by distinguishing between genuine group actions and uncoordinated shared actions of independent agents. This distinction allows us to ascribe collective responsibility and complicity to agents in a meaningful way.

3.4. Potential objections: non-interchangeable and mutually inconsistent plans

Having introduced some examples of **pSTIT**-models, we now return to discuss our definition of group plans to address two objections that could be taken to show that our definition is too weak.

Given a **pSTIT**-frame $F = (W, R_<, R_\Box, \{R_i\}_{i\in Ags}, \{U_G\}_{G\subseteq Ags})$, U_G represents G's plan. For any non-singleton group $G \in Gr$, the only requirement Definition 3.1 poses on the plan U_G is the closure under $\bigcap_{i\in G} R_i$. Therefore, our definition allows *non-interchangeable* plans. Informally, a group's plan is interchangeable iff members of the group can fulfill the plan by acting independently of one another, without any need to coordinate their actions. Consequently, non-interchangeable plans require coordination between group members. The notion of interchangeability was initially introduced in (Nash, 1951) and adopted in (Tamminga & Duijf, 2017) in the context of group plans. The interchangeability of a plan closely resembles the independence of agents constraint, since both require independence of action choices. Given a group $G = \{1, \ldots, g\}$, G's plan U_G is interchangeable iff for any worlds $w_1, \ldots, w_g \in W$, such that $w_1 R_\Box \ldots R_\Box w_g$ and $w_1, \ldots, w_g \in U_G$, the following holds:

$$R_1(w_1) \cap \ldots \cap R_g(w_g) \subseteq U_G$$

i.e., any aggregation of members' individual actions that are consistent with the plan is a group action. If a plan is interchangeable, group members can independently choose any actions of their own that are consistent with the plan and be sure that the plan will be fulfilled, without any need to communicate and coordinate their choice of actions. In (Tamminga & Duijf, 2017), it is argued that only interchangeable plans are plausible, given that the agents cannot communicate to act conditionally on the choice of others. We allow non-interchangeable plans, since we reason about individual and group actions causally: we make no assumptions about agents' knowledge and abilities to communicate. Therefore, groups may have non-interchangeable plans if they have sufficient communication sources to coordinate their actions accordingly.

Definition 3.1 allows groups with common members to have mutually inconsistent plans. For instance, $i \in Ags$ might be a member of two groups, G and J, such that, according to G's plan, at the given moment i must see to it that φ, while J's plan requires i to see to it that $\neg\varphi$ at the same moment. The models with mutually inconsistent plans allow us to model *inter-role conflicts*: situations where agents face contradictory obligations as members of different groups. For instance, i is a member of a research group ($i \in G$) and has a family ($i \in J$). Assume that the family plans a gathering, while the research group plans a research seminar on the same evening. As a result, G's plan presupposes that i will attend the research seminar, while J's plan presupposes that i will attend the family gathering, which are inconsistent. Therefore, i should decide in accordance with which plan to act (if any).

A reasonable objection against the occurrence of inter-role conflicts may be raised, based on our definition of plans as *mutually agreed-upon*. If the plans of two groups oblige common members to do mutually inconsistent actions, why would these common members agree to adopt such plans? The reason is again epistemic: the common members may believe that the plans are consistent. In our example, i's family may agree to gather on the occasion of i's birthday; i's research group may agree to organize the research seminar on the last Wednesday before the examination week, which happens to be i's birthday. On the contrary, i may falsely believe that those two are different days. Since we do not go beyond the causal modeling, such epistemic and doxastic nuances lie outside our interest, so we do not model them explicitly.

In case we do want to model idealized agents that always have enough information about the environment and their own abilities to exclude inter-role conflicts, we can use a subclass of **pSTIT**-frames that satisfy the following restriction. For any collection of sets of agents $G_1, \dots, G_n \subseteq Ags$, such that $G_1 \cap \dots \cap G_n \neq \emptyset$, and for any world w, if $R_\Box(w) \cap U_{G_1} \neq \emptyset, \dots, R_\Box(w) \cap U_{G_n} \neq \emptyset$, then $R_\Box(w) \cap U_{G_1} \cap \dots \cap U_{G_n} \neq \emptyset$. In other words, at any moment, groups with common members have mutually consistent plans, given that their plans are non-empty.

It is worth noting that non-empty plans of groups can be inconsistent only if groups in question have some common members, i.e., every time we face the conflict between plans of two groups, these groups share some members, and those members face some form of inter-role conflict.

Proposition 3.5 (Groups without common members never have simultaneously mutually inconsistent plans)**.** *Let $G_1, \dots, G_n \subseteq Ags$ be groups without common members, i.e., $G_1 \cap \dots \cap G_n = \emptyset$. Then, whenever their plans are possible to fulfill, it is possible to fulfill their plans simultaneously, i.e.:*

$$\mathbf{pSTIT} \models (\Diamond\langle G_1\rangle\top \wedge \dots \wedge \Diamond\langle G_n\rangle\top) \to \Diamond(\langle G_1\rangle\top \wedge \dots \wedge \langle G_n\rangle\top)$$

Proof. W.l.o.g., we prove it for the case of two groups, $G = \{i_1, \dots, i_k\}$ and $J = \{j_1, \dots, j_l\}$. Assume otherwise, i.e. there exists a **pSTIT**-model M and a world w, such that:

$$M, w \models \Diamond\langle G\rangle\top \wedge \Diamond\langle J\rangle\top \wedge \Box\neg(\langle G\rangle\top \wedge \langle J\rangle\top)$$

Since $M, w \models \Diamond\langle G\rangle\top \wedge \Diamond\langle J\rangle\top$, there exists worlds v and u, such that $wR_\Box v$, $wR_\Box u$ and $v \in U_G$ and $u \in U_J$. Since $G \cap J = \emptyset$, by independence of agents (i.e., by 2b of 2.2):

$$R_{i_1}(v) \cap \dots \cap R_{i_k}(v) \cap R_{j_1}(u) \cap \dots \cap R_{j_l}(u) \neq \emptyset$$

Consequently, there exists a world s, such that $s \in R_{i_1}(v) \cap \dots \cap R_{i_k}(v) \cap R_{j_1}(u) \cap \dots \cap R_{j_l}(u)$. In other words, $s \in \bigcap_{i \in G} R_i(v) \cap \bigcap_{j \in J} R_j(u)$. Since $v \in U_G$ and U_G is $\bigcap_{i \in G} R_i$-closed, $s \in U_G$. Since $u \in U_J$ and U_J is $\bigcap_{j \in J} R_j$-closed, $s \in U_J$. Therefore, $s \in U_G \cap U_J$. Moreover, since $R_x \subseteq R_\Box$ for any $x \in G \cup J$, $vR_\Box s$ and $uR_\Box s$. Then, since $wR_\Box v$, $wR_\Box u$ and $R_\Box$ is an equivalence relation, $wR_\Box s$. Since $s \in U_G \cap U_J$, $M, s \models \langle G\rangle\top \wedge \langle J\rangle\top$. Then, since $wR_\Box s$, $M, w \models \Diamond(\langle G\rangle\top \wedge \langle J\rangle\top)$. Taking it all together, $M, w \models \Diamond\langle G\rangle\top \wedge \Diamond\langle J\rangle\top \wedge \Diamond(\langle G\rangle\top \wedge \langle J\rangle\top)$, which contradicts our initial assumption. Hence, there exists no pointed **pSTIT**-model (M, w), such that $M, w \models \Diamond\langle G\rangle\top \wedge \Diamond\langle J\rangle\top \wedge \Box\neg(\langle G\rangle\top \wedge \langle J\rangle\top)$, i.e. $\mathbf{pSTIT} \models (\Diamond\langle G\rangle\top \wedge \Diamond\langle J\rangle\top) \to \Diamond(\langle G\rangle\top \wedge \langle J\rangle\top)$. $\Box$

4. Logic of pSTIT: axiomatization

In this section, we address an axiomatization of the set of validities $Log(\mathbf{pSTIT}) = \{\varphi \in \mathcal{L}_{pSTIT} \mid \mathbf{pSTIT} \models \varphi\}$. It is straightforward to see that $\mathcal{L}_{STIT} \subseteq \mathcal{L}_{pSTIT}$, and the set $Log(\mathbf{STIT}) = \{\varphi \in \mathcal{L}_{STIT} \mid \models \varphi\}$ is already proven to be not finitely axiomatizable (Herzig & Schwarzentruber, 2008). Since $Log(\mathbf{STIT}) \subseteq Log(\mathbf{pSTIT})$, it trivially follows that $Log(\mathbf{pSTIT})$ lacks finite axiomatizability. Therefore, we move to exploring axiomatizations of $\mathcal{L}_{pSTIT}$-fragments.

We give a sound and strongly complete axiomatization for a rich fragment of $\mathcal{L}^{G}_{pSTIT} \subset \mathcal{L}_{pSTIT}$, where we have $[stit]_G$ and $[G]$-modalities for individuals and disjoint non-singleton groups. Our result is new and cannot be reduced to the known completeness results for STIT logic fragments (Schwarzentruber, 2012), although we adapt similar proof techniques.

We also explore a class of subframes $SF(\mathbf{pSTIT})$ that allows us to reason under the assumption that a designated group of agents always acts in accordance with its plan. We give a sound and complete axiom system for the logic of such subframes. This subsection contains only technical results and their detailed proofs. The reader may skip it in case there is no interest in the mathematical properties of the formalisms we present in this paper.

For the whole section, we are going to axiomatize the validities w.r.t the simplified, atemporal **pSTIT**-frames. As it was shown in (Herzig & Schwarzentruber, 2008, Theorem 11), the satisfiability of $\mathcal{L}_{STIT}$-formulae on Kripke STIT frames is equivalent to their satisfiability on the following class of frames:

Definition 4.1 (Atemporal Kripke STIT frames)**.** A class of atemporal Kripke frames (for short **STIT**) consists of structures of the form $\mathcal{F} = (W, R_\Box, \{R_i\}_{i \in Ags})$, where all the components are as in Definition 2.2.

In other words, we drop the relation $R_<$ from Definition 2.2. Since $\mathcal{L}_{STIT}$ contains no temporal modalities, the satisfiability on **STIT** and standard Kripke STIT frames are equivalent. As with Kripke STIT frames, we can get rid of $R_<$ in the definition of **pSTIT** if we want to simplify our reasoning about satisfiability of $\mathcal{L}_{pSTIT}$-formulae.

Definition 4.2 (Atemporal **pSTIT**)**.** An atemporal **pSTIT**-frame is a tuple of the form $F = (W, R_\Box, \{R_i\}_{i \in Ags}, \{U_G\}_{G \subseteq Ags})$, where all the components are as in Definition 3.1.

Analogously to (Herzig & Schwarzentruber, 2008, Theorem 11), $\models$ relations on standard and atemporal **pSTIT**-frames are equivalent.

4.1. Fragment for disjoint groups

We can axiomatize the fragment of plan-restricted group STIT logic that contains group modalities only for some collection of disjoint groups.

Given $Ags = \{1, \ldots, n\}$ and $Var = \{p_1, p_2, \ldots\}$, let $\mathcal{G} \subseteq Gr$ be a set of non-singleton groups, such that for any $G_1, G_2 \in \mathcal{G}$ such that $G_1 \neq G_2$, $G_1 \cap G_2 = \emptyset$. That is, any distinct elements of $\mathcal{G}$ are disjoint. Let $Mod = Ags \cup \mathcal{G}$, i.e., Mod contains individual agents and mutually disjoint groups from $\mathcal{G}$, i.e., the entities about whose actions we are to reason. We define the language $\mathcal{L}^{\mathcal{G}}_{pSTIT}$ as:

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \vee \varphi) \mid \Box\varphi \mid \mid [stit]_G\varphi \mid [G]\varphi$$

where $G \in Mod$. Note that if $G \in 2^{Ags} \setminus Mod$, then $[G]\varphi$, $[stit]_G\varphi$ are not a well-formed formulae of $\mathcal{L}^{\mathcal{G}}_{pSTIT}$.

First, we define the axiom system $\mathbf{L}^{\mathcal{G}}$ (see Table 2 for a definition). Given any set of formulae $\Gamma \subseteq \mathcal{L}_{pSTIT}$ and any formula $\varphi \in \mathcal{L}_{pSTIT}$, by $\Gamma \vdash_{\mathbf{L}^{\mathcal{G}}} \varphi$ we mean that φ is derivable from Γ in $\mathbf{L}^{\mathcal{G}}$.

(CPL)	Tautologies of classical propositional logic
(S5)	**S5** axioms for $\Box, [stit]_G$ for every $G \in Mod$
(KB4)	**KB4** axioms for $[G]$ for every $G \in Mod$
(i-G)	$[i]\varphi \leftrightarrow [stit]_i\varphi$ for any $i \in Ags$
(G-Ind)	$(\Diamond[stit]_{G_1}\varphi_1 \wedge \ldots \wedge \Diamond[stit]_{G_m}\varphi_m) \rightarrow \Diamond([stit]_{G_1}\varphi_1 \wedge \ldots \wedge [stit]_{G_m}\varphi_m)$ for any pairwise disjoint $G_1, \ldots, G_m \in Mod$
(Incl)	$\Box\varphi \rightarrow [stit]_i\varphi$, where $i \in Ags$
(s-Add)	$[stit]_i\varphi \rightarrow [stit]_G\varphi$ for any $i \in G \in \mathcal{G}$
(Contr1)	$\langle G\rangle\top \rightarrow ([G]\varphi \leftrightarrow [stit]_G\varphi)$ for any $G \in Mod$
(Contr2)	$[G]\bot \rightarrow [stit]_G[G]\bot$ for any $G \in Mod$
(RN)	If $\vdash \varphi$, then $\vdash X\varphi$, where $X \in \{[G], [stit]_G\}_{G \in Mod} \cup \{\Box\}$
(MP)	If $\vdash \varphi$ and $\vdash \varphi \rightarrow \psi$, then $\vdash \psi$

Table 2. Logic $\mathbf{L}^{\mathcal{G}}$

S5 axioms for $\Box, [stit]_G$ modalities ensure that $R_\Box, R_i$ (and consequently $\bigcap_{i \in G} R_i$) are equivalence relations. **KB4** axioms for $[G]$ ensure that if G is a group (i.e. $\langle G\rangle\top$), then $[G]$ behaves like a normal group STIT modality, i.e. **S5**-modality. (i-G) ensures that $U_i = W$ for any $i \in Ags$. (G-Ind) ensures that independence of agents (2b of Definition 2.2) holds for individual agents and for disjoint groups. (Incl) corresponds to 2a of Def. 2.2. (s-Add) corresponds to *super-additivity* of group actions, i.e., the group is at least as effective as all of its members. (Contr1) and (Contr2) express that group actions of G are shared actions of G, i.e., it is only up to the choice of actions by members of G which group action (if any) they are to perform.

Next, we prove $\mathbf{L}^{\mathcal{G}}$ to be sound and strongly complete w.r.t. the class of **pSTIT**-models. Our completeness proof strategy is the following. First, we define a broader class of super-additive frames **SA** and show that $\mathbf{L}^{\mathcal{G}}$ is sound and strongly complete w.r.t. **SA**. Then we prove that frame classes **SA** and **pSTIT** are modally equivalent w.r.t. the language $\mathcal{L}^{\mathcal{G}}_{pSTIT}$, providing a method to transform every **SA**-frame into a $\mathcal{L}^{\mathcal{G}}_{pSTIT}$-invariant atemporal **pSTIT**-frame.

First, we define the class of *super-additive* frames **SA** as follows:

Definition 4.3 (**SA**-frames)**.** A tuple $F = (W, R_\Box, \{R_G\}_{G \in Mod}, \{\sim_G\}_{G \in Mod})$[14] is a super-additive frame – for short, **SA**-frame – where:

(1) $W \neq \emptyset$;
(2) $R_\Box, \{R_i\}_{i \in Ags}$ are equivalence relations, such that for any $i \in Ags$, $R_i \subseteq R_\Box$;
(3) For every $G \in Mod$, R_G is an equivalence relation, such that $R_G \subseteq \bigcap_{i \in G} R_i$;
(4) For any $w_1, \ldots, w_m \in W$ and any pairwise disjoint $G_1, \ldots, G_m \in Mod$: if $w_1 R_\Box \ldots R_\Box w_m$, then $R_{G_1}(w_1) \cap \ldots \cap R_{G_1}(w_n) \neq \emptyset$.
(5) For any $G \in Mod$, $\sim_G$ is a symmetric and transitive relation[15] such that for any $w \in W$:

[14]Since we do not have modalities in $\mathcal{L}^{\mathcal{G}}_{pSTIT}$ for any $G \in Gr \setminus \mathcal{G}$, we drop all the corresponding R_G and $\sim_G$-relations from the definition of the frame.

[15]Since $\sim_G$ is transitive and symmetric, for any $w \in W$, if there exists some $v \in W$, such that $w \sim_G v$, then $w \sim_G w$: from $w \sim_G v$ by symmetry it follows that $v \sim_G w$; from $w \sim_G v$ and $v \sim_G w$ by transitivity it follows that $w \sim_G w$. In other words, $\sim_G$ is just like an equivalence relation, but some worlds may have no $\sim_G$-successors.

(a) If $\sim_G (w) \neq \emptyset$, then $R_G(w) =\sim_G (w)$;
(b) If $\sim_G (w) = \emptyset$, then for any $v \in W$, if $wR_G v$, then $\sim_G (v) = \emptyset$;
(c) $\sim_i = R_i$ for any $i \in Ags$;

For every **SA**-frame, plans $\{U_G\}_{G\subseteq Ags}$ can be reconstructed from $\{\sim_G\}_{G\subseteq Ags}$ relations as follows: $U_G = \{w \in W \mid \sim_G (w) \neq \emptyset\}$. It is straightforward to see that 1) every U_G is a R_G-closed set; 2) for every $i \in Ags$, $U_{\{i\}} = W$.

In **SA**-frames, R_G is *super-additive* over the intersection of individual relations, i.e., $R_G \subseteq \bigcap_{i\in G} R_i$, not necessarily $R_G = \bigcap_{i\in G} R_i$. It is important to note that every atemporal **pSTIT**-frame is a **SA**-frame with $R_G = \bigcap_{i\in G} R_i$ and $\sim_G = R_G \cap (U_G \times U_G)$. Atemporal **pSTIT**-frames and **SA**-frames differ only in definitions of $\{R_G\}_{G\subseteq Ags}$: for the latter, $R_G \subseteq \bigcap_{i\in G} R_i$, while for the former, $R_G = \bigcap_{i\in G} R_i$. Property (4) of Definition 4.3 holds for atemporal **pSTIT**-frames, since it follows from the independence of agents and the fact that $R_G = \bigcap_{i\in G} R_i$. We interpret $\mathcal{L}^{\mathcal{G}}_{pSTIT}$-formulae on **SA**-models as in Def. 3.3.

Proposition 4.4. $\mathbf{L}^{\mathcal{G}}$ *is sound and strongly complete w.r.t. the class of* **SA**-*frames.*

Proof. All axioms of $\mathbf{L}^{\mathcal{G}}$ are Sahlqvist formulae[16], whose first-order translations directly correspond to the conditions in Definition 4.3:

$$
\begin{aligned}
\text{(S5)}:&\ R_\Box, R_i, R_G \text{ are equivalence relations}\\
\text{(KB4)}:&\ \sim_G \text{ is serial and transitive}\\
\text{(G-Ind)}:&\ \forall w_1, \ldots, w_m \in W\big(w_1 R_\Box \ldots R_\Box w_m \Rightarrow \bigcap_{i\in\{1,\ldots,m\}} R_{G_i}(w_i) \neq \emptyset\big)\\
&\ \text{for any pairwise disjoint } G_1, \ldots, G_m \in Mod\\
\text{(Incl)}:&\ R_i \subseteq R_\Box\\
\text{(s-Add)}:&\ R_G \subseteq \bigcap_{i\in G} R_i\\
\text{(Contr1)}:&\ \sim_G (w) \neq \emptyset \Rightarrow \sim_G (w) = R_G(w)\\
\text{(Contr2)}:&\ \sim_G (w) = \emptyset \Rightarrow \forall v \in W(wR_G v \Rightarrow \sim_G (v) = \emptyset)\\
\text{(i-G)}:&\ R_i = \sim_i
\end{aligned}
$$

Hence, $\mathbf{L}^{\mathcal{G}}$ is sound and strongly complete w.r.t. the class of **SA**-frames. □

Next, we are going to prove that **SA** and **pSTIT**-frames are equivalent with respect to $\mathcal{L}^{\mathcal{G}}_{pSTIT}$. Namely, we show how to construct a $\mathcal{L}^{G}_{pSTIT}$-equivalent atemporal **pSTIT**-model from any **SA**-model. The technique we are to use was seminally introduced in (Vakarelov, 1992) and then adapted for STIT logics in (Schwarzentruber, 2012, Theorem 3) and in (Lorini, 2013, Lemma 9). We use the notation from (Lorini, 2013, Lemma 9) in our adaptation of the proof.

Definition 4.5. Let $M = (W, R_\Box, \{R_G\}_{G\in Mod}, \{\sim_G\}_{G\in Mod}, V)$ be an **SA**-model. For any $i \in Ags$, let $\Delta_i = \{R_i(w) \mid w \in W\}$. Elements of Δ_i represent i's possible actions.

[16]For the technique of proving completeness via obtaining first-order translations of Sahlqvist formulae, see (Sahlqvist, 1975) or any textbook on modal logic, e.g., (Blackburn, Rijke, & Venema, 2001).

Then, let

$$\Delta = \{\delta \in \prod_{i \in Ags} \Delta_i \mid \bigcap_{i \in Ags} \delta_i \neq \emptyset\}$$

where δ_i denotes the i's projection of δ for any $i \in Ags$. Δ is the set of possible shared actions of Ags. We use $\delta, \gamma, \ldots$ as placeholders for elements of Δ. For any $G \in \mathcal{G}$, by $\delta|_G$ we mean G's projection of δ. E.g., if $\delta = (\delta_1, \ldots, \delta_n)$ and $G = \{1, 2, 3\}$, then $\delta|_G = (\delta_1, \delta_2, \delta_3)$. Note that there is a one-to-one correspondence between Δ and the set $\{\bigcap_{i \in Ags} R_i(w) | w \in W\}$. Analogously, there is a one-to-one correspondence between $\{\delta|_G \mid \delta \in \Delta\}$ and $\{\bigcap_{i \in G} R_i(w) \mid w \in W\}$.

Since $\bigcap_{i \in Ags} R_i$ is an equivalence relation, Δ partitions W, i.e. for every world w there exists a unique $\delta \in \Delta$, such that $w \in \bigcap_{i \in Ags} \delta_i$. For every world $w \in W$ we will denote its corresponding element of Δ as δ^w. With a slight abuse of notation, by $w \in \delta$ we mean that $w \in \bigcap_{i \in Ags} \delta_i$. By $w \in \delta|_G$ we mean that $w \in \bigcap_{i \in G} \delta_i$.

For every $\delta \in \Delta$ and every $G \in \mathcal{G}$, let

$$\Gamma^G_\delta = \{R_G(w) \mid w \in \delta|_G\}$$

be a set of R_G-equivalence classes that are included in $\bigcap_{i \in G} \delta_i$. Since $R_G \subseteq \bigcap_{i \in G} R_i$, it is straightforward to see that Γ^G_δ partitions $\bigcap_{i \in G} \delta_i$. By super-additivity of R_G over $\bigcap_{i \in G} R_i$, Γ^G_δ may have more than one element. Note that if for some $\gamma, \delta \in \Delta$, $\gamma|_G = \delta|_G$, then $\Gamma^G_\delta = \Gamma^G_\gamma$.

For every group $G \in \mathcal{G}$ and every $\delta \in \Delta$, we define an arbitrary surjective function

$$k^G_\delta : \mathbb{R} \to \Gamma^G_\delta$$

that ascribes an element of Γ^G_δ to every real number.[17] We only require that if $\delta|_G = \gamma|_G$, then $k^G_\delta = k^G_\gamma$.

Given that $Ags = \{1, \ldots, n\}$, we define the set W':

$$W' = \{(\delta^w, \vec{i}, w) \mid w \in W, \vec{i} \in \mathbb{R}^n, \forall G \in \mathcal{G} : w \in k^G_{\delta^w}(\sum \vec{i}|_G)\}$$

where $\vec{i}|_G$ is a G-projection of the tuple $\vec{i}$. In other words, every world $w \in W$ has infinitely many copies in W': every copy of w is labeled by its Δ-element δ^w and a tuple of n-many reals $(i_1, \ldots, i_n)$, such that for any group $G \in \mathcal{G}$:

$$w \in k^G_{\delta^w}(\sum_{j \in G} i_j)$$

For example, assume that $Ags = \{1, 2, 3, 4\}$, $\mathcal{G} = \{\{1, 2\}, \{3, 4\}\}$. Let $w \in W$ be a world, such that $w \in k^{\{1,2\}}_{\delta^w}(2)$ and $w \in k^{\{3,4\}}_{\delta^w}(3)$. Let $\vec{i} = (1, 1, 2, 1)$. Then, $(\delta^w, \vec{i}, w) \in W'$. To see why, note that $\vec{i}|_{\{1,2\}} = (1, 1)$, $\vec{i}|_{\{3,4\}} = (2, 1)$, i.e. $\sum \vec{i}|_{\{1,2\}} = \sum(1, 1) = 2$

[17] Given that $\mathcal{L}^{\mathcal{G}}_{pSTIT}$ is countable, any satisfiable set of $\mathcal{L}^{\mathcal{G}}_{pSTIT}$-formulae can be satisfied on a **SA**-model of the cardinality at most $2^{\aleph_0}$ (by a standard canonical model construction and truth lemma). Since every R_G-relation partitions the set of possible worlds, every set of $\mathcal{L}^{\mathcal{G}}_{pSTIT}$-formulae can be satisfied on a **SA**-model with at most $2^{\aleph_0}$-many R_G-equivalence classes. Therefore, $\mathbb{R}$ has a sufficient cardinality for us to be sure that k^G_δ is surjective. When the model is countable or finite, we can take a countable or large enough finite index set, respectively, instead of $\mathbb{R}$.

and $\sum \vec{i}|_{\{3,4\}} = \sum(2,1) = 3$. Since $w \in k^{\{1,2\}}_{\delta^w}(2)$ and $w \in k^{\{3,4\}}_{\delta^w}(3)$, for any $G \in \mathcal{G}$, $w \in k^G_{\delta^w}(\sum \vec{i}|_G)$, hence by definition of W', $(\delta^w, \vec{i}, w) \in W'$.

Finally, for any **SA**-model M we define its corresponding **pSTIT**-model as:

$$M' = (W', R'_\Box, \{R'_i\}_{i \in Ags}, \{U'_G\}_{G \in Mod}, V')$$

(1) $W' = \{(\delta^w, \vec{i}, w) \mid w \in W, \vec{i} \in \mathbb{R}^n, \forall G \in \mathcal{G} : w \in k^G_{\delta^w}(\sum \vec{i}|_G)\}$;
(2) $(\delta^w, \vec{j_1}, w) R'_\Box (\delta^v, \vec{j_2}, v)$ iff $w R_\Box v$;
(3) for any $i \in Ags$, $(\delta^w, \vec{j_1}, w) R'_i (\delta^v, \vec{j_2}, v)$ iff $\delta^w_i = \delta^v_i$, $\vec{j_1}(i) = \vec{j_2}(i)$;
(4) $U'_G = \{(\delta^w, \vec{i}, w) \mid \sim_G (w) \neq \emptyset\}$;
(5) $V'(p) = \{(\delta^w, \vec{i}, w) \mid w \in V(p)\}$.

from which we can derive the following definitions of $\{R'_G, \sim'_G\}_{G \in Mod}$:

(6) for any $G \in Mod$, $(\delta^w, \vec{j_1}, w) R'_G (\delta^v, \vec{j_2}, v)$ iff $\delta^w|_G = \delta^v|_G$, $\vec{j_1}|_G = \vec{j_2}|_G$;
(7) for any $G \in Mod$, $\sim'_G = R'_G \cap (U'_G \times U'_G)$.

For an example of this construction, see Figure 4.

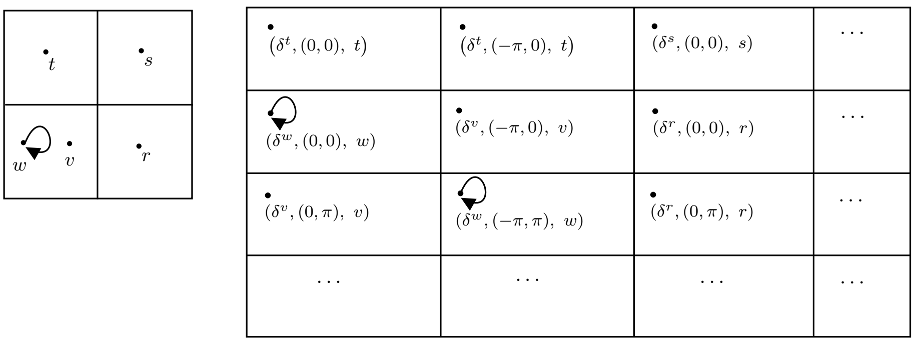

Figure 4. Construction of the atemporal **pSTIT**-model (on the right) from the **SA**-model (on the left), where $Ags = \{1,2\}$, $\mathcal{G} = \{Ags\}$. R_1-equivalence classes in the initial model and R'_1-equivalence classes in the constructed model are depicted as columns, R_2 and R'_2 –as rows. In the initial model, $R_{Ags} = \{(w,w) \mid w \in W\}$ and $\sim_{Ags} = \{(w,w)\}$. The initial model is super-additive, since $\{w\} = R_{Ags}(w) \subset \bigcap_{i \in Ags} R_i(w) = \{w,v\}$. As for Δ, $\delta^w = \delta^v$, and $\Gamma^{Ags}_{\delta^w} = \{\{w\},\{v\}\}$; for any other world $x \in \{t,s,r\}$, δ^x is unique, and $\Gamma^{Ags}_{\delta^x} = \{\{x\}\}$. For any $i \in \mathbb{Q}$, $k^{Ags}_{\delta^w}(i) = \{w\}$; for any $j \in \mathbb{R} \setminus \mathbb{Q}$, $k^{Ags}_{\delta^w}(j) = \{v\}$; for any $\delta \in \Delta \setminus \{\delta^w\}$ and any $i,j \in \mathbb{R}$, $k^{Ags}_\delta(i) = k^{Ags}_\delta(j)$.

Lemma 4.6. *For every* **SA***-model* $M = (W, R_\Box, \{R_G\}_{G \in Mod}, \{\sim_G\}_{G \in Mod}, V)$*, the model* $M' = (W', R'_\Box, \{R'_i\}_{i \in Ags}, \{U'_G\}_{G \in Mod}, V')$ *constructed as in Definition 4.5, is an atemporal* **pSTIT***-model.*

Proof. It is routine to check that $\{R'_i\}_{i \in Ags}, R'_\Box$ are equivalence relations (since they are defined via $=$ and $R_\Box$, respectively, which are equivalence relations). $R'_i \subseteq R'_\Box$ for every $i \in Ags$, since for every $(\delta^w, \vec{i}, w), (\delta^v, \vec{j}, v) \in W'$:

$$\begin{aligned}
&(\delta^w, \vec{i}, w) R'_i (\delta^v, \vec{j}, v) && \\
&\Rightarrow \delta^w_i = \delta^v_i && \text{(Def. 4.5)} \\
&\Leftrightarrow w R_i v && \text{(Def. } \Delta) \\
&\Rightarrow w R_\Box v && \text{(2 of Def. 4.3)} \\
&\Rightarrow (\delta^w, \vec{i}, w) R'_\Box (\delta^v, \vec{j}, v) && \text{(Def. 4.5)}
\end{aligned}$$

For every $w \in W$, $\delta^w|_G = \delta^v|_G$ iff for all $i \in Ags$, $\delta^w_i = \delta^v_i$; and $\vec{j_1}|_G = \vec{j_2}|_G$ iff for all $i \in G$, $\vec{j_1}(i) = \vec{j_2}(i)$. Therefore, it is straightforward that $R'_G = \bigcap_{i \in G} R'_i$.

We show that independence of agents, i.e. 2b of Definition 2.2, holds for M'. Let

$$\begin{aligned}
&\vec{w}_1 = (\delta^{w_1}, (i^1_1, \dots, i^1_n), w_1) \\
&\vec{w}_2 = (\delta^{w_2}, (i^2_1, \dots, i^2_n), w_1) \\
&\dots \\
&\vec{w}_n = (\delta^{w_n}, (i^n_1, \dots, i^n_n), w_n)
\end{aligned}$$

be arbitrary n-many $R'_\Box$-connected worlds. By definition of $R'_\Box$, $w_1 R_\Box \dots R_\Box w_n$. Then, by property (4) of Definition 4.3, there exists a world $v \in R_1(w_1) \cap \dots \cap R_n(w_n)$. In other words, $\delta^v_1 = \delta^{w_1}_1, \dots, \delta^v_n = \delta^{w_n}_n$.

Take a tuple $\vec{i} = (i^1_1, i^2_2, \dots, i^n_n) \in \mathbb{R}^n$. First, we show that there exists a world $x \in W$, such that $(\delta^v, \vec{i}, x) \in W'$. Since every function $k^G_{\delta^v}$ is surjective, for every $G \in \mathcal{G}$ there exists some R_G-equivalence class $k^G_{\delta^v}(\sum \vec{i}|_G)$. W.l.o.g., let $k^G_{\delta^v}(\sum \vec{i}|_G) = R_G(w_G)$ for every $G \in \mathcal{G}$, where $w_G \in \delta^v|_G$. Since $\mathcal{G}$ consists of mutually disjoint sets and all worlds $\{w_G \mid G \in \mathcal{G}\}$ are $R_\Box$-connected, by property (4) of Definition 4.3, these equivalence classes have a non-empty intersection, i.e.

$$\bigcap_{G \in \mathcal{G}} R_G(w_G) = \bigcap_{G \in \mathcal{G}} k^G_{\delta^v}(\sum \vec{i}|_G) \neq \emptyset$$

By definition of $\Gamma^G_{\delta^v}$, $R_G(w_G) \subseteq \bigcap_{i \in G} R_i(v)$ for any $G \in \mathcal{G}$. Therefore,

$$\bigcap_{G \in \mathcal{G}} R_G(w_G) \subseteq \bigcap_{i \in G}^{G \in \mathcal{G}} R_i(v) = \bigcap_{i \in \bigcup \mathcal{G}} R_i(v) \tag{A}$$

.

We show that there exists a possible world $x \in W$, such that:

$$\begin{aligned}
&x \in \bigcap_{G \in \mathcal{G}} R_G(w_G) \cap \bigcap_{i \in Ags \setminus \bigcup \mathcal{G}} R_i(v) && \text{(By (4) of Def. 4.3)} \\
&\Rightarrow x \in \bigcap_{i \in \bigcup \mathcal{G}} R_i(v) \cap \bigcap_{i \in Ags \setminus \bigcup \mathcal{G}} R_i(v) && \text{(By (A))} \\
&\Rightarrow x \in \bigcap_{v \in Ags} R_i(v) && (\text{By} Ags = \bigcup \mathcal{G} \cup Ags \setminus \bigcup \mathcal{G}) \\
&\Rightarrow \delta^x = \delta^v && \text{(By def. of } \Delta)
\end{aligned}$$

Since $\delta^v = \delta^x$ and $x \in k^G_{\delta^v}(\sum \vec{i}|_G) = R_G(w_G)$ for any $G \in \mathcal{G}$, $\vec{x} = (\delta^v, \vec{i}, x) \in W'$.

We can now prove our main statement. Since for every $j \in \{1, \dots, n\}$, $\delta^v_j = \delta^{w_j}_j$ and $\vec{i}(j) = i^j_j$, $\vec{x} \in R'_1(\vec{w}_1) \cap R'_2(\vec{w}_2) \cap \dots \cap R'_n(\vec{w}_n)$. Since $\vec{w}_1, \dots, \vec{w}_n$ were arbitrarily chosen, we can conclude that for any $\vec{w}_1, \dots, \vec{w}_n \in W'$, if $\vec{w}_1 R'_\Box \dots R'_\Box \vec{w}_n$, then $R'_1(\vec{w}_1) \cap \dots \cap R'_n(\vec{w}_n) \neq \emptyset$, i.e., independence of agents holds in M'.

Next, show that (2b) of Def. 3.1 holds for M', i.e., $U'_{\{i\}} = W'$ for every $i \in Ags$. Since $\sim_i = R_i$ and R_i is an equivalence relation, for any $i \in Ags$ and any world $w \in W$, $w \sim_i w$. Therefore, it is straightforward that $U'_i = \{(\delta^w, \vec{j}, w) \mid \sim_i (w) \neq \emptyset\} = W'$ for any $i \in Ags$.

It is left for us to show that for any $G \in \mathcal{G}$, U'_G is R'_G-closed. Take any pair of worlds $\vec{w} = (\delta^w, \vec{j_1}, w)$ and $\vec{v} = (\delta^v, \vec{j_2}, v)$, such that $\vec{w} \in U'_G$ and $\vec{w} R'_G \vec{v}$. We show that $\vec{v} \in U'_G$. Since $\vec{w} R'_G \vec{v}$, $\delta^w|_G = \delta^v|_G$ and $\vec{j_1}|_G = \vec{j_2}|_G$. From the former, $k^G_{\delta^w} = k^G_{\delta^v}$. From the latter, $\sum \vec{j_1}|_G = \sum \vec{j_2}|_G = m$ for some $m \in \mathbb{R}$. Then, by definition of W', $w \in k^G_{\delta^w}(m) = k^G_{\delta^v}(m) \ni v$, i.e., w and v are in the same R_G-equivalence class, i.e., $w R_G v$. Then, since $\vec{w} \in U'_G$, $\sim_G (w) \neq \emptyset$. By property 5a of Def. 4.3, $w \sim_G v$. Then, by symmetry of $\sim_G$, $v \sim_G w$, i.e., $\sim_G (v) \neq \emptyset$. Therefore, $\vec{v} \in U'_G$. This concludes our proof that M' is an atemporal **pSTIT**-model. □

Lemma 4.7. *For every* **SA**-*model* $M = (W, R_\Box, \{R_G\}_{G \in Mod}, \{\sim_G\}_{G \in Mod}, V)$ *and its constructed* **pSTIT**-*model* $M' = (W', R'_\Box, \{R'_i\}_{i \in Ags}, \{U'_G\}_{G \in Mod}, V')$, *there exists a p-morphism from* M' *onto* M.

Proof. Let $f : (\delta^w, \vec{j}, w) \mapsto w$ be a surjective function from W' onto W. We prove that f is a p-morphism.[18] Obviously, $f(\vec{w}) \in V(p)$ iff $\vec{w} \in V'(p)$. As for zig- and zag-conditions, one direction follows directly from the definitions, i.e.

$$\vec{w} R'_i \vec{v} \Rightarrow f(\vec{w}) R_i f(\vec{v})$$
$$\vec{w} R'_\Box \vec{v} \Rightarrow f(\vec{w}) R_\Box f(\vec{v})$$
$$\vec{w} R'_G \vec{v} \Rightarrow f(\vec{w}) R_G f(\vec{v})$$
$$\vec{w} \sim'_G \vec{v} \Rightarrow f(\vec{w}) \sim_G f(\vec{v})$$

The proof for R'_i follows directly from definitions of $\delta \in \Delta$ and $\delta_i \in \Delta_i$. The proof for $R'_\Box$ follows immediately from its definition. We explicitly demonstrate it only for the last two cases.

Assume $(\vec{w}, \vec{v}) \in R'_G$, where $\vec{w} = (\delta^w, \vec{i}, w)$, $\vec{v} = (\delta^v, \vec{j}, v)$. By the definition of R'_G, $\delta^w|_G = \delta^v|_G$; analogously, $\vec{i}|_G = \vec{j}|_G$. From the former, it follows that $k^G_{\delta^w} = k^G_{\delta^v}$. From the latter, it follows that $\sum \vec{i}|_G = \sum \vec{j}|_G = m$ for some $m \in \mathbb{R}$. By the definition of W', $w \in k^G_{\delta^w}(m)$ and $v \in k^G_{\delta^v}(m)$. But $k^G_{\delta^w}(m) = k^G_{\delta^v}(m)$. Hence, w and v are elements of the same R_G-equivalence class, i.e. $w R_G v$.

Assume $(\delta^w, \vec{j_1}, w) \sim'_G (\delta^v, \vec{j_2}, v)$. Then, $(\delta^w, \vec{j_1}, w) R'_G (\delta^v, \vec{j_2}, v)$, from which it follows that $\delta^w|_G = \delta^v|_G$ and $\vec{j_1}|_G = \vec{j_2}|_G$. As we have already shown in the previous

[18] Given two Kripke models $M = (W, R_1, \dots, R_n, V)$ and $M' = (W', R'_1, \dots, R'_n, V')$, a surjective function $f : W' \to W$ is a p-morhism iff 1) for all $w \in W'$, $p \in Var$: $w \in V'(p)$ iff $f(w) \in V(p)$; 2) for any $w, v \in W'$ and any $R' \in \{R'_1, \dots, R'_n\}$: if $wR'v$, then $f(w)Rf(v)$; 3) for any $w, v \in W'$ and any $R' \in \{R'_1, \dots, R'_n\}$: if $f(w)Rv$, then there exists $u \in W'$, such that $f(u) = v$ and $wR'u$. The fundamental preservation theorem: given any two Kripke models M, M', a modal formula φ and a p-morphism $f : W' \to W$, for any world $w \in W'$: $M', w' \models \varphi$ iff $M, f(w) \models \varphi$. For a detailed exposition and a proof of preservation results, see any handbook of modal logic, e.g., (Blackburn et al., 2001, Chapter 3).

case, from $\delta^w|_G = \delta^v|_G$ and $\vec{j_1}|_G = \vec{j_2}|_G$ it follows that w and v are in the same R_G-equivalence class. Hence, by properties (5a-5b of Def. 4.3), either both w and v have no $\sim_G$-successors or $w \sim_G v$. From $(\delta^w, \vec{j_1}, w) \sim'_G (\delta^v, \vec{j_2}, v)$ by definition of $\sim'_G$ and U'_G it follows that both w and v have $\sim_G$-successors. Hence, $w \sim_G v$.

Now, the more difficult, zag-direction:

$$\begin{aligned}
&\text{If } f(\vec{w})R_\Box v, \text{ then } \exists \vec{v} \in W'(f(\vec{v}) = v, \vec{w}R'_\Box \vec{v}) \\
&\text{If } f(\vec{w})R_i v, \text{ then } \exists \vec{v} \in W'(f(\vec{v}) = v, \vec{w}R'_i \vec{v}) \\
&\text{If } f(\vec{w})R_G v, \text{ then } \exists \vec{v} \in W'(f(\vec{v}) = v, (\vec{w}, \vec{v}) \in R'_G) \\
&\text{If } f(\vec{w}) \sim_G v, \text{ then } \exists \vec{v} \in W'(f(\vec{v}) = v, \vec{w} \sim'_G \vec{v})
\end{aligned}$$

Assume $f(\vec{w})R_\Box v$ for some $\vec{w} = (\delta^w, \vec{j_1}, w) \in W'$. Then, $wR_\Box v$. Take any $\vec{v} = (\delta^v, \vec{j_2}, v) \in W'$. By definition of f, $f(\vec{v}) = v$. Since $wR_\Box v$, by definition of $R'_\Box$ it follows that $\vec{w}R'_\Box \vec{v}$.

Assume $f(\vec{w})R_i v$ for some $\vec{w} = (\delta^w, \vec{j_1}, w) \in W'$. Then, $wR_i v$. By definition of Δ_i, it means that $\delta^w_i = \delta^v_i$. For any $G \in \mathcal{G}$, let $g(G) \in \mathbb{R}$ be any real, such that $R_G(v) = k^G_{\delta^v}(g(G))$. Take any $\vec{j_2} \in \mathbb{R}^n$, such that $\vec{j_2}(i) = \vec{j_1}(i)$ and for any $G \in \mathcal{G}$:

$$\sum \vec{j_2}|_G = g(G)$$

We are sure that such $\vec{j_2}$ exists. Assume that $i \notin \bigcup \mathcal{G}$. Then, let $h : \mathcal{G} \to \bigcup \mathcal{G}$ be a function that picks one agent from every group in $\mathcal{G}$, i.e., such that $h(G) \in G$. Let $\vec{j_2}$ be tuple, such that $\vec{j_2}(i) = \vec{j_1}(i)$, $\vec{j_2}(h(G)) = g(G)$ for any $G \in \mathcal{G}$ and for any other agent $l \notin \{i\} \cup \{h(G) \mid G \in Gr\}$, $\vec{j_2}(l) = 0$. It is a routine to check that $\sum \vec{j_2}|_G = g(G)$ for any $G \in \mathcal{G}$ and $\vec{j_2}(i) = \vec{j_1}(i)$. Assume that $i \in G$ for some $G \in \mathcal{G}$. Let $h : \mathcal{G} \to \bigcup \mathcal{G}$ still be a function that picks an agent from every group in $\mathcal{G}$, but such that $h(G) \neq i$. Then, $\vec{j_2}(i) = \vec{j_1}(i)$, $\vec{j_2}(h(G)) = g(G) - \vec{j_1}(i)$, for any $J \in \mathcal{G} \setminus \{G\}$, $\vec{j_2}(h(J)) = g(J)$ and for any other $l \notin \{i\} \cup \{h(G) \mid G \in \mathcal{G}\}$, $\vec{j_2}(l) = 0$. It is again a routine to check that $\sum \vec{j_2}|_G = g(G)$ for any $G \in \mathcal{G}$ and $\vec{j_2}(i) = \vec{j_1}(i)$. Then, in both cases, $\vec{v} := (\delta^v, \vec{j_2}, v) \in W'$ is such that $f(\vec{v}) = v$ and $\vec{w}R_i\vec{v}$.

Assume $f(\vec{w})R_G v$ for some $\vec{w} = (\delta, \vec{j_1}, w) \in W'$. Then, $wR_G v$, i.e. w and v are elements of the same R_G-equivalent class, i.e. $v \in k^G_{\delta^w}(\sum \vec{j_1}|_G)$. By super-additivity of R_G, from $wR_G v$ it follows that $(w, v) \in \bigcap_{i \in G} R_i$, i.e. $\delta^w|_G = \delta^v|_G$. Then, $v \in k^G_{\delta^v}(\sum \vec{j_1}|_G)$. Take any copy of v in W', i.e., any $(\delta^v, \vec{i}, v) \in W$. Let $\vec{j_2}$ be a tuple, such that $\vec{j_2}|_G = \vec{j_1}|_G$, and $\vec{j_2}|_{Ags \setminus G} = \vec{i}|_{Ags \setminus G}$. It is straightforward to see that $(\delta^v, \vec{j_2}, v) \in W'$ and $\vec{w}R'_G\vec{v}$.

Assume $f(\vec{w}) \sim_G v$ for some $\vec{w} = (\delta, \vec{j_1}, w) \in W'$. Then, $w \sim_G v$, from which it follows that $\sim_G (w) \neq \emptyset$. Since $\sim_G (w) \neq \emptyset$, $\vec{w} \in U_G$. From $w \sim_G v$ by 5a of Def. 4.3 it follows that $wR_G v$. By super-additivity of R_G over $\bigcap_{i \in G} R_i$, $(w, v) \in \bigcap_{i \in G} R_i$, i.e. $\delta^w|_G = \delta^v|_G$. Since $wR_G v$, w and v are elements of the same R_G-equivalence class. Then, by definition of k function, for any $m \in \mathbb{R}$, $w \in k^G_{\delta^w}(m)$ iff $v \in k^G_{\delta^v}(m)$ iff $m = \sum \vec{j_1}|_G$. Now, we show that there is a world $\vec{v}$, such that $\vec{w} \sim'_G \vec{v}$ and $f(\vec{v}) = v$. Take any copy of v in W', i.e., any world $(\delta^v, \vec{i}, v) \in W'$. Let $\vec{j_2} \in \mathbb{R}^n$ be a tuple, such that $\vec{j_2}|_G = \vec{j_1}|_G$ and $\vec{j_2}|_{Ags \setminus G} = \vec{i}|_{Ags \setminus G}$. It is straightforward to see that $\vec{v} = (\delta^v, \vec{j_2}, v) \in W'$. Note that since $w \sim_G v$, $\sim_G (v) \neq \emptyset$, i.e., $\vec{v} \in U'_G$. Since $\delta^w|_G = \delta^v|_G$ and $\vec{j_1}|_G = \vec{j_2}|_G$, $\vec{w}R'_G\vec{v}$. Since $\vec{w}R'_G\vec{v}$ and $\vec{w}, \vec{v} \in U'_G$, it follows that $(\vec{w}, \vec{v}) \in R'_G \cap (U'_G \times U'_G)$, i.e., $\vec{w} \sim'_G \vec{v}$, while $f(\vec{v}) = v$. This case concludes the proof that $f : W' \to W$ is a p-morphism. $\square$

Theorem 4.8 ($\mathbf{L}^{\mathcal{G}}$ is sound and complete)**.** *Given any set of formulae* $\Gamma \subseteq \mathcal{L}^{\mathcal{G}}_{pSTIT}$ *and a formula* $\varphi \in \mathcal{L}^{\mathcal{G}}_{pSTIT}$*, the following holds:*

$$\Gamma \vdash_{\mathbf{L}^{\mathcal{G}}} \varphi \Leftrightarrow \Gamma \models_{\mathbf{pSTIT}} \varphi$$

Proof. ($\Rightarrow$) By a routine check of validity of axioms and validity preservation under the rules of inference.

($\Leftarrow$) By contraposition. Assume that $\Gamma \nvdash_{\mathbf{L}^{\mathcal{G}}} \varphi$. Then, $\Gamma \cup \{\neg\varphi\}$ is $\mathbf{L}^{\mathcal{G}}$-consistent. Then, by Proposition 4.4, there exists a **SA**-model

$$M = (W, R_{\Box}, \{R_G\}_{G \in Mod}, \{\sim_G\}_{G \in Mod}, V)$$

and a world $w \in W$, such that $M, w \models \Gamma \cup \{\neg\varphi\}$. By Lemma 4.7, there exists an atemporal **pSTIT**-model $M' = (W', R'_{\Box}, \{R'_i\}_{i \in Ags}, \{U'_G\}_{G \in Mod}, V')$ and a p-morphism $f : W' \to W$. Hence, there exists a world $\vec{w} \in W'$, such that $f(\vec{w}) = w$. Then, $M', \vec{w} \models \Gamma \cup \{\neg\varphi\}$, hence $\Gamma \not\models_{\mathbf{pSTIT}} \varphi$. $\square$

4.2. Plan-based subframes

We often reason under the assumption that some group always acts in accordance with its plan. For example, in our everyday lives, we often say that we can buy food in a grocery store, get to work by public transport, and meet a colleague in their office no matter what others are going to do. In reality, we have these abilities only conditional on other agents following the plans of their groups, i.e., grocery workers, bus drivers and our colleagues being at their workplaces. We take such restrictions on others' behaviour just as a background for our practical reasoning.

To model such reasoning, we do not need the full **pSTIT**-models: the models regard some choices of group members' actions as possible, while these choices are forbidden by the group's plan. Hence, we take into account mere metaphysical possibilities of agents' actions that we are sure they will not choose to do, being restricted by the group's plan. If we want to reason about such plan-abiding behaviour of agents, we can use *plan-based submodels.*

Definition 4.9 (Plan-based subframes of **pSTIT**)**.** Given a **pSTIT**-model $M = (W, R_{\Box}, R_{<}, \{R_i\}_{i \in Ags}, \{U_G\}_{G \subseteq Ags}, V)$ and a group $G \in Gr$, a U_G*-based submodel* of M is a tuple $M_G = (\mathcal{W}, \mathcal{R}_{\Box}, \mathcal{R}_{<}, \{\mathcal{R}_i\}_{i \in Ags}, \{\mathcal{U}_J\}_{J \subseteq Ags}, \mathcal{V})$, such that:

- $\mathcal{W} = U_G$;
- $\mathcal{R}_{\Box} = R_{\Box} \cap (U_G \times U_G)$;
- $\mathcal{R}_{<} = R_{<} \cap (U_G \times U_G)$;
- For any $i \in Ags$: $\mathcal{R}_i = R_i \cap (U_G \times U_G)$. Consequently, $\mathcal{R}_G = \bigcap_{i \in G} \mathcal{R}_i$ for any $G \subseteq Ags$;
- For all $J \subseteq Ags$: $\mathcal{U}_J = U_J \cap U_G$;
- For any $p \in Var$: $\mathcal{V}(p) = V(p) \cap U_G$.

A notion of a U_G-based subframe $F_G = (\mathcal{W}, \mathcal{R}_{<}, \mathcal{R}_{\Box}, \{\mathcal{R}_i\}_{i \in Ags}, \{\mathcal{U}_J\}_{J \subseteq Ags})$ is defined accordingly. We call a set $\{M_G \mid G \in Gr\}$ *plan-based submodels of* M. We denote a class of plan-based subframes of **pSTIT** as $SF(\mathbf{pSTIT})$, i.e. $F \in SF(\mathbf{pSTIT})$ iff F is a plan-based subframe of some **pSTIT**-frame for some group $G \in Gr$.

Intuitively, given some **pSTIT**-model and a proper group with a non-empty plan $G \in Gr$, a U_G-submodel of M simulates agents' possible choices of action under the assumption that members of G will behave as a group and follow their plan. Such models are especially fruitful for reasoning from G members' point of view: group members naturally expect themselves and their peers to follow the mutually agreed-upon plan, so they might not take into account possible actions of themselves and their peers that go against the plan.

From now on, we will work with the full language, $\mathcal{L}_{pSTIT}$. The satisfiability relation on plan-based subframes is defined standardly:

Definition 4.10 (Satisfiability on $SF(\mathbf{pSTIT})$-models)**.** Given any plan-based model $M_G \in SF(\mathbf{pSTIT})$ and any fromula $\varphi \in \mathcal{L}_{STIT}$, we inductively define $\models$ relation as follows:

$$
\begin{aligned}
&M_G, w \models p \Leftrightarrow w \in \mathcal{V}(p)\\
&M_G, w \models \neg\varphi \Leftrightarrow \text{ not } M_G, w \models \varphi\\
&M_G, w \models \varphi \vee \psi \Leftrightarrow M_G, w \models \varphi \text{ or } M_G, v \models \psi\\
&M_G, w \models \Box\varphi \Leftrightarrow \forall v \in \mathcal{W}(w\mathcal{R}_\Box v \Rightarrow M_G, v \models \varphi)\\
&M_G, w \models [stit]_G\varphi \Leftrightarrow \forall v \in \mathcal{W}(w\mathcal{R}_G v \Rightarrow M_G, v \models \varphi)\\
&M_G, w \models [J]\varphi \Leftrightarrow w \in \mathcal{U}_J \Rightarrow M_G, w \models [stit]_G\varphi
\end{aligned}
$$

Interestingly, the set of validities of the plan-based subframes for the whole language $\{\varphi \in \mathcal{L}_{pSTIT} \mid SF(\mathbf{pSTIT}) \models \varphi\}$ can be finitely axiomatized. See the definition of $\mathbf{L}^p_G$ in Table 3.

(PL)	Axioms for classical propositional logic
(S5)	**S5** axioms for $\Box$, $[stit]_G$ for any $G \subseteq Ags$
(Incl)	$\Box\varphi \to [i]\varphi$ for any $i \in Ags$
(i-G)	$[i]\varphi \leftrightarrow [stit]_i\varphi$ for any $i \in Ags$
(KB4)	**KB4** axioms for $[G]$ for any $G \subseteq Ags$
(S-Add)	$[stit]_G\varphi \to [stit]_{G'}\varphi$ for every $G \subseteq G' \subseteq Ags$
(Contr1)	$\langle G\rangle\top \to ([G]\varphi \leftrightarrow [stit]_G\varphi)$ for every $G \subseteq Ags$
(Contr2)	$[G]\bot \to [stit]_G[G]\bot$ for every $G \subseteq Ags$
(Pl)	$\bigvee_{G\in Gr} \Box\langle G\rangle\top$
(RN)	If $\vdash \varphi$, then $\vdash X\varphi$, where $X \in \{[G], [stit]_G\}_{G\subseteq Ags} \cup \{\Box\}$
(MP)	If $\vdash \varphi \to \psi$ and $\vdash \varphi$, then $\vdash \psi$

Table 3. Logic $\mathbf{L}^p_G$

In comparison with $\mathbf{L}^{\mathcal{G}}$, we lose (G-Ind) axiom, since plan-based subframes do not necessarily satisfy the independence of agents constraint. We get a new axiom (Pl), which expresses the fact, in plan-based subframes, at least one group necessarily acts in accordance with its plan.

We proceed with proving that $\mathbf{L}^p_G$ is sound and strongly complete w.r.t $SF(\mathbf{pSTIT})$-models. First, we show that $\mathbf{L}^p_G$ is sound and strongly complete w.r.t. a special class of Kripke frames $\mathbf{KB4}_J$. Then we define a subclass of $\mathbf{KB4}_G \subseteq \mathbf{KB4}_J$ and show that every $\mathbf{KB4}_J$-frame is a p-morphic image of some $\mathbf{KB4}_G$-frame, which guarantees that $\mathbf{KB4}_J \models \varphi$ iff $\mathbf{KB4}_G \models \varphi$ for any $\varphi \in \mathcal{L}_{pSTIT}$. After that we show that every $\mathbf{KB4}_G$-frame is a plan-based subframe of some **pSTIT**-frame.

Definition 4.11 ($\mathbf{KB4}_J$-frames)**.** A $\mathbf{KB4}_J$-frame is a tuple of the form:

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

(1) $\{R_i\}_{i\in Ags}$, $R_\Box$ are equivalence relations;
(2) $R_i \subseteq R_\Box$ for every $i \in Ags$;
(3) $\{R_G\}_{G\subseteq Ags}$ are equivalence relations, such that $R_{G'} \subseteq R_G$ for every $G \subseteq G' \subseteq Ags$;
(4) $\sim_G$ is a symmetric transitive relation, such that:
  (a) If $\sim_G (w) \neq \emptyset$, then $\sim_G (w) = R_G(w)$;
  (b) If $\sim_G (w) = \emptyset$, then for any $v \in W$, such that wR_Gv, $\sim_G (v) = \emptyset$;
  (c) $\sim_i = R_i$ for every $i \in Ags$;
(5) At every world $w \in W$ there exists a proper group $G \in Gr$, such that $\forall v \in W(wR_\Box v \Rightarrow \sim_G (v) \neq \emptyset)$;

Just like in Def. 3.1, for every $G \subseteq Ags$, $U_G = \{w \in W \mid \sim_G (w) \neq \emptyset\}$

Definition 4.12 ($\mathbf{KB4}_G$-frames)**.** A tuple $F = (W, R_\Box, \{R_G\}_{G\subseteq Ags}, \{\sim_G\}_{G\subseteq Ags})$ is a $\mathbf{KB4}_G$-frame iff F is a $\mathbf{KB4}_J$-frame, such that for any group $G \subseteq Ags$, $R_G = \bigcap_{i\in G} R_i$.

From Definitions 4.12 and 4.11 it follows that $\mathbf{KB4}_G \subseteq \mathbf{KB4}_J$. The difference is only in the definition of $\{R_G\}_{G\subseteq Ags}$ relations.

Lemma 4.13. $\mathbf{L}^p_G$ *is sound and strongly complete w.r.t.* $\mathbf{KB4}_J$*-frames, i.e. for any set of formulae* $\Gamma \subseteq \mathcal{L}_{STIT}$ *and any formula* $\varphi \in \mathcal{L}_{STIT}$*:*

$$\Gamma \vdash_{\mathbf{L}^p_G} \varphi \Leftrightarrow \Gamma \models_{\mathbf{KB4}_J} \varphi$$

Proof. All $\mathbf{L}^p_G$ axioms are Sahlqvist formulae whose first-order translations directly correspond to (1)-(5) of Definition 4.11. By the Sahlqvist theorem (Sahlqvist, 1975), $\mathbf{L}^p_G$ is canonical and hence sound and strongly complete w.r.t. $\mathbf{KB4}_J$. $\Box$

Since $\mathbf{KB4}_G \subseteq \mathbf{KB4}_J$ and frame validities are closed under p-morphic images, if every $\mathbf{KB4}_J$-frame is a p-morphic image of some $\mathbf{KB4}_G$-frame, then $\{\varphi \in \mathcal{L}_{pSTIT} \mid \mathbf{KB4}_G \models \varphi\} = \{\varphi \in \mathcal{L}_{pSTIT} \mid \mathbf{KB4}_J \models \varphi\}$.[19] To demonstrate exactly that, we use an unraveling technique, first introduced (to the best to our knowledge) in (Fagin, Halpern, & Vardi, 1992) and later simplified in (Wáng & Ågotnes, 2020). We modify the method used in (Wáng & Ågotnes, 2020): we use it on the level of frames instead of applying it to a canonical model only.

Definition 4.14 (Paths, initial segments)**.** Take any $\mathbf{KB4}_J$-frame $F = (W, R_\Box, \{R_G\}_{G\subseteq Ags}, \{\sim_G\}_{G\subseteq Ags})$. By a path on F we mean a finite, non-empty sequence $\vec{w} = (w_1, G_1, \dots, G_{n-1}, w_n)$, such that $w_1, \dots, w_n \in W$, $G_1, \dots, G_{n-1} \subseteq$

[19] For the definitions of modally definable classes of frames, p-morphisms and the preservability results, see the classic result in (Goldblatt & Thomason, 1975, p.168) or any handbook of modal logic, e.g. (Blackburn et al., 2001, Chapter 3).

Ags and for every pair w_i, w_{i+1}: $(w_i, w_{i+1}) \in R_{G_i}$. If for some i, $G_i = \emptyset$, then $(w_i, w_{i+1}) \in R_\Box$. By $head(\vec{w})$ we mean the first world that appears in the sequence (i.e. $head((w_1, G_1, \dots, G_{n-1}, w_n)) = w_1$); $tail(\vec{w})$ is the last world that appears in $\vec{w}$ (i.e. $tail((w_1, G_1, \dots, G_{n-1}, w_n)) = w_n$). Note that for any world $w \in W$, (w) is also a path.

Given two paths $\vec{w} = (w_1, G_1, \dots, G_{n-1}, w_n)$, $\vec{v} = (v_1, J_1, \dots, J_{k-1}, v_k)$, we say that $\vec{w}$ is an initial segment of $\vec{v}$ ($\vec{w} \sqsubseteq \vec{v}$) iff $n \leq k$, $w_i = v_i$ for any $1 \leq i \leq n$ and $G_i = J_i$ for any $1 \leq i \leq n-1$. Then, $\vec{v} \setminus \vec{w} = (w_n J_n, \dots, J_{k-1}, v_k)$. If $\vec{v} = \vec{w}$, then $\vec{v} \setminus \vec{w} = (tail(\vec{v}))$.

Given a path $\vec{w} = (w_1, G_1, \dots, G_{n-1}, w_n)$ and a group $J \subseteq Ags$, we say that $\vec{w}$ is a J-path iff $J \subseteq \bigcap_{1 \leq i \leq n-1} G_i$. Note that paths of the length 1, i.e. paths of the form (w), are vacuously J-path for any $J \subseteq Ags$.

Definition 4.15 (Unraveling)**.** For any $\mathbf{KB4}_J$-frame $F = (W, R_\Box, \{R_G\}_{G \subseteq Ags}, \{\sim_G\}_{G \subseteq Ags})$ we define its unraveling as:

$$\mathfrak{F} = (\mathfrak{W}, \mathfrak{R}_\Box, \{\mathfrak{R}_G\}_{G \subseteq Ags}, \{\simeq_G\}_{G \subseteq Ags})$$

(1) $\mathfrak{W} = \{\vec{w} \mid \vec{w}$ is a path on $F\}$;
(2) $\vec{w}\mathfrak{R}_\Box\vec{v}$ iff 1) $\vec{w}$ and $\vec{v}$ have a common initial segment, i.e. $\exists \vec{g} : \vec{g} \sqsubseteq \vec{w}$ and $\vec{g} \sqsubseteq \vec{v}$;
(3) $\vec{w}\mathfrak{R}_G\vec{v}$ iff 1) $\vec{w}$ and $\vec{v}$ have some common initial segment $\vec{g}$; 2) $\vec{w} \setminus \vec{g}$ and $\vec{v} \setminus \vec{g}$ are G-paths;
(4) $\vec{w} \simeq_G \vec{v}$ iff $tail(\vec{w}) \sim_G tail(\vec{v})$;

Proposition 4.16. *Given any* $\mathbf{KB4}_J$*-frame* $F = (W, R_\Box, \{R_G\}_{G \subseteq Ags}, \{\sim_G\}_{G \subseteq Ags})$*, its unraveling* $\mathfrak{F} = (\mathfrak{W}, \mathfrak{R}_\Box, \{\mathfrak{R}_G\}_{G \subseteq Ags}, \{\simeq_G\}_{G \subseteq Ags})$ *is a* $\mathbf{KB4}_G$*-frame.*

Proof. It is a routine to check that $\mathfrak{R}_i$, $\mathfrak{R}_\Box, \mathfrak{R}_G$ are equivalence relations, s.t. $\mathfrak{R}_i \subseteq \mathfrak{R}_\Box$. $\mathfrak{F}$ also trivially satisfies (4a-4b) of Def. 4.11, given that $\simeq_G$ is defined via $\sim_G$.

Property 5 of Def. 4.11 holds for $\mathfrak{F}$: at every world $\vec{w} \in \mathfrak{W}$ there exists a non-singleton group $G \in Gr$, such that $\forall \vec{v} \in \mathfrak{W}(\vec{w}\mathfrak{R}_\Box\vec{v} \Rightarrow \simeq_G (v) \neq \emptyset)$. To see why, take any path $\vec{w} \in W$, such that $tail(\vec{w}) = w$. Since $F \in \mathbf{KB4}_J$, by Definition 4.11 there exists a non-singleton group $G \in Gr$, such that for any $v \in W$, if $wR_\Box v$, then $\sim_G (v) \neq \emptyset$. Let us fix such group as G. Take any $\vec{v} \in \mathfrak{R}_\Box(\vec{w})$. By definition of $\mathfrak{R}_\Box$, $\vec{w}$ and $\vec{v}$ have some common initial segment $\vec{g}$. Let $tail(\vec{g}) = g$. Then, there exist paths $(g, G_1, \dots, G_n, w)$ and $(g, J_1, \dots, J_k, v)$. Since $R_{G_1}, \dots, R_{G_n}, R_{J_1}, \dots, R_{J_k} \subseteq R_\Box$, $gR_\Box \dots R_\Box w$ and $gR_\Box, \dots, R_\Box v$, and by transitivity of $R_\Box$, $wR_\Box v$. Then, since $\forall v \in W(wR_\Box v \Rightarrow \sim_G (v) \neq \emptyset)$, $\sim_G (v) \neq \emptyset$. Hence, there exists a world $u \in W$, such that $v \sim_G u$. By 4a of Def.4.11, $uR_G w$. Then, a path $\vec{v}$ can be extended to a path $\vec{u}$, such that $\vec{u} \setminus \vec{v} = (v, G, u)$. Since $\vec{v}$ and $\vec{u}$ have common initial segment, $\vec{v}\mathfrak{R}_\Box\vec{u}$. Since $v \sim_G v$, $\vec{v} \simeq_G \vec{u}$. Hence, $\simeq_G (\vec{v}) \neq \emptyset$. Since $\vec{v}$ was an arbitrary $\mathfrak{R}_\Box$-successor of an arbitrary world $\vec{w} \in \mathfrak{W}$, we can conclude that for any $\vec{w} \in \mathfrak{W}$, there exists a group $G \in Gr$, such that for any $\vec{v} \in \mathfrak{W}$, $(\vec{w}\mathfrak{R}_\Box\vec{v} \Rightarrow \simeq_G (v) \neq \emptyset)$.

So it is left for us to show that $\vec{w}\mathfrak{R}_G\vec{v}$ iff $(\vec{w}, \vec{v}) \in \bigcap_{i \in G} \mathfrak{R}_i$. Left-to-right direction. Assume $\vec{w}\mathfrak{R}_G\vec{v}$. It means that $\vec{v}$, $\vec{w}$ share some initial segment $\vec{g}$ and $\vec{w} \setminus \vec{g}$, $\vec{v} \setminus \vec{g}$ are G-paths. By definition of a G-path, for every $i \in G$, every G-path is a i-path. Hence, $\vec{v}$, $\vec{w}$ share some initial segment $\vec{g}$ and $\vec{w} \setminus \vec{g}$, $\vec{v} \setminus \vec{g}$ are i-paths for every $i \in G$. By definition of $\mathfrak{R}_i$, it means that $(\vec{w}, \vec{v}) \in \bigcap_{i \in G} \mathfrak{R}_i$. Right-to-left direction. Assume $(\vec{w}, \vec{v}) \in \bigcap_{i \in G} \mathfrak{R}_i$. By definition of $\mathfrak{R}_i$, $(\vec{w}, \vec{v}) \in \bigcap_{i \in G} \mathfrak{R}_i$ means that for every $i \in G$, $\vec{w}$ and $\vec{v}$ have some initial segment $\vec{g_i}$ and $\vec{w} \setminus \vec{g_i}$, $\vec{v} \setminus \vec{g_i}$ are i-paths. By definition of a path, $\vec{w}, \vec{v}$ are finite. Hence, given a set of such common initial segments $\{\vec{g_i}\}_{i \in G}$, there is the longest path $\vec{g} \in \{\vec{g}_i\}_{i \in G}$, i.e. for every $\vec{t} \in \{\vec{g}_i\}_{i \in G}$, $\vec{g} \setminus \vec{t}$ is defined. Moreover, $\vec{w} \setminus \vec{g}$, $\vec{v} \setminus \vec{g}$ are i-paths for every $i \in G$. Hence, $\vec{w}$, $\vec{v}$ have an initial segment $\vec{g}$, such that $\vec{w} \setminus \vec{g}$, $\vec{v} \setminus \vec{g}$ are G-paths. Then, by definition of $\mathfrak{R}_G$, $\vec{w}\mathfrak{R}_G\vec{v}$. Taking it all together, we have shown that $\mathfrak{F} \in \mathbf{KB4}_G$. $\square$

Lemma 4.17. *Given a* $\mathbf{KB4}_J$*-frame* $F = (W, R_\Box, \{R_G\}_{G \subseteq Ags}, \{U_G\}_{G \subseteq Ags})$ *and its unraveling* $\mathfrak{F} = (\mathfrak{W}, \mathfrak{R}_\Box, \{\mathfrak{R}_G\}_{G \subseteq Ags}, \{\simeq_G\}_{G \subseteq Ags})$*, a function* $f : \mathfrak{W} \to W$*, such that* $f(\vec{w}) = tail(\vec{w})$*, is a p-morphism.*

Proof. The function f is obviously surjective, since every world $w \in W$ is a tail of at least one path (w). We claim that f is a p-morphism.

Assume $\vec{w}\mathfrak{R}_\Box\vec{v}$. Then, $\vec{w}$ and $\vec{v}$ share some initial segment $\vec{g}$, which means that there exists a world $tail(\vec{g}) = g$, such that there is a path from g to $f(\vec{w})$ and from g to $f(\vec{v})$, where every step of these paths is some of $\{R_G\}_{G \subseteq Ags}$ relations. Since every R_G-relation is a subset of $R_\Box$, those paths are $R_\Box$-paths. And since $R_\Box$ is transitive, $gR_\Box f(\vec{w})$ and $gR_\Box f(\vec{v})$. Since $R_\Box$ is Euclidean and symmetric, $f(\vec{w})R_\Box f(\vec{v})$.

Assume $f(\vec{w})R_\Box v$. I.e., $tail(\vec{w}) = w_n$ and $w_n R_\Box v$. Then, let $\vec{v}$ be a path that extends $\vec{w}$ with $(w_n, \emptyset, v)$, i.e. $\vec{v} \setminus \vec{w} = (w_n, \emptyset, v)$. $\vec{w}$ and $\vec{v}$ share an initial segment, namely, $\vec{w}$. Hence, $\vec{w}\mathfrak{R}_\Box\vec{v}$. Moreover, since $tail(\vec{v}) = v$, $f(\vec{v}) = v$.

Assume $\vec{w}\mathfrak{R}_G\vec{v}$. Then, $\vec{w}$, $\vec{v}$ share some initial segment $\vec{g}$, such that $\vec{w} \setminus \vec{g}$,$\vec{v} \setminus \vec{g}$ are G-paths. Let $tail(\vec{w}) = w$, $tail(\vec{v}) = v$, $tail(\vec{g}) = g$. Given that R_G is transitive and the definition of a path, it means that gR_Gw and gR_Gv. Since R_G is symmetric, wR_Gg and gR_Gv. Using transitivity again, wR_Gv, i.e. $f(\vec{w})R_Gf(\vec{v})$.

Assume $f(\vec{w})R_Gv$. Let $tail(\vec{w}) = w$. Then, wR_Gv. There exists a path $\vec{v} = (\vec{w}, G, v)$, i.e. $\vec{v} \setminus \vec{w} = (w, G, v)$. Note that $\vec{w}, \vec{v}$ have the initial segment $\vec{w}$ and $\vec{w} \setminus \vec{v}$ is a G-path ($\vec{w} \setminus \vec{w}$ is vacuously a G-path). Hence, $\vec{w}\mathfrak{R}_G\vec{v}$ and $f(\vec{v}) = v$.

Case of $\simeq_G$ and $\sim_G$ is trivial, since $\vec{w} \simeq_G \vec{v}$ iff $f(\vec{w}) \sim_G f(\vec{v})$

So that, f is a p-morphism from $\mathfrak{F}$ onto F. Since F was arbitrary, we can conclude that for every $\mathbf{KB4}_J$-frame there exists a $\mathbf{KB4}_G$-frame, such that the latter is a p-morphic image of the former. $\square$

Theorem 4.18 ($\mathbf{KB4}_G \subseteq SF(\mathbf{pSTIT})$)**.** *Every* $\mathbf{KB4}_G$*-frame is a plan-based subframe of some atemporal* **pSTIT***-frame.*

Proof. Take any $\mathbf{KB4}_G$-frame $F = (W, R_\Box, \{R_G\}_{G \subseteq Ags}, \{\sim_G\}_{G \subseteq Ags})$ for some set of agents $Ags = \{1, \dots, n\}$. W.l.o.g., we assume that F is point-generated (otherwise, the procedure we are to present should be done for every point-generated subframe of F, after which we should take their disjoint union). Note that in a point-generated frame, $R_\Box$ is a universal relation, since for any other relation R_G, $R_G \subseteq R_\Box$. Then, by (5) of Definition 4.11, there exists a group $J \in Gr$, such that $\forall w \in W : \sim_J (w) \neq \emptyset$. Fix this group as J. We are to construct a **pSTIT**-frame $\mathcal{F}$, such that F is its U_J-generated subframe.

By Definition 4.2, in order to come up with an atemporal **pSTIT**-frame we need to construct a **STIT**-frame $(W, R_\Box, \{R_i\}_{i \in Ags})$ and a $\bigcap_{i \in G} R_i$-closed set U_G for every $G \subseteq Ags$. We begin with a **STIT**-frame. Take the tuple $(W, R_\Box, \{R_i\}_{i \in Ags})$ from F. The only thing that differentiates it from a **STIT** frame is that $(W, R_\Box, \{R_i\}_{i \in Ags})$ may not satisfy the independence of agents constraint. We will reformulate the constraint in a certain way, so that it will be easier for us to build a **STIT**-frame out of the given tuple.

Every R_i is an equivalence relation, so R_i yields a partition $[R_i] = \{R_i(w) \mid w \in W\}$ of W. We can index elements of each partition as $[R_i] = \{X_1^i, \dots, X_{Card([R_i])}^i\}$. In such

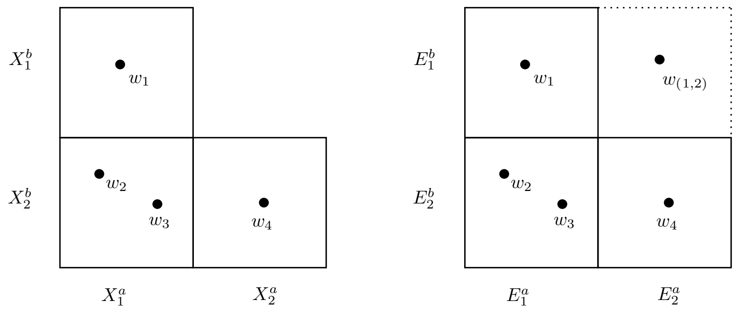

Figure 5. Example of constructing an atemporal **STIT**-frame (on the right) from a $\mathbf{KB4}_G$-frame (on the left) for two agents. In the initial $\mathbf{KB4}_G$-frame, the independence of agents fails, since $X_2^a \cap X_1^b = \emptyset$. In a constructed **STIT**-frame, $E_2^a \cap E_1^b = w_{(1,2)}$, where $w_{(1,2)}$ is fresh.

a case, the independence of agents constraint is satisfied iff $X^1_{i_1} \cap \ldots \cap X^n_{i_n} \neq \emptyset$ for every $i_1 \leq Card([R_1]), \ldots, i_n \leq Card([R_n])$.

Since we are not sure if independence of agents holds for $(W, R_\Box, \{R_i\}_{i\in Ags})$, there might be elements $X^1_{i_1} \in [R_1], \ldots, X^n_{i_n} \in [R_n]$, such that $X^1_{i_1} \cap \ldots \cap X^n_{i_n} = \emptyset$. To ensure the independence of agents and construct a **STIT**-frame, we need to make all such intersections non-empty. For this purpose, we will extend W to a larger set of possible worlds – $\mathcal{W}$ – by putting a new world for every such empty intersection. A a result, we will get an extension of the set of possible worlds W to a new set $\mathcal{W}$, and every relation R_i defined on W will be extended to $\mathcal{R}_i$ defined on $\mathfrak{W}$.

First, we create a *skeleton* of a **STIT**-frame. For every agent $i \in Ags$, let $[\mathcal{R}_i] = \{E^i_1, \ldots, E^i_{Card([R_i])}\}$ be a collection of sets of possible worlds, which are to serve as $\mathcal{R}_i$-equivalence classes. Our task is to describe elements of all these sets in a manner such that for every $E_1 \in [\mathcal{R}_i], \ldots, E_n \in [\mathcal{R}_n]$, $E_1 \cap \ldots \cap E_n \neq \emptyset$. We do it as follows. For every $E^1_{j_1} \in [\mathcal{R}_1], \ldots, E^n_{j_n} \in [\mathcal{R}_n]$:

$$E^1_{j_1} \cap \ldots \cap E^n_{j_n} = \begin{cases} X^1_{j_1} \cap \ldots \cap X^n_{j_n} & \text{if } X^1_{j_1} \cap \ldots \cap X^n_{j_n} \neq \emptyset \\ \{w_{(j_1,\ldots,j_n)}\} & \text{otherwise} \end{cases}$$

where every $w_{(j_1,\ldots,j_n)}$ is just a unique "fresh" possible world, i.e. $w_{(j_1,\ldots,j_n)} \notin W$.

Knowing the elements of every intersection of the form $E_1 \cap \ldots \cap E_n$, we know the elements of every equivalence class E_i: $w \in E_i$ iff there exists $E_1 \in [\mathcal{R}_1], \ldots, E_{i-1} \in [\mathcal{R}_{i-1}], E_{i+1} \in [\mathcal{R}_{i+1}], \ldots, E_n \in [\mathcal{R}_n]$, such that $w \in E_1 \cap \ldots \cap E_i \cap \ldots \cap E_n$.

So we define a tuple $(\mathcal{W}, \mathcal{R}_\Box, \{\mathcal{R}_i\}_{i\in Ags})$, where:

- $\mathcal{W} = \{w \in E_1 \cap \ldots \cap E_n \mid E_1 \in [\mathcal{R}_1], \ldots, E_n \in [\mathcal{R}_n]\}$, i.e. $\mathcal{W}$ consists of W and all the added "fresh" worlds;
- $\mathcal{R}_\Box = \mathcal{W} \times \mathcal{W}$;
- For any $w, v \in \mathcal{W}$, $w\mathcal{R}_i v$ iff there exists $E \in [\mathcal{R}_i]$ such that $w, v \in E$;

See Figure 5 for the example of the construction. According to the Definition 4.1, $(\mathcal{W}, \mathcal{R}_\Box, \{\mathcal{R}_i\}_{i\in Ags})$ is a **STIT**-frame. Every $[\mathcal{R}_i]$ is a partition of $\mathcal{W}$, hence, $\mathcal{R}_i$ is an equivalence relation on $\mathcal{W}$. Since $\mathcal{R}_\Box$ is a total relation, $\mathcal{R}_i \subseteq \mathcal{R}_\Box$. Since for every $E_1 \in [\mathcal{R}_1], \ldots, E_n \in [\mathcal{R}_n]$: $E_1 \cap \ldots \cap E_n \neq \emptyset$, independence of agents holds. Note that $W \subseteq \mathcal{W}$ and for any $w, v \in W$, wR_iv iff $w\mathcal{R}_iv$.

To finish our construction, we need to define group plans, i.e., a collection of sets $\{\mathcal{U}_G\}_{G\subseteq Ags}$. We do it in the following manner:

- $\forall i \in Ags : \mathcal{U}_{\{i\}} = \mathcal{W}$;
- $\mathcal{U}_J = W$;
- $\mathcal{U}_G = \{w \in \mathcal{W} \mid \exists v \in W : \sim_G (v) \neq \emptyset \text{ and } (w, v) \in \bigcap\limits_{i\in G} \mathcal{R}_i)\}$

Note that for every $w, v \in W$, if $w \sim_G v$, then $(w, v) \in \bigcap_{i\in G} \mathcal{R}_i$. Then, for any group $G \subseteq Ags$, $\mathcal{U}_G \cap W = \{w \in W \mid \sim_G (w) \neq \emptyset\}$. According to Definition 4.2, $\mathcal{F} = \{\mathcal{W}, \mathcal{R}_\Box, \{\mathcal{R}_i\}_{i\in Ags}, \{\mathcal{U}_G\}_{G\subseteq Ags}\}$ is an atemporal **pSTIT**-frame. As we have already shown, $\{\mathcal{W}, \mathcal{R}_\Box, \{\mathcal{R}_i\}_{i\in Ags}\}$ is a **STIT**-frame. By definition, every $\mathcal{U}_G$ is a $\bigcap_{i\in G} \mathcal{R}_i$-closed set (including $\mathcal{U}_J$) and $\mathcal{U}_{\{i\}} = \mathcal{W}$ for any $i \in Ags$.

Moreover, $\mathcal{U}_J$-based subframe of $\mathcal{F}$ is equal to the initial $\mathbf{KB4}_G$-frame F. $\mathcal{U}_J = W$; for every relation $\mathcal{R} \in \{\mathcal{R}_\Box, \mathcal{R}_G\}$, $\mathcal{R} \cap (\mathcal{U}_J \times \mathcal{U}_J) = R$ and for every $G \subseteq Ags$:

$$\mathcal{U}_G \cap \mathcal{U}_J = \mathcal{U}_G \cap W = \{w \in W \mid \sim_G (w) \neq \emptyset\} = U_G$$

In other words, F is a plan-generated subframe of an atemporal **pSTIT**-frame, $\mathcal{F}$.

Since F was an arbitrary chosen $\mathbf{KB4}_G$-frame, we can conclude that every $\mathbf{KB4}_G$-frame is a plan-generated subframe of some **pSTIT**-frame. □

Theorem 4.19 (Soundness and completeness)**.** *Given any set of formulae* $\Gamma \subseteq \mathcal{L}_{pSTIT}$ *and any formula* $\varphi \in \mathcal{L}_{pSTIT}$*, the following holds:*

$$\Gamma \vdash_{SF(\mathbf{pSTIT})} \varphi \Leftrightarrow \Gamma \models_{\mathbf{L}^p_G} \varphi$$

Proof. ($\Rightarrow$) By a routine check of the axioms' validity.

($\Leftarrow$) For contradiction, assume $\Gamma \not\vdash_{\mathbf{L}^p_G} \varphi$. Then, $\Gamma \cup \{\neg\varphi\}$ is $\mathbf{L}^p_G$-consistent. By Lemmas 4.13, 4.17, there exists a pointed $\mathbf{KB4}_G$-model M, w, such that $M, w \models \Gamma \cup \{\neg\varphi\}$. By Theorem 4.18, there is an atemporal **pSTIT**-model $\mathcal{M}$, such that $\mathcal{M}_G, w \models \Gamma \cup \{\neg\varphi\}$, where $\mathcal{M}_G$ is the U_G-based submodel of $\mathcal{M}$. As a consequence, $\Gamma \not\models_{SF(\mathbf{pSTIT})} \varphi$. □

5. Conclusion and further work

In this paper, we have developed a new logic of group agency — plan-restricted group STIT logic. Our logic differentiates between proper group actions and mere aggregations of individual actions. Namely, we treat simultaneous actions of individuals as a group action if and only if their actions go in line with their mutually agreed-upon plan (in case such a plan exists).

We have defined the formal language $\mathcal{L}_{pSTIT}$, where we can express the modality $[pstit]_G\varphi$, "*group G jointly sees to it that* φ", and a corresponding class of models – **pSTIT**. We have demonstrated how this framework can distinguish between genuine group actions and uncoordinated simultaneous actions of independent agents, and how this distinction may be beneficial for attributing complicity and collective responsibility.

From the metalogical side, we have axiomatized the validities of a rich fragment of $\mathcal{L}_{pSTIT}$, $\mathcal{L}^{\mathcal{G}}_{pSTIT}$, with modalities only for pairwise disjoint groups. We have also provided a sound and complete axiom system for the special class of **pSTIT**-subframes that allow us to reason under the assumption that members of a designated group follow their plan.

We have abstracted away from the two important aspects of agency: epistemics and temporality. As for epistemics, it will be interesting to investigate the reasoning about group agency under uncertainty: agents may be (un)certain about actions and abilities of themselves and other agents, as well as which plans groups have and which sets of agents act as groups. We have pointed out some points of discussion about epistemic assumptions in subsection 3.4. As for temporality, since $\mathcal{L}_{pSTIT}$ is atemporal, we can only reason about the immediate consequences of agents' and groups' actions, not the long-term ones. Moreover, we regard plans as static: our framework lacks instruments to model revisions of plans. We leave the investigation of epistemic and temporal extensions of our logic for future work.

References

Alur, R., Henzinger, T. A., & Kupferman, O. (2002). Alternating-time temporal logic. *Journal of the ACM (JACM)*, *49*(5), 672–713.

Bazargan, S. (2013). Complicitous liability in war. *Philosophical Studies*, *165*(1), 177–195.

Belnap, N., Perloff, M., & Xu, M. (2001). *Facing the future: agents and choices in our indeterminist world*. Oxford University Press.

Blackburn, P., Rijke, M. d., & Venema, Y. (2001). *Modal logic*. Cambridge University Press.

Bratman, M. (1987). *Intention, plans, and practical reason*. Harvard University Press.

Christoff, Z. (2016). *Dynamic logics of networks: information flow and the spread of opinion* (Unpublished doctoral dissertation). University of Amsterdam.

Ciuni, R., & Lorini, E. (2018). Comparing semantics for temporal stit logic. *Logique et Analyse*, *243*, 299–339.

Danecki, R. (1984). Nondeterministic propositional dynamic logic with intersection is decidable. In *Symposium on computation theory* (pp. 34–53).

de Neufville, R., & Baum, S. D. (2021). Collective action on artificial intelligence: A primer and review. *Technology in Society*, *66*, 101649.

Duijf, H. (2022). *The logic of responsibility voids*. Springer.

Duijf, H., Broersen, J., Kuncova, A., & Abarca, A. I. R. (2021). Doing without action types. *The Review of Symbolic Logic*, *14*(2), 380–410.

Fagin, R., Halpern, J. Y., & Vardi, M. Y. (1992). What can machines know? on the properties of knowledge in distributed systems. *Journal of the ACM (JACM)*, *39*(2), 328–376.

Gilbert, M. (1992). *On social facts*. Princeton University Press.

Goldblatt, R. I., & Thomason, S. K. (1975). Axiomatic classes in propositional modal logic. In *Algebra and logic: Papers from the 1974 summer research institute of the australian mathematical society, monash university, australia* (pp. 163–173).

Goranko, V., & Enqvist, S. (2018). Socially friendly and group protecting coalition logics. In *17th international conference on autonomous agents and multiagent systems (aamas 2018), stockholm, sweden, july 10-15, 2018* (pp. 372–380).

Goranko, V., & Ju, F. (2022). A logic for conditional local strategic reasoning. *Journal of Logic, Language and Information*, *31*(2), 167–188.

Goubault, É., Kniazev, R., Ledent, J., & Rajsbaum, S. (2024). Simplicial models for the epistemic logic of faulty agents. *Boletín de la Sociedad Matemática Mexicana*, *30*(3), 90.

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

Horty, J. F. (2001). *Agency and deontic logic*. Oxford University Press.

Kutz, C. (2000). *Complicity: Ethics and law for a collective age*. New York.

Lorini, E. (2013). Temporal logic and its application to normative reasoning. *Journal of Applied Non-Classical Logics*, *23*(4), 372–399.

Nash, J. (1951). Non-cooperative games. *Annals of Mathematics*, *54*(2), 286–295. Retrieved 2024-12-16, from `http://www.jstor.org/stable/1969529`

Pauly, M., & Parikh, R. (2003). Game logic-an overview. *Studia Logica*, *75*, 165–182.

Pedersen, M. Y., Smets, S., & Ågotnes, T. (2020). Further steps towards a logic of polarization in social networks. In *International conference on logic and argumentation* (pp. 324–345).

Peleg, B., & Sudhölter, P. (2007). *Introduction to the theory of cooperative games* (Vol. 34). Springer Science & Business Media.

Sahlqvist, H. (1975). Completeness and correspondence in the first and second order semantics for modal logic. In *Studies in logic and the foundations of mathematics* (Vol. 82, pp. 110–143). Elsevier.

Schwarzentruber, F. (2012). Complexity results of stit fragments. *Studia logica*, *100*(5), 1001–1045.

Searle, J. R. (1990). Collective intentions and actions.

Segerberg, K., Meyer, J.-J., & Kracht, M. (2009). The logic of action.

Sergot, M. (2021). Some forms of collectively bringing about or 'seeing to it that'. *Journal of Philosophical Logic*, *50*(2), 249–283.

Smets, S., & Velázquez-Quesada, F. R. (2017). How to make friends: a logical approach to social group creation. In *Logic, rationality, and interaction: 6th international workshop, lori 2017, sapporo, japan, september 11-14, 2017, proceedings 6* (pp. 377–390).

Sugden, R. (1993). Thinking as a team: Towards an explanation of nonselfish behavior. *Social Philosophy and Policy*, *10*(1), 69–89.

Sverdlik, S. (1987). Collective responsibility. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *51*(1), 61–76.

Tamminga, A., & Duijf, H. (2017). Collective obligations, group plans and individual actions. *Economics & Philosophy*, *33*(2), 187–214.

Tollefsen, D. P. (2015). *Groups as agents*. John Wiley & Sons.

Tuomela, R. (2005). Two basic kinds of cooperation. In *Logic, thought and action* (pp. 79–107). Springer.

Tuomela, R. (2013). *Social ontology: Collective intentionality and group agents*. Oxford University Press.

Vakarelov, D. (1992). A modal theory of arrows. arrow logics i. In *European workshop on logics in artificial intelligence* (pp. 1–24).

Van Ditmarsch, H., & Kuznets, R. (2024). Wanted dead or alive: epistemic logic for impure simplicial complexes. *Journal of Logic and Computation*, exae055.

Wáng, Y. N., & Ågotnes, T. (2020). Simpler completeness proofs for modal logics with intersection. In *International workshop on dynamic logic* (pp. 259–276).